ISBN 978-82-575-1021-3
ISSN 1503-1667

Norwegian University of Life Sciences
NO–1432 Ås, Norway
Phone +47 64 96 50 00
www.umb.no, e-mail: postmottak@umb.no

Liv Borghild Anette Andreassen Bøhle

Norwegian University of Life Sciences • Universitetet for miljø- og biovitenskap
Department of Chemistry, Biotechnology and Food Science
Philosophiae Doctor (PhD) Thesis 2011:58

Philosophiae Doctor (PhD) Thesis 2011:58

# Studies of proteins involved in environmental responses of *Enterococcus faecalis* V583

Studier av proteiner som er involvert i *Enterococcus faecalis* V583 sin respons mot omgivelsene

Liv Borghild Anette Andreassen Bøhle

# Studies of proteins involved in environmental responses of *Enterococcus faecalis* V583

Studier av proteiner som er involvert i *Enterococcus faecalis* V583 sin respons mot omgivelsene

Philosophiae doctor (Ph.d.) Thesis

Liv Borghild Anette Andreassen Bøhle

Department of Chemistry, Biotechnology and Food Sciences
Norwegian University of Life Sciences

Ås, Norway, 2011

Thesis number 2011:58
ISSN 1503-1667
ISBN 978-82-575-1021-3

**TABLE OF CONTENTS**

**ACKNOWLEDGEMENT**

The present work was carried out during the period of 2006-2011 in the Protein Engineering and Proteomics group (PEP), Department of Chemistry, Biotechnology and Food Science, The Norwegian University of Life Sciences. The work was supported by the Research Council of Norway and a Departmental fellowship.

First of all I would like to thank my supervisor Prof. Vincent Eijsink for giving me the opportunity to do a PhD in his group. Thank you for your enthusiasm, for all the help and expertise you have giving me these years. It has truly been an inspiration working with you.

Next I would like to thank my supervisor Dr. Geir Mathiesen. Thank you for encouraging me to ask "stupid" questions by saying "No questions are stupid questions", and for always finding time to discuss science with me. You have helped me a lot and I am very grateful for being supervised by you. I would also like to thank Dr. Gustav Vaaje-Kolstad for introducing me to the field of chitin-degradation during the last part of my PhD-period. You have answered my questions patiently and I appreciate your willingness to guide me. I also thank my supervisor Prof. Ingolf F. Nes for introducing me to science and the interesting field of lactic acid bacteria.

I am grateful to all former and present members of the PEP-group for creating a stimulating working athmosphere. Thank you for all the help, inspiration and discussions during these years.

At last but not least I would like to thank my family, especially my parents and my sisters for always encouraging me, and believing in me. A special thank to my boyfriend Christoffer for being a good mental support the last months, for sharing my frustrations and pleasure, and for taking such good care of our daughter. You have really done an excelent job, and I could not have done this without you. Ane Johanne, thank you for giving me such joy and something else to think about even though the time has come to short. I am really looking foreward to be a real mum and sing you good-night songs again. Thank you for showing me whats most important in life.

Ås, September, 2011

Liv Anette Bøhle

**SUMMARY**

*Enterococcus faecalis* has traditionally been considered a harmless inhabitant of the gastrointestinal tract of mammals. However, in the past decade *E. faecalis* has become a major cause of infections in hospitalized patients. Due to a dramatic increase in resistance to most antibiotics, treatment of such infections is difficult. *E. faecalis* V583 was the first clinical vancomycin-resistant isolate reported in the United States, and the sequencing of its genome was completed in 2003. Increased knowledge of enterococci is crucial for successful development of new methods and therapies to combat the infections that they cause.

The main objective of the work described in this thesis was to identify and study proteins involved in environmental responses of *E. faecalis* V583. The first part of the thesis (Papers I and II) describes a proteomics approach, while the second part describes studies of individual proteins that were selected on the basis of their potential roles in host-microbe interactions (Paper III and IV). In Paper I we identified proteins that were differently expressed after exposure of *E. faecalis* V583 to bile. Bile is a major stress factor that the bacteria have to cope with in order to survive in the intestinal tract. This study was done using two-dimensional gel electrophoresis combined with mass spectrometry-based identification of differentially expressed proteins. Of a total of 500 proteins visualized on the gels, 53 unique proteins were identified as differently expressed and among these; proteins involved in fatty acid and phospholipid metabolism were overrepresented. The study of Paper I also allowed a comparison of proteome data with (previously generated) transcriptome data.

In the study described in Paper II, an enzymatic method for "shaving" of intact cells was combined with high resolution mass spectrometry methods to identify proteins located on the surface of *E. faecalis* V583. Surface proteins are important for communication and interactions both among bacterial cells and between a bacterium and its environment. This approach led to the identification of 69 unique proteins, of which 36 were predicted to have an extracellular localization. The majority of these proteins are involved in cell wall synthesis, pheromone-regulated processes and transport of solutes. Interestingly, this study yielded several proteins of unknown function that are interesting targets for further research.

Paper III describes an endo-*β*-*N*-acetylglucosaminidase (EF2863) that could potentially enable *E. faecalis* V583 to deglycosylate host glycoproteins. It was shown that the endo-*β*-*N*-acetylglucosaminidase, *Ef*Endo18A, hydrolyses glycosidic linkages in glycoproteins that contain N-linked glycans of the high-mannose and hybrid-type. This may be relevant for enterococcal survival and behavior in the host since it is predicted that two-thirds of the eukaryotic proteins are glycoproteins, including several proteins of the immune

system. In addition to being involved in nutrient acquisition, enzymes such as *Ef*Endo18A may also have an effect on the host immune system.

*E. faecalis* V583 contains a chitinolytic machinery whose expression is known to be regulated in response to various environmental stimuli, including stress factors that may be encountered in an animal host. Paper IV show that *E. faecalis* V583 can grow on chitin and chitin-derived sugars and describes an enzymological chacterization of a chitinase (EF0361) and a chitin binding protein (EF0362). The chitinase, *Ef*Chi18A, is an endochitinase, whereas the chitin-binding protein, *Ef*CBM33A, cleaves glycosidic bonds via an oxidative mechanism and acts synergistically with the chitinase. An ultra-high resolution crystal structure of *Ef*CBM33A revealed details of a conserved binding surface that contains the metal-binding catalytic center. There are several studies in the literature indicating that the role of CBM33s could be more complex than a simple food scavenging role, and that suggest roles in virulence and interactions with host cells.

Taken together, these studies have provided novel insights into some of the proteins that may be involved in behavior and impact of enterococci and in their interactions with a human host.

## SAMMENDRAG

Tradisjonelt har *Enterococcus faecalis* blitt ansett som en harmløs bakterie som blant annet forekommer i tarm hos pattedyr. De siste tiår har imidlertid dette endret seg, og man har sett en økning i forekomsten av disse bakteriene som årsak til infeksjoner hos pasienter på sykehus. Behandling av infeksjoner forårsaket av enterokokker er vanskelig på grunn av økt resistens mot de fleste typer antibiotika. *E. faecalis* V583 var det første kliniske vancomycin-resistente isolatet som ble rapportert i USA, og genomet var ferdig sekvensert i 2003. Økt kunnskap om disse bakteriene er viktig for å kunne utvikle alternative behandlinger og for å bekjempe infeksjoner som enterokokker forårsaker.

Hovedfokuset i denne avhandlingen har vært å studere hvilke proteiner hos *E. faecalis* V583 som responderer på ulike faktorer i omgivelsene. I den første delen (Artikkel I og II) ble proteomikk brukt for å identifisere proteiner, mens i den andre delen (Artikkel III og IV) ble bestemte proteiner valgt ut og karakterisert på bakgrunn av deres potensielle rolle i interaksjonen mellom vert og mikrobe. I Artikkel I ble proteiner som viste endret ekspresjon som en respons på galle identifisert. Galle er en betydelig stressfaktor bakterier må tilpasse seg med tanke på å overleve i tarmsystemet. Dette ble gjort ved bruk av to-dimensjonell gel elektroforese og masse spektrometri-basert identifisering. Av totalt 500 synlige proteiner på gelene, ble det identifisert 53 proteiner som endret proteinuttrykk etter at galle ble tilsatt. Dette var hovedsakelig proteiner involvert i fettsyre og fosfolipid metabolismen. Dette er tidligere gjort på transkriptom-nivå, noe som gjorde det mulig å sammenligne dataene mellom disse to studiene.

I Artikkel II ble proteiner som er lokalisert på overflaten til *E. faecalis* V583 identifisert og analysert ved enzymatisk "barbering" av bakteriens overflate, væskekromatografi og masse spektrometri. Proteiner som sitter på overflaten spiller en viktig rolle ved kommunikasjon og interaksjon mellom bakterieceller og mellom bakterien og dens omgivelser. Totalt ble 69 proteiner identifisert, 36 av disse var predikert til å være lokalisert på bakteriens overflate. De fleste av disse proteinene er involvert i syntesen av cellevegg, feromonregulerte prosesser og i transport av stoffer over membranen. Flere av proteinene har også ukjent funksjon, og er derfor aktuelle mål for mer inngående studier.

Artikkel III beskriver en endo-*β*-*N*-acetylglucosaminidase (EF2863) som potensielt kan spalte av sukker fra vertens glykosylerte proteiner. Det ble vist at endo-*β*-*N*-acetylglucosaminidase, *Ef*Endo18A, kunne hydrolysere glykosidbindingen i glykoproteiner bestående av N-linkede glukaner bestående av ”high-mannose” og ”hybrid-type”. Dette kan være spesielt relevant for *E. faecalis* V583 sin evne til overlevelse og adferd i verten, siden

det er predikert at 2/3 av eukaryote proteiner er glykoproteiner, inkludert flere proteiner som hører til immunsystemet. I tillegg til å uttnytte karbohydrater/sukkermolekyler som næring kan også en hypotese være at *Ef*Endo18A har en effekt på vertens immunsystem.

*E.faecalis* V583 har et kitinolytisk system, som blir regulert som en respons på ulike stimuli fra omgivelsene, inkludert stress-faktorer som bakterien kan møte i verten. Artikkel IV viser at *E.faecalis* V583 kan vokse på kitin og sukkere av kitin-derivater og det beskrives en enzymologisk karakterisering av en kitinase (EF0361) og et kitin-bindende protein (EF0362). Kitinasen, *Ef*Chi18A, er en endokitinase, mens det kitin-bindende proteinet, *Ef*CBM33A spalter glykosidbindinger ved hjelp av en oksidativ mekanisme og virker synergistisk sammen med kitinasen. En høy-oppløselig krystallstruktur av *Ef*CBM33A synliggjorde detaljer av en konservert bindingsoverflate som inneholder ett metall-bindende katalytisk senter. Litteraturen viser flere studier som indikerer at CBM33 kan ha en mer kompleks rolle enn å bare bidra til få tak i næring, men at den i tillegg kan ha en rolle innen virulens og binding til vertens celler.

Sett under ett, har disse studiene gitt ett bedre og nytt innblikk i noen av proteinene hos enterokokker som kan være involvert i adferd og interaksjon med verten.

## ABBREVIATIONS

2D – two dimensional

Asn - asparagine

Cazy – Carbohydrate-Active Enzymes database

CBM – carbohydrate-binding module

*Ef*Endo18A – endo-*β*-*N*-acetylglucosaminidase from *Enterococcus faecalis* V583

*Ef*Chi18A – chitinase from *Enterococcus faecalis* V583

*Ef*CBM33A – chitin binding protein from *Enterococcus faecalis* V583

GH – glycoside hydrolase

GlcNAc – N-acetylglucosamine

LAB – Lactic Acid Bacteria

LC – liquid chromatography

PMF – proton motive force

Sec – secretion pathway

Ser – serine

Thr - threonine

## LIST OF PAPERS

### Paper I

Bøhle, L. A., E. M. Færgestad, E. Veiseth-Kent, H. Steinmoen, I. F. Nes, V. G. H. Eijsink and G. Mathiesen (2010). "Identification of proteins related to the stress response in *Enterococcus faecalis* V583 caused by bovine bile." Proteome Science **8**(1): 37.

### Paper II

Bøhle, L. A., T. Riaz, W. Egge-Jacobsen, M. Skaugen, O. L. Busk, V. G. H. Eijsink and G. Mathiesen (2011). "Identification of surface proteins in *Enterococcus faecalis* V583." BMC Genomics **12**: 135.

### Paper III

Bøhle, L. A., G. Vaaje-Kolstad, G. Mathiesen and V. G. H. Eijsink (2011). "An endo-*β*-*N*-acetylglucosaminidase from *Enterococcus faecalis* V583 responsible for the hydrolysis of high-mannose and hybrid-type N-linked glycans." Accepted for publication in FEMS Microbiology Letters.

### Paper IV

Vaaje-Kolstad G., L. A. Bøhle, S. Gåseidnes, B. Dalhus, M. Bjørås, G. Mathiesen and V. G. H. Eijsink (2011). "Characterization of the chitinolytic machinery of *Enterococcus faecalis* V583 and high resolution structure of its oxidative CBM33 enzyme." Submitted to Journal of Molecular Biology.

### Other papers by the author

Bøhle, L. A., D. A. Brede, D. B. Diep, H. Holo and I. F. Nes (2010). "Specific degradation of the mucus adhesion-promoting protein (MapA) of *Lactobacillus reuteri* to an antimicrobial peptide." Applied and Environmental Microbiology **76**(21): 7306-7309.

# 1. INTRODUCTION

The enterococci are generally considered as harmless commensals of the gastrointestinal tract of mammals. However, they are harsh bacteria that are able to withstand a range of different conditions in their environment and in the past decade they have become a leading cause of nosocomial infections, like bacteraemia, peritonitis, endocarditis and urinary tract infections. Because of their increasing resistance toward many antibiotics, including vancomycin, treatment of enterococcal infections is difficult. The need for alternative therapeutics is increasing, which underpins the importance of getting increased knowledge of these bacteria. In this respect, both identification of proteins that are differently expressed in response to environmental factors and targeted studies of potentially important proteins are useful. For example, such studies may yield information on how the bacteria survive in and adapt to the intestinal tract, or they may give insight in the mechanisms underlying the bacteria's ability to colonize and subsequently cause an infection.

## 1.1 *Enterococcus*

The genus *Enterococcus* consists of Gram-positive, catalase-negative, non-spore forming, facultatively anaerobic bacteria that occur as single cocci, in pairs or in short chains. They have the ability to grow at temperatures between 10 °C and 45 °C, and in the pH range from 4.6 to 9.6. The microbe survive harsh conditions like high salinity (up to 6.5 % NaCl), heating at 60 °C for 30 minutes, and bile salt stress [up to 40 % (w/v)] (Murray 1990; Fisher et al. 2009). Enterococci are chemoorganotrophic and are considered as lactic acid bacteria (LAB) due to production of L-lactic acid from hexoses by homofermentative lactic acid fermentation (Franz et al. 2003). LAB generally uses substrate-level phosphorylation in order to generate energy. However, *Enterococcus faecalis* is one of few LAB that also can generate energy by utilizing oxidative phosphorylation using a proton motive force (PMF) established by electron transport (Leblanck 2006).

The enterococci were for a long time considered to belong to the streptococci. In 1899, the term "enterococcus" was used for the first time when Thiercelin refered to bacteria of intestinal source that appeared as pairs or short chains in human feces. In 1937, Sherman (Sherman 1937) divided the streptococci into four subgroups, one of which was the *Enterococcus* group, which was shown to posess the group D antigen. However, enterococci remained classified as streptococci until 1987 (Klein 2003; Fisher et al. 2009) when the streptococci were split into the genera *Streptococcus, Enterococcus*, and *Lactococcus* based

on molecular data. The differences between these genera can be demonstrated by 16S rRNA sequence comparisons as illustrated in Figure 1. The genus *Enterococcus* now comprises 28 species (Fisher et al. 2009), which are divided in a number of different strains. The number of sequenced enterococcal genomes is increasing. Recently, draft genome sequences of 28 enterococcal species have been reported (Palmer et al. 2010). For *E. faecalis*, the enterococcal species studied here, four genomes have been completed: the first clinical vancomycin-resistant isolate V583 (Paulsen et al. 2003), the probiotic strain Symbioflor 1 (Domann et al. 2007), the human isolate OG1RF (Bourgogne et al. 2008), and the commensal *E. faecalis* 62, isolated from a healthy infant (Brede et al. 2011). These genome data provide an excellent opportunity to predict and analyse metabolic pathways harnessed by the respective bacteria, as well as many other properties, such as the presence of putative pathogenicity factors. Presumably, this will lead to a better understanding of bacterial behavior and impact in their natural environments.

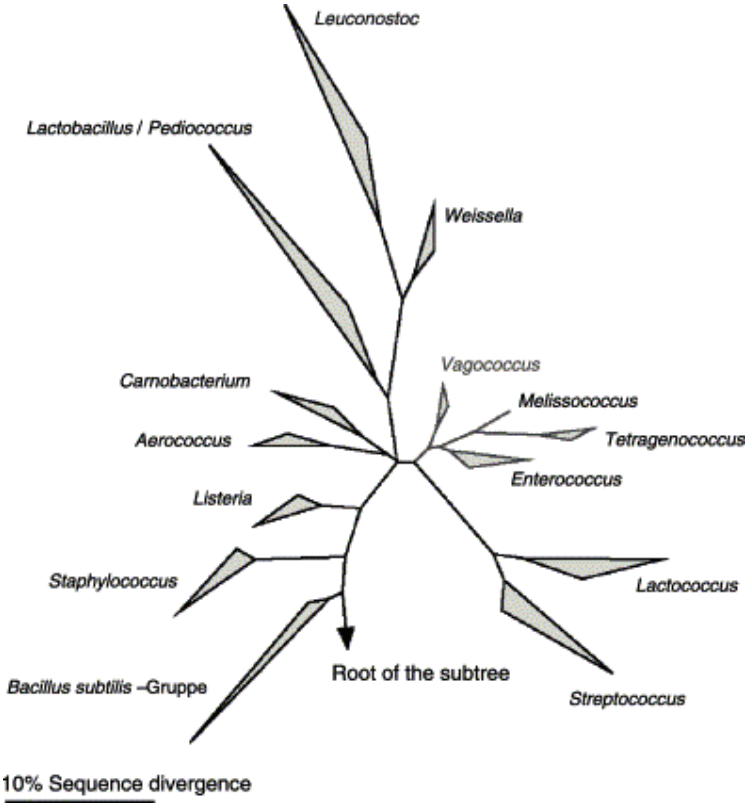



**Figure 1. Phylogenetic relationship of the Gram-positive genera.** The dendrogram is based on 16S rRNA-comparison of Gram-positive genera demonstrating the difference between *Enterococcus, Lactococcus* and *Streptococcus* (Klein 2003).

Enterococci can survive under harsh conditions, including those found in animal digestive tracts, as illustrated by the fact that they are found as a member in the intestinal tracts of nearly all animals. In humans, typical concentrations of enterococci in feces are up to $10^8$ colony-forming units per gram, with *E. faecalis* and *Enterococcus faecium* being most frequently isolated. Enterococci have also been identified in water, sewage, soil and plants, where their presence usually is ascribed to fecal contamination (Huycke et al. 1998). Enterococci also occur in food products, vegetables and insects (Foulquie Moreno et al. 2006).

**1.2 The health-promoting effects of enterococci**

Like members of most other LAB genera, some strains of the enterococci are used as starter cultures, feed supplements or as probiotics. The use of enterococci for cheese production (e.g. Cheddar, Feta, Mozarella) is important in most Southern European countries because of their proteolytic and esterolytic activities and their production of diacetyl contribute to flavor and ripening (Ogier et al. 2008). They are also present in other fermented food products like sausages and olives (Foulquie Moreno et al. 2006). In Denmark a probiotic fermented milk called Gaio is produced, which contains *E. faecium* and is claimed to decrease the cholesterol levels in blood (Bertolami et al. 1999). In Switzerland *E. feacium* SF68 has been used successfully for treatment of diarrehae in both children and adults and the bacterium is considered as an alternative to antibiotics (Cremonini et al. 2002; Foulquie Moreno et al. 2006; Ogier et al. 2008). Other probiotic products containing enterococci are CausidoR (Agerholm-Larsen et al. 2000; Franz et al. 2003), and Idoform produced in Norway.

Several of the enterococcal probiotic strains have the ability to produce bacteriocins, called enterocins, which have antimicrobial activity against other Gram-positive bacteria including. *Listeria monocytogenes*, *Staphylococcus aureus*, *Clostridium botulinum*, and other bacteria that frequently causes infection and diseases in mammals (Maisnier-Patin et al. 1996; Simonetta et al. 1997; Laukova et al. 1999; Leroy et al. 2003; Cocolin et al. 2007; Theppangna et al. 2007; Ibarguren et al. 2010). Bacteriocins from LAB are divided into different classes, and the enterocins belong to Class I (lantibiotics), Class IIa (pediocin-like), Class IIc and Class III. Most bacteriocins target the cytoplasmic membrane, where they form pores, leading to leakage and disruption of the transmembrane potential (Foulquie Moreno et al. 2006). The ability to produce bacteriocins adds to the health-promoting effects of enterococci because this leads to active supression of the growth of pathogenic bacteria.

The use of enterococci as probiotic strains is controversial, due to the risk of transfer of antimicrobial resistance and virulence genes (see below). In 2004, a total of 10 different preparations, from 9 different *E. faecium* strains, were authorized as feed additives in the European Union (Foulquie Moreno et al. 2006); in that same year Canada banned the use of enterococci as probiotics (Ogier et al. 2008).

### 1.3 Pathogenicity and virulence factors of the enterococci

The enterococci have emerged as a major cause of nosocomial infections and are ranged among the top four causes of blood infections and second as a cause of urinary tract infections both in the United States and Europe (Ogier et al. 2008; Sood et al. 2008). The majority of these infections are caused by either *E. faecalis* or *E. faecium*. Even though the bacteria are harmless to healthy individuals, enterococcal clinical isolates may become pathogenic in patients in intensive care units, in immunocompromised patients with severe underlying diseases, or in elderly people (Ogier et al. 2008). Nosocomial infections caused by enterococci have mortality rates of more than 60 % (de Fatima Silva Lopes et al. 2005). Two factors may be linked to this high rate. One factor is the intrinsic resistance of enterococci to a wide range of antibiotics including vancomycin, which is used as a last resort treatment for multiple antibiotic resistant enterococci (Huycke et al. 1998). The other factor is the large number of plasmids in *E. faecalis* and *E. faecium*. These plasmids may mediate efficient transfer of resistance, virulence and other adaptive traits, not only to enterococci, but more recently also to staphylococci (Palmer et al. 2010).

The emergence of vancomycin-resistant enterococci in Europe has been associated with the use of the glycopeptide avoparcin, which is an antibiotic that has been used as a feed additive to promote growth in pigs and poultry until the late 1990s, when this use was banned in the European Union. In the United States the use of glycopeptides such as avoparcin as feed additive has never been approved. However, in the United States, hospital use of vancomycin has been far greater than in Europe, and this is considered a major factor in the development and spread of vancomycin-resistant enterococci (Amyes 2007; Sood et al. 2008).

Historically, the ratio between infections caused by *E. faecalis* and *E. faecium* has been 10:1. However in recent years, this ratio has been changed (Mundy et al. 2000). Over the last 20 years, *E. faecalis* and *E. faecium* have been responsible for 60 % and 40 %, respectively, of hospital-acquired infections caused by enterococci (van Schaik et al. 2010). One reason for the increasing problems with *E. faecium* is the high occurrence of vancomycin resistance in this species (Mundy et al. 2000; Soderblom et al. 2010). Vancomycin resistant

enterococci were first detected in Europe in 1986, and in 1989 a vancomycin resistant *E. faecalis* clinical isolate (*E. faecalis* V583) was reported in the United States (Sahm et al. 1989). The complete genome sequence of this strain was reported in 2003 (Paulsen et al. 2003) (Figure 2).

*E. faecalis* V583 is a metabolically diverse bacterium and contains a wide range of regulatory systems. According to Paulsen et al. (2003) the strain contains a circular chromosome of 3.2 Mb with 3182 open reading frames. In addition to the chromosome, the genome of *E. faecalis* V583 contains three circular plasmids named pTEF1 (66320 bp), pTEF2 (57660 bp) and pTEF3 (17963 bp). The average G+C content of the chromosome is ~ 38 %. Interestingly, the gene cluster associated with the vancomycin resistance, positioned near 2.22 Mb (Figure 2), showed a deviating, higher G+C content, which may indicate acquisition of genetic material from other species. Generally, the genome sequence indicates that this versatile strain has collected large amounts of foreign DNA. 25 % of the *E. faecalis* V583 genome consists of mobile elements and/or exogenously acquired DNA, including integrated phage regions, insertion elements, transposons, a pathogenicity island and integrated plasmid genes. The vancomycin-resistant phenotype is encoded within a mobile genetic element, and the genetic determinant for this resistance is almost identical to the *van*B genes found in Tn*1549* which is a transposon from an enterococcal plasmid similar to pAD1 (Paulsen et al. 2003; Tendolkar et al. 2003). pAD1 is a pheromone-responsive, conjugative plasmid found in clinical isolates of *E. faecalis* that encodes cytolysin and that has been shown to contribute to virulence in animal models (Clewell 2007).

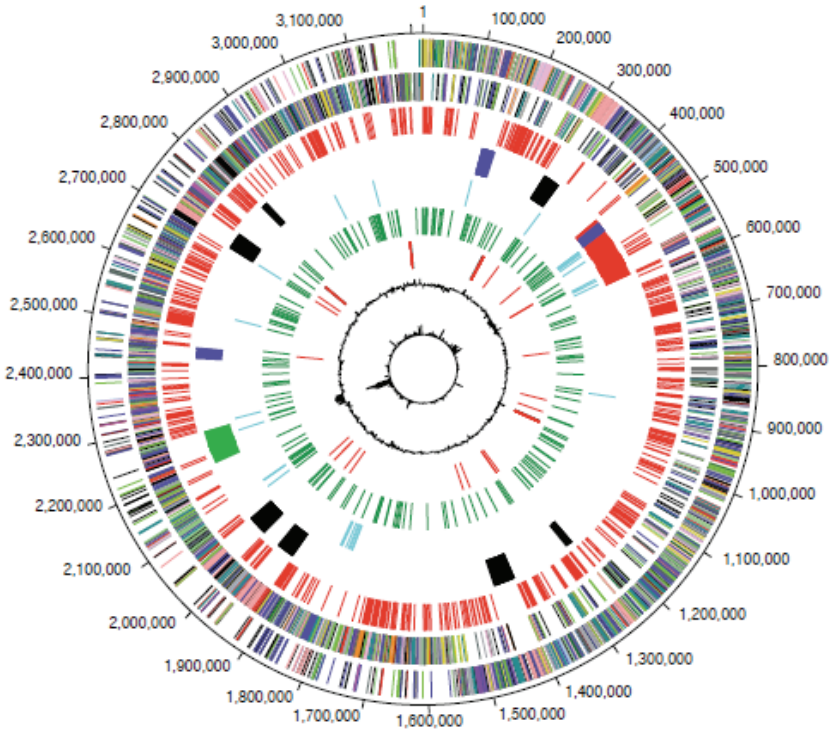



**Figure 2. Circular representation of the chromosome of *E. faecalis* V583.** The different circles represent different traits found on the bacterial chromosome. From the outer to the inner circle: 1) Predicted coding regions on the plus strand color-coded by functional role categories: salmon, amino acid biosynthesis; light blue, biosynthesis of cofactors, prosthetic groups and carriers; light green, cell envelope; red, cellular processes; brown, central intermediary metabolism; yellow, DNA metabolism; green, energy metabolism; purple, fatty acid and phospholipid metabolism; pink, protein fate/synthesis; orange, purines, pyrimidines, nucleosides, nucleotides; blue, regulatory functions; grey, transcription; teal, transport and binding proteins; black, hypothetical proteins. 2) Predicted coding regions on the minus strand color-coded by functional categories as in the outer circle. 3) Proteins conserved amongst ten sequenced low-GC Gram-positive bacteria, red. 4) Phage genes, black; genes located within a putative pathogenicity island, red; genes within the putative vancomycin resistant conjugative transposon, green; integrated plasmid genes, blue. 5) Transposase genes, blue. 6) Predicted surface exposed proteins, green. 7) tRNA, red; rRNA operons, black. 8) GC % curve in black. 9) Atypical nucleotide composition curve in black (Paulsen et al. 2003).

Because species of the enterococci, especially *E. faecalis,* are responsible for a high number of hospital-acquired infections, recent research on the species has been focused on

identifying, detecting and trying to control virulence and individual virulence factors, rather than on the beneficial properties. In order to cause infection, pathogenic bacteria including enterococci must have the capability to colonize the gastrointestinal tract, adhere to extracellular matrix proteins and epithelial cells, and subsequently cause an infection. Several enterococcal virulence factors have been identified and most of these are part of or relate to the extracellular proteome. These include four different types of factors:

1. Secreted virulence factors

- Cytolysin. This is a two-subunit toxin, related to the lantibiotic bacteriocins, that is *β*-haemolytic in humans and bactericidal against other Gram-positive bacteria. Cytolysin secretion is regulated by a quorum-sensing mechanism involving a two-component regulatory system (Coburn et al. 2003). The responsible operon in *E. faecalis* is localized on a pheromone-responsive plasmid or in a pathogenicity island. *E. faecalis* produces cytolysin as a response to the sensing of target cells such as erythrocytes. When target cells are absent, one subunit, CylL(L) forms a complex with the other subunit, CylL(S). This blocks autoinduction of the operon. When target cells are present, CylL(L) binds to the target, which leads to accumulation of CylL(S) and induction of increased expression (Coburn et al. 2004). Experimental studies have shown that enterococcal cytolysin contributes to virulence in infection models (Stevens et al. 1992; Garsin et al. 2001).
- Gelatinase, GelE. This is a secreted zinc-metallo-endopeptidase, whose production is regulated by a quorum sensing system which is controlled by the *fsr* two-component locus (Qin et al. 2001). Gelatinase activity provides nutrients to the bacterium by degrading host tissue, and has a function in biofilm formation (Hancock et al. 2004). In addition, GelE has an important impact on translocation of *E. faecalis* across human enterocyte-like T84 cells (Zeng et al. 2005).

2. Surface located virulence factors

- Surface protein, Esp. Esp is a cell-wall-associated protein (Shankar et al. 1999) that promotes adhesion, colonization and evasion of the immune system, and is involved in antibiotic resistance (Foulquie Moreno et al. 2006). Esp contributes to biofilm formation, which is an important trait among clinical *E. faecalis* isolates, and it promotes attachment of *E. faecalis* on abiotic surfaces, like catheters and implants (Toledo-Arana et al. 2001).
- Aggregation substance, Agg. Agg is a pheromone-inducible surface protein that is involved in both the spread of antibiotic resistance and in the pathogenesis of enterococcal infections (Waters et al. 2001). It promotes cell-cell contacts and adhesion to host cells and extracellular matrix proteins (Hallgren et al. 2009). The presence of this protein increases

the hydrophobicity of the enterococcal cell surface (Hirt et al. 2000), which may contribute to adherence and attachment of bacteria (Jenkinson et al. 1997).

- Adhesin, Ace. Ace is a collagen-binding cell-surface protein belonging to the "microbial surface components recognizing adhesive matrix molecules" (MSCRAMM) family (Rich et al. 1999; Sillanpaa et al. 2004). This family of bacterial adhesins recognizes extracellular matrix components, and binds to them. Several MSCRAMMs have been isolated and characterized from staphylococci and streptococci (Patti et al. 1994; Foster et al. 1998).
- Capsular polysaccharides. Two types of capsular polysaccharides have been identified in enterococci, one is encoded by the enterococcal polysaccharide biosynthesis cluster (*epa*) and the other is called serotype C capsular polysaccharide (*cps*). Both play an important role in biofilm formation, adherence, and evasion of the host immune system (Xu et al. 1998; Hancock et al. 2002; Sava et al. 2010). Cps masks bacterial surface antigens, a trait that makes it possible for the bacterium to escape the host immune response (Thurlow et al. 2009). While the *epa* gene clusters are widespread among *E.faecalis*, c*ps* has mainly been identified in clinical isolates (Hancock et al. 2002; Teng et al. 2002).

3. Extracellular superoxide production. Production of $O_2^-$ is mainly limited to *E. faecalis*, and is a trait shared with only few other prokaryotes. Superoxide is a free radical anion formed by reduction of molecular oxygen, and can lead to the formation of strong oxidants such as hydrogen peroxide and hydroxyl radicals. The production of superoxide is dependent on the activity of a membrane-associated demetylmenaquinone (Huycke et al. 2001). *E. faecalis* strains associated with bacteremia produced $O_2^-$ at a 60 % higher rate than other isolates (Huycke et al. 1996). Superoxide may be an important cause of oxidative stress in the intestinal tract, and may damage host tissue. It has been proposed that intestinal production of superoxide may be linked to the emergence of colorectal cancer (Huycke et al. 2002).
4. Pili. It has been proposed that formation of pili plays a role in colonization of human hosts. Two loci responsible for pilus formation have been identified in *E. faecalis*, designated the *ebp* locus (Nallapareddy et al. 2006), and the *bee* locus (Schluter et al. 2009). It has been shown that the *ebp* locus is highly conserved among *E. faecalis*, while only 1.2 % have the *bee* pilus locus (Nallapareddy et al. 2011). The *ebp* locus has been shown to play a role in urinary tract invasion and in mediating adherance to platelets in endocarditis (Singh et al. 2007; Nallapareddy et al. 2011).

### 1.4 The proteome of *E. faecalis* V583, with focus on stress responses

According to the LocateP database (updated March 10, 2010; (Zhou et al. 2008)) the proteome (chromosome plus plasmids encoded proteins) of *E. faecalis* V583 consists of 3264 proteins. The predicted distribution of these proteins in various subcellular localizations is depicted in Figure 3. Almost 30 % of the proteins are predicted to have membrane and extracellular locations, while the remaining proteins are intracellular.

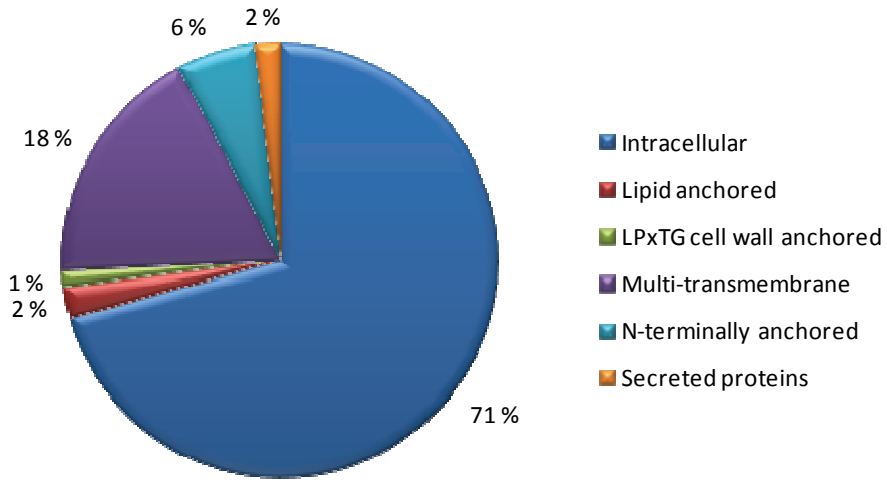



**Figure 3. Predicted localization of proteins of *E. faecalis* V583**. The numbers are based on the LocateP database (Zhou et al. 2008) which contains predictions for the subcellular localization of bacterial proteins. The frequencies of the various subcellular localizations are indicated in percent. *E. faecalis* V583 contain different types of anchored proteins: Lipid anchored (74 proteins), post-translational addition of a lipid anchor; LPxTG cell wall anchored (38 proteins), covalently coupled to the cell wall through sortase activity; Multi-transmembrane (588 proteins), contains at least one transmembrane helix; N-terminally anchored (190 proteins) and C-terminally anchored (4 proteins), contains a Sec-type leader peptide without a signal peptidase cleavage site (see also Fig 5.). Proteins that are non-covalently attached to the cell wall though specific binding domains (see text) are categorized as secreted (67 proteins) in this picture.

#### 1.4.1 The extracellular proteome

Extracellular proteins are important for uptake of nutrients, cell-to-cell communication, detoxification of the environment, adhesion, and killing of potential competitors (Tjalsma et al. 2004). In addition, they play an important role as virulence factors of pathogenic bacteria as described in section 1.3. Because proteins exposed on the bacterial surface can interact

directly with extracellular molecules like drugs and antibodies, they are potential drug targets as well as targets for passive or active immunization. The ability of surface proteins to interact with the human immune system has been exploited and vaccines based on such proteins have been commercialized, while others are under development (Pizza et al. 2000; Maione et al. 2005).

Proteins synthesized at the ribosome can be sorted to various destinations depending on the presence or absence of a signal peptide and specific retention signals. Bacteria have two major secretion pathways: the general secretion (Sec) pathway (Figure 4) and the twin-arginine translocation pathway. The major part of the extracellular proteins of Gram-positive bacteria, including enterococci, is predicted to follow the Sec pathway (Tjalsma et al. 2004). This pathway includes three stages:

1) Targeting. Proteins that are secreted (called pre-proteins) contain an N-terminal signal peptide in order to be sorted from cytoplasmic proteins, to ensure proper targeting to the translocation machinery and for initiation of translocation across the membrane. The signal peptide is recognized by chaperones like the signal recognition particle that binds and targets the pre-protein to the translocase at the membrane and that also contributes to keeping the protein in an unfolded state.

2) Translocation. The translocase comprise a membrane-embedded protein-conducting channel consisting of SecA (ATPase), SecYEG (pore) and SecDF. The pre-protein cross the membrane through this channel at the expense of ATP.

3) Folding and release. Signal peptidases (SPaseI, or SPaseII for lipoproteins) recognise a cleveage-site between the signal peptide and the secretory protein, which results in removal of the signal peptide from the secretory pre-protein when the C-domain emerges at the extracytoplasmatic side of the membrane. This is subsequently followed by folding of the polypeptide chain which thereby converts to a mature, functionally active protein. In the case of lipoproteins the pre-proteins are first modified by a diacylglyceryl-transferase that puts on a lipid anchor, followed by cleavage by SPaseII (Tjalsma et al. 2004).

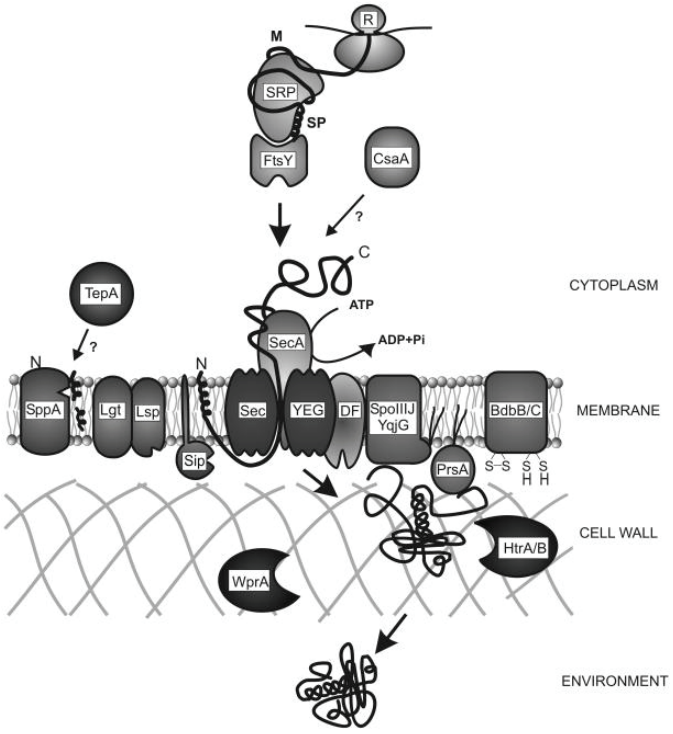



**Figure 4. Secretion of proteins by the sec-pathway in the Gram positive model organism, *Bacillus subtilis*.** Secreted proteins are synthesized as pre-proteins with an N-terminal signal peptide (SP). Chaperones like SRP, FtsY and/or CsaA bind to the pre-protein and direct it to the translocase, which consists of SecA, SecY, SecE, SecG and SecDF. SecA (ATPase) has affinity for the pre-protein-SRP complex, and functions as a molecular motor that couples the energy of ATP binding and hydrolysis to the translocation of the pre-protein through SecYEG. During or shortly after translocation, the pre-protein is cleaved by a signal peptidase type I (Sip), or lipid modified by the diacylglyceryl-transferase (Lgt) and subsequently cleaved by the signal peptidase type II (Lsp). Degradation of signal peptides may involve SppA and TepA. The protein folds in a process that may be dependent on the activities of PrsA, BdbBC and/or SpoIIIJ/YqjG, and the quality of the folded protein is controlled by HtrA, HtrB and WprA. The mature protein is subsequently released into the environment (Tjalsma et al. 2004).

As indicated in Figure 3, extracellular proteins may find different locations on the bacterial surface. They are either released to the environment or localized in the cell membrane or the cell wall, as summarized in Figure 5. Proteins that are non-covalently bound

to the cell wall contain one or more domains that have affinity for the cell wall. One example of such a domain is the LysM domain (Buist et al. 2008). Proteins that are covalently bound to the cell wall need a C-terminal cell wall sorting signal, consisting of the LPxTG motif followed by a C-terminal hydrophobic domain with a positively charged tail. A transpeptidase (sortase A) is responsible for cleveage between the threonine and the glycine of the LPxTG motif and catalyzes the formation of an amide bond between the carboxyl-group of threonine and the amino-group of peptidoglycan cross-bridges.

Lipoproteins are a group of membrane anchored proteins which have a conserved lipobox in the preprotein that contain a cystein residue which after translocation through the Sec pathway are modified by the diacylglyceryl transferase, before cleveage by the Signal peptidase II. The attachment of a diacylglyceryl group to the cysteine residue located as +1 next to the cleavage site prevents release of the lipoprotein into the environment and promotes attachment to the membrane. Membrane proteins have no signal peptidase cleveage site and because of the presence of one or more transmembrane domains they remain anchored to the membrane (Tjalsma et al. 2004). These membrane proteins may be integral membrane proteins, containing one or more hydrophobic membrane-spanning helices, or they may insert into the membrane by the N-terminal signal peptide only. The latter situation entails that the signal peptide does not contain a cleavage site for SPaseI.

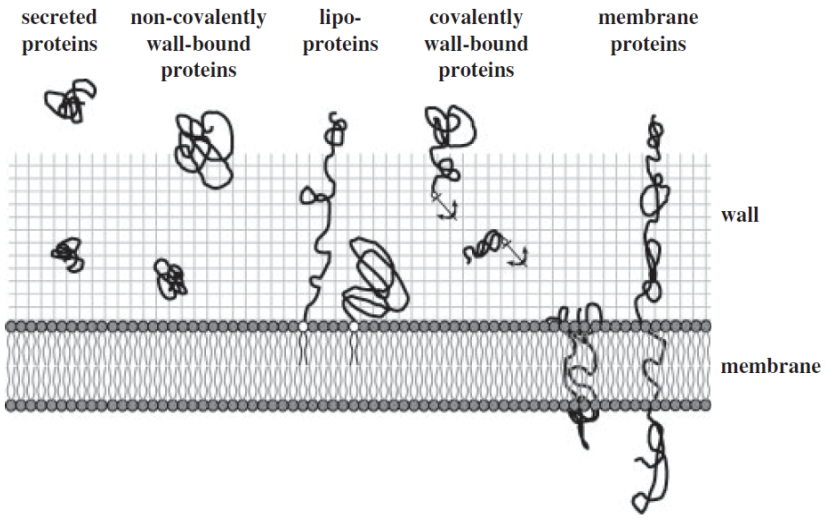



**Figure 5. Overview of the protein classes of the surface proteome.** Secreted proteins that don't have a retention signal are released into the extracellular surroundings. Non-covalently cell wall-associated proteins have specific domains that interact with specific cell wall components. Lipo-proteins are retained at the extracytoplasmic side of the membrane due to

diacyl-glyceryl modification of their N-terminal Cys residue. Covalently cell wall-associated proteins are attached to the cell wall by covalent linkage to the peptidoglycan (indicated by the anchor) through the action of sortases. Membrane proteins are anchored to the cytoplasmic membrane by hydrophobic transmembrane domains (Dreisbach et al. 2011).

### 1.4.2 Bacterial adaptation to survive and colonize in the gastrointestinal tract

The intestinal tract of mammals is the most densely populated ecosystem on earth, comprising $10^{14}$ microbes (Ley et al. 2006), belonging to an estimated 500-1000 different species (Xu et al. 2007), with the *Bacteroidetes* and *Firmicutes* being the dominant phyla (Eckburg et al. 2005).

The entire gastrointestinal tract is covered by mucosal surfaces that are composed of water and glycoproteins called mucins. This layer is important as a lubricant, as a selective barrier allowing passage of nutrients, and as a defence system that protects the underlying epithelial cells from mechanical damage or entrance of harmful substances and pathogens. This layer interacts directly with bacteria in the lumen, and represents the first line of defence against bacterial penetration (Derrien et al. 2010). The human intestinal microbiota play many beneficial roles to the host including enhancement of digestive efficiency, promotion of proper immune responses, and limiting pathogen colonization. In return, the mucin provides attachment sites to the bacteria and offers an important carbon source. Importantly, the success of this symbiotic relationship between host and microbe depends on restriction of bacterial penetration of host tissues (Duerkop et al. 2009). The ability to degrade mucin and to use released carbohydrates and amino acids as nutrients has been shown for a number of bacterial species including *Lactobacillus rhamnosus* GG (Sanchez et al. 2010), *Bifidobacterium* species (Ruas-Madiedo et al. 2008; Ruiz et al. 2011), and *Akkermansia muciniphila* (Derrien et al. 2004). While this ability may be beneficial in healthy symbiontic relationships it also presents a potential problem, since disturbance of the structure and function of the mucus layer can be deleterious for the host and is characteristic for the pathology of many diseases.

While it may seem that mucin affects the microbial ecosystem in the intestinal tract beneficially, other factors may be detrimental for survival. Both commensal and pathogenic bacteria must resist the deleterious actions of a number of potential stress factors present in the intestinal tract in order to survive. These stress factors include low pH, low oxygen levels, nutrient limitations, elevated osmolarity and the presence of bile, a powerful surfactant. The

ability to quickly sense and respond to these stress factors with appropriate alterations in gene expression and protein activity are crucial for survival (van de Guchte et al. 2002).

LAB have evolved a range of cellular defence mechanisms which allow them to withstand harsh conditions and sudden environmental changes. This includes chaperones to aid protein folding, catalases and superoxide dismutases to combat reactive oxygen species, proton pumps, decarboxylases and transporters to increase intracellular pH following acid exposure, and transport systems to maintain cellular osmolarity (Corcoran et al. 2008).

*E. faecalis* are robust bacteria that resist many kinds of stress factors, including heat, acid, hydrogen peroxide ($H_2O_2$), hyperosmolarity, NaOCl, UV irradiation and bile (Giard et al. 2001). Studies have shown that part of the bacterial stress response is of a general nature, involving general factors such as chaperones, that create simultaneous resistance towards several stress factors (van de Guchte et al. 2002). For inhabitants of the intestinal tract, like enterococci, several stress factors like reactive oxygen species, bile salts, osmolarity and acid are quite severe. One study showed that prolonged exposure to stress factors like bile salts, acid and heat, induced tolerance toward bile salts and acid that was maintained for a longer period compared to the tolerance toward heat, which reflects the conditions in their natural environment (Flahaut et al. 1996).

#### 1.4.2.1 The acid stress response

Intestinal bacteria have to survive the transit from the oral cavity to the intestinal tract, a journey characterized by acidic conditions. In addition, acidic end products as a result of fermentation by LAB accumulate and may locally create unfavourable conditions for many bacteria (van de Guchte et al. 2002). Acid exposure causes intracellular accumulation of protons, which reduces the intracellular pH and affects the transmembrane ΔpH. This alters the PMF, which is required for transport across the membrane (Corcoran et al. 2008). Acid stress can also cause structural damage to the cell membrane, to DNA and to proteins (van de Guchte et al. 2002). A number of proteins have been identified by two-dimentional- (2D) gel electrophoresis as contributing to the acid tolerance of LAB, including chaperones (e.g. GroEL, GroES, DnaK, ClpE and GrpE), proteins involved in handling oxidative stress (e.g. superoxide dismutase), and heat shock proteins (Frees et al. 2003). Additional mechanisms potentially associated with acid tolerance relate to DNA repair, changes in the fatty acid composition of the cells (Fozo et al. 2004), alkalization of the external environment, expression of transcriptional regulators, and alteration of metabolism and responses (Cotter et al. 2003). One important system that contributes to acid tolerance is the $F_0F_1$-ATPase. The

catalytic portion $F_1$ consists of the α, β, γ, δ and ε subunits for ATP hydrolysis, whereas the integral membrane portion $F_0$ includes the a, b and c subunits which function as a channel for proton translocation (Sebald et al. 1982). While this $F_0F_1$-ATPase normally is used to convert the PMF to ATP, it may also generate a PMF via proton expulsion at the expense of ATP. This system is crucial for maintaining pH homeostasis at low pH in LAB (Corcoran et al. 2008). For *E. faecalis*, it has been shown that the acid stress response is a rather specific response, since no cross-protection has been observed between this stress factor and others (Flahaut et al. 1996; Rince et al. 2003).

#### 1.4.2.2 The bile stress response

Each day the liver secretes one litre of bile into the gastrointestinal tract. Bile affects phospholipids and proteins of cell membranes and disrupts cellular homeostasis. Furthermore, bile induces secondary structure formation in RNA, induces DNA damage, activates enzymes involved in DNA repair, and alters the conformation of proteins (Begley et al. 2005). Mechanisms involved in the bile stress response include changes in the fatty acid composition of the cell, expression of bile salt hydrolases that deconjugate bile acids, as well as expression of chaperones and general stress proteins (Corcoran et al. 2008). Studies on *E. faecalis* have shown that adaptation to bile salts leads to cross-protection towards heat challenge and, to some extent vice versa. This is due to the fact that both types of stress induce production of heat shock proteins. On the basis of this observation it has been claimed that the bile salt and the heat shock responses are closely related in *E. faecalis* (Flahaut et al. 1996; Rince et al. 2003).

#### 1.4.2.3 The osmotic stress response

Intestinal bacteria are surrounded by nutrient solutions of various osmolarities. Exposure to osmosis results in a decrease in their cytoplasmic water activities which leads to changes in volume and pressure of the cell. Generally, bacteria respond to osmotic stress by increasing the concentration of osmolytes (Csonka et al. 1991). *E. faecalis* responds to osmotic shock provided by NaCl by an increase in intracellular potassium ion and glycine concentrations (Kunin et al. 1991).

#### 1.4.2.4 The oxidative stress response

Facultatively anaerobic bacteria such as the enterococci do not need oxygen for growth, and the presence of oxygen may be toxic. This toxic effect is attributed to reactive

oxygen species like hydrogen peroxide ($H_2O_2$) and superoxide ($O_2^-$) that attack proteins, lipids and nucleic acids (van de Guchte et al. 2002). Interestingly, phagocytic cells in the immune systems of mammals use a mechanism resulting in generation of oxidative stress in their attempts to kill pathogenic bacteria during infections (Klebanoff 1980; Hassett et al. 1989). For successful infection the bacteria have to respond to this oxidative stress. The genome sequence of *E. faecalis* V583 reveals several antioxidant defence systems. Using knockouts it has been shown that manganese-containing superoxide dismutase (MnSOD) is induced by oxygen and that this affects the survival of the bacterium inside macrophages due to a better capability to handle reactive oxygen species (Verneuil et al. 2006). Furthermore, despite the fact that *E. faecalis* is generally considered a catalase negative bacterium, a gene coding for a catalase is present in the genome of *E. faecalis* V583 and catalase production has indeed been detected when *E. faecalis* V583 was cultured in the presence of heme (Frankenberg et al. 2002). *E. faecalis* V583 also seems to have three peroxidases which all are important for the defence against $H_2O_2$: a NADH peroxidase which reduces $H_2O_2$ to water, an alkyl hydroperoxide reductase, and a protein (EF2932), which encodes a thiol peroxidase that is part of the regulon controlled by the hydrogen peroxide regulator HypR (La Carbona et al. 2007).

### 1.5 Classification, mechanism and function of Glycoside hydrolases

Glycoside hydrolases (GHs) are found in the three major kingdoms of life: archaebacteria, eubacteria and eukaryotes. They are key enzymes in carbohydrate metabolism (Henrissat 1991), where they catalyze the hydrolysis of O-, N- and S-linked glycosides. Carbohydrates (mono-, di-, oligo- and polysaccharides) play central roles in a diverse array of biological processes and are crucial for normal functioning of a cell. A key role for these enzymes concerns their ability to liberate energy by depolymerizing storage polysaccharides such as starch or by liberating fermentable sugars from other sources. They also play roles in processes as diverse as signaling events (Henrissat et al. 1997), anti-bacterial defence strategies (Williams et al. 2001; Collin et al. 2008) and in pathogenesis mechansims (Aristoteli et al. 2003; Marion et al. 2009).

GHs have during the years been classified in several different ways, based on their substrate specificities (expressed by the EC number for a given enzyme), mechanisms of action (inversion or retention of the anomeric configuration), and mode of attack (endo versus exo enzymes) (Henrissat et al. 1997). In 1991 GHs were grouped into families based on their

amino acid sequence (Henrissat 1991; Henrissat et al. 1997). The resulting Carbohydrate-Active Enzymes database (CaZy) (http://www.cazy.org) is now a widely used continuously updated database (Cantarel et al. 2009). By July 2011 a total of 117 different GH families had been classified.

### 1.5.1 The mechanism of glycosyl hydrolases

Catalysis of glycosidic bonds occurs by a nucleophile substitution at the anomeric carbon. There are two main enzymatic mechanisms of glycoside hydrolysis that differ in the stereochemical outcome of the reaction, which may either be retention or inversion of the configuration at the anomeric carbon atom of the hydrolysed glycoside (Figure 6). Both mechanisms depend on acid catalysis, and the participation of two carboxylates on the enzyme. In the inverting mechanism one of these functions as a proton donor (the catalytic acid) and the other as a base. In the retaining mechanism one functions as a proton donor and subsequently a base, while the other functions as a nucleophile (Koshland 1953; Henrissat 1991).

In the inverting mechanism, also called single displacement mechanism, the glycoside hydrolase catalyses an acid-base mechanism in a single step (Fig. 6A). The catalytic acid protonates the glycosidic oxygen, at the same time as a water molecule which is activated by the catalytic base carries out a nucleophilic attack on the anomeric carbon. This results in direct displacement of the leaving group by the nucleophilic water molecule and leads to inversion of the configuration at the anomeric carbon (Davies et al. 1995).

The retaining mechanism (Fig. 6B), also called double displacement mechanism, involves formation of a covalent glycosyl-enzyme intermediate. In the first step of the reaction, sometimes referred to as the glycosylation step, the catalytic acid protonates the glycosidic oxygen while a second carboxylate carries out a nucleophilic attack on the anomeric carbon. This results in cleavage of the glycosidic bond and formation of a glycosyl-enzyme intermediate. In the second step, which is the deglycosylation step, the glycosyl enzyme is hydrolysed by water that is activated by the former acid that has been deprotonated and now acts as a base (Davies et al. 1995). A few glycoside hydrolase families, including families GH18 and GH20, have an alternative retaining mechanism where a group on the substrate itself acts as the catalytic nucleophile (Tews et al., 1997; Van Aalten et al., 2001).

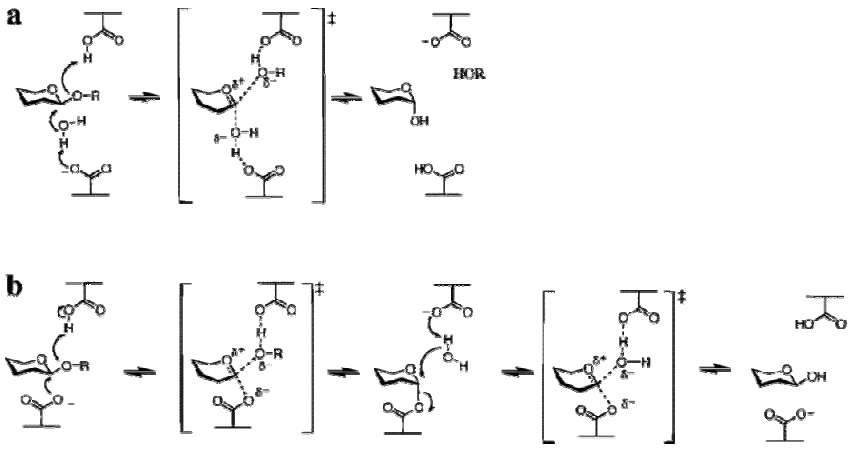


**Figure 6. Mechanisms for hydrolysis by glycoside hydrolases.** (a) The inverting mechanism, (b) the retaining mechanism. Modified from (Zechel et al. 2000).

### 1.5.2 Family 18 Glycosyl hydrolases

The GH18 family consists of catalytically active chitinases (EC 3.2.1.14), inactive chitinase-like proteins and endo-*β*-*N*-acetylglucosaminidases (EC 3.2.1.96). They perform enzymatic catalysis using a variant of the double displacement reaction. This variant involves neighboring group participation. Instead of an enzyme-derived nucleophile these enzymes utilize the *N*-acetamido carbonyl oxygen on the substrate, which acts as a nucleophile. Instead of providing the nucleophile directly, the enzyme activates and positions the *N*-acetamido group for nucleophilic attack. One carboxylate acting as catalytic acid protonates the glycosidic oxygen which, together with the nucelophilic attack by the *N*-acetamido group results in departure of the leaving group (i.e. cleavage of the glycosidic bond). In a second step, the resulting oxazolinium ion intermediate is hydrolyzed by a water molecule that is activated by the former catalytic acid that now cats as a base, see figure 7 (van Aalten et al. 2001).

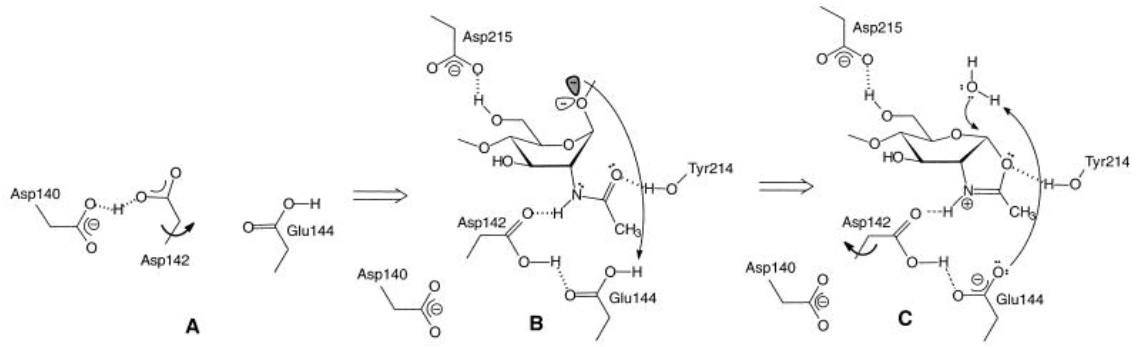



**Figure 7.** The proposed mechanism of neighboring group participation. (A) The active site contains the amino acids Asp[140], Asp[142] and Glu[144]. When there is no substrate present in the active site there is a distance between Asp[142] and Glu[144]. (B) When substrate enters the active site, Asp[142] rotates toward Glu[144] and donates a proton resulting in hydrogen bond interactions between Asp[142] and Glu[144]. This makes Glu[144] more acidic which again promotes leaving group departure. Simultaneously, the acetamido group is optimally positioned for the nucleophilic attack that promotes leaving group departure and leads to formation of the oxazolinium ion intermediate. (C) The resulting oxazolinium ion intermediate is hydrolyzed by a water molecule that is activated by the now de-protonated catalytic acid while Asp[142] rotates back to its original position sharing a proton with Asp[140] (van Aalten et al. 2001).

Enzymes belonging to the GH18 contain several conserved carboxylates that are critical for activity (Synstad et al., 2004) and that are found in a the diagnostic D-X-X-D-X-D-X-E sequence motif (D corresponds to aspartic acid, E to Glutamic acid and X to any amino acid). The glutamate at the end of this motif acts as the catalytic acid.

### 1.5.2.1 Endo-*β*-*N*-acetylglucosaminidases and glycoproteins

Endo-*β*-*N*-acetylglucosaminidases belong to family GH18 and are known to hydrolyse the N-linked glycans of glycoproteins. Some *β*-*N*-acetylglucosaminidases can act weakly as exochitinases, by cleaving monosaccharide units from the non-reducing ends of chitin chains (Gooday 1990). It has been proposed that the ability to hydrolyse the carbohydrates of glycosylated proteins provides bacteria with an important nutrient source that may be crucial for proliferation *in vivo* (Tarelli et al. 1998; Byers et al. 1999; Roberts et al. 2000; Roberts et al. 2001; Sanchez et al. 2010; Ruiz et al. 2011).

Glycosylation is the most common of all post-translational modifications of eukaryotic proteins and is crucial in many physiological processes, including signaling, cellular

differentiation and adhesion (Rudd et al. 2001; Moran et al. 2011). Protein glycosylation was first demonstrated in eukaryotes in the late 1930s, and was thought to be absent from prokaryotes until the 1970s when glycoproteins was detected on the surface layers of the archaea *Halobacterium salinarum* (Mescher et al. 1974) and of two *Clostridium* species (Sleytr et al. 1976) . Most bacterial glycoproteins appear to be either associated with the surface of the organisms as in pili or flagella where they are involved in cell-cell interactions between the bacteria and eukaryotic cells, or to be secreted into the environment (Benz et al. 2002). More than two/third of the eukaryotic proteins are predicted to by glycosylated (Apweiler et al. 1999) which includes many proteins of the immune system, such as key molecules involved in antigen recognition, and proteins covering the epithelial cell surface and the mucous layer (Rudd et al. 2001; Moran et al. 2011).

The carbohydrate unit is normally linked to the protein by either O-glycosylation or N-glycosylation (Nothaft et al. 2010). In O-glycosylation (Figure 8) the carbohydrate is attached via the hydroxyl oxygen of Serine (Ser) or Threonine (Thr). The presence of such an amino acid is the only requirement for O-glycosylation; no consensus sequence motif has been identified (Benz et al. 2002). O-linked glycans are usually heterogeneous and are classified by their core structure, which may include glucosamine, xylose, galactose, fucose or mannose as the initial sugar bound to the Ser/Thr residues. However, the most common type of O-linked glycan contain an initial GalNAc residue, and these glycans are referred to as mucin-type glycans (Spiro 2002). In N-glycosylation (Figure 8) the carbohydrate is attached via the nitrogen of an Asparagine (Asn) residue. N-linked glycosylation has a consensus sequence motif, Ser/Thr-X-Asn, where X may be any amino acid exept proline (Benz et al. 2002). N-linked glycans contain a core pentasaccharide of two *N*-acetylglucosamines (GlcNAc) and three mannoses, and are classified into three subgroups on the basis of the nature and position of sugar residues added to the common core pentasaccharide:

1) The high-mannose type contains additional α-linked mannose residues.
2) The complex type glycans have several antennae containing galactose, fucose, additional GlcNAc-residues, and one or more sialic acid residues at the end of the antennae.
3) The hybrid type glycans possess a combination of high-mannose and complex type branches (Michalski 2005).

Endo-*β*-*N*-acetylglucosaminidases catalyze the hydrolysis of the glycosidic linkage between the two GlcNAc residues in the ($Man_3$-$GlcNAc_2$)-core of an N-linked glycoprotein. During hydrolysis one GlcNAc residue of the core pentasaccharide remains attached to the protein and the rest of the oligosaccharide is released intact (Morelle 2005).

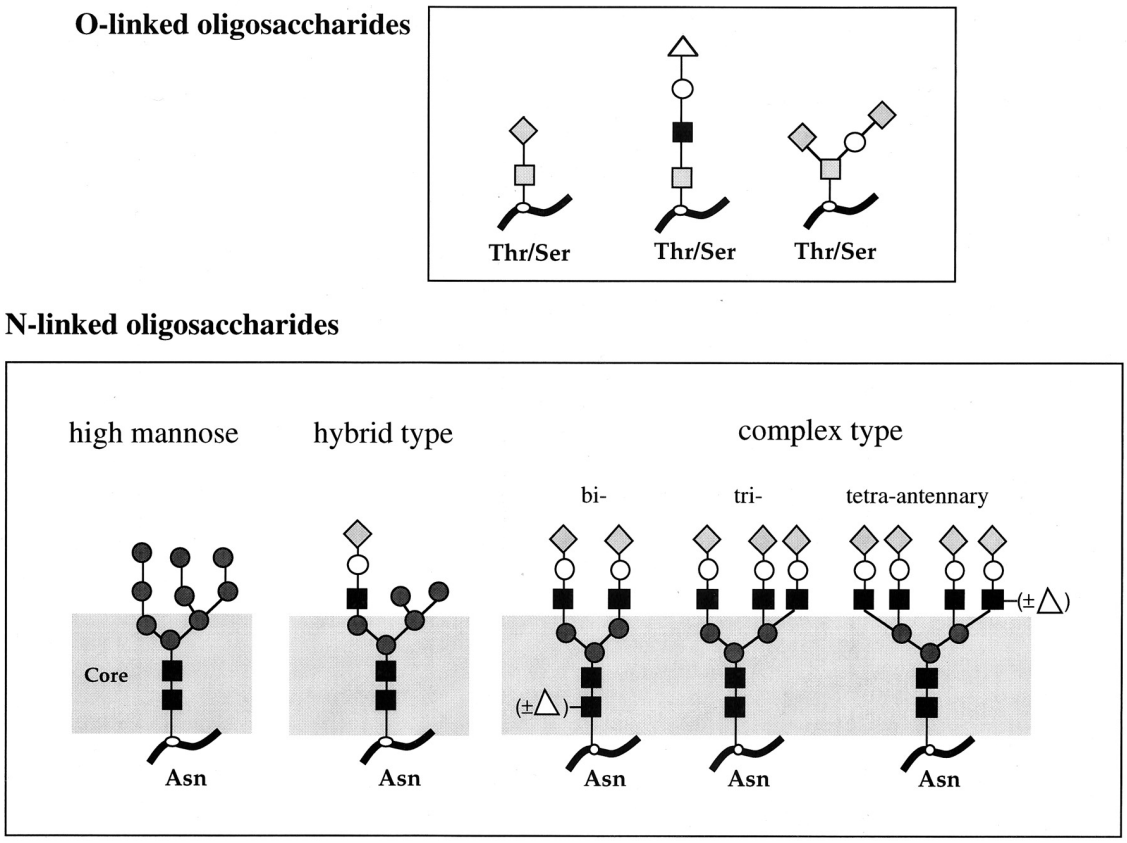



**Figure 8.** Structures of O- and N-linked oligosaccharides. The top figure shows some examples of O-linked oligosaccharides. The only requirement for O-linked glycosylation is that the carbohydrate is attached via the hydroxyl oxygen of Ser or Thr. The bottom figure shows the different sub-classes of N-linked oligosaccharides where the carbohydrate is attached via the nitrogen of an Asn residue; high-mannose, hybrid and complex-type. The common core pentasaccharide is indicated by the light grey boxes. The symbols correspond to; ▩, GalNAc; ■, GlcNAc; ○, galactose; ●, mannose; △, fucose ◆, SA; ~, polypeptide chain (Durand et al. 2000).

### 1.5.2.2 Chitin and Chitinases

Chitin is an insoluble linear $\beta$-1, 4-linked polymer of GlcNAc, that is widely distributed on Earth (Figure 9). After cellulose, chitin is the most abundant organic compound with an annual production amounting to $10^{10}$-$10^{11}$ tons (Gooday 1990). It is a major component in cell walls of fungi and algae, in the exoskeletons of arthropods and in shells and radulae of molluscs such as crabs and shrimps (Bhattacharya et al. 2007). Chitin is a solid structural

component, and its main role is protection of the organism or cell from chemical and mechanical stress (Gooday 1990).

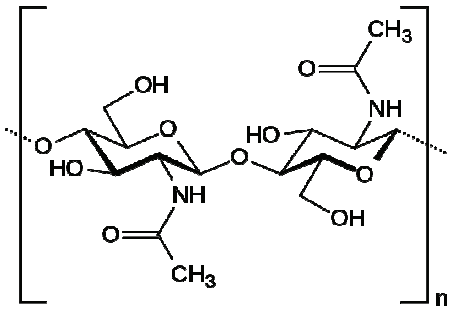



**Figure 9. The structure of chitin.** Chitin consists of repeating units of *β*-1,4 linked *N*-acetylglucosamine (GlcNAc). The monosaccharides are rotated 180° relative to each other, meaning that the repeating unit is a dimer.

Three types of chitin occur in nature and these are defined by the arrangement of the individual chitin chains. The α-form is mainly found in crab and shrimp shells, and is the most common form. The GlcNAc chains are arranged in an anti-parallel fashion, which enables formation of many strong hydrogen bonds, resulting in a densely packed, very stable and rigid material (Minke et al. 1978). The β-form is mainly obtained from molluscs and in this case the GlcNAc chains are arranged in a paralell fashion. Packing of the chains is more loose compared to α-chitin (Gardner et al. 1975). The last form, γ-chitin, is the least abundant chitin-type, and in this form two parallel strands alternate with a single anti-parallel strand. The mechanical traits of chitin make it insoluble and resistant to chemicals. Although chitin is widely distributed on earth and produced in large amounts, chitin does not accumulate in different ecosystems, suggesting that chitin-degradation in nature is an efficient process.

Chitin and, in particular, it's partially deacetylated derivative chitosan have a number of applications in a wide variety of fields. These applications include medical applications related to antimicrobial and antitumor activities, use as dietary fiber, and use in wastewater treatment, drug delivery and wound healing, as reviewed by (Khoushab et al. 2010).

Chitinases (EC 3.2.1.14) are hydrolytic enzymes, belonging to the GH18 and GH19 families. They may be divided into two major categories: endochitinases which cleave chitin at random positions on the polysaccharide chain, and exochitinases which cleave chitin from the ends of the polysaccharide chain (Horn et al. 2006; Eijsink et al. 2010; Horn S. J. 2006). Chitinases are present in a range of organisms including viruses (Hawtin et al. 1995), fungi

(Kuranda et al. 1991; Xia et al. 2001), insects (Merzendorfer et al. 2003), higher plants (Yamagami et al. 1993; Taira et al. 2005), mammals (Bussink et al. 2007) and a wide range of bacteria, including *Serratia marcescens* (Fuchs et al. 1986), *Listeria* species (Leisner et al. 2008), *Streptomyces* species (Yano et al. 2008), *Aeromonas* species (Ueda et al. 1995; Lan et al. 2006), *Bacillus cereus* (Kuo et al. 2006), *Pseudomonas aeruginosa* (Folders et al. 2001), *Lactococcus lactis* (Vaaje-Kolstad et al. 2009) and *E. faecalis* (Leisner et al. 2009). Because chitinases are produced by such different organisms, their physiological functions may differ from one organism to the next. Organisms that have chitin as part of their body structure, like insects and crustaceans, require chitinases to degrade old cuticle as body size and shape change. Organisms that consume chitin-containing organisms need chitinases to digest chitin into absorbable metabolites in order to derive energy. Organisms such as plants that are prone to infection by chitin-coated organisms express chitinases to degrade the protective outer layer of the invading pathogen.

Bacteria containing a chitinolytic machinery are ubiquitous, and the main function of bacterial chitinases is to process and digest GlcNAc-containing macromolecules in order to obtain nutrients (Bhattacharya et al. 2007). Most studies have therefore been conducted in order to characterize chitinases from bacteria living in soil and water where chitin is present (Gooday 1990). Generally there is little information on the ability of LAB to degrade chitin and chitin derivates. A screening for α- chitinolytic activity using a basic chitin medium agar lawn showed that among nine genera of LAB only species of the *Carnobacterium* genera were able to hydrolyse α-chitin (Leisner et al. 2008). However, it had previously been shown that *Lc. lactis* is able to grow on minimal media utilizing N-acetylglucosamine oligomers as main carbon source (Chen et al. 2002). Recent work has shown that two other LAB, *Lc. lactis* (Vaaje-Kolstad et al. 2009) and *E. faecalis* V583 (unpublished results; (Vaaje-Kolstad 2011) are able to hydrolyse both α- and β-chitin especially in the presence of a chitin-binding protein.

Interestingly, bacterial chitinases have been linked to virulence (DebRoy et al. 2006; Chaudhuri et al. 2010). In *Legionella pneumophila* which causes pneumonia it has been shown that a chitinase is required for optimal survival of the bacterium in the lungs. This requirement is not directly linked to chitinaseactivity, but rather to the effect the enzyme has on the host immune response. A knockout-study showed that when the chitinase was present the host developed pneumonia, however, when the chitinase was absent the bacterium was removed by the immune system (DebRoy et al. 2006). The same phenomonon was observed

in another study which showed that two chitinases from *L. monocytogenes* contributed to virulence in bloodstream infections in mice (Chaudhuri et al. 2010).

Despite the absence of chitin in mammals, humans express two enzymatically active chitinases, chitotriosidase and AMCase, and three enzymatically inactive chitinase-like proteins; CHI3L1, OVGP1 and CHI3L2 (Funkhouser et al. 2007; Shuhui et al. 2009). Both human chitinases and chitinase-like proteins are suggested to play a central role in T-helper 2-mediated inflammation and allergic diseases like asthma (Elias et al. 2005), and they have been used as indicators of inflammation and cancer (Kzhyshkowska et al. 2007).

In recent years, chitinases have gained interest because they have various biotechnological applications, including use in agriculture to control plant pathogens, use as antimicrobial or insecticidal agents, and use for bioconversion of chitin. A recent study has shown that an endochitinase from the most widely used bioinsecticide worldwide, *Bacillus thuringienis*, is able to generate chitin-derived oligosaccharides that have antimicrobial activity against food-borne pathogens (Ortiz-Rodriguez et al. 2010). More general, the enzymatic production of chito-oligosaccharides is of considerable interest (Aam et al. 2010).

### 1.5.3 Family 33 of the Carbohydrate-binding modules

Carbohydrate binding modules (CBMs) are non-catalytic sugar-binding protein domains that are classified according to amino acid similarity in the CAZy database. Currently (July 2011) CBMs are classified into 64 different families and they have been identified in a range of different organisms (Cantarel et al. 2009). The majority of chitinases are equipped with CBMs that are thought to increase substrate affinity. The CBMs are thought to enhance catalytic efficiency of the enzyme for the insoluble substrate by bringing the catalytic module into prolonged and intimate contact with the substrate (Hashimoto et al. 2000).

Family 33 CBMs (CBM33) have chitin-binding properties and are mainly found in bacteria and viruses. CBM33s are special, because they normally occur as single domain proteins or are linked to another CBM, whereas linkage to a catalytic domain (a GH) is rare. It has been suggested that, in additon to binding to chitin (Suzuki et al. 1998), chitin binding proteins may bind other GlcNAc residues such as those present on the surfaces of intestinal mucins and epithelial cells. It has been reported that CBM33 chitin binding proteins from *Lactobacillus plantarum* (Sanchez et al. 2011) and *Vibrio cholerae* (Kirn et al. 2005; Bhowmick et al. 2008) enhance bacterial colonization of the intestine through the binding of mucin. It has also been shown that the chitinolytic machinery of *L. monocytogenes* , which includes a CBM33, contributes to bloodstream infection in mice (Chaudhuri et al. 2010).

Chitin binding proteins are also frequently found in insect viruses (Li et al. 2003), were they may contribute to survival and/or infection of their chitin-containing hosts. As to the bacterial CBM33s, current data indicate that they may play roles in processes varying from chitin conversion, via host-microbe interactions to bacterial pathogenesis.

The first structure of a CBM33 was solved in 2005 for a protein from *S. marcescens* called CBP21 (Vaaje-Kolstad et al. 2005) (Figure 10). The structure consists of two β-sheets, one with three strands, and the other with four strands. These two β-sheets form a β-sandwich with a compact fibronectin III-type fold. Previously, it had been shown that conserved aromatic residues were important for the binding of CBMs to their substrates, but the structure of CBP21 showed that all aromatic residues, except one, are located in the core of the protein. One of the surfaces of CBP21 contains several highly conserved residues that are mostly of a polar nature. Using site-directed mutagenesis, the importance of these residues for substrate-binding was confirmed, showing that binding to chitin involves specific polar interactions, mediated by a cluster of conserved residues on the surface of the protein (Vaaje-Kolstad et al. 2005).

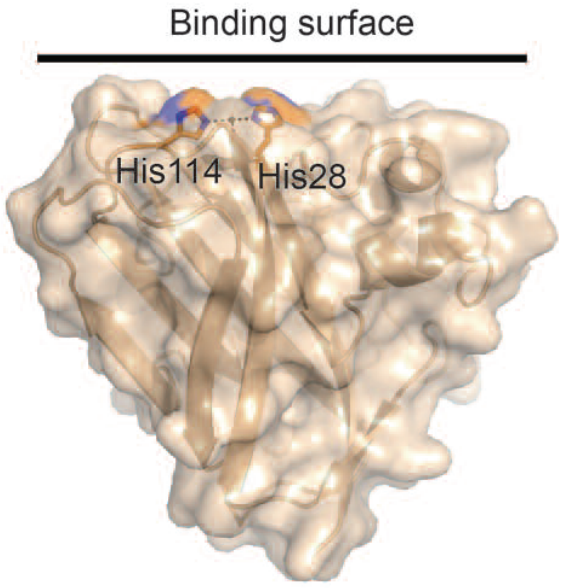



**Figure 10. The structure of CBP21**. The active site, presumably close to the two histidine side chains that are shown, is part of a flat substrate binding surface, and is not located in a groove or a pocket as seen in other enzymes. The conserved histidines are part of the active site and are important for metal binding (Vaaje-Kolstad et al. 2010).

In a further study, also from 2005 (Vaaje-Kolstad et al. 2005) it was shown that CBP21 can potentiate the activity of chitinases and it was speculated that this could be due to

the CBP21 somehow disrupting the crystalline structure of the substrate. Exactly how CBP21 would achieve this remained unknown until it was discovered, in 2010, that CBP21 is an enzyme that cleaves glycosidic bonds by a novel metal-dependent mechanism that involves hydrolysis and oxidation (Vaaje-Kolstad et al. 2010). In contrast to other enzymes, CBP21 has no substrate-binding groove or pocket; instead the active site is part of an almost flat surface (Figure 10). Two histidines are part of the catalytic center, and comprise a potentially promiscous metal binding site. Vaaje-Kolstad et al (2010) concluded that CBP21 binds to the surface of crystalline chitin and introduces cleveage of a glycosidic bond, using molecular oxygen, water and an external electron donor. The reaction results in oxidation of one of the new chain ends, namely the end that would have been the new reducing end if the CBP21 had been a normal hydrolytic enzyme (Figure 11). It is quite conceivable that the introduction of this type of chain breaks on the crystal surface leads to disruption of crystalline packing and increased substrate accessibility (Vaaje-Kolstad et al. 2010).

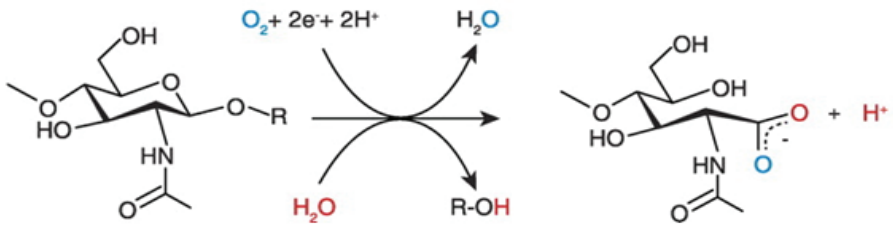



**Figure 11.** The enzymatic reaction catalyzed by CBP21. The enzyme cleaves glycosidic bonds using molecular oxygen and water. The resulting oxidized product contains one oxygen from molecular oxygen (blue), and one oxygen from water (red) (Vaaje-Kolstad et al. 2010).

It remains to be seen whether all members of the CBM33 have this catalytic activity and one might also wonder about which types of substrates may be preferred by the various members of the family. Very recently, it has been shown that *Streptomyces coelicolor* produces a CBM33-containing protein, CelS2 that is active on cellulose. It produces oxidized cello-oligosaccharides and acts synergistically with cellulases (Forsberg et al. 2011).

Clearly, it will be interesting to evaluate the occurrence of CBM33s in LAB and their possible roles in host-microbe interactions in the light of these recent findings.

## 2. OUTLINE OF THE THESIS

The thesis can be divided into two parts, the main purpose of both parts being to identify and study proteins that are involved in environmental responses of *E. faecalis* V583. In the first part (Paper I and Paper II) proteomics technologies were used to identify proteins that were differently expressed in response to bile (Paper I), and to find surface proteins that are potentially important for interaction and communication with the environment (Paper II). In the second part (Paper III and Paper IV) some proteins involved in carbohydrate degradation (a response to the presence of carbohydrates in the environment) were purified and characterized more deeply.

Paper I describes how bile affects protein expression in *E. faecalis* V583, which is an important environmental factor. Each day, intestinal bacteria are exposed to the deleterious effects of bile, and this represents one of the major stress factors that both harmless and pathogenic bacteria have to cope with in order to survive in the gastrointestinal tract.The changes in the protein profiles were analyzed using 2D-gel electrophoresis, and differently expressed proteins were identified using MALDI-TOF mass spectrometry and the Mascot server. In Paper II proteins that are localized extracellularly were identified by digesting surface located proteins using enzymatic shaving of living cells. The released peptides and protein fragments were analyzed and identified by LC-MS/MS and the Sequest server. Proteins located at the surface of the bacterial cell are important determinants of bacterial behavior, responses and pathogenicity, because these proteins interact with the surroundings.

Paper III and IV describe studies on three proteins that were selected as targets for characterization. Paper III concerns an enzyme that may help the bacterium to harvest carbohydrates from glycoproteins in the host organism, whereas paper IV concerns a putative chitinolytic system, the expression of which had been shown to be regulated in response to various stress factors. Paper III describes an endo-*β*-*N*-acetylglucosaminidase (EF2863) from *E. faecalis* V583 named *Ef*Endo18A that belongs to family 18 of the glycoside hydrolases. GH18 enzymes are known to catalyze the hydrolysis of the glycosidic linkages between two *N*-acteylglucosamines, as they occur in chitin and in some of the glycans commonly found in glycosylated proteins. Paper IV describes the chitinolytic machinery of *E. faecalis* V583comprising a chitinase (EF0361; *Ef*Chi18A) and an enzyme currently referred to as “chitin binding protein” and (erroneously) classified as a member of family 33 of Carbohydrate Binding Modules (EF0362; *Ef*CBM33A). Interestingly, *E. faecalis* V583 has these genes despite the absence of chitin in mammals. Transcriptome studies have shown that

all three of these proteins (EF2863, EF0361 and EF0362) are up-regulated when the *E. faecalis* V583 grows in blood and urine, where the bacterium frequently causes infection. EF0361 and EF0362 were also identified with one peptide each as part of the surface proteome in Paper II. There have been speculations that these proteins may have roles during colonization and invasion of the intestinal tract of the human host. The enzymological characterizations described here help in future investigations related to these speculations.

## 3. MAIN RESULTS AND DISCUSSION

*E. faecalis* is a robust bacterium that survives under harsh conditions. Such robust enterococci have become a problem in hospital environments were they are known to cause several infections e.g. in blood and the urinary tract. The presence of a pathogenicity island and plasmids encoding genes for antibiotic resistance, as well as the ablity to survive on abiotic material complicate the treatment of enterococcal infections. Increased understanding of how these bacteria survive under unfavorable conditions and how they interact with their environment, as well as identification of potential target proteins for drug design, are necessary in order to develop and establish new alternative treatments. In addition to resistance toward many antibiotics, *E. faecalis* is equipped with several mechanisms for resistance toward environmental factors like oxidative stress, bile, detergents and salt. The presence of these mechanisms and the resulting metabolic flexibility make *E. faecalis* an interesting model system for studying bacterial stress-responses.

One main stress factor for bacteria that pass or reside in the intestinal tract is the presence of bile. The deleterious actions of bile are not limited to its detergent action, but also include other effects such as lowering the pH and creation of oxidative and osmotic stress. Responses to this complicated stress factor have been studied in several bacteria at both the proteome and the transcriptome level, for example in lactobacilli (Bron et al. 2004; Lee et al. 2008; Whitehead et al. 2008; Burns et al. 2010; Wu et al. 2010), bifidobacteria (Sanchez et al. 2005; Sanchez et al. 2007), enterococci (Flahaut et al. 1996; Flahaut et al. 1996; Giard et al. 2001; Solheim et al. 2007) and propionibacteria (Leverrier et al. 2003). Despite that the enterococcal bile stress response has been analyzed previously at the proteomic level, there is still a need for (much) more work because different enterococci may have different strategies for overcoming bile stress. In a recent study (Hamon et al. 2011) a comparative proteomic analysis showed that bile tolerance properties of nine *Lb. plantarum* strains, revealed large differences between the strains. This may also be the case for enterococci, especially if one would compare commensals versus pathogens. An additional reason for further proteome work is that this technology has developed rapidly in recent years and that it, generally, is difficult to “see all” in one study.

Paper I represents the first proteomic study after the genome sequence of *E. faecalis* V583 was completed, and differs from several earlier proteome studies in that we used a time-course experiment in order to identify both immediate responses to bile exposure and more

long-term effects. The results were compared to transcriptomic data acchieved under the same growth/stress conditions (Solheim et al. 2007).

Under the conditions used there was a clear effect of the bile on growth rate of the bacterium that became visible at about one hour after bile exposure. The stressed culture did grow well and with a stable rate during the rest of the cultivation, albeit at lower rate and to lower density than the non-stressed culture. This indicates that the cells were stressed to appropriate levels. 2D-gel electrophoresis was conducted to separate the proteins according to pI and molecular mass. Since initial analyses covering the whole pI- range had shown that the vast majority of proteins had pI values in the range between 4 and 7, this pI range was used for the experiments. Approximately 500 proteins were visualized on the gel, of which 59 were considered to show significant differences in abundance between the stressed and the non-stressed cultures. Of these 59 proteins, 53 unique proteins were identified by MALDI-TOF mass spectrometry.This number represents about 2 % of the total proteome of *E. faecalis* V583. A limitation of the 2D-gel electrophoresis technique is that insoluble proteins, proteins located in the bacterial cell wall or membrane, and the majority of secreted proteins will not be detected. In addition there may be limitations concerning the sensitivity during the staining procedure. This represents drawbacks compared to gel-free approaches (such as those used in paper II) and it seems evident that the total number of regulated proteins is higher than the 59/53 that were identified in Paper I. However, for quantitative experiments, for example when comparing different conditions such as in Paper I, gel-based methods are a reliable and widely used technique.

The majority of differently expressed proteins were down-regulated in the presence of bile. The number of differently expressed proteins increased during the time course experiment, which may seem to correlate to some extent to that effects on growth were not observed before about one hour after administration of the bile. The identified proteins belonged to nine different functional classes, and one class, containing proteins of the fatty acid and phospholipid metabolism was clearly overrepresented. All of the proteins detected in this class were down-regulated, demonstrating a clear and coordinated physiological response. Down-regulation of genes involved in fatty acid and phospholipid metabolism was also dominant at the transcriptome level (Solheim et al. 2007). Interestingly, such a response has also been detected by studying membranes. A study on *Lactobacillus reuteri* showed a significant reduction in the phospholipid fraction of total membrane lipids and a lower ratio of saturated vs un-saturated fatty acids as a response to bile salt exposure (Taranto et al. 2003).

Another study showed enrichment in mono-unsaturated and longer chain fatty acids in the membrane that could reduce the permeability of bile acids (van de Guchte et al. 2002). As the bacterial membrane is among the first targets of the bile, a change in the fatty acid composition seems like a logical response.

*E. faecalis* is one of few prokaryotes that may enhance growth under various conditions through oxidative phosphorylation using a PMF established by electron transport over the bacterial membrane (Leblanck 2006). Interestingly, two of the proteins that are upregulated in response to bile are involved in oxidative phosphorylation. One of these, EF1499, is an ATPase which couples hydrolysis of ATP to the translocation of protons across the bacterial membrane. Maintenance of the PMF has been shown to be important during bile exposure due to its contribution to sustain cellular homeostasis, which is required to withstand the osmotic stress caused by bile. The importance of ATPase activity upon bile stress was confirmed by transcriptomic data, which showed upregulation of nine ATPases upon bile stress. The fact that ATPases are membrane-bound proteins may explain why only one ATPase was identified at the proteomic level, illustrating one of the drawbacks of 2D-gel electrophoresis. Increased expression of proteins related to maintenance of the PMF has also been demonstrated in other bacteria exposed to bile (Sanchez et al. 2005; Bron et al. 2006; Sanchez et al. 2006).

In contrast to other bacteria, *E. faecalis* down-regulates proteins involved in carbohydrate metabolism as a response to bile. Unlike other LAB, enterococci are not dependent on carbohydrates to obtain energy, as other substrates, such as glycerol, lactic acid, malic acid and citric acid may serve as an energy source. In addition, to obtain energy they may produce a PMF via electron transport to drive oxidative phosphorylation (Leblanck 2006).

By using oxidative phosphorylation the production of potent oxidants like superoxide ($O_2^-$) and hydrogen peroxide ($H_2O_2$) may increase. Since, as described in the introduction of this thesis, *E. faecalis* is able to tolerate oxidative stress, the choice of using oxidative phosphorylation over other pathways will not be detrimental. Interestingly, two proteins, DivIVA (EF1002) and glutathione reductase (EF3270), which play roles in managing oxidative stress, were identified as upregulated in this study, compatible with the idea that *E. faecalis* V583 establishes increased metabolic fluxes towards oxidative phosphorylation as a response to bile while protecting itself.

Bacteria are continously responding to changes in their environment in order to colonize and proliferate in the host, or, for pathogens, to cause an infection. It has been shown that as a response to bile, pathogens like *Salmonellae*, *Shigella* and *Vibrio* spp. express virulence factors, because the presence of bile in their environment correlates with the places where infections are initiated (Gunn 2000). As most of the virulence factors of *E. faecalis* V583 are extracellularly located proteins (Paulsen et al. 2003), they may have been overlooked in the present study. Generally, the approach of Paper I only gives some fragments of the puzzle because of the limitations of the 2D-electrophoresis approach discussed above. Several other proteins including several hypothetical proteins were found to be differently expressed, without providing enough pieces of the puzzle to make general statements about possible overall physiological responses (see Paper I for further discussion). However, apart from the clear effects on fatty acid and phospholipid metabolism, the data do seem to indicate that bile redirects metabolic fluxes in *E. faecalis* V583 in such a way that reactive oxygen species are produced. It is interesting to note that these reactive oxygen species may inhibit the growth of other bacterial species, while the producing *E. faecalis* V583 seems to be well protected. This may give a competitive advantage under otherwise stressful conditions.

The main object of the study described in Paper II was to identify extracellular proteins in *E. faecalis* V583, which could be targets for further studies. An additional goal was to test and develop methods for specifically detecting surface proteins. Proteins that have an extracellular localization, sense environmental factors and directly interacts with the environment of the bacterium. They are important both in communication between bacterial cells but also during interaction between bacterial and host cells, and hence essential for survival, colonization, and infection. Because of the extracellular location they may also interact with drugs and antibodies, representing interesting candidate targets for active or passive immunization. The study described in paper II is one of the first examples of a proteome study specifically targeting enterococcal surface proteins.

Extracellular proteins of *E. faecalis* V583 were identified after proteolytic shaving of the intact bacterial cells by using trypsin or trypsin bound to agarose beads, whereas non-proteolytic shedding was assessed in control experiments with no enzymatic treatment. The difference between the two proteolytic approaches is that while trypsin can reach deeper into the bacterial cell wall and cleave off proteins that are located both within the cell wall and

even the membrane, trypsin beads will only have access to proteins or parts of the proteins that are sticking out of the bacterial cell. An advantage of this shaving approach is that the proteolytic enzymes only will have access to surface-exposed proteins, limiting the risk of intracellular contamination. However, proteins that are predicted to be of intracellular nature have been identified as part of extracellular proteomes in all studied Gram-positive bacteria so far. This may be due to cellular lysis during the growth or harvesting phase or to other processes discussed further below. As many as 40-50 % of proteins identified in studies of the exoproteome have been reported to be of intracellular nature (Dreisbach et al. 2011).

In the study of Paper II, tryptic fragments from shaved and released proteins were analyzed by liquid chromatography (LC) coupled to a LTQ-Orbitrap mass spectrometer. One of the criteria the proteins had to meet in order to be considered as a significant hit is that the proteins had to be identified by at least two different peptides in total, and that at least one peptide was detected in at least two of the independent parallels (four parellels in total). This criterium may seem a bit strict, but was applied to increase the quality of the study and to avoid artifacts. The LC-MS/MS method is not a quantitative method and is a method with notorious reproducibility problems, hence these strict criteria that also imply that a larger part (namely two peptides) needs to be detected, than when one peptide had been sufficient. Other studies have also included proteins that were identified by only one peptide (Rodriguez-Ortega et al. 2006; Benachour et al. 2009). If the same criterium had been applied to the study of Paper II, more than 100 additionally surface located proteins would have been identified. Among these proteins are a predicted N-terminal anchored chitinase (EF0361) and a predicted secreted chitin binding protein (EF0362) that have been characterized in Paper IV.

Using these strict criteria, 58 proteins were identified from cells treated with free trypsin, whereas treatment with trypsin beads yielded 29 proteins, and 16 proteins were identified without any enzymatic treatment. Of these proteins, 69 were unique, and most detected proteins with predicted extracellular locations were involved in cell wall synthesis and maintenance and cell-cell communication.

The most populated group of proteins was proteins predicted to have intracellular locations. Of these 33 proteins (i.e. almost 50 % of the total number of identified proteins), at least 20 have been identified as part of the extracellular proteome in previous studies on exoproteomes of Gram-positive bacteria (Schaumburg et al. 2004; Trost et al. 2005; Severin et al. 2007; Tjalsma et al. 2008; Beck et al. 2009; Benachour et al. 2009; Dumas et al. 2009; Sanchez et al. 2009; Dreisbach et al. 2011). Most (23) of the proteins that have a predicted

intracellular location were only identified after the trypsin treatment, indicating that they are on the outside of the cells and have affinity to the cell wall. Clearly, if the detection of these proteins were to be the result of cell lysis, the data indicate that they reassociate with intact bacterial cells. Interestingly, for several such external cytoplasmic proteins, there is evidence that they have an additional extracellular function and these proteins are sometimes referred to as "moonlighting proteins" (Jeffery 1999). It has been suggested that such "anchorless proteins" may play roles in adhesion and invasion (Chhatwal 2002).

36 proteins of the identified proteins were predicted to have an extracellular localization. Of these proteins, 17 were lipoproteins and this class of surface proteins was clearly overrepresented. Lipoproteins have a variety of functions in the bacterial cell, including roles in nutrient uptake, signal transduction, conjugation, adhesion and sporulation. They may also contribute to antibiotic resistance, and to the transport and folding of proteins. In many Gram-positive bacteria lipoproteins are essential for the interaction between the bacterium and its host (Kovacs-Simon et al. 2011; Reffuveille et al. 2011), and several lipoproteins are being considered as vaccine candidates (Ayalew et al. 2009; Sardinas et al. 2009). Most of the lipoproteins identified in the study described in Paper II are proteins with unknown function.

Five of these 36 proteins are predicted to be secreted proteins. Althought secreted proteins are not directly linked to physical interaction between bacterial and host cells, they may be important during virulence. This included the gelatinase (EF1818) described in section 1.3 that was identified after all treatments. The fact that only five proteins with this predicted localization were detected is a positive result because secreted proteins are not expected to be identified in this experiment. It is clear though that studies on secreted proteins would be of major interest, and such studies are currently in progress. Preliminary results show that more than 100 proteins (both one- and two-peptides proteins) are secreted and released by *E. faecalis* V583 growing on YT-medium (G. Mathiesen et al., unpublished results).

Five N-terminally anchored proteins were identified, and these were mainly predicted to be involved in cell wall synthesis. Four LPxTG proteins were identified, one of these (EFA0052) was the plasmid encoded surface exclusion factor Sea1 involved in conjugation, and there are indications that one of the other (EF2860) may contribute to antibiotic resistance. It has been shown that LPxTG proteins in pathogens like streptococci (Mora et al. 2005; Seepersaud et al. 2005; Gianfaldoni et al. 2007) can be utilized for passive or active

immunization and this could also be the case for *E. faecalis*. In light of this, it is disappointing that only four (of a predicted total of 38) such proteins were found. Three (of totally 588) of the identified proteins are predicted to be transmembrane proteins. It has previously been shown that membrane proteins are less available for the shaving approach than lipoproteins and proteins that are covalently attached to the cell wall, due to their lack of surface-exposed domains (Dreisbach et al. 2011). The data described in Paper II show the same trend. A recent study specifically aimed at mapping the membrane proteome of *E. faecalis*, showed that it contained proteins involved in cell homeostasis, virulence and antibiotic intervention (Maddalo et al. 2011). The last group cell-wall attached proteins, with the lowest number of identified proteins, were proteins that are non-covalently attached to the cell wall. Only two such proteins were identified.

There were large differences in the amounts of peptides detected for the various identified proteins. These numbers are in part related to protein size, protein accessibility and protein proteolytic susceptibility but may of course also be related to protein abundance. One of the most prominent proteins was the lipid-anchored fumarate reductase, EF2556. This protein was identified by 2189 peptide hits in total, with 40 unique hits. As described earlier, *E. faecalis* may obtain energy by producing a PMF via electron transport to drive oxidative phosphorylation, resulting in energy gain for the bacterial cell (Huycke 2002; Leblanck 2006). Fumarate reductase is a part of this respiratory chain and contributes to generation of reactive oxygen species, which subsequently may contribute to the virulence of this bacterium (Huycke et al. 2002). As described in Paper I proteins involved in oxidative phosphorylation were up-regulated during bile stress. The results from Paper I and Paper II may indicate that oxidative phosphorylation is a highly abundant pathway in *E. faecalis* that is in general use and not only used as a response to stress.

The study of Paper II yielded only a very limited number of proteins that may be related to virulence, despite the fact that surface-exposed proteins are crucial during bacterial infection. It has been shown that *E. faecalis* is able to sense its environment and thereby regulate the expression of virulence factors (Hew et al. 2006). The hypothesis is that virulence-related proteins only are expressed when the bacterium senses the necessity to express them e.g. as a response to target cells like erythrocytes. In the absence of target cells or any competing bacteria, there is no reason for the bacteria to use energy on expressing proteins that only have a role during infection. The fact that the bacteria in the present study were cultivated in normal growth-media may have depressed the expression of proteins

contributing to infection. This would explain why most identified surface proteins have functions that are part of regular bacterial life. It would have been interesting to cultivate the bacteria in media containing e.g. bile or blood and monitor the resulting changes in the occurrence of surface-exposed proteins.

Many (15) of the identified proteins with predicted surface-locations have unknown functions, illustrating the gaps in current knowledge of the biology of *E. faecalis* V583. Obviously, these proteins are interesting targets for further research on this bacterium.

*E. faecalis* is known to have endo-*β*-*N*-acetylglucosaminidase activity and Paper III describes a study on one of the two proteins from *E. faecalis* V583 that could be responsible for this activity, EF2863. While EF2863 is predicted to be a secreted protein, the other candidate, EF0114 is likely to be N-terminally anchored. Indeed, EF0114 was identified as part of the extracellular proteome (Paper II) after treatment with trypsin beads for one hour, but the identification was based on only one peptide, and was hence not considered significant, according to the strict criteria set for identification. Since EF0114 shows 99 % sequence identity with a previously characterized endo-*β*-*N*-acetylglucosaminidase, EndoE, from *E. faecalis* HER1044 (Collin et al. 2004), the study of paper III focused on EF2863, renamed *Ef*Endo18A.

Analyses of band shifts on SDS-PAGE gels and mass spectrometry data showed that *Ef*Endo18A is able to hydrolyse N-linked glycans of the high-mannose and hybrid type, but is unable to hydrolyse N-linked glycans of the complex-type and O-linked glycans. Previously, it has been shown that EndoE (EF0114) hydrolyzes all N-linked type glycans (Collin et al. 2004). More specifically, EF0114 can glycosylate IgG, whereas EF2863 cannot. It thus seems that the two enterococcal endo-*β*-*N*-acetylglucosaminidases not only have different locations but also slightly different substrate specificities.

The results indicated that *Ef*Endo18A is unable to hydrolyse the glycans of mucin. This was expected since the glycans in mucin mainly are O-linked (Bansil 2006), and it is well known that enterococci cannot degrade these (Hoskins et al. 1985; Corfield et al. 1992). However, our own unpublished experiments have shown that treatment with *Ef*Endo18A, but not treatment with BSA or a chitinase, leads to a marked change in the sedimentation behaviour of mucin from porcine stomach. This may indicate that *Ef*Endo18A has an effect on mucin, either by hydrolyzing the minor fraction of N-linked glycans or by an unknown glycan modifying activity. One possible explanation is that *Ef*Endo18A is capable of hydrolyzing

GlcNAc-GlcNAc bonds that are known to occur in the O-linked glycans of mucin (Wopereis *et al.*, 2006). Because the mucus layer covers the entire gastrointestinal tract and provides a first line of defence against pathogens and noxious substances, modification of mucin may affect the risk of invasion and subsequent infection by pathogenic bacteria. These unpublished observations clearly indicate that *Ef*Endo18A affects mucin properties, but further studies are needed to proof this. In particular, studies with a catalytically inactive mutant of *Ef*Endo18A are needed.

As many as two thirds of all eukaryotic proteins are predicted to be glycosylated, including many proteins of the immune system. It has been proposed that the glycan units of glycoproteins offer a significant nutritional advantage to bacteria that are able to hydrolyse them (Roberts et al. 2000). However, one may speculate about additional functions of the endo-*β*-*N*-acetylglucosaminidases. One hypothesis is that, since many proteins of the human immune system are glycosylated, endo-*β*-*N*-acetylglucosaminidase activity may lead to modulation of the immune system. Removal of the glycans will change protein properties and may result in dysfunction of the protein (Durand et al. 2000). Examples of eukaryotic glycoproteins related to the immune system include toll-like receptors that play a key role in sensing microbes and development of antigen-specific adaptive immunity. It has been proposed that after removal of N-linked glycans toll-like receptors may not be able to sense pathogenic bacteria and that the bacteria therefore may escape the immune system (Kataoka et al. 2006).

Further studies are needed to unravel the physiological importance of the two enterococcal endo-*β*-*N*-acetylglucosaminidases and of the difference between the two enzymes in terms of activity and localization. Interestingly, *Ef*Endo18A is upregulated during growth in both blood and urine, whereas EF0114 is not (Vebo et al. 2009; Vebo et al. 2010). As described in the introduction, section 1.3, enterococci are ranged among the top four causes of blood infections and second as a cause of urinary tract infections both in the United States and Europe (Sood et al. 2008). One may speculate that *Ef*Endo18A is a significant factor in these infections.

Analysis of genomic data have revealed that a number of LAB species contain simple enzymatic machineries for degradation of chitin, an insoluble and recalcitrant source of carbon and nitrogen. The genome sequence of *E. faecalis* V583 indicates the presence of a chitinolytic machinery that includes a chitinase (EF0361; *Ef*Chi18A) and a chitin binding

protein (EF0362; *Ef*CBM33A). Information regarding the functionality of these enzymes and their putative importance for the species as well as knowledge of the ability of LAB to degrade chitin is limited.Previously, only one study has shown convincing evidence of a functional chitinolytic system in a LAB, namely in *Lc. lactis* (Vaaje-Kolstad et al. 2009). Paper IV describes how *E. faecalis* V583 responds to the presence of chitin in the environment as well as a characterization of *Ef*Chi18A and *Ef*CBM33A.

The two proteins of the chitinolytic machinery of *E. faecalis* V583 described in Paper IV were identified during the studies of the surface proteome of Paper II. The chitinase, *Ef*Chi18A, was identified after the SDS-PAGE step (i.e. the step for identification of shedded, trypsin-resistant or protein fragments) in samples treated with trypsin and in samples not treated with trypsin. The chitin binding protein, *Ef*CBM33A, was identified by tryptic fragments obtained after the regular "shaving" treatment with trypsin agarose beads for one hour, or with free trypsin after both one and two hours. However, both proteins were only identified with one peptide each and, hence, they were not included as a significant hit in Paper II. *Ef*Chi18A and *Ef*CBM33A are predicted to be N-terminally anchored and secreted, respectively (Zhou et al. 2008). Interestingly, like *Ef*Endo18A (Paper III), both *Ef*Chi18A and *Ef*CBM33A were significantly up-regulated when *E. faecalis* V583 was grown in blood and urine (Vebo et al. 2009; Vebo et al. 2010). In addition, these latter two proteins were significantly up-regulated as a response to bile stress (Solheim et al. 2007). These results may indicate that *Ef*Chi18A and *Ef*CBM33A have an important unidentified function linked to proliferation and infection in the host.

As described in Paper IV, *Ef*Chi18A and *Ef*CBM33A were overexpressed and characterized. The results show that *Ef*Chi18A is an endochitinase with possibly up to eight subsites, whereas *Ef*CBM33A is an oxidative "helper enzyme" that acts synergistically with the chitinase. The data also show that, together, these two enzymes can degrade chitin, especially β-chitin, which they can convert completely to soluble sugars. The importance of CBM33 proteins for efficient degradation of chitin is known for other bacterial systems (Vaaje-Kolstad et al. 2005; Vaaje-Kolstad et al. 2009), but until 2010, it was not known that these proteins in fact are oxidative enzymes who use oxygen and whose activity is boosted by the addition of an external electron donor (Vaaje-Kolstad et al. 2010). The study of Paper IV shows that addition of an external electron donor (ascorbic acid or reduced glutathione) resulted in a boost in the activity of *Ef*CBM33A. Furthermore, the expected soluble oxidized products (chitooligosaccharide aldonic acids) were indeed observed with both MALDI-TOF

mass spectrometry and UHPLC analysis. The exact mechanism of the CBM33 enzymes is not yet known, and neither is the mechanism behind the contribution of the externally added reductant. One option is that auto-oxidation of the reductants results in reactive oxygen species which are then exploited by *Ef*CBM33A. Interestingly, *E. faecalis* V583 is known to produce substantial amounts of reactive oxygen species such as superoxide and hydrogen peroxide. This could imply that *Ef*CBM33A is automatically activiated during normal growth of *E. faecalis* V583 in the intestinal tract.

Because of the current major interest in CBM33 enzymes, the X-ray chrystallographic structure of *Ef*CBM33A was solved at atomic resolution (0.95 Å) by Bjørn Dalhus and co-workers. Compared to other CBM33s, *Ef*CBM33A lacks two disulphide bridges, which, however does not affect the stability of the protein. The structure also showed that the binding surface of *Ef*CBM33A contains all the conserved residues that previously have been found to be important for chitin binding by CBM33s (Vaaje-Kolstad et al. 2005). *Ef*CBM33A was shown to have a tryptophan in a position that is known to be important for substrate binding affinity. Interestingly this residue is variable for different CBMs and may thus play a role in substrate specificity. A tryptophan in this position indicates that the protein may bind to both α- and β-chitin (Kolbe et al. 1994), whereas a Tyrosine indicates that the protein has affinity to only β-chitin, as seen for CBP21 (Vaaje-Kolstad et al. 2005). Indeed, *Ef*CBM33A, having a tryptophan in this position, bound better to α-chitin, than CBP21, having a tyrosine in this position. It should be noted though that the activity of *Ef*CBM33A, on α-chitin was not very high (only 10 % conversion; Paper IV) and not significantly higher than the activity of CBP21 on α-chitin (unpublished observations).

As described in the introduction, section 1.5.3, CBM33s bind a divalent metal in a conserved metal-binding site on their surface. Studies on CBP21 have indicated that this metal binding site is promiscuous, since several divalent metals could re-activate the (inactive) EDTA-treated enzyme (Vaaje-Kolstad et al. 2010). Considering the oxidative nature of the reaction catalyzed by the CBM33s, it is remarkable that redox metals ($Cu^{2+}$, $Fe^{3+}$) do not seem to be essential. It is also remarkable that metal ions are not generally observed in the crystal structures of CBM33s, including *Ef*CBM33A. As described in Paper IV, similar metal-reconstitution experiments were done with *Ef*CBM33A. Interestingly, when both synergy (Paper IV) and oxidized products (unpublished data) were considered, cobalt increased the activity the most, while the redox metals had smaller or even inhibitory (in the case of copper) effects. These results seem to confirm the promiscuity of the metal-binding

site and the absence of a need for a redox metal. Interestingly, for the structurally and functionally analogous fungal GH61 proteins, it has very recently been shown that these are copper oxidases (Westereng et al. 2011; Quinlan et al. 2011). Although the same does not seem to apply for CBM33s, this possibility cannot yet be fully excluded. For example, it is possible that the experiments with metals described in Paper IV and Vaaje-Kolstad et al., 2010, lead to partly artificial results due to the presence of divalent metals in the substrate.

The most obvious function of bacterial chitinases is to meet nutritional needs and the results described in Paper IV indeed show that *E. faecalis* V583 is able to utilize GlcNAc, $GlcNAc_2$, and β-chitin as the sole carbon source. This corroborates the hypothesis that *E. faecalis* V583 is fully able to utilize both soluble and insoluble chitinous substrates as a carbon source. Previous studies on the growth abilities of intestinal bacteria have shown that they generally are able to utilize different oligomers of GlcNAc as their main carbon source (Chen et al. 2002). Interestingly, the results of this 2002 study showed that *Bacteroides fragilis* and *Clostridium perfringens* was able to use GlcNAc and $GlcNAc_2$, respectively, more efficiently than glucose. Remarkably, the same was found for *E. faecalis* V583, as shown in Fig. 1 of Paper IV. One may wonder why enterococci have evolved the ability to efficiently grow on GlcNAc, since there is no chitin present in the natural environment (intestinal tract) of enterococci. Interestingly, different sources of chitin may occasionally be ingested. It has been shown that if shrimp shells are ingested, the presence of gastric juice in the stomach will release chitin (Chen 1997), which then becomes available to the complex microflora in the intestinal tract.

It should be noted that enterococci are versatile organisms occurring in a variety of ecological niches and that it thus may be wrong to interprete our observations in the light of the (human) intestinal niche only. Enterococci are frequently used as an indicator of fecal contamination in soil and water, which are both chitin-rich environments where the presence of *Ef*Chi18A and *Ef*CBM33A certainly could result in a competitive advantage. *E. faecalis* V583 has also been identified as part of the microflora of the intestinal tract of insects where it may cause infections (Cox et al. 2007). It is quite conceivable that production of chitin-degrading enzymes will contribute to the infection process in these chitin-rich organisms. A completely different explanation for the presence of a chitinolytic system could relate to the bacterium's defence against (chitin-containing) fungi; it is well known that chitinases have anti-fungal activity (Bhattacharya et al. 2007).

It has been shown that chitinases and chitin-binding proteins may modify host immune responses by binding to glycoproteins carrying GlcNAc-containing carbohydrates. Recently, a chitinase has been linked to bacterial virulence of *Lg. pneumophila* (DebRoy et al. 2006), a CBM33 has been linked to bacterial virulence of *V. cholerae* (Kirn et al. 2005; Bhowmick et al. 2008; Jude et al. 2009) and a chitinase and a CBM33 have been linked to virulence of *L. monocytogenes* (Chaudhuri et al. 2010). In addition it has been shown that a chitin binding protein from the probiotic bacterium *Lb. plantarum* has been shown to interact with GlcNAc-containing polymers such as mucin and with epithelial cells (Sanchez et al. 2011). Clearly, these observations suggest that *Ef*Chi18A and *Ef*CBM33A may have other roles than chitin degradation and further work is needed to unravel these roles. For example, the combination of the relative lack of both oxygen and chitin in the intestinal tract, with the fact that *Ef*CBM33A acts on chitin and needs oxygen, indicates that *Ef*CBM33A has an additional function. It might act on other substrates, or its main purpose may not necessarily be cleavage of GlcNAc-GlcNAc bonds, but rather to interact with GlcNAc-containing carbohydrates.

## 4. CONCLUDING REMARKS

The work presented in this thesis provides information regarding the mechanisms that *E. faecalis* V583 may exploit to respond optimally to its environment. The tolerance to bile is a crucial trait for survival and colonization in the gastrointestinal tract, whereas surface proteins and the ability to harvest various types of food are other factors that determine bacterial success and harshness.

Paper I revealed that *E. faecalis* V583 mainly regulates proteins involved in fatty acid and phospholipid metabolism as a response to bile. Furthermore, the results also indicated that under these conditions oxidative phosphorylation is the pathway of choice in order to obtain energy and perhaps to combat competing species. These physiological responses and the underlying pathways are interesting targets in further research on enterococcal harshness. Paper II describes successful use of a relatively novel method for identification of surface-located proteins. Among the 69 identified proteins there are as much as 23 proteins with unknown functions. Further work should focus on identifying a large fraction of the surface proteome, studying the effects of environmental factors on the surface proteome, and functional studies of the identified proteins, in particular those with unknown functions. These latter proteins may be interesting targets for drug design, and knockout-studies will be conducted in our laboratory in order to identify their function.

*Ef*Endo18A characterized in Paper III hydrolyses a glycosidic linkage in the high-mannose and hybrid-type N-linked glycans of glycoproteins, thus releasing the glycans, except one remaining GlcNAc, to the medium. Glycoproteins are abundant in the environment of the bacterium, and it is therefore tempting to assign a function in nutrient acquisition to this enzyme. However, a secondary function may be modulation of the immune system in a way that perhaps may help the bacteria in escaping the immune system. It would be of major interest to knock-out this gene, perhaps in combination with knocking out the previously described EF0114 gene, and look for phenotypes related to survival (nutrient exploitation) and immunological effects. The latter could be done by stuyding the effects of wild-type and knock-out bacteria on cytokine production by human cells.

Paper IV describes the chitinolytic machinery of *E. faecalis* V583 comprising of *Ef*Chi18A and *Ef*CBM33A, which is able to hydrolyse both α- and β-chitin, especially in the presence of external electron donors that boost the action of *Ef*CBM33A. A particularly important element of Paper IV is the in-depth study of the CBM33, since this protein belongs to a recently discovered class of oxidative enzymes that are of major importance for the

enzymatic conversion of biomass, a “hot topic” in current biotechnology. From the point of view of enterococcal physiology, it is interesting to wonder about the role of this chitinolytic system, since chitin is not normally present in the intestinal tract of humans. There are vague indications for other possible roles of his enzyme system. Clearly, also in this case, comparative phenotypical studies of the wild-type strain and knockout mutants would be interesting to do. Such studies should include studies of bacterial virulence and it may also be wise to see if there are any functional links between the family GH18 protein described in paper III and the enzymes described in Paper IV.

## 5. REFERENCES


Agerholm-Larsen, L., A. Raben, N. Haulrik, A. S. Hansen, M. Manders and A. Astrup (2000). "Effect of 8 week intake of probiotic milk products on risk factors for cardiovascular diseases." Eur J Clin Nutr **54**(4): 288-297.

Amyes, S. G. (2007). "Enterococci and streptococci." Int J Antimicrob Ag **29 Suppl 3**: S43-52.

Apweiler, R., H. Hermjakob and N. Sharon (1999). "On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database." Biochim Biophys Acta **1473**(1): 4-8.

Aristoteli, L. P. and M. D. Willcox (2003). "Mucin degradation mechanisms by distinct Pseudomonas aeruginosa isolates in vitro." Infect Immun **71**(10): 5565-5575.

Ayalew, S., D. L. Step, M. Montelongo and A. W. Confer (2009). "Intranasal vaccination of calves with Mannheimia haemolytica chimeric protein containing the major surface epitope of outer membrane lipoprotein PlpE, the neutralizing epitope of leukotoxin, and cholera toxin subunit B." Vet Immuno Immunop **132**(2-4): 295-302.

Bansil, R., Turner, B.S. (2006). "Mucin structure, aggregation, physiological functions and biomedical applications." Curr opin colloid in **11**: 164-170.

Beck, H. C., S. M. Madsen, J. Glenting, J. Petersen, H. Israelsen, M. R. Norrelykke, M. Antonsson and A. M. Hansen (2009). "Proteomic analysis of cell surface-associated proteins from probiotic *Lactobacillus plantarum*." FEMS Microbiol Lett **297**(1): 61-66.

Begley, M., C. G. Gahan and C. Hill (2005). "The interaction between bacteria and bile." FEMS Microbiol Rev **29**(4): 625-651.

Benachour, A., T. Morin, L. Hebert, A. Budin-Verneuil, A. Le Jeune, Y. Auffray and V. Pichereau (2009). "Identification of secreted and surface proteins from *Enterococcus faecalis*." Can J Microbiol **55**(8): 967-974.

Benz, I. and M. A. Schmidt (2002). "Never say never again: protein glycosylation in pathogenic bacteria." Mol Microbiol **45**(2): 267-276.

Bertolami, M. C., A. A. Faludi and M. Batlouni (1999). "Evaluation of the effects of a new fermented milk product (Gaio) on primary hypercholesterolemia." Eur J Clin Nutr **53**(2): 97-101.

Bhattacharya, D., A. Nagpure and R. K. Gupta (2007). "Bacterial chitinases: properties and potential." Crit Rev Biotechnol **27**(1): 21-28.

Bhowmick, R., A. Ghosal, B. Das, H. Koley, D. R. Saha, S. Ganguly, R. K. Nandy, R. K. Bhadra and N. S. Chatterjee (2008). "Intestinal adherence of *Vibrio cholerae* involves a coordinated interaction between colonization factor GbpA and mucin." Infect Immun **76**(11): 4968-4977.

Bourgogne, A., D. A. Garsin, X. Qin, K. V. Singh, J. Sillanpaa, S. Yerrapragada, Y. Ding, S. Dugan-Rocha, C. Buhay, H. Shen, G. Chen, G. Williams, D. Muzny, A. Maadani, K. A. Fox, J. Gioia, L. Chen, Y. Shang, C. A. Arias, S. R. Nallapareddy, M. Zhao, V. P. Prakash, S. Chowdhury, H. Jiang, R. A. Gibbs, B. E. Murray, S. K. Highlander and G. M. Weinstock (2008). "Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF." Genome Biol **9**(7): R110.

Brede, D. A., L. G. Snipen, D. W. Ussery, A. J. Nederbragt and I. F. Nes (2011). "Complete Genome Sequence of the Commensal *Enterococcus faecalis* 62, Isolated from a Healthy Norwegian Infant." J Bacteriol **193**(9): 2377-2378.

Bron, P. A., M. Marco, S. M. Hoffer, E. Van Mullekom, W. M. de Vos and M. Kleerebezem (2004). "Genetic characterization of the bile salt response in *Lactobacillus plantarum*

and analysis of responsive promoters in vitro and in situ in the gastrointestinal tract." J Bacteriol **186**(23): 7829-7835.

Bron, P. A., D. Molenaar, W. M. de Vos and M. Kleerebezem (2006). "DNA micro-array-based identification of bile-responsive genes in *Lactobacillus plantarum*." J Appl Microbiol **100**(4): 728-738.

Buist, G., A. Steen, J. Kok and O. P. Kuipers (2008). "LysM, a widely distributed protein motif for binding to (peptido)glycans." Mol Microbiol **68**(4): 838-847.

Burns, P., B. Sanchez, G. Vinderola, P. Ruas-Madiedo, L. Ruiz, A. Margolles, J. Reinheimer and C. G. de los Reyes-Gavilan (2010). "Inside the adaptation process of *Lactobacillus delbrueckii* subsp. *lactis* to bile." Int J Food Microbiol **142**(1-2): 132-141.

Bussink, A. P., D. Speijer, J. M. Aerts and R. G. Boot (2007). "Evolution of mammalian chitinase(-like) members of family 18 glycosyl hydrolases." Genetics **177**(2): 959-970.

Byers, H. L., E. Tarelli, K. A. Homer and D. Beighton (1999). "Sequential deglycosylation and utilization of the N-linked, complex-type glycans of human alpha1-acid glycoprotein mediates growth of *Streptococcus oralis*." Glycobiology **9**(5): 469-479.

Cantarel, B. L., P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard and B. Henrissat (2009). "The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics." Nucleic Acids Res **37**(Database issue): D233-238.

Chaudhuri, S., J. C. Bruno, F. Alonzo, 3rd, B. Xayarath, N. P. Cianciotto and N. E. Freitag (2010). "Contribution of chitinases to *Listeria monocytogenes* pathogenesis." Appl Environ Microbiol **76**(21): 7302-7305.

Chen, H. C., C. C. Chang, W. J. Mau and L. S. Yen (2002). "Evaluation of N-acetylchitooligosaccharides as the main carbon sources for the growth of intestinal bacteria." FEMS Microbiol Lett **209**(1): 53-56.

Chen, H. C. a. H., C.C. (1997). "Production of chitin and its partial deacetylation from shrimp shells by microbial degradation." J Chin Agric Chem Soc **35**: 342-353.

Chhatwal, G. S. (2002). "Anchorless adhesins and invasins of Gram-positive bacteria: a new class of virulence factors." Trends Microbiol **10**(5): 205-208.

Clewell, D. B. (2007). "Properties of *Enterococcus faecalis* plasmid pAD1, a member of a widely disseminated family of pheromone-responding, conjugative, virulence elements encoding cytolysin." Plasmid **58**(3): 205-227.

Coburn, P. S. and M. S. Gilmore (2003). "The *Enterococcus faecalis* cytolysin: a novel toxin active against eukaryotic and prokaryotic cells." Cell Microbiol **5**(10): 661-669.

Coburn, P. S., C. M. Pillar, B. D. Jett, W. Haas and M. S. Gilmore (2004). "*Enterococcus faecalis* Senses Target Cells and in Response Expresses Cytolysin." Science **306**(5705): 2270-2272.

Cocolin, L., R. Foschino, G. Comi and M. Grazia Fortina (2007). "Description of the bacteriocins produced by two strains of *Enterococcus faecium* isolated from Italian goat milk." Food Microbiol **24**(7-8): 752-758.

Collin, M. and V. A. Fischetti (2004). "A novel secreted endoglycosidase from *Enterococcus faecalis* with activity on human immunoglobulin G and ribonuclease B." J Biol Chem **279**(21): 22558-22570.

Collin, M., O. Shannon and L. Bjorck (2008). "IgG glycan hydrolysis by a bacterial enzyme as a therapy against autoimmune conditions." Proc Natl Acad Sci U S A **105**(11): 4265-4270.

Corcoran, B. M., C. Stanton, G. Fitzgerald and R. P. Ross (2008). "Life under stress: the probiotic stress response and how it may be manipulated." Curr Pharm Design **14**(14): 1382-1399.

Corfield, A. P., S. A. Wagner, J. R. Clamp, M. S. Kriaris and L. C. Hoskins (1992). "Mucin degradation in the human colon: production of sialidase, sialate O-acetylesterase, N-acetylneuraminate lyase, arylesterase, and glycosulfatase activities by strains of fecal bacteria." Infect Immun **60**(10): 3971-3978.

Cotter, P. D. and C. Hill (2003). "Surviving the acid test: responses of gram-positive bacteria to low pH." Microbiol Mol Biol Rev **67**(3): 429-453, table of contents.

Cox, C. R. and M. S. Gilmore (2007). "Native microbial colonization of *Drosophila melanogaster* and its use as a model of *Enterococcus faecalis* pathogenesis." Infect Immun **75**(4): 1565-1576.

Cremonini, F., S. Di Caro, L. Santarelli, M. Gabrielli, M. Candelli, E. C. Nista, A. Lupascu, G. Gasbarrini and A. Gasbarrini (2002). "Probiotics in antibiotic-associated diarrhoea." Dig Liver Dis **34 Suppl 2**: S78-80.

Csonka, L. N. and A. D. Hanson (1991). "Prokaryotic osmoregulation: genetics and physiology." Annu Rev Microbiol **45**: 569-606.

Davies, G. and B. Henrissat (1995). "Structures and mechanisms of glycosyl hydrolases." Structure **3**(9): 853-859.

de Fatima Silva Lopes, M., T. Ribeiro, M. Abrantes, J. J. Figueiredo Marques, R. Tenreiro and M. T. Crespo (2005). "Antimicrobial resistance profiles of dairy and clinical isolates and type strains of enterococci." Int J Food Microbiol **103**(2): 191-198.

DebRoy, S., J. Dao, M. Soderberg, O. Rossier and N. P. Cianciotto (2006). "*Legionella pneumophila* type II secretome reveals unique exoproteins and a chitinase that promotes bacterial persistence in the lung." Proc Natl Acad Sci U S A **103**(50): 19146-19151.

Derrien, M., M. W. van Passel, J. H. van de Bovenkamp, R. G. Schipper, W. M. de Vos and J. Dekker (2010). "Mucin-bacterial interactions in the human oral cavity and digestive tract." Gut Microbes **1**(4): 254-268.

Derrien, M., E. E. Vaughan, C. M. Plugge and W. M. de Vos (2004). "*Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium." Int J Syst Evol Microbiol **54**(Pt 5): 1469-1476.

Domann, E., T. Hain, R. Ghai, A. Billion, C. Kuenne, K. Zimmermann and T. Chakraborty (2007). "Comparative genomic analysis for the presence of potential enterococcal virulence factors in the probiotic *Enterococcus faecalis* strain Symbioflor 1." Int J Med Microbiol **297**(7-8): 533-539.

Dreisbach, A., J. M. van Dijl and G. Buist (2011). "The cell surface proteome of *Staphylococcus aureus*." Proteomics **11** (15): 3154-3168.

Duerkop, B. A., S. Vaishnava and L. V. Hooper (2009). "Immune responses to the microbiota at the intestinal mucosal surface." Immunity **31**(3): 368-376.

Dumas, E., M. Desvaux, C. Chambon and M. Hebraud (2009). "Insight into the core and variant exoproteomes of *Listeria monocytogenes* species by comparative subproteomic analysis." Proteomics **9**(11): 3136-3155.

Durand, G. and N. Seta (2000). "Protein glycosylation and diseases: blood and urinary oligosaccharides as markers for diagnosis and therapeutic monitoring." Clin Chem **46**(6 Pt 1): 795-805.

Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson and D. A. Relman (2005). "Diversity of the human intestinal microbial flora." Science **308**(5728): 1635-1638.

Eijsink, V., I. Hoell and G. Vaaje-Kolstada (2010). "Structure and function of enzymes acting on chitin and chitosan." Biotechnol Genet Eng **27**: 331-366.

Elias, J. A., R. J. Homer, Q. Hamid and C. G. Lee (2005). "Chitinases and chitinase-like proteins in T(H)2 inflammation and asthma." J Allergy Clin Immunol **116**(3): 497-500.

Fisher, K. and C. Phillips (2009). "The ecology, epidemiology and virulence of *Enterococcus*." Microbiology **155**(Pt 6): 1749-1757.

Flahaut, S., J. Frere, P. Boutibonnes and Y. Auffray (1996). "Comparison of the bile salts and sodium dodecyl sulfate stress responses in *Enterococcus faecalis*." Appl Environ Microb **62**(7): 2416-2420.

Flahaut, S., A. Hartke, J. C. Giard, A. Benachour, P. Boutibonnes and Y. Auffray (1996). "Relationship between stress response toward bile salts, acid and heat treatment in *Enterococcus faecalis*." FEMS Microbiol Lett **138**(1): 49-54.

Folders, J., J. Algra, M. S. Roelofs, L. C. van Loon, J. Tommassen and W. Bitter (2001). "Characterization of *Pseudomonas aeruginosa* chitinase, a gradually secreted protein." J Bacteriol **183**(24): 7044-7052.

Forsberg, Z., G. Vaaje-Kolstad, B. Westereng, A. C. Bunaes, Y. Stenstrom, A. Mackenzie, M. Sorlie, S. J. Horn and V. G. Eijsink (2011). "Cleavage of cellulose by a CBM33 protein." Protein sci **20**(9): 1479-1483.

Foster, T. J. and M. Hook (1998). "Surface protein adhesins of *Staphylococcus aureus*." Trends Microbiol **6**(12): 484-488.

Foulquie Moreno, M. R., P. Sarantinopoulos, E. Tsakalidou and L. De Vuyst (2006). "The role and application of enterococci in food and health." Int J Food Microbiol **106**(1): 1-24.

Fozo, E. M. and R. G. Quivey, Jr. (2004). "Shifts in the membrane fatty acid profile of *Streptococcus mutans* enhance survival in acidic environments." Appl Environ Microbiol **70**(2): 929-936.

Frankenberg, L., M. Brugna and L. Hederstedt (2002). "*Enterococcus faecalis* heme-dependent catalase." J Bacteriol **184**(22): 6351-6356.

Franz, C. M., M. E. Stiles, K. H. Schleifer and W. H. Holzapfel (2003). "Enterococci in foods--a conundrum for food safety." Int J Food Microbiol **88**(2-3): 105-122.

Frees, D., F. K. Vogensen and H. Ingmer (2003). "Identification of proteins induced at low pH in *Lactococcus lactis*." Int J Food Microbiol **87**(3): 293-300.

Fuchs, R. L., S. A. McPherson and D. J. Drahos (1986). "Cloning of a *Serratia marcescens* Gene Encoding Chitinase." Appl Environ Microbiol **51**(3): 504-509.

Funkhouser, J. D. and N. N. Aronson, Jr. (2007). "Chitinase family GH18: evolutionary insights from the genomic history of a diverse protein family." BMC Evol Biol **7**: 96.

Gardner, K. H. and J. Blackwell (1975). "Refinement of the structure of beta-chitin." Biopolymers **14**(8): 1581-1595.

Garsin, D. A., C. D. Sifri, E. Mylonakis, X. Qin, K. V. Singh, B. E. Murray, S. B. Calderwood and F. M. Ausubel (2001). "A simple model host for identifying Gram-positive virulence factors." Proc Natl Acad Sci U S A **98**(19): 10892-10897.

Gianfaldoni, C., S. Censini, M. Hilleringmann, M. Moschioni, C. Facciotti, W. Pansegrau, V. Masignani, A. Covacci, R. Rappuoli, M. A. Barocchi and P. Ruggiero (2007). "*Streptococcus pneumoniae* pilus subunits protect mice against lethal challenge." Infect Immun **75**(2): 1059-62.

Giard, J. C., J. M. Laplace, A. Rince, V. Pichereau, A. Benachour, C. Leboeuf, S. Flahaut, Y. Auffray and A. Hartke (2001). "The stress proteome of *Enterococcus faecalis*." Electrophoresis **22**(14): 2947-2954.

Gooday, G. W. (1990). "Physiology of microbial degradation of chitin and chitosan." Biodegradation **1**: 177-190.

Gunn, J. S. (2000). "Mechanisms of bacterial resistance and response to bile." Microbes Infect **2**(8): 907-913.

Hallgren, A., C. Claesson, B. Saeedi, H. J. Monstein, H. Hanberger and L. E. Nilsson (2009). "Molecular detection of aggregation substance, enterococcal surface protein, and cytolysin genes and in vitro adhesion to urinary catheters of *Enterococcus faecalis* and *E. faecium* of clinical origin." Int J Med Microbiol **299**(5): 323-332.

Hamon, E., P. Horvatovich, E. Izquierdo, F. Bringel, E. Marchioni, D. Aoude-Werner and S. Ennahar (2011). "Comparative proteomic analysis of *Lactobacillus plantarum* for the identification of key proteins in bile tolerance." BMC Microbiol **11**: 63.

Hancock, L. E. and M. S. Gilmore (2002). "The capsular polysaccharide of *Enterococcus faecalis* and its relationship to other polysaccharides in the cell wall." Proc Natl Acad Sci U S A **99**(3): 1574-1579.

Hancock, L. E. and M. Perego (2004). "The *Enterococcus faecalis fsr* two-component system controls biofilm development through production of gelatinase." J Bacteriol **186**(17): 5629-5639.

Hashimoto, M., T. Ikegami, S. Seino, N. Ohuchi, H. Fukada, J. Sugiyama, M. Shirakawa and T. Watanabe (2000). "Expression and characterization of the chitin-binding domain of chitinase A1 from Bacillus circulans WL-12." J Bacteriol **182**(11): 3045-3054.

Hassett, D. J. and M. S. Cohen (1989). "Bacterial adaptation to oxidative stress: implications for pathogenesis and interaction with phagocytic cells." The FASEB journal : official publication of the Federation of American Societies for Experimental Biology **3**(14): 2574-2582.

Hawtin, R. E., K. Arnold, M. D. Ayres, P. M. Zanotto, S. C. Howard, G. W. Gooday, L. H. Chappell, P. A. Kitts, L. A. King and R. D. Possee (1995). "Identification and preliminary characterization of a chitinase gene in the *Autographa californica* nuclear polyhedrosis virus genome." Virology **212**(2): 673-685.

Henrissat, B. (1991). "A classification of glycosyl hydrolases based on amino acid sequence similarities." Biochem J **280 ( Pt 2)**: 309-316.

Henrissat, B. and G. Davies (1997). "Structural and sequence-based classification of glycoside hydrolases." Curr Opin Struct Biol **7**(5): 637-44.

Hew, C. M., M. Korakli and R. F. Vogel (2006). "Expression of virulence-related genes by *Enterococcus faecalis* in response to different environments." Syst Appl Microbiol.

Hirt, H., S. L. Erlandsen and G. M. Dunny (2000). "Heterologous inducible expression of Enterococcus faecalis pCF10 aggregation substance asc10 in *Lactococcus lactis* and *Streptococcus gordonii* contributes to cell hydrophobicity and adhesion to fibrin." J Bacteriol **182**(8): 2299-2306.

Horn S. J., S. M., Vaaje-Kolstad G., ( 2006). "Comparative studies of chitinases A, B and C from Serratia marcescens " Biocatal Biotransfor **24**(1-2): 39-53.

Horn, S. J., A. Sorbotten, B. Synstad, P. Sikorski, M. Sorlie, K. M. Varum and V. G. Eijsink (2006). "Endo/exo mechanism and processivity of family 18 chitinases produced by *Serratia marcescens*." FEBS J **273**(3): 491-503.

Hoskins, L. C., M. Agustines, W. B. McKee, E. T. Boulding, M. Kriaris and G. Niedermeyer (1985). "Mucin degradation in human colon ecosystems. Isolation and properties of fecal strains that degrade ABH blood group antigens and oligosaccharides from mucin glycoproteins." J Clin Invest **75**(3): 944-953.

Huycke, M. M. (2002). The Enterococci: Pathogenesis, molecular biology and antibiotic resistance, Am Soc Microbiol. Ch 4; 133-140

Huycke, M. M., V. Abrams and D. R. Moore (2002). "*Enterococcus faecalis* produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA." Carcinogenesis **23**(3): 529-536.

Huycke, M. M., W. Joyce and M. F. Wack (1996). "Augmented production of extracellular superoxide by blood isolates of *Enterococcus faecalis*." J Infect Dis **173**(3): 743-746.

Huycke, M. M., D. Moore, W. Joyce, P. Wise, L. Shepard, Y. Kotake and M. S. Gilmore (2001). "Extracellular superoxide production by *Enterococcus faecalis* requires demethylmenaquinone and is attenuated by functional terminal quinol oxidases." Mol Microbiol **42**(3): 729-740.

Huycke, M. M., D. F. Sahm and M. S. Gilmore (1998). "Multiple-drug resistant enterococci: the nature of the problem and an agenda for the future." Emerg Infect Dis **4**(2): 239-249.

Ibarguren, C., R. R. Raya, M. C. Apella and M. C. Audisio (2010). "*Enterococcus faecium* isolated from honey synthesized bacteriocin-like substances active against different *Listeria monocytogenes* strains." J Microbiol **48**(1): 44-52.

Jeffery, C. J. (1999). "Moonlighting proteins." Trends Biochem Sci **24**(1): 8-11.

Jenkinson, H. F. and R. J. Lamont (1997). "Streptococcal adhesion and colonization." Crit Rev Oral Biol Med **8**(2): 175-200.

Jude, B. A., R. M. Martinez, K. Skorupski and R. K. Taylor (2009). "Levels of the secreted *Vibrio cholerae* attachment factor GbpA are modulated by quorum-sensing-induced proteolysis." J Bacteriol **191**(22): 6911-6917.

Kataoka, H., M. Yasuda, M. Iyori, K. Kiura, M. Narita, T. Nakata and K. Shibata (2006). "Roles of N-linked glycans in the recognition of microbial lipopeptides and lipoproteins by TLR2." Cell Microbiol **8**(7): 1199-1209.

Khoushab, F. and M. Yamabhai (2010). "Chitin research revisited." Mar Drugs **8**(7): 1988-2012.

Kirn, T. J., B. A. Jude and R. K. Taylor (2005). "A colonization factor links *Vibrio cholerae* environmental survival and human infection." Nature **438**(7069): 863-866.

Klebanoff, S. J. (1980). "Oxygen metabolism and the toxic properties of phagocytes." Ann Intern Med **93**(3): 480-489.

Klein, G. (2003). "Taxonomy, ecology and antibiotic resistance of enterococci from food and the gastro-intestinal tract." Int J Food Microbiol **88**(2-3): 123-131.

Koshland, D. E. (1953). "Stereochemistry and the mechanism of enzymatic reactions." Biol Rev **28**(4): 416-436.

Kovacs-Simon, A., R. W. Titball and S. L. Michell (2011). "Lipoproteins of bacterial pathogens." Infect Immun **79**(2): 548-561.

Kunin, C. M. and J. Rudy (1991). "Effect of NaCl-induced osmotic stress on intracellular concentrations of glycine betaine and potassium in *Escherichia coli*, *Enterococcus faecalis*, and staphylococci." J Lab Clin Med **118**(3): 217-224.

Kuo, C. Y., Y. J. Wu, Y. C. Hsieh, H. H. Guan, H. J. Tsai, Y. H. Lin, Y. C. Huang, M. Y. Liu, Y. K. Li and C. J. Chen (2006). "Purification, crystallization and preliminary X-ray crystallographic analysis of chitinase from *Bacillus cereus* NCTU2." Acta Crystallogr Sect F Struct Biol Cryst Commun **62**(Pt 9): 916-919.

Kuranda, M. J. and P. W. Robbins (1991). "Chitinase is required for cell separation during growth of *Saccharomyces cerevisiae*." J Biol Chem **266**(29): 19758-19767.

Kzhyshkowska, J., A. Gratchev and S. Goerdt (2007). "Human chitinases and chitinase-like proteins as indicators for inflammation and cancer." Biomark Insights **2**: 128-146.

La Carbona, S., N. Sauvageot, J. C. Giard, A. Benachour, B. Posteraro, Y. Auffray, M. Sanguinetti and A. Hartke (2007). "Comparative study of the physiological roles of three peroxidases (NADH peroxidase, Alkyl hydroperoxide reductase and Thiol peroxidase) in oxidative stress response, survival inside macrophages and virulence of *Enterococcus faecalis*." Mol Microbiol **66**(5): 1148-1163.

Lan, X., X. Zhang, J. Hu and M. Shimosaka (2006). "Cloning, expression, and characterization of a chitinase from the chitinolytic bacterium *Aeromonas hydrophila* strain SUWA-9." Biosci Biotechnol Biochem **70**(10): 2437-2442.

Laukova, A. and S. Czikkova (1999). "The use of enterocin CCM 4231 in soy milk to control the growth of *Listeria monocytogenes* and *Staphylococcus aureus*." J Appl Microbiol **87**(1): 182-186.

Leblanck, D. J. (2006). Enterococcus. The Prokaryotes, Springer Science. **4:** 175-204.

Lee, K., H. G. Lee and Y. J. Choi (2008). "Proteomic analysis of the effect of bile salts on the intestinal and probiotic bacterium *Lactobacillus reuteri*." J Biotech **137**(1-4): 14-19.

Leisner, J. J., M. H. Larsen, H. Ingmer, B. O. Petersen, J. O. Duus and M. M. Palcic (2009). "Cloning and comparison of phylogenetically related chitinases from *Listeria monocytogenes* EGD and *Enterococcus faecalis* V583." J Appl Microbiol. 107 (6): 2080-2087.

Leisner, J. J., M. H. Larsen, R. L. Jorgensen, L. Brondsted, L. E. Thomsen and H. Ingmer (2008). "Chitin hydrolysis by *Listeria* spp., including *L. monocytogenes*." Appl Environ Microbiol **74**(12): 3823-3830.

Leisner, J. J., F. K. Vogensen, J. Kollmann, B. Aideh, P. Vandamme, M. Vancanneyt and H. Ingmer (2008). "alpha-Chitinase activity among lactic acid bacteria." Syst Appl Microbiol **31**(2): 151-156.

Leroy, F., M. R. Foulquie Moreno and L. De Vuyst (2003). "*Enterococcus faecium* RZS C5, an interesting bacteriocin producer to be used as a co-culture in food fermentation." Int J Food Microbiol **88**(2-3): 235-240.

Leverrier, P., D. Dimova, V. Pichereau, Y. Auffray, P. Boyaval and G. Jan (2003). "Susceptibility and adaptive response to bile salts in *Propionibacterium freudenreichii*: physiological and proteomic analysis." Appl Environ Microbiol **69**(7): 3809-3818.

Ley, R. E., D. A. Peterson and J. I. Gordon (2006). "Ecological and evolutionary forces shaping microbial diversity in the human intestine." Cell **124**(4): 837-848.

Li, Z., C. Li, K. Yang, L. Wang, C. Yin, Y. Gong and Y. Pang (2003). "Characterization of a chitin-binding protein GP37 of *Spodoptera litura* multicapsid nucleopolyhedrovirus." Virus Research **96**(1-2): 113-122.

Maddalo, G., P. Chovanec, F. Stenberg-Bruzell, H. V. Nielsen, M. I. Jensen-Seaman, L. L. Ilag, K. A. Kline and D. O. Daley (2011). "A reference map of the membrane proteome of *Enterococcus faecalis*." Proteomics.

Maione, D., I. Margarit, C. D. Rinaudo, V. Masignani, M. Mora, M. Scarselli, H. Tettelin, C. Brettoni, E. T. Iacobini, R. Rosini, N. D'Agostino, L. Miorin, S. Buccato, M. Mariani, G. Galli, R. Nogarotto, V. Nardi Dei, F. Vegni, C. Fraser, G. Mancuso, G. Teti, L. C. Madoff, L. C. Paoletti, R. Rappuoli, D. L. Kasper, J. L. Telford and G. Grandi (2005). "Identification of a universal Group B streptococcus vaccine by multiple genome screen." Science **309**(5731): 148-150.

Maisnier-Patin, S., E. Forni and J. Richard (1996). "Purification, partial characterisation and mode of action of enterococcin EFS2, an antilisterial bacteriocin produced by a strain of *Enterococcus faecalis* isolated from a cheese." Int J Food Microbiol **30**(3): 255-270.

Marion, C., D. H. Limoli, G. S. Bobulsky, J. L. Abraham, A. M. Burnaugh and S. J. King (2009). "Identification of a pneumococcal glycosidase that modifies O-linked glycans." Infect Immun **77**(4): 1389-1396.

Merzendorfer, H. and L. Zimoch (2003). "Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases." J Exp Biol **206**(Pt 24): 4393-4412.

Mescher, M. F., J. L. Strominger and S. W. Watson (1974). "Protein and carbohydrate composition of the cell envelope of *Halobacterium salinarium*." J Bacteriol **120**(2): 945-954.

Michalski, W. M. a. J.-C. (2005). "The mass spectrometric analysis of glycoproteins and their glycan structures." Curr Anal Chem **1**(1): 29-57.

Minke, R. and J. Blackwell (1978). "The structure of alpha-chitin." J Mol Biol **120**(2): 167-181.

Mora, M., G. Bensi, S. Capo, F. Falugi, C. Zingaretti, A. G. Manetti, T. Maggi, A. R. Taddei, G. Grandi and J. L. Telford (2005). "Group A Streptococcus produce pilus-like structures containing protective antigens and Lancefield T antigens." Proc Natl Acad Sci U S A **102**(43): 15641-15646.

Moran, A. P., A. Gupta and L. Joshi (2011). "Sweet-talk: role of host glycosylation in bacterial pathogenesis of the gastrointestinal tract." Gut. **60**(10): 1412-1425.

Morelle, W., Michalski, J.-C. (2005). "The mass spectrometric analysis of glycoproteins and their glycan structures." Curr Anal Chem **1**: 29-57.

Mundy, L. M., D. F. Sahm and M. Gilmore (2000). "Relationships between enterococcal virulence and antimicrobial resistance." Clin Microbiol Rev **13**(4): 513-522.

Murray, B. E. (1990). "The life and times of the *Enterococcus*." Clin Microbiol Rev **3**(1): 46-65.

Nallapareddy, S. R., J. Sillanpaa, J. Mitchell, K. V. Singh, S. A. Chowdhury, G. M. Weinstock, P. M. Sullam and B. E. Murray (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**(7): 2911-2920.

Nallapareddy, S. R., K. V. Singh, J. Sillanpaa, D. A. Garsin, M. Hook, S. L. Erlandsen and B. E. Murray (2006). "Endocarditis and biofilm-associated pili of *Enterococcus faecalis*." J Clin Invest **116**(10): 2799-2807.

Nothaft, H. and C. M. Szymanski (2010). "Protein glycosylation in bacteria: sweeter than ever." Nat Rev Microbiol **8**(11): 765-778.

Ogier, J. C. and P. Serror (2008). "Safety assessment of dairy microorganisms: the *Enterococcus* genus." Int J Food Microbiol **126**(3): 291-301.

Ortiz-Rodriguez, T., N. de la Fuente-Salcido, D. K. Bideshi, R. Salcedo-Hernandez and J. E. Barboza-Corona (2010). "Generation of chitin-derived oligosaccharides toxic to pathogenic bacteria using ChiA74, an endochitinase native to *Bacillus thuringiensis*." Lett Appl Microbiol **51**(2): 184-190.

Palmer, K. L., K. Carniol, J. M. Manson, D. Heiman, T. Shea, S. Young, Q. Zeng, D. Gevers, M. Feldgarden, B. Birren and M. S. Gilmore (2010). "High-quality draft genome sequences of 28 *Enterococcus* sp. isolates." J Bacteriol **192**(9): 2469-2470.

Palmer, K. L., V. N. Kos and M. S. Gilmore (2010). "Horizontal gene transfer and the genomics of enterococcal antibiotic resistance." Curr Opin Microbiol **13**(5): 632-639.

Patti, J. M., B. L. Allen, M. J. McGavin and M. Hook (1994). "MSCRAMM-mediated adherence of microorganisms to host tissues." Annu Rev Microbiol **48**: 585-617.

Paulsen, I. T., L. Banerjei, G. S. Myers, K. E. Nelson, R. Seshadri, T. D. Read, D. E. Fouts, J. A. Eisen, S. R. Gill, J. F. Heidelberg, H. Tettelin, R. J. Dodson, L. Umayam, L. Brinkac, M. Beanan, S. Daugherty, R. T. DeBoy, S. Durkin, J. Kolonay, R. Madupu, W. Nelson, J. Vamathevan, B. Tran, J. Upton, T. Hansen, J. Shetty, H. Khouri, T. Utterback, D. Radune, K. A. Ketchum, B. A. Dougherty and C. M. Fraser (2003). "Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*." Science **299**(5615): 2071-2074.

Pizza, M., V. Scarlato, V. Masignani, M. M. Giuliani, B. Arico, M. Comanducci, G. T. Jennings, L. Baldi, E. Bartolini, B. Capecchi, C. L. Galeotti, E. Luzzi, R. Manetti, E.

Marchetti, M. Mora, S. Nuti, G. Ratti, L. Santini, S. Savino, M. Scarselli, E. Storni, P. Zuo, M. Broeker, E. Hundt, B. Knapp, E. Blair, T. Mason, H. Tettelin, D. W. Hood, A. C. Jeffries, N. J. Saunders, D. M. Granoff, J. C. Venter, E. R. Moxon, G. Grandi and R. Rappuoli (2000). "Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing." Science **287**(5459): 1816-1820.

Qin, X., K. V. Singh, G. M. Weinstock and B. E. Murray (2001). "Characterization of fsr, a regulator controlling expression of gelatinase and serine protease in *Enterococcus faecalis* OG1RF." J Bacteriol **183**(11): 3372-3382.

Quinlan, R. J., M. D. Sweeney, L. Lo Leggio, H. Otten, J. C. Poulsen, K. S. Johansen, K. B. Krogh, C. I. Jorgensen, M. Tovborg, A. Anthonsen, T. Tryfona, C. P. Walter, P. Dupree, F. Xu, G. J. Davies and P. H. Walton (2011). "Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components." Proc Natl Acad Sci U S A.

Reffuveille, F., C. Leneveu, S. Chevalier, Y. Auffray and A. Rince (2011). "Lipoproteins of *Enterococcus faecalis*: bioinformatic identification, expression analysis and relation to virulence." Microbiology.

Rich, R. L., B. Kreikemeyer, R. T. Owens, S. LaBrenz, S. V. Narayana, G. M. Weinstock, B. E. Murray and M. Hook (1999). "Ace is a collagen-binding MSCRAMM from *Enterococcus faecalis*." J Biol Chem **274**(38): 26939-26945.

Rince, A., Y. Le Breton, N. Verneuil, J. C. Giard, A. Hartke and Y. Auffray (2003). "Physiological and molecular aspects of bile salt response in *Enterococcus faecalis*." Int J Food Microbiol **88**(2-3): 207-213.

Roberts, G., K. A. Homer, E. Tarelli, J. Philpott-Howard, L. A. Devriese and D. Beighton (2001). "Distribution of endo-beta-N-acetylglucosaminidase amongst enterococci." J Med Microbiol **50**(7): 620-626.

Roberts, G., E. Tarelli, K. A. Homer, J. Philpott-Howard and D. Beighton (2000). "Production of an endo-*β*-*N*-acetylglucosaminidase activity mediates growth of *Enterococcus faecalis* on a high-mannose-type glycoprotein." J Bacteriol **182**(4): 882-890.

Rodriguez-Ortega, M. J., N. Norais, G. Bensi, S. Liberatori, S. Capo, M. Mora, M. Scarselli, F. Doro, G. Ferrari, I. Garaguso, T. Maggi, A. Neumann, A. Covre, J. L. Telford and G. Grandi (2006). "Characterization and identification of vaccine candidate proteins through analysis of the group A *Streptococcus* surface proteome." Nature Biotechnology **24**(2): 191-197.

Ruas-Madiedo, P., M. Gueimonde, M. Fernandez-Garcia, C. G. de los Reyes-Gavilan and A. Margolles (2008). "Mucin degradation by *Bifidobacterium* strains isolated from the human intestinal microbiota." Appl Environ Microbiol **74**(6): 1936-1940.

Rudd, P. M., T. Elliott, P. Cresswell, I. A. Wilson and R. A. Dwek (2001). "Glycosylation and the immune system." Science **291**(5512): 2370-2376.

Ruiz, L., M. Gueimonde, Y. Coute, S. Salminen, J. C. Sanchez, C. G. de los Reyes-Gavilan and A. Margolles (2011). "Evaluation of the ability of *Bifidobacterium longum* to metabolize human intestinal mucus." FEMS Microbiol Lett **314**(2): 125-130.

Sahm, D. F., J. Kissinger, M. S. Gilmore, P. R. Murray, R. Mulder, J. Solliday and B. Clarke (1989). "In vitro susceptibility studies of vancomycin-resistant *Enterococcus faecalis*." Antimicrob Agents Ch **33**(9): 1588-1591.

Sanchez, B., S. Arias, S. Chaignepain, M. Denayrolles, J. M. Schmitter, P. Bressollier and M. C. Urdaci (2009). "Identification of surface proteins involved in the adhesion of a probiotic *Bacillus cereus* strain to mucin and fibronectin." Microbiology **155**(Pt 5): 1708-1716.

Sanchez, B., M. C. Champomier-Verges, P. Anglade, F. Baraige, C. G. de Los Reyes-Gavilan, A. Margolles and M. Zagorec (2005). "Proteomic analysis of global changes

in protein expression during bile salt exposure of *Bifidobacterium longum* NCIMB 8809." J Bacteriol **187**(16): 5799-5808.

Sanchez, B., M. C. Champomier-Verges, B. Stuer-Lauridsen, P. Ruas-Madiedo, P. Anglade, F. Baraige, C. G. de los Reyes-Gavilan, E. Johansen, M. Zagorec and A. Margolles (2007). "Adaptation and response of *Bifidobacterium animalis* subsp. *lactis* to bile: a proteomic and physiological approach." Appl Environ Microb **73**(21): 6757-6767.

Sanchez, B., C. G. de los Reyes-Gavilan and A. Margolles (2006). "The F1F0-ATPase of Bifidobacterium animalis is involved in bile tolerance." Environ Microbi **8**(10): 1825-1833.

Sanchez, B., C. Gonzalez-Tejedo, P. Ruas-Madiedo, M. C. Urdaci and A. Margolles (2011). "*Lactobacillus plantarum* extracellular chitin-binding protein and its role in the interaction between chitin, Caco-2 cells, and mucin." Appl Environ Microbiol **77**(3): 1123-1126.

Sanchez, B., N. Saad, J. M. Schmitter, P. Bressollier and M. C. Urdaci (2010). "Adhesive properties, extracellular protein production, and metabolism in the *Lactobacillus rhamnosus* GG strain when grown in the presence of mucin." J Microbiol Biotechnol **20**(6): 978-984.

Sardinas, G., Y. Climent, Y. Rodriguez, S. Gonzalez, D. Garcia, K. Cobas, E. Caballero, Y. Perez, C. Brookes, S. Taylor, A. Gorringe, M. Delgado, R. Pajon and D. Yero (2009). "Assessment of vaccine potential of the Neisseria-specific protein NMB0938." Vaccine **27**(49): 6910-6917.

Sava, I. G., E. Heikens and J. Huebner (2010). "Pathogenesis and immunity in enterococcal infections." Clin Microbiol Infect **16**(6): 533-540.

Schaumburg, J., O. Diekmann, P. Hagendorff, S. Bergmann, M. Rohde, S. Hammerschmidt, L. Jansch, J. Wehland and U. Karst (2004). "The cell wall subproteome of *Listeria monocytogenes*." Proteomics **4**(10): 2991-3006.

Schluter, S., C. M. Franz, F. Gesellchen, O. Bertinetti, F. W. Herberg and F. R. Schmidt (2009). "The high biofilm-encoding Bee locus: a second pilus gene cluster in *Enterococcus faecalis*?" Curr Microbiol **59**(2): 206-211.

Sebald, W., P. Friedl, H. U. Schairer and J. Hoppe (1982). "Structure and genetics of the H+-conducting F0 portion of the ATP synthase." Ann N Y Acad Sci **402**: 28-44.

Seepersaud, R., S. B. Hanniffy, P. Mayne, P. Sizer, R. Le Page and J. M. Wells (2005). "Characterization of a novel leucine-rich repeat protein antigen from group B streptococci that elicits protective immunity." Infect Immun **73**(3): 1671-1683.

Severin, A., E. Nickbarg, J. Wooters, S. A. Quazi, Y. V. Matsuka, E. Murphy, I. K. Moutsatsos, R. J. Zagursky and S. B. Olmsted (2007). "Proteomic analysis and identification of *Streptococcus pyogenes* surface-associated proteins." J Bacteriol **189**(5): 1514-1522.

Shankar, V., A. S. Baghdayan, M. M. Huycke, G. Lindahl and M. S. Gilmore (1999). "Infection-derived *Enterococcus faecalis* strains are enriched in esp, a gene encoding a novel surface protein." Infect Immun **67**(1): 193-200.

Sherman, J. M. (1937). "The Streptococci." Bacteriol Rev **1**(1): 3-97.

Shuhui, L., Y. K. Mok and W. S. Wong (2009). "Role of mammalian chitinases in asthma." Int Arch Allergy Immunol **149**(4): 369-377.

Sillanpaa, J., Y. Xu, S. R. Nallapareddy, B. E. Murray and M. Hook (2004). "A family of putative MSCRAMMs from *Enterococcus faecalis*." Microbiology **150**(Pt 7): 2069-2078.

Simonetta, A. C., L. G. Moragues de Velasco and L. N. Frison (1997). "Antibacterial activity of enterococci strains against *Vibrio cholerae*." Lett Appl Microbiol **24**(2): 139-143.

Singh, K. V., S. R. Nallapareddy and B. E. Murray (2007). "Importance of the ebp (endocarditis- and biofilm-associated pilus) locus in the pathogenesis of *Enterococcus faecalis* ascending urinary tract infection." J Infect Dis **195**(11): 1671-1677.

Sleytr, U. B. and K. J. Thorne (1976). "Chemical characterization of the regularly arranged surface layers of *Clostridium thermosaccharolyticum* and *Clostridium thermohydrosulfuricum*." J Bacteriol **126**(1): 377-383.

Soderblom, T., O. Aspevall, M. Erntell, G. Hedin, D. Heimer, I. Hokeberg, K. Kidd-Ljunggren, A. Melhus, B. Olsson-Liljequist, I. Sjogren, J. Smedjegard, J. Struwe, S. Sylvan, K. Tegmark-Wisell and M. Thore (2010). "Alarming spread of vancomycin resistant enterococci in Sweden since 2007." Euro Surveill **15**(29).

Solheim, M., A. Aakra, H. Vebo, L. Snipen and I. F. Nes (2007). "Transcriptional responses of *Enterococcus faecalis V583* to bovine bile and sodium dodecyl sulfate." Appl Environ Microb **73**(18): 5767-5774.

Sood, S., M. Malhotra, B. K. Das and A. Kapil (2008). "Enterococcal infections & antimicrobial resistance." Indian J Med Res **128**(2): 111-121.

Spiro, R. G. (2002). "Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds." Glycobiology **12**(4): 43R-56R.

Stevens, S. X., H. G. Jensen, B. D. Jett and M. S. Gilmore (1992). "A hemolysin-encoding plasmid contributes to bacterial virulence in experimental *Enterococcus faecalis* endophthalmitis." Invest Ophthalmol Vis Sci **33**(5): 1650-1656.

Suzuki, K., M. Suzuki, M. Taiyoji, N. Nikaidou and T. Watanabe (1998). "Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170." Biosci Biotech Bioch **62**(1): 128-135.

Taira, T., N. Toma and M. Ishihara (2005). "Purification, characterization, and antifungal activity of chitinases from pineapple (Ananas comosus) leaf." Biosci Biotech Bioch **69**(1): 189-196.

Taranto, M. P., M. L. Fernandez Murga, G. Lorca and G. F. de Valdez (2003). "Bile salts and cholesterol induce changes in the lipid cell membrane of *Lactobacillus reuteri*." J Appl Microb **95**(1): 86-91.

Tarelli, E., H. L. Byers, K. A. Homer and D. Beighton (1998). "Evidence for mannosidase activities in Streptococcus oralis when grown on glycoproteins as carbohydrate source." Carbohydr Res **312**(3): 159-164.

Tendolkar, P. M., A. S. Baghdayan and N. Shankar (2003). "Pathogenic enterococci: new developments in the 21st century." Cell Mol Life Sci **60**(12): 2622-2636.

Teng, F., K. D. Jacques-Palaz, G. M. Weinstock and B. E. Murray (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**(4): 2010-5.

Theppangna, W., T. Murase, N. Tokumaru, H. Chikumi, E. Shimizu and K. Otsuki (2007). "Screening of the enterocin genes and antimicrobial activity against pathogenic bacteria in *Enterococcus* strains obtained from different origins." J Vet Med Sci **69**(12): 1235-1239.

Thurlow, L. R., V. C. Thomas, S. D. Fleming and L. E. Hancock (2009). "*Enterococcus faecalis* capsular polysaccharide serotypes C and D and their contributions to host innate immune evasion." Infect Immun **77**(12): 5551-5557.

Tjalsma, H., H. Antelmann, J. D. Jongbloed, P. G. Braun, E. Darmon, R. Dorenbos, J. Y. Dubois, H. Westers, G. Zanen, W. J. Quax, O. P. Kuipers, S. Bron, M. Hecker and J. M. van Dijl (2004). "Proteomics of protein secretion by *Bacillus subtilis*: separating the "secrets" of the secretome." Microbiol Mol Biol R **68**(2): 207-233.

Tjalsma, H., L. Lambooy, P. W. Hermans and D. W. Swinkels (2008). "Shedding & shaving: disclosure of proteomic expressions on a bacterial face." Proteomics **8**(7): 1415-1428.

Toledo-Arana, A., J. Valle, C. Solano, M. J. Arrizubieta, C. Cucarella, M. Lamata, B. Amorena, J. Leiva, J. R. Penades and I. Lasa (2001). "The enterococcal surface protein, Esp, is involved in *Enterococcus faecalis* biofilm formation." Appl Environ Microbiol **67**(10): 4538-4545.

Trost, M., D. Wehmhoner, U. Karst, G. Dieterich, J. Wehland and L. Jansch (2005). "Comparative proteome analysis of secretory proteins from pathogenic and nonpathogenic *Listeria* species." Proteomics **5**(6): 1544-1557.

Ueda, M., A. Fujiwara, T. Kawaguchi and M. Arai (1995). "Purification and some properties of six chitinases from *Aeromonas* sp. no. 10S-24." Biosci Biotech Bioch **59**(11): 2162-2164.

van de Guchte, M., P. Serror, C. Chervaux, T. Smokvina, S. D. Ehrlich and E. Maguin (2002). "Stress responses in lactic acid bacteria." Anton Van Leeuwe **82**(1-4): 187-216.

van Schaik, W. and R. J. Willems (2010). "Genome-based insights into the evolution of enterococci." Clin Microbiol Infect. **16**(6): 527-532.

van Aalten, D. M., D. Komander, B. Synstad, S. Gaseidnes, M. G. Peter and V. G. Eijsink (2001). "Structural insights into the catalytic mechanism of a family 18 exo-chitinase." Proc Natl Acad Sci U S A **98**(16): 8979-8984.

Vebo, H. C., L. Snipen, I. F. Nes and D. A. Brede (2009). "The transcriptome of the nosocomial pathogen *Enterococcus faecalis* V583 reveals adaptive responses to growth in blood." PLoS One **4**(11): e7660.

Vebo, H. C., M. Solheim, L. Snipen, I. F. Nes and D. A. Brede (2010). "Comparative genomic analysis of pathogenic and probiotic *Enterococcus faecalis* isolates, and their transcriptional responses to growth in human urine." PLoS One **5**(8): e12489.

Verneuil, N., A. Maze, M. Sanguinetti, J. M. Laplace, A. Benachour, Y. Auffray, J. C. Giard and A. Hartke (2006). "Implication of (Mn)superoxide dismutase of *Enterococcus faecalis* in oxidative stress responses and survival inside macrophages." Microbiology **152**(Pt 9): 2579-2589.

Vaaje-Kolstad, G., A. C. Bunaes, G. Mathiesen and V. G. Eijsink (2009). "The chitinolytic system of *Lactococcus lactis* ssp. *lactis* comprises a nonprocessive chitinase and a chitin-binding protein that promotes the degradation of alpha- and beta-chitin."FEBS J **276**(8): 2402-2415.

Vaaje-Kolstad, G., S. J. Horn, D. M. van Aalten, B. Synstad and V. G. Eijsink (2005). "The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation." J Biol Chem **280**(31): 28492-28497.

Vaaje-Kolstad, G., D. R. Houston, A. H. Riemen, V. G. Eijsink and D. M. van Aalten (2005). "Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21."J Biol Chem **280**(12): 11313-11319.

Vaaje-Kolstad, G., L. A. Bøhle, S. Gåseidnes, B. Dalhus, M. Bjørås, G. Mathiesen and V. G.H. Eijsink. (2011). "Characterization of the chitinolytic machinery of Enterococcus faecalis V583 and high resolution structure of its oxidative CBM33 enzyme." Submitted for publication to Journal of Molecular Biology.

Vaaje-Kolstad, G., B. Westereng, S. J. Horn, Z. Liu, H. Zhai, M. Sorlie and V. G. Eijsink (2010). "An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides." Science **330**(6001): 219-222.

Waters, C. M. and G. M. Dunny (2001). "Analysis of functional domains of the *Enterococcus faecalis* pheromone-induced surface protein aggregation substance." J Bacteriol **183**(19): 5659-5667.

Westereng, B., T. Ishida, G. Vaaje-Kolstad, M. Wu, V. G. H. Eijsink, K. Igarashi, M. Samejima, J. Stålberg, S. J. Horn, M. Sandgren (2011). "The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose." PLOS One, submitted for publication.

Whitehead, K., J. Versalovic, S. Roos and R. A. Britton (2008). "Genomic and genetic characterization of the bile stress response of probiotic *Lactobacillus reuteri* ATCC 55730." Appl Environ Microb **74**(6): 1812-1819.

Williams, S. J. and G. J. Davies (2001). "Protein-carbohydrate interactions: learning lessons from nature." Trends Biotechnol **19**(9): 356-362.

Wu, R., Z. Sun, J. Wu, H. Meng and H. Zhang (2010). "Effect of bile salts stress on protein synthesis of *Lactobacillus casei* revealed by 2-dimensional gel electrophoresis." J Dairy Sci **93**(8): 3858-3868.

Xia, G., C. Jin, J. Zhou, S. Yang and S. Zhang (2001). "A novel chitinase having a unique mode of action from *Aspergillus fumigatus* YJ-407." Eur J Biochem **268**(14): 4079-4085.

Xu, J., M. A. Mahowald, R. E. Ley, C. A. Lozupone, M. Hamady, E. C. Martens, B. Henrissat, P. M. Coutinho, P. Minx, P. Latreille, H. Cordum, A. Van Brunt, K. Kim, R. S. Fulton, L. A. Fulton, S. W. Clifton, R. K. Wilson, R. D. Knight and J. I. Gordon (2007). "Evolution of symbiotic bacteria in the distal human intestine." PLoS Biol **5**(7): e156.

Xu, Y., B. E. Murray and G. M. Weinstock (1998). "A cluster of genes involved in polysaccharide biosynthesis from *Enterococcus faecalis* OG1RF." Infect Immun **66**(9): 4313-4323.

Yamagami, T. and G. Funatsu (1993). "Purification and some properties of three chitinases from the seeds of rye (Secale cereale)." Biosci Biotech Bioch **57**(4): 643-647.

Yano, S., N. Rattanakit, A. Honda, Y. Noda, M. Wakayama, A. Plikomol and T. Tachiki (2008). "Purification and characterization of chitinase A of *Streptomyces cyaneus* SP-27: an enzyme participates in protoplast formation from *Schizophyllum commune* mycelia." Biosci Biotech Bioch **72**(1): 54-61.

Zechel, D. L. and S. G. Withers (2000). "Glycosidase mechanisms: anatomy of a finely tuned catalyst." Accounts of chemical research **33**(1): 11-18.

Zeng, J., F. Teng and B. E. Murray (2005). "Gelatinase is important for translocation of *Enterococcus faecalis* across polarized human enterocyte-like T84 cells." Infect Immun **73**(3): 1606-1612.

Zhou, M., J. Boekhorst, C. Francke and R. J. Siezen (2008). "LocateP: genome-scale subcellular-location predictor for bacterial proteins." BMC Bioinformatics **9**: 173.

Aam, B. B., E. B. Heggset, A. L. Norberg, M. Sorlie, K. M. Varum and V. G. Eijsink (2010). "Production of chitooligosaccharides and their potential applications in medicine." Marine Drugs **8**(5): 1482-1517.

# PAPER I

PROTEOME SCIENCE

RESEARCH Open Access

# Identification of proteins related to the stress response in *Enterococcus faecalis* V583 caused by bovine bile

Liv Anette Bøhle[1], Ellen M Færgestad[2], Eva Veiseth-Kent[2], Hilde Steinmoen[1], Ingolf F Nes[1], Vincent GH Eijsink[1] and Geir Mathiesen*[1]

## Abstract

**Background:** *Enterococcus faecalis* is an opportunistic pathogen and one of the most important causes of hospital infections. Bile acids are a major stress factor bacteria have to cope with in order to colonize and survive in the gastrointestinal tract. The aim of this study was to investigate the effects of bile acids on the intracellular proteome of *E. faecalis* V583.

**Results:** The proteomes of cells challenged with 1% bile were analyzed after 20 - 120 minutes exposure, using 2D gel electrophoresis and mass spectrometry. Among the approximately 500 observed proteins, 53 unique proteins were found to be regulated in response to bile and were identified with mass spectrometry. The identified proteins belonged to nine different functional classes, including fatty acid- and phospholipid-biosynthesis, energy metabolism, and transport and binding. Proteins involved in fatty acid and phospholipid biosynthesis pathways were clearly overrepresented among the identified proteins and all were down-regulated upon exposure to bile. The proteome data correlated reasonably well with data from previous transcriptome experiments done under the same conditions, but several differences were observed.

**Conclusion:** The results provide an overview of potentially important proteins that *E. faecalis* V583 needs to regulate in order to survive and adapt to a bile-rich environment, among which are several proteins involved in fatty acid and phospholipid biosynthesis pathways. In addition, this study reveals several hypothetical proteins, which are both abundant and clearly regulated and thus stand out as targets for future studies on bile stress.

## Background

*Enterococcus faecalis* is a wide-spread Gram-positive lactic acid bacterium, and is a natural inhabitant of the gastrointestinal tract (GIT) of humans and animals. The bacterium is also commonly found in soil, sewage, water and food. *E. faecalis* V583 is an opportunistic pathogen that can cause diseases like urinary tract infections, bacteremia, and infective endocarditis in immunocompromised patients. These infections may be problematic because *E. faecalis* strains tend to be resistant toward many antibiotics, including vancomycin [1,2]. Vancomycin-resistant enterococci were first found among clinical isolates in the late 1980s, and antibiotic resistance has increased since. Infections by enterococci have become a major problem in the hospital environments and enterococci are now ranked among the most prevalent nosocomial pathogens [3,4].

*E. faecalis* is able to grow and colonize many hostile environments including the GIT, and is considered as an interesting model for studying bacterial stress responses [5]. It is important to understand such responses in enterococci because the ability to survive in a wide range of environments obviously contributes to enterococcal prevalence in e.g. hospital environments. In order to survive in the human GIT, bacteria must overcome several adverse environmental stresses such as low pH, low oxygen levels, nutrient limitations, elevated osmolarity and the deleterious actions of bile. The liver daily secretes about one liter of bile, which consists mainly of bile acids,

* Correspondence: geir.mathiesen@umb.no
[1] Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Chr. M. Falsensvei 1, N-1432 Ås, Norway
Full list of author information is available at the end of the article

cholesterol, phospholipids, and the pigment biliverdin. In the human GIT, bile acts as a biological detergent, emulsifying and solubilising fats [6]. Exposure to bile may lead to changes in the fatty acid- and phospholipid-composition of bacterial membranes and to distortion of the cell surface [7,8]. It has also been shown that bile can induce secondary structure formation in RNA, induce DNA damage and activate enzymes involved in DNA repair [6].

Various studies indicate that bacteria thriving in the GIT, such as lactic acid bacteria and bifidobacteria, have evolved mechanisms to protect themselves from the noxious effects of bile. Genome-wide and gene-by-gene studies have shown that Gram-positive bacteria such as listeria, lactobacilli and bifidobacteria carry genes coding for transporters able to extrude bile salts, the expression of which is regulated by bile salts [9-11]. Additional genes whose expression is regulated by bile include genes involved in more general stress responses and genes involved in carbohydrate metabolism and fatty acid biosynthesis [12-15].

While genome and transcriptome analyses have provided interesting clues as to how enterococci manage bile stress, so far, proteome information is limited [16-18]. Since transcriptome and proteome data do not necessarily correlate, knowledge of proteome responses is a prerequisite for obtaining a more complete picture of the bile salt response. Therefore, we have analyzed how the intracellular proteome of *E. faecalis* V583 responds to bovine bile in a time course experiment, assessing both immediate responses and longer term effects. We used two-dimensional- (2D) gel electrophoresis combined with mass spectrometry-based protein identification to identify the bacterial proteins whose abundances change during growth in presence of 1% bovine bile. The experiments were conducted using conditions identical to those used in a previous transcriptome study on bile stress [15] permitting comparison of proteome and transcriptome data.

## Materials and methods

### Bacterial strain and sample collection

*E. faecalis* V583 [19] was grown in brain heart infusion (BHI) medium (Oxoid Ltd., Hampshire, England) aerobically, over-night, with shaking, 300 rpm, at 37°C. The over-night cultures (50 ml) were diluted 50-fold in 100 ml BHI (37°C) and grown further to an $OD_{600}$ of ~0.2, after which the cells were collected by centrifugation (9800 × g, 10 min, at room-temperature). After resuspending the cells in 50 ml of fresh prewarmed BHI (37°C) another 50 ml of prewarmed BHI-medium with or without bovine bile (Sigma-Aldrich Inc, St. Louis, MO) was added. Thus, the final volume was 100 ml for each condition and the final concentration for the samples with bovine bile was 1% (w/v). The cultures were incubated with shaking, 300 rpm, at 37°C, until harvesting. After incubation for 20, 60 or 120 minutes cells were harvested from the cultures by centrifugation (9800 × g, 10 min, 4°C). The cell pellets were washed three times with ice-cold 0.9% (w/v) NaCl, and subsequently resuspended in a lysis- and rehydration-solution containing 8 M urea, 2 M thiourea, 50 mM DTT and 2% CHAPS. The cells were disrupted with glass-beads (106 micron, Sigma-Aldrich) using a Fast-Prep-24 instrument (MP Biomedicals, Solon, OH) (speed 6, three treatments of 45 seconds each, with 60 second pauses in between, at 4°C). The resulting cell-free extracts were stored at -20°C until the 2D analyses were performed (see below). The protein concentration of the cell-free extracts was measured using the Bradford Microassay (Bio-Rad Laboratories, Inc, Hercules, CA).

To check the effect of bile on the growth rate of *E. faecalis* V583, the bacteria were grown in liquid BHI with or without 1% bile, at 300 rpm, 37°C. The growth was followed by measuring the optical density ($OD_{600}$) of the cultures every 20 minutes in the first 3 hours, and then every 30 minutes.

### Two-dimensional gel electrophoresis and analysis

Isoelectric focusing (IEF) was performed using 24 cm IPG-strips (Bio-Rad) covering the pH-area 4-7. The cell-free protein sample was diluted in a lysis-and rehydration-solution together with ampholytes in a total volume of 450 μl, containing 75 μg proteins. This protein-sample was loaded on the strips, and rehydration was performed for 16 hours at 50 V using a Protean IsoElectric Focusing Cell II unit (Bio-Rad). Then, isoelectric focusing was carried out using the following voltage program: a linear ramp to 250 V over 30 minutes, followed by a rapid ramp to 500 V over 1 hour, a rapid ramp to 1000 V over 1 hour, and then a rapid ramp to 10000 V, where the voltage is fixed until the total reached 70000 Vh.

To prepare for electrophoresis in the second dimension, the IPG-strips were first equilibrated in a buffer containing 6 M urea, 50 mM Tris-HCl pH 8.8, 30% glycerol, 2% SDS and 1% dithiothreitol (DTT) for 15 minutes to reduce sulfhydryl groups, and then in a buffer containing 6 M urea, 50 mM Tris-HCl pH 8.8, 30% glycerol, 2% SDS and 5% iodoacetamide for an additional 15 minutes to alkylate the reduced sulfhydryl groups. Subsequently, the strips were loaded onto a 12.5% SDS-gel for the SDS-PAGE step. The electrophoresis was run at 5 mAmp/gel for 3 hours and then at 15 mAmp/gel for 12 hours. Normally, 12 gels were run simultaneously, using an Ettan-DALT Electrophoresis system (Amersham Biosciences, USA). The gels were silver stained according to a previously described method [20].

For each growth condition/time point (six combinations in total), five biological replicates were produced. These 30 samples were run twice through the 2D-electrophoresis, yielding a total of 60 gels. Spot detection and gel alignment were performed using Delta2D software (DECODON, Greifswald, Germany). In this method all pixels are compiled into a fused image that is used for setting common spot boundaries for all gel samples, whereas the raw data for the spot intensities are obtained from each individual image [21]. All proteins were visually checked and only spots that did not appear as streaking were included (this procedure was repeated after preselection of regulated proteins; see Results section). In the further analyses only spots showing a fold change of at least 1.9 (with bile versus without bile) at at least one time point were included. This somewhat arbitrary threshold value is comparable to the threshold factors used in other studies [22,23]. Lower threshold factors were not considered in order to maximize the chance that the discovered changes in protein levels are not only statistically but also biologically significant. Based on this criterion, 115 spots were selected and subjected for statistical analysis. Analysis of variance was performed where the p-values were adjusted for multiple comparisons by False Discovery Rate (FDR) using rotation test [24,25] with a significance level of $p < 0.05$. Upon this procedure, the number of spots was reduced to 91.

### Protein identification

Spots representing differently expressed proteins were excised from the gel and the gel pieces were washed 2 × 15 minutes with a 1:1 mixture of acetonitrile (ACN) and 50 mM ammonium bicarbonate. The gel-pieces were dried in a speed-vac and the dried gel-pieces were reswollen with 0.03 μg/μl trypsin at 4°C for 30 minutes (approximately 5 μl on average). Subsequently, 25 μl 50 mM ammonium bicarbonate was added, and the digestion of proteins was performed by incubating overnight at 37°C with shaking at 300 rpm. The peptides were extracted from the gel-pieces by incubating the pieces in 1% trifluoroacetic acid (TFA), 0.1% (v/v) TFA in 50% (v/v) ACN, and 100% ACN, in three consecutive steps. The liquid phases were pooled and the extracted peptides were dried in a speed-vac, rehydrated in 0.1% (v/v) TFA, and sonicated prior to desalting using C18 STAGE tips [26]. After eluting the peptides with 1 μl 65% (v/v) ACN, 0.5 μl of the eluate was mixed with 0.5 μl matrix solution (α-cyano-4-hydroxycinamic acid), and spotted onto a MALDI target plate (Bruker Daltonics, Billerica, MA).

Peptide mass fingerprints (PMF) and MS/MS fragmentation spectra were determined using an Ultraflex-MALDI-TOF/TOF (Bruker Daltonics) instrument. Protein identification was carried out with Mascot (Matrix Science Inc., Boston, MA), limiting the search to bacteria belonging to the "Other Firmicutes" in the NCBI database. The searches were limited to only consider tryptic fragments, with carbamidomethylation of cysteine (fixed modification) and possible oxidation of methionine. The error tolerance was set to 50 ppm, and the number of allowed missed cleavage sites was set to 2.

Some proteins giving weak spots could not be identified by MALDI-TOF/TOF. These proteins were analyzed by Nano LC coupled to ESI-MS/MS, using an LC-LTQ Orbitrap-MS at The Biotechnology Centre of Oslo, Norway.

## Results

The effect of bovine bile on the growth rate of *E. faecalis* V583 was tested by comparing growth in the absence and presence of 1% (w/v) bovine bile in liquid BHI medium (Fig. 1). The growth curves showed that the presence of bile reduces both the growth rate and the maximum cell density, but also showed that the levels of bile used in this study are clearly sub-lethal, permitting accumulation of sufficient cells in the samples with bile. Cells were harvested at 20, 60 and 120 minutes after the addition of bile.

Additional file 1 shows six representative gels from the 2D-gel electrophoresis experiments, one gel for each time point at each condition (presence or absence of bile). Visual inspection of the about 500 distinguishable spots on the gels showed some clear differences between the bile and non-bile samples. The data-set used for image analysis and identification of differentially expressed proteins was obtained from 60 gels. These gels represented five biological replicates for each growth condition/time point (30 samples in total), which were run twice. Statistical validation was performed based on the pixel values for

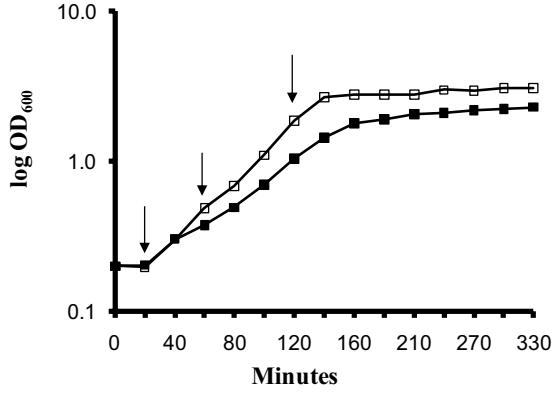



**Figure 1 Growth-curve of *E. faecalis* V583 in liquid BHI medium with (black square) or without 1% bovine bile (open square)**. Growth was measured as absorbance at 600 nm. Each time point represents the mean value of three biological replicates. The arrows indicate the time points (20, 60 and 120 minutes) where the cells were harvested for proteome analysis.

each protein spot. In total 91 spots met the two criteria of (1) showing significant changes according to the FDR adjusted p-value (<0.05) and (2) showing an increase or decrease of at least 1.9-fold in response to bile at at least one of the three time points. The 91 spots were visually controlled in more detail and proteins that appeared as streaking were removed from the analysis (32 spots; these are spots for which the image analysis is prone to errors). The resulting 59 spots were subjected to protein identification by mass spectroscopy. This analysis identified 55 proteins whose abundances were considered to change upon bile stress while four proteins could not be identified. Two proteins were identified twice: spot 16 and 19 were identified as formate acetyltransferase (EF1613), and spot 49 and 51 were identified as ribose-phosphate pyrophosphokinase (EF3163). In both cases, the two protein spots belonging to the same protein represented different masses and pI. Such differences may be explained by one of the spots being a proteolytic fragment or carrying another type of post-translational modification. Within each pair, the spots showed similar trends in terms of the change in protein abundance.

Of the 53 identified unique proteins, 17 were up-regulated, 32 were down-regulated, and four were both up- and down-regulated during the time course experiment (Table 1). The number of regulated proteins increased during the time course experiment, from eight after 20 minutes, to 21 after 60 minutes, to 38 after 120 minutes (Table 1). Only 14 proteins showed significantly altered abundances at more than one time point, indicating that the observed up- and down-regulations are relatively transient. Among the 53 unique proteins, 14 are annotated in the J. Craig Venter Institute (JCVI) Comprehensive microbial resource as hypothetical proteins or proteins with unknown functions (Table 1). The other 39 proteins belong to nine different functional classes (Table 1). Proteins involved in energy metabolism and in the transcription-translation process were most abundant, with eight and nine identified proteins in each functional class, but their relative abundance was comparable with that of proteins in most of the other functional classes (Fig. 2). These two most numerous groups did not show clear overall trends with respect to up- or down-regulation. Proteins involved in fatty acid and phospholipid metabolism were clearly overrepresented among the identified proteins (Fig. 2) and all of these were down-regulated by the presence of 1% bovine bile (Fig. 3). The rest of the identified proteins are predicted to be involved in processes related to protein fate, transport and binding, nucleotide metabolism, coenzyme transport and metabolism, amino acid biosynthesis, and cell wall/membrane biogenesis, as well as in "other cellular processes" related to bacterial adaptation to atypical conditions.

The growth conditions used in the present proteome study were similar to those used in a previous transcriptome study with the same bacterium [15]. The expression profiles from the transcriptome study [15] were compared with the data from the present study for the 19 proteins that were found to be significantly regulated in both studies (Fig. 4). While the expression profiles derived from the two methods show the same tendencies for about eleven of these 19 proteins, discrepancies are observed in several cases, either at one time point (e.g. EF0020 & EF1499 at 120 minutes) or in the form of an overall trend (e.g. EF3184 and EF3186 where an increase in transcription does not seem to lead to an increase in protein).

## Discussion

In the present study we used 2D gel electrophoresis to identify proteins in *E. faecalis* V583 whose levels are regulated in response to bile. The cells were analysed at three different time points. By using this set up both the initial and more long-term effects of bile could be observed.

The gels showed approximately 500 spots, which is 20% of the intracellular proteome of *E. faecalis* V583. Use of visual inspection, a threshold fold ratio of 1.9, and rigorous statistical analysis led to the identification of 53 unique proteins that are regulated in response to bile stress, representing about 10% of the proteins on the gels. This number of regulated proteins is higher than what has been obtained in similar proteomic studies in other Gram-positive bacteria. Previous studies have identified 28 proteins in *Lactobacillus reuteri* [27], 24 proteins in *Propionibacterium freudenreichii* [22], 45 proteins in *Bifidobacterium animalis* [28], and 34 proteins in *Bifidobacterium longum* [23]. The high number obtained in the present study may partly be due to the fact that we have studied the proteome in a time-course experiment, which, to the best of our knowledge, has not been done before in studies of bile responses in Gram-positive bacteria.

Bile stress in *Enterococcus faecalis* was among the first bacterial stress responses to be studied using the 2D electrophoresis approach. In an early study aimed at identifying stress-induced proteins, 45 protein spots were found to be up-regulated upon exposure to bile for 30 minutes [16]. Only two proteins, DnaK and GroEL, were identified and it is therefore difficult to compare this work with the present study (note that DnaK was also found to be upregulated in our study; see below). In a later study by Giard et al. [18], peptide sequences for four additional bile-induced proteins were presented which correspond to EF0453, a protein involved in hydrogen peroxide resistance, and to three hypothetical proteins (EF0770, EF1560, EF2798). None of these four proteins were found

**Table 1: Identified proteins of E. faecalis V583 whose abundance is affected by the presence of 1% bovine bile—the three columns to the right show the results for cells harvested at three different time points (20, 60 and 120 minutes)**

| Spot no. | Functional class[a] | Putative function[a] | ORF | Mass (kDa) | pI[b] | Coverage[c] | No. of peptides matched | Fold ratio[d] T(20) | T(60) | T(120) |
|---|---|---|---|---|---|---|---|---|---|---|
| **1**[e] | Fatty acid and phospholipid metabolsim | Enoyl-(acyl carrier protein) reductase | EF0282 | 26.9 | 5.29 | 42 | 14 | 1.01 | 0.67 | **0.3** |
| **2**[e] | | Acetyl_coA carboxylase, biotin carboxyl carrier protein | EF2879 | 17.6 | 4.23 | 40 | 8 | 0.92 | **0.39** | **0.28** |
| **3**[e] | | (3R)-hydroxymyristoyl-(acyl-carrier-protein) dehydratase | EF0284 | 16.4 | 5.73 | 22 | 5 | 1.09 | 0.71 | **0.49** |
| **4**[e] | | 3-oxoacyl-(acyl-carrier-protein) synthase II | EF0283 | 43.5 | 5.11 | 40 | 12 | **0.51** | 0.87 | **0.41** |
| **5**[e] | | 3-ketoacyl-(acyl-carrier-protein) reductase | EF2881 | 26.1 | 5.92 | 35 | 10 | 1.70 | **0.37** | 0.55 |
| **6** | Transport and binding protein | Phosphocarrier protein HPr | EF0709 | 9.3 | 4.92 | 13 | 11 | 0.99 | 0.70 | **0.52** |
| **7**[e] | | PTS system, mannose-specific IIAB components | EF0020 | 35.5 | 5.11 | 50 | 22 | 1.62 | **3.06** | 0.90 |
| **8** | | peptide ABC transporter, ATP-binding protein | EF0912 | 35.9 | 5.97 | 20 | 7 | 1.10 | 1,00 | **0.46** |
| **9**[e] | Amino acid biosynthesis | Ornithine carbamoyltransferase | EF0105 | 38.1 | 5.02 | 18 | 5 | 1.03 | 0.64 | **3.48** |
| **10**[e] | | Decarboxylase, putative | EF0634 | 71.96 | 5.14 | 47 | 29 | **0.52** | 1.21 | 1.86 |
| **11** | Other cellular processes | Glutathione reductase | EF3270 | 49.6 | 5.23 | 21 | 10 | **2.42** | 0.96 | 0.87 |
| **12** | | Dps family protein | EF3233 | 17.9 | 4.56 | 46 | 6 | **0.51** | 0.81 | 1.23 |
| **13** | | General stress protein, putative | EF1744 | 20.5 | 4.61 | 31 | 6 | 0.88 | 0.66 | **0.52** |
| **14** | | Cell division protein DivIVA | EF1002 | 26.6 | 4.53 | 62 | 16 | 0.61 | 1.21 | **3.5** |
| **15**[e] | Energy metabolism | Deoxyribose-phosphate aldolase | EF0174 | 23.3 | 4.65 | 48 | 10 | 1.39 | 1.28 | **0.47** |
| **16**[f] | | Formate acetyltransferase | EF1613 | 84.5 | 5.31 | 11 | 8 | 1.00 | 0.95 | **2.87** |
| **17**[e] | | Fumarate reductase flavoprotein subunit | EF2556 | 53.8 | 5.26 | 10 | 6 | 0.86 | 1.57 | **2.26** |
| **18** | | Phosphoglycerate mutase 1 | EF0195 | 26.0 | 5.09 | 41 | 11 | 0.87 | 1.13 | **0.36** |
| **19**[f] | | Formate acetyltransferase | EF1613 | 84.7 | 5.31 | 14 | 11 | 0.94 | 1.21 | **2.09** |
| **20**[e] | | V-type ATP synthase subunit B | EF1499 | 51.3 | 5.03 | 18 | 8 | 0.59 | **2.02** | 1.29 |
| **21** | | Pyruvate kinase | EF1046 | 62.6 | 4.99 | 41 | 25 | **2.99** | **3.61** | 1.09 |
| **22** | | Enolase | EF1961 | 46.5 | 4.56 | 54 | 22 | 1.46 | 0.98 | **0.45** |
| **23** | | Thioredoxin | EF1405 | 11.7 | 4.35 | 66 | 9 | 1.49 | **0.40** | 0.60 |
| **24**[e] | Cell wall/membrane biogenesis | D-fructose-6-phosphate amidotransferase | EF2151 | 65.7 | 4.93 | 24 | 12 | **1.93** | 1.13 | **2.12** |
| **25** | Coenzyme transport and metabolism | Naphthoate synthase | EF0445 | 30.0 | 5.24 | 39 | 10 | 0.89 | 1.46 | **0.43** |
| **26**[e] | | 2-dehydropantoate 2-reductase | EF2445 | 34.7 | 5.08 | 3 | 1 (39)[g] | 1.32 | 1.40 | **0.46** |
| **27**[e] | Transcription and translation | Transcriptional regulator, AraC family | EF0432 | 34.6 | 6.76 | 25 | 7 | 0.91 | **2.64** | **0.18** |

**Table 1: Identified proteins of E. faecalis V583 whose abundance is affected by the presence of 1% bovine bile—the three columns to the right show the results for cells harvested at three different time points (20, 60 and 120 minutes) (Continued)**

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **28** | | Cold-shock domain-contain protein | EF2925 | 7.3 | 4.35 | 60 | 5 | 0.97 | **0.33** | **0.37** |
| **29** | | Peptide deformylase | EF3066 | 21.0 | 5.08 | 48 | 6 | 1.15 | 0.77 | **0.5** |
| **30** | | 30S ribosomal protein S2 | EF2398 | 29.5 | 5 | 24 | 8 | **2.82** | 1.87 | 0.81 |
| **31**[e] | | Ribosomal protein L31 | EF1171 | 10.1 | 5.57 | 95 | 10 | 1.65 | **0.36** | **0.29** |
| **32** | | Elongation factor G | EF0200 | 76.6 | 4.8 | 23 | 19 | 0.93 | 0.87 | **1.97** |
| **33** | | 30S ribosomal protein S3 | EF0212 | 24.4 | 9.8 | 12 | 3 | 0.86 | **2.19** | 0.68 |
| **34** | | Phenylalanyl-tRNA synthetase subunit beta | EF1116 | 88.8 | 4.76 | 12 | 11 | 0.87 | 0.97 | **2.18** |
| **35** | | Methionyl-tRNA formyltransferase | EF3123 | 34.3 | 6.06 | 3 | 1 (48)[g] | 0.96 | 1.29 | **0.41** |
| **36** | Hypothetical protein | Hypothetical protein EF1967 | EF1967 | 20.7 | 5.98 | 20 | 5 | **0.51** | 0.86 | **0.25** |
| **37** | | Hypothetical protein EF2909 | EF2909 | 12.1 | 4.33 | 52 | 7 | 0.99 | **0.47** | 0.69 |
| **38** | | Hypothetical protein EF2763 | EF2763 | 12.1 | 4.7 | 43 | 5 | 1.25 | **0.43** | 0.66 |
| **39** | | Hypothetical protein EF2888 | EF2888 | 8.9 | 4.54 | 42 | 6 | 1.07 | **0.32** | **0.43** |
| **40**[e] | | Hypothetical protein EF3184 | EF3184 | 26.4 | 4.86 | 28 | 7 | 0.80 | **0.31** | **0.38** |
| **41**[e] | | Hypothetical protein EF3186 | EF3186 | 25.7 | 4.91 | 28 | 8 | 1.44 | 1.80 | **0.44** |
| **42** | | Hypothetical protein EF0123 | EF0123 | 85.5 | 7.23 | 4 | 3 | 1.89 | **2.11** | **0.43** |
| **43** | | Hypothetical protein EF0352 | EF0352 | 29.3 | 6.91 | 4 | 1 (61)[g] | 1.44 | **2.18** | **0.48** |
| **44** | | Hypothetical protein EF2174 | EF2174 | 99.6 | 8.68 | 36 | 33 | 1.54 | 1.41 | **0.32** |
| **45** | | Hypothetical protein EF 2104 | EF2104 | 43.6 | 4.88 | 18 | 6 | 0.82 | **2.36** | 1.17 |
| **46** | Protein fate | Glutamyl-aminopeptidase | EF3037 | 39.4 | 5.68 | 22 | 8 | 0.87 | **1.90** | 1.01 |
| **47**[e] | | DnaK protein | EF1308 | 65.5 | 4.59 | 56 | 31 | 1.11 | 1.20 | **2.21** |
| **48**[e] | | Heat shock protein GrpE | EF1307 | 20.1 | 4.5 | 26 | 8 | 1.10 | **0.30** | **0.30** |
| **49**[f] | Nucleotid metabolism | Ribose-phosphate pyrophosphokinase | EF3163 | 35.4 | 6.16 | 22 | 9 | 1.21 | **2.5** | **0.47** |
| **50** | | Inositol-5-monophosphate dehydrogenase | EF3293 | 52.8 | 5.7 | 50 | 84 | 1.24 | **2.0** | 1.05 |
| **51**[f] | | Ribose-phosphate pyrophosphokinase | EF3163 | 35.5 | 6.16 | 15 | 7 | 0.95 | 1.21 | **0.38** |
| **52** | Unknown function | DNAbinding response regulator VicR | EF1193 | 26.9 | 5.17 | 15 | 6 | 0.87 | 1.20 | **0.51** |
| **53** | | Glyoxalase family protein | EF2591 | 31.7 | 4.85 | 50 | 15 | 1.04 | 0.99 | **0.30** |
| **54** | | PhnA protein | EF1374 | 12.4 | 5.01 | 66 | 6 | 0.9 | 0.55 | **0.47** |
| **55** | | Oxidoreductase, aldo/keto reductase family | EF1138 | 31.0 | 5.28 | 18 | 5 | 0.97 | 1.22 | **0.44** |

[a] The putative function is based on the JCVI Comprehensive Microbial Resource database http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi.

[b] The pI values are theoretical values calculated from the protein sequences

[c] The coverage shows the percentage of the protein that is covered by the identified peptides.

[d] Proteins with values over 1 are up-regulated in response to bile, proteins with values below 1 are down-regulated. Values representing a change larger than 1.9-fold are printed in bold face.

[e] Proteins that also were shown to be regulated in a transcriptome study of bile responses in E. faecalis V583 [15]; see text for details.

[f] Note that proteins EF3163 (spots 49 & 51) and EF1613 (spots 16 & 19) were identified twice.

[g] The value in parenthesis shows the probability-based Mowse score for proteins that were identified by only one peptide (in all cases using the LC-LTQ Orbitrap). The Mowse score equals -10*Log (P), where P is the probability that the observed match is a random event; Mowse scores greater than 38 indicate identity or extensive homology ($p < 0.05$).

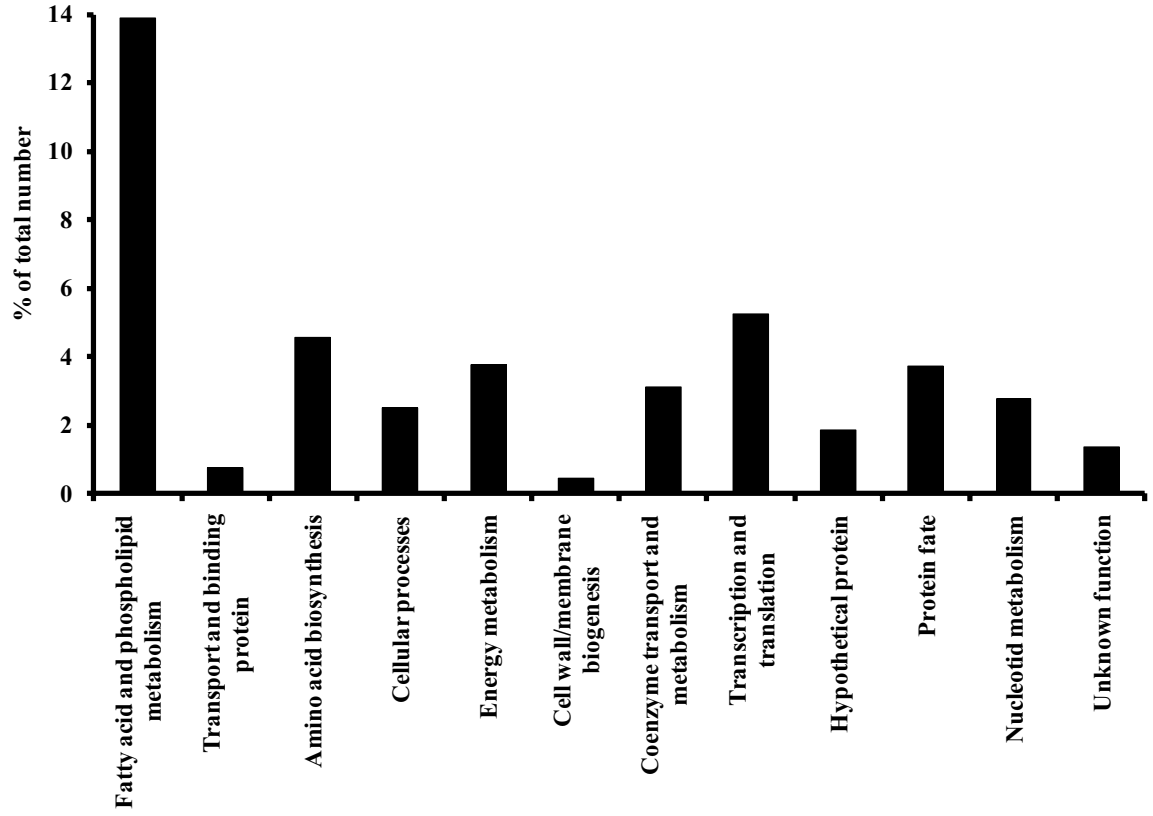



**Figure 2 Relative abundance of identified regulated proteins in *E. faecalis* V583**. Proteins (53 in total) are grouped according to their functional class as defined by JVCI http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi. The values for each class represent the fraction (in %) of the total expected proteome in that class that was identified in the present study.

to be significantly regulated in the present study. This difference may be due to the fact that the bile salt concentrations used in these previous studies were lower (0.08 or 0.3%). Furthermore, these studies used another *E. faecalis* strain which is less resistant to bile than the V583 strain.

The 2D-PAGE experiments provide information on changes in protein levels only and do not provide direct information on the underlying gene expression levels (mRNA production levels). The previous transcriptional study [15] on the same bacterium and under the same conditions, revealed 308 genes whose expression was affected by bile during the time course experiment, representing about 10% of the genes in the genome. The fraction of regulated genes in each functional class varied between 5.6% and 16.7% of the total number of genes per class, with the exception of genes involved in the fatty acid- and phospholipid metabolism, of which 41.7% were found to be regulated [15]. Compared to the transcriptome study, the present proteome study revealed fewer regulated genes/proteins; we identified about 2% of the total proteome and the representation of each functional class varied between 0.5% and 5.3% (Fig. 2). Again, proteins involved in fatty acid- and phospholipid metabolism formed the exception, since 14.0% of these were found to be regulated. The lower number of genes/proteins found in the proteome study is not only due to sensitivity and dynamic range issues but also to the fact that insoluble proteins and the major part of proteins that are secreted out of the cell are not expected to be found in the cytoplasmic fraction. The transcriptome and proteome data correlate reasonably well for 11 of the 19 genes/proteins that were identified with both methods (Fig. 4). Not unexpectedly, in several cases changes in protein levels lag behind changes in RNA levels. Several proteins show clear discrepancies between the proteome and the transcriptome data and in almost all of these cases an upregulation of the mRNA is not accompanied by higher protein levels (EF1499, EF2151, EF1171, EF0020, EF3184, EF3186). One may speculate that other regulatory mechanisms come into play, such as translational regulation or specific protein degradation. Two of these genes (EF3184, EF3186) are part of an operon consisting of five genes, of which four encode proteins with putative N-terminal signal-peptides. None of these proteins show homology with other proteins with known function. Clearly, the fact that these proteins are secreted may lead to discrepancies

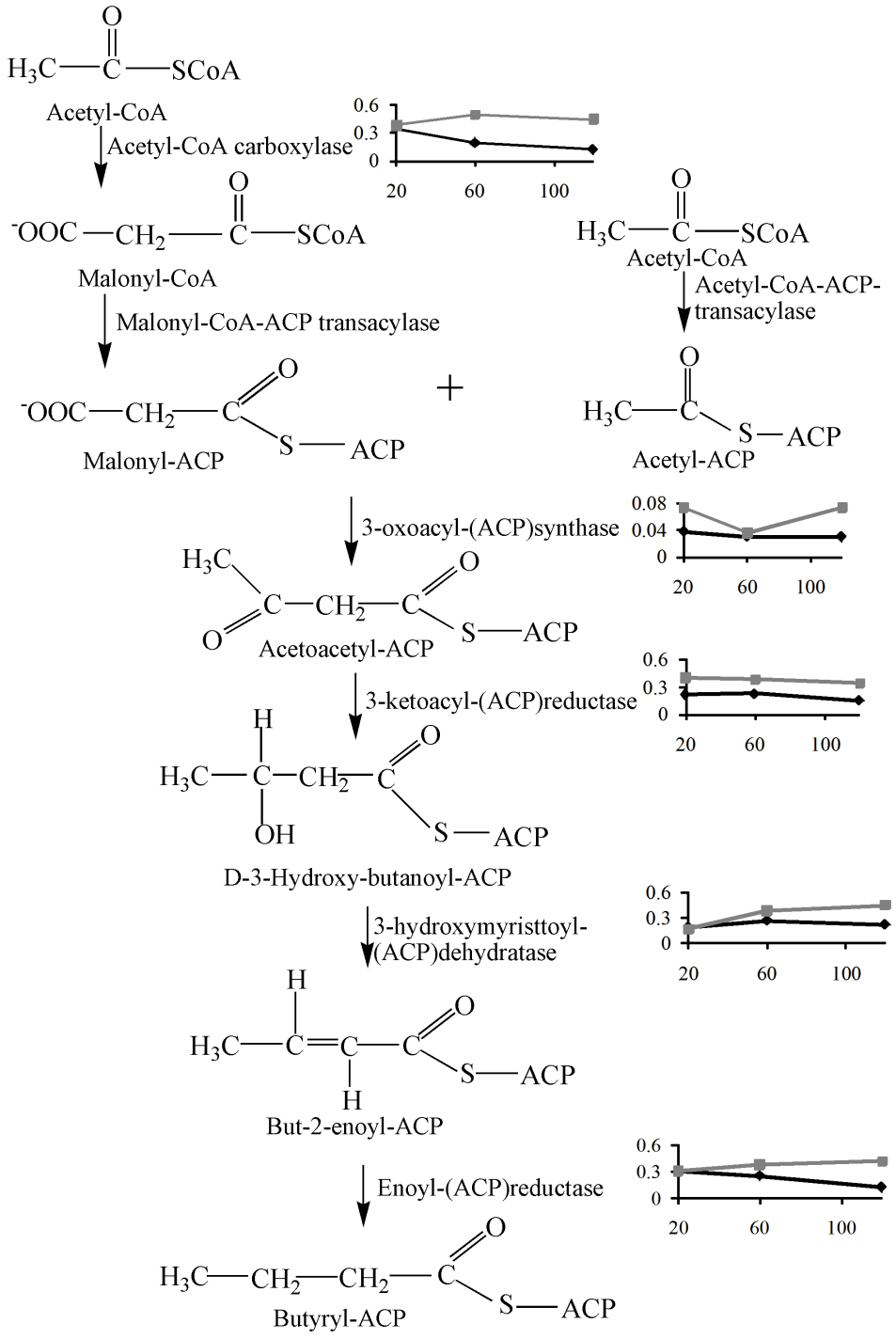



**Figure 3** Overview of fatty acid biosynthesis **http://www.genome.jp/kegg/pathway/map/map00061.html and regulation of key enzymes**. The starting reactant is acetyl-CoA. The principal reaction products of the fatty acid biosynthesis are straight-chain C16 and C18 fatty acids, which can be used in the synthesis and the repair of damaged phospholipid membranes. The resulting fatty acids are therefore important constituents of the cell membrane. The graphs to the right show the development of protein abundance during the time course experiment. The grey line corresponds to protein produced in bacteria grown in media without bovine bile, while the black line corresponds to bacteria grown in media containing 1% bovine bile. The x-axis indicates the time (minutes) and the y-axis the spot intensity (normalized raw spot volume) of the protein as a mean value of the parallels. The lines are drawn for illustration purposes only, connecting the three time points that were analyzed (20, 60 and 120 minutes).

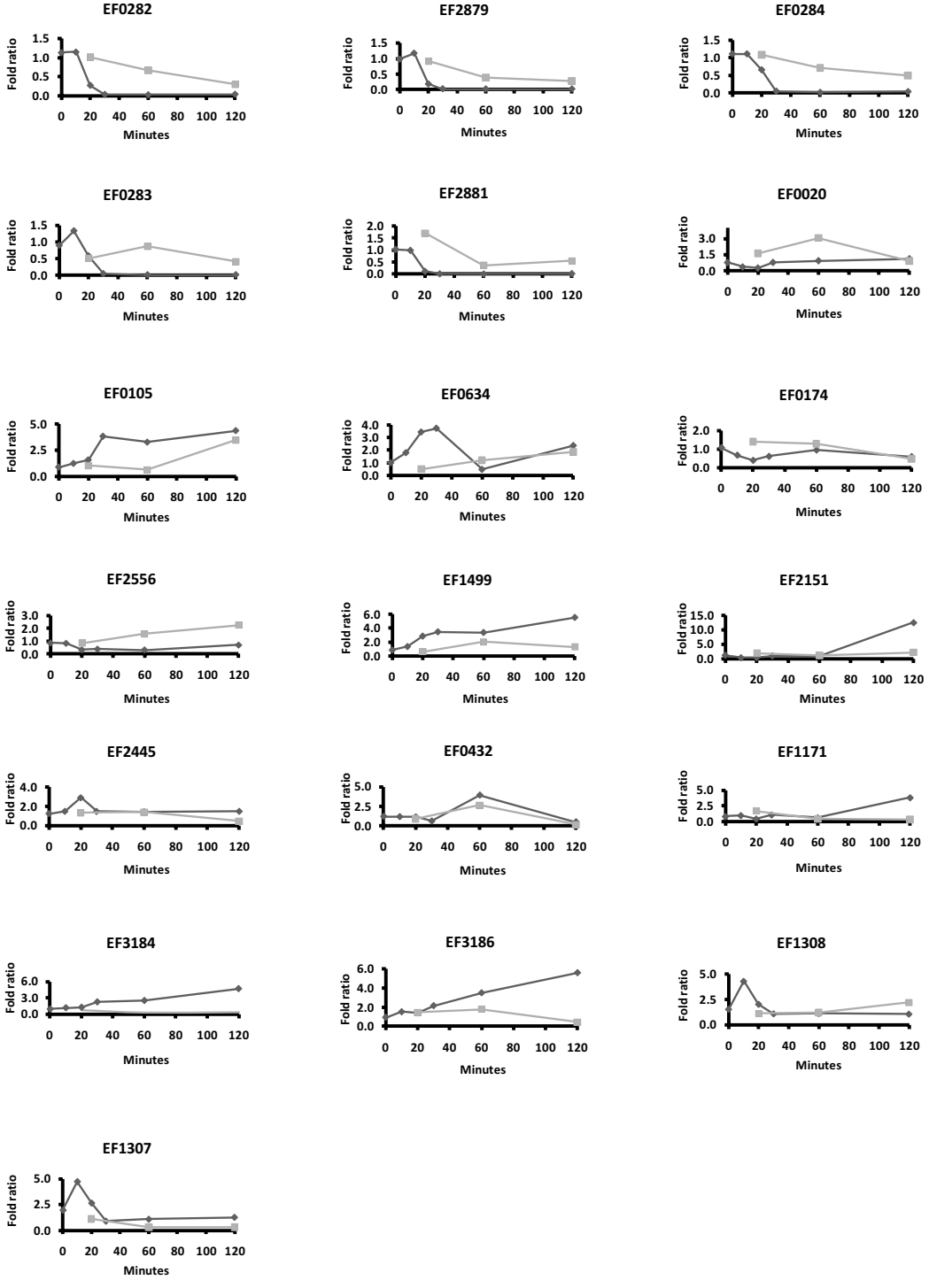



**Figure 4 Expression-profiles of the 19 proteins that were found to be regulated by bile at both the transcriptome** [15]**and the proteome level**. The dark lines show the data from the transcriptome study, while the grey lines show the data from the present proteome study. The y-axis indicates the fold ratio, while the x-axis shows the culturing time in minutes.

between the transcriptome data and data for the intracellular proteome.

Proteins related to general bacterial stress responses primarily occur in the functional classes "Protein fate" and "Other cellular processes". We identified two general stress proteins, the chaperone DnaK (EF1308) and a heat-shock protein GrpE (EF1307). It has previously been shown that a number of stress factors, including bile, have an effect of the expression of DnaK and/or GrpE in *Enterococcus* and other Gram-positive bacteria [16,18,29,30]. Three of the four identified proteins in "Other cellular processes" are related to stress; glutathione reductase (EF3270), Dps (DNA protecting protein under starved conditions; EF3233), and a putative general stress protein of unknown function (EF1744). Whereas EF3233 and EF1744 both are down-regulated, the glutathione reductase, a cellular antioxidant involved in oxidative defence, is significantly up-regulated after 20 minutes, indicating fast response to the deleterious effects of bovine bile. The fourth protein, DivIVA (EF1002) is strongly up-regulated after 120 minutes. DivIVA plays a crucial role in cell division [31]. Recently, it has been shown that DivIVA may also play a role in managing oxidative stress [32]. Other (transcriptome) studies have also shown that expression of general stress proteins is influenced by bile acids [12,13].

The five proteins involved in the fatty acid biosynthesis affected by bile correspond to 14% of the total number of proteins in the functional class of "Fatty acid and phospholipid metabolism", meaning that this class is clearly overrepresented among the regulated proteins (Fig. 2). These five proteins, as well as nine of the ten additional affected genes within this class identified in the previous transcriptome study [15], were all down-regulated upon exposure to bile and the down-regulation increased during the time course experiment (Table 1). All five proteins identified here are involved in fatty acid biosynthesis (Fig. 3), and our data thus clearly show that this process is down-regulated in response to bile. Previous studies have suggested that bile changes the membrane surface as a response to the actions of bile [7,33], but so far, there is limited proteomic data that has supported this notion. Proteomic studies of bile stress in *B. animalis* [28] and *P. freudenreihii* [22] each led to the identification of only one (down-regulated) protein in the "Fatty acid and phospholipid metabolism" class. Interestingly, a recent transcriptome study showed that expression of genes involved in fatty acid and phospholipid biosynthesis was up-regulated in *E. faecalis* V583 grown on blood [34]. This indicates that the down-regulation of fatty acid biosynthesis is a bile-specific response and that the bacteria respond differently when grown in blood.

Eight regulated proteins play a role in energy metabolism, representing 4% of this functional class (Fig. 2). The proteins are involved in different pathways, including glycolysis (EF 0195, EF1046, EF1961), pentose phosphate pathway (EF0174), pyruvate metabolism (EF1613), oxidative phosphorylation (EF2556, EF1499), and the electron transport chain (EF1405). Proteins identified in the first two pathways are members of the carbohydrate metabolism, and most of these proteins were down-regulated in contrast to what has been observed in other studies regarding bile stress [23,27]. Interestingly, a V-type ATPase (EF1499), which couples hydrolysis of ATP to the translocation of protons across bacterial membranes, was up-regulated. Maintenance of the proton motive force is an important factor during bile stress, because it contributes to sustain cellular homeostasis [6]. Pyruvate kinase (EF1046), which converts phosphoenolpuruvate to pyruvate, and formate acetyltransferase (EF1613), which is responsible for the transfer of an acetyl group from acetyl-CoA to formate, yielding CoA and pyruvate, were up-regulated. Pyruvate is a metabolic key molecule that can be used in a number of different reactions to increase the ATP levels. The combined up-regulation of formate acetyltransferase and down-regulation of fatty acid biosynthesis genes indicate that bile stress changes the flux of acetyl-CoA from fatty acid synthesis towards generation of ATP.

Three transporter and binding proteins were identified (EF0020, EF0709 and EF0912). The IIAB component of a mannose-specific phosphotransferase system (PTS; EF0020) was strongly up-regulated. Interestingly, there are studies showing that this type of transporter is important for the survival of saprophytic and pathogenic bacteria on the mucosal surfaces of animals [35]. Thus, it is conceivable that these proteins are also important for adapting to bile-rich environments. Both the phosphocarrier protein HPr (EF0709) and the peptide ABC transporter (EF0912) were down regulated. The phosphocarrier protein HPr is a component of the phosphoenolpyruvate-dependendent PTS, which is a major carbohydrate transport system in bacteria. The mechanism involves the transfer of a phosphoryl group from phosphoenolpyruvate which are formed by the actions of enolase (EF1961) in the glycolysis. The fact that both enolase (EF1961), and phophocarrier protein HPr (EF0709) are down- regulated, correlates with the down-regulation of other proteins involved in carbohydrate metabolism (see above). In bifidobacteria, bile stress resulted in the up-regulation of some sugar transport proteins [23,28]. This may indicate that *E. faecalis* V583 and bifidobacteria use different strategies in how to handle bile.

Four of the identified proteins have unknown functions, and another ten are hypothetical. Most of these fourteen "unknown" proteins were down-regulated by bile. Only two of these proteins (EF3184, EF3186) were also found to be down-regulated in the transcriptome

study [15]. A Blast search with the hypothetical proteins did not reveal significant sequence similarity with hypothetical proteins found in studies of bile responses in other Gram positive bacteria. It must be emphasized that several of these hypothetical proteins are strongly regulated, indicating that they may play important roles in the bile stress response.

Two of the proteins that were classified as proteins with unknown function according to JCVI, are a DNA binding response regulator (EF1193) that is part a two-component regulatory system (EF1194 is a kinase), and an oxidoreductase (EF1138). Two-component regulatory systems are important for responding to environmental changes, whereas the oxidoreductase could play a role in bile modification [13].

Bile salt hydrolases (BSH) catalyze the deconjugation of bile salts, which may be a detoxification mechanism [36]. Experimental data suggests that BSHs indeed play a role in bile tolerance in Gram-positive bacteria [37-39]. The genome of *E. faecalis* contains two genes that putatively encode for BSHs. However, in response to bile no effect on the BSH expression was observed, neither at the transcriptome [15] nor at the proteome level, which is in agreement with other studies on bile-induced gene regulation in Gram-positive bacteria [13,27]. Several gene-by-gene studies indicate that at least part of the known (putative) *bsh*-genes play minor roles in the bile response of lactobacilli [40-42]. All in all, the role of BSHs in bile responses remains somewhat elusive.

## Conclusions

The present study showed that of approximately 500 observed proteins, 53 proteins were significantly regulated in response to bile which provides several leads for further analyses of how *E. faecalis* responds to bile and, perhaps, to stress in general. The clear bile-induced regulation of fatty acid biosynthesis shown at both the proteome and transcriptome level in *E. faecalis* V583 has not been shown previously in Gram-positive bacteria, including enterococci, bifidobacteria and lactobacilli. Proteins involved in fatty acid metabolism were overrepresented among the regulated proteins. In addition, several hypothetical proteins also stand out as targets for further work, including EF1967 and EF2104, which are both abundant and clearly regulated. Such further work is currently in progress in our laboratories.

## Additional material

**Additional file 1 Figure S1**. Silver stained 2D-electrophoresis gels of the intracellular proteome of *E. faecalis* V583 grown in liquid BHI with and without 1% bovine bile. The gels show protein extracts from cells harvested 20, 60 or 120 minutes after the addition of bile. The numbered spots indicate proteins that were identified as being regulated in response to bile stress, using statistical methods and cut-off values described in the main manuscript.

**Competing interests**
The authors declare that they have no competing interests.

**Authors' contributions**
IFN, VGHE and GM developed the initial concept for this study. LAB, EMF, IFN, VGHE and GM participated in experimental design. LAB carried out the 2D electrophoresis and did the mass spectrometric analyses. EVK and HS participated in the set-up and standardization of the 2D-electrophoresis. LAB and EMF did the statistical analyses and the initial interpretation of the results. LAB, VGHE and GM drafted the paper, implementing contributions from all other authors. All authors have read, corrected and approved the final manuscript.

**Acknowledgements**
This work was supported by grants 183627 and 183637 from the Norwegian Research Council. Thanks to Morten Skaugen, and Øyvind Busk for technical help during the MS-analysis. We also thank Bernd Thiede at the Biotechnology Centre of Oslo for LC-LTQ Orbitrap-MS analyses.

**Author Details**
[1]Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Chr. M. Falsensvei 1, N-1432 Ås, Norway and [2]Nofima Mat AS, Norwegian Institute of Food, Fisheries and Aquaculture Research, N-1430 Ås, Norway

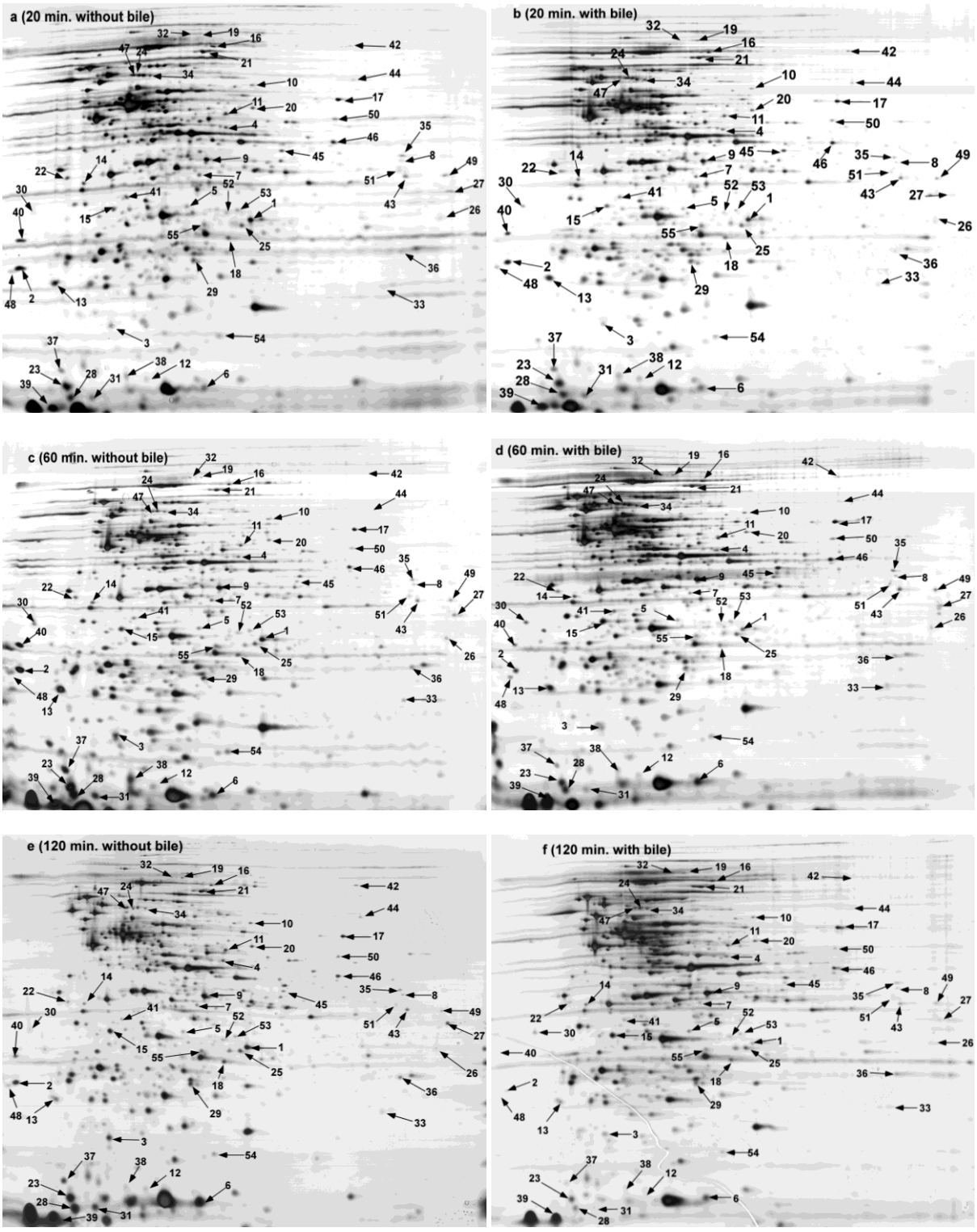
Received: 4 March 2010 Accepted: 25 June 2010
Published: 25 June 2010

**References**

1. Paulsen IT, Banerjei L, Myers GS, Nelson KE, Seshadri R, Read TD, Fouts DE, Eisen JA, Gill SR, Heidelberg JF, *et al.*: **Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*.** *Science* 2003, **299:**2071-2074.
2. Tacconelli E, Cataldo MA: **Vancomycin-resistant enterococci (VRE): transmission and control.** *Int J Antimicrob Agents* 2008, **31:**99-106.
3. Cetinkaya Y, Falk P, Mayhall CG: **Vancomycin-resistant enterococci.** *Clin Microbiol Rev* 2000, **13:**686-707.
4. Fisher K, Phillips C: **The ecology epidemiology and virulence of Enterococcus.** *Microbiology* 2009, **155:**1749-1757.
5. Rince A, Le Breton Y, Verneuil N, Giard JC, Hartke A, Auffray Y: **Physiological and molecular aspects of bile salt response in *Enterococcus faecalis*.** *Int J Food Microbiol* 2003, **88:**207-213.
6. Begley M, Gahan CG, Hill C: **The interaction between bacteria and bile.** *FEMS Microbiol Rev* 2005, **29:**625-651.
7. Taranto MP, Fernandez Murga ML, Lorca G, de Valdez GF: **Bile salts and cholesterol induce changes in the lipid cell membrane of *Lactobacillus reuteri*.** *J Appl Microbiol* 2003, **95:**86-91.
8. Ruiz L, Sanchez B, Ruas-Madiedo P, de Los Reyes-Gavilan CG, Margolles A: **Cell envelope changes in *Bifidobacterium animalis* ssp. *lactis* as a response to bile.** *FEMS Microbiol Lett* 2007, **274:**316-322.
9. Bron PA, Marco M, Hoffer SM, Van Mullekom E, de Vos WM, Kleerebezem M: **Genetic characterization of the bile salt response in *Lactobacillus plantarum* and analysis of responsive promoters in vitro and in situ in the gastrointestinal tract.** *J Bacteriol* 2004, **186:**7829-7835.
10. Kristoffersen SM, Ravnum S, Tourasse NJ, Okstad OA, Kolsto AB, Davies W: **Low concentrations of bile salts induce stress responses and reduce motility in Bacillus cereus ATCC 14579.** *J Bacteriol* 2007, **189:**5302-5313.
11. Gueimonde M, Garrigues C, van Sinderen D, de los Reyes-Gavilan CG, Margolles A: **Bile-inducible efflux transporter from *Bifidobacterium longum* NCC2705, conferring bile resistance.** *Appl Environ Microbiol* 2009, **75:**3153-3160.
12. Whitehead K, Versalovic J, Roos S, Britton RA: **Genomic and genetic characterization of the bile stress response of probiotic *Lactobacillus reuteri* ATCC 55730.** *Appl Environ Microbiol* 2008, **74:**1812-1819.
13. Pfeiler EA, Azcarate-Peril MA, Klaenhammer TR: **Characterization of a novel bile-inducible operon encoding a two-component regulatory system in *Lactobacillus acidophilus*.** *J Bacteriol* 2007, **189:**4624-4634.
14. Bron PA, Molenaar D, de Vos WM, Kleerebezem M: **DNA micro-array-based identification of bile-responsive genes in *Lactobacillus plantarum*.** *J Appl Microbiol* 2006, **100:**728-738.
15. Solheim M, Aakra A, Vebo H, Snipen L, Nes IF: **Transcriptional responses of *Enterococcus faecalis V583* to bovine bile and sodium dodecyl sulfate.** *Appl Environ Microbiol* 2007, **73:**5767-5774.

16. Flahaut S, Frere J, Boutibonnes P, Auffray Y: **Comparison of the bile salts and sodium dodecyl sulfate stress responses in *Enterococcus faecalis*.** *Appl Environ Microbiol* 1996, **62:**2416-2420.
17. Flahaut S, Hartke A, Giard JC, Benachour A, Boutibonnes P, Auffray Y: **Relationship between stress response towards bile salts acid and heat treatment in *Enterococcus faecalis*.** *FEMS Microbiol Lett* 1996, **138:**49-54.
18. Giard JC, Laplace JM, Rince A, Pichereau V, Benachour A, Leboeuf C, Flahaut S, Auffray Y, Hartke A: **The stress proteome of *Enterococcus faecalis*.** *Electrophoresis* 2001, **22:**2947-2954.
19. Sahm DF, Kissinger J, Gilmore MS, Murray PR, Mulder R, Solliday J, Clarke B: **In vitro susceptibility studies of vancomycin-resistant Enterococcus faecalis.** *Antimicrob Agents Chemother* 1989, **33:**1588-1591.
20. Shevchenko A, Wilm M, Vorm O, Mann M: **Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels.** *Anal Chem* 1996, **68:**850-858.
21. Berth M, Moser FM, Kolbe M, Bernhardt J: **The state of the art in the analysis of two-dimensional gel electrophoresis images.** *Appl Microbiol Biotechnol* 2007, **76:**1223-1243.
22. Leverrier P, Dimova D, Pichereau V, Auffray Y, Boyaval P, Jan G: **Susceptibility and Adaptive Response to Bile Salts in Propionibacterium freudenreichii: Physiological and Proteomic Analysis.** *Appl Environ Microbiol* 2003, **69:**3809-3818.
23. Sanchez B, Champomier-Verges MC, Anglade P, Baraige F, de Los Reyes-Gavilan CG, Margolles A, Zagorec M: **Proteomic analysis of global changes in protein expression during bile salt exposure of *Bifidobacterium longum* NCIMB 8809.** *J Bacteriol* 2005, **187:**5799-5808.
24. Moen B, Oust A, Langsrud O, Dorrell N, Marsden GL, Hinds J, Kohler A, Wren BW, Rudi K: **Explorative multifactor approach for investigating global survival mechanisms of *Campylobacter jejuni* under environmental conditions.** *Appl Environ Microbiol* 2005, **71:**2086-2094.
25. Langsrud Ø: **Rotation tests.** *Statistics and computing* 2005, **15:**53-60.
26. Rappsilber J, Ishihama Y, Mann M: **Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics.** *Analyt Chem* 2003, **75:**663-670.
27. Lee K, Lee HG, Choi YJ: **Proteomic analysis of the effect of bile salts on the intestinal and probiotic bacterium *Lactobacillus reuteri*.** *J Biotechnol* 2008, **137:**14-19.
28. Sanchez B, Champomier-Verges MC, Stuer-Lauridsen B, Ruas-Madiedo P, Anglade P, Baraige F, de los Reyes-Gavilan CG, Johansen E, Zagorec M, Margolles A: **Adaptation and response of *Bifidobacterium animalis* subsp. *lactis* to bile: a proteomic and physiological approach.** *Appl Environ Microbiol* 2007, **73:**6757-6767.
29. Leverrier P, Vissers JP, Rouault A, Boyaval P, Jan G: **Mass spectrometry proteomic analysis of stress adaptation reveals both common and distinct response pathways in *Propionibacterium freudenreichii*.** *Arch Microbiol* 2004, **181:**215-230.
30. Schmidt G, Zink R: **Basic features of the stress response in three species of bifidobacteria: *B. longum, B. adolescentis, and B. breve*.** *Int J Food Microbiol* 2000, **55:**41-45.
31. Ramirez-Arcos S, Liao M, Marthaler S, Rigden M, Dillon JA: ***Enterococcus faecalis* divIVA: an essential gene involved in cell division, cell growth and chromosome segregation.** *Microbiology* 2005, **151:**1381-1393.
32. Mukherjee P, Sureka K, Datta P, Hossain T, Barik S, Das KP, Kundu M, Basu J: **Novel role of Wag31 in protection of mycobacteria under oxidative stress.** *Mol Microbiol* 2009, **73:**103-119.
33. Ruiz L, Coute Y, Sanchez B, de los Reyes-Gavilan CG, Sanchez JC, Margolles A: **The cell-envelope proteome of Bifidobacterium longum in an in vitro bile environment.** *Microbiology* 2009, **155:**957-967.
34. Vebø HC, Snipen L, Nes IF, Brede DA: **The transcriptome of the nosocomial pathogen *Enterococcus faecalis* V583 reveals adaptive responses to growth in blood.** *PLoS One* 2009, **4:**e7660.
35. Zuniga M, Comas I, Linaje R, Monedero V, Yebra MJ, Esteban CD, Deutscher J, Perez-Martinez G, Gonzalez-Candelas F: **Horizontal gene transfer in the molecular evolution of mannose PTS transporters.** *Mol Biol Evol* 2005, **22:**1673-1685.
36. Begley M, Hill C, Gahan CG: **Bile salt hydrolase activity in probiotics.** *Appl Environ Microbiol* 2006, **72:**1729-1738.
37. Jones BV, Begley M, Hill C, Gahan CG, Marchesi JR: **Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome.** *Proc Natl Acad Sci USA* 2008, **105:**13580-13585.
38. Elkins CA, Moser SA, Savage DC: **Genes encoding bile salt hydrolases and conjugated bile salt transporters in *Lactobacillus johnsonii* 100-100 and other *Lactobacillus* species.** *Microbiology* 2001, **147:**3403-3412.
39. Tanaka H, Doesburg K, Iwasaki T, Mierau I: **Screening of lactic acid bacteria for bile salt hydrolase activity.** *J Dairy Sci* 1999, **82:**2530-2535.
40. Lambert JM, Bongers RS, de Vos WM, Kleerebezem M: **Functional analysis of four bile salt hydrolase and penicillin acylase family members in *Lactobacillus plantarum* WCFS1.** *Appl Environ Microbiol* 2008, **74:**4719-4726.
41. McAuliffe O, Cano RJ, Klaenhammer TR: **Genetic analysis of two bile salt hydrolase activities in *Lactobacillus acidophilus* NCFM.** *Appl Environ Microbiol* 2005, **71:**4925-4929.
42. Fang F, Li Y, Bumann M, Raftis EJ, Casey PG, Cooney JC, Walsh MA, O'Toole PW: **Allelic variation of bile salt hydrolase genes in *Lactobacillus salivarius* does not determine bile resistance levels.** *J Bacteriol* 2009, **191:**5743-5757.

doi: 10.1186/1477-5956-8-37
**Cite this article as:** Bøhle *et al.*, Identification of proteins related to the stress response in *Enterococcus faecalis* V583 caused by bovine bile *Proteome Science* 2010, **8**:37




**Additional file 1. Figure S1.** Silver stained 2D-electrophoresis gels of the intracellular proteome of *E. faecalis* V583 grown in liquid BHI with and without 1% bovine bile. The gels show protein extracts from cells harvested 20, 60 or 120 minutes after the addition of bile.The numbered spots indicate proteins that were identified as being regulated in response to bile stress, using statistical methods and cut-off values described in the main manuscript.

# PAPER II

BMC Genomics

RESEARCH ARTICLE Open Access

# Identification of surface proteins in *Enterococcus faecalis* V583

Liv Anette Bøhle[1], Tahira Riaz[1], Wolfgang Egge-Jacobsen[2], Morten Skaugen[1], Øyvind L Busk[1], Vincent GH Eijsink[1], Geir Mathiesen[1*]

## Abstract

**Background:** Surface proteins are a key to a deeper understanding of the behaviour of Gram-positive bacteria interacting with the human gastro-intestinal tract. Such proteins contribute to cell wall synthesis and maintenance and are important for interactions between the bacterial cell and the human host. Since they are exposed and may play roles in pathogenicity, surface proteins are interesting targets for drug design.

**Results:** Using methods based on proteolytic "shaving" of bacterial cells and subsequent mass spectrometry-based protein identification, we have identified surface-located proteins in *Enterococcus faecalis* V583. In total 69 unique proteins were identified, few of which have been identified and characterized previously. 33 of these proteins are predicted to be cytoplasmic, whereas the other 36 are predicted to have surface locations (31) or to be secreted (5). Lipid-anchored proteins were the most dominant among the identified surface proteins. The seemingly most abundant surface proteins included a membrane protein with a potentially shedded extracellular sulfatase domain that could act on the sulfate groups in mucin and a lipid-anchored fumarate reductase that could contribute to generation of reactive oxygen species.

**Conclusions:** The present proteome analysis gives an experimental impression of the protein landscape on the cell surface of the pathogenic bacterium *E. faecalis*. The 36 identified secreted (5) and surface (31) proteins included several proteins involved in cell wall synthesis, pheromone-regulated processes, and transport of solutes, as well as proteins with unknown function. These proteins stand out as interesting targets for further investigation of the interaction between *E. faecalis* and its environment.

## Background

Enterococci are versatile Gram-positive bacteria that can survive under harsh conditions. Most enterococci are non-virulent and commonly found in fermented food and in the gastrointestinal (GI) tract of humans and animals. Other strains are opportunistic pathogens that contribute in a large number of nosocomial infections worldwide [1]. The mechanism underlying the switch from a harmless microbe into a life-threatening pathogen entering the host bloodstream is not well known. It is believed that the bacteria normally are well controlled in the GI tract of healthy individuals, whereas a weakened host immune system and/or development of bacterial traits to occupy new niches may lead to translocation to the bloodstream [2]. The past decade has shown a dramatic increase in antibiotic resistance of *Enterococcus* species, creating an increased need for developing new ways to combat these bacteria. To achieve this, in-depth insight in the physiology, virulence and epidemiology of enterococci is required.

*Enterococcus faecalis* is one of the most frequent *Enterococcus* species in the GI tract [3,4] and accounts for at least 60% of the bacteraemia caused by *Enterococcus* species [1]. The genome sequence of three *E. faecalis* strains (V583; [5], OG1RF [6], Symbioflor 1 [7]) have been completed, and several genome projects are ongoing. In the genome sequence of *E. faecalis* V583, a vancomycin resistant clinical isolate, over a quarter of the genome consists of mobile or foreign DNA, including pathogenicity islands. The abundance of mobile elements in *E. faecalis* probably contributes to accumulation of virulence and drug determinants. Several studies have revealed proteins

* Correspondence: geir.mathiesen@umb.no
[1]Department of Chemistry, Biotechnology, and Food Science, The Norwegian University of Life Sciences, 1432 Ås, Norway
Full list of author information is available at the end of the article

that contribute to the virulence of *E. faecalis* [8-10], but it has so far not been possible to link virulence to one or very few key gene products. Since virulence depends on the ability to colonize the GI tract and interact with host cells and proteins in the GI tract, secreted proteins and proteins located on the cell surface are thought to be important. One well-studied secreted virulence factor is cytolysin, which is toxic or lytic to bacterial and human cells [9,11]. Several adhesion factors facilitating binding to mucosal and epithelial surfaces have been reported [2]. In addition to involvement in adhesion, surface proteins may affect virulence in other ways, for example by involvement in cell-cell signalling [9], interactions with the host immune system, sensing environmental factors, or protection from environmental factors.

To understand the success of bacterial pathogens and their adaption to the GI tract it is important to get an impression of the repertoire of surface associated proteins. According to the LocateP database [12], which contains genome-wide predictions for the subcellular locations of bacterial proteins, 306 proteins in *E. faecalis* are predicted to be covalently anchored to cell surface, primarily via N-terminal or lipid anchors. Another 67 are predicted to be secreted or non-covalently attached to the surface. There is only limited experimental data supporting these predicted locations [13]. Furthermore, despite their expected importance for bacterial behaviour and impact, the function of most of the predicted surface and secreted proteins remains unknown.

In the past decade, the extracellular proteomes of several Gram-positive bacteria have been analyzed using proteomics approaches. Many of these studies employed some kind of protein extraction methods from culture supernatants and/or cell wall fractions followed by two-dimensional electrophoresis and mass spectrometry-based protein identification (e.g. [14-18]). Recently, more direct and rapid methods for the "in situ" identification of surface proteins have been developed which are based on "shaving" the surface of intact bacteria with proteolytic enzymes, followed by identification of the released peptides by liquid chromatography (LC) and tandem mass spectrometry (MS/MS) [19]. An advantage of this approach is that the proteolytic enzymes will only have access to proteins that are exposed on the surface of the bacterial cell, which could limit contamination with intracellular proteins. This approach should in principle be biased towards proteins that are of particular importance for bacterial interactions with the environment. Indeed, the "shaving" approach has been applied successfully in the search for new bacterial epitopes. In a landmark study, Rodriguez-Ortega et al [19] identified in total 72 proteins in *Streptococcus pyogenes* M1_SF370 by shaving the bacterial surface with trypsin or proteinase K. The identified proteins included known protective antigens and also revealed new promising candidate antigens for vaccine development.

In the present study, we have applied the "shaving" approach to identify the surface proteome of *E. faecalis* V583. Cells were treated with free trypsin or trypsin anchored to agarose beads to shave off and digest surface-exposed proteins. Using a combination of experiments, 69 surface-located proteins were identified, including proteins assumed to be involved in pathogenicity and several proteins with unknown function. We also identified proteins that are thought to be cytoplasmic, but which tend to be found at bacterial surfaces too. We discuss the putative roles and relevance of several of the identified proteins and we compare the various approaches. The results provide a basis for the identification and further study of novel proteins putatively involved in pathogenicity and adaptability of *E. faecalis* V583.

## Methods

### Culture conditions and surface shaving

Overnight cultures of *E. feacalis* V583 [5] were diluted in fresh prewarmed brain heart infusion (BHI) medium (Oxoid Ltd., Hampshire, England) to an $OD_{600}$~0.2 and incubated at 37°C without agitation. The bacteria were harvested on the transition between late exponential and stationary phase ($OD_{600}$~1.7) by centrifugation (3000 × g, 10 min, 4°C). The cell pellet from 100 ml of culture was washed three times with 10 ml PBS by centrifugation (3000 × g, 10 min, 4°C) and subsequently resuspended in PBS containing 40% sucrose. Three different shaving reactions were set-up, all containing 5 mM DTT and all with the same final concentration of cells: (1) addition of 20 μg trypsin (Promega, Mannheim, Germany), (2) addition of trypsin-agarose (100 units; Invitrogen, Karlsruhe, Germany), or (3) no addition of trypsin (untreated). The samples were incubated for 1 or 2 hours at 37°C with shaking at 300 rpm. After incubation the cells were pelleted by centrifugation (3000 × g, 10 min) and the supernatants were collected for further protein digestion with 1 μg freshly added trypsin over night (16-18 h) at 37°C with agitation at 400 rpm. Cell samples taken before (i.e. after resuspending in PBS) and after the different enzymatic treatments were used to test cell viability by plating appropriate dilutions on BHI agar plates and counting of colony forming units (CFU). The overnight trypsin digestion of the supernatants was stopped by adding formic acid to a final concentration of 0.1% (v/v). Prior to nanoLC-MS/MS analysis, peptides were concentrated and purified in two steps using $C_{18}$ Dynabeads (Invitrogen) in the first step and C18 StageTips [20] in the second step. For each treatment, samples from four biological replicates were analysed.

### Nanoflow on-line liquid chromatographic MS analysis of trypticpeptides

Reverse phase (C18) nano online liquid chromatographic MS/MS analyses of tryptic peptides were performed using a HPLC system consisting of two Agilent 1200 HPLC binary pumps (nano and capillary) with corresponding autosampler, column heater and integrated switching valve. This LC system was coupled via a nanoelectrospray ion source to a LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). For the analyses, the peptide solution was injected onto the 5 × 0.3-mm extraction column filled with Zorbax 300 SB-C18 of 5-μm particle size (Agilent, Waldbronn, Germany). Samples were washed with a mobile phase consisting of 97% 0.1% formic acid & 3% acetonitrile. The flow rate of 4 μl/min provided by the capillary pump. After 7 min, the switching valve of the integrated switching valve was activated, and the peptides were eluted in the back-flush mode from the extraction column onto a 150 × 0.075-mm $C_{18}$, 3-μm resin, column (GlycproSIL C18-80Å, Glycpromass, Stove, Germany). The mobile phase consisted of acetonitrile and MS grade water, both containing 0.1% formic acid. Chromatographic separation was achieved using a binary gradient from 5 to 55% of acetonitrile in 120 min. The flow rate of 0.2 μl $min^{-1}$ was provided by the nanoflow pump.

Mass spectra were acquired in the positive ion mode applying a data-dependent automatic switch between survey scan and tandem mass spectra (MS/MS) acquisition. Peptide samples were analyzed by collision induced dissociation (CID) in the LTQ ion trap by acquiring one Orbitrap survey scan in the mass range of *m/z* 380-2000 followed by CID of the six most intense ions in the ion trap. The target value in the LTQ-Orbitrap was 1,000,000 for survey scan at a resolution of 60,000 at *m/z* 400 using lock masses for recalibration to improve the mass accuracy of precursor ions. Fragmentation was performed with a target value of 5,000 ions. The ion selection threshold was 500 counts. Selected sequenced ions were dynamically excluded for 180 s.

### MS data analysis

Mass spectrometric data were first analyzed by generating msf files from raw MS and MS/MS spectra using the Proteome Discoverer 1.0 software (Thermo Fisher Scientific) and the database searches were then performed with an in house maintained *E. faecalis* V583 protein sequence database, using the SEQUEST search engine. The following criteria were applied; database decoy, true; Enzyme name, trypsin (full); Missed cleavage sites, 2; Precursor mass tolerance, 10 ppm; fragment mass tolerance: 0.6 Da; dynamic modifications: N-term acetyl (any N-terminus), oxidation (M), carboxymethyl (C), deamidated (N, Q). Proteins were considered as significant hits if the following conditions were met: XCorr higher than 2.0; false discovery rate less than 5%; identified by at least two different peptides; identified in at least two of the independent parallels by at least one peptide in each.

### SDS-PAGE analysis

To visualize proteins or protein fragments that were resistant to trypsin, 20 μl of the supernatant from the over-night trypsination were applied to 10% NuPAGE Novex Bis-Tris gels (Invitrogen). Only samples after two-hour incubation were studied. The gels were stained using SilverSNAP Stain for Mass Spectrometry (Pierce, Rockford, IL) following the manufacturer's procedure. After the silver staining, the gel-lane was sliced into 12 pieces, and destained using the protocol included in SilverSNAP Stain for Mass Spectrometry kit. Each gel piece was then incubated with 0.1 μg trypsin in 25 μl 25 mM ammonium bicarbonate, over night at 37°C and 400 rpm. The trypsin reactions were stopped by adding 0.1% formic acid. The supernatants were transferred to new tubes, and the rest of the peptides were extracted from the gel pieces by incubating with 0.1% (v/v) trifluoroacetic acid in 60% (v/v) acetonitrile, at 37°C, 400 rpm for 10 min. The extracts from three gel-pieces were pooled together (giving four samples from each treatment). The peptides were dried in a speed-vac, and rehydrated in 30 μl 0.1% (v/v) TFA. The peptide samples were desalted using C18 stage tips [20] prior to nanoLC-MS/MS. Proteins were considered as significant hits if the following conditions were met: XCorr higher than 2.0; false discovery rate less than 5%; identified by at least two different peptides; identified in at least one of the three samples from each treatment.

### Bioinformatic analysis of protein localization

Protein sequences used for *in computo* analysis of the localization of the identified proteins were extracted from the LocateP database [12] and analyzed using several bioinformatic tools. Putative N-terminal signal sequences and cleavage sites were predicted using the Signal P 3.0 server [21] and LipoP v 1.0 [22]. The TMHMM Server v. 2.0 [23] was used to predict proteins with multiple transmembrane helices or N-terminal transmembrane anchors. Proteins with features indicating non-classical secretion were predicted using the SecretomeP 2.0 Server [24]. Domain annotations were done using Pfam [25]. After these verifications, the predicted localizations for 62 of the 69 proteins discussed below correspond to those given in the LocateP database (updated March 10, 2010). For seven proteins, we reached a different conclusion than LocateP, as described in results and discussion.

## Results and discussion

### Protein identification

Before carrying out the experiments, we performed extensive tests to find optimal conditions for the trypsin treatment. Most importantly, we checked the effect of incubation time (30 min to 24 hours) on cell viability. Incubation times of 2 hours or less did not lead to significant reductions in the CFU number (Additional file 1), whereas longer incubation times led to decreased viability (data not shown). Based on these observations, incubation times were set to one or two hours. Two hour incubations led to a higher number of identified proteins (Additional file 2). Generally, the longer incubations did not lead to an increase in the fraction of cytoplasmic proteins, confirming the absence of cell lysis during the enzymatic treatment (Additional file 2).

Intact bacterial cells were harvested and treated with either trypsin or trypsin beads (trypsin bound to agarose beads) for one or two hours. Because trypsin beads are less likely to penetrate the cell wall than free trypsin, they are more likely to only act on proteins that protrude from the cell wall. As a control, cells were incubated for one or two hours without adding trypsin. Direct analysis of released tryptic peptides led to the identification of 57 unique proteins (Table 1). Subsequent analysis of solubilized proteins and large protein fragments using SDS-PAGE followed by MS-based identification (Figure 1; see Materials and Methods) led to the identification of another 12 unique proteins (Table 2). The sequences of all identified proteins were analysed using a variety of bioinformatic tools (LocateP, SignalP, Pfam, LipoP, pSORT and TMHMM) to verify or (for EF2860, EF0071, EF0123, EF0164, EF0394, EF0417 & EF_B0004) to adjust the localization given in the LocateP database. The results are incorporated in Table 1 & 2 and are discussed in appropriate sections, below.

Analysis with Signal P 3.0 [21] suggested the presence of a signal peptidase I (SPase I) cleavage site in ten of the identified proteins. Four of these proteins contained a putative C-terminal LPxTG motif, whereas one additional protein (EF2860) is likely to be cell-wall anchored because it contains a putative peptidoglycan binding domain (Pfam PF12229). It should be noted that the sequences deposited in the GenBank database for two of the four LPxTG-containing proteins (EF1033 and EF2713) lack a predicted N-terminal signal sequence. A closer look at the upstream sequences showed that the start codons probably are located 72 and 96 nucleotides upstream of the start codon suggested in the GenBank entries, for EF1033 and EF2713, respectively (Additional file 3). After this N-terminal "extension" SignalP and LipoP detected a putative SPase I cleavage site in both sequences.

Table 3 gives an overview over the predicted localizations of the 69 identified unique proteins and shows that the methods yielded a strong bias towards identifying proteins that are predicted to be covalently anchored to the cell wall or to carry lipid anchors. Several proteins were identified in more than one experiment and an overview is provided in Figure 2. Treatment with free trypsin yielded 58 proteins and treatment with trypsin beads yielded 29 proteins. Analysis of samples from untreated cells yielded 16 proteins. More detailed information concerning the numbers of proteins identified after the various treatments is provided in additional file 2. Details of the proteomic analysis are provided in Additional file 4 and Additional file 5 containing Tables S5 and S6, respectively.

The number of proteins only identified after a "shaving" treatment amounted to 53. Nine of these were only found after treatment with trypsin beads, 38 were only found after treatment with free trypsin and six were found after both treatments (Figure 2). The 15 proteins identified with only trypsin-beads or with both trypsin and trypsin beads are likely to be exposed on the surface of the cell wall, whereas the 38 unique proteins found in the free trypsin samples fraction are probably localized deeper in the cell wall.

Of the 16 proteins identified from untreated cells, 12 were identified in all three treatments (Figure 2). While three of these are predicted to be secreted and one (EF0201) is probably cytoplasmic, the others are predicted to be attached to the bacterial surface through an anchor (five lipo-anchors and one LPxTG anchor) or cell wall binding domain (one, EF2860) or even as integral membrane protein (one, EF1264). The trypsin-independent release of these proteins may be a result of natural shedding, a phenomenon that indeed has been observed previously, in particular for lipoproteins [15,26]. According to a TMHMM topology prediction EF1264, annotated as membrane protein, contains five N-terminal transmembrane helices and a huge extracellular domain with putative sulfatase activity of 523 residues (starting at amino acid 179). EF1264 was identified by many significant peptide hits spread over all treatments. Figure 3 shows that all identified peptides stem from the extracellular domain and that there is a 115 amino acid gap between the predicted integral membrane domain and the first identified tryptic fragment. Perhaps the extracellular domain is shedded after natural cleavage of EF1264. It is conceivable that such (apparently rather abundant) shedding is a physiologically relevant phenomenon since the sulfatase may remove sulphate from mucin, which would allow more easy degradation of mucin by glycosidases [27] and perhaps also could facilitate bacterial adhesion.

**Table 1 Proteins identified by LC-MS/MS analysis of tryptic fragments obtained after different treatments of intact *E. faecalis* V583**

| Gene | Pfam[a] | Gene product[b] | Predicted localization[c] | Peptide hits[d] | | | | | | Total cover-age % |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Un-treated 1 hour | Un-treated 2 hours | Trypsin 1 hour | Trypsin 2 hours | Beads 1 hour | Beads 2 hours | |
| EF_B0004[e] | Bacterial extracellular solute-binding proteins, family 5 | TraC protein | Cell wall | | | 2 | 2 | (1) | (1) | 13.0 |
| EF0071 | Contains trehalase domain | Putative lipoprotein | Lipid anchor SP-II, WS-CF | | | (3) | 2 | | | 10.0 |
| EF0123 | Contains nine Clostridial hydro-phobic W domain | Hypothetical protein | Secreted SP-I, AYA-LE | 5 | 2 | 9 | 7 | (1) | 4 | 23.2 |
| EF0164 | | Putative lipoprotein | Lipid anchor SP-II, FTS-CG | 2 | 3 | 2 | 3 | | 2 | 32.3 |
| EF0176[f] | Basic membrane protein | Hypothetical protein | Lipid anchor SP-II, LAA-CG | | | 3 | 5 | | 2 | 25.7 |
| EF0177[f] | Basic membrane protein | Hypothetical protein | Lipid anchor SP-II, LAA-CG | | | 5 | 7 | 2 | | 32.1 |
| EF0195[g] | | Phospho-glycerate mutase 1 | Cytoplasmic | | | 2 | (1) | | | 10.5 |
| EF0199[g] | | 30S ribosomal protein S7 | Cytoplasmic | | | | 2 | | | 30.1 |
| EF0200[g] | | Elongation factor G | Cytoplasmic | (1) | | 7 | 7 | | | 22.4 |
| EF0201[fg] | | Elongation factor Tu | Cytoplasmic | (1) | 2 | 11 | 8 | 5 | 4 | 47.1 |
| EF0205 | | 30S ribosomal protein S10 | Cytoplasmic | | | 2 | | | (1) | 29.4 |
| EF0206[g] | | 50S ribosomal protein L3 | Cytoplasmic | | | 2 | 5 | | | 23.4 |
| EF0207[g] | | 50S ribosomal protein L4 | Cytoplasmic | | | (1) | 2 | | | 15.9 |
| EF0211 | | 50S ribosomal protein L22 | Cytoplasmic | (1) | 2 | | 2 | (1) | (1) | 17.4 |
| EF0218[g] | | 50S ribosomal protein L5 | Cytoplasmic | | | (1) | 3 | | | 21.2 |
| EF0221[g] | | 50S ribosomal protein L6 | Cytoplasmic | | | (1) | 3 | | | 27.0 |
| EF0223 | | 50S ribosomal protein L18 | Cytoplasmic | | | (1) | 2 | | | 39.0 |
| EF0226 | | 50S ribosomal protein L15 | Cytoplasmic | (1) | (2) | 2 | (1) | (1) | (1) | 21.2 |
| EF0228 | | Adenylate kinase | Cytoplasmic | | | 2 | 2 | | | 19.4 |
| EF0234[g] | | 50S ribosomal protein L17 | Cytoplasmic | | | 2 | (1) | | | 26.0 |
| EF0304 | - | Putative lipoprotein | Lipid anchor SP-II, LSA-CS | | | (1) | 2 | | 2 | 20.5 |
| EF0394 | Cysteine-rich secretory protein family | Secreted antigen, putative | Secreted SP-I, ALA-DN | 2 | | | | 3 | (1) | 17.8 |
| EF0417 | - | Hypothetical protein | Secreted SP-I, VNA-LN | 3 | 2 | 4 | 4 | | 3 | 11.2 |
| EF0502 | - | Hypothetical protein | Multiple transmembrane proteins | | | | | | 2 | 7.9 |
| EF0633 | | tyrosyl-tRNA synthetase | Cytoplasmic | | | | 2 | | | 6.9 |
| EF0685[f] | | Rotamase family protein | Lipid anchor SP-II, LAA-CS | | | | 2 | | | 10.8 |
| EF0737 | | Amidase | N-terminal anchor | | | | 2 | | | 6.8 |

**Table 1 Proteins identified by LC-MS/MS analysis of tryptic fragments obtained after different treatments of intact *E. faecalis* V583** *(Continued)*

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| EF0907[f] | - | Peptide ABC transporter peptide-binding protein | Lipid anchor SP-II, LAA-CG | | | | 2 | | | 16.4 |
| EF0916 | | 50S ribosomal protein L20 | Cytoplasmic | (1) | (1) | 2 | | | | 27.7 |
| EF0970[e] | | 50S ribosomal protein L27 | Cytoplasmic | | | 2 | 2 | | | 34.7 |
| EF0991 | | Penicillin-binding protein C | N-terminal anchor | | | (1) | 2 | | | 11.2 |
| EF1033 | | Lipoamidase | Cell wall anchored LPxTG SP-I, AQE-SI | | | 2 | 2 | | | 4.7 |
| EF1046[g] | | Pyruvate kinase | Cytoplasmic | (2) | | (2) | 2 | | | 11.6 |
| EF1167[g] | | Fructose-bisphosphate aldolase | Cytoplasmic | (1) | (2) | 2 | 4 | 3 | (1) | 30.8 |
| EF1264[f] | | Sulfatase domain-containing protein | Multiple trans-membrane proteins | 2 | | 10 | 11 | (1) | 5 | 29.8 |
| EF1308[fg] | | Dnak protein | Cytoplasmic | (1) | | 2 | 3 | | | 23.7 |
| EF1319 | - | Hypothetical protein | N-terminal anchor | | | (1) | 3 | | | 23.3 |
| EF1379 | | Alanyl-tRNA synthetase | Cytoplasmic | | | | | 2 | | 3.3 |
| EF1420 | - | Hypothetical protein | Lipid anchor SP-II, MTA-CS | | | 2 | (1) | | | 15.6 |
| EF1523[e] | - | Hypothetical protein | Cytoplasmic | | | (1) | 3 | | | 6.4 |
| EF1613[fg] | | Formate acetyltransferase | Cytoplasmic | | | 3 | (2) | | | 13.4 |
| EF1818[f] | | Coccolysin | Secreted SP-I, VAA-EE | 8 | 3 | 8 | 5 | (1) | 2 | 45.1 |
| EF1898 | | 50S ribosomal protein L19 | Cytoplasmic | | | (1) | 2 | | | 30.4 |
| EF1961[fg] | | Enolase | Cytoplasmic | | | 2 | 4 | (1) | | 25.7 |
| EF1964[fg] | | Glycer-aldehyde-3-phosphate dehydrogenase (GAPDH) | Cytoplasmic | | | 3 | 8 | | | 36.0 |
| EF2144 | - | Putative lipoprotein | Lipid anchor SP-II LTA-CS | (1) | 2 | (1) | 2 | | | 22.4 |
| EF2174 | Glycosyl hydrolases family 25 | Hypothetical protein | Secreted SP-I, ASG-EE | | | | 3 | | | 4.5 |
| EF2398[g] | | 30S ribosomal protein S2 | Cytoplasmic | (1) | | 2 | 4 | 2 | (1) | 19.2 |
| EF2556[f] | | Fumarate reductase flavoprotein subunit | Lipid anchor SP-II, ATG-CT | 40 | 39 | 40 | 41 | 22 | 37 | 79.2 |
| EF2718[ge] | | 50S ribosomal protein L1 | Cytoplasmic | | | 2 | 2 | | | 24.0 |
| EF2746 | | DltD protein | N-terminal anchor | (1) | 2 | 2 | 3 | (1) | 2 | 10.4 |

**Table 1 Proteins identified by LC-MS/MS analysis of tryptic fragments obtained after different treatments of intact *E. faecalis* V583** *(Continued)*

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| EF2860[f] | | YkuD putative, peptido-glycan binding domain | Cell wall, SP-I VYF-QS, | 2 | 6 | 5 | 9 | (1) | 7 | 36.3 |
| EF2864 | - | Hypothetical protein | Lipid anchor SP-II, LTA-CR | 2 | 2 | 3 | 4 | 2 | 2 | 25.7 |
| EF2925[g] | | Cold-shock domain-contain protein | Cytoplasmic | (1) | 2 | | | (1) | 2 | 45.5 |
| EF3041[f] | | Pheromone binding protein | Lipid anchor SP-II, LAA-CG | 2 | 3 | 2 | 7 | | 2 | 24.3 |
| EF3256[f] | | Pheromone cAD1 precursor lipoprotein | Lipid anchor SP-II, LAA-CG | | | (1) | 2 | | | 19.4 |
| EFA0003 | | TraC protein | Lipid anchor SP-II, LGA-CN | | | | 2 | | | 9.4 |

Further details on the Sequest-based protein identification process are provided in the Materials and Methods section and in Table S5 (Additional file 4).

[a]Significant hits (or absence thereof, indicated by -) obtained after searches in Pfam [25] for putative and hypothetical proteins.

[b]All data extracted from the LocateP database [12] with two exceptions: EF1033 (annotated as 6-aminohexanoate-cyclic-dimer hydrolase; annotated as putative, in LocateP) and EF2860 annotated as ErfK/YbiS/YcfS/YnhG family protein; annotated as putative, in LocateP).

[c]Predicted localization and potential cleavage site. Localization is based on LocateP annotations, with seven exceptions (for Tables 1 and 2 in total) that are all explicitly mentioned in the text. See also Table 3.

[d]The column shows the number of peptide hits from four biological replicates. Protein identifications were considered significant using the criteria described in Materials and Methods. One criterium was the detection of at least two different peptides; another the detection of peptides in at least two independent parallels; these criteria are met for all listed proteins. Putative, non-significant additional identifications of these proteins (based on just one peptide and/or on just one parallel) are indicated in parentheses.

[e]SecretomeP value >0.5; this means that the protein is predicted to be secreted via a "non-classical" pathway.

[f]Proteins that have been identified as being localised on the surface in a previous study of *E. faecelis* JH2-2 [13].

[g]Cytoplasmic proteins that have been identified in other studies of the surface proteomes of Gram-positive bacteria. See text for references.

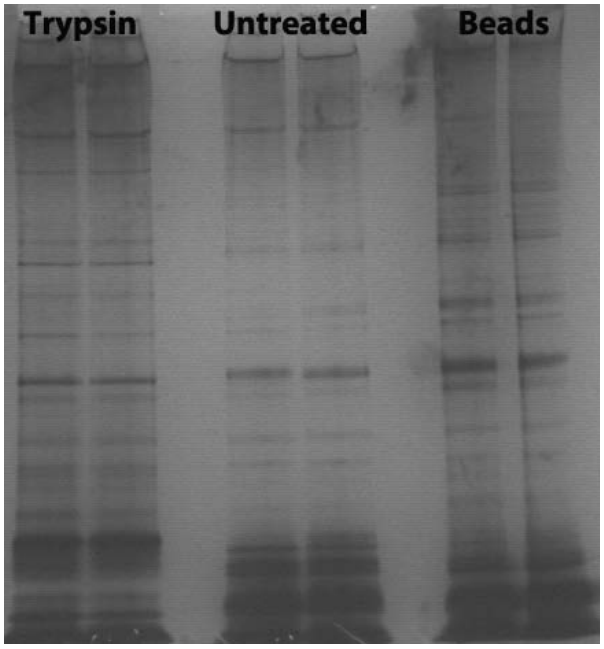



**Figure 1 SDS-PAGE analysis of the supernatants obtained after treating intact cells with trypsin; see materials and methods for details**. The gel shows the results from cells treated with trypsin, cells treated with trypsin beads and a control sample ("untreated") where no trypsin was added. Samples sizes represent approximately the same amount of cells in all lanes.

**Table 3 Summary of the identified proteins grouped according to predicted localization**

| Predicated localization[a] | Number of identified unique proteins | Number in the *E. faecalis* V583 genome[b] | Percent identified |
|---|---|---|---|
| Cytoplasmic | 33 | 2303 | 1.4 |
| Membrane[c] | 3 | 588 | 0.5 |
| Lipid anchor | 17 | 74 | 23.0 |
| N-terminal anchor | 5 | 190 | 2.6 |
| LPxTG proteins | 4 | 38 | 10.5 |
| Cell wall associated | 2 | ND[d] | ND[d] |
| Secreted | 5 | 59 | 8.5 |
| C-terminal anchor | 0 | 4 | 0 |
| Secreted via minor pathway | 0 | 8 | 0 |
| Sum | 69 | 3264 | 2.1 |

[a] Localization data are from LocateP (Zhou et al., 2008), with the seven corrections described in the text. According to Locate P, the numbers of identified cytoplasmic, lipid anchor, N-terminal anchor, cell wall associated and secreted proteins.

[b] Data from the LocateP database.

[c] Containing multiple transmembrane helices.

[d] Not detected (ND), LocateP only predicts cell wall proteins with LPxTG motifs.

### Cytoplasmic proteins

According to the LocateP database 34 of the 69 identified proteins lack an N-terminal signal sequence and are therefore predicted to be cytoplasmic proteins. All these proteins were analysed using PSORTb v.3.0 [28] and the SecretomeP (SecP) 2.0 Server [24,29]. SecP predict proteins that are putatively secreted without having a detectable N-terminal signal sequence, i.e. by "non-classical secretion". Analysis with SecP indicated that four of the proteins predicted to be cytoplasmic (EF0970, EF1523, EF2718 and EF_B0004) may follow non-classical secretion. PSORTb predicted EF_B0004 (TraC protein) to be a cell wall protein and Pfam gave a significant hit against "Bacterial extracellular solute-binding proteins, family 5" (PF00496). It has been shown that TraC proteins play a role as surface pheromone receptor and are thereby involved in the regulation of the conjugation process [30,31]. Therefore, EF_B0004 was classified as a cell wall protein in this study. The other three proteins were retained as cytoplasmic, despite the fact that PSORTb predicted EF0970 (ribosomal protein L27) to have an extracellular location.

Identification of cytoplasmic proteins at extracellular locations is not unusual and at least 20 of the 33 proteins found in this study have been identified in previous studies of the secretomes or surface proteomes of Gram-positive bacteria (Tables 1 and 2) [13,18,26,32-36]. The majority of the identified cytoplasmic proteins were unique to the trypsin fraction and/or the beads fraction, indicating that these proteins bind to the cell envelope and need to be "shaved" from the surface, despite the lack of known binding motifs or domains. Many of the identified cytoplasmic proteins are highly abundant proteins like ribosomal proteins (Rbps; more than 20% of all identified proteins), EF-Tu/G, GADPH and chaperones, which suggests that cell lysis rather than an unknown active secretion process determines their extracellular presence. While the cell viability checks described above indicate that cell lysis due to the trypsin treatment is unlikely, it is conceivable that cell lysis in the cell culture prior to the trypsin treatment may have released intracellular proteins that somehow have re-associated with the cell envelope and escaped proteolytic degradation [26,37].

While cytoplasmic proteins found in studies such as the present generally must be considered contaminants, there have been speculations in the literature that some of these actually may have extracellular functions. One example is Rbp L7/L12 (EF2715) which has been identified at the surface of several Gram positive bacteria [19,26,34,35,38]. We found Rbp L7/L12 in the beads fraction only and this implies an exposed localization, similar to the localization suggested in *B. subtilis* [26]. Bacterial Rbp L7/L12 has immunogenic properties in humans [38,39] and is being explored as candidate

**Table 2 Additional proteins identified using the SDS-PAGE approach after different treatments**

| Gene | Pfam[a] | Gene products[b] | Predicted localization[c] | Peptide hits[d] | | | Coverage % |
|---|---|---|---|---|---|---|---|
| | | | | Untreated 2 hours | Trypsin 2 hours | Beads 2 hours | |
| EF0517[f] | | 2-dehydropantoate 2-reductase | Cytoplasmic | | | 2 | 13.1 |
| EF0968[f] | | 50S ribosomal protein L21 | Cytoplasmic | | | 2 | 35.3 |
| EF2221 | | ABC transporter. substrate-binding protein | Lipid anchor SP-II, LSA-CG | | | 2 | 8.8 |
| EF2224 | Four DUF11 repeats | Cell wall surface anchor family protein | Cell wall, LPxTG, SP-I, MNA-FA | | | 2 | 1.4 |
| EF2633[f] | | Chaperonin. GroEL | Cytoplasmic | | | 4 | 9.2 |
| EF2713 | Gram positive anchor | Cell wall surface anchor family protein | Cell wall, LPxTG, SP-I VWA-ED | | 2 | | 9.4 |
| EF2715[f] | | Ribosomal protein L7/L12 | Cytoplasmic | | | 2 | 29.5 |
| EF2857[e] | | Penicillin-binding protein 2B | N-terminal anchor | | | 3 | 4.8 |
| EF2903[e] | | ABC transporter, substrate-binding protein | Lipid anchor SP-II, LGA-CG | | 6 | 2 | 20.2 |
| EF3106[e] | | Peptide ABC transporter. peptide-binding protein | Lipid anchor SP-II, LAA-CG | | 2 | | 3.4 |
| EF3257 | | Pyridine nucleotide-disulfide family oxidoreductase | Multiple transmembrane proteins | | 2 | | 7.3 |
| EFA0052 | | Surface exclusion protein Sea1 | Cell wall LPxTG SP-1, VQA-AE | 2 | 2 | 2 | 6.3 |

The gel approach yielded 25 unique proteins in total but only 12 of these were novel compared to the list of Table 1; see Additional file 2 for more details. Further details on the Sequest-based protein identification process are provided in the Materials and Methods section and in Table S6 (Additional file 5).

[a]Significant hits obtained after searches in Pfam [25] for putative and hypothetical proteins.

[b]Data extracted from the LocateP database [12].

[c]Predicted localization and potential cleavage site. Localization is based on LocateP annotations, with seven exceptions (for Tables 1 and 2 in total) that are all explicitly mentioned in the text. See also Table 3.

[d]Number of peptide hits in each of the three treatments. Protein identifications were considered significant using the criteria described in Materials and Methods. Proteins were only considered a significant hit if at least two unique peptides were found.

[e] Proteins that have been identified as being localised on the surface in a previous study of *E. faecelis* JH2-2 [13].

[f]Cytoplasmic proteins that have been identified in other studies of the surface proteomes of Gram-positive bacteria. See text for references.

antigen for vaccine purposes [40,41]. Recent data showed that exposure of bacterial Rbp L7/L12 is a risk factor for colorectal cancer and stimulates progression of adenomas into carcinomas [42]. There has also been some speculation about possible adhesive roles of extracellular EF-Tu, DnaK, enolase and GAPDH, all identified in the present study and in a recent study of the laboratory strain *Enterococcus faecalis* JH2-2 [13], since these proteins bind strongly to human plasminogen [32].

### Secreted proteins

Proteins were annotated as being secreted to the culture medium if the bioinformatic analyses showed the presence of a SPaseI cleavage site and did not reveal any sequence or domain known to be involved in covalent or non-covalent binding to the cell wall. Only five such proteins were identified indicating that few secreted proteins are closely associated to the microbe. Three of the secreted proteins (EF0123, EF0394, EF0417) are annotated as N-terminally anchored in the LocateP database. It is, however, not easy to differentiate between secreted proteins with processed signal peptides and proteins that retain their signal peptides as N-terminal membrane anchors [12]. The Signal P server predicted all three proteins to contain a unique cleavage site, when using both algorithms in the program. All three were detected even without trypsin treatment, suggesting a loose association with the cell envelope. Taken together, we conclude that EF0123, EF0394 and EF0417 are secreted proteins.

EF2174 was detected after trypsin treatment only and putatively encodes a glycoside hydrolase belonging to family GH25 [43]. This family comprises enzymes with lysozyme activity that are thought to be involved in peptidoglycan remodelling during cell division [44].

The well known secreted metalloprotease coccolysin, a gelatinase (EF1818; GelE) was identified after all treatments (Table 1), indicating that GelE is loosely associated to the cell surface. Coccolysin is known to be associated with virulence and is capable of degrading cellular tissues during infection, cleaving substrates such as haemoglobin, collagen and fibrin [45,46]. There are also indications that GelE is required for biofilm formation [47].

### Proteins on the surface

Among the 31 non-secreted non-cytoplasmic proteins found in this study (Table 3), three are annotated as

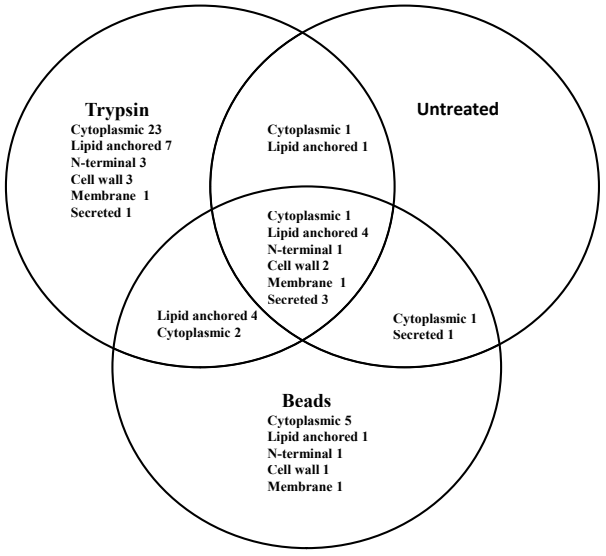



**Figure 2 Overview over the identifcation of 69 surface-located proteins using various treatments of enterococcal cells**. The putative cellular localizations of the identified proteins, as provided in Table 3 are indicated. In total 58, 16 and 29 proteins were identified in the trypsin, untreated and beads samples, respectively. Note that this figure combines the data from the one and two hour treatments and includes the data from both the direct LC-MS/MS analysis and the SDS-PAGE-based approach. More details about the numbers of identified proteins in the various experiments are provided in Additional file 2. Proteins containing multiple transmembrane helices are indicated by "Membrane"; proteins indicated by "Cell Wall" include 4 proteins with LPxTG anchors and two proteins containing domains known to display cell wall binding affinity.

integral membrane proteins. We identified only 0.5% of the predicted transmembrane proteins in the genome of *E. faecalis* V583 (Table 3). Such low numbers of identified transmembrane proteins are not unusual [19,26,37,48] and are likely to be due to limited accessibility of the proteins and/or a limited ability of trypsin to penetrate the cell wall. One of these transmembrane proteins, the sulfatase (EF1264), was detected by many peptides after all treatments (see above and Figure 3). The other two were only detected after trypsin treatment and only by a minimal number of peptides. EF0502 is a 781-residue hypothetical protein which according to topology predictions contains several extracellular domains. Both detected peptides are predicted to have an extracellular location. EF3257 is a 648-residue oxidoreductase belonging to the pyridine nucleotide-disulfide family with probably only two transmembrane helices and a large extracellular domain. Both detected peptides stem from extracellular domains.

Five identified proteins are thought to be N-terminally anchored to the cell membrane via a Sec-type signal peptide that is not cleaved off during secretion. Relatively few such proteins were identified (Table 3), suggesting that they are expressed at low levels or that they generally have low accessibility for trypsin. Among these proteins are two penicillin-binding proteins (EF2857 & EF0991), one amidase (EF0737) and one protein of unknown function (EF1319). EF2857 and EF0991 are class B penicillin binding proteins (PBP), which are transpeptidases involved in the final stage of cell wall synthesis [49]. *E. faecalis* has three class B PBPs that have low affinity for β-lactams and can take over the transpeptidase activity of more high affinity PBPs when these are inhibited by antibiotics [50]. We also identified a L,D-transpeptidase (EF2860; YkuD domain), in all treatments and with a high number of peptide hits, indicating that this protein is abundant at the surface of V583. Studies on *E. faecium* have indicated that L,D-transpeptidase activity may represent another way to bypass inhibition of PBPs [51,52].

The fifth protein, DltD (EF2746), is also involved in the biosynthesis of surface structures. The *dltD* gene is part of an operon consisting four genes (*dltA-dltD*) whose gene products are all necessary to incorporate D-alanyl residues into lipoteichoic acids (LTA) [53,54]. It has previously been shown that disruption of genes in the *dlt* operon of *E. faecalis* lead to diminished adhesion to eukaryotic cells and less biofilm formation, indicating that the *dlt* operon is involved in pathogenicity [55]. Interestingly, it is not fully established whether the N-terminal anchor tethers DltD to the inner or outer leaflet of the membrane [53,56]. The fact that DltD was detected after all types of treatments may be taken to suggest that the protein is attached to the outer leaflet, as one would expect for proteins using a Sec-type signal peptide as membrane anchor.

Of the four proteins containing a putative LPxTG anchor, two (EF2224 & EF2713) have unknown functions. EF2713 is up-regulated when *E. faecalis* V583 is grown in the presence of blood [57], indicating a putative role of this protein in infection processes. The LPxTG anchor protein EF2224 contains five copies of a so-called DUF11 repeat with unknown function and is a putative member of the MSCRAMM (mirobial surface component recognizing adhesive matrix molecules) family of proteins [58]. These proteins contain tandemly repeated immunoglobulin-like folds as observed for staphylococcal adhesins. The *E. faecalis* V583 genome contains seventeen proteins belonging to the MSCRAMM family [59]. Interestingly, EF2224 is highly expressed during the infection process in humans [59].

The LPxTG anchor protein EF1033 is a lipoamidase (Lpa) cleaving lipoic acids from lipoylated molecules [60,61]. EF1033 was only detected after trypsin treatment, indicative of covalent cell wall attachment. Lipoic acid is an essential sulphur containing cofactor of several

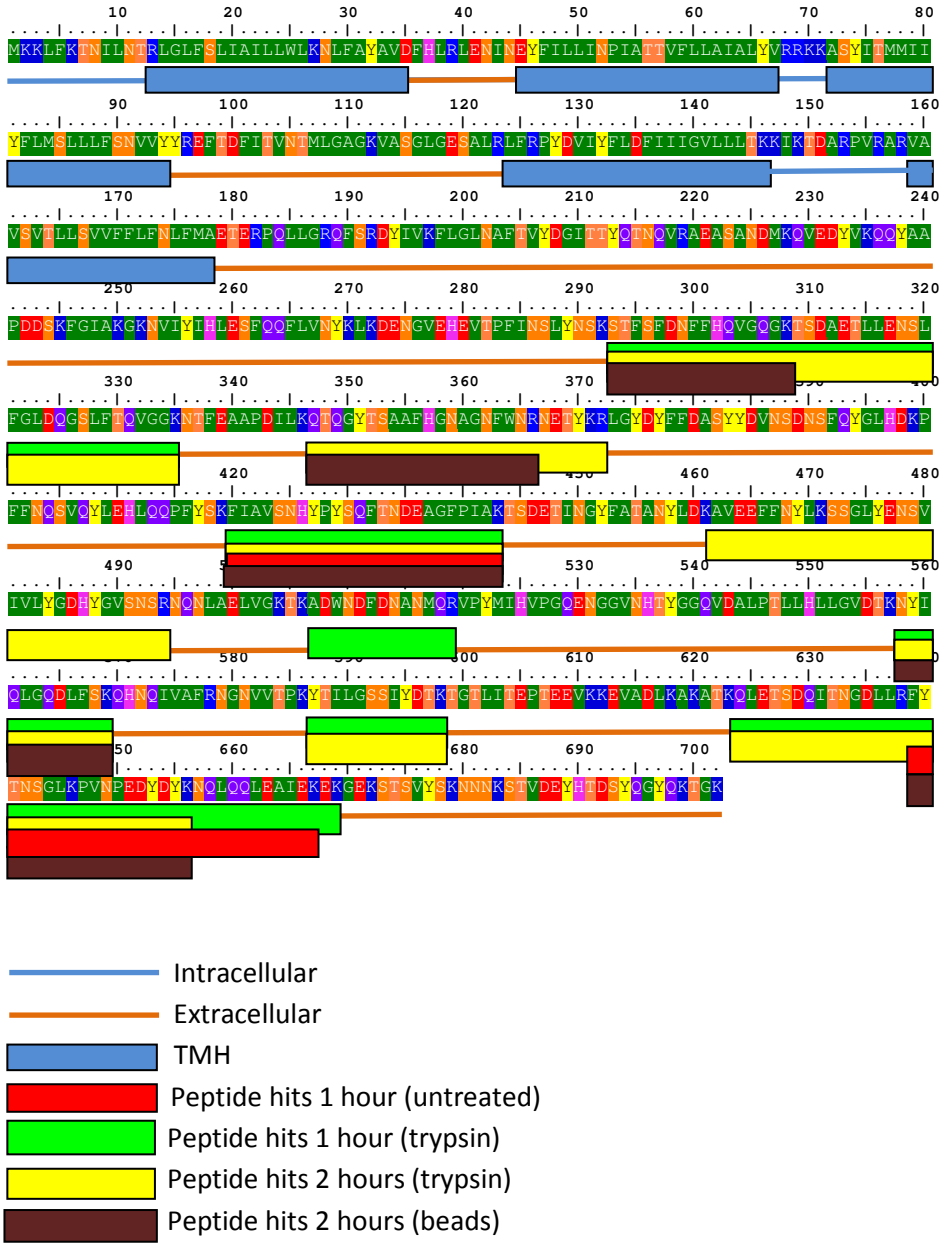



**Figure 3 Amino acid sequence and predicted topology of EF1264**. The predicted trans-memberane helices (TMH) and the identified tryptic peptides after different treatments are mapped on the sequence, using the color coding indicated in the figure.

enzymes. Interestingly, so far, *E. faecalis* is the only bacterial species in which Lpa activity has been detected [61]. Jiang and Cronan [61] speculate that the Lpa is a cytoplasmic salvage enzyme, but our experimental and bioinformatic results indicate that the enzyme is a cell wall anchored protein. Most likely, Lpa recruits its substrates from the environment, such as the GI tract.

The fourth LPxTG protein is the plasmid encoded surface exclusion factor Sea1 (EFA0052) which is involved in the regulation of sex pheromone-controlled conjugation [62].

EF2860 is linked to the cell wall by a peptidoglucan binding domain and encodes a cell wall modifying transpeptidase homologous with YkuD from *Bacillus subtilis*. This protein was found after all three treatments and identified with relatively many peptide hits (Table 1), indicating that EF2860 is abundant on the surface and may show a relatively large extent of shedding. This protein may contribute to the antibiotic resistance of *E. faecalis*, as discussed above.

The most populated group of proteins identified in this study are the lipoproteins, of which 17 were

detected, representing 23% of the lipoproteins putatively encoded on the *E. faecalis* genome (Table 3). Twelve of these were detected only after a "shaving" treatment. Seven of the detected lipoproteins are proteins with no predicted function. Two of these unknown lipoproteins (EF0176 & EF0177) are located on the same operon, share 70% sequence identity, and contain a "Basic membrane protein" domain (PF02608) belonging to Clan CL0144 in the Pfam database. This clan consists of proteins that are involved in chemotaxis and membrane transport of sugars as well as outer membrane proteins that are known for their antigenicity in pathogenic bacteria. Both proteins are homologous with a CD4+ T-cell-stimulating antigen in *Listeria* [63]. One of the other proteins with unknown function, EF0164, is annotated as N-terminally anchored in the LocateP database, but is classified as a lipoanchored protein on the basis of our analyses with LipoP.

Of the ten lipoproteins with predicted functions, four lipoproteins resemble the substrate-binding domains of multi-component ABC transporters for the import of peptides (EF0907, EF3106) or sugars (EF2221, EF2903). Three proteins (EF3041, EF3256, EFA0003) are involved in pheromone-regulated processes that include conjugation [31], adding to the two cell wall associated proteins involved in these processes that are discussed above (EFB0004 & EFA0052). EF0071 seems to encode a glycoside hydrolase belonging to clan GH-G in the CAZy database. The protein is classified as N-terminally anchored in the LocateP database, but our analyses with the LipoP program clearly indicated that EF0071 is a lipoanchored protein. EF0685 belongs to the rotamase family and is thus likely to be involved in extracellular protein folding, possibly by exerting prolyl-peptidyl isomerase activity.

The final lipo-anchored protein is EF2556, fumarate reductase, which was detected by remarkably large numbers of peptides after all treatments (Table 1). *E. faecalis* is one of few bacteria that produces substantial amounts of extracellular superoxide. Fumarate reductase is likely to be involved in superoxide production and may thus be an important source of oxidative stress for the host [64]. It has been demonstrated that the superoxide from *E. faecalis* promotes chromosomal instability in mammalian cells and that this can lead to colorectal cancer [65,66].

## Conclusions

In recent years, several analyses of bacterial surface proteomes have been described. Despite the improvements in the mass spectrometry methods, the numbers of identified proteins are normally in the order of 30 - 80, meaning that only a minority of the putative surface-located proteins is being found. In this type of studies, it is common to find a significant fraction of proteins that are thought to be cytoplasmic and there is some evidence that this is not just the result from artefacts such as cell lysis. We show that the large majority of the identified cytoplasmic proteins are only found after treatment with trypsin. This is an important observation, since it shows that these proteins bind tightly to the cell envelope.

In a recent published proteomics-based analysis 38 proteins were identified on the surface of *E. faecalis* JH2-2 [13]. Seventeen of these proteins were found after using a method similar to the one used here (i.e. surface shaving with trypsin) and seven of these were also found in the present study (EF0177, EF0201, EF0907, EF1613, EF1964, EF2556 and EF3256). Disparities between this type of studies may be due to many factors, e.g. differences between the strains and growth conditions or differences in the confidence of protein identification. Benachour et al [13] allowed protein identification on the basis of only one peptide hit, while we required at least two peptides for confident identification. Large inter-strain variation has been observed in several previous studies, both for secreted proteins and for proteins detected by a trypsin-shaving approach [34,67,68].

In conclusion, our studies reveal 69 surface-located proteins in *E. faecalis* V583 with varying roles in bacterial behaviour. Several of the identified proteins are involved in cell wall synthesis and maintenance as well as in cell-cell communication and seem interesting targets for drug design. We detected only a few proteins with known or conceivable functions in adhesion, but such proteins may be among the many identified proteins with unknown function. Clearly, the identified proteins with unknown function stand out as targets for more in-depth investigations and several of these are currently subjected to knock-out studies in our laboratory.

## Additional material

**Additional file 1: Figure S1: Control of viability of the cells before and after incubation for one or two hours with trypsin, trypsin beads or without any enzyme.**

**Additional file 2: Table S1-S4: Number of identified proteins in each treatment PDF.**

**Additional file 3: Figure S2: Nucleotide and amino acid sequences of EF1033 and EF2713 after adjustment of the start codon.**

**Additional file 4: Table S5: Proteome data of the proteins identified by LC-MS analysis after different treatments.**

**Additional file 5: Table S6: Proteome data of the proteins identified using the SDS-PAGE approach after different treatments.**

#### Acknowledgements

We thank Dr. Manuel Josè Rodriquez Ortega for helpful discussions. This work was funded by grants 159058, 183627 and 183637 and from The Norwegian Research Council.

**Author details**
[1]Department of Chemistry, Biotechnology, and Food Science, The Norwegian University of Life Sciences, 1432 Ås, Norway. [2]Department of Molecular Biosciences, Glyconor Mass Spectrometry, University of Oslo, 0316 Oslo, Norway.

**Authors' contributions**
LAB, GM, TR and VE developed the initial concept for this study. All authors participated in experimental design and coordination of the study. LAB and TR carried out shaving experiments. MS, WEJ and ØB contributed to the experimental design of the mass spectrometry experiments. LAB, TR, WEJ carried out the LC-MS/MS analysis. GM and LAB did the bioinformatics analysis. LAB, GM and VE drafted the paper, implementing contributions from all other authors. All authors read, corrected and approved the final manuscript.

Received: 2 December 2010 Accepted: 1 March 2011
Published: 1 March 2011

**References**

1. Fisher K, Phillips C: **The ecology, epidemiology and virulence of Enterococcus.** *Microbiology* 2009, **155**:1749-1757.
2. Koch S, Hufnagel M, Theilacker C, Huebner J: **Enterococcal infections: host response, therapeutic, and prophylactic possibilities.** *Vaccine* 2004, **22**:822-830.
3. Klein G: **Taxonomy, ecology and antibiotic resistance of enterococci from food and the gastro-intestinal tract.** *Int J Food Microbiol* 2003, **88**:123-131.
4. Kayser FH: **Safety aspects of enterococci from the medical point of view.** *Int J Food Microbiol* 2003, **88**:255-262.
5. Paulsen IT, Banerjei L, Myers GS, Nelson KE, Seshadri R, Read TD, Fouts DE, Eisen JA, Gill SR, Heidelberg JF, *et al*: **Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*.** *Science* 2003, **299**:2071-2074.
6. Bourgogne A, Garsin DA, Qin X, Singh KV, Sillanpaa J, Yerrapragada S, Ding Y, Dugan-Rocha S, Buhay C, Shen H, *et al*: **Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF.** *Genome Biol* 2008, **9**:R110.
7. Domann E, Hain T, Ghai R, Billion A, Kuenne C, Zimmermann K, Chakraborty T: **Comparative genomic analysis for the presence of potential enterococcal virulence factors in the probiotic *Enterococcus faecalis* strain Symbioflor 1.** *Int J Med Microbiol* 2007, **297**:533-539.
8. Lebreton F, Riboulet-Bisson E, Serror P, Sanguinetti M, Posteraro B, Torelli R, Hartke A, Auffray Y, Giard JC: **ace, which encodes an adhesin in *Enterococcus faecalis*, is regulated by Ers and is involved in virulence.** *Infect Immun* 2009, **77**:2832-2839.
9. Coburn PS, Pillar CM, Jett BD, Haas W, Gilmore MS: ***Enterococcus faecalis* senses target cells and in response expresses cytolysin.** *Science* 2004, **306**:2270-2272.
10. Engelbert M, Mylonakis E, Ausubel FM, Calderwood SB, Gilmore MS: **Contribution of gelatinase, serine protease, and fsr to the pathogenesis of *Enterococcus faecalis* endophthalmitis.** *Infect Immun* 2004, **72**:3628-3633.
11. Coburn PS, Gilmore MS: **The *Enterococcus faecalis* cytolysin: a novel toxin active against eukaryotic and prokaryotic cells.** *Cell Microbiol* 2003, **5**:661-669.
12. Zhou M, Boekhorst J, Francke C, Siezen RJ: **LocateP: genome-scale subcellular-location predictor for bacterial proteins.** *BMC Bioinformatics* 2008, **9**:173.
13. Benachour A, Morin T, Hebert L, Budin-Verneuil A, Le Jeune A, Auffray Y, Pichereau V: **Identification of secreted and surface proteins from *Enterococcus faecalis*.** *Can J Microbiol* 2009, **55**:967-974.
14. Cole JN, Ramirez RD, Currie BJ, Cordwell SJ, Djordjevic SP, Walker MJ: **Surface analyses and immune reactivities of major cell wall-associated proteins of group a streptococcus.** *Infect Immun* 2005, **73**:3137-3146.
15. Antelmann H, Tjalsma H, Voigt B, Ohlmeier S, Bron S, van Dijl JM, Hecker M: **A proteomic view on genome-based signal peptide predictions.** *Genome Res* 2001, **11**:1484-1502.
16. Voigt B, Schweder T, Sibbald MJJB, Albrecht D, Ehrenreich A, Bernhardt J, Feesche J, Maurer KH, Gottschalk G, Dijl JMv, *et al*: **The extracellular proteome of *Bacillus licheniformis* grown in different media and under different nutrient starvation conditions.** *Proteomics* 2006, **6**:268-281.
17. Mujahid S, Pechan T, Wang C: **Improved solubilization of surface proteins from *Listeria monocytogenes* for 2-DE.** *Electrophoresis* 2007, **28**:3998-4007.
18. Dumas E, Desvaux M, Chambon C, Hebraud M: **Insight into the core and variant exoproteomes of *Listeria monocytogenes* species by comparative subproteomic analysis.** *Proteomics* 2009, **9**:3136-3155.
19. Rodriguez-Ortega MJ, Norais N, Bensi G, Liberatori S, Capo S, Mora M, Scarselli M, Doro F, Ferrari G, Garaguso I, *et al*: **Characterization and identification of vaccine candidate proteins through analysis of the group A *Streptococcus* surface proteome.** *Nat Biotechnol* 2006, **24**:191-197.
20. Rappsilber J, Ishihama Y, Mann M: **Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics.** *Analyt Chem* 2003, **75**:663-670.
21. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340**:783-795.
22. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A: **Prediction of lipoprotein signal peptides in Gram-negative bacteria.** *Protein Sci* 2003, **12**:1652-1662.
23. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.** *J Mol Biol* 2001, **305**:567-580.
24. Bendtsen JD, Jensen LJ, Blom N, von Heijne G, Brunak S: **Feature-based prediction of non-classical and leaderless protein secretion.** *Protein Eng Des Sel* 2004, **17**:349-356.
25. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, *et al*: **The Pfam protein families database.** *Nucleic Acids Res* 2010, **38**:D211-222.
26. Tjalsma H, Lambooy L, Hermans PW, Swinkels DW: **Shedding & shaving: Disclosure of proteomic expressions on a bacterial face.** *Proteomics* 2008, 1415-1428.
27. Rho JH, Wright DP, Christie DL, Clinch K, Furneaux RH, Roberton AM: **A novel mechanism for desulfation of mucin: identification and cloning of a mucin-desulfating glycosidase (sulfoglycosidase) from *Prevotella* strain RS2.** *J Bacteriol* 2005, **187**:1543-1551.
28. Gardy JL, Laird MR, Chen F, Rey S, Walsh CJ, Ester M, Brinkman FSL: **PSORTb v.2.0: Expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis.** *Bioinformatics* 2005, **21**:617-623.
29. Bendtsen JD, Kiemer L, Fausboll A, Brunak S: **Non-classical protein secretion in bacteria.** *BMC Microbiology* 2005, **5**:58.
30. Weaver KE, Clewell DB: **Regulation of the pAD1 sex pheromone response in *Enterococcus faecalis*: effects of host strain and traA, traB, and C region mutants on expression of an E region pheromone-inducible lacZ fusion.** *J Bacteriol* 1990, **172**:2633-2641.
31. Chandler JR, Dunny GM: **Enterococcal peptide sex pheromones: synthesis and control of biological activity.** *Peptides* 2004, **25**:1377-1388.
32. Schaumburg J, Diekmann O, Hagendorff P, Bergmann S, Rohde M, Hammerschmidt S, Jansch L, Wehland J, Karst U: **The cell wall subproteome of *Listeria monocytogenes*.** *Proteomics* 2004, **4**:2991-3006.
33. Trost M, Wehmhoner D, Karst U, Dieterich G, Wehland J, Jansch L: **Comparative proteome analysis of secretory proteins from pathogenic and nonpathogenic *Listeria* species.** *Proteomics* 2005, **5**:1544-1557.
34. Severin A, Nickbarg E, Wooters J, Quazi SA, Matsuka YV, Murphy E, Moutsatsos IK, Zagursky RJ, Olmsted SB: **Proteomic analysis and identification of *Streptococcus pyogenes* surface-associated proteins.** *J Bacteriol* 2007, **189**:1514-1522.
35. Beck HC, Madsen SM, Glenting J, Petersen J, Israelsen H, Norrelykke MR, Antonsson M, Hansen AM: **Proteomic analysis of cell surface-associated proteins from probiotic *Lactobacillus plantarum*.** *FEMS Microbiol Lett* 2009, **297**:61-66.
36. Sanchez B, Bressollier P, Chaignepain S, Schmitter JM, Urdaci MC: **Identification of surface-associated proteins in the probiotic bacterium *Lactobacillus rhamnosus* GG.** *Int Dairy J* 2009, **19**:85-88.
37. Rodriguez-Ortega M, Luque I, Tarradas C, Barcena J: **Overcoming function annotation errors in the Gram-positive pathogen *Streptococcus suis* by a proteomics-driven approach.** *BMC Genomics* 2008, **9**:588.
38. Tjalsma H, Lasonder E, Schöller-Guinard M, Swinkels DW: **Shotgun immunoproteomics to identify disease-associated bacterial antigens: Application to human colon cancer.** *PROTEOMICS-Clin Appl* 2007, **1**:429-434.
39. Voland P, Weeks DL, Vaira D, Prinz C, Sachs G: **Specific identification of three low molecular weight membrane-associated antigens of *Helicobacter pylori*.** *Aliment Pharmacol Ther* 2002, **16**:533-544.

40. Ribeiro LA, Azevedo V, Le Loir Y, Oliveira SC, Dieye Y, Piard JC, Gruss A, Langella P: **Production and targeting of the *Brucella abortus* antigen L7/L12 in *Lactococcus lactis*: a first step towards food-grade live vaccines against brucellosis.** *Appl Environ Microbiol* 2002, **68**:910-916.
41. Mallick AI, Singha H, Chaudhuri P, Nadeem A, Khan SA, Dar KA, Owais M: **Liposomised recombinant ribosomal L7/L12 protein protects BALB/c mice against Brucella abortus 544 infection.** *Vaccine* 2007, **25**:3692-36704.
42. Boleij A, Roelofs R, Schaeps RMJ, Schülin T, Glaser P, Swinkels DW, Kato I, Tjalsma H: **Increased exposure to bacterial antigen RpL7/L12 in early stage colorectal cancer patients.** *Cancer* 2010, **116**:4014-4022.
43. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B: **The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics.** *Nucleic Acids Res* 2009, **37**:D233-238.
44. **Glycoside Hydrolase Family 25.** *[CAZypedia]* [http://www.cazypedia.org/].
45. Mäkinen PL, Clewell DB, An F, Mäkinen KK: **Purification and substrate specificity of a strongly hydrophobic extracellular metalloendopeptidase ("gelatinase") from *Streptococcus faecalis* (strain 0G1-10).** *J Biol Chem* 1989, **264**:3325-3334.
46. Kayaoglu G, Ørstavik D: **Virulence factors of *Enterococcus faecalis*: Relationship to endodontic disease.** *Crit Rev Oral Biol Med* 2004, **15**:308-320.
47. Mohamed JA, Huang DB: **Biofilm formation by enterococci.** *J Med Microbiol* 2007, **56**:1581-1588.
48. Doro F, Liberatori S, Rodriguez-Ortega MJ, Rinaudo CD, Rosini R, Mora M, Scarselli M, Altindis E, D'Aurizio R, Stella M, *et al*: **Surfome analysis as a fast track to vaccine discovery: Identification of a novel protective antigen for Group B *Streptococcus* hypervirulent strain COH1.** *Mol Cell Proteomics* 2009, **8**:1728-1737.
49. Sauvage E, Kerff F, Terrak M, Ayala JA, Charlier P: **The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis.** *FEMS Microbiol Rev* 2008, **32**:234-258.
50. Zapun A, Contreras-Martel C, Vernet T: **Penicillin-binding proteins and β-lactam resistance.** *FEMS Microbiol Rev* 2008, **32**:361-385.
51. Mainardi JL, Legrand R, Arthur M, Schoot B, van Heijenoort J, Gutmann L: **Novel mechanism of β-lactam resistance due to bypass of DD-transpeptidation in *Enterococcus faecium*.** *J Biol Chem* 2000, **275**:16490-16496.
52. Biarrotte-Sorin S, Hugonnet JE, Delfosse V, Mainardi JL, Gutmann L, Arthur M, Mayer C: **Crystal structure of a novel beta-lactam-insensitive peptidoglycan transpeptidase.** *J Mol Biol* 2006, **359**:533-538.
53. Neuhaus FC, Baddiley J: **A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in Gram-positive bacteria.** *Microbiol Mol Biol Rev* 2003, **67**:686-723.
54. Velez MP, Verhoeven TLA, Draing C, Von Aulock S, Pfitzenmaier M, Geyer A, Lambrichts I, Grangette C, Pot B, Vanderleyden J, *et al*: **Functional analysis of D-alanylation of lipoteichoic acid in the probiotic strain *Lactobacillus rhamnosus* GG.** *Appl Environ Microbiol* 2007, **73**:3595-3604.
55. Fabretti F, Theilacker C, Baldassarri L, Kaczynski Z, Kropec A, Holst O, Huebner J: **Alanine esters of enterococcal lipoteichoic acid play a role in biofilm formation and resistance to antimicrobial peptides.** *Infect Immun* 2006, **74**:4164-4171.
56. Debabov DV, Kiriukhin MY, Neuhaus FC: **Biosynthesis of lipoteichoic acid in *Lactobacillus rhamnosus*: role of DltD in D-alanylation.** *J Bacteriol* 2000, **182**:2855-2864.
57. Vebø HC, Snipen L, Nes IF, Brede DA: **The transcriptome of the nosocomial pathogen *Enterococcus faecalis* V583 reveals adaptive responses to growth in blood.** *PLoS One* 2009, **4**:e7660.
58. Patti JM, Allen BL, McGavin MJ, Hook M: **MSCRAMM-mediated adherence of microorganisms to host tissues.** *Annu Rev Microbiol* 1994, **48**:585-617.
59. Sillanpaa J, Xu Y, Nallapareddy SR, Murray BE, Hook M: **A family of putative MSCRAMMs from *Enterococcus faecalis*.** *Microbiology* 2004, **150**:2069-2078.
60. Reed LJ, Koike M, Levitch ME, Leach FR: **Studies on the nature and reactions of protein-bound lipoic acid.** *J Biol Chem* 1958, **232**:143-158.
61. Jiang Y, Cronan JE: **Expression cloning and demonstration of *Enterococcus faecalis* lipoamidase (pyruvate dehydrogenase inactivase) as a Ser-Ser-Lys triad amidohydrolase.** *J Biol Chem* 2005, **280**:2244-2256.
62. Galli D, Friesenegger A, Wirth R: **Transcriptional control of sex-pheromone-inducible genes on plasmid pAD1 of *Enterococcus faecalis* and sequence analysis of a third structural gene for (pPD1-encoded) aggregation substance.** *Mol Microbiol* 1992, **6**:1297-1308.
63. Sanderson S, Campbell DJ, Shastri N: **Identification of a CD4+ T cell-stimulating antigen of pathogenic bacteria by expression cloning.** *J Exp Med* 1995, **182**:1751-1757.
64. Huycke MM, Moore D, Joyce W, Wise P, Shepard L, Kotake Y, Gilmore MS: **Extracellular superoxide production by *Enterococcus faecalis* requires demethylmenaquinone and is attenuated by functional terminal quinol oxidases.** *Mol Microbiol* 2001, **42**:729-740.
65. Xingmin W, Mark MH: **Extracellular superoxide production by *Enterococcus faecalis* promotes chromosomal instability in mammalian cells.** *Gastroenterology* 2007, **132**:551-561.
66. Huycke MM, Abrams V, Moore DR: ***Enterococcus faecalis* produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA.** *Carcinogenesis* 2002, **23**:529-536.
67. Dreisbach A, Hempel K, Buist G, Hecker M, Becher D, van Dijl JM: **Profiling the surfacome of *Staphylococcus aureus*.** *Proteomics* 2010.
68. Ziebandt AK, Kusch H, Degner M, Jaglitz S, Sibbald MJJB, Arends JP, Chlebowicz MA, Albrecht D, Pantuček R, Doškar J, *et al*: **Proteomics uncovers extreme heterogeneity in the *Staphylococcus aureus* exoproteome due to genomic plasticity and variant gene regulation.** *Proteomics* 2010, **10**:1634-1644.

doi:10.1186/1471-2164-12-135
**Cite this article as:** Bøhle *et al.*: Identification of surface proteins in *Enterococcus faecalis* V583. *BMC Genomics* 2011 **12**:135.

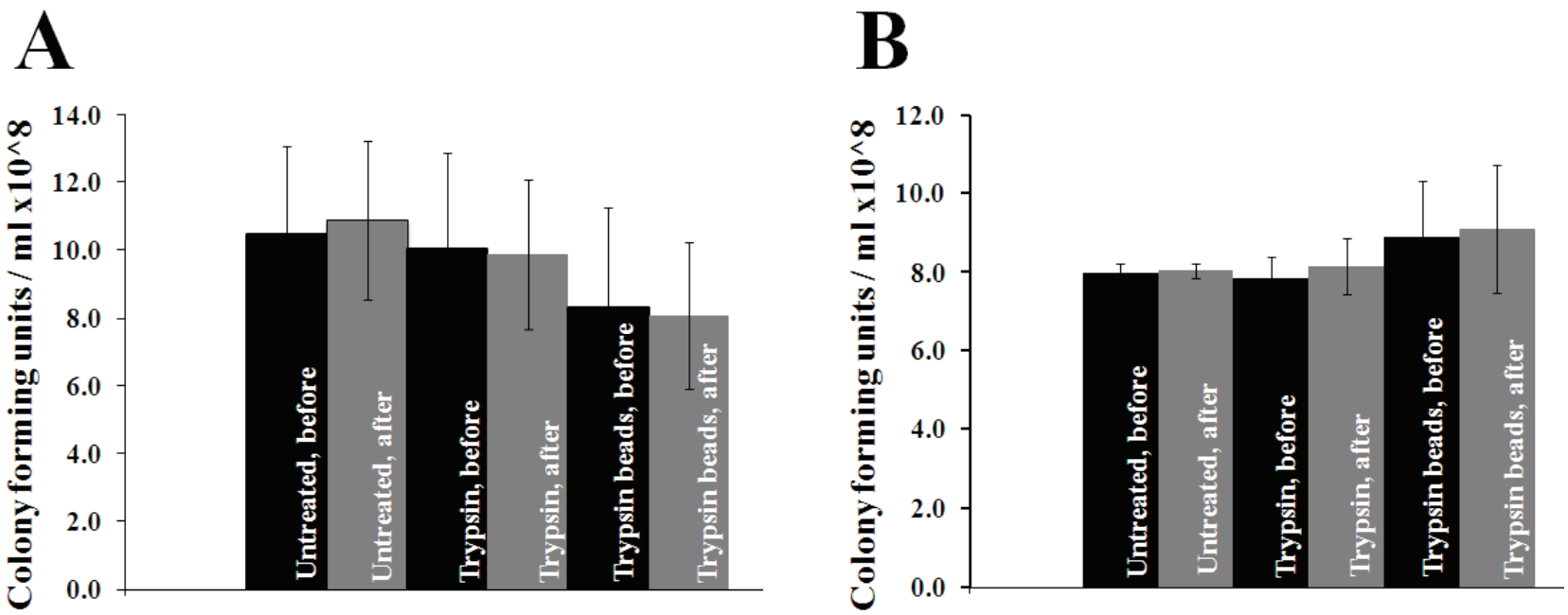



**Figure S1:** Control of viability of the cells before and after incubation for one (A) or two (B) hours with trypsin or trypsin beads or without any enzyme. Colony forming units were determined by plating appropriate dilutions of the cell suspensions on MRS plates. Note that the various treatments were done with different cell cultures and that this explains why the numbers of cells vary between the different treatments.

**Table S1.** Proteins identified in the untreated fractions after one or two hours, grouped according to predicted localization.

| Protein class | 1 hour | 2 hours | Total unique proteins |
|---|---|---|---|
| Cytoplasmic | 0 | 3 | 3 |
| Membrane[a] | 1 | 0 | 1 |
| N-terminal anchor | 0 | 1 | 1 |
| Lipid anchor | 4 | 5 | 5 |
| Cell wall | 1 | 1 | 1 |
| Secreted | 4 | 3 | 4 |
| Sum | 10 | 13 | 15 |

[a]Containing multiple transmembrane helices

**Table S2.** Proteins identified after treating with free trypsin for one or two hours, grouped according to predicted localization.

| Protein class | 1 hour | 2 hours | Total unique proteins |
|---|---|---|---|
| Cytoplasmic | 17 | 21 | 27 |
| Membrane[a] | 1 | 1 | 1 |
| Lipid anchor | 8 | 13 | 14 |
| N-terminal anchor | 1 | 4 | 4 |
| Cell wall | 3 | 3 | 3 |
| Secreted | 3 | 4 | 4 |
| Sum | 33 | 46 | 53 |

[a]Containing multiple transmembrane helices

**Table S3.** Proteins identified after treating with trypsin beads for one or two hours, grouped according to predicted localization.

| Protein class | 1 hour | 2 hours | Total unique proteins |
|---|---|---|---|
| Cytoplasmic | 4 | 2 | 5 |
| [a]Membrane | 0 | 2 | 2 |
| Lipid anchor | 3 | 6 | 7 |
| N-terminal anchor | 0 | 1 | 1 |
| Cell wall | 0 | 1 | 1 |
| Secreted | 1 | 3 | 4 |
| Sum | 8 | 15 | 20 |

[a]Containing multiple transmembrane helices

**Table S4.** Proteins identified via the SDS-PAGE approach. This approach was only used for samples obtained after two hours of incubation.

| Protein class | No trypsin | Free trypsin | Trypsin-beads | Total unique proteins from gel approach | Novel compared to direct LC-MS/MS |
|---|---|---|---|---|---|
| Cytoplasmic | 2 | 4 | 7 | 9 | 4 |
| MTH | 0 | 1 | 0 | 1 | 1 |
| Lipid anchor | 3 | 7 | 6 | 8 | 3 |
| N-terminal anchor | 0 | 0 | 1 | 1 | 1 |
| Cell wall | 1 | 2 | 2 | 3 | 3 |
| Secreted | 3 | 0 | 2 | 3 | 0 |
| Sum | 9 | 14 | 18 | 25 | 12 |
| Only found via gels | 1 | 5 | 9 | 12 | |

[a]Containing multiple transmembrane helices

**EF1033**

```
   1 CTGTATTAAT AGATAAAGGA TGTGACTTAT GAAGAAGTCG ATTTTGTTCA
                                     M   K  K  S   I  L  F  K
  51 AAAAGTTAGG GATTATTTTA TTAATTAGCC AAACACTGGT AGGGGTACCA
       K  L  G   I  I  L   L  I  S  Q   T  L  V   G  V  P
 101 ATGTTGGCAC AAGAAAGTAT ACTAGAAACA ACCGTTCAAA CAGAGACGGA
     M  L  A  Q   E  S  I   L  E  T   T  V  Q  T   E  T  E
 151 ATCAGTAACA ACAGAAACCA GTCAGACTGT AGCTAACTTG GAATCTGAAA
      S  V  T   T  E  T  S   Q  T  V   A  N  L   E  S  E  T
 201 CTACTAGCCA AACGGTGATG CAGGAAAAAG AATCCTCTTC GGCAATCGCC
       T  S  Q   T  V  M   Q  E  K  E   S  S  S   A  I  A
 251 GAAAGCAGTA GCGGAAATGC GGTTGCAGTA ACTACTGAAA CCACAAATGA
     E  S  S  S   G  N  A   V  A  V   T  T  E  T   T  N  E
-------------------------------------------------------------
2101 TCAGGGAACC AAATCAACAG AAAGTACTCA TGCTTTTTCT GATAAAAATA
      Q  G  T   K  S  T  E   S  T  H   A  F  S   D  K  N  M
2151 TGATTGGGAA AAAAGAACAG CTTCCCAAAA AAGTATTACC AAAAGCAGGT
       I  G  K   K  E  Q   L  P  K  K   V  L  P   K  A  G
2201 GCAGAAGTAC CTAGTACTTT CTGGGTTGTT TTAGGAGGAG CTTTCTTAGT
     A  E  V  P   S  T  F   W  V  V   L  G  G  A   F  L  V
2251 AACGAGTGGA ACGATTTATA TAAGGAAAAC TAGAAAATGA
      T  S  G   T  I  Y  I   R  K  T   R  K  *
```

**EF2713**

```
   1  TACAATAGGT AACGAAAGGA ATGATTAAAA TGAAAAAAAT GTTTTTAGGA
                                      M   K  K  M   F  L  G
  51 CTCTTTTGTT TTGTTTCGAT TACCACACTA TCAGTTTCTA ATGTATGGGC
     L  F  C  F   V  S  I   T  T  L   S  V  S  N   V  W  A
 101 AGAAGATTCG GCTAATGAAG CTTCTGTGGA AAAGTCAACA GTTACTTCTT
      E  D  S   A  N  E  A   S  V  E   K  S  T   V  T  S  L
 151 TAACTCAAGA AACTTCAGCA ACAACAATTA ATGCTTCAAC TGATTCAACA
       T  Q  E   T  S  A   T  T  I  N   A  S  T   D  S  T
 201 GCACTTACAG CTGAAAGCGA GGAACTTCCA TCGCTTCGTC AAACACTTCT
     A  L  T  A   E  S  E   E  L  P   S  L  R  Q   T  L  L
 251 TAATTACGTA GGTATGTATG GTTTAACAGA AACCTTGATA AATCGTTTAT
      N  Y  V   G  M  Y  G   L  T  E   T  L  I   N  R  L  S
 301 CTGATGACGA ACTGGACTAC GCAAAAAAAG TTTCGTTTCA TTTTGTAAAC
       D  D  E   L  D  Y   A  K  K  V   S  F  H   F  V  N
 351 CAAGACATCA GTGGAACAGC GAGAATGATT ACTAAATTAT ATGGTGAGAA
     Q  D  I  S   G  T  A   R  M  I   T  K  L  Y   G  E  K
 401 ACCTATCCCG GAGGACTCCT ACTCTACAGA TTACTCAACA TTAACGATTG
      P  I  P   E  D  S  Y   S  T  D   Y  S  T   L  T  I  D
 451 ATGACTTAAA AAATTATTTA CCTCAAATTC GACTTTCGTT AATTTATGTA
       D  L  K   N  Y  L   P  Q  I  R   L  S  L   I  Y  V
 501 TATGATTTAA ATAGTGATGT TGTAAATAAT TTAAGCGATC AAACACTCGT
     Y  D  L  N   S  D  V   V  N  N   L  S  D  Q   T  L  V
 551 CGATTTAATT AATCAAGTCA AGGTAGACTA CGCTAACCAA AATTACCCAT
      D  L  I   N  Q  V  K   V  D  Y   A  N  Q   N  Y  P  S
 601 CTGATGTCCG TGGTGATTAT GGTTTAGCAG CAATGGCAGA TAAAATAAAA
       D  V  R   G  D  Y   G  L  A  A   M  A  D   K  I  K
 651 GCAAACGATT ATACGTCAAT TAATCAAAGT GCTGAAAGTG TTTCTTCGGA
     A  N  D  Y   T  S  I   N  Q  S   A  E  S  V   S  S  D
 701 TACTACTAAT ACAGAATCTA CATTACAAAC AACTACGAGC AGTTCGAAGA
      T  T  N   T  E  S  T   L  Q  T   T  T  S   S  S  K  K
 751 AAGCTACTAC ATCAAGTTCA ACCGAGCACA AAAAGGGGAT ATTTCCCAGT
       A  T  T   S  S  S   T  E  H  K   K  G  I   F  P  S
 801 ACTGGCGAAA AAAAATCAGT TTTATTTACT ATCATTGGAA TCATCTTACT
     T  G  E  K   K  S  V   L  F  T   I  I  G  I   I  L  L
 851 ATCTTTAGTT AGTATATTCA TTATAAAAAA TAAAAAGAAA TAG
      S  L  V   S  I  F  I   I  K  N   K  K  K   *
```

**Figure S2. N**ucleotide and amino acid sequences of EF1033 and EF2713 after adjustment of the start codon. Note that the middle part (representing 600 residues) of EF1033 is not shown, as indicated by the dotted horizontal line. The N-terminal "extensions" of the proteins due to adjustment of the start codon are printed in bold face. Both new start codons are preceded by putative ribosome binding sites, which are underlined. Predicted signal peptidase I cleavage sites are indicated by vertical arrows. The putative LPxTG motifs are printed bold face and underlined. In both proteins the LPxTG motif is followed by a hydrophobic (italic) region (identified with the TMHMM Server) and positively charged residues at the C-terminus (double underlined), which is characteristic for functional C-terminal anchors that are attached to the cell wall by sortase.

Table S5. Proteome data of the proteins identified by LC-MS analysis after the different treatments (untreated, trypsin and trypsin beads for one or two hours) from Enterococcus faecalis V583.

| Sample | Gene | Gene product | Accession nr | Cover-age % | No of unique peptides | Peptide sequence | Best proba-bility score | Best Xcorr score | No of +2H spectra | No of +3H spectra | No of +4H spectra | Δ M ppm | Modification |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Untreated 1h | EF2556 | fumarate reductase flavoprotein subunit | gi29377044 | 69.70 | 40 | AIDFYDQK | 38.82 | 2.43 | 1 | | | -0.70 | |
| | | | | | | AIDFYDQKGFVEK | 77.33 | 3.24 | 2 | | | 0.73 | |
| | | | | | | AIDFYDQKGFVEKGETIEELAEK | 300.00 | 5.10 | | 3 | 1 | 1.45 | |
| | | | | | | AKAIDFYDQK | 34.02 | 2.34 | 1 | | | -0.03 | |
| | | | | | | AKAVVVTTGGFGANEK | 87.37 | 2.92 | 2 | 3 | | -0.07 | |
| | | | | | | AKAVVVTTGGFGANEKLITQYKPELK | 127.27 | 4.52 | | 2 | 2 | 0.82 | |
| | | | | | | ATIDTWNQDVNAK | 88.20 | 3.91 | 3 | | | 0.93 | |
| | | | | | | AVVVTTGGFANEK | 84.41 | 3.91 | 2 | | | -0.58 | |
| | | | | | | AVVVTTGGFGANEKLITQYKPELK | 97.62 | 3.58 | | 1 | 2 | 0.34 | |
| | | | | | | DDKQFGRTTGMEADLSTAPYYAIK | 24.42 | 2.70 | | 1 | | -1.71 | |
| | | | | | | DKVSAAINALPEK | 58.45 | 3.54 | 2 | | | -0.25 | |
| | | | | | | DKVSAAINALPEKSAYLVFDQGVR | 156.54 | 6.27 | | 1 | 4 | 0.05 | |
| | | | | | | EAGMNPVILEK | 49.31 | 2.34 | 2 | | | 0.08 | |
| | | | | | | EDGTPIKGLYAAGELTGGLHGQNR | 70.07 | 5.07 | 1 | 4 | 3 | 0.79 | |
| | | | | | | EEKIPLFVDADVTDLVEENGQIDGVK | 92.36 | 6.33 | | 4 | 1 | 2.14 | |
| | | | | | | EEKIPLFVDADVTDLVEENGQIDGVKVK | 124.18 | 3.77 | | 2 | 2 | -0.25 | |
| | | | | | | EIQIHPTVQQSDAFLIGEAVR | 86.11 | 3.51 | 2 | 3 | | 0.57 | |
| | | | | | | EIQIHPTVQQSDAFLIGEAVRGEGAILASQKGER | 1.19 | 3.44 | | | 2 | -0.54 | |
| | | | | | | FVNELDTR | 56.29 | 2.66 | 2 | | | 0.02 | |
| | | | | | | GITLSNLTITGGMSEK | 137.27 | 4.26 | 3 | | | 0.22 | |
| | | | | | | GITLSNLTITGGMSEKR | 103.26 | 2.74 | 1 | | | 2.25 | |
| | | | | | | GLYAAGELTGGLHGQNR | 111.26 | 3.96 | 3 | 2 | | -0.54 | |
| | | | | | | IGGNAIADIIIYGR | 131.29 | 4.77 | 172 | 42 | | 0.17 | |
| | | | | | | IGGNAIADIIIYGRQAGTQSAEFASAQK | 23.44 | 3.20 | | 1 | | -3.29 | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | IGMPADTLKATIDTWNQDVNAK | 131.13 | 5.92 | 4 | 5 | | 1.02 | |
| | | | | | IPLFVDADVTDLVEENGQIDGVKVK | 300.00 | 5.89 | | 3 | 1 | 4.45 | |
| | | | | | LITQYKPELK | 17.59 | 2.26 | 1 | | | 0.20 | |
| | | | | | LITQYKPELKNYVTTNQEGTTGDGIQMIQK | 98.23 | 5.39 | | 3 | 5 | 3.08 | |
| | | | | | MPVAGGNTIK | 67.47 | 2.50 | 2 | | | -0.57 | |
| | | | | | MPVAGGNTIKSSSGMNASQTKFQEK | 48.78 | 3.46 | | 2 | | 5.42 | |
| | | | | | NYVTTNQEGTTGDGIQMIQK | 116.73 | 5.52 | 6 | 11 | | -0.80 | |
| | | | | | QAGTQSAEFASAQK | 13.06 | 2.06 | 1 | | | 0.30 | |
| | | | | | RTHRPADGSAIGGYLVDGLVR | 93.50 | 4.84 | | 5 | 2 | 0.50 | |
| | | | | | SAYLVFDQGVR | 78.38 | 3.29 | 2 | | | 0.48 | |
| | | | | | SAYLVFDQGVRDR | 23.14 | 2.13 | 1 | | | -0.15 | |
| | | | | | TEVLREDGTPIKGLYAAGELTGGLHGQNR | 137.14 | 6.64 | | 2 | 4 | 0.02 | |
| | | | | | THRPADGSAIGGYLVDGLVR | 95.76 | 5.02 | 1 | 4 | 5 | 0.16 | |
| | | | | | TTGMEADLSTAPYYAIK | 105.87 | 3.88 | 6 | | | 0.03 | |
| | | | | | VGGALVDMK | 70.41 | 2.97 | 2 | | | -0.10 | |
| | | | | | VSAAINALPEKSAYLVFDQGVR | 63.41 | 3.78 | | 1 | | 3.09 | |
| EF0123 | hypothetical protein EF0123 | gi29374774 | 14.81 | 5 | GSYGYNVNPVSVATR | 97.82 | 3.46 | 3 | | | 1.23 | |
| | | | | | RPDEIKPNVNYQTHVQNIGWQGVVK | 97.11 | 4.64 | | 1 | 3 | -0.64 | |
| | | | | | THVQEIGWQGYVK | 51.15 | 2.59 | | 1 | | -1.38 | |
| | | | | | VPDINYQTHIQDIGWQGVVK | 116.14 | 4.23 | 1 | 4 | | 0.78 | |
| | | | | | YFGASATDLVITAQSYGR | 144.11 | 6.05 | 2 | | | -0.08 | |
| EF1818 | coccolysin | gi29376362 | 30.59 | 8 | DLVFLAIDKRVnNEGqLFYK | 15.29 | 2.39 | | | 1 | 22.31 | N12(Deamidated) Q16(Deamidated) |
| | | | | | GMPILSVVDEQHPDAYDNAFWDGK | 69.17 | 2.60 | | 1 | | 1.38 | |
| | | | | | IqSVDAIGEEGVK | 6.13 | 2.04 | 1 | | | 4.75 | Q2(Deamidated) |
| | | | | | NSFQVAFNVPVEK | 38.54 | 3.21 | 2 | | | -0.03 | |
| | | | | | QTEGVTVDSDNVIIHLDR | 50.57 | 2.81 | | 1 | | 1.00 | |
| | | | | | TGIRNLQTPSKHGQPETMAQYDDR | 23.86 | 3.59 | | | 2 | 2.36 | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | VQYGDEAASVVSAAFNSAGIGAK | 47.39 | 5.17 | 2 | | | 3.93 | |
| | | | | | YKGTPYYDQGGVHYNSGIINR | 67.45 | 3.58 | | 1 | 1 | -0.92 | |
| EF2864 | hypothetical protein EF2864 | gi29377332 | 11.03 | 2 | LLITQDSGNYPAEEYYYR | 139.96 | 4.23 | 3 | 3 | | -0.03 | |
| | | | | | WGPTGAALLLYK | 74.72 | 3.37 | 2 | | | 0.34 | |
| EF3041 | pheromone binding protein | gi29377499 | 5.26 | 2 | TLKGDFQIAVR | 23.46 | 2.68 | | 1 | | 0.78 | |
| | | | | | WSDGKPVTAADYVYGWQR | 85.32 | 3.73 | 2 | 4 | | 0.74 | |
| EF2860 | YkuD putative, pewptidoglycan binding protein | gi29377328 | 10.55 | 2 | FKNNGSYGWSIDGAK | 31.59 | 2.68 | | 1 | | 0.77 | |
| | | | | | IFDVSYDGmPVIIYGHIYDDAPGEFDKPVDYGEEV | 149.55 | 5.45 | | 4 | | 4.98 | M9(Oxidation) |
| EF0394 | secreted antigen, putative | gi29375030 | 6.46 | 2 | ASLALEQSSAESSK | 103.09 | 3.68 | 2 | | | 1.93 | |
| | | | | | VGFGYSGSTIVGHSA | 63.14 | 3.14 | 3 | | | 1.26 | |
| EF0226 | 50S ribosomal protein L15 | gi29374870 | 8.22 | 1 | LGFEGGQTPLFR | 77.39 | 3.05 | 4 | | | 0.13 | |
| EF0417 | hypothetical protein EF0417 | gi29375051 | 12.50 | 3 | EGEQAYVLVNDFGTIIR | 38.98 | 3.18 | | 1 | | 1.92 | |
| | | | | | NGYHMQATIDLGDLGAIELPKK | 1.52 | 3.04 | | | 1 | -0.53 | |
| | | | | | RANIYNKWNR | 76.70 | 3.27 | | 2 | | -0.48 | |
| EF0201 | elongation factor Tu | gi29374847 | 4.05 | 1 | LLDYAEAGDNIGALLR | 121.69 | 4.42 | 4 | | | 1.15 | |
| EF1264 | sulfatase domain-containing protein | gi29375833 | 7.55 | 2 | FIAVSNHYPYSQFTNDEAGFPIAK | 111.99 | 3.75 | | 2 | | 1.78 | |
| | | | | | FYTNSGLKPVNPEDYDYKNQLQQLEAIEK | 40.69 | 2.89 | | | 1 | 2.24 | |
| EF2746 | dltD protein | gi29377221 | 2.83 | 1 | YVPFFGSSELSR | 52.21 | 2.27 | 3 | | | 0.93 | |
| EF0916 | 50S ribosomal protein L20 | gi29375500 | 10.08 | 1 | EQVMNSYYYAFR | 93.01 | 2.90 | 2 | | | 0.97 | |
| EF0164 | putative lipoprotein | gi29374814 | 14.52 | 2 | IISHVGDLYDEK | 54.38 | 2.94 | | 1 | | -1.14 | |
| | | | | | NKIISHVGDLYDEKYQEK | 64.90 | 3.13 | | | 1 | -4.27 | |
| EF2925 | cold-shock domain-contain protein | gi29377389 | 19.70 | 1 | WFNAEKGFGFISR | 37.57 | 2.77 | 2 | | | 1.34 | |
| EF2144 | putative lipoprotein | gi29376653 | 8.04 | 1 | EANVEIHITPQSADSGLVEIIDY | 116.55 | 3.65 | 2 | | | 4.23 | |
| EF1046 | pyruvate kinase | gi29375625 | 4.79 | 2 | AVVAATAEEAVAK | 69.20 | 3.46 | 1 | | | 1.06 | |
| | | | | | LVQGQGVGEEAIIAK | 76.59 | 3.29 | 1 | | | 1.54 | |
| EF0200 | elongation factor G | gi29374846 | 2.45 | 1 | VYSGVLESGSYVLNASK | 125.70 | 3.98 | 1 | | | 2.79 | |
| EF0211 | 50S ribosomal protein L22 | gi29374856 | 9,57 | 1 | TSHITVVVTEK | 40.34 | 2.20 | | 1 | | -0.26 | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EF1167 | fructose-bisphosphate aldolase | gi29375743 | 9.34 | 1 | KGGYAVGGYNTNNLEWTQAILEAAEAK | 71.77 | 5.02 | | 1 | | 3.27 | |
| | EF1308 | dnak protein | gi29375876 | 4.6 | 1 | ePNKSVNPDEVVAmGAAIqGGVITGDVK | 1 | 2.55 | | 1 | | 8.87 | N-Term(Acetyl) M14(Oxidation) Q19(Deamidated) |
| | EF2398 | 30S ribosomal protein S2 | gi29376895 | 3.83 | 1 | FLGGIADMPR | 21.89 | 2.05 | 1 | | | 0.6 | |
| Untreated 2h | EF2556 | fumarate reductase flavoprotein subunit | gi29377044 | 58.81 | 39 | AIDFYDQK | 29.30 | 2.03 | 1 | | | 0.89 | |
| | | | | | | AIDFYDQKGFVEK | 87.86 | 4.09 | 4 | | | 1.83 | |
| | | | | | | AIDFYDQKGFVEKGETIEELAEK | 144.11 | 4.94 | | 2 | 1 | 4.08 | |
| | | | | | | AKAIDFYDQK | 29.65 | 2.78 | 2 | 2 | | -0.47 | |
| | | | | | | AKAIDFYDQKGFVEK | 88.13 | 3.49 | 1 | 2 | | 0.15 | |
| | | | | | | AKAVVVTTGGFGANEK | 61.40 | 3.80 | 2 | 2 | | 0.11 | |
| | | | | | | AKAVVVTTGGFGANEKLITQYKPELK | 63.38 | 2.93 | | | 1 | 3.26 | |
| | | | | | | ATIDTWNQDVNAKDDKQFGR | 72.20 | 3.59 | | 1 | | 3.94 | |
| | | | | | | AVVVTTGGFGANEK | 98.28 | 4.24 | 5 | | | -0.42 | |
| | | | | | | AVVVTTGGFGANEKLITQYKPELK | 50.42 | 4.39 | | 1 | | 9.02 | |
| | | | | | | DKVSAAINALPEK | 91.57 | 4.08 | 3 | | | 0.48 | |
| | | | | | | EAGmNPVILEK | 41.97 | 2.43 | 3 | | | 0.16 | M4(Oxidation) |
| | | | | | | EAGmNPVILEKmPVAGGNTIK | 43.61 | 3.53 | | 2 | | 3.29 | M4(Oxidation) M12(Oxidation) |
| | | | | | | EDGTPIKGLYAAGELTGGLHGQNR | 49.23 | 3.80 | | 1 | 2 | 1.13 | |
| | | | | | | EDGTPIKGLYAAGELTGGLHGQNRIGGNAIADIIIYGR | 144.36 | 4.51 | | | 2 | 3.71 | |
| | | | | | | EEKIPLFVDADVTDLVEENGQIDGVK | 97.46 | 7.15 | | 3 | | 3.24 | |
| | | | | | | EEKIPLFVDADVTDLVEENGQIDGVKVK | 73.29 | 3.74 | | 3 | 3 | -0.01 | |
| | | | | | | GLYAAGELTGGLHGQNR | 123.87 | 5.02 | 3 | 1 | | 1.11 | |
| | | | | | | GLYAAGELTGGLHGQNRIGGNAIADIIIYGR | 77.32 | 3.86 | | | 1 | 3.13 | |
| | | | | | | IGGNAIADIIIYGR | 120.52 | 4.54 | 172 | 17 | | -0.15 | |
| | | | | | | IGGNAIADIIIYGRQAGTQSAEFASAQK | 300.00 | 6.36 | 1 | 4 | | 3.66 | |
| | | | | | | IGmPADTLKATIDTWNQDVNAK | 95.11 | 4.73 | | 3 | | 2.66 | M3(Oxidation) |
| | | | | | | IPLFVDADVTDLVEENGQIDGVK | 85.00 | 5.85 | 1 | | | 7.03 | |
| | | | | | | IPLFVDADVTDLVEENGQIDGVKVK | 150.51 | 6.08 | | 4 | | 3.72 | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | LITQYKPELK | 40.64 | 2.83 | 2 | | | 0.81 | |
| | | | | | LITQYKPELKNYVTTNQEGTTGDGIQmIQK | 111.11 | 4.38 | | 1 | 7 | 2.55 | M27(Oxidation) |
| | | | | | MPVAGGNTIK | 80.73 | 2.35 | 4 | | | 0.10 | |
| | | | | | MPVAGGNTIKSSSGmNASQTKFQEK | 31.17 | 2.87 | | 5 | | 4.40 | M15(Oxidation) |
| | | | | | NYVTTNQEGTTGDGIQmIQK | 114.31 | 5.15 | 5 | 6 | | 1.66 | M17(Oxidation) |
| | | | | | QAGTQSAEFASAQK | 66.85 | 3.10 | 4 | | | 1.40 | |
| | | | | | RTHRPADGSAIGGYLVDGLVR | 132.73 | 4.57 | | 5 | 2 | -0.05 | |
| | | | | | SAYLVFDQGVR | 76.11 | 3.16 | 4 | | | 0.97 | |
| | | | | | SAYLVFDQGVRDR | 10.15 | 2.23 | 1 | | | 1.44 | |
| | | | | | TEVLREDGTPIKGLYAAGELTGGLHGQNR | 300.00 | 7.67 | 2 | 2 | | 1.74 | |
| | | | | | THRPADGSAIGGYLVDGLVR | 103.39 | 5.39 | 2 | 4 | | 0.18 | |
| | | | | | THRPADGSAIGGYLVDGLVRNVR | 43.41 | 2.17 | | | 1 | -0.15 | |
| | | | | | TTGmEADLSTAPYYAIK | 84.68 | 4.09 | 1 | | | 3.29 | M4(Oxidation) |
| | | | | | VGGALVDmK | 59.43 | 2.82 | 2 | | | -0.63 | M8(Oxidation) |
| | | | | | VSAAINALPEKSAYLVFDQGVR | 110.62 | 4.43 | | 1 | | 3.64 | |
| EF2860 | YkuD putative, pewptidoglycan binding protein | gi29377328 | 19.62 | 6 | FKNNGSYGWSIDGAK | 87.92 | 2.16 | 3 | 2 | | 0.47 | |
| | | | | | IANNYIEIDLKDQK | 47.33 | 2.63 | | 1 | | -0.64 | |
| | | | | | IFDVSYDGMPVIIYGHIYDDAPGEFDKPVDYGEEV | 300.00 | 5.67 | | 7 | 2 | 3.44 | |
| | | | | | RFKNNGSYGWSIDGAK | 72.84 | 3.89 | | 4 | 1 | 0.17 | |
| | | | | | VKLPLNEAFKK | 18.42 | 2.09 | | 1 | | -0.16 | |
| | | | | | YNKGTATVPGFHTILYR | 94.39 | 4.29 | | 2 | | 1.98 | |
| EF2864 | hypothetical protein EF2864 | gi29377332 | 11.03 | 2 | LLITQDSGNYPAEEYYYR | 142.39 | 5.07 | 4 | 3 | | 4.80 | |
| | | | | | WGPTGAALLLYK | 71.24 | 3.59 | 2 | | | 0.71 | |
| EF1818 | coccolysin | gi29376362 | 10.59 | 3 | GmPILSVVDEQHPDAYDNAFWDGK | 79.02 | 3.16 | | 2 | | 6.40 | M2(Oxidation) |
| | | | | | TYYEDHFQR | 48.10 | 2.24 | 1 | | | -2.55 | |
| | | | | | YKGTPYYDQGGVHYNSGIINR | 82.59 | 4.63 | | 2 | 3 | 1.59 | |
| EF3041 | pheromone binding protein | gi29377499 | 6.35 | 3 | LREESKWSDGKPVTAADYVYGWQR | 66.67 | 2.87 | | | 2 | 2.51 | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | TLKGDFQIAVR | 32.15 | 2.99 | | 2 | -0.04 | |
| | | | | | | WSDGKPVTAADYVYGWQR | 92.16 | 3.18 | 2 | | 1.37 | |
| | EF0417 | hypothetical protein EF0417 | gi29375051 | 8.16 | 2 | AIQKEGEQAYVLVNDFGTIIRR | 41.98 | 2.39 | | 1 | 3.68 | |
| | | | | | | RANIYNKWNR | 64.35 | 3.37 | 1 | 4 | -0.39 | |
| | EF0164 | putative lipoprotein | gi29374814 | 25.00 | 3 | IISHVGDLYDEKYQEK | 41.58 | 2.25 | | | 1 -0.40 | |
| | | | | | | SHGNYEVIYK | 44.35 | 2.30 | 1 | | 1.60 | |
| | | | | | | SHGNYEVIYKSGKFK | 22.09 | 2.64 | | 1 | 2 -0.23 | |
| | EF2144 | putative lipoprotein | gi29376653 | 12.59 | 2 | EANVEIHITPQSADSGLVEIIDY | 128.03 | 4.66 | 4 | | 3.99 | |
| | | | | | | HQFNDmYPYKGSK | 19.55 | 2.88 | | 1 | -0.66 | M6(Oxidation) |
| | EF0123 | hypothetical protein EF0123 | gi29374774 | 5.75 | 2 | RPDEIKPNVNYQTHVQNIGWQGVVK | 109.73 | 4.03 | | | 2 2.29 | |
| | | | | | | VPDINYQTHIQDIGWQGVVK | 87.05 | 3.78 | | 2 | -1.78 | |
| | EF1167 | fructose-bisphosphate aldolase | gi29375743 | 9.69 | 2 | GGYAVGGYNTNNLEWTQAILEAAEAKK | 1.00 | 2.27 | | 1 | 3.82 | |
| | | | | | | KGGYAVGGYNTNNLEWTQAILEAAEAK | 119.91 | 6.13 | | 2 | 3.82 | |
| | EF2925 | cold-shock domain-contain protein | gi29377389 | 30.30 | 2 | mEqGTVKWFnAEKGFGFISR | 1.00 | 3.29 | | 1 | 14.54 | N-Term(Acetyl) Q3(Deamidated) N10(Deamidated) |
| | | | | | | WFNAEKGFGFISR | 70.24 | 3.72 | 2 | | 2.81 | |
| | EF0226 | 50S ribosomal protein L15 | gi29374870 | 12.33 | | LGFEGGQTPLFR | 64.05 | 2.59 | 1 | | 1.47 | |
| | | | | | | SGGGVRLGFEGGQTPLFR | 75.74 | 3.76 | | 1 | -0.60 | |
| | EF0201 | elongation factor Tu | gi29374847 | 7.59 | 2 | FKAEVYVLSK | 29.06 | 2.76 | | 1 | -1.19 | |
| | | | | | | LLDYAEAGDNIGALLRGVAR | 12.52 | 2.17 | | 1 | 4.45 | |
| | EF2746 | dltD protein | gi29377221 | 6.37 | 2 | AINNNKFEISNGFYR | 37.52 | 2.23 | | 1 | 0.25 | |
| | | | | | | YVPFFGSSELSR | 18.08 | 2.30 | 1 | | 0.81 | |
| | EF0211 | 50S ribosomal protein L22 | gi29374856 | 17.39 | 2 | GSASPINKRTSHITVVVTEK | 38.27 | 2.21 | | | 1 -1.55 | |
| | | | | | | TSHITVVVTEK | 25.47 | 2.65 | 1 | | 0.28 | |
| | EF0916 | 50S ribosomal protein L20 | gi29375500 | 10.08 | 1 | EQVmNSYYYAFR | 77.98 | 2.84 | 1 | | 3.37 | M4(Oxidation) |
| Trypsin 1h | EF2556 | fumarate reductase flavoprotein subunit | gi29377044 | 67.72 | 40 | AIDFYDQK | 60.47 | 2.40 | 3 | | -0.15 | |
| | | | | | | AIDFYDQKGFVEKGETIEELAEK | 125.70 | 4.76 | | 2 | 3 1.94 | |
| | | | | | | AKAIDFYDQK | 31.09 | 2.05 | 1 | | -0.28 | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AKAIDFYDQKGFVEK | 1.00 | 2.45 | | | 1 | -1.26 | |
| AKAVVVTTGGFGANEK | 63.46 | 3.64 | 3 | 3 | | -0.99 | |
| ATIDTWNQDVNAK | 84.42 | 3.14 | 4 | | | 1.23 | |
| ATIDTWNQDVNAKDDKQFGR | 28.01 | 2.32 | | | 1 | -0.53 | |
| AVVVTTGGFGANEK | 78.16 | 3.74 | 8 | 1 | | -0.03 | |
| DDKQFGRTTGMEADLSTAPYYAIK | 95.87 | 3.88 | | 2 | | 1.77 | |
| DKVSAAINALPEK | 96.69 | 3.92 | 2 | | | 0.24 | |
| DKVSAAINALPEKSAYLVFDQGVR | 300.00 | 6.92 | 2 | 3 | 3 | 0.66 | |
| EAGMNPVILEK | 35.82 | 2.29 | 1 | | | 0.94 | |
| EDGTPIKGLYAAGELTGGLHGQNR | 137.14 | 4.97 | 2 | 5 | 3 | 0.72 | |
| EEKIPLFVDADVTDLVEENGQIDGVK | 87.53 | 6.13 | | 4 | 1 | 1.04 | |
| EEKIPLFVDADVTDLVEENGQIDGVKVK | 93.37 | 3.98 | | 2 | 1 | 0.23 | |
| EGIKDSNDKFFEETLK | 40.49 | 2.82 | | 1 | | 2.65 | |
| EIQIHPTVQQSDAFLIGEAVR | 83.00 | 4.42 | 4 | 3 | | 0.02 | |
| EIQIHPTVQQSDAFLIGEAVRGEGAILASQKGER | 3.91 | 2.79 | | | 1 | -14.58 | |
| FVNELDTR | 32.99 | 2.68 | 2 | | | -0.72 | |
| GEGAILASQK | 43.78 | 2.41 | 1 | | | -0.23 | |
| GFVEKGETIEELAEK | 80.34 | 4.30 | 2 | 2 | | -0.23 | |
| GITLSNLTITGGMSEK | 121.92 | 4.04 | 4 | | | 1.03 | |
| GITLSNLTITGGMSEKR | 97.43 | 2.82 | 2 | 2 | | -0.27 | |
| GLYAAGELTGGLHGQNR | 124.90 | 4.52 | 5 | 3 | | -0.03 | |
| IGGNAIADIIIYGR | 121.17 | 4.56 | 188 | 63 | | 0.04 | |
| IGmPADTLKATIDTWNQDVNAK | 300.00 | 4.74 | | 6 | | 1.75 | M3(Oxidation) |
| IPLFVDADVTDLVEENGQIDGVK | 106.12 | 5.12 | 4 | 9 | | 2.99 | |
| IPLFVDADVTDLVEENGQIDGVKVK | 138.68 | 5.29 | | 5 | 2 | -0.75 | |
| LITQYKPELKNYVTTNQEGTTGDGIQmIQK | 52.69 | 3.39 | | | 3 | -0.38 | M27(Oxidation) |
| NYVTTNQEGTTGDGIQMIQK | 140.62 | 4.96 | 8 | 7 | | -0.17 | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | RTHRPADGSAIGGYLVDGLVR | 121.96 | 4.52 | | 5 | 5 | -0.02 | |
| | | | | | SAYLVFDQGVR | 73.16 | 3.35 | 5 | | | 0.24 | |
| | | | | | SAYLVFDQGVRDR | 24.70 | 2.20 | 1 | | | 0.58 | |
| | | | | | TEVLREDGTPIKGLYAAGELTGGLHGQNR | 300.00 | 6.80 | | 2 | 4 | -0.09 | |
| | | | | | THRPADGSAIGGYLVDGLVR | 107.35 | 5.42 | 3 | 5 | 5 | -0.01 | |
| | | | | | TTGMEADLSTAPYYAIK | 99.49 | 4.06 | 8 | 2 | | 1.13 | |
| | | | | | TTGMEADLSTAPYYAIKIAPGIHHTMGGVK | 45.60 | 3.02 | | 1 | 1 | 3.07 | |
| | | | | | VGGALVDmK | 71.69 | 3.10 | 2 | | | 3.10 | M8(Oxidation) |
| | | | | | VSAAINALPEK | 45.52 | 2.73 | 1 | | | -0.47 | |
| | | | | | VSAAINALPEKSAYLVFDQGVR | 71.51 | 4.69 | | 2 | | 1.63 | |
| EF0123 | hypothetical protein EF0123 | gi29374774 | 21.07 | 9 | AFLVGDEAR | 37.11 | 2.23 | 1 | | | -1.60 | |
| | | | | | GSYGYNVNPVSVATR | 97.22 | 3.69 | 4 | | | 1.23 | |
| | | | | | LTGEIANAYDVYYR | 112.42 | 3.42 | 3 | | | 1.55 | |
| | | | | | NGYTLTYDPYGR | 56.05 | 2.64 | 3 | | | 0.89 | |
| | | | | | RPDEIKPNVNYQTHVQNIGWQGVVK | 125.36 | 5.12 | | 1 | 4 | -0.26 | |
| | | | | | THVQEIGWQGYVK | 55.00 | 2.94 | 2 | 2 | | -0.64 | |
| | | | | | VPDINYQTHIQDIGWQGVVK | 64.37 | 3.73 | | 2 | | 0.41 | |
| | | | | | YFGASATDLVITAQSYGR | 141.16 | 4.57 | 4 | 1 | | 2.00 | |
| | | | | | YYDVYYR | 41.92 | 2.38 | 1 | | | 0.45 | |
| EF0201 | elongation factor Tu | gi29374847 | 40.00 | 11 | DLLSEYDFPGDDVPVIAGSALK | 148.41 | 4.52 | 1 | | | 5.11 | |
| | | | | | FKAEVYVLSK | 52.32 | 2.10 | 1 | 2 | | 0.19 | |
| | | | | | GITINTSHIEYETETR | 101.30 | 4.12 | 1 | 1 | | 1.66 | |
| | | | | | GQVLAKPATITPHTKFKAEVYVLSK | 14.05 | 2.89 | | | 1 | 1.69 | |
| | | | | | HYAHVDCPGHADYVK | 37.86 | 2.72 | | | 1 | -1.39 | |
| | | | | | KLLDYAEAGDNIGALLR | 83.99 | 4.47 | 2 | 2 | | 0.90 | |
| | | | | | LLDYAEAGDNIGALLR | 121.69 | 4.97 | 3 | 1 | | 1.39 | |
| | | | | | NmITGAAQmDGAILVVSAADGPmPQTR | 17.58 | 3.08 | | 1 | | 1.90 | M2(Oxidation) M9(Oxidation) M23(Oxidation) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | TTLTAAIATVLSK | 80.00 | 3.12 | 1 | | | 0.07 | |
| | | | | | TVGSGVVTEIVK | 20.36 | 2.30 | 1 | | | 0.15 | |
| | | | | | VGDEVEIVGIK | 56.84 | 2.36 | 1 | | | -0.87 | |
| EF1818 | coccolysin | gi29376362 | 26.47 | 8 | GmPILSVVDEQHPDAYDNAFWDGK | 127.13 | 3.26 | | 3 | | 2.84 | M2(Oxidation) |
| | | | | | IGYTIIQNLGIEK | 90.74 | 4.22 | 2 | | | 0.50 | |
| | | | | | IqSVDAIGEEGVK | 10.76 | 2.00 | 1 | | | 4.02 | Q2(Deamidated) |
| | | | | | NSFQVAFNVPVEK | 37.13 | 2.95 | 1 | | | 1.44 | |
| | | | | | QTEGVTVDSDNVIIHLDR | 56.01 | 3.26 | | 1 | | 0.82 | |
| | | | | | VNNEGQLFYK | 41.97 | 2.23 | 1 | | | 1.19 | |
| | | | | | VQYGDEAASVVSAAFNSAGIGAK | 113.17 | 5.30 | 3 | | | 3.44 | |
| | | | | | YKGTPYYDQGGVHYNSGIINR | 63.60 | 3.76 | | 2 | 3 | 0.06 | |
| EF1264 | sulfatase domain-containing protein | gi29375833 | 21.37 | 10 | ADWNDFDNANmQR | 99.94 | 3.05 | 1 | | | 2.69 | M11(Oxidation) |
| | | | | | FIAVSNHYPYSQFTNDEAGFPIAK | 121.35 | 3.90 | 1 | 2 | | 4.53 | |
| | | | | | FYTNSGLKPVNPEDYDYK | 80.12 | 3.50 | | 1 | | 1.82 | |
| | | | | | FYTNSGLKPVNPEDYDYKNQLQQLEAIEK | 39.58 | 3.92 | | | 1 | 6.39 | |
| | | | | | nqLqqLEAIEKEK | 1.00 | 2.14 | 1 | | | 13.80 | N-Term(Acetyl)<br>Q2(Deamidated)<br>Q4(Deamidated)<br>Q5(Deamidated) |
| | | | | | NYIQLGQDLFSK | 59.75 | 3.20 | 2 | | | 0.17 | |
| | | | | | QLETSDQITNGDLLR | 103.92 | 2.93 | 2 | | | 2.52 | |
| | | | | | STFSFDNFFHQVGQGK | 93.34 | 3.96 | 2 | | | 3.52 | |
| | | | | | TSDAETLLENSLFGLDQGSLFTQVGGK | 72.87 | 4.80 | | 1 | | 7.58 | |
| | | | | | YTILGSSIYDTK | 33.85 | 3.05 | 1 | | | 0.50 | |
| EF2860 | YkuD putative, pewptidoglycan binding protein | gi29377328 | 20.04 | 5 | FKNNGSYGWSIDGAK | 17.96 | 2.91 | | 1 | | -0.33 | |
| | | | | | IFDVSYDGMPVIIYGHIYDDAPGEFDKPVDYGEEV | 300.00 | 5.06 | | 4 | | 3.15 | |
| | | | | | RFKNNGSYGWSIDGAK | 25.00 | 2.69 | | 1 | | -0.56 | |
| | | | | | TQELLVnALnSqEqTNAITAPLVGDTK | 1.00 | 2.58 | | | 4 | 18.27 | |
| | | | | | YNKGTATVPGFHTILYR | 73.92 | 4.18 | | 2 | 1 | -0.03 | |
| EF2864 | hypothetical protein EF2864 | gi29377332 | 20.22 | 3 | LLITQDSGNYPAEEYYYR | 137.27 | 4.59 | 4 | 4 | | 0.70 | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | TLTDAENNDTNLGFLGmNGNDFFFR | 46.45 | 3.87 | | 1 | 4.13 | M17(Oxidation) |
| | | | | | WGPTGAALLLYK | 86.58 | 3.60 | 2 | | -0.27 | |
| EF0200 | elongation factor G | gi29374846 | 16.88 | 7 | GQYGHVWVEFTPNEEGK | 11.31 | 2.64 | | 1 | -2.26 | |
| | | | | | IGADFFYSVESLHDR | 50.47 | 3.00 | | 1 | -1.55 | |
| | | | | | IGETHEGASQMDWMEQEQER | 50.23 | 4.02 | | 2 | -2.14 | |
| | | | | | LYDGSYHDVDSNETAFR | 35.19 | 2.58 | | 1 | 2.89 | |
| | | | | | VNIIDTPGHVDFTIEVQR | 114.69 | 3.53 | | 2 | -0.96 | |
| | | | | | VYSGDIAAAVGLK | 35.59 | 2.50 | 1 | | 1.95 | |
| | | | | | VYSGVLESGSYVLNASK | 101.96 | 4.18 | 1 | | 3.52 | |
| EF3041 | pheromone binding protein | gi29377499 | 5.26 | 2 | TLKGDFQIAVR | 39.99 | 2.95 | | 3 | -0.41 | |
| | | | | | WSDGKPVTAADYVYGWQR | 117.87 | 3.88 | 2 | 1 | 0.88 | |
| EF0417 | hypothetical protein EF0417 | gi29375051 | 15.31 | 4 | EGEQAYVLVNDFGTIIR | 92.53 | 3.89 | 1 | 1 | -1.56 | |
| | | | | | MAIESGLESADR | 39.17 | 2.46 | 1 | | 1.48 | |
| | | | | | NGYHMQATIDLGDLGAIELPK | 19.53 | 3.04 | | 1 | -2.47 | |
| | | | | | RANIYNKWNR | 51.12 | 3.17 | | 2 | -0.67 | |
| EF1964 | glyceraldehyde-3-phosphate dehydrogenase | gi29376486 | 15.92 | 3 | IQDVEGIEVVAINDLTDAK | 300.00 | 5.87 | 2 | | 2.37 | |
| | | | | | TVAWYDNEmSYTAQLVR | 114.74 | 3.98 | 2 | | 5.10 | M9(Oxidation) |
| | | | | | VPVATGSLTELVTVLDK | 145.40 | 5.33 | 2 | | 1.34 | |
| EF0177 | hypothetical protein EF0177 | gi29374827 | 21.05 | 5 | ALAAAmYQNGVDIIFHASGATGQGVFQEAK | 97.36 | 6.35 | | 2 | 1.21 | M6(Oxidation) |
| | | | | | FNTIFGIGYLLK | 43.30 | 2.56 | 1 | | 1.79 | |
| | | | | | GVGTAVQDIANR | 37.08 | 2.28 | 1 | | 3.09 | |
| | | | | | VGFVGGEEGVVIDR | 55.06 | 2.86 | 1 | | 0.66 | |
| | | | | | VWVIGVDR | 49.32 | 2.07 | 1 | | -0.40 | |
| EF0228 | adenylate kinase | gi29374872 | 15.28 | 2 | IIDTYGIPHISTGDMFR | 46.21 | 3.14 | | 2 | -1.02 | |
| | | | | | LAVNIESSAPILAFYK | 124.95 | 4.52 | 3 | | 1.19 | |
| EF1961 | enolase | gi29376483 | 8.80 | 2 | AAADYLEVPLYHYLGGFNTK | 109.75 | 3.80 | 1 | 3 | 0.97 | |
| | | | | | GNPTIEVEVYTESGAFGR | 43.45 | 2.74 | 1 | | 1.52 | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EF0176 | hypothetical protein EF0176 | gi29374826 | 14.85 | 3 | ALASSmYQAGADIIYHAAATTGQGIFQEAK | 109.32 | 5.46 | | 1 | -4.22 | M6(Oxidation) |
| | | | | | SFNQSAWEGMQEWGK | 106.43 | 3.41 | 3 | | 3.42 | |
| | | | | | VWVIGVDR | 49.32 | 2.07 | 1 | | -0.40 | |
| EF2718 | 50S ribosomal protein L1 | gi29377196 | 14.41 | 2 | LVENFNTINDVLLK | 83.34 | 2.48 | 1 | | 0.92 | |
| | | | | | NISVTTTFGPGIHVDQASF | 104.17 | 3.84 | 3 | | 1.49 | |
| EF0205 | 30S ribosomal protein S10 | gi29374850 | 12.75 | 2 | LDLPSGVNIEIK | 91.86 | 2.79 | 3 | | 0.36 | |
| | | | | | LDLPSGVNIEIKL | 34.62 | 2.51 | 1 | | 1.49 | |
| EF0226 | 50S ribosomal protein L15 | gi29374870 | 17.12 | 2 | AGIKVLADGELTK | 16.48 | 2.72 | | 1 | 0.06 | |
| | | | | | LGFEGGQTPLFR | 79.26 | 3.21 | 3 | | 0.98 | |
| EF0164 | putative lipoprotein | gi29374814 | 17.74 | 2 | IISHVGDLYDEK | 40.27 | 2.78 | | 2 | 0.69 | |
| | | | | | SHGNYEVIYK | 28.94 | 2.41 | 1 | 1 | -0.36 | |
| EF_B0004 | TraC protein | gi29377898 | 6.43 | 2 | ELSATTmFLEVNqRnK | 18.45 | 2.51 | | 2 | -18.71 | M7(Oxidation) Q13(Deamidated) N15(Deamidated) |
| | | | | | WADGTDITADDFVTAWQR | 113.41 | 4.79 | 2 | | 2.65 | |
| EF1167 | fructose-bisphosphate aldolase | gi29375743 | 13.84 | 2 | KGGYAVGGYNTNNLEWTQAILEAAEAK | 42.15 | 2.29 | | 1 | 3.27 | |
| | | | | | VNVNTEFQLSFAK | 85.79 | 3.34 | 3 | | -1.07 | |
| EF0970 | 50S ribosomal protein L27 | gi29375553 | 34.74 | 2 | IYPGVNVGIGGDDTLFAK | 83.77 | 2.51 | 2 | | 2.71 | |
| | | | | | SADGQTVTGGSILYR | 22.76 | 2.66 | 1 | | 1.08 | |
| EF0916 | 50S ribosomal protein L20 | gi29375500 | 27.73 | 2 | EQVMNSYYYAFR | 81.48 | 3.31 | 2 | | -0.50 | |
| | | | | | MLADLAVNDAAAFTALAEQAK | 148.41 | 3.90 | 1 | | 3.42 | |
| EF1319 | hypothetical protein EF1319 | gi29375887 | 7.76 | 1 | NLTTLYQNPVNTPYMTK | 86.68 | 4.34 | 3 | | -1.31 | |
| EF2144 | putative lipoprotein | gi29376653 | 8.04 | 1 | EANVEIHITPQSADSGLVEIIDY | 107.61 | 4.13 | 2 | 1 | 3.42 | |
| EF1420 | hypothetical protein EF1420 | gi29375987 | 10.40 | 2 | SMSSNLGTLIK | 28.62 | 2.10 | 1 | | -0.24 | |
| | | | | | TSAEIQLGISKYGIQ | 72.52 | 3.49 | 2 | | 1.19 | |
| EF1613 | formate acetyltransferase | gi29376172 | 4.81 | 3 | DmQFFGAR | 36.20 | 2.42 | 1 | | 0.75 | M2(Oxidation) |
| | | | | | FRDLTPEQQADVISR | 44.50 | 2.32 | | 1 | -1.30 | |
| | | | | | IALYGIDYLMEQK | 100.36 | 2.79 | 1 | | 1.30 | |
| EF1308 | dnak protein | gi29375876 | 7.22 | 2 | ePNKSVNPDEVVAmGAAIqGGVITGDVK | 1.00 | 2.20 | | 1 | 25.32 | N-Term(Acetyl) M14(Oxidation) Q19(Deamidated) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | SYTPQEVSAmILQYLK | 98.70 | 3.64 | 1 | | | 2.63 | M10(Oxidation) |
| EF0206 | 50S ribosomal protein L3 | gi29374851 | 16.75 | 2 | VDVFQAGDVVDVTGTTK | 27.65 | 2.05 | 1 | | | 1.52 | |
| | | | | | YHRRPGSMGPVAPNRVFK | 25.26 | 2.70 | | | 1 | -1.40 | |
| EF3256 | pheromone cAD1 precursor lipoprotein | gi29377699 | 3.88 | 1 | NGYRAVFEMTVK | 21.53 | 2.24 | 1 | 1 | | -0.25 | |
| EF0234 | 50S ribosomal protein L17 | gi29374878 | 14.96 | 2 | DITTDLIINER | 59.17 | 2.98 | 1 | | | 0.08 | |
| | | | | | LFNDLGPR | 38.59 | 2.18 | 1 | | | -0.39 | |
| EF2746 | dltD protein | gi29377221 | 4.95 | 2 | FEISNGFYR | 36.37 | 2.09 | 1 | | | -0.18 | |
| | | | | | YVPFFGSSELSR | 70.23 | 2.32 | 1 | | | 1.18 | |
| EF0195 | phosphoglycerate mutase 1 | gi29374841 | 10.53 | 2 | ALPFWQDEIAPALK | 50.54 | 2.25 | 1 | | | 1.27 | |
| | | | | | YGDEQVHIWR | 51.67 | 2.87 | | 1 | | -0.53 | |
| EF0071 | putative lipoprotein | gi29374726 | 8.47 | 3 | ELLGGFAGPLIIAEEYPVNLAASLNK | 6.40 | 3.10 | | | | 7.41 | |
| | | | | | GKGTEGWLPLWAK | 9.07 | 2.26 | | | | 0.08 | |
| | | | | | TNGVYDTNYFNNFSDLGAWHGYYLPEK | 300.00 | 4.77 | | | | 5.81 | |
| EF1523 | hypothetical protein EF1523 | gi29376088 | 2.03 | 1 | GYVVPGGYSLEPAK | 38.69 | 2.57 | 2 | | | 0.05 | |
| EF0304 | putative lipoprotein | gi29374943 | 8.52 | 1 | TAIYGIQLNVEEVAK | 71.77 | 3.55 | 2 | | | 2.35 | |
| EF0223 | 50S ribosomal protein L18 | gi29374867 | 19.49 | 1 | NIYAQVIDDVAGVTLASASALDK | 92.98 | 5.76 | | 2 | | 1.15 | |
| EF1046 | pyruvate kinase | gi29375625 | 4.79 | 2 | AVVAATAEEAVAK | 69.20 | 2.83 | 1 | | | 1.43 | |
| | | | | | LVQGQGVGEEAIIAK | 85.64 | 3.25 | 1 | | | 1.29 | |
| EF1033 | lipoamidase | gi29375612 | 4.66 | 2 | EGLPLGIQFNSALNEDR | 34.25 | 2.2 | 1 | | | 1.89 | |
| | | | | | ALQDTGQPFLGVPLLLK | 47.61 | 3.24 | 2 | | | 2.3 | |
| EF2398 | 30S ribosomal protein S2 | gi29376895 | 11.49 | 2 | INAmEEDGTFEVLPK | 38.67 | 3.84 | 1 | | | 2.01 | M4(Oxidation) |
| | | | | | WLGGTLTNWDTIQKR | 55.36 | 2.77 | 1 | | | 4.55 | |
| EF1898 | 50S ribosomal protein L19 | gi29376426 | 23.48 | 1 | mNPLIQELTQEQLRTDIPAFRPGDTVR | 1.00 | 2.54 | | 1 | | 28.83 | N-Term(Acetyl) |
| EF0207 | 50S ribosomal protein L4 | gi29374852 | 5.31 | 1 | GGGVVFGPTPR | 9.66 | 2.07 | 1 | | | 0.20 | |
| EF0991 | penicillin-binding protein C | gi29375573 | 2.56 | 1 | LILITNGTNYmPDLTGWSK | 103.37 | 4.10 | 1 | | | -8.82 | M11(Oxidation) |
| EF0221 | 50S ribosomal protein L6 | gi29374866 | 8.99 | 1 | ANFNNmVVGVSEGFQK | 67.81 | 2.85 | 1 | | | 0.06 | M6(Oxidation) |
| EF0218 | 50S ribosomal protein L5 | gi29374863 | 6.70 | 1 | ELLAQLGmPFQK | 32.78 | 2.83 | 1 | | | -0.18 | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | WADGTDITADDFVTAWQR | 138.54 | 5.00 | 3 | | | 2.40 | 2.40 |
| Trypsin 2h | EF2556 | fumarate reductase flavoprotein subunit | gi29377044 | 72.87 | 41 | AIDFYDQK | 39.00 | 2.76 | 3 | | | 0.04 | |
| | | | | | | AIDFYDQKGFVEK | 80.48 | 3.65 | 1 | | | 2.19 | |
| | | | | | | AIDFYDQKGFVEKGETIEELAEK | 92.12 | 3.78 | | 2 | 2 | 1.16 | |
| | | | | | | AKAIDFYDQK | 37.63 | 2.64 | 1 | 2 | | -0.57 | |
| | | | | | | AKAVVVTTGGFGANEK | 99.95 | 3.56 | 3 | 2 | | -0.07 | |
| | | | | | | ATIDTWNQDVNAK | 110.81 | 3.31 | 4 | | | 0.87 | |
| | | | | | | ATIDTWNQDVNAKDDKQFGR | 30.89 | 2.20 | | 2 | | 0.28 | |
| | | | | | | AVVVTTGGFGANEK | 112.21 | 3.77 | 4 | | | -0.42 | |
| | | | | | | DKVSAAINALPEK | 89.94 | 4.15 | 4 | 2 | | 0.07 | |
| | | | | | | DKVSAAINALPEKSAYLVFDQGVR | 148.41 | 6.36 | 1 | 3 | 1 | 2.02 | |
| | | | | | | DSNDKFFEETLK | 39.12 | 2.76 | 2 | 2 | | -0.09 | |
| | | | | | | EDGTPIKGLYAAGELTGGLHGQNR | 68.21 | 5.40 | 1 | 3 | 3 | 1.63 | |
| | | | | | | EEKIPLFVDADVTDLVEENGQIDGVK | 108.30 | 6.92 | | 4 | | 5.44 | |
| | | | | | | EEKIPLFVDADVTDLVEENGQIDGVKVK | 39.50 | 2.20 | | 1 | | 7.27 | |
| | | | | | | EGIKDSNDKFFEETLK | 53.63 | 2.29 | 1 | | | 2.79 | |
| | | | | | | EIQIHPTVQQSDAFLIGEAVR | 107.72 | 4.38 | 4 | 6 | 1 | 2.21 | |
| | | | | | | FVNELDTR | 52.04 | 2.78 | 4 | | | 0.14 | |
| | | | | | | GEGAILASQK | 34.17 | 2.49 | 2 | | | 0.02 | |
| | | | | | | GFVEKGETIEELAEK | 106.16 | 4.29 | 3 | 1 | | -0.13 | |
| | | | | | | GITLSNLTITGGmSEK | 91.69 | 4.56 | 4 | | | 2.01 | M13(Oxidation) |
| | | | | | | GITLSNLTITGGMSEKR | 98.11 | 2.16 | 1 | | | 5.79 | |
| | | | | | | GLYAAGELTGGLHGQNR | 124.90 | 4.62 | 7 | 5 | | -0.17 | |
| | | | | | | IGGNAIADIIIYGR | 127.37 | 4.68 | 159 | 41 | | 0.03 | |
| | | | | | | IGGNAIADIIIYGRQAGTQSAEFASAQK | 148.41 | 5.10 | | 1 | | 5.86 | |
| | | | | | | IGMPADTLKATIDTWNQDVNAK | 91.78 | 4.02 | | 4 | | 2.58 | |
| | | | | | | IPLFVDADVTDLVEENGQIDGVK | 130.52 | 5.77 | 4 | 15 | | -1.04 | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | IPLFVDADVTDLVEENGQIDGVKVK | 300.00 | 5.76 | | 2 | | 5.19 | |
| | | | | | LITQYKPELK | 49.79 | 2.72 | 2 | | | -0.28 | |
| | | | | | MPVAGGNTIK | 65.67 | 2.47 | 3 | | | -0.20 | |
| | | | | | NYVTTNQEGTTGDGIQMIQK | 137.27 | 5.62 | 7 | 7 | | 0.93 | |
| | | | | | RTHRPADGSAIGGYLVDGLVR | 102.07 | 4.81 | | 13 | 6 | 0.44 | |
| | | | | | SAYLVFDQGVR | 75.16 | 3.37 | 11 | 1 | | -0.13 | |
| | | | | | SAYLVFDQGVRDR | 12.03 | 2.36 | 1 | | | -5.03 | |
| | | | | | TEVLREDGTPIKGLYAAGELTGGLHGQNR | 127.96 | 5.76 | | 1 | 2 | 2.57 | |
| | | | | | THRPADGSAIGGYLVDGLVR | 120.40 | 5.56 | 3 | 12 | 2 | 0.41 | |
| | | | | | TTGMEADLSTAPYYAIK | 110.26 | 4.19 | 7 | 1 | | 1.74 | |
| | | | | | TTGMEADLSTAPYYAIKIAPGIHHTmGGVK | 10.89 | 2.24 | | | 1 | 2.47 | M26(Oxidation) |
| | | | | | VGGALVDmK | 58.20 | 2.55 | 2 | | | -0.09 | M8(Oxidation) |
| | | | | | VSAAINALPEK | 45.58 | 2.17 | | 1 | | 1.92 | |
| | | | | | VSAAINALPEKSAYLVFDQGVR | 109.94 | 3.75 | | | 1 | 4.01 | |
| | | | | | YFVDHSAEAIDWLDTK | 106.62 | 3.39 | 3 | | | 2.96 | |
| EF1264 | sulfatase domain-containing protein | gi29375833 | 26.07 | 11 | AVEEFFNYLK | 45.29 | 2.50 | 2 | | | 0.81 | |
| | | | | | FIAVSNHYPYSQFTNDEAGFPIAK | 98.22 | 4.59 | 1 | 4 | | 4.53 | |
| | | | | | FYTNSGLKPVNPEDYDYK | 16.15 | 3.13 | | 2 | | -0.74 | |
| | | | | | NYIQLGQDLFSK | 39.21 | 2.74 | 2 | | | 0.66 | |
| | | | | | QLETSDQITNGDLLR | 65.37 | 2.43 | 1 | | | 4.83 | |
| | | | | | QTQGYTSAAFHGNAGNFWNR | 61.20 | 2.97 | 1 | 2 | | 3.89 | |
| | | | | | QTQGYTSAAFHGNAGNFWNRNETYKR | 16.89 | 2.10 | | | 1 | 2.67 | |
| | | | | | SSGLYENSVIVLYGDHYGVSNSR | 69.88 | 3.85 | | 2 | | 6.25 | |
| | | | | | STFSFDNFFHQVGQGK | 110.22 | 4.53 | 2 | 1 | | 1.00 | |
| | | | | | TSDAETLLENSLFGLDQGSLFTQVGGK | 56.18 | 4.51 | | 1 | | 0.99 | |
| | | | | | YTILGSSIYDTK | 43.07 | 3.51 | 1 | | | -11.22 | |
| EF0201 | elongation factor Tu | gi29374847 | 29.37 | 8 | DLLSEYDFPGDDVPVIAGSALK | 139.38 | 4.94 | 3 | | | 4.86 | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | FKAEVYVLSK | 81.09 | 3.30 | 1 | 2 | -0.55 | |
| | | | | | GITINTSHIEYETETR | 33.46 | 2.86 | | 1 | 2.02 | |
| | | | | | KLLDYAEAGDNIGALLR | 97.35 | 4.63 | 4 | 1 | 1.74 | |
| | | | | | LLDYAEAGDNIGALLR | 128.50 | 4.44 | 5 | | 1.03 | |
| | | | | | NmITGAAQmDGAILVVSAADGPmPQTR | 53.56 | 3.98 | | 1 | -20.44 | M2(Oxidation) M9(Oxidation) M23(Oxidation) |
| | | | | | NVGVPYIVVFLNK | 61.78 | 2.59 | 2 | | 2.45 | |
| | | | | | VGDEVEIVGIK | 74.12 | 2.62 | 3 | | 0.35 | |
| EF2860 | YkuD putative, pewptidoglycan binding protein | gi29377328 | 23.21 | 9 | FKNNGSYGWSIDGAK | 71.77 | 3.19 | 1 | 1 | 1.68 | |
| | | | | | FLKDHLHSSK | 36.12 | 2.44 | | 2 | -0.37 | |
| | | | | | GTATVPGFHTILYR | 19.28 | 2.09 | 1 | | 4.18 | |
| | | | | | IANNYIEIDLK | 69.96 | 2.81 | 1 | | 0.55 | |
| | | | | | IFDVSYDGMPVIIYGHIYDDAPGEFDKPVDYGEEV | 300.00 | 5.49 | | 5 | 2 4.25 | |
| | | | | | NNGSYGWSIDGAK | 67.75 | 2.68 | 2 | | 0.17 | |
| | | | | | RFKNNGSYGWSIDGAK | 38.35 | 3.05 | | 2 | -1.11 | |
| | | | | | RGNGTFEIVPEEQGTVVDTQR | 74.68 | 3.85 | | 1 | 1.38 | |
| | | | | | YNKGTATVPGFHTILYR | 73.19 | 2.79 | | 1 | 2.71 | |
| EF0123 | hypothetical protein EF0123 | gi29374774 | 14.43 | 7 | AFLVGDEAR | 55.44 | 2.16 | 1 | | -0.45 | |
| | | | | | GSYGYNVNPVSVATR | 109.12 | 3.76 | 4 | | 1.23 | |
| | | | | | LTGEIANAYDVYYR | 93.89 | 3.46 | 1 | | 3.14 | |
| | | | | | NGYTLTYDPYGR | 68.91 | 3.64 | 2 | | 2.11 | |
| | | | | | RPDEIKPNVNYQTHVQNIGWQGVVK | 101.71 | 5.07 | | | 2 3.27 | |
| | | | | | VPDINYQTHIQDIGWQGVVK | 114.79 | 4.07 | | 4 | 1.15 | |
| | | | | | YFGASATDLVITAQSYGR | 152.56 | 5.32 | 4 | 1 | 2.00 | |
| EF1964 | glyceraldehyde-3-phosphate dehydrogenase | gi29376486 | 31.53 | 8 | AIGLVIPELNGKLDGAAQR | 31.96 | 2.16 | | 1 | -13.34 | |
| | | | | | AIGLVIPELNGKLDGAAQRVPVATGSLTELVTVLDK | 76.73 | 3.86 | | | 3 -0.39 | |
| | | | | | FNGTVEVHEGSFNVNGK | 46.59 | 3.27 | 2 | 1 | 0.30 | |
| | | | | | IQDVEGIEVVAINDLTDAK | 154.77 | 5.08 | 1 | | -12.03 | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | TLEYFANL | 32.82 | 2.37 | 1 | | -0.94 | |
| | | | | | TVAWYDNEmSYTAQLVR | 149.55 | 3.36 | 4 | | 3.87 | M9(Oxidation) |
| | | | | | VGINGFGR | 28.04 | 2.28 | 1 | | -0.39 | |
| | | | | | VPVATGSLTELVTVLDK | 137.73 | 4.88 | 3 | | 3.05 | |
| EF3041 | pheromone binding protein | gi29377499 | 20.51 | 7 | AASSFYLEmNQADEK | 52.43 | 2.95 | 1 | | 1.94 | M9(Oxidation) |
| | | | | | ATVSEDGLVYK | 49.84 | 2.19 | 1 | | 0.26 | |
| | | | | | GWSADYSDPINFLDLLESSTSNNRGR | 44.78 | 2.71 | | 1 | 8.18 | |
| | | | | | nAEKISKGELPK | 1.00 | 2.25 | | 1 | -0.38 | N-Term(Acetyl) |
| | | | | | NILANGSLPSQGFVPVDVAK | 96.18 | 4.24 | 2 | | 2.05 | |
| | | | | | TLKGDFQIAVR | 43.25 | 2.51 | | 3 | 0.23 | |
| | | | | | WSDGKPVTAADYVYGWQR | 121.01 | 4.47 | 4 | 2 | 3.67 | |
| EF2864 | hypothetical protein EF2864 | gi29377332 | 25.74 | 4 | LLITQDSGNYPAEEYYYR | 300.00 | 4.89 | 4 | 3 | -0.22 | |
| | | | | | QMAANGGTGTVGFFR | 79.12 | 2.68 | 1 | | 2.34 | |
| | | | | | TLTDAENNDTNLGFLGmNGNDFFFR | 109.49 | 4.21 | 1 | 1 | 5.41 | M17(Oxidation) |
| | | | | | WGPTGAALLLYK | 68.56 | 2.48 | 2 | | 1.81 | |
| EF0200 | elongation factor G | gi29374846 | 17.03 | 7 | GLEDSmNNGVLAGYPLVDIK | 60.87 | 3.52 | 1 | | -0.09 | M6(Oxidation) |
| | | | | | IGADFFYSVESLHDR | 99.84 | 4.10 | | 1 | -13.46 | |
| | | | | | LYDGSYHDVDSNETAFR | 81.81 | 2.79 | | 2 | 1.79 | |
| | | | | | QATINVEFFPVLAGSAFK | 67.42 | 3.47 | 1 | | 1.88 | |
| | | | | | VNIIDTPGHVDFTIEVQR | 45.22 | 3.02 | | 1 | -6.46 | |
| | | | | | VYSGDIAAAVGLK | 55.19 | 3.02 | 2 | | -0.01 | |
| | | | | | VYSGVLESGSYVLNASK | 92.19 | 3.57 | 3 | | 1.32 | |
| EF1961 | enolase | gi29376483 | 17.36 | 4 | AAADYLEVPLYHYLGGFNTK | 87.30 | 4.13 | 1 | 4 | 1.88 | |
| | | | | | AVDNVNNIIAEAIIGYDVR | 32.91 | 2.76 | 1 | | 3.56 | |
| | | | | | GmVPSGASTGEYEAVELR | 74.95 | 2.37 | 2 | | 3.28 | M2(Oxidation) |
| | | | | | GNPTIEVEVYTESGAFGR | 123.36 | 5.23 | 2 | | -3.82 | |
| EF1818 | coccolysin | gi29376362 | 18.43 | 5 | GmPILSVVDEQHPDAYDNAFWDGK | 81.71 | 3.34 | | 3 | 1.01 | M2(Oxidation) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | IGYTIIQNLGIEK | 45.93 | 2.74 | 1 | | | 0.99 | |
| | | | | | NSFQVAFNVPVEK | 53.49 | 2.89 | 1 | | | -11.02 | |
| | | | | | VQYGDEAASVVSAAFNSAGIGAK | 115.72 | 3.77 | 1 | 1 | | -11.94 | |
| | | | | | YKGTPYYDQGGVHYNSGIINR | 88.95 | 4.15 | | 2 | 1 | 3.96 | |
| EF0176 | hypothetical protein EF0176 | gi29374826 | 23.53 | 5 | ALASSMYQAGADIIYHAAATTGQGIFQEAK | 141.16 | 7.10 | | 2 | | -15.65 | |
| | | | | | DIQITSTYAGTFADASK | 6.36 | 2.61 | 1 | | | 2.13 | |
| | | | | | SFNQSAWEGmQEWGK | 90.59 | 3.77 | 4 | | | 2.24 | M10(Oxidation) |
| | | | | | VGFIGGVEGPVIGR | 66.84 | 3.57 | 1 | | | 1.32 | |
| | | | | | VWVIGVDR | 51.51 | 2.09 | 1 | | | 0.03 | |
| EF0177 | hypothetical protein EF0177 | gi29374827 | 29.36 | 7 | ALAAAmYQNGVDIIFHASGATGQGVFQEAK | 51.93 | 4.11 | | | 1 | 2.84 | M6(Oxidation) |
| | | | | | DKVISGDVKVPEKPE | 35.60 | 2.29 | | 1 | | 2.15 | |
| | | | | | FNTIFGIGYLLK | 67.31 | 3.08 | 2 | | | 1.67 | |
| | | | | | GVGTAVQDIANR | 57.93 | 3.22 | 1 | | | 2.72 | |
| | | | | | SFNQSSWEGLQAWGK | 109.02 | 3.89 | 2 | | | 2.20 | |
| | | | | | VGFVGGEEGVVIDR | 91.57 | 3.70 | 1 | | | -10.57 | |
| | | | | | VWVIGVDR | 51.51 | 2.09 | 1 | | | 0.03 | |
| EF2174 | hypothetical protein EF2174 | gi29376682 | 4.48 | 3 | LTEGTWYVNPYAAGQIR | 67.94 | 4.09 | 3 | 1 | | 1.07 | |
| | | | | | NDALYVGKWWLTNGYIDTSAFGR | 81.20 | 3.90 | | 1 | | 8.16 | |
| | | | | | WWLTNGYIDTSAFGR | 104.25 | 3.93 | 3 | | | 1.80 | |
| EF0417 | hypothetical protein EF0417 | gi29375051 | 15.31 | 4 | EGEQAYVLVNDFGTIIR | 97.18 | 4.10 | 1 | 2 | | 1.18 | |
| | | | | | MAIESGLESADR | 77.44 | 3.03 | 3 | | | 0.50 | |
| | | | | | NGYHmQATIDLGDLGAIELPK | 64.50 | 4.25 | | 1 | | 4.56 | M5(Oxidation) |
| | | | | | RANIYNKWNR | 51.13 | 3.16 | | 1 | | -0.58 | |
| EF2144 | putative lipoprotein | gi29376653 | 12.59 | 2 | EANVEIHITPQSADSGLVEIIDY | 125.88 | 3.50 | 2 | 2 | | 2.69 | |
| | | | | | HQFNDMYPYKGSK | 38.26 | 2.51 | | 3 | | 0.91 | |
| EF2398 | 30S ribosomal protein S2 | gi29376895 | 15.33 | 4 | FLGGIADMPR | 55.14 | 2.57 | 2 | | | 0.36 | |
| | | | | | INAmEEDGTFEVLPK | 40.64 | 3.08 | 1 | | | -10.20 | M4(Oxidation) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | WLGGTLTNWDTIQK | 7.35 | 2.26 | 1 | | | -0.04 | |
| | | | | | WLGGTLTNWDTIQKR | 75.95 | 3.53 | 2 | | | 3.45 | |
| EF0206 | 50S ribosomal protein L3 | gi29374851 | 23.44 | 5 | EIKVDVFQAGDVVDVTGTTK | 36.99 | 2.92 | | 1 | | 1.42 | |
| | | | | | NVELGEYEVGK | 29.09 | 2.45 | 1 | | | 0.91 | |
| | | | | | VDVFQAGDVVDVTGTTK | 80.12 | 4.11 | 1 | | | -12.64 | |
| | | | | | YHRRPGSmGPVAPNR | 45.36 | 2.32 | | | 1 | -0.47 | M8(Oxidation) |
| | | | | | YHRRPGSmGPVAPNRVFK | 21.64 | 2.11 | | | 1 | -0.22 | M8(Oxidation) |
| EF1319 | hypothetical protein EF1319 | gi29375887 | 21.92 | 3 | NLTTLYQNPVNTPYmTK | 82.81 | 3.71 | 2 | | | 4.87 | M15(Oxidation) |
| | | | | | TNSPIDFSVTILNKQ | 76.03 | 3.35 | 2 | | | 2.91 | |
| | | | | | YRVTPVYNGNDLLAEK | 60.94 | 3.43 | 2 | | | 6.84 | |
| EF1308 | dnak protein | gi29375876 | 9.85 | 3 | DLSGVTSTQISLPFITAGEAGPLHLEmNLTR | 22.11 | 2.69 | | 1 | | 3.59 | M27(Oxidation) |
| | | | | | GVPQIEVSFDIDK | 61.27 | 2.67 | 1 | | | -9.35 | |
| | | | | | SYTPQEVSAmILQYLK | 115.90 | 3.88 | 3 | | | -0.50 | M10(Oxidation) |
| EF0970 | 50S ribosomal protein L27 | gi29375553 | 34.74 | 2 | IYPGVNVGIGGDDTLFAK | 100.80 | 2.74 | 4 | | | 1.98 | |
| | | | | | SADGQTVTGGSILYR | 31.28 | 2.54 | 1 | | | 2.43 | |
| EF0164 | putative lipoprotein | gi29374814 | 16.94 | 3 | IISHVGDLYDEK | 38.49 | 2.77 | 1 | 2 | | -0.31 | |
| | | | | | NKIISHVGDLYDEK | 25.98 | 2.39 | | 1 | | 0.64 | |
| | | | | | YFYNVNK | 25.69 | 2.09 | 1 | | | -0.56 | |
| EF1167 | fructose-bisphosphate aldolase | gi29375743 | 26.30 | 4 | GGYAVGGYTNNLEWTQAILEAAEAKK | 6.76 | 2.83 | | 1 | | -14.67 | |
| | | | | | GLAFDHLQAIAEAVGSDVPLVLHGGSGIPQEQIEK | 32.04 | 3.91 | | | 1 | -24.70 | |
| | | | | | KGGYAVGGYNTNNLEWTQAILEAAEAK | 99.78 | 4.88 | | 1 | | -14.67 | |
| | | | | | VNVNTEFQLSFAK | 82.25 | 3.86 | 2 | | | 1.86 | |
| EF0304 | putative lipoprotein | gi29374943 | 20.45 | 2 | TAIYGIQLNVEEVAK | 107.68 | 4.32 | 4 | | | 2.47 | |
| | | | | | TLENDNDVISFITPYTNGNDR | 96.58 | 3.54 | 1 | | | 6.55 | |
| EF1898 | 50S ribosomal protein L19 | gi29376426 | 19.13 | 2 | mnPLIQELTQEqLR | 1.00 | 2.80 | 3 | | | -0.56 | N-Term(Acetyl) N2(Deamidated) Q12(Deamidated) |
| | | | | | VAQIEVVR | 28.62 | 2.37 | 1 | | | 0.03 | |
| EF2718 | 50S ribosomal protein L1 | gi29377196 | 12.66 | 2 | FDATVEVAYK | 63.68 | 3.40 | 1 | | | -3.84 | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | NISVTTTFGPGIHVDQASF | 127.27 | 4.69 | 3 | | 0.64 | |
| EF0223 | 50S ribosomal protein L18 | gi29374867 | 38.98 | 2 | GGYLYHGRVQALAEAARENGLEF | 38.68 | 2.89 | | 1 | 5.05 | |
| | | | | | NIYAQVIDDVAGVTLASASALDK | 142.39 | 5.37 | 1 | 2 | 2.87 | |
| EF0228 | adenylate kinase | gi29374872 | 11.57 | 2 | GFLLDGFPR | 51.40 | 2.38 | 2 | | 0.02 | |
| | | | | | LAVNIESSAPILAFYK | 106.54 | 3.93 | 2 | | 2.29 | |
| EF3256 | pheromone cAD1 precursor lipoprotein | gi29377699 | 11.00 | 2 | FVKGFAAIALSSLVLAAcGADK | 1.00 | 2.42 | | 1 | -6.49 | C18(Carboxymethyl) |
| | | | | | NGYRAVFEMTVK | 29.78 | 2.84 | 2 | 1 | 1.30 | |
| EF_B0004 | TraC protein [Enterococcus faecalis V583] | gi29377898 | 6.81 | 2 | LILEDAGVIPLLQIGNAK | 93.67 | 2.59 | 1 | | -13.64 | |
| | | | | | WADGTDITADDFVTAWQR | 138.54 | 5.00 | 3 | | 2.40 | |
| EFA0003 | traC protein | gi29377806 | 5.71 | 2 | TASPSVELFSAIK | 33.53 | 2.31 | 1 | | -8.47 | |
| | | | | | WADGTDITADDFVTAWQR | 138.54 | 5.00 | 3 | | 2.40 | |
| EF2746 | dltD protein | gi29377221 | 7.55 | 3 | AINNNKFEISNGFYR | 75.06 | 3.53 | | 2 | -0.11 | |
| | | | | | FEISNGFYR | 27.87 | 2.24 | 1 | | -0.30 | |
| | | | | | FYQTDWQQQNPLVLPQF | 37.85 | 2.43 | | 1 | 0.20 | |
| EF1033 | lipoamidase | gi29375612 | 4.66 | 2 | ALQDTGQPFLGVPLLLK | 57.70 | 3.51 | 3 | | 2.71 | |
| | | | | | EGLPLGIQFNSALNEDR | 122.62 | 3.43 | 1 | | -11.35 | |
| EF0218 | 50S ribosomal protein L5 | gi29374863 | 21.23 | 3 | ELLAQLGmPFQK | 36.36 | 2.54 | 1 | | -2.14 | M8(Oxidation) |
| | | | | | FNYSSVmQTPK | 69.40 | 2.99 | 1 | | 0.88 | M7(Oxidation) |
| | | | | | IVINmGVGDAVSNAK | 45.23 | 3.13 | 1 | | 1.44 | M5(Oxidation) |
| EF0195 | phosphoglycerate mutase 1 | gi29374841 | 6.14 | 1 | ALPFWQDEIAPALK | 79.88 | 2.48 | 3 | | 1.27 | |
| EF0226 | 50S ribosomal protein L15 | gi29374870 | 8.22 | 1 | LGFEGGQTPLFR | 71.76 | 3.69 | 3 | | 0.86 | |
| EF0207 | 50S ribosomal protein L4 | gi29374852 | 13.53 | 2 | GGGVVFGPTPR | 10.02 | 2.02 | 1 | | 0.69 | |
| | | | | | VLVVLENGNDFAALSAR | 99.93 | 4.83 | 2 | | -4.38 | |
| EF0221 | 50S ribosomal protein L6 | gi29374866 | 20.79 | 3 | ALELIGVGYR | 45.90 | 2.49 | 1 | | 0.09 | |
| | | | | | ANFNNmVVGVSEGFQK | 38.33 | 2.86 | 1 | | 2.87 | M6(Oxidation) |
| | | | | | EVVGELAANIR | 51.91 | 2.35 | 1 | | 0.83 | |
| EF1523 | hypothetical protein EF1523 | gi29376088 | 6.37 | 3 | GYVVPGGYSLEPAK | 61.17 | 3.03 | 1 | | 2.13 | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | GYVVPGGYSLEPAKIVNGEGYYNLYATNNQSK | 79.63 | 3.14 | | 1 | 9.44 | |
| | | | | | | NVHMYLVTINAK | 9.35 | 2.10 | | 1 | -0.54 | |
| | EF0071 | putative lipoprotein | gi29374726 | 6.80 | 2 | ELLGGFAGPLIIAEEYPVNLAASLNK | 93.89 | 3.76 | 1 | 1 | -10.70 | |
| | | | | | | TNGVYDTNYFNNFSDLGAWHGYYLPEK | 143.21 | 5.17 | | 1 | 6.18 | |
| | EF1046 | pyruvate kinase | gi29375625 | 4.79 | 2 | AVVAATAEEAVAK | 28.27 | 2.28 | 1 | | 0.57 | |
| | | | | | | LVQGQGVGEEAIIAK | 49.66 | 3.62 | 2 | | 1.05 | |
| | EF0211 | 50S ribosomal protein L22 | gi29374856 | 17.39 | 2 | GSASPINKRTSHITVVVTEK | 47.45 | 3.08 | | | 1 -1.31 | |
| | | | | | | TSHITVVVTEK | 26.20 | 3.07 | | 1 | -0.77 | |
| | EF0991 | penicillin-binding protein C | gi29375573 | | 2 | LFTTAASmEQGQFNPNELFNR | 53.69 | 2.76 | 1 | | 6.61 | M8(Oxidation) |
| | | | | | | LYPNGQFASHFIGYTK | 42.65 | 2.57 | 1 | | 3.19 | |
| | EF0199 | 30S ribosomal protein S7 | gi29374845 | 21.15 | 2 | GIAANIIYNSFDIIK | 70.17 | 4.09 | 1 | | -9.53 | |
| | | | | | | RVGGSNYQVPVEVRPERR | 4.93 | 2.14 | | | 1 0.80 | |
| | EF0737 | amidase | gi29375331 | 5.30 | 2 | EQTLFEQAYSFEQSTK | 84.21 | 3.23 | 1 | | -11.21 | |
| | | | | | | HGQTLNPYGPLK | 47.87 | 2.49 | 1 | | 0.49 | |
| | EF1613 | formate acetyltransferase | gi29376172 | 7.09 | 2 | ILHTLTNmGPSPEPNLTVLYSSHLPEGFR | 8.11 | 4.69 | | | 1 7.84 | M8(Oxidation) |
| | | | | | | NGVYDmDSDIPATITSHEPGYLIK | 30.43 | 2.90 | | 1 | 8.82 | M6(Oxidation) |
| | EF0234 | 50S ribosomal protein L17 | gi29374878 | 11.02 | 1 | RGDAAPmVVIEFVK | 26.57 | 2.69 | 1 | | 2.56 | M7(Oxidation) |
| | EF1420 | hypothetical protein EF1420 | gi29375987 | 4.40 | 1 | TSAEIQLGISK | 57.34 | 3.36 | 1 | | -0.13 | |
| | EF0633 | tyrosyl-tRNA synthetase | gi29375234 | 5.02 | 2 | ISEALFSGNIK | 29.27 | 2.12 | 2 | | 0.96 | |
| | | | | | | MNIIDELAWR | 58.85 | 3.26 | 1 | | -1.1 | |
| | EF0685 | rotamase family protein | gi29375282 | 10.82 | 2 | QLKqRAAYDAGLK | 9.8 | 2.04 | 1 | | -6.76 | Q4(Deamidated) |
| | | | | | | TAWASFHPEVEAQIIQVASEDDAK | 26.35 | 2.03 | | 1 | 1.8 | |
| | EF0907 | peptide ABC transporter, peptide-binding protein | gi29375491 | 7.01 | 2 | WSDGKPVTANDYVYGWQR | 21.85 | 3.34 | | 2 | 0.68 | |
| | | | | | | iSFIALnnVYEGIYRLDKDnK | 1 | 2.03 | | 1 | 2.03 | N-Term(Acetyl) N7(Deamidated) N8(Deamidated) N20(Deamidated) |
| Beads 1h | EF2556 | fumarate reductase flavoprotein subunit | gi29377044 | 47.13 | | AIDFYDQKGFVEK | 58.58 | 2.97 | 2 | | 2.31 | |
| | | | | | | AIDFYDQKGFVEKGETIEELAEK | 81.90 | 3.55 | | 2 | 2 -2.46 | |
| | | | | | | AKAIDFYDQK | 25.95 | 2.41 | 2 | | 0.33 | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | AKAVVVTTGGFGANEK | 29.41 | 2.95 | | 1 | | 1.02 | |
| | | | | | ATIDTWNQDVNAK | 48.23 | 2.62 | 2 | | | 0.42 | |
| | | | | | AVVVTTGGFANEK | 68.56 | 3.65 | 3 | | | -0.13 | |
| | | | | | DKVSAAINALPEK | 65.81 | 3.22 | 1 | 2 | | -0.11 | |
| | | | | | DKVSAAINALPEKSAYLVFDQGVR | 146.12 | 5.44 | | 2 | 1 | 0.79 | |
| | | | | | EDGTPIKGLYAAGELTGGLHGQNR | 55.73 | 3.57 | | 2 | 1 | 1.38 | |
| | | | | | FVNELDTR | 38.92 | 2.40 | 3 | | | 0.08 | |
| | | | | | GLYAAGELTGGLHGQNR | 89.97 | 4.32 | 4 | | | 0.82 | |
| | | | | | IGGNAIADIIIYGR | 126.24 | 4.77 | 131 | 24 | | 0.03 | |
| | | | | | IPLFVDADVTDLVEENGQIDGVK | 91.82 | 5.33 | 3 | | | -1.03 | |
| | | | | | MPVAGGNTIK | 43.71 | 2.31 | 1 | | | 0.35 | |
| | | | | | NYVTTNQEGTTGDGIQmlQK | 98.36 | 4.28 | 4 | 1 | | 0.30 | M17(Oxidation) |
| | | | | | QAGTQSAEFASAQK | 35.26 | 2.24 | 1 | | | 1.40 | |
| | | | | | RTHRPADGSAIGGYLVDGLVR | 76.02 | 4.45 | | 2 | 1 | 1.41 | |
| | | | | | SAYLVFDQGVR | 77.30 | 2.86 | 1 | | | 0.24 | |
| | | | | | THRPADGSAIGGYLVDGLVR | 47.17 | 4.10 | 1 | | 1 | -3.01 | |
| | | | | | TTGMEADLSTAPYYAIK | 78.60 | 4.41 | 5 | | | 2.55 | |
| | | | | | VGGALVDmK | 55.15 | 3.26 | 4 | | | -0.09 | M8(Oxidation) |
| | | | | | VSAAINALPEKSAYLVFDQGVR | 88.67 | 4.29 | | 2 | | 0.34 | |
| EF0201 | elongation factor Tu | gi29374847 | 19.24 | 5 | DLLSEYDFPGDDVPVIAGSALK | 128.76 | 4.93 | 2 | | | 1.20 | |
| | | | | | FKAEVYVLSK | 23.46 | 2.67 | | 1 | | -0.18 | |
| | | | | | GITINTSHIEYETETR | 38.80 | 3.10 | 1 | 1 | | 1.66 | |
| | | | | | LLDYAEAGDNIGALLR | 127.40 | 4.66 | 2 | | | 1.15 | |
| | | | | | TVGSGVVTEIVK | 33.48 | 2.69 | 2 | | | 0.27 | |
| EF0394 | secreted antigen, putative | gi29375030 | 11.14 | 3 | ASLALEQSSAESSK | 117.49 | 4.46 | 1 | | | 1.68 | |
| | | | | | QSLGLRPVVWDAGLAASATAR | 45.01 | 4.36 | | 1 | | 3.60 | |
| | | | | | VGFGYSGSTIVGHSA | 62.63 | 2.78 | 3 | | | 0.53 | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EF1961 | enolase | gi29376483 | 4.63 | 1 | AAADYLEVPLYHYLGGFNTK | 69.19 | 3.58 | | 3 | | 1.70 | |
| | EF1167 | fructose-bisphosphate aldolase | gi29375743 | 14.19 | 3 | GGYAVGGYNTNNLEWTQAILEAAEAKK | 1.38 | 2.29 | | 1 | | -14.49 | |
| | | | | | | KGGYAVGGYNTNNLEWTQAILEAAEAK | 70.82 | 4.47 | | 1 | | -14.49 | |
| | | | | | | VNVNTEFQLSFAK | 64.92 | 3.79 | 1 | | | 2.72 | |
| | EF0123 | hypothetical protein EF0123 | gi29374774 | 1.79 | 1 | LTGEIANAYDVYYR | 92.06 | 4.20 | 3 | | | 2.53 | |
| | EF1264 | sulfatase domain-containing protein | gi29375833 | 3.42 | 1 | FIAVSNHYPYSQFTNDEAGFPIAK | 79.16 | 4.61 | | 3 | | -0.60 | |
| | EF0177 | hypothetical protein EF0177 | gi29374827 | 8.03 | 2 | SFNQSSWEGLQAWGK | 75.07 | 2.85 | 1 | | | -3.30 | |
| | | | | | | VGFVGGEEGVVIDR | 69.44 | 3.60 | 2 | | | -2.02 | |
| | EF2398 | 30S ribosomal protein S2 | gi29376895 | 9.20 | 2 | FLGGIADmPR | 19.91 | 2.17 | 1 | | | -0.74 | M8(Oxidation) |
| | | | | | | WLGGTLTnWDTIqK | 1.00 | 2.12 | 1 | | | 9.71 | N8(Deamidated) Q13(Deamidated) |
| | EF1379 | alanyl-tRNA synthetase | gi29375946 | 3.30 | 2 | FHETINEGLSmLnEVIK | 4.61 | 2.04 | 1 | | | 9.70 | M11(Oxidation) N13(Deamidated) |
| | | | | | | IVSESGIGAGVR | 18.43 | 2.40 | 1 | | | 0.12 | |
| | EF2864 | hypothetical protein EF2864 | gi29377332 | 11.03 | 2 | LLITQDSGNYPAEEYYYR | 105.45 | 4.71 | | 1 | | 4.54 | |
| | | | | | | WGPTGAALLLYK | 68.21 | 2.67 | 1 | | | -8.69 | |
| | EF2860 | YkuD putative, pewptidoglycan binding protein | gi29377328 | 7.38 | 1 | IFDVSYDGmPVIIYGHIYDDAPGEFDKPVDYGEEV | 136.28 | 5.54 | | 2 | | 4.61 | M9(Oxidation) |
| | EF0226 | 50S ribosomal protein L15 | gi29374870 | 8.22 | 1 | LGFEGGQTPLFR | 59.91 | 3.06 | 1 | | | 0.50 | |
| | EF1818 | coccolysin | gi29376362 | 4.71 | 1 | GMPILSVVDEQHPDAYDNAFWDGK | 38.17 | 2.14 | | 1 | | -10.71 | |
| | EF0211 | 50S ribosomal protein L22 | gi29374856 | 9.57 | 1 | TSHITVVVTEK | 30.00 | 2.20 | | 1 | | -0.95 | |
| | EF2925 | cold-shock domain-contain protein | gi29377389 | 19.70 | 1 | WFNAEKGFGFISR | 49.33 | 2.53 | | 1 | | -3.37 | |
| | EF2746 | dltD protein | gi29377221 | 2.83 | 1 | YVPFFGSSELSR | 42.56 | 2.96 | 1 | | | -2.36 | |
| | EF_B0004 | TraC protein | gi29377898 | 3.40 | 1 | WADGTDITADDFVTAWQR | 125.45 | 4.77 | 1 | | | 0.45 | |
| Beads 2h | EF2556 | fumarate reductase flavoprotein subunit | gi29377044 | 64.95 | 37 | AIDFYDQK | 45.18 | 2.47 | 2 | | | 0.34 | |
| | | | | | | AIDFYDQKGFVEK | 23.89 | 2.19 | 1 | | | 5.61 | |
| | | | | | | AIDFYDQKGFVEKGETIEELAEK | 119.71 | 3.58 | | 2 | 2 | 0.47 | |
| | | | | | | AKAIDFYDQK | 21.29 | 3.08 | | 1 | | -1.02 | |
| | | | | | | AKAVVVTTGGFGANEK | 54.49 | 3.36 | 1 | | | 4.21 | |
| | | | | | | ATIDTWNQDVNAK | 73.73 | 4.14 | 2 | | | 1.96 | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ATIDTWNQDVNAKDDKQFGR | 61.36 | 2.86 | | 1 | | 5.22 | |
| AVVVTTGGFGANEK | 91.39 | 3.69 | 8 | | | 0.07 | |
| DKVSAAINALPEK | 95.21 | 4.24 | 2 | | | 1.58 | |
| DKVSAAINALPEKSAYLVFDQGVR | 93.68 | 4.56 | | 2 | 2 | 2.98 | |
| DSNDKFFEETLK | 105.70 | 3.26 | 1 | 2 | | -0.09 | |
| EAGMNPVILEK | 55.28 | 2.77 | 2 | | | 0.77 | |
| EDGTPIKGLYAAGELTGGLHGQNR | 100.53 | 5.59 | 1 | 3 | 3 | 2.35 | |
| EEKIPLFVDADVTDLVEENGQIDGVK | 88.88 | 5.69 | | 2 | | 5.07 | |
| EEKIPLFVDADVTDLVEENGQIDGVKVK | 105.41 | 2.88 | 1 | | 1 | 0.72 | |
| EGIKDSNDKFFEETLK | 43.61 | 2.47 | 1 | | | 0.22 | |
| EIQIHPTVQQSDAFLIGEAVR | 105.77 | 4.49 | 1 | 4 | | 2.03 | |
| FVNELDTR | 18.43 | 2.26 | 1 | | | 0.26 | |
| GETIEELAEK | 66.96 | 2.91 | 1 | | | 0.90 | |
| GFVEKGETIEELAEK | 61.69 | 3.58 | 2 | 2 | | -0.68 | |
| GITLSNLTITGGmSEK | 94.23 | 4.59 | 4 | | | 2.42 | M13(Oxidation) |
| GITLSNLTITGGmSEKR | 77.69 | 2.65 | 1 | | | 4.40 | M13(Oxidation) |
| GLYAAGELTGGLHGQNR | 142.06 | 4.41 | 10 | 7 | | 0.38 | |
| IGGNAIADIIIYGR | 122.46 | 4.83 | 170 | 37 | | 0.03 | |
| IGMPADTLKATIDTWNQDVNAK | 60.30 | 3.55 | | 1 | | 3.50 | |
| IPLFVDADVTDLVEENGQIDGVK | 99.58 | 5.26 | 3 | | | 4.83 | |
| IPLFVDADVTDLVEENGQIDGVKVK | 147.78 | 4.46 | | 3 | | 1.16 | |
| LITQYKPELK | 25.70 | 2.57 | 1 | | | 0.69 | |
| LITQYKPELKNYVTTNQEGTTGDGIQMIQK | 72.43 | 3.83 | | | 1 | 7.48 | |
| NYVTTNQEGTTGDGIQMIQK | 108.16 | 5.90 | 4 | 5 | | 2.28 | |
| QAGTQSAEFASAQK | 36.35 | 2.51 | 1 | | | 4.46 | |
| RTHRPADGSAIGGYLVDGLVR | 84.11 | 4.77 | | 6 | | -0.02 | |
| SAYLVFDQGVR | 75.92 | 3.50 | 3 | | | 1.21 | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | TEVLREDGTPIKGLYAAGELTGGLHGQNR | 152.56 | 7.33 | | 1 | 1 | 5.95 | |
| | | | | | THRPADGSAIGGYLVDGLVR | 122.26 | 5.98 | | 8 | 6 | -0.01 | |
| | | | | | TTGmEADLSTAPYYAIK | 92.35 | 4.65 | 7 | | | 2.43 | M4(Oxidation) |
| | | | | | VGGALVDmK | 46.38 | 2.38 | 1 | | | -0.76 | M8(Oxidation) |
| EF2860 | YkuD putative, pewptidoglycan binding protein | gi29377328 | 23.21 | 7 | FKNNGSYGWSIDGAK | 77.56 | 3.35 | 2 | 3 | | 0.58 | |
| | | | | | IFDVSYDGmPVIIYGHIYDDAPGEFDKPVDYGEEV | 300.00 | 5.66 | | 6 | 4 | 0.14 | M9(Oxidation) |
| | | | | | LNQQIIADVEAGKGNYQYNAK | 30.88 | 2.20 | | 1 | | 1.48 | |
| | | | | | NNGSYGWSIDGAK | 56.08 | 3.41 | 1 | | | 2.61 | |
| | | | | | RFKNNGSYGWSIDGAK | 45.85 | 3.85 | 1 | 3 | | 1.64 | |
| | | | | | RGNGTFEIVPEEQGTVVDTQR | 27.63 | 3.30 | | 1 | | 0.10 | |
| | | | | | YNKGTATVPGFHTILYR | 73.60 | 2.78 | 1 | 3 | 1 | 0.83 | |
| EF0201 | elongation factor Tu | gi29374847 | 16.20 | 4 | DLLSEYDFPGDDVPVIAGSALK | 300.00 | 5.13 | 4 | | | -2.95 | |
| | | | | | FKAEVYVLSK | 47.05 | 2.46 | | 2 | | -0.27 | |
| | | | | | GITINTSHIEYETETR | 25.85 | 2.32 | | 2 | | 1.29 | |
| | | | | | LLDYAEAGDNIGALLR | 119.39 | 4.79 | 4 | | | 1.03 | |
| EF2864 | hypothetical protein EF2864 | gi29377332 | 11.03 | 2 | LLITQDSGNYPAEEYYYR | 159.55 | 4.98 | 3 | 3 | | 2.16 | |
| | | | | | WGPTGAALLLYK | 53.05 | 3.02 | 2 | | | 0.71 | |
| EF1264 | sulfatase domain-containing protein | gi29375833 | 12.82 | 5 | FIAVSNHYPYSQFTNDEAGFPIAK | 78.59 | 3.71 | | 2 | | 4.16 | |
| | | | | | FYTNSGLKPVNPEDYDYK | 20.16 | 3.19 | | 1 | | -0.56 | |
| | | | | | NYIQLGQDLFSK | 60.95 | 2.83 | 1 | | | 0.17 | |
| | | | | | QTQGYTSAAFHGNAGNFWNR | 104.80 | 2.83 | | 3 | | 0.23 | |
| | | | | | STFSFDNFFHQVGQGK | 96.39 | 3.15 | 1 | | | 4.37 | |
| EF1818 | coccolysin | gi29376362 | 8.82 | 2 | GmPILSVVDEQHPDAYDNAFWDGK | 76.01 | 3.27 | | 2 | | 3.84 | M2(Oxidation) |
| | | | | | YKGTPYYDQGGVHYNSGIINR | 63.53 | 4.28 | 2 | 2 | | 0.31 | |
| EF2925 | cold-shock domain-contain protein | gi29377389 | 30.30 | 2 | mEqGTVKWFnAEKGFGFISR | 1.00 | 2.84 | | 1 | | 13.81 | N-Term(Acetyl) Q3(Deamidated) N10(Deamidated) |
| | | | | | WFNAEKGFGFISR | 63.21 | 3.07 | 3 | 1 | | -0.26 | |
| EF0123 | hypothetical protein EF0123 | gi29374774 | 9.83 | 4 | LTGEIANAYDVYYR | 50.20 | 2.48 | 1 | | | 4.12 | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | RPDEIKPNVNYQTHVQNIGWQGVVK | 66.45 | 4.13 | | | 2 | 3.51 | |
| | | | | | VPDINYQTHIQDIGWQGVVK | 28.77 | 3.37 | | 1 | | 3.34 | |
| | | | | | YFGASATDLVITAQSYGR | 94.25 | 3.29 | 1 | | | 4.68 | |
| EF3041 | pheromone binding protein | gi29377499 | 5.26 | 2 | TLKGDFQIAVR | 39.43 | 3.41 | 1 | 2 | | 0.03 | |
| | | | | | WSDGKPVTAADYVYGWQR | 128.81 | 3.75 | 1 | 1 | | 3.57 | |
| EF0164 | putative lipoprotein | gi29374814 | 20.97 | 2 | IISHVGDLYDEKYQEK | 77.22 | 3.73 | | 1 | 1 | 0.21 | |
| | | | | | SHGNYEVIYK | 23.55 | 2.95 | 1 | 1 | | 0.01 | |
| EF_B0004 | TraC protein | gi29377898 | 3.40 | 1 | WADGTDITADDFVTAWQR | 134.41 | 5.11 | 4 | | | 0.94 | |
| EF2746 | dltD protein | gi29377221 | 6.84 | 2 | FYQTDWQQQNPLVLPQF | 51.51 | 2.73 | | 2 | | 2.77 | |
| | | | | | YVPFFGSSELSR | 71.39 | 2.62 | 2 | | | 1.42 | |
| EF0176 | hypothetical protein EF0176 | gi29374826 | 8.12 | 2 | SFNQSAWEGMQEWGK | 95.61 | 3.87 | 2 | | | 3.70 | |
| | | | | | VGFIGGVEGPVIGR | 64.76 | 4.03 | 1 | | | 1.56 | |
| EF0417 | hypothetical protein EF0417 | gi29375051 | | 3 | NGYHmQATIDLGDLGAIELPK | 32.24 | 3.03 | 1 | | | 4.70 | M5(Oxidation) |
| | | | | | RANIYNKWNR | 60.61 | 3.12 | | 1 | | -0.48 | |
| EF0502 | hypothetical protein EF0502 | gi29375130 | 7.94 | 2 | VVISSKKPVKVGDIVESDAAIASDESATNESMTDASK | 1.00 | 2.23 | | | 1 | 17.60 | |
| | | | | | RFFGIFLMTILIFTGLSVLKDANTSNSLFDmMFSVDK | 21.56 | 2.17 | | | 1 | 2.63 | |
| | | | | | MDQVKqVTqVDVPqTHSTPQRVqSK | 1.38 | 2.11 | | 4 | | 7.13 | |
| EF0304 | putative lipoprotein | gi29374943 | 20.45 | 2 | TAIYGIQLNVEEVAK | 111.82 | 4.32 | 2 | | | 2.11 | |
| | | | | | TLENDNDVISFITPYTNGNDR | 82.53 | 3.35 | 1 | | | -0.04 | |
| EF0211 | 50S ribosomal protein L22 | gi29374856 | 9.57 | 1 | TSHITVVVTEK | 39.37 | 2.37 | | 2 | | -0.86 | |
| EF0205 | 30S ribosomal protein S10 | gi29374850 | 11.76 | 1 | LDLPSGVNIEIK | 59.24 | 2.44 | 2 | | | 0.11 | |
| EF0226 | 50S ribosomal protein L15 | gi29374870 | 8.22 | 1 | LGFEGGQTPLFR | 78.94 | 3.19 | 2 | | | 0.37 | |
| EF2398 | 30S ribosomal protein S2 | gi29376895 | 5.36 | 1 | WLGGTLTNWDTIQK | 59.01 | 2.18 | 1 | | | 4.72 | |
| EF1167 | fructose-bisphosphate aldolase | gi29375743 | 9.34 | 1 | KGGYAVGGYNTNNLEWTQAILEAAEAK | 71.25 | 3.97 | | 1 | | 8.58 | |
| EF0394 | secreted antigen, putative | gi29375030 | 4.45 | 1 | aAAEAEQARLAAEQKAAAEK | 1.00 | 2.54 | | 1 | | -4.03 | N-Term(Acetyl) |

Table S6. Proteome data of the proteins identified using the SDS-PAGE approach after different treatments (untreated, trypsin and trypsin beads: incubation 2 hours).

| Sample | Gene | Gene product | Accession nr | Cover-age % | No of unique peptides | Peptide sequence | Best probability score | Best Xcorr score | No of +2H spectra | No of +3H spectra | Δ M ppm | Modification |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Untreated | EFA0052 | surface exclusion protein Sea1 | gi29377848 | 2.58 | 2 | LENAQPTYEK | 20.09 | 2.29 | 1 | | -0.21 | |
| | | | | | | VADAQAIEQTSAK | 15.96 | 2.05 | 1 | | 1.73 | |
| Trypsin | EF2903 | ABC transporter, substrate-binding protein | gi29377367 | 14.32 | 6 | DGIEAGYFR | 55.69 | 2.62 | 2 | | 0.16 | |
| | | | | | | DSQEVVDYYR | 57.57 | 2.31 | 2 | | 2.76 | |
| | | | | | | INSTLTSPK | 5.48 | 2.16 | 1 | | -0.38 | |
| | | | | | | TAAFEFmK | 29.44 | 2.03 | 1 | | 0.47 | M7(Oxidation) |
| | | | | | | VELQNQSAYPDLQAK | 68.95 | 2.36 | 1 | | 4.23 | |
| | | | | | | VPAQLENAVK | 48.29 | 2.52 | 1 | | 0.06 | |
| | EFA0052 | surface exclusion protein Sea1 | gi29377848 | 2.58 | 2 | DLEAQQAEQR | 56.73 | 3.12 | 1 | | 1.90 | |
| | | | | | | VADAQAIEQTSAK | 66.96 | 2.36 | 1 | | 1.85 | |
| | EF2713 | cell wall surface anchor family protein | gi29377193 | 9.41 | 2 | GDYGLAAMADK | 63.23 | 3.79 | 1 | | 0.42 | |
| | | | | | | VDYANQNYPSDVR | 58.75 | 2.69 | 1 | | 3.28 | |
| | EF3106 | peptide ABC transporter, peptide-binding protein | gi29377562 | 3.37 | 2 | LDEDANTATIK | 19.28 | 2.08 | 1 | | 3.33 | |
| | | | | | | NEVLPVNDR | 18.07 | 2.00 | 1 | | 0.58 | |
| | EF3257 | pyridine nucleotide-disulfide family oxidoreductase | gi29377700 | 3.70 | 2 | ATSDAADFGLEAAR | 79.46 | 2.67 | 1 | | 2.79 | |
| | | | | | | LVANEYmQAK | 53.94 | 2.46 | 1 | | 1.04 | M7(Oxidation) |
| Beads | EF2633 | chaperonin, GroEL | gi29377118 | 9.24 | 4 | GTFNVVAVK | 29.77 | 2.20 | 1 | | 0.18 | |
| | | | | | | VGNDGVITIEESK | 39.22 | 3.08 | 1 | | 3.75 | |
| | | | | | | VGQLIADAMEK | 18.87 | 2.19 | 2 | | -0.02 | |
| | | | | | | VVVDKDNTTIVEGAGSK | 12.96 | 2.63 | | 1 | 2.36 | |
| | EF2857 | penicillin-binding protein 2B | gi29377325 | 4.78 | 3 | AFAEYGmGTK | 18.41 | 2.23 | 1 | | 0.78 | M7(Oxidation) |
| | | | | | | ANLAITYTR | 36.52 | 2.64 | 1 | | 1.55 | |
| | | | | | | TGIDIPGETTGIQNK | 11.28 | 2.22 | 1 | | 2.57 | |
| | EF2221 | ABC transporter, substrate-binding protein | gi29376728 | 5.12 | 2 | TEITNVATVMNR | 20.80 | 2.38 | 2 | | 0.75 | |
| | | | | | | VDGSYESATEVLK | 63.63 | 2.17 | 1 | | 2.04 | |
| | EF2715 | ribosomal protein L7/L12 | gi29377194 | 18.85 | 2 | AVVDGAPAPVK | 31.09 | 2.21 | 1 | | 0.27 | |
| | | | | | | EAVSKEEAEALK | 31.84 | 2.15 | | 1 | -0.05 | |
| | EF0517 | 2-dehydropantoate 2-reductase | gi29375143 | 7.69 | 2 | ANYNGEEITVK | 18.54 | 2.12 | 1 | | -0.92 | |
| | | | | | | LAETLSASGLNAK | 46.39 | 2.18 | 1 | | 2.80 | |
| | EF2224 | cell wall surface anchor family protein | gi29376731 | 1.40 | 2 | IEDTSSQEDIK | 45.17 | 2.40 | 1 | | -0.60 | |
| | | | | | | TVTYEVTNTR | 38.84 | 2.30 | 1 | | 0.87 | |
| | EFA0052 | surface exclusion protein Sea1 | gi29377848 | 2.47 | 2 | ATQTTEQAITEK | 15.35 | 2.21 | 1 | | 2.59 | |
| | | | | | | LENAQPTYEK | 66.40 | 2.43 | 1 | | 0.53 | |
| | EF2903 | ABC transporter, substrate-binding protein | gi29377367 | 5.16 | 2 | DATSQFEQAWNQ | 44.93 | 2.29 | 1 | | 2.63 | |
| | | | | | | VPAQLENAVK | 76.72 | 2.88 | 1 | | 0.52 | |
| | EF0968 | 50S ribosomal protein L21 | gi29375551 | 18.63 | 2 | LNVEAGEK | 47.34 | 2.31 | 1 | | -0.72 | |
| | | | | | | VEVGQAIYVEK | 8.45 | 2.57 | 1 | | 0.45 | |

# PAPER III

RESEARCH LETTER

# An endo-β-*N*-acetylglucosaminidase from *Enterococcus faecalis* V583 responsible for the hydrolysis of high-mannose and hybrid-type N-linked glycans

Liv Anette Bøhle, Geir Mathiesen, Gustav Vaaje-Kolstad & Vincent G.H. Eijsink

Department of Chemistry, Biotechnology, and Food Science, The Norwegian University of Life Sciences, Ås, Norway

**Correspondence:** Geir Mathiesen, Department of Chemistry, Biotechnology, and Food Science, The Norwegian University of Life Sciences, 1432 Ås, Norway. Tel.: +47 64 965899; fax: +47 64 965901; e-mail: geir.mathiesen@umb.no

Received 4 July 2011; revised 11 August 2011; accepted 13 September 2011.

DOI: 10.1111/j.1574-6968.2011.02419.x

Editor: Marco Moracci

**Keywords**
endo-β-*N*-acetylglucosaminidase; *Enterococcus faecalis* V583; glycoprotein; N-linked glycan.

## Abstract

It has been demonstrated previously that *Enterococcus faecalis* produces secreted endoglycosidases that enable the bacteria to remove N-linked glycans from glycoproteins. One enzyme potentially responsible for this activity is EF0114, comprising a typical GH18 endoglycosidase domain and a GH20 domain. We have analyzed the other candidate, EF2863, and show that this predicted single domain GH18 protein is an endo-β-*N*-acetylglucosaminidase. EF2863 hydrolyzes the glycosidic bond between two *N*-acetylglucosamines (GlcNAc) in N-linked glycans of the high-mannose and hybrid type, releasing the glycan and leaving one GlcNAc attached to the protein. The activity of EF2863 is similar to that of the well known deglycosylating enzyme EndoH from *Streptomyces plicatus*. According to the CAZy nomenclature, the enzyme is designated *Ef*Endo18A.

## Introduction

Enterococci are Gram-positive lactic acid bacteria that traditionally have been considered harmless inhabitants of the intestinal tract of mammals. However, in recent years, enterococci have emerged as nosocomial pathogens, causing urinary tract infections, bacteraemia and infective endocarditis. Most clinical infections caused by enterococci are due to *Enterococcus faecalis* (Tendolkar *et al.*, 2003; Fisher & Phillips, 2009). *Enterococcus faecalis* V583 was the first clinical vancomycin-resistant enterococcal isolate reported in the USA (Sahm *et al.*, 1989) and sequencing of the genome was completed in 2003 (Paulsen *et al.*, 2003). Due to increasing problems with antibiotic resistance, treatment of enterococcal infections is difficult, and there is a need for alternative strategies for overcoming pathogen-related diseases. There is thus an interest in identifying enterococcal proteins that may affect the bacteria's persistence and/or virulence.

More than two/thirds of the eukaryotic proteins are predicted to be glycosylated (Apweiler *et al.*, 1999), including many proteins of the immune system. It has been shown that pathogenic and commensal Gram-positive bacteria produce enzymes that hydrolyze the glycans of glycoproteins, thus liberating nutrients supporting bacterial growth (Roberts *et al.*, 2000, 2001; Collin & Fischetti, 2004; Sanchez *et al.*, 2010; Ruiz *et al.*, 2011). Glycan hydrolysis may also serve the purpose of modifying the functionality of specific host proteins, e.g. in the immune system (Collin and Olsen, 2001, 2003). *Enterococcus faecalis* is known to have extracellular endo-β-*N*-acetylglucosaminidase activity that enables growth on RNaseB by releasing high-mannose type glycans from a single N-glycosylation site (Roberts, *et al.*, 2000, 2001; Collin & Fischetti, 2004).

There are three main types of N-linked glycans that are all build on a common core pentasaccharide ($Man_3$-$GlcNAc_2$) that is linked to the asparagine through one of two consecutive GlcNAc units (Supporting Information, Fig. S1): (1) the 'high-mannose' type found in RNaseB, containing additional α-linked mannose residues, (2) the 'complex' type, which has no additional mannose residues but may have as many as five antennae containing a variety of additional sugar types, and (3) the 'hybrid' type, comprising a combination of high-mannose and complex-type branches. Hydrolysis of the N-linked glycans by endo-β-*N*-acetylglucosaminidases entails hydrolysis of the





FEMS MICROBIOLOGY LETTERS

glycosidic bond between the two GlcNAc residues in the ($Man_3$-$GlcNAc_2$) core, which implies that the deglycosylated protein retains one GlcNAc (Morelle and Michalski, 2005, 2007).

The ability of *E. faecalis* to release high-mannose type glycans from RNaseB has been ascribed to a protein called EndoE from *E. faecalis* HER1044, corresponding to EF0114 of *E. faecalis* V583 (99% sequence identity; Collin & Fischetti, 2004). This is a two-domain protein consisting of a family 18 glycoside hydrolase (GH18) with the originally detected endo-β-*N*-acetylglucosaminidase activity and a GH20 domain that hydrolyzes complex-type glycans of IgG. In addition to EF0114, the genome of *E. faecalis* V583 encodes two other GH18 proteins, EF0361 and EF2863 (Cantarel *et al.*, 2009). EF0361 is a chitinase (Leisner *et al.*, 2009) and its gene is followed by *ef0362*, encoding a CBM33 protein that belongs to a family of enzymes known to play a critical role in chitin degradation (Vaaje-Kolstad *et al.*, 2010). EF2863 is predicted to be a secreted endo-β-*N*-acetylglucosaminidase that has not been investigated so far, despite its potential importance for the ability of *E. faecalis* V583 to exploit host glycoproteins.

In this study, we have characterized the putative endo-β-*N*-acetylglucosaminidase, EF2863, from *E. faecalis* V583. The results confirm the predicted activity and provide information concerning the ability of this novel enzyme to hydrolyze different types of glycoproteins. Hereafter, the enzyme is named *Ef*Endo18A, in accordance with the CAZy nomenclature (Henrissat *et al.*, 1998; Cantarel *et al.*, 2009).

## Materials and methods

### Cloning, expression and purification of the endo-β-*N*-acetylglucosaminidase, *Ef*Endo18A

Genomic DNA from *E. faecalis* V583 and pBAD/HisB expression plasmid (Invitrogen, Karlsruhe, Germany) from *Escherichia coli* were isolated, using the E.Z.N.A.® Bacterial DNA Kit (Omega Bio-Tek Inc., Norcross, GA), and the E.Z.N.A.® Plasmid Miniprep kit I (Omega), respectively. The gene corresponding to EF2863 (without the part encoding a predicted N-terminal signal peptide) was amplified by PCR (forward primer, 5′-AGATCTG CATCAACTGTTACACC-3′; reverse primer, 5′-GAATTC TTAAGGTGTTGGAACAGTT-3′; restriction sites are underlined). Amplified fragments were digested with BglII and EcoRI and cloned into a BglII/EcoRI-digested pBAD/HisB-vector (Invitrogen) using Quick Ligation Kit (New England Biolabs, Ipswich, MA). Transformation of the ligation mix into *E. coli* TOP10 compentent cells followed by selective plating on brain heart infusion (BHI) plates containing 0.1 mg $mL^{-1}$ ampicillin yielded transformants containing the pBad/HisB-EF plasmid for *Ef*Endo18A expression. The gene sequence was verified by DNA sequencing using a BigDye® Terminator v3.1 Cycle Sequencing Kit (Perkin Elmer/Applied Biosystems, Foster City, CA).

A 10-mL overnight culture of *E. coli* harbouring pBAD/HisB-EF was added to 500 mL fresh BHI broth (Oxoid Ltd., Hampshire, UK) containing 0.1 mg $mL^{-1}$ ampicillin, and the culture was incubated at 37 °C with shaking. At an $OD_{600\,nm}$ of 0.7, expression was induced by the addition of L-arabinose to a final concentration of 0.002% (w/v). The culture was further incubated at 30 °C overnight, after which the cells were harvested by centrifugation (7700 ***g***, 10 min, 4 °C) and resuspended in 20 mL Buffer A (100 mM TrisHCl pH 8, 20 mM imidazole). The cells were lysed by sonication, using a Vibra cell Ultrasonic Processor converter (Sonics, Newton, CT), at 20% amplitude with 5-s pulses (with a 5-s delay between pulses) for 15 min on ice. The sonicated cells were centrifuged (17 400 ***g*** 15 min, 4 °C), and the supernatant was applied to a Ni-NTA column equilibrated with Buffer A. *Ef*Endo18A was eluted with Buffer B (100 mM TrisHCl pH 8, 100 mM imidazole) and concentrated using a centricon Plus-20 unit (Millipore, Billerica, MA). Protein purity was analyzed by SDS-PAGE, and the protein concentration was determined using the Bradford micro-assay (Bio-Rad Laboratories Inc., Hercules, CA) according to the suppliers' procedure. Purified *Ef*Endo18A was stored in 20 mM Tris-HCl pH 8 at 4 °C until use. The enterococcal chitinase EF0361, cloned and purified by nickel affinity chromatography in the same way as *Ef*Endo18A (G. Vaaje-Kolstad, L.A. Bøhle, G. Mathiesen, V.G.H. Eijsink, unpublished results), was used as negative control.

### Analysis of glycosidase activity

Glycosidase activity was measured by incubating 500 μg fetuin (Sigma, St. Louis, MO), 500 μg ovalbumin (Sigma), 250 μg human IgG (Sigma), 50 μg RNaseB (NEB), 5 mg $mL^{-1}$ mucin from porcine stomach (Sigma) or 10% (v/v) defibrinated horse serum (Fisher Scientific) with 0.1 mg $mL^{-1}$ *Ef*Endo18A (~3 μM) at 37 °C in 50 mM ammonium acetate buffer pH 6 for 16 h. A 5-μL aliquot of the supernatant was mixed with 7 μL loading buffer (NuPAGE; Invitrogen) and 3 μL reducing agent (NuPAGE; Invitrogen). The protein solutions were boiled for 10 min and analyzed by SDS-PAGE. The rate of hydrolytic activity was determined by incubating 50 μg RNaseB with 25 nM *Ef*Endo18A or 25 nM EndoH from *Streptomyces plicatus* (NEB) in 50 mM ammonium acetate buffer pH 6 at 37 °C. Samples were then taken every fifth minute over a period of 30 min and analyzed by SDS-PAGE. Carbohydrates in the supernatants of the

reactions were analyzed by mass spectrometry (MS) using an Ultraflex MALDI-TOF/TOF instrument (Bruker Daltonics GmbH, Bremen, Germany) controlled by FLEXCONTROL v.3.3. For analysis with MS, 1 μL supernatant (diluted 5× in $dH_2O$) was mixed with 2 μL of a 9 mg $mL^{-1}$ solution of 2.5-dihydroxybenzoic acid in 30% acetonitrile and applied as a droplet to a MTP 384 target plate ground steel TF (Bruker Daltonics). After drying under a stream of air, mass spectra were recorded in the range from *m/z* 0–3000, and from an average of 300 laser shots with the lowest laser energy necessary to obtain sufficient signal-to-noise ratios. The following settings were used: reflectron mode with an acceleration voltage of 25 kV, reflector voltage of 26 kV and pulsed ion extraction of 40 ns in the positive ion mode. Peak lists were generated using Bruker FLEXANALYSIS software v.3.3.

## Analysis of glycolytic activity using GlcNAc oligomers, and their analogues as substrates

Possible hydrolytic activity of *Ef*Endo18A towards oligosaccharides was tested using the chito-oligomer analogues, 4-methylumbelliferyl-β-D-*N*,*N*′-diacetylchitobioside [4-MU-$(GlcNAc)_2$] and 4-methylumbelliferyl-β-D-*N*-acetylglucosamine (4MU-NAG) as substrates. A 50-μL reaction mixture contained: 0.1 mg $mL^{-1}$ bovine serum albumin (BSA), 50 μM 4-MU$(GlcNAc)_2$ or 4MU-NAG, 0–50 nM *Ef*Endo18A in 50 mM citrate phosphate buffer pH 6. After incubation at 37 °C for 10 min, the reaction was stopped by adding 1.95 mL 0.2 M $Na_2CO_3$. The amount of released 4-MU was measured using a DyNA 200 Fluorimeter (Hoefer Pharmacia Biotech, San Francisco, CA). The hydrolysis of GlcNAc oligomers was analyzed in a reaction volume of 200 μL containing 0.1 mg $mL^{-1}$ BSA, 200 μM $(GlcNAc)_4$ or $(GlcNAc)_6$ and 50 nM *Ef*Endo18A in 50 mM ammonium acetate buffer pH 6. After incubation at 37 °C overnight, the reaction was stopped by adding 1 : 1 of 20 mM $H_2SO_4$ and reaction products were analyzed using a Dionex Ultimate 3000 HPLC system set up with a Rezex column (Phenomenex, Torrance, CA). The conditions used for the HPLC analysis were: mobile phase, 5 mM $H_2SO_4$; flow rate 1 mL $min^{-1}$, detection of eluted oligosaccharides by recording absorption at 195 nm. The only detectable product, $(GlcNAc)_2$,

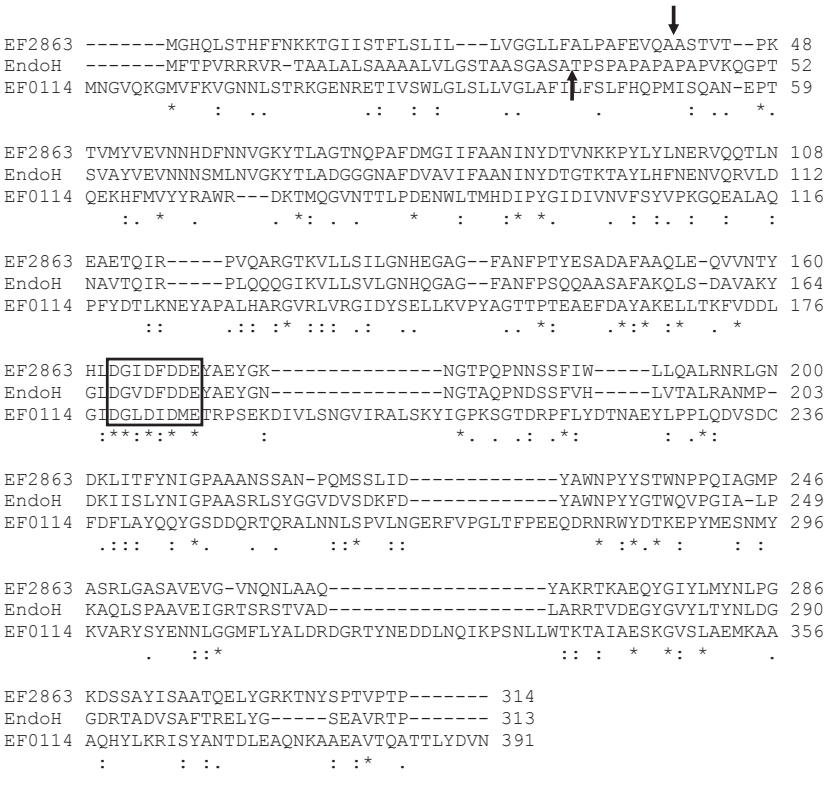



**Fig. 1.** Alignment of *Ef*Endo18A (EF2863), EndoH and the GH18 part (aa 1–391) of EF0114 using CLUSTALW. The diagnostic conserved sequence motif of GH18 enzymes (DXXDXDXE) is framed; the glutamic acid (E) in this motif serves as the proton donor in the hydrolysis. Predicted cleavage sites for signal peptidase I are indicated by arrows: *Ef*Endo18A, VQA-AS; EndoH, ASA-TP. There is no predicted cleavage site for EF0114 (see text for details). *Amino acids that are identical in *Ef*Endo18A, EndoH and EF0114. (:) Positions where the aligned amino acids are not conserved but have highly similar properties. (.) Positions where the aligned amino acids have weakly similar properties.

was quantified using external standards and the CHROMELEON 7.0 chromatography software (Dionex).

## Results and discussion

Figure 1 shows an alignment of *Ef*Endo18A with the commercial endoglycosidase EndoH from *S. plicatus* (83% sequence identity) and amino acid 1–391 of EF0114, which correspond to the GH18 part of the protein (58% sequence identity). The diagnostic conserved GH18 sequence motif DXXDXDXE, which contains important catalytic residues (Synstad *et al.*, 2004), including the glutamic acid (E) acting as catalytic acid which is present in all three proteins (Fig. 1). The strong similarity between *Ef*Endo18A and EndoH was also shown at the structural level, using MODELLER v9.4 (Eswar *et al.*, 2006) to construct a model of *Ef*Endo18A using EndoH as template (data not shown). Production of *Ef*Endo18A in *E. coli* and subsequent purification were straightforward. The procedure described in Materials and methods typically yielded 30–40 mg of highly pure protein per litre of culture.

The functionality of *Ef*Endo18A was tested using RNaseB, human IgG, fetuin, ovalbumin, mucin and serum proteins as substrates. Possible deglycosylation of these substrates was analyzed by looking at band shifts in SDS-PAGE gels and by analyzing released carbohydrates using MALDI-TOF MS. RNaseB is a well known substrate for this type of study; it has one N-linked glycosylation site containing glycans of the high-mannose type (Fu *et al.*, 1994). Ovalbumin is N-glycosylated at one site that may carry both high-mannose and hybrid-type glycans (Nisbet *et al.*, 1981; An *et al.*, 2003). Human IgG, fetuin and serum proteins contain N-linked glycans of the complex type as well as O-linked glycans (Spiro, 2002; Chu *et al.*, 2009), whereas mucin mainly contains O-linked glycans (Bansil & Turner, 2006). Figure 2 shows clear band shifts for RNaseB and ovalbumin, one shift for the former and two shifts for the latter, indicating hydrolysis of N-linked glycans of the high mannose and possibly also the hybrid type (see below). The absence of band-shifts after *Ef*Endo18A-treatment of fetuin, human IgG, mucin and serum proteins indicates that *Ef*Endo18A is unable to hydrolyze O-linked glycans and N-linked glycans of the complex type. It is well known that enterococci cannot release O-linked glycans from mucin (Hoskins *et al.*, 1985; Corfield *et al.*, 1992).

To verify that the observed band shifts are the result of hydrolysis of glycans, the supernatants were analyzed using MALDI-TOF MS. The spectra for RNase B (Fig. 3a and b) showed four signals with masses corresponding to 1 GlcNAc plus five to eight mannoses. This corresponds to four of the known glycoforms of the glycans of RNaseB, in which between two and five mannose residues can be coupled to the core $GlcNAc_2Man_3$ pentasaccharide

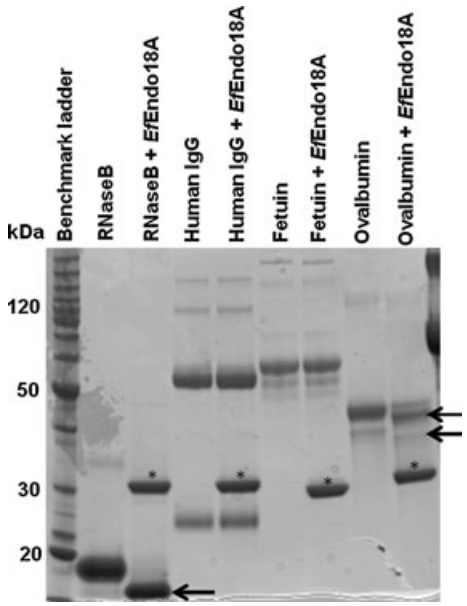



**Fig. 2.** SDS-PAGE analysis after incubation of different glycoproteins with *Ef*Endo18A. The arrows indicate bands that appear after *Ef*Endo18A treatment and that have shifted position relative to the non-treated protein visible in the adjacent gel lane. *The band representing purified *Ef*Endo18A. Note that the size marker proteins are non-glycosylated, whereas the analyzed proteins contain glycans. Experiments with mucin and serum proteins yielded more smeared band patterns that were identical for samples incubated with and without *Ef*Endo18A.

(Fu *et al.*, 1994). The MALDI-TOF MS spectrum of released oligosaccharides from ovalbumin (Fig. 3c and d) showed three clear signals corresponding to $GlcNAcMan_5$, $GlcNAcMan_6$ and $GlcNAcMan_5HexNAc_2$ (Fig. 3c and d). This is in accordance with the notion that both high-mannose and hybrid type glycans are linked to the N-glycosylation site of ovalbumin (An *et al.*, 2003). No such glycan signals were observed in the supernatants of *Ef*Endo18A-treated human IgG, fetuin, mucin or serum, confirming the absence of hydrolysis by *Ef*Endo18A (data not shown).

To determine whether *Ef*Endo18A could hydrolyze GlcNAc oligomers in the absence of any protein and links to other sugars, *Ef*Endo18A was incubated with 4MU-GlcNAc, 4MU-$(GlcNAc)_2$ and different GlcNAc oligomers under conditions that would lead to massive substrate conversion if *Ef*Endo18A were a chitinase such as the enterococcal chitinase EF0361 (G. Vaaje-Kolstad, L.A. Bøhle, G. Mathiesen, V.G.H. Eijsink, unpublished results). *Ef*Endo18A did not release 4MU from the fluorogenic substrates, but showed a low but significant activity towards $(GlcNAc)_4$ and $(GlcNAc)_6$. After overnight incubation, about 0.1% of the substrate was converted, whereas chitinases such as EF0361 (G. Vaaje-Kolstad, L.A. Bøhle,

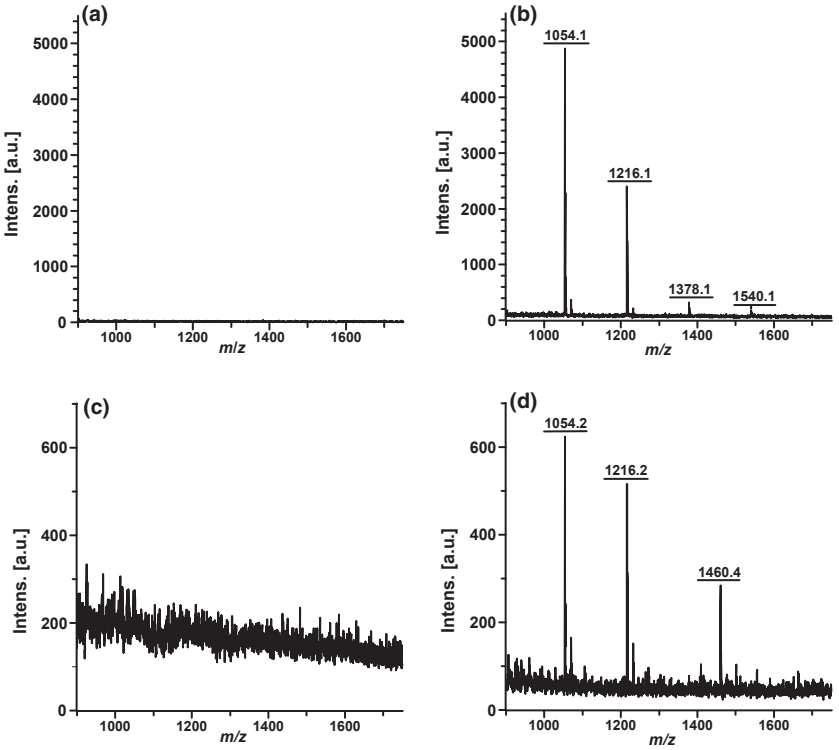



**Fig. 3.** MALDI-TOF MS spectra showing released oligosaccharides. RNaseB was incubated in the absence (a) or the presence (b) of *Ef*Endo18A. The high-mannose glycan on RNaseB occurs in four different glycoforms with varying numbers of attached mannose residues to the core pentasaccharide: GlcNAc-$Man_5$ (1054.1 Da), GlcNAc-$Man_6$ (1216.1 Da), GlcNAc-$Man_7$ (1378.1 Da), GlcNAc-$Man_8$ (1540.1 Da). The mass difference between each glycoform is 162 Da, which corresponds to one mannose residue; the calculated total masses are as follows: 203 (for GlcNac) + $n \times 162$ (for mannose) + 18 (for water) + 23 (for sodium), with $n$ varying from five to eight for the annotated peaks. Ovalbumin, known to contain both high-mannose and hybrid type glycans, was incubated in the absence (c) or presence (d) of *Ef*Endo18A. Three clear peaks appeared after *Ef*Endo18A treatment, two (1054.2 Da, 1216.2 Da) corresponding to the high-mannose type, and one (1460.4 Da) corresponding to a hybrid-type of glycan comprising five hexoses of 162 Da and three HexNAc of 203 Da. Similar MS analyses of reaction mixtures obtained after incubating human IgG, fetuin, mucin and serum with *Ef*Endo18A under identical conditions did not yield any oligosaccharide signals.

G. Mathiesen, V.G.H. Eijsink, unpublished results) or for example the family 18 chitinases from *Serratia marcescens* (Horn *et al.*, 2006) would convert most of the substrate under these conditions. The only detectable product was $(GlcNAc)_2$. This indicates that *Ef* Endo18A is not a chitinase and that its glycosidase activity depends on the scissile GlcNAc-GlcNAc being linked to a protein. Likewise, control experiments with various family 18 chitinases, including the enterococcal EF0361 cloned and purified in the same way as *Ef* Endo18A, did not release glycans from RNase B.

In agreement with results obtained for other endoglycosidases, the present data show that *Ef* Endo18A hydrolyzes the glycosidic bond of the *N*,*N*′-diacetylchitobiose core structure which is N-linked to asparagine. After hydrolysis, one GlcNAc residue remains attached to the protein and the other GlcNAc is released with the rest of the oligosaccharide. The activities of *Ef* Endo18A and its close relative EndoH (Tarentino & Maley, 1974) are limited to the high mannose and hybrid glycans occurring in RNaseB and ovalbumin. There exist GH18 endoglycosidases that act on complex N-linked glycans and that deglycosylate protein such as IgG. However, these endoglycosidases are multi-domain proteins and it has been shown that the additional domains are essential for the deglycosylating activity on IgG (Collin & Olsen, 2001; Collin & Fischetti, 2004).

To compare the rate of glycan hydrolysis by *Ef* Endo18A and EndoH, RNaseB was used as a substrate. Figure 4 shows that EndoH and *Ef* Endo18A hydrolyze RNaseB at similar rates. Both enzymes, at a concentration of 25 nM, were able to hydrolyze the glycans in 50 μg RNaseB within 20 min.

## Concluding remarks

So far, the ability of *E. faecalis* to release high-mannose glycans from glycoproteins (Roberts *et al.*, 2000, 2001) has been linked to EndoE/EF0144 (Collin & Fischetti,

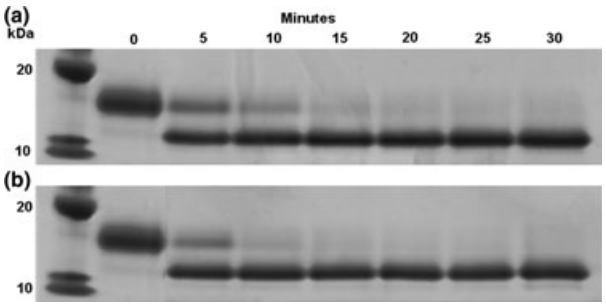



**Fig. 4.** Rate of RNaseB hydrolysis. RNaseB 50 μg was hydrolyzed with 25 nM *Ef*Endo18A (a) or 25 nM EndoH (b) in 50 μL 50 mM ammonium acetate buffer pH 6 at 37 °C over a period of 30 min. The gels show samples taken over time.

2004). However, although the activity of recombinantly produced EndoE/EF0144 is well documented (Collin & Fischetti, 2004), there is to the best of our knowledge no hard evidence justifying the claim that the observed endo-β-*N*-acetylglucosaminidase activity in supernatants of *E. faecalis* is due (solely) to this protein. The present data show that EF2863 is another source of endo-β-*N*-acetylglucosaminidase activity, with narrower substrate specificity (N-linked high-mannose and hybrid type glycans) than the two-domain EndoE/EF0144 (all N-linked glycans). Interestingly, our own predictions of enzyme localization using SIGNALP 3.0 (Bendtsen *et al.*, 2004) and LIPOP v. 1.0 (Juncker *et al.*, 2003), as well as the LOCATEP database (Zhou *et al.*, 2008) indicate that EF2863 is a secreted protein, whereas the leader peptide of EF0114 seems to have no signal peptidase I cleavage site, meaning that this protein may be N-terminally anchored to the cell membrane. Different localization of the two endoglycosidases may reflect different physiological roles.

Proteins with high-mannose N-linked glycans are frequently found in human glycoproteins (Fujiwara *et al.*, 1988, Furukawa *et al.*, 1989). Even though the release of nutrients from these glycoproteins seems to be a physiologically important role of enzymes such as *Ef*Endo18A, one may speculate about additional physiological roles such as modulation of the host immune system. Interestingly, it has been shown that *Ef*Endo18A from *E. faecalis* V583 is up-regulated in blood and urine (Vebo *et al.*, 2009, 2010), where *E. faecalis* frequently causes infection. The prevalence of endoglycosidases that exploit, alter or inactivate host glycoproteins may give pathogenic bacteria an advantage during infection.

## Acknowledgements

This work was supported by grant 183637/S10 from the Research Council of Norway. We thank Britt Dahl for technical assistance during the cloning experiments.

## References

An HJ, Peavy TR, Hedrick JL & Lebrilla CB (2003) Determination of N-glycosylation sites and site heterogeneity in glycoproteins. *Anal Chem* **75**: 5628–5637.

Apweiler R, Hermjakob H & Sharon N (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* **1473**: 4–8.

Bansil R& Turner BS (2006) Mucin structure, aggregation, physiological functions and biomedical applications. *Curr Opin Colloid Interface Sci* **11**: 164–170.

Bendtsen JD, Nielsen H, von Heijne G & Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**: 783–795.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V & Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* **37**: D233–D238.

Chu CS, Ninonuevo MR, Clowers BH, Perkins PD, An HJ, Yin H, Killeen K, Miyamoto S, Grimm R & Lebrilla CB (2009) Profile of native N-linked glycan structures from human serum using high performance liquid chromatography on a microfluidic chip and time-of-flight mass spectrometry. *Proteomics* **9**: 1939–1951.

Collin M & Fischetti VA (2004) A novel secreted endoglycosidase from *Enterococcus faecalis* with activity on human immunoglobulin G and ribonuclease B. *J Biol Chem* **279**: 22558–22570.

Collin M & Olsen A (2001) EndoS, a novel secreted protein from *Streptococcus pyogenes* with endoglycosidase activity on human IgG. *EMBO J* **20**: 3046–3055.

Collin M & Olsen A (2003) Extracellular enzymes with immunomodulating activities: variations on a theme in *Streptococcus pyogenes*. *Infect Immun* **71**: 2983–2992.

Corfield AP, Wagner SA, Clamp JR, Kriaris MS & Hoskins LC (1992) Mucin degradation in the human colon: production of sialidase, sialate O-acetylesterase, N-acetylneuraminate lyase, arylesterase, and glycosulfatase activities by strains of fecal bacteria. *Infect Immun* **60**: 3971–3978.

Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U & Sali A (2006) Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics*, Ch. 5, Unit 5 6.

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

Fu D, Chen L & O'Neill RA (1994) A detailed structural characterization of ribonuclease B oligosaccharides by 1H NMR spectroscopy and mass spectrometry. *Carbohydr Res* **261**: 173–186.

Fujiwara S, Shinkai H, Deutzmann R, Paulsson M & Timpl R (1988) Structure and distribution of N-linked oligosaccharide chains on various domains of mouse tumour laminin. *Biochem J* **252**: 453–461.

Furukawa K, Roberts DD, Endo T & Kobata A (1989) Structural study of the sugar chains of human platelet thrombospondin. *Arch Biochem Biophys* **270**: 302–312.

Henrissat B, Teeri TT & Warren RA (1998) A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. *FEBS Lett* **425**: 352–354.

Horn SJ, Sikorski P, Cederkvist JB, Vaaje-Kolstad G, Sorlie M, Synstad B, Vriend G, Varum KM & Eijsink VG (2006) Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *P Natl Acad Sci USA* **103**: 18089–18094.

Hoskins LC, Agustines M, McKee WB, Boulding ET, Kriaris M & Niedermeyer G (1985) Mucin degradation in human colon ecosystems. Isolation and properties of fecal strains that degrade ABH blood group antigens and oligosaccharides from mucin glycoproteins. *J Clin Invest* **75**: 944–953.

Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H & Krogh A (2003) Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* **12**: 1652–1662.

Leisner JJ, Larsen MH, Ingmer H, Petersen BO, Duus JO & Palcic MM (2009) Cloning and comparison of phylogenetically related chitinases from *Listeria monocytogenes* EGD and *Enterococcus faecalis* V583. *J Appl Microbiol* **6**: 2080–2087.

Morelle W & Michalski JC (2005) The mass spectrometric analysis of glycoproteins and their glycan structures. *Curr Anal Chem* **1**: 29–57.

Morelle W & Michalski JC (2007) Analysis of protein glycosylation by mass spectrometry. *Nat Protoc* **2**: 1585–1602.

Nisbet AD, Saundry RH, Moir AJ, Fothergill LA & Fothergill JE (1981) The complete amino-acid sequence of hen ovalbumin. *Eur J Biochem* **115**: 335–345.

Paulsen IT, Banerjei L, Myers GS *et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**: 2071–2074.

Roberts G, Tarelli E, Homer KA, Philpott-Howard J & Beighton D (2000) Production of an endo-beta-N-acetylglucosaminidase activity mediates growth of *Enterococcus faecalis* on a high-mannose-type glycoprotein. *J Bacteriol* **182**: 882–890.

Roberts G, Homer KA, Tarelli E, Philpott-Howard J, Devriese LA & Beighton D (2001) Distribution of endo-beta-*N*-acetylglucosaminidase amongst enterococci. *J Med Microbiol* **50**: 620–626.

Ruiz L, Gueimonde M, Coute Y, Salminen S, Sanchez JC, de los Reyes-Gavilan CG & Margolles A (2011) Evaluation of the ability of *Bifidobacterium longum* to metabolize human intestinal mucus. *FEMS Microbiol Lett* **314**: 125–130.

Sahm DF, Kissinger J, Gilmore MS, Murray PR, Mulder R, Solliday J & Clarke B (1989) *In vitro* susceptibility studies of vancomycin-resistant *Enterococcus faecalis*. *Antimicrob Agents Chemother* **33**: 1588–1591.

Sanchez B, Saad N, Schmitter JM, Bressollier P & Urdaci MC (2010) Adhesive properties, extracellular protein production, and metabolism in the *Lactobacillus rhamnosus* GG strain when grown in the presence of mucin. *J Microbiol Biotechnol* **20**: 978–984.

Spiro RG (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**: 43R–56R.

Synstad B, Gaseidnes S, Van Aalten DM, Vriend G, Nielsen JE & Eijsink VG (2004) Mutational and computational analysis of the role of conserved residues in the active site of a family 18 chitinase. *Eur J Biochem* **271**: 253–262.

Tarentino AL & Maley F (1974) Purification and properties of an endo-beta-N-acetylglucosaminidase from *Streptomyces griseus*. *J Biol Chem* **249**: 811–817.

Tendolkar PM, Baghdayan AS & Shankar N (2003) Pathogenic enterococci: new developments in the 21st century. *Cell Mol Life Sci* **60**: 2622–2636.

Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sorlie M & Eijsink VG (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330**: 219–222.

Vebo HC, Snipen L, Nes IF & Brede DA (2009) The transcriptome of the nosocomial pathogen *Enterococcus faecalis* V583 reveals adaptive responses to growth in blood. *PLoS ONE* **4**: e7660.

Vebo HC, Solheim M, Snipen L, Nes IF & Brede DA (2010) Comparative genomic analysis of pathogenic and probiotic *Enterococcus faecalis* isolates, and their transcriptional responses to growth in human urine. *PLoS ONE* **5**: e12489.

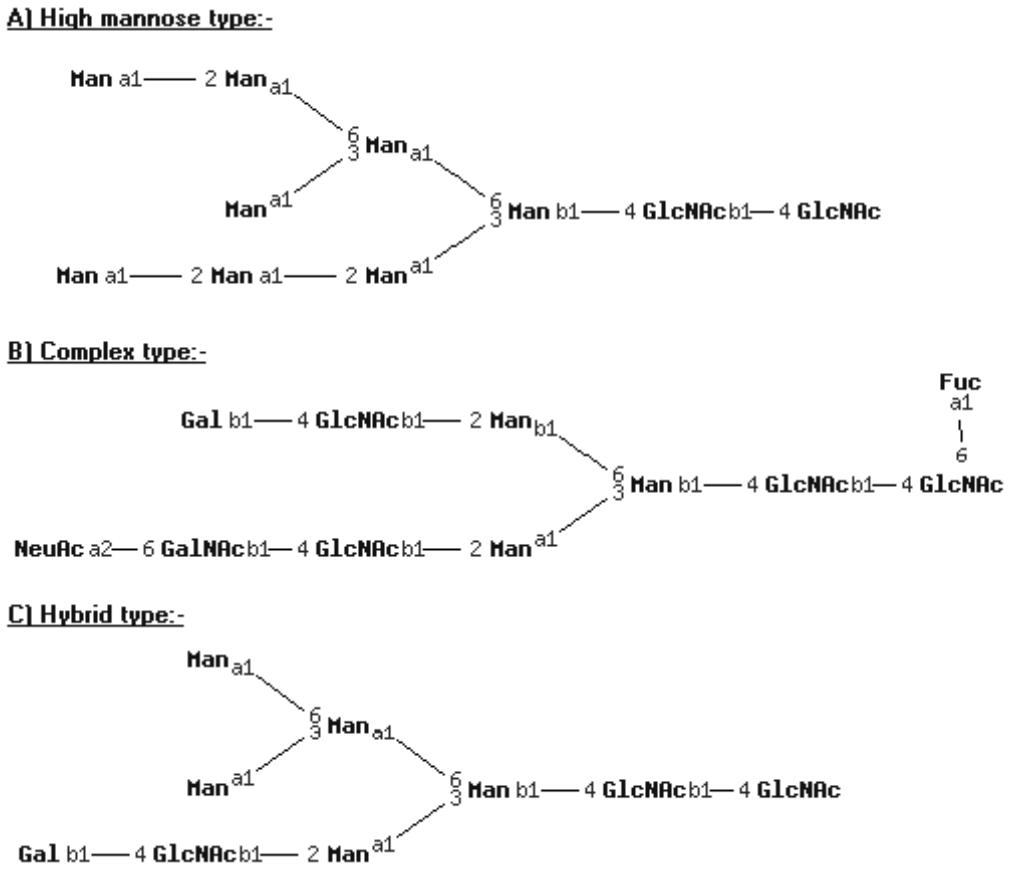
Zhou M, Boekhorst J, Francke C & Siezen RJ (2008) LocateP: genome-scale subcellular-location predictor for bacterial proteins. *BMC Bioinformatics* **9**: 173.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Fig. S1.** Overview of N-linked glycan types.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.




**Supplementary Figure 1.** Overview of N-linked glycan types. N-linked glycans are normally subgrouped into three types: (A) the high-mannose type, (B) the complex type, and (C) the hybrid type. The glycans are linked to an asparagine in the protein via the GlcNac unit to the right. Common to all N-linked glycans is the pentasaccharide core of two *N*-acetylglucosamines (GlcNAc), and three mannoses (Man). Note that the branches pointing to the left may be longer than indicated in the Figure. Adapted from http://au.expasy.org/cgi-bin/glycomod.pl

# PAPER IV

# Characterization of the chitinolytic machinery of *Enterococcus faecalis* V583 and high resolution structure of its oxidative CBM33 enzyme

Gustav Vaaje-Kolstad*, Liv Anette Bøhle, Sigrid Gåseidnes, Bjørn Dalhus[1], Magnar Bjørås[1], Geir Mathiesen and Vincent G.H. Eijsink

Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P. O. Box 5003, N-1432 Ås, Norway

[1]Department of Medical Microbiology and Department of Medical Biochemistry, Oslo University Hospital, National Hospital, P. O. Box 4950, Nydalen, N-0424 Oslo, Norway.

*Address correspondence to: Gustav Vaaje-Kolstad, Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P. O. Box 5003, N-1432 Ås, Norway. Telephone: +47 64965905, Fax: +47 64965901. Email: gustko@umb.no

ABSTRACT

Little information exists for the ability of enterococci to utilize chitin as a carbon source. We show that *Enterococcus faecalis* V583 can grow on β-chitin as result of an operon encoding a family 18 chitinase (*ef0361; Ef*Chi18A) and a family 33 carbohydrate binding module (*ef0362; Ef*CBM33A). Various types of enzyme activity assays showed that *Ef*Chi18A has functional properties characteristic of an endochitinase. *Ef*CBM33A belongs to a recently discovered family of enzymes that cleave glycosidic bonds via an oxidative mechanism and that act synergistically with classical hydrolytic enzymes (i.e. chitinases). The structure and function of this protein were probed in detail. An ultra-high resolution crystal structure of *Ef*CBM33A revealed details of a conserved binding surface that is optimized to interact with chitin and contains the catalytic center. Chromatography and mass spectrometry analyses of product formation showed that *Ef*CBM33A cleaves chitin via the oxidative mechanism previously described for CBP21 from *Serratia marcescens*. In the presence of an external electron donor *Ef*CBM33A boosted the activity of *Ef*Chi18A and combining the two enzymes led to rapid and complete conversion of β-chitin to chitobiose. This study provides insight into the structure and function of the CBM33 family of enzymes, which, together with their

fungal counterpart called GH61, currently receive considerable attention in the biomass processing field.

Keywords: CBM33, GH61, chitin, *Enterococcus faecalis*, chitinase, oxidation

DP: degree of polymerization, CBM: carbohydrate binding module, GH: glycoside hydrolase, GlcNAc: *N*-acetylglucosamine, GlcNAcA: *N*-acetylglucosamine aldonic acid, MALDI-TOF MS: Matrix-assisted laser desorption/ionization time of flight mass spectrometry.

INTRODUCTION

Enterococci are Gram-positive cocci that occur in several ecological niches, including the gut of mammals. These bacteria are normally considered harmless or even beneficial for humans, but in recent years enterococci such as *Enterococcus faecalis* commonly found in humans, have emerged as opportunistic pathogens. Enterococci easily spread in water where the bacteria can attach to chitinous surfaces of chitin containing plankton or copepods,[1, 2] a trait also observed for other pathogenic bacteria.[3] All currently available enterococcal genomes contain what appear to be minimal enzymatic systems for utilization of chitin as a carbon source. This system includes a family 18 chitinase, a family 20 *N*-acetylhexosaminidase, a family 33 carbohydrate binding module (CBM33), an *N*-acetylglucosamine-specific phosphotransferase sugar transporter and enzymes required for the deacetylation (N-acetylglucosamine-6-phosphate deacetylase) and deamination (glucosamine-6-phosphate isomerase) of the *N*-acetylated sugar that is required prior to downstream processing in the glycolytic pathway.[4] By analogy with known microbial chitinolytic systems, the GH18 chitinase and the CBM33 protein (see Ref. 5 for family nomenclature) are expected to be responsible for depolymerizing the insoluble chitinous substrate to soluble oligomeric products where $GlcNAc_2$ is the dominant product (usually >95%). The GH20 *N*-acetylhexosaminidase (or chitobiase) cleaves $GlcNAc_2$ to GlcNAc which is then taken up by the bacterium. *E. faecalis* harbours genes putatively encoding the secreted GH18 and CBM33 proteins organized in a predicted operon. This operon comprises *ef0361* and *ef0362,* both encoding single-domain proteins hereafter referred to as *Ef*Chi18A and *Ef*CBM33A, respectively. The GH20 hexosaminidase is putatively encoded by *ef0114*.

Little is known about the ability of enterococci to utilize chitin as a carbon source, but recently it was shown that the *ef0361* gene indeed encodes an active chitinase.[6] The presence of a CBM33 is of particular interest due to the recent discovery that some members of this protein family in fact are enzymes that cleave glycosidic bonds.[7] Originally, proteins from the CBM33 family were thought to be involved in substrate recognition because they are secreted in abundance during microbial chitin degradation, bind to chitin, and have no detectable hydrolytic activity.[8, 9, 10] However, the ability of a CBM33 (CBP21 from *S. marcescence*) to substatially boost the degradation of chitin by chitinases demonstrated that CBM33s rather has a role in chitin degradation.[11, 12] The nature of this activity remained enigmatic until a recent study showed that CBP21 could cleave chitin chains in their ordered/ crystalline state

through a mechanism involving both hydrolysis and oxidation.[7] A follow-up study revealed that some members of the CBM33 family are able to cleave crystalline cellulose[13] indicating that CBM33 enzymes have evolved to also act on insoluble substrates other than chitin. It also seems that structurally similar fungal proteins classified as GH61 may have similar activities as the (bacterial) CBM33s.[7, 13, 14, 15]

The catalytic mechanism of these novel enzymes has not yet been resolved in detail, but it is evident that their activity depends on the presence of a divalent metal ion which is thought to bind to a fully conserved structural motif on the substrate binding surface of the proteins (see Refs 7 and 14 and below). Furthermore, and most importantly from a biotechnological point of view, it has been shown that the addition of external electron donors substantially amplifies the activity of CBM33 enzymes and thereby increase the synergy with "classic" hydrolases in the depolymerisation of insoluble substrates.

We have studied the chitinolytic system of *E. faecalis* V583 in more detail, with particular focus on the role and structure of the CBM33. Both cultivation studies and functional studies of recombinantly produced *Ef*Chi18A and *Ef*CBM33A clearly show that *E. faecalis* is capable of degrading and metabolizing chitin. Deeper analysis of the chitinolytic system showed that combination of the chitinase (*Ef*Chi18A) and *Ef*CBM33A leads to rapid depolymerization of crystalline chitin, especially with β-chitin as a substrate. The catalytic action of *Ef*CBM33A was investigated in detail, showing that this protein acts by oxidative cleavage of glycosidic bonds, analogous to what has been found for CBP21. Finally we present an ultra high resolution crystal structure of *Ef*CBM33A that provides further insight into structural features of the CBM33 family and that allows detailed comparisons with CBP21, the only other CBM33 family member with a known crystal structure.

RESULTS

## General microbiology, cloning and protein expression

### *Cultivation of E. faecalis V583 on chitinous substrates*

Growth of *E. faecalis* was assessed using soluble sugars resulting from enzymatic chitin hydrolysis as carbon sources (GlcNAc or $GlcNAc_2$). In terms of growth rate in the exponential phase and the cell densities that were reached, the growth curves resembled those obtained when growing the bacterium on a rich medium such as BHI. However, cultures growing on GlcNAc or $GlcNAc_2$ showed an extended lag phase before reaching exponential growth relative to cultures grown on BHI, especially for $GlcNAc_2$ (where there also was a delay compared to growth on minimal medium grown on glucose; Fig. 1A). Supernatants from cultures grown in minimal media supplemented with various soluble carbon sources all showed substantial β-hexosaminidase/ chitobiase activity, whereas essentially no background activity was observed in the uninoculated media that served as a control. The highest activity was observed in cultures with $GlcNAc_2$ as a carbon source (Fig. 1B).

To determine whether *E. faecalis* V583 was able to utilise insoluble chitin as carbon source the bacterium was cultivated in minimal media containing β-, α-chitin or cellulose (Fig. 1C). Substantial chitinolytic activity was detected after 24 hours in media containing β-chitin, whereas for α-chitin only weak chitinolytic activity could be detected and only after prolonged cultivation (48 hours for α-chitin). Chitinolytic activity could not be detected in uninoculated culture medium or in cultures grown on glucose or cellulose. The presence of substantially higher chitinolytic activity in cultures with β-chitin correlated with higher growth rates as detected by measurement of colony forming units (CFU) in the cultures (Fig. 1C; CFU counts). In fact, under these conditions, growth on β-chitin was comparable to growth on glucose.

## Characterization of *Ef*Chi18A

### *Chitinolytic activity and subsite binding preferences of EfChi18A*

Purified recombinant *Ef*Chi18A showed activity towards soluble chitooligosaccharides and α- and β-chitin, as exemplified below.

To further explore the substrate binding and hydrolyzing properties of *Ef*Chi18A, initial reaction rates towards chitooligosaccharides (DP3-6) were determined in combination with analysis of initial reaction products. The results showed that in the DP range of three to

six the reaction rate is strongly correlated with the DP of the substrate (initial rates: $GlcNAc_3$ = 0.6 ± 0.03 $s^{-1}$, $GlcNAc_4$ = 13.9 ± 0.1 $s^{-1}$ and $GlcNAc_6$ = 79.3 ± 0.4 $s^{-1}$).

Binding preferences for the various chitooligosaccharides were estimated by analysis and quantification of initial reaction products prior to these products reaching their standard equilibrium α:β-anomer ratio (Table 1 and Fig. 2). Hydrolysis of $GlcNAc_3$ resulted in the production of $GlcNAc_2$ and GlcNAc (GlcNAc could not be directly observed in the HPLC method used). $GlcNAc_2$ was found to occur in the β-anomeric form only, indicating that productive binding of $GlcNAc_3$ always occurs in the -2 to +1 subsites. Hydrolysis of $GlcNAc_4$ resulted in the production of $GlcNAc_2$ with 80% β-anomers, indicating binding from -2 to +2, and $GlcNAc_3$ and GlcNAc, which indicates binding from -1 to +3 or -3 to +1. The $GlcNAc_3$ product had an anomer ratio close to that of the standard, indicating that this productive binding mode exclusively involves subsites -1 to +3. Hydrolysis of $GlcNAc_5$ gave almost exclusively $GlcNAc_3$ and $GlcNAc_2$ as products. The $GlcNAc_2$ product was dominated by the β–anomeric form (97%), suggesting binding of $GlcNAc_5$ from the -2 to the +3 subsites. Some productive binding involving the -1 to +4 subsites was also observed. Hydrolysis of $GlcNAc_6$ resulted in products ranging from $GlcNAc_5$ to GlcNAc. $GlcNAc_5$ was only present in small amounts (~5% of the total reaction products) and $GlcNAc_5$ showed only a small increase in the fraction of β-anomers compared to the standard; this indicates productive binding to subsites -1 to +5. The $GlcNAc_2$ product was exclusively found in the β-anomeric form, which is indicative of binding to subsites -2 to +4. $GlcNAc_3$ was also observed, suggesting productive binding to the -3 to +3 subsites. Clearly, analysis of hexamer degradation is subject to minor errors because the longer products that are being formed are substrates themselves. However, since experiments were conducted under conditions that led to less than 30 % of the substrate being converted and since the hexamer is a considerably better substrate than the shorter oligomers (see above), it seems safe to assume that the present data for hexamer degradation provide a reasonably reliable view of preferred productive binding modes.

**Characterization of *Ef*CBM33A**

*Atomic structure of EfCBM33A at 0.95 Å resolution*

The structure determination of *Ef*CBM33A at 0.95 Å resolution revealed electron density for all 166 residues in the protein, as well as 286 solvent water molecules. The fully refined structure had R-factors converging to $R_{work}$ = 0.09 and $R_{free}$ = 0.11 (Table 1). The atomic

coordinates and structure factors have been deposited in the Protein Data Bank under accession code **4A02**. The structural fold of *Ef*CBM33A is a distorted β-sandwich comprised of a three- and a four-stranded β-sheet (Fig. 3A). A distinct 72 residue protuberance of mainly small α-helices and loops positioned between β-strands 1 and 2 stretches out on the side of the β-sandwich (Fig. 3A). CBM33s contain several highly conserved aromatic residues in the interior of the protein[12] and this is also the case for *Ef*CBM33A (Fig. 3B). Like in CBP21, only one aromatic amino acid is solvent exposed (Trp58, corresponding to Tyr54 in CBP21; Fig. 3C and below).

*Similarity to other structures*

Comparison to 3D protein structures available in the PDB database using the Dali server[16] shows highest structural similarity to CBP21 (Z-score 29.2 and RMSD 0.9 Å for all back bone atoms), the only other 3D structure available for CBM33 proteins. The following most similar proteins are fungal enzymes from the GH61 family *Tt*GH61E (Z-score 9.2 and RMSD 2.9 Å for all back bone atoms) and *Hj*GH61B (Z-score 9.0 and RMSD 3.2 Å for all back bone atoms).

Detailed comparison of *Ef*CBM33A and CBP21 shows high overall similarity (Fig. 3A), but some noticeable differences: The solvent exposed and functionally important Tyr54 in CBP21[11, 12] is substituted with a tryptophan in *Ef*CBM33A (Trp58; Fig. 3C). Otherwise all conserved solvent exposed residues on the substrate binding surface (as described in Refs 11 and 12), are also conserved on the *Ef*CBM33A surface (Fig. 3C). The conserved metal binding motif of the active site (Figs 3D&E), comprising His29, the N-terminal amino group, and His114 is not binding a metal in *Ef*CBM33A (Fig. 3D). Instead, these groups are forming hydrogen bonds to other highly conserved amino acids in the active site (the backbone of Ala112 and the side chain of Asp180, respectively; the binding of metals is discussed more extensively below). Interestingly, *Ef*CBM33A lacks two disulphide bridges that are conserved in several members of the CBM33 family, including CBP21 (the corresponding residue pairs in *Ef*CBM33A are Gly42-Val53 and Ala143-Ile160).

*Apparent melting points ($T_m^{app}$) of EfCBM33A and CBP21*

Lack of two disulphide bridges in *Ef*CBM33A prompted the question whether this would have consequences for protein stability. Determination of the apparent melting points (Fig. 4) showed that *Ef*CBM33A in fact is slightly more stable ($T_m^{app}$ = 72.0°C) than CBP21 ($T_m^{app}$ =

70.3°C). In the presence of DTT, the apparent melting point of the disulphide containing CBP21 was reduced by ~6 °C (Fig. 4), whereas *Ef*CBM33A was less affected ($\Delta T_m^{app}$ = -1.5°C). These results indicate that the disulphide bridges do contribute to the stability of CBP21 and that the lack of such disulphide bonds in *Ef*CBM33A is compensated for by other stabilizing interactions.

*Binding studies*

The binding preferences of *Ef*CBM33A were probed by substrate binding assays using α-chitin, β-chitin and cellulose as substrates. The data show that *Ef*CBM33A binds both α- and β-chitin and that slightly more protein binds to β-chitin (Fig. 5). Since family 33 CBMs have a metal binding site on the substrate binding surface, the effects of a chelating agent (5 mM EDTA) on binding was also tested. Binding to α-chitin seemed to be slightly affected by EDTA, whereas binding to β-chitin was unaffected.

*Oxidative cleavage of crystalline chitin*

In order to explore the ability of *Ef*CBM33A to cleave chitin the enzyme was incubated with α- or β-chitin in the presence of an external electron donor (reduced glutathione or ascorbic acid). MALDI-TOF MS analysis of soluble products revealed that *Ef*CBM33A generates chitooligosaccharide aldonic acids (i.e. oxidized oligosaccharides) from both α- and β-chitin (Fig. 6A). Spectra obtained for both substrates show the same striking dominance of even numbered products as previously observed for CBP21 (Fig. 6A and Ref. 7), suggesting that the reaction occurs on the crystalline/ ordered parts of the substrate (see Discussion for further explanation). Soluble products were also analyzed using an UHPLC method and standards previously developed for CBP21. The results (Fig. 6B) underpin the dominance of even-numbered products and also show that in this initial phase of the reaction, product concentrations increase linearly with time. The reactions with *Ef*CBM33A did not yield detectable amounts of non-oxidized soluble products.

*Degradation of α- and β-chitin by EfChi18A in the presence or absence of EfCBM33A*

The chitin degrading efficiency of *Ef*Chi18A was estimated in the absence and presence of *Ef*CBM33A and/or an external electron donor (reduced glutathione). The chitinase was able to convert both substrates to soluble sugars (mainly $GlcNAc_2$, plus minor amounts of GlcNAc amounting to <5% of the total amount of released sugar). Fig. 7 shows that the rate of degradation increased in the presence of *Ef*CBM33A, especially in the presence of an external

electron donor. Combining both enzymes and the external electron donor led to total solubilisation of β-chitin within ~24 hours, whereas only ~10% of the α-chitin was solubilised under the same conditions. The reactions shown in Fig. 7 were sampled for up to three weeks following the 24 h time point. All reactions containing β-chitin reached complete solubilisation after about 2 weeks of incubation, whereas none of the samples containing α-chitin were degraded beyond ~10% (results not shown).

*The influence of divalent cations on activity*

The importance of divalent cations for the activity of family 33 CBMs and the functionally related GH61 proteins has been demonstrated in several studies,[7, 13, 14] but so far, the identity of the metal involved in the reaction has remained elusive. In an attempt to shed light on this ambiguity, *Ef*CBM33A was produced in a metal free form (*Ef*CBM33A$_{apo}$) and enzyme activity was investigated in the presence or absence of various metals. Metal concentrations used were much lower (3 μM) than in previous studies (typically 1 mM) to reduce the risk for artefacts. The results show that the metal-free *Ef*CBM33A has lost almost all activity and that several bivalent metals can re-activate the enzyme, in particular $Co^{2+}$, $Zn^{2+}$, and $Fe^{3+}/Fe^{2+}$.

DISCUSSION

The present data show that *E. faecalis* grows well on β-chitin and products from its degradation, GlcNAc and $GlcNAc_2$ (Fig. 1A&B). Although knock-out studies are needed to confirm this, it seems highly likely that the ability to grow on chitin is due to the production of *Ef*Chi18A and *Ef*CBM33A, which together allow rapid conversion of insoluble chitin, especially β-chitin, to soluble sugars. The only essential component of the chitinolytic system not cloned and characterized in this study is the GH20 enzyme responsible for converting $GlcNAc_2$ to GlcNAc, called *N*-acetyl hexosaminidase or chitobiase (E.C. 3.2.1.52). However, the presence of such an activity was clearly demonstrated by growth of *E. faecalis* on $GlcNAc_2$ as a sole carbon source, combined with detection of substantial *N*-acetyl hexosaminidase activity in the culture supernatants. The genome sequence of *E. faecalis* V583 contains one GH20 module that occurs as part of a bimodular enzyme. This enzyme also contains a GH18 module and it has been shown that both catalytic modules possess endoglycosidase activity directed at cleaving off glycans from N-glycosylated proteins.[17] Thus, this GH20 may not be involved in chitin metabolism. The genome of *E. faecalis* V583 also contains a gene in the GH3 family (*ef1238*) of which several members have been shown to exhibit *N*-acetyl hexosaminidase activity. While most *N*-acetyl hexosaminidases in this family seem to have functions in peptidoglycan recycling[18] there is at least one example of a GH3 from a Gram positive bacterium having chitobiase activity related to chitin metabolism.[19] More studies on these enzymes are needed in order to determine the enzyme responsible for hydrolyzing $GlcNAc_2$ in *E. faecalis* V583.

Judged by its sequence, *Ef*Chi18A resembles the well studied endo-acting non-processive chitinases from *S. marcescens* (ChiC; 59% sequence identity)[20] and *Lactococcus lactis* (*Ll*Chi18A; 57% sequence identity),[21] but lacks the two C-terminal chitin-binding modules (CBM5/12 and FnIII-like) contained by these two chitinases. Like ChiC and *Ll*Chi18A, *Ef*Chi18A lacks the so-called α+β domain that is inserted in the catalytic domain of the more processive and exo-acting members of family 18 chitinases and that conveys these latter enzymes with deep substrate-binding grooves.[22] The activity of *Ef*Chi18A on chitooligosaccharides increased considerably with DP (the initial rate of hydrolysis for $GlcNAc_6$ is ~130 fold higher than for $GlcNAc_3$), which is a common feature amongst non-processive endochitinases.[21, 23, 24] The mapping of preferential productive binding modes depicted in Fig. 2 suggests the presence of perhaps up to 8 subsites with detectable sugar

affinities (-3 to +5). *Ef*Chi18A is special in that it depicts productive binding modes where the only occupied non-reducing end subsite is subsite -1. This is not commonly observed for family 18.[25, 26] It is also unexpected, because it is generally assumed that strong binding to the -2 subsite is necessary to facilitate the energetically unfavorable distortion that accompanies sugar binding to the -1 subsite.[27]

The most important feature of the chitinolytic system of *E. faecalis* V583 is the presence of a catalytically active member of the CBM33 family for which a high resolution crystal structure was determined. The importance of family 33 CBMs for the efficiency of chitinolytic systems has been demonstrated in the past,[11, 21, 28] but only very recently it has become clear that at least some CBM33s are enzymes that cleave insoluble polysaccharides such as chitin[7] and cellulose.[13] The functional studies of *Ef*CBM33A described above show that its enzymatic activity is similar to that of CBP21, the only chitin-active CBM33 that has been characterized in detail to date.[7] The enzyme cleaves chitin in an oxidative reaction that depends on divalent metal ions and the presence of an electron donor, and it acts synergistically with the chitinase, *Ef*Chi18A.

To gain further insight into this potentially highly important class of enzymes, the X-ray crystallographic structure of *Ef*CBM33A was solved at atomic resolution (0.95 Å; Fig. 3). Overall, the structures of *Ef*CBM33A and CBP21 are very similar (Fig. 3A). One difference concerns the absence of disulfide bridges in *Ef*CBM33A, which are conserved in many CBM33s, including CBP21, where they seem to contribute to stability (Fig. 4). Most remarkably, while the purified enzyme was active without addition of metal ions (showing their presence) and while its activity was abolished by removal of metal ions (showing their necessity), the X-ray data for *Ef*CBM33A showed no evident residual electron density near the histidine motif that could be interpreted as a metal ion. Structure determinations of CBP21 WT (three protein molecules in the asymmetric unit) and CBP21 Y54A (two protein molecules in the asymmetric unit) showed no trace of metals bound either, except in one protein molecule in the CBP21 WT structure (chain C, PDB-ID 2BEM; Fig. 3D&E; see Ref. 12). It should be noted that all three proteins were crystallized by different conditions (*Ef*CBM33A: 1.0 M K/Na tartrate, 0.1 M imidazole pH 8.0 and 0.2 M NaCl, CBP21 WT: 1.26 M $(NH_4)_2SO_4$, 0.1 M HEPES pH 7.5 and 20% ethylene glycol and CBP21 Y54A: 35% (w/v) PEG 3000 and 0.5 M CHES buffer, pH 9.5) and that the protein batches used for crystallization all were active in chitin oxidation reactions (verified by MALDI-TOF-MS, results not shown). For comparison, the two published structures of the structurally similar

GH61 proteins[14, 29] were determined using high concentrations of divalent cations in the crystallization conditions and both have metal ions bound to the histidine motif.

The various available structures of apo-CBM33s show considerable plasticity of the histidines, as illustrated for *Ef*CBM33A vs. CBP21 in Fig. 3D&E. Fig. 9 shows a surface representation of CBP21 and *Ef*CBM33A illustrating how the movement of the histidine affects the surface outline, in particular the shape of a small crevice close to the histidines. In addition to affecting the nature of the binding surface, the conformational flexibility of the histidines, which in the structure of *Ef*CBM33A also is reflected in high B-factors, may be part of the catalytic mechanism itself. In *Ef*CBM33A, His114 is hydrogen bonding to Asp180, which is another fully conserved residue in CBM33s.

The metal dependency of CBM33s is evident as enzyme activity was strongly reduced when experiments were conducted with a metal free enzyme. In our experimental set-up, we were able to reactivate the metal-free enzyme by addition specific metals at concentrations that were much lower, and thus more realistic, than the concentrations used in our previous studies on CBP21.[7] The results showed that at a concentration as low as 3 μM of several metals can reactivate *Ef*CBM33A, suggesting that the enzyme has no strong preference for a specific metal ligand. Considering the redox activity of *Ef*CBM33A, it is remarkable that a metal not commonly known for a role in redox enzymes ($Zn^{2+}$) is among the best metals for re-activation. It should be noted that it cannot be ruled out that metal contaminants from the metal solutions or present in the substrate may have contributed to enzyme reactivation. Notably, the addition of ions such as $Zn^{2+}$ may lead to displacement of bound ions from the substrate. Indeed, chitin is known to have considerable affinity for heavy metals including e.g. copper.[30] In a very recent study that became available after this paper was ready for submission, it was claimed that at least some of the GH61 are copper-dependent oxidases.[31]

The putative substrate-binding surface of *Ef*CBM33A contains essentially all conserved residues that are known to be important for substrate binding by family 33 CBMs (Figs. 3C, 9 and Refs 12 and 32). However, the residue that so far seems to have the largest impact on binding affinity[12, 32] (residues W56 and Y54 in *Ef*CBM33A and CBP21, respectively) varies between these proteins, and has been suggested to also play a role in substrate specificity. CBP21, which has a clear preference for binding β-chitin,[12, 33] has a Tyr in this position (Tyr54), whereas CBM33s thought to specifically bind α-chitin (CHB1 and CHB2 from

*Streptomyces avermitilis* and *Streptomyces reticuli*, respectively) have a Trp in this position (W57 and W56, respectively; Refs 32 and 34). *Ef*CBM33A shows a new situation in that it has a Trp in this position (W58) and binds to both β- and α-chitin (Fig. 5). Thus, substrate binding specificity seems to be governed by additional, as yet not identified surface residues. While the highly conserved metal binding motif is an important part of the binding surface (Fig. 9), we did not see conclusive effects of EDTA on binding (whereas EDTA does abolish activity).

There are currently no known structures of CBM33s that bind to and cleave cellulose. Fig. 9D shows an alignment of CBP21 and *Ef*CBM33A with the CBM33 domain of CelS2, which is active of cellulose.[13] In Fig. 9C, the differences between this domain and the two chitin-active enzymes are mapped on the *Ef*CBM33A structure. The figure shows that several residues in the binding surface and the crevice close to the histidine motif differ between the two chitin-active CBM33s and the cellulose active CBM33s.

The functionality of *Ef*CBM33A was demonstrated by showing release of oxidized products upon incubation with chitin (Fig. 6) and by showing that the presence of the enzyme boosts chitinase activity (Fig. 7). As previously discussed,[7] the soluble products seen in Fig. 6 result from *Ef*CBM33A cleaving the same polysaccharide chain twice. The fact that most oligomeric products are even numbered, combined with the structural periodicity in the substrate, where the dimer is the repeating unit, indicates that the chains are attacked from one side only, as would necessarily be the case for chains packed in a crystalline matrix with only one accessible side.[7] As no dimeric products (GlcNAcGlcNAcA) were observed, it is likely that productive binding of *Ef*CBM33A requires at least two native GlcNAc moieties on the reducing side of the cleavage point for successful catalysis to occur.

The results indicate that *Ef*CBM33A is capable of cleaving both polymorphs of chitin through the oxidative mechanism employed by these enzymes. It should be noted though, that while ~100 % conversion of β-chitin was obtained, a large series of experiments (mostly not shown) demonstrated that conversion of α-chitin never exceeded approximately 10 %. Accordingly, the growth and enzyme activity experiments presented in Fig. 1 show that α-chitin is not a good substrate for *E. faecalis*. It is well established that α-chitin is of a more recalcitrant nature than β-chitin and this in itself (i.e. lower accessibility) could explain the low conversion rate. In addition, while the ordered surface of the α-chitin substrate is compatible

with binding of *Ef*CBM33A (Fig. 2), this binding may be non-productive for the particular arrangement of polysaccharide chains in this substrate. It is quite possible that the substrate is heterogeneous, e.g. containing fractions with varying degrees of crystallinity, and that this explains why a subfraction can be converted. In fact, CBP21, which shows low binding to α-chitin, does degrade a minor fraction of this substrate too (unpublished data). All in all, the specificity issue remains largely unanswered, but the important point is that the various members of the CBM33 family do show differences in their substrate (binding) preferences.[7, 12, 13, 21, 32, 34] It should also be noted that the substrates typically used in these studies differ from natural chitinous substrates, since these laboratory substrates were prepared by subjecting natural chitinous materials to various chemical and mechanical treatments.

In conclusion, this study shows that *E. faecalis* has a fully functional chitinolytic machinery that enables the bacterium to grow on both soluble and insoluble chitinous substrates. This machinery contains an oxidative chitin-cleaving enzyme that is currently classified as a CBM33. The most obvious use for having such an enzyme system would be food scavenging for example in habitats in which *E. faecalis* is attached to a copepod or another chitin containing organisms.[1, 2] However, gene regulation studies indicate that the genes encoding *Ef*CBM33A and *Ef*Chi18A are co-regulated by several conditions that are more related to stress and host-microbe interactions in animals. Both enzymes are down-regulated in the presence of bile, whereas they are both up-regulated in the presence of urine and, in particular, serum.[35, 36] Another observation casting doubt on a simple food-scavenging role of *Ef*CBM33A is the fact that several members of the CBM33 family are present in multi-modular proteins (GlcNAc binding proteins; Gbp's) that are thought to be involved in bacterial attachment to both aquatic transfer vectors (such as chitin containing plankton) and host intestinal cells.[3, 37, 38, 39] Finally, a recent study claims that a single module CBM33 from *Lactobacillus plantarum* is able to bind both chitinous and cell surface type substrates.[40] It is tempting to speculate that some CBM33s may be bifunctional with roles in both carbohydrate metabolism and host-microbe interactions.

MATERIALS AND METHODS

**Growth of *E. faecalis* V583 in the presence of different carbon sources**

The ability of *E. faecalis* V583 to grow on different soluble (glucose, $GlcNAc_2$ and GlcNAc) and insoluble (β-chitin, α-chitin and cellulose) substrates was determined as follows. Cultures of *E. faecalis* V583 were grown overnight in LM17ent medium (2.5 g/l Maritex Fish peptone, 2.5 g/L bacto yeast, 0.5 g/l ascorbic acid, 0.25 g/L magnesium sulfate, 19 g/l disodium glycerolphosphate, 0.05 g/l manganese sulfate) containing 0.4% (w/v) glucose. The cultures were then diluted to $OD_{600}$ ~0.1 in LM17ent containing different carbon sources, 0.4% (w/v) $GlcNAc_2$, 0.4% (w/v) GlcNAc, 0.4% (w/v) glucose, 1% (w/v) α-chitin isolated from shrimp shells (Hov-Bio, Tromsø, Norway), 1% (w/v) β-chitin from squid pen (France Chitin), 1% (w/v) crystalline cellulose (Avicel; Sigma), or no additional carbon source. Growth on soluble substrates was measured by following optical density using a microtiter plate format. 200 µl of each condition was transferred to different wells in the microtiter plate, and the plate was covered by Adhesive Plate Seals (ABgene House, Surrey, UK). The microtiter plate was incubated at 37°C with slow background shaking, and the OD was measured at 595 nm each 15 min over a period of 13 hours, using a Multiscan FC version 1.00.79 (Thermo Fisher Scientific). Growth in cultures containing insoluble substrates was monitored by counting bacteria on agar plates. In short, cultures were incubated at 37°C and harvested at various time points ranging from 8 to 144 h. Appropriate dilutions were plated on Brain Heart Infusion (BHI) agar plates and colony forming units (CFU) were counted after one day of incubation at 37°C. All experiments were done in triplicates or quadruplicates.

**Determination of chitinolytic and *N*-acetylhexosaminidase activity in culture supernatants**

Chitinase and *N*-acetylhexosaminidase activities in culture supernatants of *E. faecalis* V583 were measured using the $GlcNAc_3$ analogue 4-methylumbelliferyl-β-D-*N*,*N*`-diacetylchitobioside (4-MU-$GlcNAc_2$) and the $GlcNAc_2$ analogue 4-methylumbelliferyl-β-D-*N*-acetylglucosamine (4MU-GlcNAc), respectively. After preparing 40 µl reaction mixtures containing 0.1 mg/ml BSA, 50 µM substrate and 50 mM citrate phosphate buffer pH 6.0 that had been prewarmed at 37°C, reactions were started by adding 10 µl of the supernatant. After 10 min the reactions were stopped by adding 1.95 ml 0.2 M $Na_2CO_3$ and the amount of

released 4-MU was measured using a DyNA 200 Fluorimeter (Hoefer Pharmacia Biotech, San Francisco, CA, USA). All experiments were done in triplicates.

**Cloning, expression and purification of *Ef*Chi18A**

Genomic DNA from *E. faecalis* V583 and the pBAD/HisB expression plasmid (Invitrogen, Karlsruhe, Germany) from *E. coli* were isolated, using the E.Z.N.A. Bacterial DNA Kit (Omega Bio-Tek Inc, Norcross, USA) and the E.Z.N.A. Plasmid Miniprep Kit I (Omega) respectively. The gene corresponding to the chitinase *Ef*Chi18A, *ef0361* (GenBank Accession no: **AAO80224.1**) was amplified by PCR (forward primer, 5`-AGATCTGATGCTGCGGATAC-3`; reverse primer, 5`-CGGTACCTTATTTTACAAGGTTACTATAA-3`; restriction sites underlined). Amplified fragments were digested with *Bgl*II and *Kpn*I restriction enzymes (New England Biolabs), ligated into a *Bgl*II/*Kpn*I-digested pBAD/HisB vector, using the Quick Ligation Kit (New England Biolabs), and transformed into *E. coli* TOP10 competent cells, yielding the pBAD/HisB-0361 construct. The amplified sequence was confirmed by DNA sequencing using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Perkin Elmer/Applied Biosystems, Foster City, CA, USA) and an ABI PRISM® 3100 Genetic Analyser (Perkin-Elmer/Applied Biosystems) for sequencing. For enzyme production, 10 ml over-night-culture of *E. coli* TOP10 containing pBAD/HisB-0361 was added to 500 ml BHI broth containing 0.1 mg/ml ampicillin, and the resulting culture was incubated at 37°C and 250 rpm. When the $OD_{600}$ reached 0.7, expression was induced by the addition of L-arabinose to a final concentration of 0.002% (w/v). The culture was further incubated at 30°C over-night. The cell pellet was harvested by sentrifugation (7700 x g, 10 min, 4°C), and resuspended in 20 ml Buffer A (100 mM Tris-HCl pH 8.0 and 20 mM Imidazole). The cells were then lysed by sonication, using a Vibra cell Ultrasonic Processor (Sonics, Newton, CT, USA), at 20% amplitude with 5 s pulses (with 5 s delay between pulses) in a total of 15 minutes on ice. The sonicated cells were centrifuged (17400 x g, 15 min, 4°C), and the supernatant was transferred to a new tube. This extract was applied to a Ni-NTA column equilibrated with Buffer A. The protein was eluted with Buffer B (100 mM Tris-HCl pH 8.0, 100 mM Imidazole), and the peak containing the chitinase was collected and concentrated using a Centricon Plus-20 unit (Millipore, Billerica, MA, USA). Protein purity was analyzed by SDS-PAGE, and protein concentration was determined using the Bradford micro-assay provided by Bio-Rad (Bio-Rad Laboratories Inc, Hercules, CA), according to the suppliers' procedure. Purified and concentrated chitinase was stored in 20 mM Tris-HCl pH 8 at 4°C until use.

**Cloning, expression and purification of EfCBM33A and CBP21**

Genomic DNA from *E. faecalis* V583 was isolated as described above. The cloning vector was the pRSETB-based expression construct used for production of CBP21,[12] which is referred to as pT7-CBP21. The gene encoding *Ef*CBM33A (GenBank Accession no: **AAO80225.1**) was amplified by PCR (forward primer, 5`-CATACGGAATGCTCATGGTTACGTAGCAAGTC -3`; reverse primer, 5`-GGAAGACGTCTTACTGAATGTTCACATCAATCGCTT -3`; restriction sites underlined). Amplified fragments were digested with *Bsm*I and *Aat*II (New England Biolabs) and ligated into *Bsm*I/*Aat*II- digested pT7-CBP21 using the Quick Ligation Kit (New England Biolabs). By doing so the part of the *cbp21* gene that encodes the mature protein is exchanged by the mature part of *Ef*CBM33A, while the signal peptide of CBP21 remains part of the construct and will drive translocation of *Ef*CBM33A. Ligation mixtures were transformed into *E. coli* TOP10 competent cells and transformants were verified by DNA sequencing as described above. A correct construct was then transformed into BL21 (DE3) Star cells (Invitrogen). The protein was produced by growing the cells in 150 ml Luria-Bertani broth (Oxoid) containing 50 µg/ml ampicillin at 37°C and 250 rpm over-night (i.e. a procedure without induction). The cells were lysed by cold osmotic shock[41] and the resulting periplasmic fraction was filtered through a 0.2µM filter. 25 ml periplasmic extract was adjusted to 1.0 M ammoniumsulfate and 20 mM Tris-HCl pH 8 (Buffer A) before being applied to a 1.5 x 10-cm (18 ml) chitin bead column (New England Biolabs) which had been equilibrated with Buffer A. The column was washed with 2 volumes of running buffer before the protein was eluted with Buffer B (20 mM acetic acid). The protein was concentrated using a Centricon Plus-20 unit (Millipore), and the protein concentration was determined using the Bradford micro-assay as described above. Purified and concentrated *Ef*CBM33A (~ 5 mg/ml) was stored in 20 mM Tris-HCl pH 8.0 at 4°C until use. CBP21 was produced according to the procedure previously described.[12]

**Protein crystallization**

*Ef*CBM33A at 25 mg/ml concentration was used in sitting drop vapor diffusion crystallization experiments (0.5 µl protein solution mixed with 0.5 µl crystallization buffer) using the Wizard I random sparse matrix crystallization screen (Emerald BioSystems). Crystals were obtained in several conditions after one week of incubation at room temperature. Crystals of appropriate size, shape and quality for X-ray crystallographic experiments were located in

conditions containing 20% (w/v) PEG-8000 and 0.1 M HEPES pH 7.5 or 1.0 M K/Na tartrate, 0.1 M imidazole pH 8.0 and 0.2 M NaCl. Only crystals obtained in the latter condition were used for protein structure elucidation.

**Data acquisition and Structure Refinement**

Crystals were transferred to a cryoprotectant solution containing glucose to a final concentration of 30% (v/v) in addition to the components in the reservoir solution for ~10 seconds before being captured up by a Litho loop (Molecular Dimensions) and flash-frozen in liquid $N_2$. Diffraction images were collected at the ID14-1 beamline at the European Synchrotron Radiation Facility in Grenoble, France. A total of 90 diffraction images, each with $\Delta\varphi = 1°$ oscillation and 3 s of X-ray exposure, were collected at 100 K. Images were processed with Mosflm,[42] scaled and merged using Scala/CCP4[43] to give a complete data set to 0.95 Å resolution (Table 1). Crystals belonged to the trigonal $P3_2$ space group with cell parameters a = b = 52.11 Å and c = 59.65 Å. The structure was solved with molecular replacement using MolRep/CCP4[43] with atomic coordinates from the known structure of CBP21 as a search template (PDB code: **2BEM**).[12] The structure was initially refined with isotropic *B* factors in Refmac5/CCP4[43] interspersed with model adjustments in Coot.[44] 5 % of randomly distributed reflections were flagged for cross-validation during all steps of the refinement and not included in the model refinement. Solvent water molecules were added using both the built-in 'find waters' procedure of Coot[44] as well as manual addition. The solvated model was further refined and adjusted in several rounds using Shelxl.[45] All protein atoms and solvent oxygens were refined anisotropically. Hydrogen atoms were added to all protein residues and refined as riding atoms.

**Degradation of chitooligosaccharides by *Ef*Chi18A**

The initial rates of hydrolysis of chitooligosaccharides of varying DP were determined by quantifying the product release over time. Reactions were set up with a total volume of 200 µl and contained 0.1 mg/ml BSA, 200 µM $GlcNAc_3$, $GlcNAc_4$, or $GlcNAc_6$, 50 mM ammonium acetate buffer pH 6.0 and 15 nM, 3.0 nM or 1.5nM purified *Ef*Chi18A, respectively. The reactions were incubated at 37°C and samples were taken over a period of 16 minutes, stopping the reaction by adding 1 volume of 20 mM $H_2SO_4$. Reaction products were analyzed using a Dionex Ultimate 3000 HPLC system set up with a Rezex RFQ-Fast Fruit H+ 7.8 x 100 mm column (Phenomenex) with a Carbo-H, 4 x 3.0 mm guard column fitted in front. The mobile phase consisted of 5 mM $H_2SO_4$. The flow rate was 1 ml/min and eluted

oligosaccharides were monitored by recording absorption at 195 nm. The increase of product $GlcNAc_2$ was quantified at each time point for determination of initial rates (see below for further discussion) using Chromeleon chromatography software (Dionex). All experiments were performed in triplicate and rates were determined by linear regression.

Some experiments were set up for determining the anomeric configuration of the resulting products. In these cases the reactions contained the same reactants as above, but the enzyme concentration was always 15 nM. The reactions were incubated for 9, 6, 4 or 2 minutes at room temperature for the substrates ranging from DP3 to DP6, respectively, before immediate injection into an Agilent 1290 Infinity UHPLC instrument, equipped with a Waters Acquity UPLC BEH amide column (2.1x50 mm) with a BEH Amide VanGuard guard column (2.1x5mm) fitted in front. The mobile phase consisted of 24% $dH_2O$ and 76% acetonitrile, the flow rate was 0.3 ml/min and eluted oligosaccharides were monitored by recording absorption at 195 nm. Chromatograms were analyzed using the ChemStation rev.B.04.02 chromatography software (Agilent Technologies).

**Degradation of α- and β-chitin**

Chitin degradation was analyzed by setting up 1.0 ml reactions containing 0.5 mg/ml chitin, 1 µM *Ef*Chi18A, 50 mM ammonium acetate buffer pH 6.0 and 1.0 mM reduced glutathione and 0.3 µM *Ef*CBM33A as indicated. Reactions were incubated at 37°C with vigorous shaking (800 rpm). For reaction monitoring, 50 µl samples were taken out regularly over a period of 24 hours (thereafter twice a week for three weeks in total) and mixed with one volume 20 mM $H_2SO_4$. All reactions were run in triplicate, and $GlcNAc_2$ concentrations were analyzed by HPLC as described above.

**Substrate-binding assays**

The substrates, α-chitin, β-chitin or cellulose were suspended in $dH_2O$ yielding a concentration of 10 mg/ml. 250 µl portions of this suspension were centrifuged and the supernatants were removed followed by addition of 250 µl of a solution containing 0.2 mg/ml *Ef*CBM33A in 50 mM ammonium acetate buffer pH 6.0 in the presence or absence of 5 mM EDTA. The binding reactions were incubated at room temperature under slow rotation for three hours. Subsequently, the substrates were pelleted by sentrifugation for 5 min, 16000 x g. 10 µl (i.e. 1 %) of the supernatants were analyzed by SDS-PAGE. The pellets were washed with 1 ml ammonium acetate pH 6.0 and resuspended in 50 µl SDS-PAGE loading buffer. After boiling for 10 min, 20 µl was analyzed by SDS-PAGE. Gels were run for 45 min at

200V, stained with Coomassie brilliant blue, and destained in a solution containing 10% (v/v) methanol, and 10% (v/v) acetic acid.

**Product analysis by mass spectrometry (MS) and UHPLC**

1 μM *Ef*CBM33A was incubated with 2 mg/ml α-chitin or β-chitin in the presence or absence of 1 mM ascorbic acid in 20 mM Tris-HCl pH 8.0 or 50 mM ammonium acetate pH 6.0 at 37°C over-night, with stirring (800 rpm). For MS analysis, 2 μl droplets of a 9 mg/ml solution of 2.5-dihydroxybenzoic acid (DHB) in 30% acetonitrile placed on a MTP 384 target plate ground steel TF (Bruker Daltonics), mixed with 1 μl of the samples and dried by hot air. The samples were analysed with an Ultraflex MALDI-TOF/TOF instrument (Bruker Daltonics GmbH, Bremen, Germany) using the same procedure and parameter settings as previously described.[7] Analysis of soluble products by UHPLC was done using a Dionex Ultimate RC-LC set up with a Waters Acquity UPLC BEH amide column (2.1 x 50 mm) with a BEH Amide VanGuard guard column (2.1 x 5mm) fitted in front, using the same procedure as described in Ref. 7.

**Production, analysis and re-activation of metal-free recombinant proteins**

In order to remove divalent cations, 1 mg/ml *Ef*CBM33A was incubated with 1 mM EDTA for four hours at 4°C in a total volume of 1 ml. The buffer containing EDTA was changed to a Chelex-100 treated 50 mM ammonium acetate-buffer pH 6.0 using a centricon Plus-20 unit. The protein concentration was measured, and assays were set up containing 0.5 mg/ml β-chitin, 1 mM reduced glutathione, 1 μM metal-free *Ef*CBM33A and 50 mM ammonium acetate buffer pH 6.0, in the presence of 0.75 μM *Ef*Chi18A and in the presence or absence of 3 μM of different divalent metals (i.e. the *Ef*CBM33A$_{apo}$:metal ratio was 1:3). Reactions were run for six hours and stopped by adding 1 volume of 20 mM $H_2SO_4$, after which product formation ($GlcNAc_2$, the by far dominant product) was analyzed by HPLC as described above. All experiments were done in triplicates.

**Use of intrinsic fluorescence to assay thermo-stability**

2 μM protein in 20 mM Tris-HCl pH 8.0 was heated in 10 mm path length quartz cells from 20 to 90 ºC at 1ºC/ min in a Cary Eclipse Fluorescence Spectrophotometer (Agilent, Santa Clara CA, US) equipped with a peltier thermostatted multicell cuvette holder. Intrinsic fluorescence was measured at 295 nm excitation and 345 nm emission wavelengths. The

apparent fraction of unfolded protein ($F_d$ ) was calculated as described by Pace (1990)[46] using the equation $F_d=(y_n - y_{obs})/(y_n - y_d)$ where $y_{obs}$ is the fluorescence intensity and $y_n$ and $y_d$ are the values of y for the native and denatured state of the protein, respectively. After each run the cuvettes were carefully cleaned by rinsing in warm tap water, washing with a cotton swab, soaking overnight in 3.0 M acetic acid and rinsing with 70% ethanol and $dH_2O$.

ACCESSION NUMBERS: Coordinates and structure factors have been deposited in the Protein Data Bank with accession number **4A02**.

TABLE LEGENDS

Table 1. Binding preferences for hydrolysis of chitooligosaccharides by *Ef*Chi18A. Four different substrates were analyzed yielding the initial products listed in the "product" column. For each substrate, one column shows the ratio of the α- and β-anomers (%) for the various products. The "standard" column shows the anomeric ratios at equilibrium for the chito-oligosaccharides indicated in the "product" column. N.d.: not determined (the analytical HPLC method did not allow reliable detection of GlcNAc). Note that differences between the anomeric ratios found for the standard and for the non-cleaved substrates (e.g for the trimer the standard ratio is 57:43, whereas the uncleaved trimeric substrate shows 58:42) are generally small. Such differences reflect inaccuracy of the method and may also reflect a preference that the enzyme may have for one of the anomeric forms of the substrate.

Table 2. Crystal data and refinement statistics for the X-ray analysis of *Ef*CBM33A.

FIGURE LEGENDS

Fig. 1. The ability of *E. faecalis* V583 to degrade and grow on chitin. (A) Cultivation of *E. faecalis* V583 on BHI or minimal medium (labelled M) supplemented with 0.4% Glc (labelled Glc), 0.4 % GlcNAc (labelled A) or 0.4 % $GlcNAc_2$ (labelled A2). The $OD_{595}$ readings for uninoculated media are labelled "M cnt." or "BHI cnt.". All data points are mean values of four parallels. For the sake of clarity, standard deviations, all below 12% of the value, are not shown. Note that the maximum densities reached on glucose were considerably lower than those reached on GlcNAc or $GlcNAc_2$. Addition of four times as much glucose (1.6 %) yielded a growth curve very similar to that obtained for the culture with 0.4 % GlcNAc (not shown). (B) *N*-acetylhexosaminidase activity in the culture supernatants harvested from the cultures shown in panel (A; m – uninoculated minimal medium, cnt – inoculated minimal medium and Glc, A2 and A – glucose, GlcNAc or $GlcNAc_2$ supplemented minimal medium, respectively). Enzyme activity was measured at the end of the cultivation experiments (13 hours) by determining the ability to hydrolyse the fluorogenic substrate 4MU-GlcNAc and activity is expressed as relative fluorescence per volume of supernatant. (C) Chitinolytic activity in culture supernatants from *E. faecalis* V583 cultures grown for 72 hours in minimal media containing various insoluble carbon-sources (labels: cnt, Glc, α, β and cel: inoculated minimal medium supplemented with water, glucose, α-chitin, β-chitin or cellulose, respectively. Activity was recorded by measuring hydrolysis of the fluorogenic substrate 4MU-$(GlcNAc)_2$ and is expressed as relative fluorescence per volume of supernatant. Figures above bars indicate the CFU/ml count for each culture.

Fig. 2. Preferred chitooligosaccharide binding to *Ef*Chi18A as judged from observed productive binding modes for oligomeric substrates. The reducing ends of the substrates are indicated by α/β. Shading indicates binding preferences: black, strong; gray, moderate; light grey, weak. Binding preferences were derived from product distributions as described in the text and Table 1. The observed frequencies of the various preferred binding modes are expressed in percent and indicated to the left.

Fig. 3. The X-ray crystallographic structure of *Ef*CBM33A. (A) Superposition *Ef*CBM33A (gray color; PDB ID: 4A02) and CBP21 (cyan color; PDB ID: 2BEM[12]), shown in cartoon representation. (B) Cluster of highly conserved tryptophanes in the core of *Ef*CBM33A surrounded by $2F_oF_c$ electron density contoured at 2.0 σ. This tryptophane cluster is characteristic for CBM33 proteins. (C) Cartoon representation of *Ef*CBM33A where

conserved residues on the substrate binding surface are shown in stick representation and are labeled. The view on the right represents a 90° rotation around the x-axis compared to the view shown on the left and illustrates the "flat" binding substrate binding surface. (D and E) Detailed view of the highly conserved histidine metal binding motif. The *Ef*CBM33A structure (D) contains no bound metal and the histidines show a different positioning compared to the CBP21 structure (E) where the histidines bind an unknown metal. Note the involvement of the N-terminal amino group in metal binding (His28 in panel E corresponds to residue 1 in the mature secreted protein). All figures were made using PyMOL.[47]

Fig 4. Thermal unfolding of CBP21 and *Ef*CBM33. The pictures show thermal unfolding of 2 μM CBP21 (diamonds) or 2 μM *Ef*CBM33A (triangles) in 20 mM Tris-HCl pH 8.0, and in the same conditions including 2 mM DTT (CBP21: circles, EfCBM33: not shown for clarity reasons as it essentially overlaps that of the non-reducing conditions of CBP21). The heating rate used was 1 °C/min.

Fig. 5. Substrate binding preferences for *Ef*CBM33A. *Ef*CBM33A was incubated with the indicated substrates (2.0 mg/ml) in the presence or absence of EDTA (5 mM) at pH 6.0, as described in Materials and Methods. The presence of *Ef*CBM33A in the supernatant (A) or bound to the substrate (B) after 3 hours of incubation was verified by SDS-PAGE. Sample sizes were such that the substrate samples of panel B represent a 5 times as large fraction of the reaction mixture than the supernatant samples of panel A. Abbreviations: cnt., control (no substrate); α, α-chitin; β, β-chitin; cel, cellulose (Avicel).

Figure 6. Soluble products generated by *Ef*CBM33A upon incubation with chitin. *Ef*CBM33 was incubated with α- or β-chitin in the presence of 1.0 mM ascorbic acid in 25 mM Tris-HCl pH 8.0. (A) MALDI-TOF MS analysis. Peaks in the MALDI-TOF MS spectra are labeled by the degree of polymerization (DP), an "ox" subscript in case the compound is oxidized, and the observed mass in Da. The insert shows the ion cluster for the hexameric aldonic acid ($DP6_{ox}$ = $GlcNAc_5GlcNAcA$) and contains the following adducts: Na, sodium adduct (m/z 1275.31); K, potassium adduct (m/z 1291.28); Na+Na, sodium adduct of the sodium salt (m/z 1297.29); Na+K, sodium adduct of the potassium salt (m/z 1313.26). Spectra obtained for both substrates were essentially identical and only one spectrum (for β-chitin degradation) is shown. Essentially identical product spectra were obtained at pH 6.0; not shown. (B) Production of oxidized oligosaccharides from β-chitin over time, as quantified by UHPLC analysis (complete line: $GlcNAc_5GlcNAcA$, widely spaced dashed line: $GlcNAc_7GlcNAcA$,

closely spaced dashed line: $GlcNAc_3GlcNAcA$ and dotted line and combined dotted and dashed lines [overlapping], $GlcNAc_4GlcNAcA$ and $GlcNAc_6GlcNAcA$, respectively).

Figure 7. Degradation of β-chitin (A) and α-chitin (B) by *Ef*Chi18A in the presence (circles on complete line) or absence (triangles on dashed line) of *Ef*CBM33 and reduced glutathione. Reactions containing *Ef*Chi18A only is shown by squares on dotted line. In the presence of reduced glutathione *Ef*CBM33 increased chitinase (*Ef*Chi18A) activity considerably in both reactions (conditions are provided in the Materials and methods section). The $(GlcNAc)_2$ concentration obtained after complete conversion of the substrate is in the order of 1.2 mM (note that ~10% of the total dissolved products from β-chitin degradation is not observable with the analytical method used due to deacetylation or oxidation of the sugars). The presence of reductant did not affect the outcome of reactions without *Ef*CBM33.

Figure 8. Effect of metals on the function of *Ef*CBM33A. The first bar ("cnt") shows the relative degree of product release ($GlcNAc_2$) after incubating 2.0 mg/ml β-chitin with 0.75 μM *Ef*Chi18A and 1.0 mM reduced glutathione in 50 mM ammonium acetate pH 6.0 for 6 hours. The remaining bars show the relative degree of product release for similar reactions that also contained 1.0 μM purified *Ef*CBM33A (Ef), 1.0 μM metal free purified *Ef*CBM33A (Ef-a) or 1.0 μM *Ef*CBM33apo and 3.0 μM of the indicated metal ions (see subtitle on the x-axis). Note that the presence of reductant will lead to some degree of reduction of the added $Cu^{2+}$ and $Fe^{3+}$.

Fig. 9. CBM33 substrate binding- and active site. The structures of CBP21 (A) and *Ef*CBM33A (B&C) are visualized using a semi-transparent surface representation with amino acid side chains shown as sticks. The degree of conservation is indicated by residue coloring using a color scale ranging from blue (not conserved), via white, to magenta (highly conserved). Levels of residue conservation were determined using the "Consurf" server[48] with default parameters, using the respective 3D structures as template (CBP21; PDB ID: 2BEM, *Ef*CBM33A; PDB ID: 4A02). Conspicuous residues in the two chitin active CBM33s that are not conserved in the cellulose active CBM33 CelS2[13] are colored orange and labeled in the *Ef*CBM33A structure shown in panel C. The arrow pointing towards I178 indicates a conserved Ile that lines the wall of the small crevice next to the metal binding site and that is substituted with an Arg in CelS2. Note that Glu59 occurs in two conformations in *Ef*CBM33A (panels B & C). Panel (D) shows an alignment of all three sequences highlighting the important substitutions in CelS2 (orange) as well as surface residues that are conserved in

all three sequences (magenta). The multiple alignment was made by the stand-alone ClustalX software[49] using default parameters.

ACKNOWLEDGEMENTS

This work was funded by grants 178428, 183627, 183637 and 196885 from the Norwegian research Council. M.B. and B.D. were supported by the South-Eastern Norway Regional Health Authority (Grant no. 2009100 and 2011040 for a Technology Core Facility for Structural Biology and Bioinformatics). The X-ray diffraction experiment was performed on the ID14-1beamline at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to the beamline staff at ESRF for providing assistance in using beamline ID14-1 and greatly appreciate Silje Bergum for assisting with melting point assays, Britt Dahl in cloning *Ef*Chi18A and Bjørge Westereng for support on chromatographic methods and data analysis.

REFERENCES


1. Signoretto, C., Burlacchini, G., Lleo, M. M., Pruzzo, C., Zampini, M., Pane, L., Franzini, G. & Canepari, P. (2004). Adhesion of Enterococcus faecalis in the nonculturable state to plankton is the main mechanism responsible for persistence of this bacterium in both lake and seawater. *Appl. Environ. Microbiol.* **70**, 6892-6896.
2. Signoretto, C., Burlacchini, G., Pruzzo, C. & Canepari, P. (2005). Persistence of Enterococcus faecalis in aquatic environments via surface interactions with copepods. *Appl. Environ. Microbiol.* **71**, 2756-2761.
3. Zampini, M., Pruzzo, C., Bondre, V. P., Tarsi, R., Cosmo, M., Bacciaglia, A., Chhabra, A., Srivastava, R. & Srivastava, B. S. (2005). Vibrio cholerae persistence in aquatic environments and colonization of intestinal cells: involvement of a common adhesion mechanism. *FEMS Microbiol. Lett.* **244**, 267-273.
4. Paulsen, I. T., Banerjei, L., Myers, G. S., 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. & Fraser, C. M. (2003). Role of mobile DNA in the evolution of vancomycin-resistant Enterococcus faecalis. *Science* **299**, 2071-2074.
5. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **37**, 233-238.
6. Leisner, J. J., Larsen, M. H., Ingmer, H., Petersen, B. O., Duus, J. O. & Palcic, M. M. (2009). Cloning and comparison of phylogenetically related chitinases from Listeria monocytogenes EGD and Enterococcus faecalis V583. *J. App. Microbiol.* **107**, 2080-2087.
7. Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z. L., Zhai, H., Sorlie, M. & Eijsink, V. G. H. (2010). An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. *Science* **330**, 219-222.
8. Chater, K. F., Biro, S., Lee, K. J., Palmer, T. & Schrempf, H. (2010). The complex extracellular biology of Streptomyces. *FEMS Microbiol. Rev.* **34**, 171-98.
9. Schrempf, H. (2001). Recognition and degradation of chitin by streptomycetes. *A. van. Leeuw.* **79**, 285-289.

10. Siemieniewicz, K. W. & Schrempf, H. (2007). Concerted responses between the chitin-binding protein secreting Streptomyces olivaceoviridis and Aspergillus proliferans. *Microbiology* **153**, 593-600.
11. Vaaje-Kolstad, G., Horn, S. J., van Aalten, D. M.F., Synstad, B. & Eijsink, V. G. H. (2005). The non-catalytic chitin-binding protein CBP21 from Serratia marcescens is essential for chitin degradation. *J. Biol. Chem.* **280**, 28492-28497.
12. Vaaje-Kolstad, G., Houston, D. R., Riemen, A. H. K., Eijsink, V. G. H. & van Aalten, D. M. F. (2005). Crystal structure and binding properties of the Serratia marcescens chitin-binding protein CBP21. *J. Biol. Chem.* **280**, 11313-11319.
13. Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunaes, A. C., Stenstrom, Y., Mackenzie, A., Sorlie, M., Horn, S. J. & Eijsink, V. G. H. (2011). Cleavage of cellulose by a CBM33 protein. *Protein Sci.* **20**, 1479-83.
14. Harris, P. V., Welner, D., McFarland, K. C., Re, E., Navarro Poulsen, J. C., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., Xu, F., Cherry, J., Larsen, S. & Lo Leggio, L. (2010). Stimulation of Lignocellulosic Biomass Hydrolysis by Proteins of Glycoside Hydrolase Family 61: Structure and Function of a Large, Enigmatic Family. *Biochemistry* **49**, 3305-3316.
15. Langston, J. A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E. & Sweeney, M. D. (2011). Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl. Environ. Microbiol.* DOI: 10.1128/AEM.05815-11.
16. Holm, L. & Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, 545-549.
17. Collin, M. & Fischetti, V. A. (2004). A novel secreted endoglycosidase from Enterococcus faecalis with activity on human immunoglobulin G and ribonuclease B. *J. Biol. Chem.* **279**, 22558-22570.
18. Mayer, C., Vocadlo, D. J., Mah, M., Rupitz, K., Stoll, D., Warren, R. A. & Withers, S. G. (2006). Characterization of a beta-N-acetylhexosaminidase and a beta-N-acetylglucosaminidase/beta-glucosidase from Cellulomonas fimi. *FEBS J.* **273**, 2929-2941.
19. Li, H., Morimoto, K., Katagiri, N., Kimura, T., Sakka, K., Lun, S. & Ohmiya, K. (2002). A novel beta-N-acetylglucosaminidase of Clostridium paraputrificum M-21 with high activity on chitobiose. *Appl. Microbiol. Biotechnol.* **60**, 420-427.

20. Synstad, B., Vaaje-Kolstad, G., Cederkvist, F. H., Saua, S. F., Horn, S. J., Eijsink, V. G. H. & Sorlie, M. (2008). Expression and characterization of endochitinase C from Serratia marcescens BJL200 and its purification by a one-step general chitinase purification method. *Biosci. Biotechnol. Biochem.* **72**, 715-723.

21. Vaaje-Kolstad, G., Bunaes, A. C., Mathiesen, G. & Eijsink, V. G. H. (2009). The chitinolytic system of Lactococcus lactis ssp. lactis comprises a nonprocessive chitinase and a chitin-binding protein that promotes the degradation of alpha- and beta-chitin. *FEBS J.* **276**, 2402-2415.

22. Hoell, I. A., Vaaje-Kolstad, G. & Eijsink, V. G. H. (2010). Structure and function of enzymes acting on chitin and chitosan. *Biotechnol. Genet. Eng.* **27**, 331-366.

23. Hoell, I. A., Klemsdal, S. S., Vaaje-Kolstad, G., Horn, S. J. & Eijsink, V. G. H. (2005). Overexpression and characterization of a novel chitinase from Trichoderma atroviride strain P1. *Biochim. Biophys. Acta* **1748**, 180-190.

24. Terwisscha van Scheltinga, A. C., Hennig, M. & Dijkstra, B. W. (1996). The 1.8 A resolution structure of hevamine, a plant chitinase/lysozyme, and analysis of the conserved sequence and structure motifs of glycosyl hydrolase family 18. *J. Mol. Biol.* **262**, 243-257.

25. Aronson, N. N., Jr., Halloran, B. A., Alexyev, M. F., Amable, L., Madura, J. D., Pasupulati, L., Worth, C. & Van Roey, P. (2003). Family 18 chitinase-oligosaccharide substrate interaction: subsite preference and anomer selectivity of Serratia marcescens chitinase A. *Biochem. J.* **376**, 87-95.

26. Eijsink, V. G. H., Horn, S. J., Sorlie, M., Vaaje-Kolstad, G., Norberg, A. L., Synstad, B. & Varum, K. M. (2006). Comparative studies of chitinases A, B and C from Serratia marcescens. *Biocatal.Biotransfor.* **24**, 39-53.

27. Tews, I., Terwisscha van Scheltinga, A. C., Perrakis, A., Wilson, K. S. & Dijkstra, B. W. (1997). Substrate-assisted catalysis unifies two families of chitinolytic enzymes. *J. Am. Chem. Soc.* **119**, 7954-7959.

28. Moser, F., Irwin, D., Chen, S. L. & Wilson, D. B. (2008). Regulation and characterization of Thermobifida fusca carbohydrate-binding module proteins E7 and E8. *Biotechnol. Bioeng.* **100**, 1066-1077.

29. Karkehabadi, S., Hansson, H., Kim, S., Piens, K., Mitchinson, C. & Sandgren, M. (2008). The First Structure of a Glycoside Hydrolase Family 61 Member, Cel61B from Hypocrea jecorina, at 1.6 angstrom Resolution. *J. Mol.Biol.* **383**, 144-154.

30. Galli, U., Schuepp, H. & Brunold, C. (1994). Heavy-Metal Binding by Mycorrhizal Fungi. *Physiol. Plantarum* **92**, 364-368.
31. Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J. C., Johansen, K. S., Krogh, K. B., Jorgensen, C. I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C. P., Dupree, P., Xu, F., Davies, G. J. & Walton, P. H. (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc. Natl. Acad. Sci. USA*. DOI: 10.1073/pnas.1105776108.
32. Zeltins, A. & Schrempf, H. (1997). Specific interaction of the Streptomyces chitin-binding protein CHB1 with alpha-chitin-the role of individual tryptophan residues. *Eur. J. Biochem.* **246**, 557-564.
33. Suzuki, K., Suzuki, M., Taiyoji, M., Nikaidou, N. & Watanabe, T. (1998). Chitin binding protein (CBP21) in the culture supernatant of Serratia marcescens 2170. *Biosci. Biotechnol. Biochem.* **62**, 128-135.
34. Kolbe, S., Fischer, S., Becirevic, A., Hinz, P. & Schrempf, H. (1998). The Streptomyces reticuli alpha-chitin-binding protein CHB2 and its gene. *Microbiology* **144**, 1291-1297.
35. Vebo, H. C., Snipen, L., Nes, I. F. & Brede, D. A. (2009). The transcriptome of the nosocomial pathogen Enterococcus faecalis V583 reveals adaptive responses to growth in blood. *Plos One* **4**, e7660.
36. Vebo, H. C., Solheim, M., Snipen, L., Nes, I. F. & 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.
37. Bhowmick, R., Ghosal, A., Das, B., Koley, H., Saha, D. R., Ganguly, S., Nandy, R. K., Bhadra, R. K. & Chatterjee, N. S. (2008). Intestinal Adherence of Vibrio cholerae Involves a Coordinated Interaction between Colonization Factor GbpA and Mucin. *Infect. Immun.* **76**, 4968-4977.
38. Jude, B. A., Martinez, R. M., Skorupski, K. & Taylor, R. K. (2009). Levels of the secreted Vibrio cholerae attachment factor GbpA are modulated by quorum-sensing-induced proteolysis. *J. Bacteriol.* **191**, 6911-6917.
39. Kirn, T. J., Jude, B. A. & Taylor, R. K. (2005). A colonization factor links Vibrio cholerae environmental survival and human infection. *Nature* **438**, 863-866.
40. Sanchez, B., Gonzalez-Tejedo, C., Ruas-Madiedo, P., Urdaci, M. C. & Margolles, A. (2011). Lactobacillus plantarum extracellular chitin-binding protein and its role in the

interaction between chitin, Caco-2 cells, and mucin. *Appl. Environ. Microbiol.* **77**, 1123-1126.

41. Brurberg, M. B., Nes, I. F. & Eijsink, V. G. H. (1996). Comparative studies of chitinases A and B from *Serratia marcescens*. *Microbiology* **142**, 1581-1589.
42. Leslie, A. G. (2006). The integration of macromolecular diffraction data. *Acta Crystallogr. D* **62**, 48-57.
43. Collaborative Computational Project, N. (1994). The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D* **50**, 760-763.
44. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126-2132.
45. Sheldrick, G. M. (2008). A short history of SHELX. *Acta Crystallogr. A* **64**, 112-122.
46. Pace, C. N. (1990). Measuring and increasing protein stability. *Trends Biotechnol.* **8**, 93-98.
47. DeLano, W. L. & Lam, J. W. (2005). PyMOL: A communications tool for computational models. *Abstr. Pap. Am. Chem. S.* **230**, 1371-1372.
48. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, 529-33.
49. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). Clustal W and clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.

TABLES

Table 1.

| | | Substrate | | | | |
|---|---|---|---|---|---|---|
| | | **$GlcNAc_3$** | **$GlcNAc_4$** | **$GlcNAc_5$** | **$GlcNAc_6$** | **Standard** |
| Product | **GlcNAc** | N.d. | N.d. | N.d. | N.d. | N.d. |
| | **$GlcNAc_2$** | 0:100 | 20:80 | 3:97 | 0:100 | 54:46 |
| | **$GlcNAc_3$** | 58:42 | 54:46 | 54:46 | 33:67 | 57:43 |
| | **$GlcNAc_4$** | | 64:36 | 0:100 | 46:54 | 60:40 |
| | **$GlcNAc_5$** | | | 65:35 | 59:41 | 63:37 |
| | **$GlcNAc_6$** | | | | 69:31 | 65:35 |

Table 2.

| | *Ef*CBM33A |
|---|---|
| **Data collection** | |
| Beamline / Synchrotron | ID14-1 / ESRF |
| Detector | ADSC Q210 CCD |
| Wavelength (Å) | 0.9334 |
| Temperature (K) | 100 |
| | |
| **Crystal data** | |
| Space group | $P3_2$ |
| Unit cell dimensions | |
| *a*, *b*, *c* (Å) | a = b = 52.11, c = 59.65 |
| α, β, γ (°) | 90.0, 90.0, 120.0 |
| Resolution (Å) | 23.9 – 0.95 (1.00 – 0.95) * |
| $R_{sym}$ | 0.053 (0.221) |
| $I / \sigma I$ | 12.1 (3.4) |
| Completeness (%) | 96.0 (86.6) |
| Redundancy | 2.6 (2.0) |
| | |
| **Refinement** | |
| Resolution (Å) | 23.9 – 0.95 |
| No. reflections (work) | 92025 / 106339 ** |
| No. reflections (free) | 4816 / 5601 |
| No. refined parameters | 14448 |
| $R_{work}$ / $R_{free}$ | 0.091 / 0.11 for data with $F_O > 4\ \sigma F_O$ |
| No. atoms | |
| Protein | 2595 |
| Water | 286 |
| *B*-factors | |
| Protein | 20.3 |
| Water | 29.2 |
| Ramachandran analysis | |
| Favored regions (%) | 97.5 |
| Allowed regions (%) | 2.5 |
| Outliers (%) | 0 |

*Values in parentheses are for the highest-resolution shell.

** For structure factors with $F_O > 4\ \sigma F_O$ and for all data, respectively

FIGURES

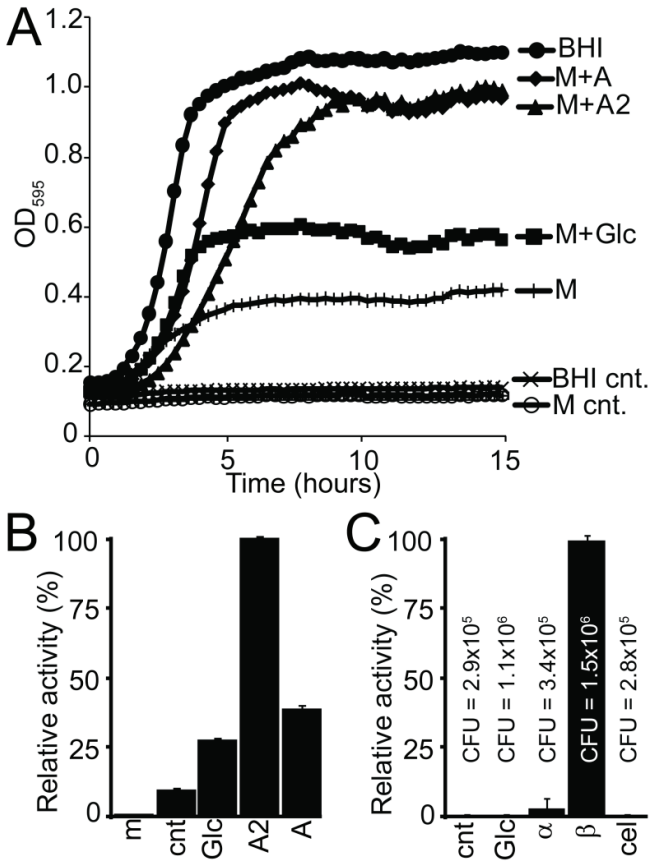



Fig. 1.

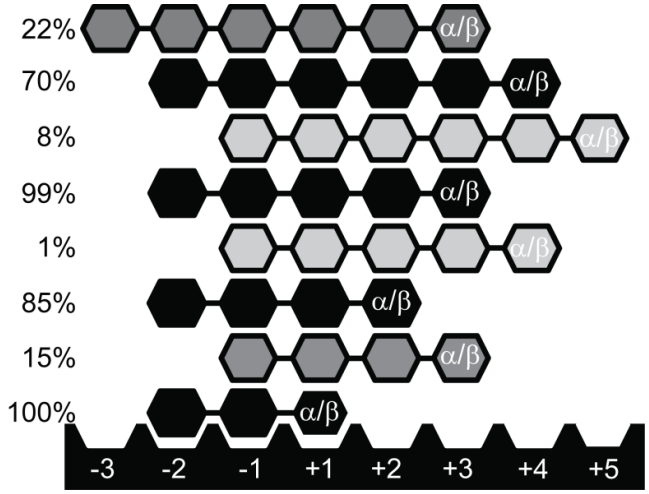



Fig. 2.

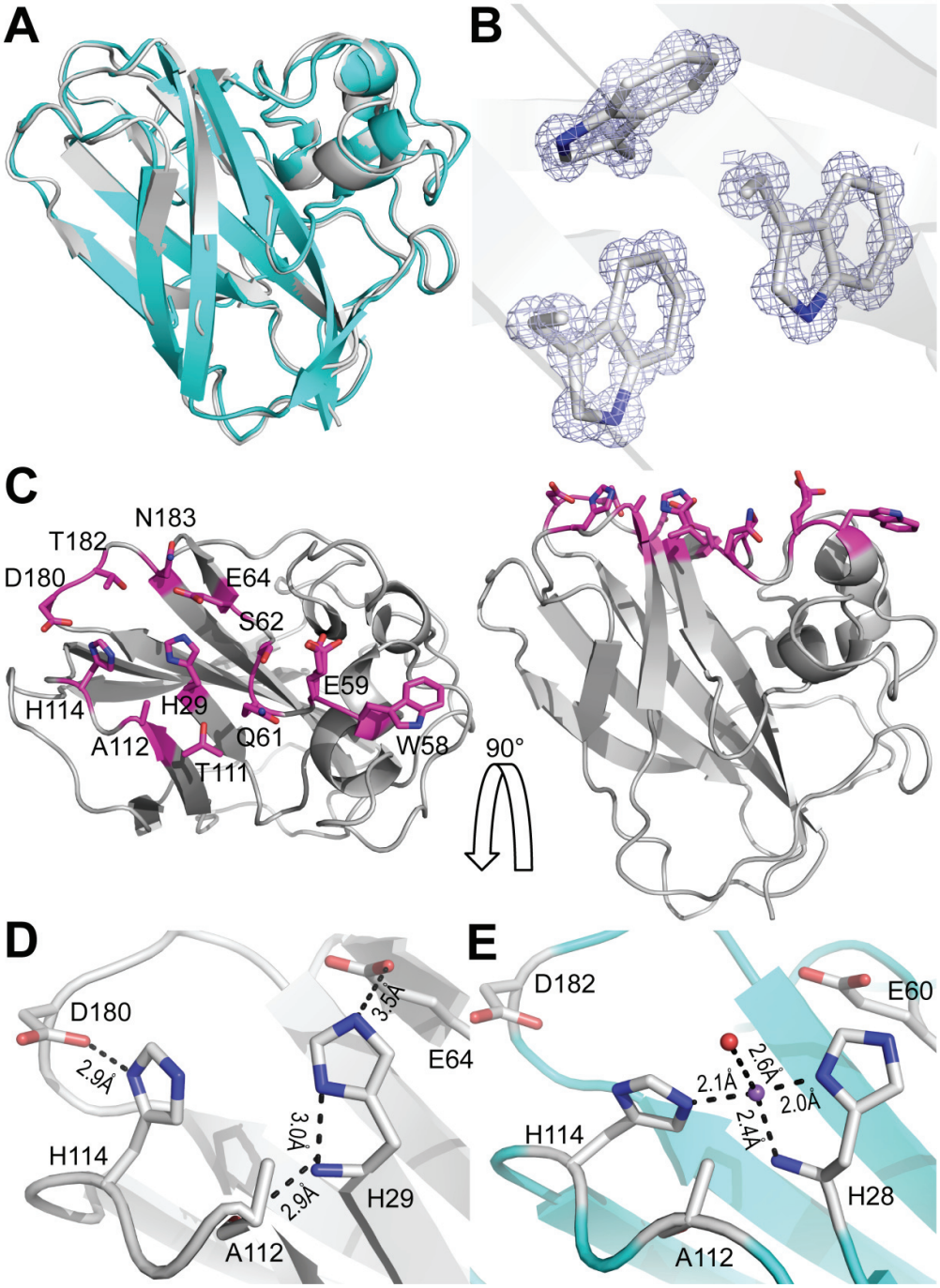

A
B
C
N183
T182
D180
E64
S62
E59
H114
H29
Q61
A112
T111
W58
90°
D
D180
2.9Å
3.5Å
E64
3.0Å
2.9Å
H114
H29
A112
E
D182
E60
2.6Å
2.1Å
2.0Å
2.4Å
H114
H28
A112

Fig. 3.

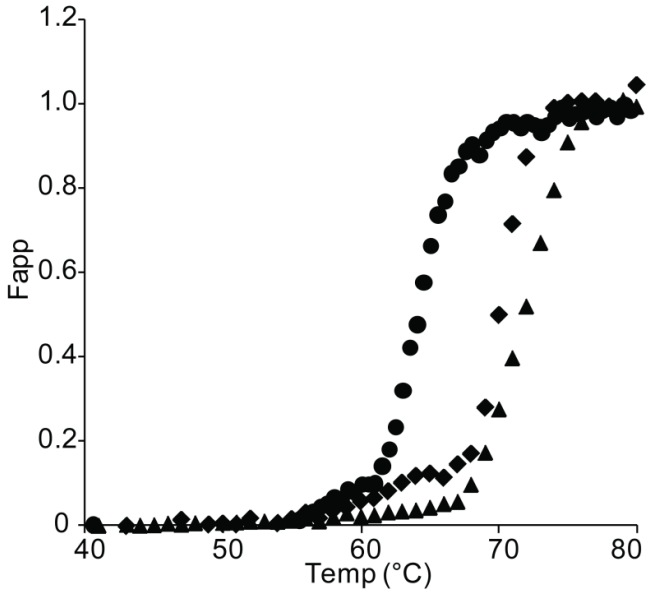



Fig. 4.

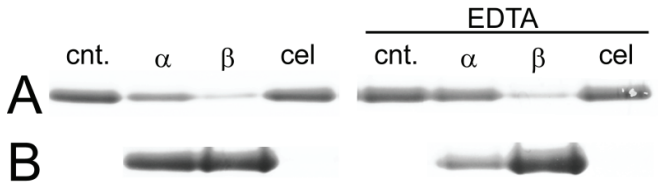



Fig. 5.

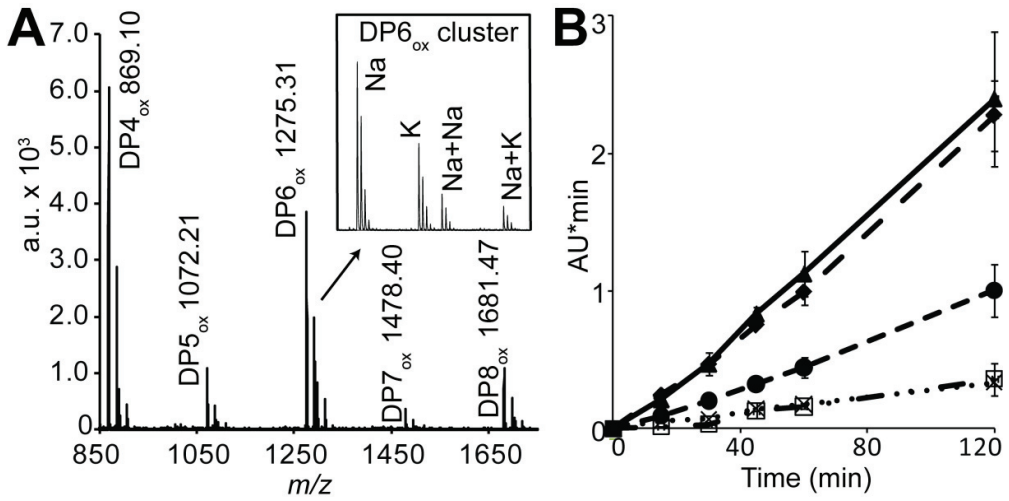

A
B
a.u. x 10³
m/z
DP4ox 869.10
DP5ox 1072.21
DP6ox 1275.31
DP7ox 1478.40
DP8ox 1681.47
DP6ox cluster
Na
K
Na+Na
Na+K
AU*min
Time (min)

Fig. 6.

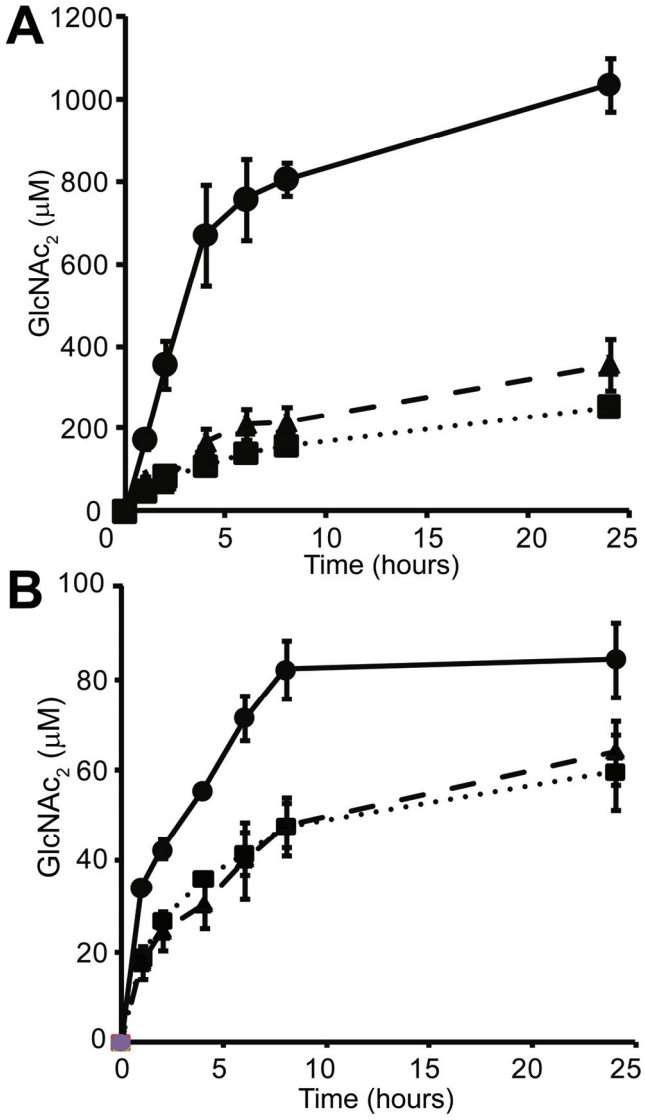

A
1200
1000
800
600
400
200
0
GlcNAc2 (μM)
0
5
10
15
20
25
Time (hours)
B
100
80
60
40
20
0
GlcNAc2 (μM)
0
5
10
15
20
25
Time (hours)

Fig. 7.

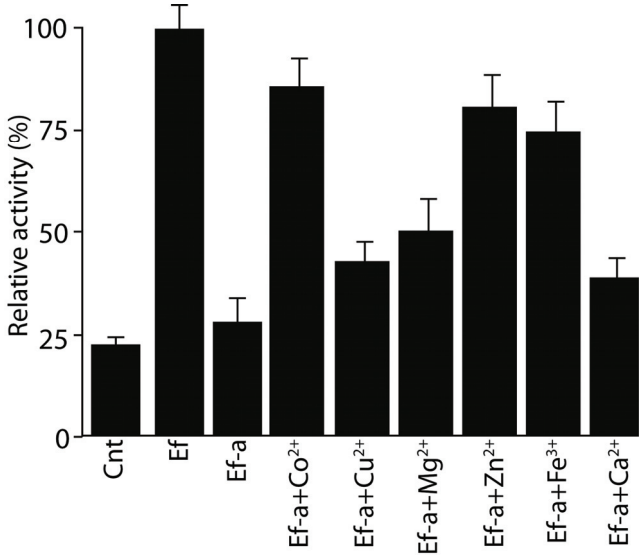

Relative activity (%)
100
75
50
25
0
Cnt
Ef
Ef-a
Ef-a+Co2+
Ef-a+Cu2+
Ef-a+Mg2+
Ef-a+Zn2+
Ef-a+Fe3+
Ef-a+Ca2+

Fig. 8.

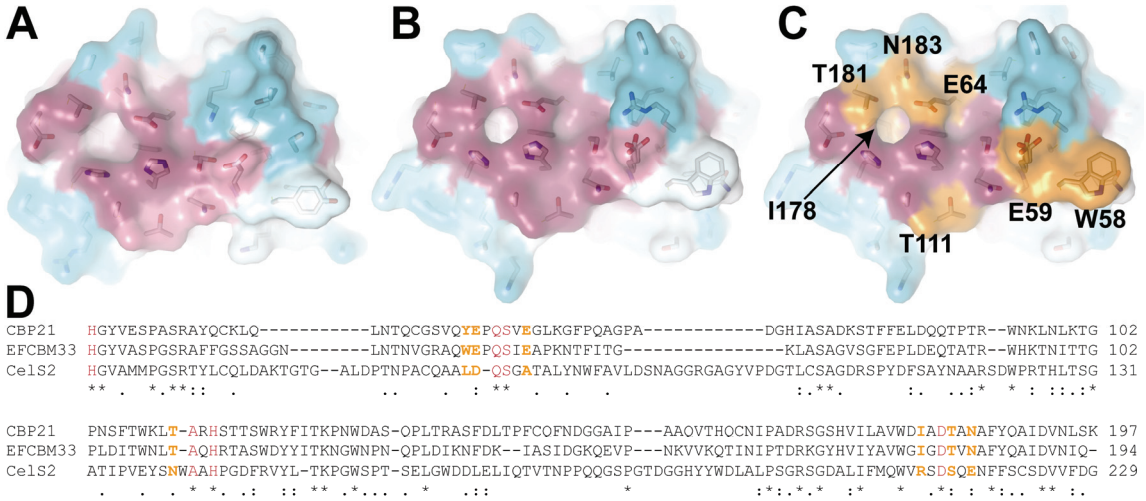

A
B
C
N183
T181
E64
I178
E59
W58
T111
D
CBP21 HGYVESPASRAYQCKLQ----------LNTQCGSVQYEPQSVEGLKGFPQAGPA----------DGHIASADKSTFFELDQQTPTR--WNKLNLKTG 102
EFCBM33 HGYVASPGSRAFFGSSAGGN--------LNTNVGRAQWEPQSIEAPKNTFITG----------KLASAGVSGFEPLDEQTATR--WHKTNITTG 102
CelS2 HGVAMMPGSRTYLCQLDAKTGTG--ALDPTNPACQAALD-QSGATALYNWFAVLDSNAGGRGAGYVPDGTLCSAGDRSPYDFSAYNAARSDWPRTHLTSG 131
CBP21 PNSFTWKLT-ARHSTTSWRYFITKPNWDAS-QPLTRASFDLTPFCQFNDGGAIP---AAQVTHQCNIPADRSGSHVILAVWDIADTANAFYQAIDVNLSK 197
EFCBM33 PLDITWNLT-AQHRTASWDYYITKNGWNPN-QPLDIKNFDK--IASIDGKQEVP---NKVVKQTINIPTDRKGYHVIYAVWGIGDTVNAFYQAIDVNIQ- 194
CelS2 ATIPVEYSNWAAHPGDFRVYL-TKPGWSPT-SELGWDDLELIQTVTNPPQQGSPGTDGGHYYWDLALPSGRSGDALIFMQWVRSDSQENFFSCSDVVFDG 229

Fig. 9.