"The extracellular biology of the lactobacilli"

Kleerebezem, Michiel ; Hols, Pascal ; Bernard, Elvis ; Rolain, Thomas ; Zhou, Miaomiao ; Siezen, Roland J. ; Bron, Peter A.

ABSTRACT

Lactobacilli belong to the lactic acid bacteria, which play a key role in industrial and artisan food raw-material fermentation, including a large variety of fermented dairy products. Next to their role in fermentation processes, specific strains of Lactobacillus are currently marketed as health-promoting cultures or probiotics. The last decade has witnessed the completion of a large number of Lactobacillus genome sequences, including the genome sequences of some of the probiotic species and strains. This development opens avenues to unravel the Lactobacillus-associated health-promoting activity at the molecular level. It is generally considered likely that an important part of the Lactobacillus effector molecules that participate in the proposed health-promoting interactions with the host (intestinal) system resides in the bacterial cell envelope. For this reason, it is important to accurately predict the Lactobacillus exoproteomes. Extensive annotation of these exoproteomes, combined with comparative analysis of species- or strain-specific exoproteomes, may identify candidate effector molecules, which may support specific effects on host physiology associated with particular Lactobacillus strains. Candidate health-promoting effector molecules of lactobacilli can then be validated via mutant approaches, which will allow for improved strain selection procedures, improved product quality control criteria and molecular science-based health claims.

CITE THIS VERSION


DIAL is an institutional repository for the deposit and dissemination of scientific documents from UCLouvain members. Usage of this document for profit or commercial purposes is strictly prohibited. User agrees to respect copyright about this document, mainly text integrity and source mention. Full content of copyright policy is available at Copyright policy

Available at: http://hdl.handle.net/2078.1/34192
The extracellular biology of the lactobacilli

Michiel Kleerebezem1,2,3,4, Pascal Hols5, Elvis Bernard5, Thomas Rolain5, Miaomiao Zhou6, Roland J. Siezen1,2,4,6 & Peter A. Bron1,2

1TI Food and Nutrition, Wageningen, The Netherlands; 2NIZO food research, Ede, The Netherlands; 3Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands; 4Kluyver Centre for Genomics of Industrial Fermentation, Delft, The Netherlands; 5Unité de Génétique, Institut des Sciences de la Vie, Université catholique de Louvain, Louvain-la-Neuve, Belgium; and 6Centre for Molecular and Biomolecular Informatics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

Abstract

Lactobacilli belong to the lactic acid bacteria, which play a key role in industrial and artisan food raw-material fermentation, including a large variety of fermented dairy products. Next to their role in fermentation processes, specific strains of Lactobacillus are currently marketed as health-promoting cultures or probiotics. The last decade has witnessed the completion of a large number of Lactobacillus genome sequences, including the genome sequences of some of the probiotic species and strains. This development opens avenues to unravel the Lactobacillus-associated health-promoting activity at the molecular level. It is generally considered likely that an important part of the Lactobacillus effector molecules that participate in the proposed health-promoting interactions with the host (intestinal) system resides in the bacterial cell envelope. For this reason, it is important to accurately predict the Lactobacillus exoproteomes. Extensive annotation of these exoproteomes, combined with comparative analysis of species- or strain-specific exoproteomes, may identify candidate effector molecules, which may support specific effects on host physiology associated with particular Lactobacillus strains. Candidate health-promoting effector molecules of lactobacilli can then be validated via mutant approaches, which will allow for improved strain selection procedures, improved product quality control criteria and molecular science-based health claims.

Introduction: the lactobacilli

Lactobacilli belong to the lactic acid bacteria (LAB), which are Gram-positive organisms with a low G+C content that belong to the phylum of the Firmicutes, and are members of the Clostridium-Bacillus subdivision of Gram-positive eu-bacteria (Pot et al., 1994). The genus Lactobacillus currently includes 148 recognized species (NCBI taxonomy database), and encompasses an unusually high phylogenetic and functional diversity. Lactobacilli encompass aero-tolerant and anaerobic species and strains and are classically regarded as strictly fermentative. They have traditionally been divided into three groups based on their fermentation characteristics: obligately homofermentative, facultatively heterofermentative and obligately heterofermentative (Pot et al., 1994; Hammers & Vogel, 1995; Claesson et al., 2008). However, the presence of heme and/or menaquinone can stimulate aerobic respiration, leading to increased biomass formation without acidification in a subset of Lactobacillus species (Brooijmans et al., 2009).

Many lactobacilli are associated with food and feed fermentation, mainly because they contribute to raw-material preservation due to acidification, but also because of their capacity to contribute to product characteristics such as flavor and texture. The natural habitat of lactobacilli ranges from dairy, meat and plant material fermentations to the oral cavity, and the genital and gastrointestinal tracts of humans and animals (Hammes & Vogel, 1995; Vaughan et al., 2002). Lactobacilli have been recognized as potential health beneficial microorganism in the human gastrointestinal tract, which is clearly reflected by the probiotic products that are currently being marketed. A broadly accepted definition of ‘probiotics,’ formulated by the World Health Organization, states that probiotics are ‘live microorganisms...
which when administered in adequate amounts confer a health benefit on the host’ (FAO/WHO, 2002). Notably, under this definition, the endogenous intestinal tract bacteria are not considered as probiotics unless they are isolated, cultured and subsequently administered to the host. Although this definition specifies neither the mode of application nor the site of action within the host body, the most common probiotic applications use oral administration (mostly as fresh fermentation products or dried bacterial supplements) and are proposed to provide their health benefits through interactions within the gastrointestinal tract. The continuing growth of markets addressing health and well-being for consumers has strongly stimulated molecular research into the metabolic behavior and potential health beneficial effects of lactobacilli, including (post)genomic research.

The extracellular characteristics of different lactobacilli are of great importance for their capacity to interact with and influence different factors encountered within the gastrointestinal tract (for reviews, see Lebeer et al., 2008; Kleerebezem & Vaughan, 2009). This review focuses on the genome-based prediction of the extracellular proteome of lactobacilli and the comparative analysis of their genes and proteins. In addition, it addresses the nonproteinaceous building blocks of the Lactobacillus cell wall because they play a key role in interactions with the host. We also discuss the current state of our knowledge of the molecular interaction of specific extracellular components of lactobacilli with the host intestinal system, combined with a short overview of postgenomic in vivo approaches to unravel host responses to lactobacilli.

Genomics of lactobacilli

Following the initial focus of bacterial genomics on pathogenic and paradigm laboratory species, the focus has shifted to encompass many industrially relevant and benign bacteria, including lactobacilli. The current public databases contain 18 complete Lactobacillus genomes, while at least 50 Lactobacillus genome sequencing projects are ongoing at present (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi). Extensive comparative analyses of the Lactobacillus (and other LAB) genomes already revealed the molecular basis for some phylogenetic, phenotypic and ecological diversities of the different species encompassed within the genus (Chanchaya et al., 2006; Makarova et al., 2006; Cleasson et al., 2007; O’Sullivan et al., 2009). In general, Lactobacillus genome annotation and metabolic reconstruction revealed a considerable degree of auxotrophy for amino acids and/or other cellular building blocks. Lactobacilli appear to compensate for these metabolic ‘gaps’ by encoding a large variety of import functions to incorporate environmental nutrients into their metabolism. Niche-specific genomic adaptations are clearly reflected within the Lactobacillus genomes. The typical milk-adapted Lactobacillus bulgaricus and Lactobacillus helveticus genomes (Makarova et al., 2006; Callanan et al., 2008) are characterized by so-called genome decay and contain many pseudogenes related to the utilization of several carbohydrates, reflecting their dedication to growth on lactose. Notably, these characteristics are shared with Streptococcus thermophilus, another LAB that is strongly adapted to the milk habitat (Botolin et al., 2004; Hols et al., 2005). In contrast, the lactobacilli associated with the intestinal niche commonly encode a large array of sugar import and utilization functions (Kleerebezem et al., 2003; Makarova et al., 2006; Ventura et al., 2009). Other functions that appear to be typically enriched in intestinal lactobacilli include the (mucus binding) cell-surface proteins and specific extracellular enzyme complexes that may be involved in complex carbohydrate degradation (Boekhorst et al., 2006a, b; Siezen et al., 2006). Analogously, the distribution of genes encoding bile salt hydrolase (BSH) among lactobacilli, as well as a recent metagenomic study (Jones et al., 2008), suggests a clear association of this function with the intestinal habitat (Lambert et al., 2008a; O’Sullivan et al., 2009). Such genes are essential for bile tolerance of Lactobacillus plantarum and Lactobacillus salivarius (Lambert et al., 2008b; Fang et al., 2009). The BSH-encoding gene has recently been proposed to be an intestinal niche-specific molecular marker for lactobacilli, as deduced from the detailed comparison of three dairy, five intestinal and three multiniche Lactobacillus genomes. Next to the bsh gene, the intestinal lactobacilli appear to exclusively encode two specific sugar transport functions, while the dairy lactobacilli exclusively contain a set of six genes encoding functions related to proteolytic capacities and restriction modification systems (O’Sullivan et al., 2009). Analogously, a recent comparative phylogenetic and single-gene marker study proposed reclassification of the Lactobacillus genus, and identified some key taxonomic lactobacilli whose genome sequencing would provide advanced molecular depth for such reclassification (Claesson et al., 2008), and some of these are targeted by ongoing whole-genome sequencing projects (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi).

The current shift of bacterial genomics from single-strain genomics to the pan-genomics of a species, including postgenomic approaches such as comparative genome hybridization using whole-genome DNA micro-arrays to assess genomic diversity in relation to phenotypic diversity, is illustrated by the lactobacilli L. plantarum (Molenaar et al., 2005; Siezen et al., 2010) and Lactobacillus sakei (McLeod et al., 2008). In addition, strain diversity can nowadays also be addressed by the determination of multiple genome sequences of individual isolates of a particular species (for a review, see Tettelin & Feldblum, 2009). For several Lactobacillus species, currently, there are multiple genome sequences available, including L. plantarum, Lactobacillus casei, Lactobacillus delbrueckii, Lactobacillus reuteri and Lactobacillus rhamnosus, while some of the ongoing
Lactobacillus genome sequencing projects target multiple strains of a particular species, including six strains of Lactobacillus crispatus and Lactobacillus jensenii (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi). This trend is bound to facilitate function assignment, including the identification of potential probiotic ‘effector molecules’ as has been illustrated for the mannose-specific adhesin function encoded by L. plantarum (Pretzer et al., 2005). Similarly, comparative genomics may directly enable the identification of strain-specific probiotic ‘effector molecules.’ In this respect, the recent completion of the genome sequence of the best-documented probiotic strain, L. rhamnosus GG (Kankainen et al., 2009), and its comparison with the closely related LC705 illustrates the potential of this approach. The two L. rhamnosus genomes (both approximately 3.0 Mbp) display high levels of similarity and synteny, but contain strain-specific genomic islands. The genomic islands specific for strain GG encode approximately 80 proteins, including those involved in sugar metabolism and transport, and exopolysaccharide biosynthesis. One of the L. rhamnosus GG-specific genome islands encodes a pilin-like surface structure that is important in adherence to intestinal mucus and is proposed to aid the persistence of L. rhamnosus GG in vivo in the intestine (Kankainen et al., 2009). Analogously, the genome of the probiotic Lactobacillus johnsonii strain NCC533 is predicted to encode fimbriae-like surface structures that may also play a role in epithelial cell adhesion (Pridmore et al., 2004).

Genome mining aiming to identify probiotic effector molecules is commonly focused on functions that are targeted toward the cell surface, because these functions are considered to be plausible candidates for probiotic interactions with the intestinal system. As an example, in silico exoproteome prediction for L. plantarum WCFS1 revealed at least 12 proteins that are putatively involved in adherence to host components such as collagen and mucin (Boekhorst et al., 2006a). Mutational analysis of predicted extracellular fibronectin and mucin-binding proteins of Lactobacillus acidophilus NCFM confirmed their role in human epithelial cell binding in vitro (Buck et al., 2005). Therefore, it is of great importance that analysis of Lactobacillus genomes includes the accurate predictions of surface-associated functions, and encompasses the prediction of subcellular location (SCL) and correlated membrane or cell-wall anchoring mechanisms.

**Lactobacillus** protein transport pathways

Seven main protein secretion mechanisms have been characterized in Gram-positive bacteria, namely the secretion (Sec), twin-arginine translocation (Tat), flagella export apparatus (FEA), fimbriin-protein exporter (FPE), holin (pore-forming), peptide-efflux ABC and the WXG100 secretion system (Wss) pathways (for reviews, see van Wely et al., 2001; Lee et al., 2006; Driessen & Nouwen, 2008; Desvaux et al., 2009). These pathways are commonly conserved in many Gram-positive bacteria, and by applying sequence homology and protein-domain searches, we have evaluated the presence of these protein secretion pathways in 13 published genomes of lactobacilli (Supporting Information, Table S1). This targeted mining of the Lactobacillus genomes revealed that these species do not encode the main factors involved in the Tat, FEA and Wss protein secretion pathways, but do contain genes encoding the Sec, FPE, peptide-efflux ABC and holin systems (Fig. 1; Table S1).

**The major secretion pathway: Sec**

The Sec translocase (Fig. 1) is the major system that mediates protein transfer across the cytoplasmic membrane in Gram-positive bacteria (for a review, see Driessen & Nouwen, 2008). The translocase consists of a membrane-embedded protein-conducting channel (SecYEG) and an ATPase motor protein (SecA). The Sec translocase is usually associated with the heterotrimeric complex SecDF-YajC, which is involved in SecA activity regulation. The SecDF-YajC may also bind to the YidC protein, which is relevant for membrane insertion of integral membrane proteins (Driessen & Nouwen, 2008). All Lactobacillus genomes encode single copies of SecA, SecE and SecY, whereas genes for double copies of YidC (Table S1), while no genes encoding SecDF proteins could be found. In addition, all Lactobacillus genomes encode single copies of the components of the signal-recognition pathway, which is involved in targeting of precursor proteins to the Sec translocase, while the alternative signal-capturing pathways depending on SecB (or its functional analogue in Bacillus subtilis CsaA) appear to be absent in all Lactobacillus genomes (Table S1).

All proteins targeted to the Sec translocase contain an N-terminal signal peptide, which typically consists of three regions: (1) the N region: a positively charged N terminus; (2) the H region: a stretch of 15–25 hydrophobic residues; and (3) the C region that may contain a signal peptidase cleavage site (Driessen & Nouwen, 2008). During or after translocation of the precursor protein across the cytoplasmic membrane, these signal peptides can be removed by signal peptidases (SPases). Type-I SPase recognizes the canonical AxAA cleavage site (van Roosmalen et al., 2004), while Type-II SPase recognizes the L-x-x-C or the so-called lipobox cleavage site (Sutcliffe & Harrington, 2002). All Lactobacillus genomes encode a single Type-II SPase, while the number of Type-I SPases ranged from one (in most species) to three (in L. plantarum) (Table S1). This variable number of Type-I SPases has also been found in other Gram-positive genera (van Roosmalen et al., 2004).
Holins

Holins (Wang et al., 2000) are small integral membrane proteins that are primarily involved in the secretion of muralytic enzymes that lack a signal peptide and play a role in autolysis (Fig. 1). Holins are frequently encoded by bacteriophage genomes, but can also be found in Lactobacillus genomes. Identification of holins is hampered by their low sequence similarity, but holins do share overall structural and functional features that are commonly conserved (Wang et al., 2000). Holins encoded within the Lactobacillus genomes were identified on the basis of the following criteria: (1) size range of 60–150 amino acids; (2) at least one, but less than four transmembrane segments; (3) a hydrophilic N terminus; (4) a polar, charge-rich C-terminal domain; (5) reside in a gene context encoding cell-lysis-associated proteins; (6) display at least 50% sequence similarity to known holin sequences; and/or (7) harbor a holin-family domain. These analyses revealed that holins are generally encoded by Lactobacillus genomes as a part of the cell lysis system, although no holin could be identified using these criteria in some Lactobacillus strains (Table S1).

FPE

The FPE pathway (Fig. 1) is part of the competence development (Com) pathway, allowing exogenous DNA uptake across the bacterial cytoplasmic membrane (Chen & Dubnau, 2004). The prepilin(-like) precursors involved in this process are proposed to be translocated via a cleavage event at the cytoplasmic side of the membrane by the prepilin-specific SPase or transmethylase ComC (Chen & Dubnau, 2004). In B. subtilis, the FPE system consists of seven comG genes (comGA-GG operon) and a genetically unlinked comC gene. These genes are involved in the assembly of the pilin-like structure involved in DNA recognition at the cell surface, including the export of the prepilins ComGC-GE and GG and the DNA-binding surface protein ComGF and the ComC-mediated prepilin cleavage (Chen & Dubnau, 2004).

All Lactobacillus genomes have single copies of the comGA-GC operon, and most species also have a comC homologue (all except L. delbrueckii and Lactobacillus fermentum), suggesting that the major constituents of the FPE pathway are present in these lactobacilli. In addition, L. delbrueckii ssp. bulgaricus American Type Culture Collection (ATCC) 111842 appears to encode an additional ComGD prepilin, while L. delbrueckii ssp. bulgaricus ATCC BAA-365 and Lactobacillus brevis ATCC 367 encode a ComGF homologue. Besides the FPE pathway, single copies of comE and comF genes, which are also involved in the DNA-uptake process (Chen & Dubnau, 2004), were also identified in all lactobacilli, except L. delbrueckii ssp. bulgaricus ATCC BAA-365 (Table S1).
Peptide efflux ABC transporters

Specific ABC transporter (Fig. 1) subfamilies that are predominantly involved in export of antimicrobial peptide (e.g. lantibiotics, bacteriocins and competence peptides) (Havarstein et al., 1995) are capable of exporting proteinaceous substrates. For example, ABC exporters are responsible for bacteriocin secretion in L. acidophilus (Dobson et al., 2007) and L. plantarum (Diep et al., 1996). Using the bacteriocin predictor BAGEL (de Jong et al., 2006), we identified 3–12 putative bacteriocins in each Lactobacillus genome. Most of the genes encoding predicted bacteriocins appear to be genetically linked to genes encoding ABC exporters, supporting the notion that peptide export via ABC exporters can be commonly found in lactobacilli.

Lactobacillus exoproteome prediction

The term ‘secretome’ has been used to encompass components of the translocation systems and their protein substrates (Desvaux et al., 2009). However, in this review, we prefer to use the term ‘exoproteome’ to only encompass the chromosomal gene products that are transported across the Lactobacillus cytoplasmic membrane, including molecules that become surface localized, are parts of surface appendages or are released into the environment (Greenbaum et al., 2001). To enable a comprehensive overview of extracellular protein components of the lactobacilli, our exoproteome definition excludes integral membrane proteins with multiple membrane-spanning regions (i.e. transport proteins, sensor kinases, etc.), although it is clear that this extensive group of Lactobacillus proteins may expose significantly sized domains to the extracellular environment and play a key role in bacterial interaction with the environment.

Dedicated efforts to predict the secretome/exoproteome of individual Lactobacillus species have been reported before (e.g. L. plantarum WCFS1; Boekhorst et al., 2006a). To predict the SCL for each of the proteins encoded by the lactobacilli (Table 1), the integrated SCL prediction pipeline provided by LocateP (Zhou et al., 2008) was used. To date, LocateP is the only SCL-pipeline that has successfully dealt with the separation problem of the N-terminally anchored proteins and the truly secreted proteins (defined as proteins with a cleaved signal peptide that are released from the bacterial cell), by incorporating a novel HMM-based N-terminal anchor recognition system into the prediction pipeline, which improved the accuracy of the differentiation of these two groups of proteins to approximately 90% (Zhou et al., 2008).

The predicted Lactobacillus exoproteomes contained two main groups of proteins: the secreted proteins that are released from the bacterial cell and the surface-associated proteins. The latter group could be divided into several subcategories based on different binding mechanisms: (1) proteins that are anchored in the cytoplasmic membrane via a single hydrophobic N- or C-terminal domain; (2) lipid-anchored proteins (lipoproteins) that are N-terminally anchored to long-chain fatty acids of the membrane; (3) proteins covalently anchored to the peptidoglycan via a C-terminal LPxTG motif; and (4) proteins noncovalently bound to the cell surface by various binding domains or

Table 1. Predicted number of extracellular proteins encoded by 13 Lactobacillus genomes

<table>
<thead>
<tr>
<th>Lactobacillus species and strains</th>
<th>Genome size (kbp)</th>
<th>Total number proteins</th>
<th>Total extracellular proteins A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>SCL CWA</th>
<th>CWA</th>
</tr>
</thead>
<tbody>
<tr>
<td>L. acidophilus NCFM</td>
<td>1993</td>
<td>1862 214</td>
<td>5 10 41 54 93</td>
<td>12</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. brevis ATCC 367</td>
<td>2291</td>
<td>2185 239</td>
<td>2 3 16 27 74 105</td>
<td>12</td>
<td>7</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. casei ATCC 334</td>
<td>2895</td>
<td>2751 306</td>
<td>4 11 18 46 46 160</td>
<td>19</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. delbrueckii bulgaricus ATCC 11842</td>
<td>1865</td>
<td>1562 150</td>
<td>1 3 11 25 41 67</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. delbrueckii bulgaricus ATCC-BAA-365</td>
<td>1857</td>
<td>1721 167</td>
<td>1 6 15 22 42 80</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. fermentum IFO3956</td>
<td>2099</td>
<td>1843 128</td>
<td>1 10 1 13 32 66</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. gasseri ATCC 33323</td>
<td>1894</td>
<td>1755 146</td>
<td>1 3 4 31 16 79</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. helveticus DPC4571</td>
<td>2081</td>
<td>1610 149</td>
<td>3 2 22 37 83 12</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. johnsonii NCC533</td>
<td>1993</td>
<td>1821 172</td>
<td>2 9 5 38 17 85</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. plantarum WCFS1</td>
<td>3308</td>
<td>3007 313</td>
<td>6 10 10 47 57 149</td>
<td>27</td>
<td>21</td>
<td>11</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. reuteri DSM20016</td>
<td>2000</td>
<td>1900 117</td>
<td>10 5 14 31 80 5 8</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. sakei 23K</td>
<td>1885</td>
<td>1879 178</td>
<td>2 3 13 27 45 83</td>
<td>3</td>
<td>14</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. salivarius UCC118</td>
<td>1827</td>
<td>1717 172</td>
<td>4 3 17 11 32 101</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The SCL of these proteins and the number of proteins with cell wall anchoring (CWA) domains is predicted (including pseudogenes, but excluding plasmids encoded genes).

SCL: A, C-terminally anchored; B, secreted via minor pathways (bacteriocin-like) (no CS*); C, N-terminally anchored (with CS*); D, lipid-anchored; E, secreted (released) (with CS*); F, N-terminally anchored (no CS*); CS*, cleavage site.

CWA: G, LPxTG cell-wall anchor; H, choline binding domain; I, S-layer protein domain; J, WxL domain; K, LysM domain; L, peptidoglycan-binding domain; M, SH3 domain.
attached to other cell-wall protein(s) via protein–protein interactions (Fig. 1). Several of the cell-wall-binding domains that have been described were searched in the Lactobacillus proteins (Table 1) that were predicted to be secreted according to LocateP (which already includes a search engine for LPxTG anchoring motifs).

**Cova-demtly anchored proteins**

**N- or C-terminally anchored proteins**

The N-terminal signal peptides that target proteins to the Sec translocation pathway contain the characteristic N, H and C regions. Provided that the C region contains a Type-I or a Type-II SPase target sequence, the signal peptide can be cleaved and the mature protein is then released. However, many of the C regions of Sec-translocated proteins do not possess this cleavage motif [or contain a motif similar to the Type-I motif that is not cleaved (Zhou et al., 2008)] and will remain N-terminally anchored in the cell membrane. Many of the proteins that are predicted to be N-terminally anchored contain typical extracellular domains or functionalities and their location at the extracellular side of the cell membrane is highly plausible. In Lactobacillus genomes, the N-terminally anchored proteins constitute the largest group of membrane-anchored proteins (Table 1). These proteins are mainly involved in extracellular bio-processes such as transport, cell-envelope metabolism, competence, signal transduction and protein turnover (Table S2).

In case a signal peptide C region contains a typical Type-I SPase cleavage site and is thus processed, it may still be anchored within the cytoplasmic membrane by a C-terminal transmembrane domain, thereby exposing the mature domain to the extracellular side of the membrane. Lactobacillus genomes encode a variable number of C-terminally anchored proteins, many of which have no known function (Table S2).

**Lipoproteins**

Lipoproteins are the second largest membrane-anchored group in the predicted Lactobacillus exoproteomes (Table 1). These proteins possess a signal peptide and are transported via the Sec pathway. The lipoprotein signal peptides also contain the characteristic N, H and C domains, although the H region is shorter than that in the Type-I signal []peptides (Sutcliffe & Harrington, 2002) and the C region contains the lipobox motif [L-(A/S)-(A/G)-C] that directs them to the lipoprotein biogenesis machinery after transport (Hutchings et al., 2009). The covalent binding of the lipoprotein is generally achieved via diacylglycerol modification of the indispensable Cys-residue in the lipobox by the lipoprotein diacylglycerol transferase. Following lipidation, cleavage occurs N-terminal of the Cys-residue by the Type-II SPase, thereby anchoring the mature protein to the membrane via thiocether linkage (Hutchings et al., 2009). The 13–47 lipoproteins predicted to be encoded by the Lactobacillus genomes mainly encompass the substrate-binding proteins of ABC transporters, but also some proteins that are involved in adhesion, antibiotic resistance, sensory processes, cell-envelope homeostasis and protein secretion, folding and translocation (Table S2).

**LPxTG-anchored proteins**

A well-studied family of proteins that is covalently attached to the peptidoglycan by the activity of the sortase (SrtA) enzyme is characterized by the C-terminal LPxTG (based on the main conserved residues) cell-wall-sorting motif (Boekhorst et al., 2005; Marraffini et al., 2006). LPxTG-containing proteins typically contain an N-terminal signal sequence that contains a Type-I SPase cleavage site in its C region. The LPxTG motif is located in the C-terminal region of the mature domain and is followed by a C-terminal membrane anchor domain, consisting of a stretch of hydrophobic residues and a positively charged tail (Marraffini et al., 2006). The sortase (SrtA) enzyme is a transpeptidase that recognizes the LPxTG motif, cleaves the motif between the T and G residues and covalently attaches the threonine carboxyl group to the peptidoglycan (Marraffini et al., 2006). The Lactobacillus genomes encode a single copy of the sortase (SrtA) and a variable number of LPxTG-motif containing proteins, ranging from two proteins in L. delbrueckii bulgaricus ATCC-BAA-365 and ATCC 11842 and to 27 proteins in L. plantarum WCFS1 (Table 1; Table S2). Although there is some species-specific variation in the amino acids of the LPxTG motifs (Boekhorst et al., 2005), most of the sortase substrates could be readily detected in the Lactobacillus genomes using the HMM from Boekhorst et al. (2005) and have the conserved composition of the motif (Table S2).

**Noncovalent cell-wall-binding domain detection**

Domains that have been described to be involved in cell-wall binding were searched using the Pfam database (http://pfam.sanger.ac.uk/) and the protein sequences identified in this way were inspected manually to verify their accurate detection.

**LysM domains**

The LysM (lysin motif) domain (Pfam PF01476) has been found in many extracellular enzymes that are involved in bacterial cell-wall metabolism, and is suggested to confer a
general peptidoglycan-binding function (Buist et al., 2008). In all Lactobacillus genomes studied here, extracellular proteins were found that contain at least one LysM domain and almost all of these proteins perform cell-wall-related enzymatic functions, in agreement with the proposed role of LysM in peptidoglycan binding (Table 1).

**Choline-binding domains**

The choline-binding domains (Pfam PF01473) are a stretch of 20 amino acids that include multiple conserved tandem copies of aromatic residues and glycines. They are mainly found in extracellular enzymes such as autolysins and muramidases, and are able to bind to choline residues of cell-wall teichoic and lipoteichoic acids (LTA), thereby anchoring the protein to the cell surface (Wren, 1991). In Lactobacillus, these choline-binding domains appear to be present only in L. reuteri, L. fermentum and L. salivarius (Table 1).

**Putative peptidoglycan-binding domains**

Another peptidoglycan-binding domain is composed of three \( \alpha \) helices located at the N or the C terminus of cell-wall-degrading enzymes (Pfam PF00395). A single extracellular protein containing this domain was found in L. plantarum, L. johnsonii, L. casei, L. brevis, L. helveticus and L. gasseri (Table 1).

**S-layer proteins with SLH domains**

S-layer proteins can form a paracrystalline monolayer that coats the surface of bacteria, and are believed to be relevant to cell-wall polysaccharide pyruvlation (Mesnage et al., 2000; Avall-Jääskeläinen & Palva, 2005). In recent years, a number of S-layer proteins have been identified experimentally in L. acidophilus (especially the major S-layer protein SlpA), L. helveticus and L. brevis (Avall-Jääskeläinen & Palva, 2005; Hollmann et al., 2007; Goh et al., 2009; Vilen et al., 2009). The Pfam database contains different HMMs that correspond to S-layer protein domains responsible for noncovalent anchoring to the cell wall (SLAP or PF03217, SLH or PF00395, S_layer_C or PF05124 and S_layer_N or PF05123). Several putative S-layer proteins were found in the genomes of L. acidophilus (14 proteins) and L. helveticus (12 proteins), while a single SlpA protein was identified in L. delbrueckii ssp. bulgaricus ATCC 11842, but not in L. delbrueckii ssp. bulgaricus ATCC-BAA-365 using Pfam and homology searches (Table 1).

**WxL domains**

The C-terminal cell-wall-binding domain designated WxL was first identified in proteins of Lactobacillus and other LAB based on in silico analysis (Chaillou et al., 2005; Boekhorst et al., 2006a; Siezen et al., 2006). WxL domain-containing proteins were found in gene clusters that also encode additional extracellular proteins with C-terminal membrane anchors and LPxTG-type peptidoglycan anchors, suggesting that they form an extracellular protein complex (Siezen et al., 2006). Recently, this domain has been shown to be responsible for noncovalent interactions between certain extracellular proteins and the bacterial cell wall in Enterococcus faecalis (Brinster et al., 2007). In the Lactobacillus exoproteomes, in total, 51 proteins containing a WxL domain were identified, supporting an interaction of these proteins with the peptidoglycan layer via their protein C terminus (Table 1).

**SH3 domains**

The prokaryotic counterparts (SH3b) of the eukaryotic SH3 domains have been proposed to be involved in targeting and binding to the peptidoglycan layer and are thought to recognize specific sequences within the cross-linking peptide bridges (Baba & Schneewind, 1996; Lu et al., 2006; Xu et al., 2009). A search of the Lactobacillus genomes for these SH3b domains identified several proteins in some lactobacilli (Table 1). These proteins appear to function predominantly in cell wall turnover.

**Comparative exoproteomics of Lactobacillus**

In total, 2451 putative extracellular proteins of 13 Lactobacillus genomes were extracted from the LocateP-generated database (Table 1, details in Table S2). The largest predicted exoproteomes are found in L. casei (306 proteins) and L. plantarum (313 proteins), and represent 11.1% and 10.4% of all proteins encoded in these genomes, respectively. The smallest exoproteomes were predicted for L. fermentum (128 proteins, 6.9%) and L. reuteri (117 proteins, 6.1%). The most frequently found subcellular localizations of proteins in these predicted exoproteomes are N-terminal anchoring and secreted proteins, while the smallest category is the C-terminally anchored proteins (Table 1). On average, the functions of up to 60% of these extracellular proteins are unknown. The proteins with a known (putative) function are mostly involved in processes related to cell-envelope metabolism, cell division, transport, competence, signal transduction, protein turnover, exopolysaccharides biosynthesis, secretion, signaling/regulation and extracellular enzymatic or binding functions.

Clustering of all proteins from 12 genomes of nonpathogenic lactic acid bacteria from the order Lactobacillales (including 15 119 proteins from six completed Lactobacillus genomes) using the method of clusters of orthologous groups of proteins (COGs) resulted in 3465 Lactobacillales-specific orthologous protein clusters (LaCOGs)(Makarova et al., 2006; Makarova & Koonin, 2007). These LaCOGs...
included 1335 putative secreted proteins from Lactobacillus, distributed over 338 orthology clusters. We have recently extended these existing LaCOGs with the exoproteomes of 18 newly published LAB genomes (including seven new Lactobacillus genomes) using BLASTP (Altschul et al., 1990), Inparanoid (O’Brien et al., 2005) and in-house protein clustering algorithms (Zhou et al., unpublished data), and stored all information in the LAB-Exoproteome database (http://www.cmbi.ru.nl/lab_exoproteome/). Here, we restrict our analysis to the LaCOG information relevant for the predicted Lactobacillus exoproteomes (see also Tables S2 and S3).

The predicted exoproteins in lactobacilli were clustered into the 338 LaCOGs, placing approximately 76% of the total exoproteome into these orthologous groups. In most of the clusters, the majority of the member proteins have identical predicted SCLs and similar functionalities (including 209 LaCOGs of conserved hypothetical proteins). Clusters with known functions are mainly involved in typical cell wall- or surface-associated functionalities (Table S3). A total of 28 orthologous groups were found to be conserved in all Lactobacillus genomes, and include, for example, the housekeeping protease HtrA (LaCOG01440) and proteins involved in cell-wall biosynthesis [LaCOG00243, penicillin-binding protein (PBP) 2B], cell division (LaCOG01506, cell shape-determining protein MreC) and competence (LaCOG00097, DNA-entry nuclease) (Tables S2 and S3).

Conserved clusters represented within the majority of the Lactobacillus genomes consist of extracellular enzymes, such as carboxy-terminal proteinase (LaCOG01825), ATP synthase (LaCOG01172), Zn-dependent protease (LaCOG01979) and linoleic acid isomerase (LaCOG00663). Moreover, multiple homologous proteins from one genome are found in some of these LaCOGs, such as the four different members of the cell-wall protease Prt family (LaCOG 90024) in L. casei, and the four paralogous genes for cell-surface hydrolases (LaCOG01138) in L. plantarum. These distributions of LaCOG representative proteins provide important insights toward understanding the molecular evolution, diversity, function and adaptation of the lactobacilli to specific environments (Makarova et al., 2006; Makarova & Koonin, 2007). An example is provided by the LaCOG distribution of the mucus-binding proteins in Lactobacillus genomes. In total, 47 proteins with mucus-binding domain(s) were found in the exoproteomes of six Lactobacillus genomes, distributed over six separate LaCOGs. The largest cluster, LaCOG 01470, contains 14 proteins that possess either the MucBP (Pfam PF06458) domain or the recently defined extended mucus-binding domain MUB (Boekhorst et al., 2006b), or both domains. LaCOG00885 contains proteins that have only MucBP domains. In LaCOG 01470, most proteins contain a YSIRK signal peptide in their N terminus, which is a typical characteristic of the gut L. acidophilus group lactobacilli (Baë & Schneewind, 2003; Boekhorst et al., 2006b). The mucus-binding proteins in group LaCOG00885 contain no YSIRK signal peptide, and the cluster contains only proteins from the typical plant lactobacilli L. plantarum and L. brevis (Fig. 2).

**Lactobacillus cell-wall molecular biology**

The Lactobacillus exoproteomes contain a variety of proteins that are proposed to be anchored (covalently or noncovalently) to the basic components of the bacterial cell wall, such as peptidoglycan, teichoic acid or polysaccharide. In addition, these basic cell-wall components play an essential role in communication mechanisms with the host environment encountered within the gastrointestinal tract, including direct signaling with the host tissues. These notions support a more extensive review of the Lactobacillus cell-wall building blocks.

**Peptidoglycan**

In lactobacilli, like in all eubacteria, peptidoglycan is an essential and specific cell-wall polymer found outside of the cytoplasmic envelope. Its main function is to preserve cell integrity from internal turgor pressure, which is of the order of 20 atm in Gram-positive bacteria. In addition, peptidoglycan is an important determinant of cell shape and serves as a scaffold for the covalent anchoring of other cell-wall polymers, wall teichoic acids (WTA) and wall polysaccharides (WPS), and some surface proteins (Fig. 3) (Delcour et al., 1999; Vollmer et al., 2008a).

Peptidoglycan is composed of glycan strands consisting in their unmodified form of alternating residues of β-1-4-linked N-acetyl muramic acid (MurNAc) and N-acetylglucosamine (GlcNAc) cross-linked by short peptides. The D-lactoyl residue of the MurNAc is substituted by a pentapeptide ending in D-Ala-D-Ala or pentadepsipeptide ending in D-Ala-D-Lac (D-Lac, D-lactate), whose composition in lactobacilli in its unmodified form is L-Ala\(^{3}\)-β-D-Glu\(^{2}\)-(L-Lys or meso-A\(^{2}\)-pm or L-Orn)\(^{3}\)-D-Ala\(^{4}\)-(D-Ala or D-Lac)\(^{5}\) [2,6 diaminopimelate (A\(^{2}\)-pm); ornithine (Orn)] (Kandler & Weiss, 1986; Delcour et al., 1999; Lebeer et al., 2008). Many modifications of this basic composition are found in the glycan strands and its associated stem peptides (Fig. 3).

In lactobacilli, N-deacetylation of GlcNAc/MurNAc (L. fermentum) and 6-O-acetylation of MurNAc (L. plantarum, L. casei, L. acidophilus and L. fermentum) of glycan strands has been reported (Fig. 3) (Delcour et al., 1999; Vollmer, 2008; E. Bernard, unpublished data). Both modifications play important roles in the physiology of Gram-positive bacteria and in their interactions with their hosts, such as an increased resistance to lysozyme that could help to escape the innate immune system (for a recent review, see Vollmer, 2008). Besides the well-recognized resistance to lysozyme of lactobacilli, the functional role of these two modifications has not yet been investigated in this group. In silico analysis of complete genome sequences of...
12 *Lactobacillus* species (Table 2) revealed the absence of a GlcNac deacetylase gene (*pgdA*) in most lactobacilli, while nearly all of them contain at least one copy of a putative MurNac O-acetyltransferase gene (*oatA*) (Table 2). Notably, two paralogues were found in *L. plantarum* WCFS1 and *L. sakei* 23K, which suggest an important role of O-acetylation in these two species. On the other hand, the absence of *oatA* in *L. delbrueckii* ssp. *bulgaricus* ATCC 11842 suggests a lack of importance of this function in the milk niche. Preliminary results of the analysis of an OatA-deficient strain of *L. plantarum* WCFS1 confirmed its contribution to lysozyme resistance (E. Bernard & P.A. Bron, unpublished data). The contribution of O-acetylation to the recognition of peptidoglycan fragments by host receptors [Toll-like receptors (TLR), nucleotide-binding oligomerization domain proteins (NOD)] and/or escape of innate immune defenses remains to be investigated in lactobacilli.

Variations in the composition, cross-linking and postmodifications of stem peptides in lactobacilli are mostly present at positions 2, 3 and 5. In most lactobacilli, an L-Lys residue is found in position 3 connected to a D-Asp included in the cross bridge (L-Lys-D-Asp type) of two adjacent stem peptides, but L-Orn-D-Asp or meso-A2pm direct types of linkage are also found (e.g. in *L. fermentum* or *L. plantarum*, respectively) (Fig. 3 and Table 2) (Schleifer & Kandler, 1972; Kandler & Weiss, 1986). The D-Asp residue is generated from L-Asp by an aspartate racemase (RacD) before its ligation to L-Lys, while AslA is remarkably absent in *L. plantarum* (meso-A2pm direct) (Table 2). Surprisingly, *L. plantarum* contains an *racD* orthologue, suggesting that aspartate racemase could be required for a metabolic...
function other than peptidoglycan biosynthesis in this species. In terms of postmodifications, amidations of d-Glu at position 2 (yielding d-iso-Gln), of meso-A2pm and of d-Asp (yielding d-iso-Asn) have been identified in L. casei or L. plantarum (Fig. 3) (Billot-Klein et al., 1997; Asong et al., 2009). However, little is known about the functional role of these amidations in lactobacilli. In Lactococcus lactis, the asnH gene encoding asparagine synthase was recently shown to be responsible for d-Asp amidation, and the deficiency of d-Asp amidation in this species resulted in an increased sensitivity to cationic antimicrobials, and affects the activity of endogenous autolysins (Veiga et al., 2009). Orthologues of asnH were detected in all genome sequences of lactobacilli with the l-Lys-d-Asp type, suggesting that amidation of d-Asp is a general feature in lactobacilli. Intriguingly, an asnH orthologue was also found in L. plantarum WCFS1 (meso-A2pm direct), which could be involved in amidation of a residue of the stem peptide other than d-Asp, possibly meso-A2pm or d-Glu (E. Bernard, unpublished data). Notably, most of these modifications (meso-A2pm vs. l-Lys, amidation of d-Glu and meso-A2pm) of stem peptides of peptidoglycan fragments affect recognition by the host receptors [e.g. NOD1, NOD2, peptidoglycan recognition proteins (PGRPs), TLR2] of the host innate immune system (illustrated for NOD1 and NOD2 in Fig. 3) (Girardin et al., 2003; Roychowdhury et al., 2005; Wolffert et al., 2007; Asong et al., 2009). For example, amidations of meso-A2pm and d-Glu of L. plantarum were shown to modulate TLR2 recognition (Asong et al., 2009). These variations among lactobacilli could significantly impact on their immunomodulatory properties and thus affect their probiotic function.

A remarkable feature of many lactobacilli is their intrinsic resistance to the glycopeptide vancomycin (VanR) (Table 2). In L. plantarum and L. casei, where vancomycin resistance (VanR) has been investigated, this antibiotic resistance is the result of the 100% incorporation of D-Lac instead of D-Ala at position 5 of the stem peptide (Fig. 3) (Ferain et al., 1996; Billot-Klein et al., 1997; Delcour et al., 1999; Goffin et al., 2005; Degoerain et al., 2007). The D-Ala-D-Lac terminus has a 1000-fold decreased affinity for vancomycin compared with the D-Ala-D-Ala terminus. In enterococci, vancomycin resistance by D-Ala/D-Lac substitution is acquired in most cases by the transfer of a large transposon (e.g. TnJ546), encoding the van genes responsible for the reprogramming of the biosynthesis of peptidoglycan precursors (for a recent review, see Mainardi et al., 2008). By contrast, the lactobacilli that are intrinsically resistant to vancomycin produce D-Lac in variable amounts as an end-product of fermentation (Table 2). However, this feature is also found in vancomycin-sensitive species (e.g. L. helveticus, L. delbrueckii spp. bulgaricus). A D-lactate dehydrogenase gene (ldhD) or a D-hydroxyisocaproate dehydrogenase gene (hicD) is present in nearly all lactobacilli, the latter being responsible for the production of a low amount of D-Lac in the vancomycin-resistant L. casei species (Viana et al., 2005). D-Lac production can also be achieved from the conversion of L-Lac into D-Lac by a lactate racemase (lar operon) as shown recently for D-Lac production in L. plantarum (Goffin et al., 2005). The lar operon is also present in L. brevis, L. fermentum and L. sakei (Table 2). The lack of an identifiable ldhD gene in L. sakei suggests that lactate racemization is the only route for D-Lac production (Malter et al., 1998). In L. plantarum and L. sakei, altering D-Lac production results in the loss of the VanR phenotype (Goffin et al., 2005; P. Hols, unpublished data). Furthermore, in L. plantarum, complete abolition of D-Lac production (ldhD, lar double mutant) results in a growth arrest that can be fully restored by external D-Lac supplementation or
Table 2. Specific features of peptidoglycan and TA and their associated genes in 12 complete genome sequences of lactobacilli

<table>
<thead>
<tr>
<th></th>
<th>L. acidophilus NCFM</th>
<th>L. brevis ATCC 367</th>
<th>L. casei ATCC 334</th>
<th>L. delbrueckii ATCC 11842</th>
<th>L. fermentum IFO 3956</th>
<th>L. gasseri ATCC 33323</th>
<th>L. helveticus DPC 4571</th>
<th>L. johnsonii NCC 533</th>
<th>L. plantarum WCFS1</th>
<th>L. reuteri DSM 20016</th>
<th>L. sakei 23K</th>
<th>L. salivarius UCC118</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peptidoglycan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>oatA1</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>oatA2</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pgdA</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>aslA</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>racD</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>asnH</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Van* phenotype*</td>
<td>S</td>
<td>R</td>
<td>R</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>α-lact production*</td>
<td>(α, β)</td>
<td>(α, β)</td>
<td>(α, β)</td>
<td>(α, β)</td>
<td>(α, β)</td>
<td>(α, β)</td>
<td>(α, β)</td>
<td>(α, β)</td>
<td>(α, β)</td>
<td>(α, β)</td>
<td>(α, β)</td>
<td>(α, β)</td>
</tr>
<tr>
<td>LdhD (hicD)</td>
<td>+</td>
<td>+</td>
<td>hicD</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+ hicD</td>
</tr>
<tr>
<td>lar operon</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ddl (Y- or F-type)</td>
<td>Y</td>
<td>F</td>
<td>F</td>
<td>Y</td>
<td>F</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Teichoic acids</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tag genes</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>LTA type*</td>
<td>Gro?</td>
<td>Gro?</td>
<td>Gro-P</td>
<td>Gro-P</td>
<td>?</td>
<td>Gro-P</td>
<td>Gro-P</td>
<td>Gro-P</td>
<td>Gro-P</td>
<td>Gro-P</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>luxS1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>luxS2</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>α-alanylated TA*</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>dlt operon</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

*Specific features or phenotype as reported in the literature for the species.
†Variable among genomes sequences of the same species.
+ presence; ‒ absence; ?, unavailable or unclear; S, sensitive; R, resistant; (α, β), racemic mixture of L- and D-Lac; (α), small amount of D-Lac; (α), exclusive production of D-Lac; Gro?, TA containing glycerol.
only partially by expressing a D-Ala-D-Ala-forming ligase (Goffin et al., 2005). In this species, the cell-wall biosynthesis machinery is specifically dedicated to the production of D-Lac-ended peptidoglycan precursors. Mutation analysis has shown that the specificity of the Ddl ligases for either D-Ala-D-Ala or D-Ala-D-Lac is associated with either a phenylalanine (F type) or a tyrosine (Y type), respectively, at a specific position in the enzyme [position 216 in DdlB of Escherichia coli (F type; Park et al., 1996) and 261 in Ddl of Leuconostoc mesenteroides (Y type; Park & Walsh, 1997)]. Interestingly, Ddl enzymes from all VanR lactobacilli are of the F type, while the VanS species possess Y-type enzymes (Table 2). This observation strongly suggests that vancomycin resistance in lactobacilli takes place principally by a reprogramming of the biosynthesis of peptidoglycan precursors. The vancomycin resistance among lactobacilli may reflect the selective advantage of this phenotype in niches that also contain glycopeptid antibiotic producers, which may especially hold true for Lactobacillus species with a broader niche specificity or those that are associated with plant fermentations such as L. plantarum, L. brevis, L. casei and L. fermentum.

Peptidoglycan is continuously remodeled during growth by the action of a variety of peptidoglycan hydrolases (PGH). These enzymes are involved in separation of daughter cells, peptidoglycan turnover and autolysis in the stationary phase. They are also involved in many other processes such as adhesion, biofilm formation, resuscitation of dormant cells or alloysis in genetic transformation (for a recent review, see Vollmer et al., 2008b). Through autolysis in the host and cell-wall turnover, lactobacilli could release muramyl-peptides that are known to interact with receptors of the immune system. For instance, muramyl-peptides from L. plantarum ATCC 8014 display immunoadjuvant activity, but the in vivo role of peptidoglycan fragments of lactobacilli remains largely unexplored (Kotani et al., 1975). In silico analysis of the PGH content of lactobacilli shows that besides low-molecular-weight PBPs (carboxypeptidases) that are mainly involved in peptidoglycan maturation, they display a variety of PGH, from 14 members in L. acidophilus to 26 in L. reuteri (Layec et al., 2008). These PGH are distributed into four classes: N-acetyl-glucosaminidases/-muramidases and lytic transglycosylases hydrolyzing the glycan strands; N-acetylmuramyl-L-alanine amidases separating the stem peptides from the glycan strands; and endopeptidases of the NLPC/P60 or CHAP families hydrolyzing a range of bonds of the cross-linked stem peptides (Fig. 3) (Layec et al., 2008; Vollmer et al., 2008b). A more detailed examination of the PGH complement (16 genes) of L. plantarum WCFS1 revealed a high level of redundancy in lytic transglycosylases (six members), glucosaminidases/muramidases (five members) and NLPC/P60 endopeptidases (four members), while a single L-alanine amidase is present (Table 3). Redundancy in these three classes is a general feature in lactobacilli, with some variations such as an overrepresentation of lytic transglycosylases in L. plantarum and endopeptidases in L. reuteri. A recent systematic inactivation of nine PGHs of L. plantarum WCFS1 shows that the inactivation of only two [lp_2645 (acm2) glucosaminidase/muramidase and lp_3421 endopeptidase] has a significant impact on cell morphology (T. Rolain, unpublished data). These two PGH and the endopeptidase lp_2162 were recently identified as cell-wall-associated proteins of L. plantarum using a proteomic approach, reinforcing their functional role in this species (Beck et al., 2009). Remarkably, 12 out of 16 PGH of L. plantarum display a modular organization, where the catalytic domain is associated with a peptidoglycan-binding domain (one to five SH3 motifs or one to two LysM motifs) and systematically to a domain rich in alanine, serine and threonine (AST motif) (Table 3). This last domain is suspected to be glycosylated, but its functional role is unexplored. The modular organization of PGH in lactobacilli is a general feature, because at least seven types of domains in addition to LysM and SH3 have been identified (for a recent review, see Layec et al., 2008). In addition to cell-wall binding and targeting PGH to their site of action, these domains could fulfill other biological functions such as adhesion by binding to receptors on eukaryotic cells (Layec et al., 2008; Vollmer et al., 2008b). Besides the functional role of Acm2 from L. plantarum WCFS1 in cell separation (Palumbo et al., 2006) and the endopeptidase activity of the S-layer protein of L. acidophilus ATCC 4356 (Prado et al., 2008), this important class of exoenzymes is poorly characterized in lactobacilli.

To conclude, small variations in the composition and modifications of peptidoglycan, as well as the endogenous capacity to release muramyl-peptides among lactobacilli, are strongly suggested to be important in host–microorganism interactions and adaptation to the ecological niche. Future work aiming to modulate the fine structure of peptidoglycan and/or the content of PGH could help to better understand these important roles.

**Teichoic acids (TA)**

Besides peptidoglycan, the cell wall of lactobacilli comprises TA, which are essential anionic polymers of Gram-positive bacteria and represent up to 50% of the cell-wall dry weight. TA are involved in various aspects of the functionality of the cell wall. Together with peptidoglycan, they form a polyanionic matrix or gel contributing to the porosity, elasticity and electrostatic steering of the cell envelope. They are also involved in cation homeostasis, in particular of Mg$^{2+}$ and protons, the latter being important in the maintenance of a proton gradient in the cell wall. Among a large range of identified biological functions, TA participate in the
modulation of the activity of PGHs, the binding of surface proteins, phage adsorption, cell adhesion and interaction with the immune system (counterpart of Gram-negative lipopolysaccharide) (Delcour et al., 1999; Neuhaus & Baddiley, 2003).

Lactobacilli deserve specific mention in terms of TA since they were initially discovered by Baddiley and colleagues in *L. plantarum* (formerly *Lactobacillus arabinosus*) (Baddiley, 1989). Most lactobacilli investigated for their TA content possess two types of anionic polymers: WTA that are covalently bound to MurNac of peptidoglycan glycan strands via a linkage unit (Fig. 3) and LTA that are anchored on the cytoplasmic membrane through a glycolipid, but that are also found to be loosely bound to peptidoglycan and even released into the extracellular medium. Both TA types are decorated by D-alanyl esters associated or not with glycosyl (mainly glucose) substitutions (Sharpe et al., 1964; Kandler & Weiss, 1986; Fischer et al., 1990; Delcour et al., 1999; Neuhaus & Baddiley, 2003; Lebeer et al., 2008).

The characterized WTA of lactobacilli are generally composed of polyglycerophosphate [poly(Gro-P)], but with some variations such as the presence of polyribitolphosphate [poly(Rbo-P)] in around half of the strains of *L. plantarum* (e.g. ATCC 8014) (Tomita et al., 2008, 2009) or a complete absence in *L. casei* (Kandler & Weiss, 1986) (Table 2). Although the genetic determinants of WTA biosynthesis have not been investigated in lactobacilli, the tag and tar genes and their products responsible for poly (Gro-P) and poly(Rbo-P) WTAs, respectively, have been characterized extensively in *B. subtilis* and *Staphylococcus aureus* (Lazarevic et al., 2002; Brown et al., 2008). Although initially reported as essential, depletion of WTA was recently achieved in both species by inactivation of the gene encoding the first enzyme of its biosynthetic pathway (D’Elia et al., 2006a, b). Nevertheless, WTA play a critical role in cell-shape maintenance in *B. subtilis* and Mg$^{2+}$ dependence for growth (Schirner et al., 2009). Orthologues of tagO/tarO are present in most species, with the remarkable exceptions of *L. casei*, which was previously reported to be WTA deficient and also in *L. fermentum* and *L. reuteri*. In these three species, all the tag/tar orthologues are completely absent, showing that the presence of this secondary anionic polymer is a variable feature among lactobacilli and suggesting that LTA in these species is sufficient to confer all the important biological functions of TA. Among lactobacilli, WTA ultrastructures of *L. plantarum* are the best characterized including the nature of the linkage unit (Kojima et al., 1985; Tomita et al., 2008, 2009). The WTA structure in this species was reinvestigated recently by nuclear magnetic resonance (NMR), and the monomeric units of poly(Gro-P)/poly(Rbo-P) WTAs were shown to be decorated not only by D-alanyl substitutions but also by multiple glucose residues in a range of configurations, including kojibiose and one or two glucose intercalated in the main chain of poly(Gro-P) (at least nine different types) (Tomita et al., 2008, 2009). These analyses revealed a high structural diversity in *L. plantarum*, suggesting that the WTA structure is important for its lifestyle. Although the *in silico* annotation of tag/tar genes is complicated due to similarities between the different transferases/polymerases, *L. plantarum* WCFS1 contains all the necessary tag/tar genes, which are scattered in eight different loci, in contrast to *B. Subtilis*, where all tag or tar genes are clustered.

Table 3. *In silico* analysis of the PGH content in *Lactobacillus plantarum* WCFS1

<table>
<thead>
<tr>
<th>Gene</th>
<th>Family</th>
<th>Size (aa)</th>
<th>SS</th>
<th>CBD</th>
<th>AST domain</th>
<th>LaCOG</th>
</tr>
</thead>
<tbody>
<tr>
<td>acm2 (lp_2645)</td>
<td>Glucoaminidase/muramidase</td>
<td>785</td>
<td>+</td>
<td>5 SH3</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>lp_3093</td>
<td></td>
<td>860</td>
<td>+</td>
<td>5 SH3</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>acm3 (N-M-C)$^*$</td>
<td></td>
<td>612</td>
<td>+</td>
<td>3 SH3</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>acm1 (lp_1138)</td>
<td></td>
<td>213</td>
<td>+</td>
<td></td>
<td>–</td>
<td>00918</td>
</tr>
<tr>
<td>lygA (lp_1158)</td>
<td></td>
<td>258</td>
<td>+</td>
<td></td>
<td>–</td>
<td>01725</td>
</tr>
<tr>
<td>lygH (lp_1982)</td>
<td>i-Ala amidase</td>
<td>282</td>
<td>+</td>
<td>1 SH3</td>
<td>–</td>
<td>01848</td>
</tr>
<tr>
<td>lp_3421</td>
<td>Endopeptidase Nlpc/P60</td>
<td>370</td>
<td>+</td>
<td>1 LysM</td>
<td>+</td>
<td>90015</td>
</tr>
<tr>
<td>lp_2162</td>
<td></td>
<td>496</td>
<td>+</td>
<td>2 LysM</td>
<td>+</td>
<td>90015</td>
</tr>
<tr>
<td>lp_2520</td>
<td></td>
<td>297</td>
<td>+</td>
<td></td>
<td>–</td>
<td>00646</td>
</tr>
<tr>
<td>lp_1242</td>
<td></td>
<td>243</td>
<td>+</td>
<td></td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>lp_0302</td>
<td>Lytic transglycosylase</td>
<td>267</td>
<td>+</td>
<td>1 LysM</td>
<td>+</td>
<td>01094</td>
</tr>
<tr>
<td>lp_3014</td>
<td></td>
<td>204</td>
<td>+</td>
<td>1 LysM</td>
<td>+</td>
<td>01094</td>
</tr>
<tr>
<td>lp_3015</td>
<td></td>
<td>220</td>
<td>+</td>
<td>1 LysM</td>
<td>+</td>
<td>01094</td>
</tr>
<tr>
<td>lp_0304</td>
<td>Lytic transglycosylase (WY domain)</td>
<td>212</td>
<td>+</td>
<td>1 LysM</td>
<td>+</td>
<td>01589</td>
</tr>
<tr>
<td>lp_2845</td>
<td></td>
<td>314</td>
<td>+</td>
<td>1 LysM</td>
<td>+</td>
<td>01589</td>
</tr>
<tr>
<td>lp_2847</td>
<td></td>
<td>354</td>
<td>+</td>
<td>1 LysM</td>
<td>+</td>
<td>01589</td>
</tr>
</tbody>
</table>

*Shown as cell-wall associated (Beck et al., 2009).

$^*$Pseudogene in three fragments (N-M-C).

+, presence; –, absence; SS, signal sequence; CBD, cell-wall binding domain.
together (Lazarevic et al., 2002). The redundancy of some genes (e.g. two putative tagD and tagF, three putative tagB and for some a higher similarity to tar genes) suggests that L. plantarum has the genetic content to produce both types of WTAs. Furthermore, six genes code for putative glucosyltransferases similar to TagE, which was shown in B. subtilis to be responsible for WTA glucosylation (Honeyman & Weiss, 1986; Fischer et al., 1990; Delcour et al., 1999; Neuhaus & Baddiley, 2003). Until recently, no gene was identified that could specifically abolish the biosynthesis of this important cell-wall constituent. The recent discovery of the LTA synthase LtaS [poly(Gro-P) polymerase] in S. aureus and B. subtilis (four paralogues) and their inactivation allowed specific blocking of LTA biosynthesis (Grundling & Schneewind, 2007; Schirner et al., 2009). The LtaS-deficient mutants of both species are viable, but display several defects, including a strongly altered cell morphology and problems in septa positions in S. aureus and a lack of control of cell elongation and separation in B. subtilis (Grundling & Schneewind, 2007; Oku et al., 2009; Schirner et al., 2009). The latter morphological defect is proposed to result from a modified Mg2+/proton homeostasis, which affects enzymes of peptidoglycan biosynthesis and the role of LTA in the control of PGHs (Schirner et al., 2009). The lethality of a double depletion of WTA and LTA in both species shows that at least one anionic polymer is essential for cell survival (Oku et al., 2009; Schirner et al., 2009). All complete genomes of lactobacilli contain at least one copy of ltaS, with a second copy identified in most species (Table 2). In contrast to the apparent variability of WTA synthesis among lactobacilli, the conservation of at least one ltaS gene suggests that these species produce LTA consistently. The LTA structure of four Lactobacillus strains (L. plantarum WCFS1, L. rhamnosus GG, L. reuteri 100-23 and L. delbrueckii ssp. lactis ATCC 15808) was investigated recently by NMR (Palumbo et al., 2006; Perea Velez et al., 2007; Walter et al., 2007; Raisanen et al., 2007; Schirner et al., 2009). The chains comprise 20–22 Gro-P in L. plantarum and L. reuteri, and 33 and 50 residues in L. delbrueckii and L. rhamnosus, respectively. D-Alanyl esters are unique substituents in L. plantarum (42% D-Ala:GroP) and L. rhamnosus (74% D-Ala:GroP), and are associated with a low level of glucosyl residues in L. reuteri (76% D-Ala:GroP and 6% Glc:GroP) and L. delbrueckii (24% D-Ala:GroP and 3% Glc:GroP). This detailed analysis shows that LTA of lactobacilli are variable in chain length, percentage and composition of substitutions, and possibly in the nature of the lipid anchor.

D-Alanyl esters of LTA (and WTA) strongly contribute to the function of TA because the positively charged amino groups partially counteract the negative charges of the backbone phosphate groups. D-Alanylation of TA is performed by at least four proteins encoded in the dlt operon. The biochemistry of D-alanylation was principally studied by Neuhaus and colleagues using L. rhamnosus 7469 as a model (for an extensive review, see Neuhaus & Baddiley, 2003). The importance of D-alanyl esters in lactobacilli is reinforced by the presence of a dlt operon in all complete genomes of lactobacilli (Table 2). The impact of a depletion of D-alanyl esters of TA was investigated in three of the four strains cited above by the inactivation of the dlt operon (Palumbo et al., 2006; Perea Velez et al., 2007; Walter et al., 2007); the dlt mutants had either strongly reduced or no D-alanyl esters, but the chain length was, respectively, threefold increased in a subpopulation of LTA of L. plantarum WCFS1, 1.7- and 8-fold reduced in the two subpopulations of LTA of L. rhamnosus GG and unaffected in L. reuteri 100-23. Interestingly, the level of glucosylation was increased from undetectable in LTA of the wild-type L. plantarum WCFS1 to 24% substitution in the dlt mutant. A fivefold increase was also observed in L. reuteri 100-23, but no change could be detected in L. rhamnosus GG. In the latter, a modification of the lipid anchor was reported. These modifications suggest that D-alanylation contributes directly or indirectly to the chemical composition of LTA in these three species. The more negatively charged TA in the absence of D-alanyl esters results in an increased sensitivity to positively charged antimicrobial peptides (e.g. nisin) in all three mutants. An increased sensitivity to human β-defensin-2 is also reported for L. rhamnosus GG. This suggests that the positive charges of D-alanyl esters in lactobacilli contribute to the general defense against positively charged antimicrobial molecules. The depletion of D-alanyl esters also results in elongated cells in L. plantarum and L. rhamnosus, a defect in cell separation in L. rhamnosus, perforations of the cell wall at division sites in L. plantarum, an increased autolysis in the three mutants and a strongly affected growth in L. plantarum. All these phenotypic modifications suggest that the machinery of peptidoglycan assembly/degradation is affected, probably due to a charge modification affecting the proton gradient, homeostasis of cations and the control of positively charged PGHs. The Acm2 PGH of L. plantarum, involved in cell separation, was shown to participate to this pleiotropic phenotype (Palumbo et al., 2006). Acid tolerance is strongly reduced in the mutants. Thus, in L. reuteri and L. Rhamnosus, growth at a low pH was affected, and a higher sensitivity to gastric juice challenge was reported for L. rhamnosus, but the resistance to acidic conditions per se appeared to be unchanged in L. reuteri. In L. reuteri, which is a true resident
of the rodent gut and forms a biofilm in the forestomach of Lactobacillus-free mice, the depletion of ω-alanyl esters strongly impairs gut colonization and the formation of biofilms in the forestomach. Mutant cells present in the mouse forestomach display damaged cell envelope structures, showing that the high level of ω-alanylation protects this species under hostile conditions prevailing in the rodent forestomach (Walter et al., 2007). In L. johnsonii La1, cell-surface-associated LTA is important for adhesion to human enterocyte-like Caco-2 cells, possibly through hydrophobic interactions (Granato et al., 1999). In L. plantarum WCFS1, ω-alanyl esters modulate specific immune responses (Grangette et al., 2005; Rigaux et al., 2009). Notably, the dlt mutant is more protective compared with the wild-type strain in a mouse model of colitis. In vitro, the mutant induces a dramatically increased level of the anti-inflammatory cytokine IL-10 in peripheral blood monocytes. The immunomodulation effect of purified LTA of both wild type and mutant is largely TLR-2-dependent (Grangette et al., 2005). However, this interesting impact of a variation in the LTA composition observed with the L. plantarum mutant seems species-specific, because the dlt mutant of L. rhamnosus is not affected compared with the wild type in similar in vitro immune modulation experiments (Perea Velez et al., 2007). In addition, the expression of the dlt operon of L. plantarum WCFS1 was modulated in the mouse cecum following a modification of the diet, showing that this species adapts its level of ω-alanylation in vivo in response to environmental conditions (Marco et al., 2009).

Overall, TA of lactobacilli contribute to many aspects of the extracellular characteristics of lactobacilli and affect many phenotypic traits, including cell adhesion, biofilm formation, survival under hostile conditions and interaction with the immune system. In the future, the selective complete depletion of WTA or LTA by gene inactivation in lactobacilli would certainly help to better understand their specific contribution in commensal–host interactions and adaptation to this ecological niche.

**Extracellular polysaccharides**

Cell-wall polysaccharides are ubiquitous components of the cell envelope of lactobacilli. These polysaccharides, which are generally neutral, can be found to be covalently attached to MurNAC of peptidoglycan glycan strands (WPS) (Fig. 3), loosely associated with the cell envelope or released into the extracellular medium (exopolysaccharides, EPS). When the polysaccharides form a thick shell (capsule) closely associated or bound to the cell envelope, they are generally named capsular polysaccharides (CPS). However, the distinction between WPS, EPS and CPS appears to be somewhat artificial because the abundance and localization of polysaccharides is for instance strongly dependent on the growth conditions (Delcour et al., 1999; Lebeer et al., 2008). Polysaccharides of lactobacilli are generally heteropolysaccharides of complex structures differing in the nature of sugar monomers, the modes of linkage, branching and substitutions. The following sugar moieties are present in Lactobacillus polysaccharides: glucose, galactose, rhamnose, GlcNAC, N-acetylgalactosamine, Glucuronate and Gro-P, along with phosphate, acetyl and pyruvyl modifications (for reviews, see De Vuyst & Degeest, 1999; Lebeer et al., 2008). Polysaccharides biosynthetic genes of lactobacilli are grouped into clusters (up to 25 kb). These clusters are highly variable in terms of glycosyltransferases responsible for the biosynthesis of the repetition unit. This high variability of polysaccharides clusters from strain to strain was recently revealed in the L. plantarum species and in members of the L. acidophilus group using microarray approaches (Molenaar et al., 2005; Berger et al., 2007). The complexity of polysaccharides biosynthesis is even increased by the presence of multiple polysaccharides gene clusters per strain, for example, four gene clusters were identified in the genome of L. plantarum WCFS1.

To date, polysaccharides of lactobacilli have been shown to be involved in phage absorption in L. plantarum and L. casei, in the attachment of the S layer in L. buchneri and in immunomodulation using for instance polysaccharides extracted from L. kefiransfaciens and L. rhamnosus RW-9595M (for reviews, see Delcour et al., 1999; Lebeer et al., 2008). Recently, knockouts of polysaccharides geneclusters have been reported in L. casei Shirot, L. johnsonii NCC533 and L. rhamnosus GG (Denou et al., 2008; Yasuda et al., 2008; Lebeer et al., 2009). In L. casei Shirot, the inactivation of 8 (cps1A to J) out of 10 genes included in a polysaccharides gene cluster led to a large reduction of a high-molecular-mass polysaccharides in the cell-wall fraction. In vitro experiments performed with heat-killed cells on mouse macrophage cell lines suggest a role of this polysaccharides as a relevant immune suppressive modulator of macrophage activation (Yasuda et al., 2008). In L. johnsonii NCC533, deletion of the entire polysaccharides gene cluster eliminates a fuzzy layer decorating the outer rim of the cell wall. Interestingly, this mutant has a slightly increased persistence time in mice, suggesting that surface polysaccharides could hinder other cell adhesions (Denou et al., 2008). Similarly, the inactivation of the priming glycosyltransferase gene weL of one of the polysaccharides gene clusters of L. rhamnosus GG, responsible for the biosynthesis of long galactose-rich polysaccharides, results in increases in adherence to epithelial cells and mucus and in biofilm formation. A compensatory mechanism is present because the mutant shows an increase in the amount and length of a second glucose-rich surface polysaccharides as indicated by atomic force microscopy studies using functionalized tips with lectins (Francius et al., 2008; Francius et al., 2009; Lebeer et al., 2009). Interestingly, electron microscopy of mutant cells not only
revealed the absence of the polysaccharide layer but also an increased exposure of fimbriae or pili (Lebeer et al., 2009). Fimbrial genes have been reported in L. johnsonii NCC533, but they are uncommon among lactobacilli, and the direct visualization of pili on Lactobacillus cells has only been shown for L. rhamnosus GG (Kankainen et al., 2009), where they become more prominently exposed upon removal of the polysaccharides.

Although the above studies all revealed the importance of polysaccharides of lactobacilli in commensal–host interactions, much remains to be learned about biosynthetic pathways, regulation, locations, compositions, size and conformation of polysaccharides in lactobacilli. Moreover, the apparent biological diversity of polysaccharides among Lactobacillus species and strains, combined with its potential role in host–microorganism interaction, supports further research on these interesting molecules.

**Extracellular function in adaptation to the host environment**

**In vitro approaches**

Several biological barriers are encountered by bacteria during residence in and travel through the different parts of the host’s gastrointestinal tract (Fig. 4), such as gastric acidity encountered in the stomach, bile salt and digestive enzyme challenges in the duodenum, a relatively high osmolarity in the colon, as well as stress conditions associated with oxygen gradients that are steep at the mucosal surface, while the colonic lumen is virtually anoxic. Moreover, considerable bacterial competition is encountered throughout the gastrointestinal tract and is most severe in the colon where cell numbers are the highest (Zoetendal et al., 2006a). Several studies describe the *in vitro* response of intestinal bacteria to a simplified model that mimics (components of) the stress encountered in the host’s gastrointestinal tract (for a review, see Lebeer et al., 2008). However, these simplified systems fail to accurately address the physiochemical or the microbial complexity of the intestinal environment *in vivo* (de Vos et al., 2004; Kleerebezem & Vaughan, 2009). Nevertheless, some common responses are apparent, including activation of DNA/protein protection and repair mechanisms, differential regulation of two-component and other regulatory mechanisms and induction of systems for the active removal of acid and bile-related stress compounds (Corcoran et al., 2008; Lebeer et al., 2008).

One particularly relevant aspect in light of this review is the fact that many lactobacilli appear to modify the different macromolecules that constitute the cell envelope, thereby contributing to maintenance of cell integrity during the *in vitro* exposure to a variety of gastrointestinal-tract–relevant stress conditions (Lebeer et al., 2008). For example, lower pH and bile salts influence the fatty acid and phospholipid composition in the cell membranes of L. casei (Fozo et al., 2004) and L. reuteri (Taranto et al., 2003). Moreover, Tween addition to the growth medium of L. rhamnosus, Lactobacillus paracasei and L. salivarius results in up to a 3-log increased survival during exposure to gastric juice. A 55-fold higher oleic acid content in the fatty acid composition of L. rhamnosus GG led to the suggestion that the resultant more rigid structure caused by increased membrane fatty acid saturation levels may explain the observed enhanced survival characteristics (Corcoran et al., 2007).

The genetic factors important during low pH and bile stress have also been investigated in several lactobacilli. Acid shock induces several genes potentially involved in membrane fluidity regulation or peptidoglycan biosynthesis and...
organization in *L. reuteri* ATCC 55730, including a putative phosphatidyl glycerophosphatase and a putative esterase gene, encoding a parologue of PBPs (Wall et al., 2007). Mutation of the latter gene causes increased sensitivity to both gastric juice (Wall et al., 2007) and bile exposure (Whitehead et al., 2008). Similarly, several genes potentially involved in cell-envelope and surface protein biosynthesis are induced by bile exposure of *L. acidophilus* NCFM (Pfeiler et al., 2007) or *L. plantarum* WCFS1 (Bron et al., 2004b, 2006). Furthermore, *L. acidophilus* gene disruption mutants in a cell-division protein (cdpA) and surface layer protein A (slpA) have an increased bile resistance and reduced osmotolerance (Altermann et al., 2004; Klaenhammer et al., 2005), further highlighting the importance of subtle modifications in cell-envelope composition in the resilience of LAB to persist under different stress conditions relevant for gastrointestinal-tract survival.

**In vivo approaches**

Several postgenomics approaches have addressed the molecular adaptation of lactobacilli to the intestinal environment using animal model systems, for example differential promoter activity studies focusing on four *L. casei* promoters revealed that this LAB is metabolically active in the mouse intestine and initiates de novo protein synthesis to adapt to this niche (Oozeer et al., 2005). Furthermore, two genomewide genetic screens in *L. reuteri* (Walter et al., 2003) and *L. plantarum* (Bron et al., 2004a) exploiting *in vivo* expression technology revealed the induction of 3 and 72 genes in the mouse intestinal tract, respectively. Strikingly, the *L. plantarum* protein encoded by lp_2718 is a homologue of the only conserved hypothetical protein that was identified as being *in vivo* induced (ivi) in *L. reuteri* (Walter et al., 2003; Bron et al., 2004a). A subsequent dedicated mutagenesis approach demonstrated that the ivi gene, encoding a methionine sulfoxide reductase B, contributes to the ecological performance of *L. reuteri* in the murine gut (Walter et al., 2005). The *L. plantarum* ivi genes identified include four predicted extracellular proteins (lp_0141, lp_0800, lp_1403 and lp_2940), several sugar PTS transport systems and a copper-transporting ATPase (*copA*) (Bron et al., 2004a). Subsequent qRT-PCR analysis unraveled the spatial and temporal *in vivo* expression patterns of a subset of the identified ivi genes (Marco et al., 2007). Furthermore, a dedicated mutagenesis approach underlined the critical contribution of lp_2940 and *copA* to murine gut persistence (Bron et al., 2007).

More recently, technical advances have allowed the isolation of high-quality bacterial RNA derived from intestinal samples (Zoetendal et al., 2006b). The targeted studies described above were followed by transcriptome approaches to describe bacterial behavior in the gastrointestinal tract. One such approach revealed the expression of specific sets of genes in *L. johnsonii* when it resides in different compartments of the mouse gastrointestinal tract (Denou et al., 2007). While no colon-specific genes were identified, the induction of expression of specific sugar PTS transport systems could be established in the jejenum, the stomach and the cecum. Furthermore, the stomach-specific genes include several multidrug transport systems, a cation efflux protein, as well as a copper-transporting ATPase. This gene induction pattern closely resembles the ivi genes identified in *L. plantarum* using the *in vivo* expression technology approach described above (Bron et al., 2004a; Denou et al., 2007; Marco et al., 2007). Three genetic loci potentially important for intestinal persistence have been identified by correlating differences in murine intestine persistence of two *L. johnsonii* strains with differential genome content and *in vivo* transcriptome data, and subsequent mutagenesis demonstrated the importance in murine intestinal performance of a mannose PTS system and a protein that shares 30% identity with immunoglobulin A proteases from pathogenic bacteria (Denou et al., 2008).

Recently, the transcriptomes of *L. plantarum* WCFS1 in the cecum of mono-associated mice fed either a western-style (high fat, low fiber) or a standard chow diet (low fat, high fiber) were compared with *in vitro* transcriptomes (Marco et al., 2009). Notably, 9 and 32 genes encoding cell envelope- and wall-localized functions were induced in the ceca of the chow- and western diet-fed mice, respectively. *These in vivo* induced genes primarily encode putative cell-wall-anchored and lipoproteins with unknown functions, rather than polysaccharide, peptidoglycan or teichoic acid biosynthetic capacities, illustrating the limited current knowledge of the factors important for *L. plantarum* in this niche. Five genes in this category are induced *in vivo* independent of the dietary regime, including the cell-surface complex operons *cscI, cscVII* and *cscVIII* (Siezen et al., 2006), potentially facilitating the ability of *L. plantarum* to utilize host or dietary glycans in the distal gut habitat (Marco et al., 2009). The expression level of lp_2940 is induced in mice irrespective of the dietary regime used, while lp_0800 is only induced in mice fed a normal mouse chow diet, reiterating the importance of these proteins for the *in vivo* fitness of *L. plantarum* and illustrating the pronounced effect of diet on *in situ* microbial responses. Moreover, the cation efflux protein previously identified utilizing an *in vivo* expression technology approach is upregulated in mice fed a western diet, and induction of genes encoding functions involved in carbohydrate transport and metabolism is the largest induced functional category, independent of the mice dietary regime (Marco et al., 2009).

All studies mentioned above exploit the murine gut as a model for the human gastrointestinal system. However, one recent study has determined the transcriptome profile of...
L. plantarum 299v, a strain closely related to WCFS1, in large intestinal biopsies obtained from patients with possible colon cancer who volunteered to participate in a probiotic trial before surgery (de Vries et al., 2006). Subsequently, the transcriptome profiles obtained were compared with over 100 in vitro transcriptomes available for L. plantarum WCFS1. Like the mouse gastrointestinal transcriptome datasets described above for L. plantarum WCFS1, the in situ profiles for L. plantarum 299v in the human colon suggest that this organism strongly adapts its capacity for carbohydrate acquisition and cell-surface composition. The L. plantarum response in all in vivo samples includes upregulation of the capsular polysaccharide biosynthesis operon cps3 and the cell-surface protein clusters cscI and cscVIII (Marco et al., 2009; M.L. Marco, M. Wels, W.M. de Vos, E.E. Vaughan & M. Kleerebezem, unpublished data). Moreover, lp_0800 and lp_2940 were upregulated in the human colon, corroborating the many observations suggesting the importance of these genes in the intestine (Bron et al., 2004a; Marco et al., 2007, 2009; M.L. Marco, M. Wels, W.M. de Vos, E.E. Vaughan & M. Kleerebezem, unpublished data). These overlapping responses for both L. plantarum WCFS1 and 299v in different gastrointestinal compartments using different mammalian model systems (mono-associated mice fed a western or a chow diet, colonized mice fed a chow diet and human volunteers) suggest a diet-, host- and microbiota-independent core response in multiple L. plantarum strains. Hence, the associated extracellular molecules are robust key factors strongly affecting the (probiotic) functionality of this LAB in the gastrointestinal tract.

**Adhesion to mucus and host mucosa**

The environmental adaptations described above can be anticipated to contribute mainly to the in situ survival and/or the persistence characteristics of the lactobacilli during their passage through the intestinal tract. Another aspect considered important for probiotics is the capacity to adhere to mucosal surfaces and/or tissues. Therefore, a variety of studies have specifically addressed the LAB extracellular adhesins and their contribution in direct microbial interactions with host cells or compounds. These studies typically use in vitro models such as epithelial cell lines, immobilized intestinal mucus or extracellular matrix molecules, including collagen and fibronectin to determine adherence capacities (Velez et al., 2007; Lebeer et al., 2008).

Several Lactobacillus extracellular compounds described as important for adhesion appear to contribute in a rather unspecific manner. For example, LTA provides the main component of hydrophobicity to the Lactobacillus cell envelope, which appears to be the reason for its involvement in the adhesive characteristics of L. johnsonii (Granato et al., 1999), L. rhamnosus (Lebeer et al., 2007) and L. reuteri (Walter et al., 2007). Similarly, exopolysaccharides, which contribute to cell-surface physiochemical properties and shielding of other cell-surface adhesins, is important for the adhesive characteristics of L. acidophilus (Lorca et al., 2002) and L. rhamnosus (Ruos-Madiedo et al., 2006). Moreover, the authors of several papers suggest that pleiotrophic effects could contribute to the altered adhesive characteristics of their constructed mutants. The L. acidophilus cdpA gene encodes a cell-wall-modifying enzyme that promotes cell division, and a mutant has reduced adhesion, which could be explained by a loss of anchoring or translocation or integrity of important adhesins (Altermann et al., 2004).

Genome mining for candidate extracellular adhesins encoded by L. acidophilus led to the identification of five proteins potentially involved in adhesion to epithelial cells (Buck et al., 2005). Mutant studies confirmed the involvement in adhesion to Caco2 cells in vitro, for three of the five proteins selected (Mub, FbpA, SlpA) (Buck et al., 2005). The authors argue that the observed phenotype of the L. acidophilus slpA (encoding surface layer protein) mutant is likely due to the loss of multiple surface proteins that may be embedded in the S layer, whereas the contributions of Mub and FbpA (encoding a mucin-binding and a fibronectin-binding protein, respectively) to adhesion are more likely to be via a specific interaction with epithelial cells. Nevertheless, a similar role in adhesion to intestinal epithelial cells could be established for the S-layer proteins in L. brevis (Hynonen et al., 2002), L. crispatus (Toba et al., 1995; Antikainen et al., 2002) and L. helveticus (Johnson-Henry et al., 2007), whereas the L. brevis SlpA protein also mediates adhesion to extracellular matrices such as fibronectin (Hynonen et al., 2002). These variations to a theme surrounding the surface layer proteins of lactobacilli, including variations in surface layer-associated protein domains, support a relevant role for these proteins in the interaction with the host environment. Such a role in immune cell recognition was recently exemplified for the L. acidophilus SlpA.

A homologue of the L. acidophilus Mub protein described above, with 25% identity at the amino acid level, is present in L. reuteri 1063 (Roos & Jonsson, 2002). The proteins are of similar size and contain a similar number of mucus-binding domains (MucBP domain; 17 and 14 domains, respectively). Interestingly, intact L. reuteri cells can adhere to pig and hen mucus. A direct role for Mub protein in mucin binding was shown using purified fusion proteins consisting of maltose-binding protein fused to various MucBP domain-containing repeats (Roos & Jonsson, 2002). Another adhesion identified in L. reuteri (strain NCIB11951) is the collagen-binding protein (CnBP), which can adhere to solubilized Type-I collagen (Aleljung et al., 1994). Notably, CnBP has sequence similarities to the solute-binding domain of bacterial ABC transporters, a domain also found in the mucus adhesion-promoting...
protein (MapA) that was reported to mediate the binding of *L. reuteri* 104R to Caco-2 cells and mucus (Miyoshi *et al*., 2006). The Mub proteins of *L. acidophilus* and *L. reuteri* mentioned above contain all the required elements for sortase-dependent anchoring to the cell wall. The *L. salivarius* UCC118 genome was predicted to encode 10 sortase-dependent extracellular proteins, and three of the encoding genes were shown to be expressed in vitro (van Pijkeren *et al*., 2006; Table 1). Mutation of either the *ispA* or the sortase-encoding *srtA* gene in *L. salivarius* significantly reduced the capacity to bind to HT-29 cells, whereas *ispC* and *ispD* mutants adhered at similar levels as the wild-type strain. Remarkably, *ispA* encodes a protein containing seven MucBP domains (van Pijkeren *et al*., 2006), and this repeated domain structure appears to be a feature shared by many extracellular proteins that contain MUB or MucBP domains (Fig. 2). The broad distribution of MUB- or MucBP-domain-containing proteins suggests that they may play a conserved role in intestinal adhesion in many lactobacilli, which is corroborated by their relative enrichment in species that are associated with the intestinal niche (Fig. 2). However, it remains to be seen whether the commonly used in vitro cell line models accurately mimic the actual in vivo situation, where host defense systems, competition with the resident microbiota, mucosal shedding and peristaltic flow are likely to modify adhesion (Lebeer *et al*., 2008). Notably, in this respect, sortase-deficient mutants of both *L. plantarum* (Bron *et al*., 2004c; Pretzer *et al*., 2005) and *L. johnsonii* (Denou *et al*., 2008) display a wild-type persistence phenotype in the intestinal tract of mice. These data further emphasize the importance of experiments that would allow translation of the in vitro adhesion data obtained so far to the actual in situ situation in the gastrointestinal tract. A clear example for such translation experiments was recently presented for the *L. rhamnosus* GG pilin-like structures (Kankainen *et al*., 2009). In view of the annotation of *Lactobacillus* exoproteomes and prediction of adherence capacities for individual proteins, it is important to note that the nomenclature of the binding domains mentioned above is debatable, because their assignments have been primarily based on binding ligands present in the in vitro assay in which they were tested. The actual substrate specificity of the mucus-, fibronectin-, collagen-, etc., binding domains remains to be established, and they may recognize specific features (e.g. attached glycosyl residues or other molecular features) present in the currently assigned substrates rather than particular features of the specific proteins per se.

**Probiotic effector molecules**

In contrast to the impressive amount of information on the *Lactobacillus* adaptive response to the host gastrointestinal environment, and the discovery of several adhesins, there is very limited knowledge on the molecular mechanisms by which probiotics exert their health-beneficial effects on the host (Marco *et al*., 2006; Kleerebezem & Vaughan, 2009). To date, only a few candidate probiotic effector molecules have been discovered, and while for some there is convincing evidence for their proposed role in vivo, some others still require validation in situ.

**Inhibition of intestinal pathogens**

In densely populated niches, such as the gastrointestinal tract, lactobacilli are in constant competition for nutrients with each other and other bacteria. Within the framework of probiotic applications, there is a growing interest in the use of dietary lactobacilli for their potential for controlling the gastrointestinal microbial ecosystem to support human health (Corr *et al*., 2009). This possibility is exemplified by the studies that support the suppressive effect of *L. johnsonii* LA1 on *Helicobacter pylori* colonization (Gotteland & Cruchet, 2003), or indicate the modulatory capacity of *L. acidophilus* LB on *H. pylori* infection (Coconnier *et al*., 1998). Several mechanisms for these suppressive effects of lactobacilli toward pathogens have been suggested, including competitive exclusion, immunomodulation, inhibition of virulence expression and/or direct killing or inhibition by antimicrobial peptides (Corr *et al*., 2009). The antimicrobial peptides produced by LAB can be classified into different classes and subclasses, encompassing the lantibiotics (Class I) and the canonical double-Glycine leader bacteriocins (Class II) (for reviews, see Klaenhammer, 1993; Kleerebezem & Quadri, 2001; Eijsink *et al*., 2002). Although the majority of the studies on LAB antimicrobial peptides primarily aim at their possible application as preservative agents in food products ( Cotter *et al*., 2005), some studies also address their potentially important role in intestinal competitiveness and their potential relevance as effector molecules in vivo in probiotic applications such as preventing pathogen infections or more general microbial control in this complex ecosystem. A recently published study elegantly demonstrated this role of bacteriocins at the molecular level (Corr *et al*., 2007). *Lactobacillus salivarius* UCC118 produces Abp118, a Class Ib bacteriocin that peaks in the early stationary phase, a process regulated through quorum sensing via the induction peptide AbpIP. Administration of the Class Ib bacteriocin (Abp118) producing *L. salivarius* UCC118 protected mice against oral infection with *Listeria monocytogenes*, while mice that were administered an Abp118-negative *L. salivarius* mutant were not protected from listeriosis. Moreover, a direct linkage of the protective effects to the Abp118 bacteriocin was proven by the complete abolition of the protective effect of the Abp118-producing *L. salivarius* strain when mice were infected with...
a L. monocytogenes strain expressing the Abp118-immunity protein, AbpIM. This study provides convincing evidence for the potential of bacteriocins produced by lactobacilli and other LAB to prevent diseases associated with intestinal pathogens (Cort et al., 2007).

Induction of the plantaricin immunity protein PlnI during murine gastrointestinal transit of L. plantarum, observed using three independent approaches (Bron et al., 2004a; Marco et al., 2007, 2009), suggests that bacteriocins are also important for the in vivo performance of this species. Moreover, preliminary observations suggest that plantaricin-related functions are important in the capacity of L. plantarum WCFS1 to modulate specific cytokine responses in human blood-derived immune cells in vitro (M. Meijerink, S. van Hemert, N. Taverne, M. Wels, P. de Vos, H. Savelkoul, J. van Bilsen, M. Kleerebezem, P.A. Bron, & J.M. Wells, unpublished data). These experiments further emphasize the importance of bacteriocins in probiotic function, which may involve multiple mechanisms, including immunomodulation and direct killing of pathogens.

The competitive exclusion concept of probiotic functioning proposes that lactobacilli or other health benefit cultures may adhere to epithelial sites that also function as site of adherence of pathogenic bacteria, and thereby lactobacilli may prevent the docking of pathogens onto epithelia and may thus inhibit their infection efficacy. As an example, certain L. plantarum strains have been reported to adhere specifically to mannose-containing moieties present on human intestinal cells (Adlerberth et al., 1996), which is proposed to competitively exclude adhesion of enterotoxigenic E. coli (ETEC) and thereby prevent infection. Using a gene-trait matching approach (Molenaar et al., 2005; Pretzer et al., 2005), the absence or the presence of the L. plantarum WCFS1 lp_1229 gene could be correlated to the capability of 14 L. plantarum strains to agglutinate yeast, an in vitro assay utilized to determine the mannose-specific adherence phenotype (Pretzer et al., 2005). Subsequently, an L. plantarum lp_1229 mutant was found to have lost its ability to agglutinate yeast and, therefore, lp_1229 was designated msa (mannose-specific adhesin). The encoded protein contains two potential MucBP domains and a SasA domain, associated with a ConA-type lectin, further supporting a role for Msa in carbohydrate recognition and binding (Pretzer et al., 2005). Intriguingly, specific regions of the Msa protein sequence vary in different L. plantarum strains; the number of MucBP domains in the protein can range from a single domain up to four or five domains, which may be correlated to the difference observed for their quantitative capacity to adhere to mannose or mannose polymers (Gross et al., 2009). However, no in vivo experiments have been reported that either establish or disprove the competitive exclusion hypothesis and a potential role for Msa therein. Nevertheless, studies using a pig model system revealed that the msa mutant has a decreased association with intestinal epithelia, and increased jejunal fluid absorption. Moreover, preliminary results indicate that expression of the host pancreatitis-associated protein, a protein with proposed bactericidal properties, is only induced by the wild-type strain (Gross et al., 2008). Although these studies do not directly support the competitive exclusion model, they do show that specific adhesin capacities of lactobacilli may affect mucosal biology in situ, which may strengthen defense mechanisms in the host mucosa.

Mucosal integrity

Probiotic applications are also proposed to reinforce mucosal barrier function, by maintaining or supporting epithelial integrity. The molecules of lactobacilli that may be involved in this process remain largely unknown, and their functioning may involve complex host signaling pathways and multiple mechanisms. Interestingly, two abundantly secreted proteins of L. rhamnosus GG, P40 and P75, were recently purified and demonstrated to activate the Akt pathway in epithelial cells. This pathway plays a central role in promoting host cell survival by inactivation of several proapoptotic pathways. In addition, P40 and P75 inhibited cytokine-induced epithelial cell apoptosis, promoted cell growth in human and mouse colon epithelial cells and decreased tumor necrosis factor (TNF)-induced colon epithelial damage (Yan et al., 2007; Seth et al., 2008). Although these elegant in vitro studies illustrate the potential of these molecules to promote intestinal homeostasis through specific signaling pathways and the prevention of, or protection against damage, validation of their in vivo role is still not available.

Immune system modulation

An important conceptual health benefit of probiotics is based on their capacity to promote immunotolerance by priming dendritic cells (DCs) to stimulate the differentiation of IL-10 producing regulatory T-cells (Kleerebezem & Vaughan, 2009). A recent paper supporting this concept reports L. reuteri and L. casei strains that can interact with the C-type lectin receptor DC-SIGN (DC-specific intercellular adhesion molecule 3-grabbing nonintegrin), which leads to the subsequent induction of regulatory T cells (Smits et al., 2005). Moreover, an slpA mutant in L. acidophilus, which was known to have a decreased capacity to adhere to Caco-2 cells (Buck et al., 2005), also displayed reduced binding to DC-SIGN and consequently led to a more proinflammatory cytokine profile (Konstantinov et al., 2008). SlpA was postulated to be the first probiotic bacterial DC-SIGN ligand (Dsl) identified that is functionally involved
in the modulation of DC and T-cells function. Interestingly, the ConA (recognizing mannose) and AAL (specific for α6 fucose) lectins can bind purified SlpA, and because mannose and fucose are the glycans recognized by DC-SIGN, these preliminary results have led the authors to further investigate the carbohydrate moieties present on different S-layer proteins (Konstantinov et al., 2008). As already concluded above (see the section on noncovalent anchoring motifs), only a few lactobacilli produce S-layer proteins, implying that species that lack Slp proteins interact with DC-SIGN via an alternative ligand. To this end, the gene-trait matching approach enabled the identification of a candidate Dsl in L. plantarum (Molenaar et al., 2005). Dsl is predicted to be a C-terminally anchored extracellular protein, rich in threonine and serine residues. The overrepresentation of these amino acids may suggest that this protein could be O-glycosylated, which may be involved in the postulated DC-SIGN recognition (D. Remus, M. Meijerink, J.M. Wells, M.L. Marco, P.A. Bron, and M. Kleerebezem, unpublished data). Protein glycosylation mechanisms and their potential implications on the functionality of lactobacilli are discussed below.

**Lactobacillus glycoproteins?**

Although protein glycosylation was long believed to be restricted to Eukarya, it is now clear that bacteria can also form protein-attached N-glycans via asparagine residues, as well as O-glycans via threonine and/or serine residues (Messner, 2004; Szymanski & Wren, 2005; Weerapana & Imperiali, 2006; Abu-Qarn et al., 2008). Most studies on bacterial protein glycosylation revealed specific O-glycan attachment to one or two abundant polymeric surface proteins such as flagellins, pilins and S-layer proteins (Schaffer et al., 2001; Benz & Schmidt, 2002; Fletcher et al., 2007). Moreover, a limited number of general glycosylation systems have been described in bacteria, including the N-glycosylation machinery in *Campylobacter jejuni* and related species (Fig. 5b). This intestinal pathogen harbors the pgl gene cluster, which encodes the PglFED enzymes responsible for modifying the nucleotide-energized sugar UDP-GlcNaC to di-N-acetylglabillosamine (Sharon, 2007). This modified monosaccharide is attached to a phosphorylated undecaprenyl lipid carrier on the cytosolic side of the cell membrane by PglC, one of the five glycosyltransferases also encoded in the

---

**Fig. 5.** Schematic representation of polysaccharide biosynthesis systems in different bacteria. (a) Displays the typical PS synthesis system of *Lactobacillus rhamnosus* GG that is basically shared by other lactobacilli (although with different sugar compositions of the polysaccharides) and is primarily involved in the production of a cell-wall-anchored polysaccharide (or exopolysaccharide in some other lactobacilli), but may play a role in delivering repeating unit oligosaccharides to a so far unknown transferase that can transfer these oligosaccharides to extracellular proteins. (b) Displays the *Campylobacter jejuni* (and relative species) biosynthesis pathway, proposed for oligosaccharide production and transfer to Asparagine residues of extracellular proteins. (c) Displays the *Neisseria gonorrhoeae* biosynthesis pathway, proposed for oligosaccharide production and transfer to Serine (and/or Threonine) residues of extracellular proteins.
pgl gene cluster. Subsequently, a heptasaccharide is formed by the consecutive addition of nonmodified, nucleotide-energyized sugars to di-N-acetylglcosamine by the activity of the other four glycosyltransferases (PglA, I, H and J). This heptasaccharide is translocated across the membrane by the flippase PglK and covalently attached to asparagine residues by the activity of the oligosaccharyltransferase PglB, which recognizes a specific acceptor sequence motif in target proteins (Linton et al., 2003; Kelly et al., 2006). The O-glycosylation machinery in the Gram-negative human pathogen Neisseria gonorrhoeae appears to function in the same way (Stimson et al., 1995) (Fig. 5c). By contrast, in Eukarya, O-glycans are assembled via sequential coupling of nucleotide-energyized monosaccharides to proteins, and downstream glycan elaboration (Abu-Qarn et al., 2008). In N. gonorrhoeae, more than 15 proteins are potentially glycosylated by this system (Vik et al., 2009). Notably, the subcellular location predicted for these proteins is in the periplasm, and they encompass diverse functions involved in protein folding, disulfide bond formation and respiration. For the majority of the identified proteins, a glycan-bearing peptide could be identified, and all glycopeptides contained a serine residue that appears to be essential for glycosylation. Interestingly, all these Neisseria glycoproteins share domains with signatures of low complexity that include their glycosylation sites, while the N-glycosylation system described for C. jejuni uses a specific acceptor sequence motif (Vik et al., 2009).

The importance of glycan biology in host–microorganism communication in the intestine is exemplified by the observation that Bacteroides species can stimulate the expression of fucosylated glycoconjugates by the intestinal epithelium and can subsequently harvest these glycans by the production of specific degradative enzymes. Moreover, Bacteroides species have a rare bacterial pathway for the incorporation of exogenous fucose into their capsular polysaccharides and glycoproteins (Comstock & Coyne, 2003). Recently, eight glycoproteins were identified in Bacteroides fragilis, which were all predicted to be located in the periplasm or the outer membrane and involved in protein folding, protein–protein interactions, peptide degradation and surface lipoproteins (Fletcher et al., 2009). Three threonine residues of one of the glycoproteins, embedded in the conserved Asp-Thr-X motif (where X is a methyl-containing amino acid), were shown to be involved in glycoprotein formation. Notably, the Threonine can be exchanged by a Serine residue without loss of glycosylation (Fletcher et al., 2009). The same study identified the locus (lfg) involved in glycosylation in B. fragilis, which encodes five glycosyltransferases and a flippase. Consistently, mutation of the lfg locus resulted in a loss of protein glycosylation, and the mutant displayed a significant defect in in vitro growth and was not able to compete with the wild-type strain during colonization of gnotobiotic mice (Fletcher et al., 2009).

The importance of sugar moieties on the surface of lactobacilli in functional traits such as adhesion (Lebeer et al., 2009), and gastrointestinal persistence (Denou et al., 2008) and adaptation (M.L. Marco, M. Wels, W.M. de Vos, E.E. Vaughan & M. Kleerebezem, unpublished data) has been suggested in several studies. However, the polysaccharides involved have also been proposed to contribute to gastrointestinal functionality in a nonspecific manner by shielding of adhesins (Denou et al., 2008; Lebeer et al., 2009). In addition, a recent study using lectin-based glycoarrays (Hsu et al., 2006) demonstrated significant differences in sugar decoration on the surface of intact cells of L. plantarum WCFS1 harvested from either the logarithmic or the stationary phase of growth (Fig. 6). As glycan recognition by eukaryal receptors is a well-established phenomenon, it is tempting to speculate that these observed differences could (partially) explain the different in vivo responses of human intestinal mucosa towards L. plantarum cells harvested in different growth phases (van Baarlen et al., 2009). However, this and other suggestions (Konstantinov et al., 2008) that sugar moieties might be covalently attached to specific extracellular proteins of lactobacilli are at this stage based on lectin-affinity studies and are, to the best of our knowledge, not yet supported by biochemical studies that address the biosynthesis and molecular structure of the glycans, and the sugar attachment motifs in the potential target proteins. Nevertheless, many lactobacilli encode the capacity to produce several nucleotide sugars, including UDP-glucose, UDP-galactose, UDP-GlcNAc, sialic acid and dTDP-rhamnose. Moreover, the genome of L. salivarius encodes a locus (RLSL00992-995) homologous with the pglFED region of C. jejuni, suggesting that it might be able to produce the modified monosaccharide di-N-acetylglcosamino (Fig. 5b). Notably, eight of the complete Lactobacillus genomes in the ERGO database harbor at least one homologue of PglC, indicating their potential capacity to attach modified monosaccharides to a lipid carrier on the cytosolic side of the cell membrane. Moreover, capsular polysaccharide biosynthesis pathways, including several glycosyltransferases and flippases, have been identified in a variety of lactobacilli, indicating a machinery to synthesize lipid-linked oligosaccharides that mechanistically resembles the general glycosylation systems described above (Fig. 5). However, identification of the dedicated transferases, responsible for coupling of these oligosaccharides to specific amino acids in potential target proteins, is less straightforward, as these oligosaccharyltransferases might display relatively low homology and typically have rather broad substrate specificities (Faridmoayer et al., 2007). Overall, we are only beginning to appreciate the potential importance of glycans in the functionality of bacteria, and...
proteo-glycomics remains virtually unexplored in lactobacilli, even though it may play a key role in host–microorganism interactions and may eventually explain (part of) the proposed health beneficial effects conferred by these bacteria.

**Molecular analysis of host responses to lactobacilli**

Postgenomic approaches to unravel host responses to lactobacilli offer novel avenues to unravel host responses to lactobacilli at the molecular level. The use of gnotobiotic mouse models has proven its value as a reductionist model to specifically address the response of intestinal tissues to a single bacterial species. This approach was pioneered by the analysis of the transcriptome responses in gnotobiotic mice following their mono-association with the gut commensal *Bacteroides thetaiotaomicron*, which underlined the broad functional impact and developmental importance of host–microorganism communication (Hooper *et al.*, 2001). Subsequent work included co-colonization with *B. thetaiotaomicron* and *L. casei* DN-114 (or *Bifidobacterium longum* or *Bifidobacterium animalis*), showing that coccolonization impacted on the behavior of *B. thetaiotaomicron* in situ, including a shift in the glycan repertoire targeted by the vast capacity of this commensal bacterium. In addition, cocolonization with *Bacteroides* and *B. longum* elicits highly connected TNF-α and IFN-γ networks, showing that both microorganism–microorganism and host–microorganism interactions shape the intestinal ecosystem (Sonnenburg *et al.*, 2006). Host responses to microorganism in the intestine extend beyond local transcriptome responses in the mucosal tissues, and include significant influences on local and systemic mammalian biochemistry (Nicholson *et al.*, 2005). The influences of probiotic interventions on local and systemic germ-free mouse metabolism have been addressed by extensive metabolic profiling of different...
mouse tissues. These interventions, which used *L. paracasei* or *L. rhamnosus*, induce microbial population changes in the intestine and altered metabolite profiles in intestinal tissues (Martin et al., 2007). The *Lactobacillus* supplementation in these mice induces remarkably different bile acid/fecal microbiota correlation networks and significantly affects many host metabolic pathways, including lipid profiles, gluconeogenesis and amino acid and methyamine metabolism (Martin et al., 2008a, b). The unprecedented resolution and explorative strength of these holistic approaches may unravel unknown mechanisms of interactions that are ongoing in the intestine, and how they influence local and systemic physiology by affecting a multitude of molecular parameters. It is unclear whether specific extracellular functions of the lactobacilli play a key role in the measured host effects, especially because the response analyses using lactobacilli have focused on the impact at the metabolic level, which may not predominantly involve the extracellular functions of the lactobacilli. However, specific extracellular functions may be important for the metabolic impact of lactobacilli in the intestinal niche, such as the protein complexes predicted to be involved in extracellular carbohydrate recognition and breakdown (Siezen et al., 2006). Additionally, extracellular functions of the lactobacilli may (in part) be responsible for the microbiota composition changes observed in the humanized germ-free mice (Martin et al., 2007) and may have had more subtle effects on the host immune system status, which may not be prominently reflected in the metabolic state of the tissues. Although these studies present break-through achievements in the area of host–microorganism interaction at the molecular level, it is certainly not a trivial task to extrapolate these results to the human situation, and human responses to dietary lactobacilli will by definition require human intervention studies. Nevertheless, these may be strengthened through concepts or mechanistic insights derived from these mouse studies.

Few of the many clinical studies with *Lactobacillus* interventions in humans have addressed local molecular responses in the intestine. One pioneering study in this field used a duodenal perfusion system in healthy volunteers to evaluate mucosal transcriptome responses to perfusion with an *L. plantarum* WCFS1 suspension, in comparison with a placebo perfusion (Troost et al., 2008). Both short-term (1 h) and long-term (6 h) perfusions elicit differential expression patterns in the human duodenal mucosa. Short-term exposure to *L. plantarum* inhibits fatty acid metabolism and cell cycle progression in the host epithelia, while long-term perfusions launch responses associated with increased lipid metabolism, cellular growth and development (Troost et al., 2008). Proteome analysis of the biopsies taken after prolonged perfusion revealed the induction of microsomal triglyceride transfer protein, which is known to play a role in lipid transport as well as in immune modulation (Troost et al., 2008). A subsequent double-blind, placebo-controlled, randomized, cross-over study protocol was used to determine the small intestinal transcriptional responses in human volunteers after consumption of a series of small-sized, bacteria-containing ‘sports drinks.’ These studies demonstrated that functionally coherent and distinct transcriptome responses are launched by the human intestinal mucosa following the consumption of different *Lactobacillus* species (*L. acidophilus, L. casei* and *L. rhamnosus*), which appear to be correlated with the previously reported effects of these lactobacilli on host physiology and health (P. van Baarlen, F. Troost, van der C. Meer, G. Hooiveld, M. Boekschoten, & M. Kleerebezem, unpublished data). The same intervention protocol showed that remarkably distinct human mucosal transcriptome responses can be detected for *L. plantarum* cells harvested from the logarithmic or the stationary phase of growth. Specifically, stationary-phase-harvested cells can elicit extensive immune-modulation-related transcriptional networks centered on the nuclear factor κB (NF-κB), including both activating and antagonizing pathways. These are virtually unaffected by exposure to *L. plantarum* harvested from the logarithmic growth phase. The molecular response patterns surrounding NF-κB induced by stationary-phase *L. plantarum* exposure could be associated with processes such as tolerance and adjuvanticity, supporting a role of this bacterium in the modulation of the mucosal immune system (van Baarlen et al., 2009). The authors speculate on a role of cell-surface-associated functions in this growth-phase-dependent host-response modulation, based on molecular studies that indicate that cell-wall functions are significantly different in *L. plantarum* cells obtained from the logarithmic and stationary phase of growth (van Baarlen et al., 2009). These differences include the differential recognition of whole cells by specific lectins, suggesting growth-phase-dependent glycan decorations of the bacterial cell wall in this species (Fig. 6). These unique studies in healthy human volunteers exemplify the highly specific mucosal responses to dietary lactobacilli (probiotics) and underline the importance of the mode of production and delivery of these health-promoting cultures.

**Concluding remarks**

The lactobacilli form a heterogeneous group of bacteria that display considerable variation both in terms of molecular characteristics and in terms of their preferred natural habitats. This variation definitely also includes variation of bacterial cell-surface properties that are probably very important for the functioning of these bacteria, especially in relation to communication with their environment, including the communication with diet- and host-derived factors encountered in the gastrointestinal tract. Although genomics has contributed to our understanding of the
extracellular biology of the members of this genus, many of the proteins that are targeted to the extracellular SCL in lactobacilli lack a functional annotation, and in many cases, their sequence analysis does not proceed beyond the detection of specific domains, illustrating our limited understanding of the functional properties of lactobacilli that can confer to their environment. Accurate (and consistent) predictions of the exoproteomes of lactobacilli on the basis of their genome sequence, combined with domain detection and comparative analysis among strains and species, and (high-throughput) functional analyses will be essential to enhance our understanding of this subcategory of *Lactobacillus* functions.

The postgenomics era has strongly stimulated the identification of candidate effector molecules of lactobacilli that are proposed to be involved in conferring a health benefit to the host via interactions with the intestinal system, including direct interactions with host epithelial or immune cells. Notably, many of these candidate effector molecules are extracellular, including both proteinaceous factors as well as components of the cell wall itself (Kleerebezem & Vaughan, 2009). However, our understanding of the molecular mode of action of host–*Lactobacillus* interactions is still in its infancy, as is illustrated by the very small number of truly validated effector molecules identified to date. Moreover, the suggestion that lactobacilli may produce extracellular proteins that are decorated with sugar moieties through glycosylation is highly intriguing, especially with regard to the predominance of host receptors that recognize glycan moieties. In view of the complexity of the host-cell signaling and response-regulation pathways, it does not seem plausible that single *Lactobacillus* molecules drive the entire host response. These effector molecules should be seen within a ‘background’ of molecular properties that includes a potentially crucial role for the basic building blocks of the *Lactobacillus* cell wall (Grangette et al., 2005). Expansion of the candidate effector molecule repertoire and their validation *in vivo* will be essential to specify the molecular mechanisms that underlie the physiological benefits associated with the consumption of these bacteria. In addition, such knowledge would strengthen the concept of the strain specificity of probiotics and would contribute to the development of advanced procedures and criteria for product quality control and/or selection of novel probiotic strains. Identification and validation of health-benefit effector molecules could facilitate research programs that aim to obtain improved probiotic strains with enhanced health benefits, either through targeted genetic engineering or through screening or adaptive evolution under selective conditions that stimulate an increase in the relevant function.

Molecular research into host–microorganism interactions has been dominated by animal model studies that allow stringent control of potentially confounding factors such as the host genotype, age and diet. Although this type of research has yielded a wealth of novel insights and may lead to highly detailed mechanistic insights, it should be realized that the human population is characterized by a tremendous variation in these very parameters. Nevertheless, consistent molecular responses in the mucosa of healthy human volunteers can be elicited by *Lactobacillus* consumption, indicating that despite the individuality of humans, they share conserved responses to specific bacteria (van Baarlen et al., 2009). This does not exclude the possibility that the physiological consequences of these responses may be highly individual. Nevertheless, these studies illustrate that validation of (extracellular) effector molecules of lactobacilli can be performed by advanced molecular response analyses in relatively small groups of volunteers. Because there is a considerable degree of variation among different *Lactobacillus* strains as well as among human volunteers, it seems essential that effector molecule validation strategies use isogenic strains and preferably human studies that follow a cross-over protocol to increase the chance of successfully elucidating the consequences of the effector molecule mutation on the molecular host response.

**Acknowledgements**

This work was supported by the BioRange program of the Netherlands Bioinformatics Centre (NBIC), which is supported by a BSIK grant through the Netherlands Genomics Initiative (NGI). In addition, support was also provided by the National Foundation for Scientific Research (FNRS), the Foundation for Training in Industrial and Agricultural Research (FRIA) and the Research Department of the ’Communauté française de Belgique’ (Concerted Research Action). Pascal Hols is research associate of the FNRS. The authors thank Daniel Teunissen for his help in constructing Fig. 2.

**References**


Baddiley J (1989) Bacterial cell walls and membranes. Discovery

Bae T & Schneewind O (2003) The YSIRK-G/S motif of

Bellais S, Arthur M, Dubost L, Hugonnet JE, Gutmann L, van


Boekhorst J, de Been MW, Kleerebezem M & Siezen RJ (2005)

Boekhorst J, Helmer Q, Kleerebezem M & Siezen RJ (2006b)

Comparative analysis of proteins with a mucus-binding domain found exclusively in lactic acid bacteria. Microbiology 152: 273–280.


The extracellular biology of lactobacilli


Kandler O & Weiss N (1986) Regular, nonsporing Gram-positive rods. Bergey’s Manual of Systematic Bacteriology (Sneath PHA,


Sharon N (2007) Celebrating the golden anniversary of the discovery of bacillosamine, the diamino sugar of a *Bacillus*. *Glycobiology* 17: 1150–1155.


Supporting Information

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

Table S1. Inventory of Lactobacillus proteins involved in protein sorting.

Table S2. Inventory of automatically identified extracellular proteins in complete Lactobacillus genomes, using LocateP software suite (Zhou et al., 2008).

Table S3. Inventory of the distribution of LaCOGs in the Lactobacillus exoproteomes, grouped by conserved function of the LaCOGs.

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.