Research Article

Journal of Molecular Microbiology and Biotechnology

J Mol Microbiol Biotechnol 2016;26:345–358
DOI: 10.1159/000447091

Received: March 8, 2016
Accepted: May 25, 2016
Published online: July 28, 2016

Systematic Exploration of the Glycoproteome of the Beneficial Gut Isolate *Lactobacillus rhamnosus* GG

Hanne L.P. Tytgat[a, b] Geert Schoofs[a] Jos Vanderleyden[a] Els J.M. Van Damme[c] Ruddy Wattiez[d] Sarah Lebeer[a, b] Baptiste Leroy[d]

[a]Centre of Microbial and Plant Genetics, KU Leuven, Leuven, [b]Department of Bioscience Engineering, Environmental Ecology and Applied Microbiology (ENdEMIC), University of Antwerp, Antwerp, [c]Laboratory of Biochemistry and Glycobiology, Department of Molecular Biotechnology, Ghent University, Ghent, and [d]Research Institute for Biosciences, Proteomic and Microbiology, Interdisciplinary Center of Mass Spectrometry (CISMa), University of Mons, Mons, Belgium

Key Words
Glycobiology · Glycoproteins · Gram-positive bacteria · *Lactobacillus rhamnosus* GG · Probiotics · Proteomics

Abstract
Glycoproteins form an interesting class of macromolecules involved in bacterial-host interactions, but they are not yet widely explored in Gram-positive and beneficial species. Here, an integrated and widely applicable approach was followed to identify putative bacterial glycoproteins, combining proteome fractionation with 2D protein and glycostained gels and lectin blots. This approach was validated for the microbiota isolate *Lactobacillus rhamnosus* GG. The approach resulted in a list of putative glycosylated proteins receiving a 'glycosylation score'. Ultimately, we could identify 41 unique glycosylated proteins in *L. rhamnosus* GG (6 top-confidence, 10 high-confidence and 25 putative hits; classification based on glycosylation score). Most glycoproteins are associated with the cell wall and membrane. Identified glycoproteins include proteins involved in transport, translation, and sugar metabolism processes. A robust screening resulted in a comprehensive mapping of glycoproteins in *L. rhamnosus* GG. Our results reflect the glycosylation of sugar metabolism enzymes, transporters, and other proteins crucial for cell physiology. We hypothesize that protein glycosylation can confer an extra level of regulation, for example by affecting enzyme functions. This is the first systematic study of the glycoproteome of a probiotic and beneficial gut isolate.

Introduction

It is the era of the microbiome: a plethora of big data projects is mapping the bacterial load of eukaryotic species. The human intestine hosts a complex bacterial community important for human health and nutrition. Large sequencing studies have led to a better understanding of this bacterial community and have shown links between several dysbiosis states and susceptibility to diseases. Therefore, the exploration of ways to beneficially modulate this bacterial load to influence host health holds great promise for the treatment of dysbiosis-related diseases.

S. Lebeer and B. Leroy share senior authorship.

KARGER

E-Mail karger@karger.com
www.karger.com/mmb

1464–1801/16/0265–0345$39.50/0

Sarah Lebeer
Department of Bioscience Engineering, Environmental Ecology and Applied Microbiology (ENdEMIC), University of Antwerp
Groenenborgerlaan 171, BE–2020 Antwerp (Belgium)
E-Mail sarah.lebeer@uantwerpen.be

An interesting strategy is the use of intrinsic members of the microbiota as probiotics. Probiotics are 'live microorganisms which, when administered in adequate amounts, confer a health benefit on the host' [FAO/WHO, 2001]. Well-known probiotics are *Lactobacillus* sp., *Bifidobacterium* sp., beneficial strains of *Escherichia coli,* and even *Saccharomyces boulardii*, although this yeast species does not seem to be an intrinsic member of the microbiota. To allow a targeted application of the best probiotic strains for specific dysbiosis conditions, in-depth knowledge of the nature of the (beneficial) bacterial-host interactions and the key molecules modulating these interactions is crucial.

From the bacterial side, the molecules in the secreted protein (SP) fraction and on the cell wall are interesting candidates to investigate as these molecules can directly mediate contacts with host cells. In this context, glycosylated proteins are intriguing molecules. Research on mainly pathogenic glycoproteins has uncovered an enormous diversity of glycans present on bacterial proteins. These unique glycans form, together with other surface saccharides, a species-specific barcode on bacterial cell surfaces and are ideal candidates to establish specific interactions with the environment [Tytgat and Lebeer, 2014]. Glycoproteins can for instance interact with specific immune lectin receptors, as was shown for *Campylobacter jejuni* glycoproteins that interact with the macrophage galactose lectin receptor [van Sorge et al., 2009]. On the other hand, glycans on bacterial surfaces can also be used to shield immunogenic surface factors from detection by the host immune system. Moreover, the attachment of glycans to proteins can modulate the biochemical properties of proteins, enhancing their activity, specificity, and stability [Tytgat and Lebeer, 2014]. Taken together, this makes glycoproteins potentially important microbiota-host interaction factors. Indeed, in the major Gram-negative commensal *Bacteroides fragilis,* the general protein *O*-glycosylation system was shown to be essential for the competitive colonization of the mammalian intestine [Coyne et al., 2013; Fletcher et al., 2009, 2011], although specific interactions have not been mapped yet. Also, in several Gram-positive *Lactobacillus* strains, glycoproteins have been reported [Anzengruber et al., 2014; Fredriksen et al., 2012, 2013; Lebeer et al., 2012].

To facilitate more systematic analyses of bacterial glycoproteomes, we report here on an extended and widely applicable analysis of the glycoproteome of a model microbiota isolate and documented probiotic strain [Doron et al., 2005]. This paper complements our earlier work on the glycosylation potential of *Lactobacillus rhamnosus* GG. In this earlier work, we performed a systematic screening of genes encoding glycosyltransferases and predicted the presence of glycoproteins using the GlycoPP webserver, but did not yet experimentally confirm the occurrence of a genuine glycoproteome containing multiple glycoproteins in *L. rhamnosus* GG [Sanchez-Rodriguez et al., 2014].

Results and Discussion

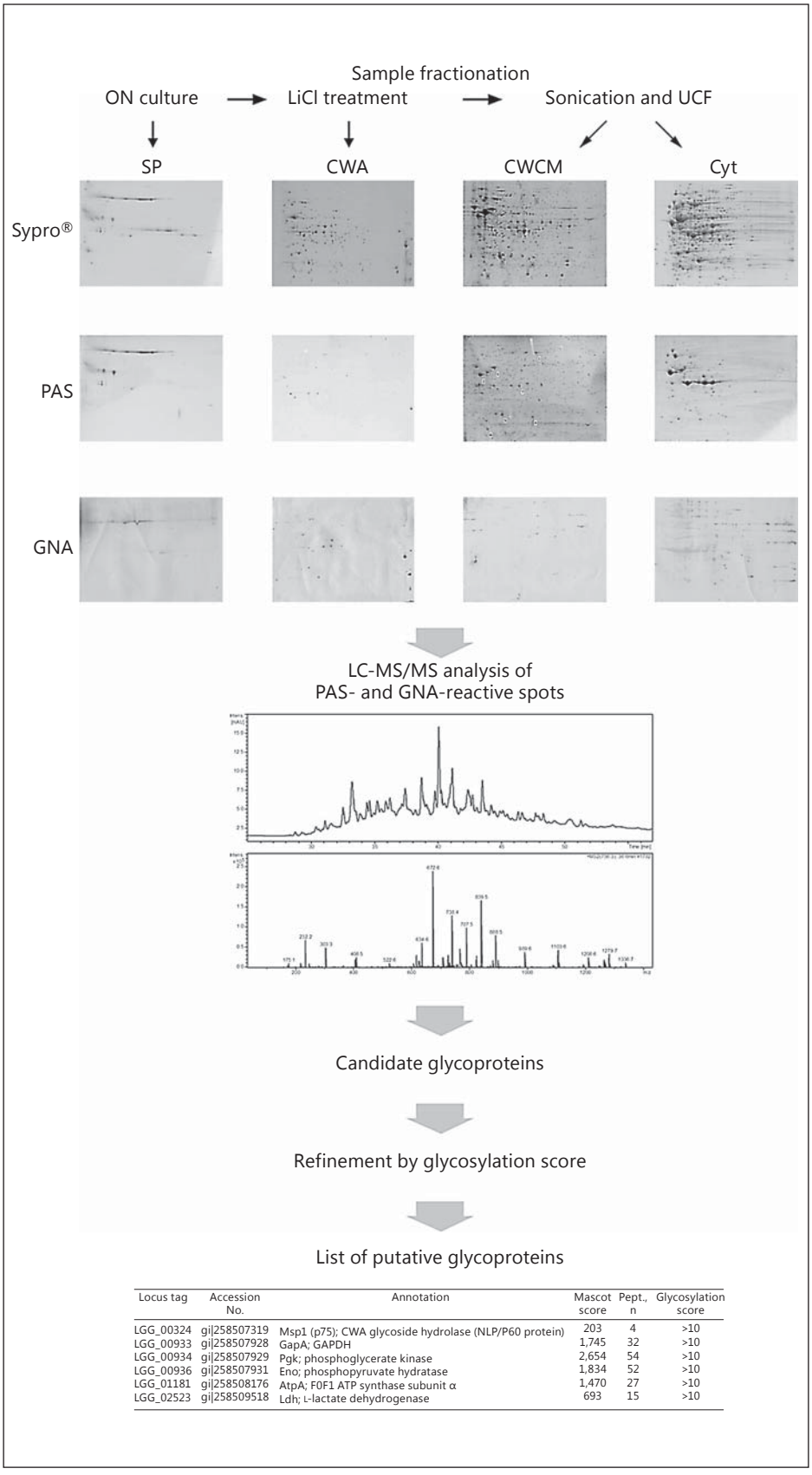
Design of the Workflow

Our workflow to map the glycoproteome of a bacterium is depicted in figure 1. In short, the fractionated proteome was screened for glycosylation using a combination of 2D periodic acid-Schiff (PAS) gels and lectin-probed 2D blots (fig. 1).

As the proteome of bacteria is quite large (e.g. >3,000 proteins in *L. rhamnosus* GG [Kankainen et al., 2009]), 2D electrophoretic images can become (over)crowded with spots. In order to end up with interpretable gels and to generate clues on the potential cellular localization of glycosylated proteins, we designed a protocol to fractionate the proteome of *L. rhamnosus* GG in four subcellular fractions: cytosol (Cyt), cell wall/membrane (CWCM), cell wall-associated (CWA), and SP fractions.

After 2D electrophoresis, the gels were stained using PAS. This glycostain oxidizes the alcohol groups of the sugars, forming aldehydes that react with the Schiff base resulting in a color reaction [Fairbanks et al., 1971]. The commercial version of this stain, the Pro-Q® Emerald 488 (Molecular Probes®), was applied, as this stain has a high sensitivity [Hart et al., 2003]. The same gels were subsequently also stained with nonspecific Sypro® Ruby protein stain to visualize all proteins (fig. 2).

The robustness of the PAS screening was increased by the design of a set of rules to select PAS-reactive spots for further identification by mass spectrometry (MS). Abundantly present proteins might, for instance, cause an overinterpretation of the PAS signal of the spot. This problem was bypassed by calculating the ratio of the relative intensity and volume percentage of the glycostained versus the protein-stained spot. These values are generated by the analysis software, and represent the intensity and distribution of this intensity relative to the area of the spot, respectively. Only spots in which the ratio of the relative intensity or volume percentage was >1, i.e. a significantly higher Pro-Q® Emerald 488 than Sypro® signal, were picked for further analysis. This list was further extended by addition of PAS-stained spots with a relative



Locus tag	Accession No.	Annotation	Mascot score	Pept., n	Glycosylation score
LGG_00324	gi|258507319	Msp1 (p75); CWA glycoside hydrolase (NLP/P60 protein)	203	4	>10
LGG_00933	gi|258507928	GapA; GAPDH	1,745	32	>10
LGG_00934	gi|258507929	Pgk; phosphoglycerate kinase	2,654	54	>10
LGG_00936	gi|258507931	Eno; phosphopyruvate hydratase	1,834	52	>10
LGG_01181	gi|258508176	AtpA; F0F1 ATP synthase subunit α	1,470	27	>10
LGG_02523	gi|258509518	Ldh; L-lactate dehydrogenase	693	15	>10

Fig. 1. Workflow overview. The presented workflow relies on the combination of proteome fractionation and 2D screening for potential glycoproteins. The proteome was fractionated in such a way that it was subdivided in four fractions. Centrifugation of an overnight culture of *L. rhamnosus* GG separated SP from the cells, which were subjected to LiCl treatment resulting in the CWA. The final sonication and ultracentrifugation (UCF) steps result in the proteins present in CWCM and the cytosol (Cyt). 2D gel electrophoresis resulted in a distinct 2D pattern for each fraction (cf. Sypro®-stained gel pictures). The 2D separated proteomes were subjected to both PAS staining and lectin probing with GNA, which specifically binds mannose residues. The combination of positive hits from both assays was identified via LC-MS/MS, which resulted in a list of candidate glycoproteins. This list was further refined by scoring all hits with a 'glycosylation scoring', which rendered a list of 41 glycosylated proteins in *L. rhamnosus* GG. The glycosylation scoring system enabled the classification of the glycosylated proteins in top-confidence hits (6), high-confidence ones (11), and putative glycosylated proteins (25). Additionally, the initial fractionation of the proteome resulted in clues on the localization of these glycoproteins. The full list of identified glycoproteins and all additional generated data can be consulted in table 1.

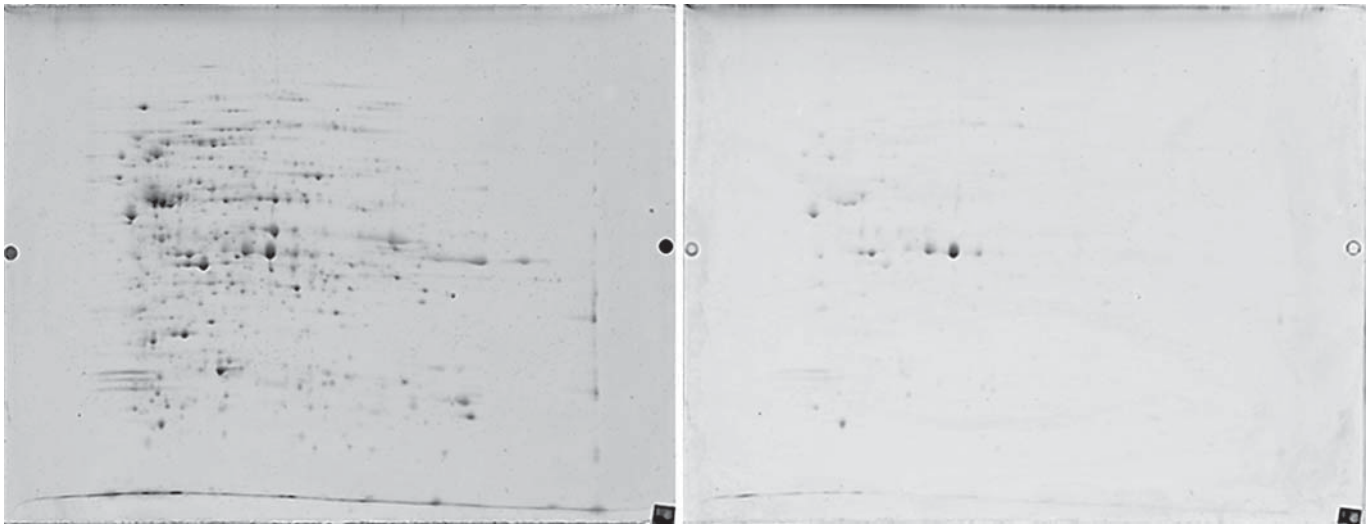


Fig. 2. Proteome and glycoproteome of the CWA fraction of the *L. rhamnosus* GG proteome. 2D electrophoresis of the CWA fraction of the proteome of *L. rhamnosus* GG is shown. On these gels, the proteins are first separated according to their isoelectric point in the first dimension, followed by molecular weight separation in the second dimension. The left panel shows the gel stained with the total protein Sypro® and thus depicts the proteome. The same gel was screened for glycoproteins using the Pro-Q® Emerald 488 staining kit (PAS staining) (right panel).

intensity >20%, to ensure that heavily glycostained spots are not missed. This last step largely corroborated the earlier selected spots, but ensured the inclusion of spots that were heavily stained with both stains.

The same subproteome fractions were also subjected to lectin probing to complement the PAS screening. Lectins are known to bind glycans in a more selective manner, i.e. recognizing specific sugars in specific configurations and linkages, while the PAS staining method detects all glycoconjugates [Van Damme et al., 2011]. The mannose-specific lectin *Galanthus nivalis* agglutinin (GNA) was applied to the Western-blotted subproteome fractions. Mannose is a typical sugar incorporated in surface glycans by *L. rhamnosus* GG, i.e. the Msp1 glycoprotein and exopolysaccharides [Lebeer et al., 2009, 2012]. The lectin GNA was chosen over concanavalin A (glucose and mannose specific), as GNA resulted in a lower background signal (results not shown). Positive spots were picked after matching them to Sypro®-stained gels. Overall, the GNA-reactive spots accorded well with the results of the PAS staining (results not shown).

In this work, we chose to validate results using lectin blotting, but further complementation of methods using gel-free techniques relying only on MS, as proposed by [Fredriksen et al., 2013], would further provide substantiation of the results.

Proteome Fractionation

Application of the workflow to *L. rhamnosus* GG cells resulted in 2D gels with distinct patterns for the four fractions, which illustrates the efficacy of the fractionation process (fig. 2, insets in fig. 1). The occurrence of the specific glycoproteins in each fraction was found to be reproducible (table 1). In table 1, an overview is given of all identified glycoproteins, and a filled circle indicates the occurrence of a glycoprotein in the same fraction in each repetition of the experiment. The actual robustness of the proteome fractionation is probably even higher, as only hits with a glycosylation score >2 were now included in the table 1.

An illustration of the robustness of the fractionation process is the fact that all 5 transport-related glycoproteins (table 2) were identified in the CWCM fractions (table 1). Another example is the three closely related glycoproteins Msp1 (LGG_00324), Msp2 (LGG_00031), and LGG_02016, being cell wall hydrolases. The function of Msp1 and Msp2 as cell wall hydrolases has previously been validated experimentally [Claes et al., 2012], in agreement with their main occurrence on the cell surface and SP fractions of the proteome.

The presence of Eno (LGG_00936), GapA (LGG_00933), Pgk (LGG_00934), and Tig (LGG_01351) in the extracellular fractions (CWCM, CWA, and SP) is remarkable. Although the main functions of these proteins

Table 1. Overview of the glycoproteins identified in *L. rhamnosus* GG and their cellular localization

Locus tag	Accession number	Annotation	Mascot score	Protein MW, Da	Pept., n	Fraction				High glyc. score	2nd glyc. score	Indep. observ.	Extra evidence
						Cyt	CWCM	CWA	SP				
Top confidence glycoproteins													
LGG_00324	gi|258507319	Msp1 (p75); cell wall-associated glycoside hydrolase (NLP/P60 protein)	203	49,766	4		○		●	>10	5	Yes	Other lactobacilli
LGG_00933	gi|258507928	GapA; glyceraldehyde-3-phosphate dehydrogenase	1,745	36,929	32	●	●	●	●	>10	>10	Yes	This work
LGG_00934	gi|258507929	Pgk; phosphoglycerate kinase	2,654	42,187	54	○	○	○		>10		Yes	
LGG_00936	gi|258507931	Eno; phosphopyruvate hydratase	1,834	47,098	52	●	●	●	○	>10	>10	Yes	
LGG_01181	gi|258508176	AtpA; F0F1 ATP synthase subunit alpha	1,470	55,253	27		●	○		>10	7	Yes	
LGG_02523	gi|258509518	Ldh; L-lactate dehydrogenase	693	35,481	15	○	●			>10	5	Yes	
High confidence glycoproteins													
LGG_00031	gi|258507026	Msp2 (p40); surface antigen	700	42,631	15		○		○	6		Yes	Other lactobacilli
LGG_00078	gi|258507073	OpuCa; glycine/betaine/L-proline ABC transporter ATP-binding protein	917	46,435	13	○	●			5	4	Yes	
LGG_01375	gi|258508370	Pyk; pyruvate kinase	2,567	62,866	56		○	○		9		Yes	
LGG_01604	gi|258508599	DnaK; molecular chaperone	1,342	67,179	25		●			6	5	Yes	*L. plantarum* WCFS1 and *F. tularensis*
LGG_02016	gi|258509011	Surface antigen NLP/P60	1,815	40,947	39			○	○	4			Other lactobacilli
LGG_02039	gi|258509034	RmlC; dTDP-4-dehydrorhamnose 3,5-epimerase	248	22,807	4		○	○		4		Yes	
LGG_02138	gi|258509133	GpmA; phosphoglyceromutase	399	25,938	5		○	○		4			
LGG_02239	gi|258509234	GroEL; chaperonin GroEL	310	57,357	6	○				4			*F. tularensis*
LGG_02493	gi|258509488	FusA; elongation factor G (EF-G)	837	76,888	15		●	○		6	5	Yes	*C. jejuni*
LGG_02806	gi|258509801	HtrA; serine protease	1,222	45,175	22	○	○	●		7	7	Yes	
Putative glycoproteins													
LGG_00255	gi|258507220	SerA; 2-hyroxyacid dehydrogenase	110	33,861	2		○			2			
LGG_00257	gi|258507252	MurE; UDP-N-acetylmuramoylanalyl-D-glutamate-2,6-diaminopimelate ligase	125	56,763	2	○				2			
LGG_00418	gi|258507413	Tal; transaldolase	594	25,533	7		○					Yes	
LGG_00634	gi|258507629	Dyp-type peroxidase family protein	355	35,060	5		○			2		Yes	
LGG_00951	gi|258507946	MalE; sugar ABC transporter substrate-biding protein	827	48,326	11		○			2		Yes	
LGG_01351	gi|258508346	Tig; trigger factor	874	49,751	19			○		2			
LGG_01374	gi|258508369	PfkA; 6-phosphofructokinase	297	34,176	6		○			2		Yes	
LGG_01474	gi|258508469	NrdE; ribonucleotide-diphosphate reductase subunit alpha	143	82,395	2	○		○		2	2	Yes	
LGG_01615	gi|258508610	NusA; transcription elongation factor	782	44,546	16		○	○		2		Yes	
LGG_01628	gi|258508623	RpsB; 30S ribosomal protein S2	414	29,623	5	○				2			
LGG_01670	gi|258508665	Fmt; methionyl-tRNA formyltransferase	110	34,226	2	○				2			
LGG_01717	gi|258508712	PheT; phenylalanyl-tRNA synthetase subunit beta	342	88,016	11		○			2			
LGG_01900	gi|258508895	Hypothetical protein	583	20,900	9	○				2		Yes	
LGG_01997	gi|258508992	RmlB; dTDP-glucose 4,6-dehydratase	124	38,560	3	○				2		Yes	
LGG_02009	gi|258509004	GlnQ; amino acid ABC transporter ATP-binding protein	307	27,436	6		○			2		Yes	
LGG_02234	gi|258509229	MutL; DNA mismatch repair protein	398	71,723	9		○			2		Yes	
LGG_02277	gi|258509272	RplJ; 50S ribosomal protein L10	383	18,237	5		○			2		Yes	
LGG_02332	gi|258509327	GltX; glutamyl-tRNA synthetase	245	56,901	7			○		2			
LGG_02419	gi|258509414	MtsA; manganase ABC transporter substrate-binding protein	1,172	34,996	28		○	○		2			
LGG_02424	gi|258509419	YmjC; NAD-dependent epimerase/dehydratase	81	23,701	2	○				2			

J Mol Microbiol Biotechnol 2016;26:345–358
DOI: 10.1159/000447091

Table 1 (continued)

Locus tag	Accession number	Annotation	Mascot score	Protein MW, Da	Pept., n	Fraction				High glyc. score	2nd glyc. score	Indep. observ.	Extra evidence
						Cyt	CWCM	CWA	SP				
LGG_02472	gi|258509467	RplF; 50S ribosomal protein L6	477	19,304	8			○		2			
LGG_02475	gi|258509470	RplE; 50S ribosomal protein L5	146	20,198	2	○				2			
LGG_02498	gi|258509493	RpoB; DNA-directed RNA polymerase subunit beta	209	133,938	6		○			2			
LGG_02575	gi|258509570	LacD; tagatose-1,6-diphosphate aldolase	1,887	36,334	37		○	○		2		Yes	
LGG_02643	gi|258509638	RecA; recombinase A	686	42,775	13		○			2			*H. pylori*

For each glycoprotein the data for the hit with the highest confidence is shown out of the two independent repetitions of the experiment. Top confidence glycoproteins are attributed a glycosylation score >10, high confidence hits have scores ranging between 4 and 9. Lower glycosylation scores were put under the 'putative glycoprotein' denominator (>2). Locus tag: gene identifier of the glycoprotein. Accession number: gi identifier of the glycoprotein. Annotation: annotation as in current genome release of GenBank (NC_013198.1). Mascot score: Highest mascot score obtained for the protein. Peptides, n: Highest number of peptides identified. Fraction: A filled circle indicates the presence of the glycoprotein in both experiments in this fraction. An empty circle represents the identification of a glycoprotein in the fraction in one repetition of the experiment. High glyc. score: the highest glycan score that could be attributed to a glycoprotein in one of the repetitions of the experiment. 2nd glyc. score: the glycosylation score of the glycoprotein in the repetition of the experiment. Indep. observ.: Yes was indicated in this column if the glycoprotein was observed at least once in a reactive spot, regardless of the concomitant presence of other glycoproteins in this spot and/or when the protein was identified in an independent GNA screening. Extra evidence: extra evidence supporting our results, e.g. hits and trends published in other species.

reside in the cytosol, their surface localization was already reported earlier [Deepika et al., 2012; Saad et al., 2009]. This 'aberrant' localization might point towards a moonlighting role for these proteins [Deepika et al., 2012; Izquierdo et al., 2009; Jeffery, 2003; Sanchez et al., 2009a], which has already been documented in other species for the glycolytic enzymes Eno, GapA, and Pgk [Castaldo et al., 2009; Granato et al., 1999; Henderson, 2014; Kinoshita et al., 2008a, b; Sanchez et al., 2009a].

In total, we identified 15 putative cytosolic glycoproteins, 28 glycoproteins in the CWCM fraction, 17 CWA proteins, and 5 SP (table 1). Nevertheless, although we illustrated that these fractions are quite robust and confirm earlier protein localizations, we are aware that some leakiness of fractions cannot be ruled out. For instance, the occurrence of glycosylated proteins in the cytosol is generally regarded to be unlikely. However, implementation of controls for the different fractions is not straightforward since various proteins that were once thought to be typical for the cytosol were discovered in other fractions and appear to be moonlighting proteins (as we also observed) [Deepika et al., 2012; Henderson and Martin, 2014; Saad et al., 2009].

Glycoprotein Identification

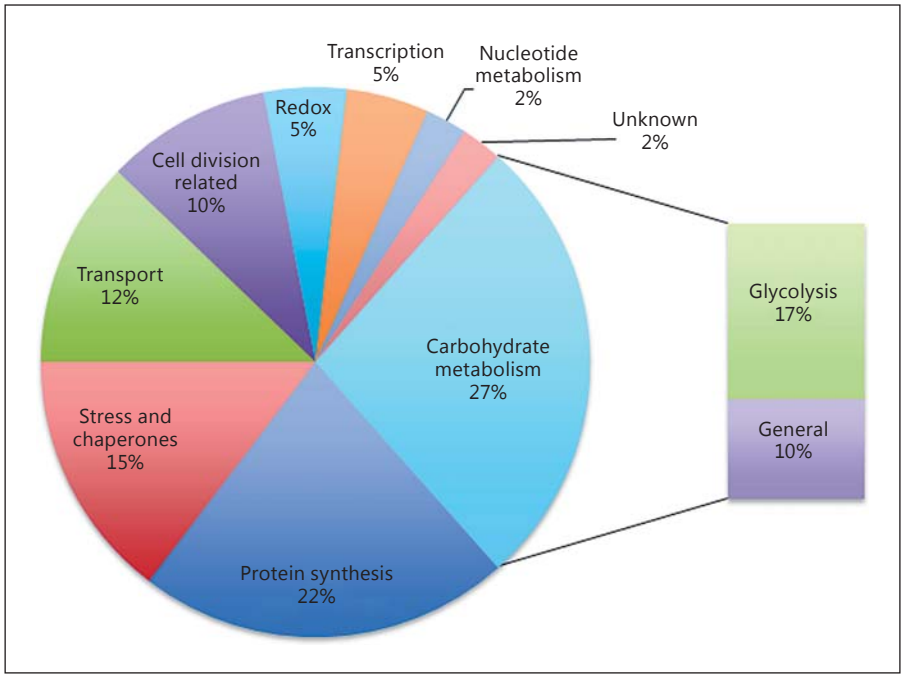
To circumvent the fact that most picked spots contain multiple proteins, a 'glycosylation scoring' system was implemented to identify genuinely glycosylated targets. Only proteins identified based on the presence of at least two peptides were taken into account in our dataset. The 'glycosylation score' was conceived as a cumulative index: the confidence of the glycosylation status of a protein increases when (a) the candidate glycoprotein is detected in both repetitions of the experiment (3 points), (b) a protein is identified alone in a PAS-/GNA-reactive spot (2 points), and (c) a protein is a top hit (highest mascot score) in a reactive spot in which no other glycoproteins are present (1 point). Based on this score, the putative glycoproteins were divided into three categories: top-confidence hits (>10), high-confidence glycoproteins (>4), and putative glycosylated proteins (>2). Proteins with a lower score were discarded from the data set.

In total, our workflow (fig. 1) resulted in the identification of one or more proteins in more than 230 PAS-/lectin-reactive spots through liquid chromatography (LC) MS/MS. The glycosylation score allowed filtering this data set and selecting 41 glycosylated proteins in *L. rhamnosus* GG, which are presented in table 1. Among the 6 glycoproteins that were identified with top confidence, the earlier identified and experimentally validated Msp1

Table 2. General trends in the glycoproteome of *L. rhamnosus* GG

Locus tag	Accession No.	Annotation	Confidence
Carbohydrate: glycolysis (-related enzymes)			
LGG_00933	gi|258507928	GapA; glyceraldehyde-3-phosphate dehydrogenase	Top
LGG_00934	gi|258507929	Pgk; phosphoglycerate kinase	Top
LGG_00936	gi|258507931	Eno; phosphopyruvate hydratase	Top
LGG_01374	gi|258508369	PfkA; 6-phosphofructokinase	Putative
LGG_01375	gi|258508370	Pyk; pyruvate kinase	High
LGG_02138	gi|258509133	GpmA; phosphoglyceromutase	High
LGG_02523	gi|258509518	Ldh; L-lactate dehydrogenase	Top
Carbohydrate metabolism: general			
LGG_00418	gi|258507413	Tal; transaldolase	Putative
LGG_01997	gi|258508992	RmlB; dTDP-glucose 4,6-dehydratase	Putative
LGG_02039	gi|258509034	RmlC; dTDP-4-dehydrorhamnose 3,5-epimerase	High
LGG_02575	gi|258509570	LacD; tagatose-1,6-diphosphate aldolase	Putative
Protein synthesis: amino acid metabolism and translation			
LGG_00255	gi|258507220	SerA; 2-hydroxyacid dehydrogenase	Putative
LGG_01628	gi|258508623	RpsB; 30S ribosomal protein S2	Putative
LGG_01670	gi|258508665	Fmt; methionyl-tRNA formyltransferase	Putative
LGG_01717	gi|258508712	PheT; phenylalanyl-tRNA synthetase subunit β	Putative
LGG_02277	gi|258509272	RplJ; 50S ribosomal protein L10	Putative
LGG_02332	gi|258509327	GltX; glutamyl-tRNA synthetase	Putative
LGG_02493	gi|258509488	FusA; elongation factor G (EF-G)	High
LGG_02472	gi|258509467	RplF; 50S ribosomal protein L6	Putative
LGG_02475	gi|258509470	RplE; 50S ribosomal protein L5	Putative
Stress and chaperones			
LGG_01351	gi|258508346	Tig; trigger factor	Putative
LGG_01604	gi|258508599	DnaK; molecular chaperone	High
LGG_02234	gi|258509229	MutL; DNA mismatch repair protein	Putative
LGG_02239	gi|258509234	GroEL; chaperonin GroEL	High
LGG_02643	gi|258509638	RecA; recombinase A	Putative
LGG_02806	gi|258509801	HtrA; serine protease	High
Transport			
LGG_00078	gi|258507073	OpuCa; glycine/betaine/L-proline ABC transporter ATP-binding protein	High
LGG_00951	gi|258507946	MalE; sugar ABC transporter substrate-binding protein	Putative
LGG_01181	gi|258508176	AtpA; F0F1 ATP synthase subunit α	Top
LGG_02009	gi|258509004	GlnQ; amino acid ABC transporter ATP-binding protein	Putative
LGG_02419	gi|258509414	MtsA; manganase ABC transporter substrate-binding protein	Putative
Cell division-related proteins			
LGG_00031	gi|258507026	Msp2 (p40); surface antigen	High
LGG_00257	gi|258507252	MurE; UDP-N-acetylmuramoyl alanyl-D-glutamate–2,6-diaminopimelate ligase	Putative
LGG_00324	gi|258507319	Msp1 (p75); cell wall-associated glycoside hydrolase (NLP/P60)	Top
LGG_02016	gi|258509011	Surface antigen NLP/P60	High
Redox			
LGG_00634	gi|258507629	Dyp-type peroxidase family protein	Putative
LGG_02424	gi|258509419	YmjC; NAD-dependent epimerase/dehydratase	Putative
Transcription			
LGG_01615	gi|258508610	NusA; transcription elongation factor	Putative
LGG_02498	gi|258509493	RpoB; DNA-directed RNA polymerase subunit β	Putative
Nucleotide metabolism			
LGG_01474	gi|258508469	NrdE; ribonucleotide-diphosphate reductase subunit α	Putative
Other			
LGG_01900	gi|258508895	Hypothetical protein	Putative

Locus tag = Gene identifier of the glycoprotein; Accession No. = gi identifier of the glycoprotein; Annotation = annotation as in current genome release of GenBank (NC_013198.1); Confidence = level of certainty with which the glycoprotein was identified (cf. table 1).



Color version available online

Fig. 3. General trends among glycoproteins identified in *L. rhamnosus* GG. This pie chart depicts the relative abundance of glycoproteins of each general trend that can be delineated in the glycoproteome of *L. rhamnosus* GG. The largest class of glycoproteins plays a role in the carbohydrate metabolism: a striking 17% are enzymes of glycolysis, the remaining 10% are proteins that can be linked to the general sugar metabolism.

protein was found [Lebeer et al., 2012]. Moreover, 24 of the identified glycoproteins could be validated in independent experiments, i.e. confirmation of the glycosylation status of proteins in independently performed lectin blots and PAS screenings (results not shown, cf. 'Independent observation' column in table 1) or by other MS-based approaches.

Of note, it is important to realize that our approach, i.e. the implementation of strict rules at several steps of the screening process, most probably neglects some glycoproteins, rendering some false negatives.

General Trends in Protein Glycosylation

When looking at the list of identified glycoproteins, some major trends can be delineated. For instance, we already mentioned the high number of glycolytic enzymes that were found to be glycosylated (table 1). The list of glycoproteins was mined for glycoproteins involved in similar processes (cf. KEGG database), which were grouped in functional classes (fig. 3; table 2).

The identified glycoproteins were organized in eight classes reflecting their functional role in *L. rhamnosus* GG. The largest set of glycoproteins (27%) comprises carbohydrate metabolism-related proteins (fig. 3). The second largest subset of glycoproteins is involved in protein synthesis (22%), particularly in amino acid metabolism and translation processes. Other apparent trends are the glycosylation transport- (12%) and stress-related (15%) proteins. Smaller classes contain cell division-related glycosylated proteins (10%), redox enzymes (5%), members of the transcription machinery (5%), and an enzyme involved in nucleotide metabolism.

Carbohydrate Metabolism

A striking 27% of the identified glycoproteins (11/41) can be linked to the metabolism of carbohydrates. Within this group, the largest portion of hits originates from glycosylated glycolytic enzymes (17%). An additional set of glycosylated targets (10%) are involved in the general carbohydrate metabolism. This high number of glycosylated carbohydrate metabolizing enzymes might point towards a feedback loop in glycoprotein production and general sugar metabolism. Intuitively, the presence of such a level of 'self' regulation would make sense. Whether this hypothesis is indeed correct and if this trend can also be validated in other species remains to be investigated.

For some glycolytic enzymes (GapA, Eno, and Pgk), an extra moonlighting function has been reported previously. Next to their glycolytic activity, these enzymes also play a role in binding to the extracellular environment [Castaldo et al., 2009; Granato et al., 1999; Henderson, 2014; Kinoshita et al., 2008a, b; Sanchez et al., 2009a]. These proteins have potentially also a moonlighting role

in *L. rhamnosus* GG as many observations were made in closely related lactobacilli [Castaldo et al., 2009; Granato et al., 1999; Kinoshita et al., 2008a, b]. This hypothesis is substantiated by the occurrence of these glycoproteins in the extracellular fractions of the proteome of *L. rhamnosus* GG in this (table 1) and previous studies [Deepika et al., 2012; Saad et al., 2009; Sanchez et al., 2009a, b].

Protein Synthesis: Amino Acid Metabolism and Translation

Proteins involved in protein metabolism, or more precisely in amino acid metabolism and translation processes, were largely represented among the glycosylated proteins identified (22%; fig. 3). We identified 4 glycosylated ribosomal proteins, as well as enzymes involved in the regulation of methionine, phenylalanine, and glutamate metabolism (putative glycoproteins). The last protein in this class is the elongation factor G or FusA protein. Glycosylation of this elongation factor was already reported earlier in the model organism for bacterial glycosylation *C. jejuni* [Young et al., 2002] and in *Helicobacter pylori* [Champasa et al., 2013]. It is intriguing that members of the largest bacterial multienzyme complex, the ribosome, are glycosylated, since glycosylation of members of protein complexes has previously been observed or suggested for several other complexes [Tytgat and Lebeer, 2014].

Stress-Related Proteins and Chaperones

Glycosylation of chaperones was already reported in several bacteria: *Lactobacillus plantarum* WCFS1 [Fredriksen et al., 2013], *B. fragilis* [Fletcher et al., 2009], *H. pylori* [Champasa et al., 2013], and *Francisella tularensis* [Balonova et al., 2010]. Our analysis resulted in the identification of 6 glycosylated proteins that have a chaperone or stress response-related activity (table 2). Glycosylation of the DnaK chaperone confirms the earlier reports in *L. plantarum* WCFS1 [Fredriksen et al., 2013], *H. pylori* [Champasa et al., 2013], and *F. tularensis* [Balonova et al., 2010]. Another chaperonin, GroEL, was also found to be glycosylated in *F. tularensis* [Balonova et al., 2010] and *H. pylori* [Champasa et al., 2013]. Taken together, this might point towards a specific role or importance of the glycosylation of chaperones. It is conceivable that glycosylation of these proteins enhances their stability. Nonetheless, further confirmation of this emerging trend is needed with functional data in different bacteria.

RecA, encoding a recombinase important for DNA repair processes, was also found to be glycosylated in *L. rhamnosus* GG. Our result corroborates a report from 2004 on the glycosylation of the RecA recombinase of *H. pylori*, which was one of the first reports on the occurrence of intracellular glycoproteins [Fischer and Haas, 2004]. We nevertheless want to stress that leakiness of the fractions is a factor that needs to be taken into account and further validation is required.

Transporters

Of the retrieved glycoproteins, 12% (5/41) are transporters or transport-related proteins (fig. 3; table 2). Glycosylation of ABC transporters and their related proteins seems thus to be a general trend among our results. Two of the 5 glycosylated transporters are moreover linked to the transport of amino acids. This is in agreement with the reports on the glycosylation of membrane-associated transporters in *C. jejuni* [Scott et al., 2011a] and ABC transporters in *Neisseria gonorrhoea* [Vik et al., 2009] and *H. pylori* [Champasa et al., 2013]. Moreover, some of the glycosylated transporters, such as the CjaA transporter of *C. jejuni* (CJ0982c) [Scott et al., 2011b] reported in these species are also linked to the amino acid metabolism of the cells.

Cell Division-Related Proteins

A common trend in lactobacilli, which are known to glycosylate proteins, is the glycosylation of cell wall hydrolases like the Msp1 glycoprotein of *L. rhamnosus* GG [Lebeer et al., 2012] and related autolysins and cell wall hydrolases in *L. plantarum* WCFS1 [Fredriksen et al., 2012; Rolain et al., 2013], *Lactococcus lactis* [Huard et al., 2003] and *Lactococcus buchneri* [Anzengruber et al., 2014]. In this work, the glycosylation of the cell wall hydrolase and probiotic effector Msp1 was further independently confirmed. Furthermore, we report on the glycosylation of the closely related probiotic effector Msp2 surface antigen (p40) and LGG_02016, both identified with high confidence. Our group proved earlier that the rather extensive glycosylation of Msp1 improves its stability, but might hamper its activity compared to the closely related Msp2 protein [Lebeer et al., 2012; Yan et al., 2007].

Redox-Related Proteins and Other Classes

Glycosylation of redox-related proteins has been shown already for a number of species, including *C. jejuni* [Ding et al., 2009; Scott et al., 2009; Young et al., 2002], *N. gonorrhoea* [Vik et al., 2009], *H. pylori* [Champasa et al., 2013], and *F. tularensis* [Balonova et al., 2010]. Here we report on the glycosylation of 2 redox-related proteins in *L. rhamnosus* GG (table 2).

Smaller classes of glycosylated proteins in *L. rhamnosus* GG include transcription- (2 glycoproteins) and nu-

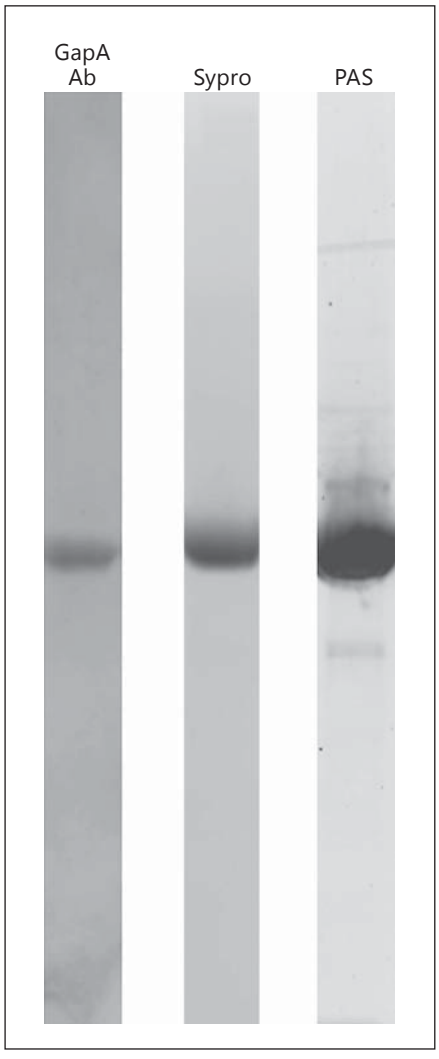


Fig. 4. The LGG_00933 protein, GapA, is glycosylated. GapA was isolated from all four fractions of the *L. rhamnosus* GG proteome (secreted fraction depicted). The first panel shows the positive reaction of the protein with the GapA antibody (Ab). The protein-stained (Sypro® stain) gel fragment shows the abundant presence of the enriched protein on 1D SDS-PAGE after successful purification of the protein from the proteome. Glycosylation of the GapA protein can be inferred from the strong positive signal of the protein with the Pro-Q® Emerald 488 dye, a commercial PAS stain.

cleotide metabolism-related proteins (1). One protein could not be attributed to a specific class (LGG_01900), as the role of this protein is currently unknown.

Confirmation of GAPDH Glycosylation Underlines the Strength of the Workflow

The *LGG_00933*-encoded GAPDH (glyceraldehyde 3-phosphate dehydrogenase), GapA, is a well-conserved glycolytic enzyme, catalyzing the conversion of glyceraldehyde 3-phosphate to D-glycerate 1,3-phosphate. Apart from this canonical function, several moonlighting roles have been attributed to this enzyme, especially in eukaryotes [Henderson, 2014; Henderson and Martin, 2014]. In Prokarya, and more in particular in two *L. plantarum* strains, this protein has been implicated in binding of the bacterium to mucin, human blood antigens, and the intestinal epithelial cell line Caco-2 [Kinoshita et al., 2008a, b; Ramiah et al., 2008]. Moreover, in *Streptococcus pyogenes,* it was shown that a specific cell surface form of GAPDH mediates its bacterial adherence and antiphagocytic moonlighting functions [Boel et al., 2005]. Also in *L. plantarum,* important biochemical differences were shown for the two 'geographical' forms of GAPDH, whereby the soluble GAPDH does not bind to the surface of the intact organism [Saad et al., 2009]. These findings might be valid in *L. rhamnosus* GG, too, as *L. plantarum* bacteria are close relatives, and earlier papers also report on the presence of GapA in the extracellular fractions of the *L. rhamnosus* GG proteome (surface and secretome) [Deepika et al., 2012; Sanchez et al., 2009a, 2009b]. Nevertheless, further research is necessary to validate this hypothesis.

Our results reflect the ubiquitous presence of GapA, as this protein was identified in all four proteome fractions for both repetitions of the experiment (table 1). These findings also corroborate earlier reports on the extracellular presence of GapA [Deepika et al., 2012; Saad et al., 2009; Sanchez et al., 2009a, b], which further substantiates the moonlighting data [Henderson, 2014; Kinoshita et al., 2008a, b; Ramiah et al., 2008].

The GapA enzyme was subsequently purified from all four proteome fractions to confirm its glycosylation status as a validation of the used approach. Gel filtration chromatography resulted in GAPDH-enriched samples for each of the four different fractions (fig. 4; Sypro stain and GapA antibody data of secreted GAPDH). The protein reacts strongly with the PAS stain. GapA purified from all fractions exhibited the same trend (results not shown). This is, to our knowledge, the first report on the glycosylation of this enzyme.

Conclusions

In this work, the workflow is presented to map the glycoproteome of a bacterium in a way that provides information on the potential cellular localization of the glycosylated targets. We were able to identify 41 glycoproteins,

which are important for the cellular physiology. Most glycoproteins were recovered from the CWCM fraction. Our research generated new insights into the role glycosylation can play in enzyme regulation. This analysis provided a list of targets that might play an important role in the interactions between the well-documented microbiota isolate and the model probiotic strain *L. rhamnosus* GG on the one side and the human gut on the other side, since most glycoproteins were isolated from the exoproteome. Our results strengthen the notion that further research is needed to explore the interaction between glycoproteins on the bacterial cell surface and the environment in order to fully grasp the finesses of host-microbial interactions.

Experimental Procedures

Bacterial Strains and Culture Conditions

L. rhamnosus GG was grown at 37°C in lactobacillus AOAC medium (BD Difco™, Erembodegem Belgium) in nonshaking conditions.

Glycoproteome Isolation and Fractionation

L. rhamnosus GG was grown for 24 h in AOAC medium. The supernatant of the culture after centrifugation at 6,000 *g* for 20 min resulted in SP (i.e. fraction 1). This fraction was filtered to remove any remaining bacteria (0.45 μm) followed by precipitation with trichloroacetic acid (TCA, 20% final concentration). The precipitated proteins were washed twice with cold acetone (100%). The protein pellet was air dried and resuspended in lysis buffer (2 M thiourea, 7 M urea, 4% CHAPS, 2% DTT). The pellet of the overnight culture, containing *L. rhamnosus* GG, was washed twice with PBS. The cells were then incubated in a solution of 1.5 M LiCl in 10 mM Tris-HCl (pH 8) for 1 h at 4°C. Centrifugation (6,000 *g*, 15 min) resulted in a second fraction, i.e. the CWA fraction, which was precipitated with TCA [Bauerl et al., 2010]. The remaining pellet was dissolved in PBS and sonicated at an amplitude of 20% (2 × 2 min) to ensure cell lysis (Branson sonifier; Emerson, Danbury, Conn., USA). The lysate was cleared by centrifugation (several runs at 4,000 *g* to remove cell debris). Cytosolic proteins (fraction 3, Cyt) could be found in the supernatant after centrifugation (22,000 *g*, 20 min). The pellet was washed three times with 50 mM cold Tris-HCl, pH 8, and 500 mM NaCl. The resulting pellet was dissolved in PBS and contained the CWCM proteins (4th fraction). Similarly to the other fractions, both the Cyt and CWCM fractions were precipitated and dissolved in lysis buffer.

2D Gel Electrophoresis

2D gel electrophoresis is a method to separate proteins in two dimensions according to their intrinsic isoelectric point (i.e. isoelectric focusing) and their molecular weight (SDS-PAGE). 2D gel electrophoresis was performed as described earlier by Sonck et al. [2009]. Briefly, 24-cm Immobiline DryStrips (GE Healthcare, Diegem, Belgium) with a pH range from 3 to 11 were rehydrated overnight in reswelling buffer (6 M urea, 2 M thiourea, 2% CHAPS, 0.4% DTT, 0.002% bromophenol blue) containing 0.5% (pH 3–11) IPG buffer (GE Healthcare). Prior to anodic cup sample loading on the DryStrips, 0.8% (pH 3–11) IPG buffer (GE Healthcare) was added to the samples dissolved in lysis buffer (cf. supra). Isoelectric focusing of the samples on the Immobiline DryStrips was performed on a Multiphor II instrument (GE Healthcare). Prior to SDS-PAGE, the focused proteins were reduced using equilibration buffer (6 M urea, 30% glycerol, 2% SDS, 0.002% bromophenol blue in 1.5 M Tris-HCl, pH 8.8, 1% DTT). The same procedure was repeated with equilibration buffer containing 4.5% IAA. Second dimension separation (SDS-PAGE) was performed by application of the DryStrips with the isofocused samples on fresh 15% polyacrylamide gels (Amresco®, Solon, Ohio, USA), using the Ettan DALT*six* instrument (GE Healthcare).

Detection of Glycosylated Proteins

Glycosylated proteins were detected by PAS staining [Fairbanks et al., 1971]. The 2D gels were fixed and stained with the Pro-Q® Emerald 488 glycoprotein gel stain kit (Molecular Probes®, Paisley, UK), a fluorescent and sensitive PAS stain (detection limit of 4 ng of glycoprotein) [Hart et al., 2003]. After analysis of the PAS-stained gels, the same gels were submerged overnight in Sypro® Ruby protein gel stain (Molecular Probes®). Scanning of the stained gels was performed on a Typhoon 9400 laser scanner (GE Healthcare).

Image Analysis and Spot Picking

Gel images were analyzed using ImageMaster software (GE Healthcare). The positive hits with PAS stain were matched to the Sypro®-stained spots. The relative intensity and volume percentage on glyco- and protein-stained gels were compared for each spot. More precisely, the intensity of the spots was recalculated in comparison to the spot with the highest intensity, both for the PAS- and Sypro®-stained gels. The same was done for the volume percentage. Then the ratio was calculated of the relative intensity values (value Pro-Q® Emerald 488/value Sypro®) and of the relative volume percentages. Spots that had a ratio >1 were selected for further identification. In order not to miss any highly glycosylated spots, all hits with a relative intensity >20% on the PAS-stained gels were also picked for further identification. Spot picking was executed automatically with the Ettan SpotPicker (GE Healthcare).

2D lectin blot-2D gels of the four fractions of the proteome of *L. rhamnosus* GG were transferred to PVDF membranes (Millipore) using a TE77 Semi-Dry transfer unit (GE Healthcare). These Western blots were probed with digoxigenin-labeled GNA, which specifically binds mannose residues [Tytgat et al., 2015; Van Damme and Peumans, 1988]. To reduce background, polyvinylalcohol was used as a blocking agent [Thompson et al., 2011]. Spots reacting with the lectin were matched to the Sypro®-stained gels and picked for identification.

In-Gel Tryptic Digestion and MS

Reactive spots were submitted to in-gel trypsin digestion. Spots were first washed twice in 50 mM NH_4HCO_3 for 15 min followed by dehydration during 25 min in 50% (v/v) acetonitrile in 50 mM NH_4HCO_3. 10 μl of 50 mM NH_4HCO_3 containing 0.05 μg of trypsin were then added to the gel pieces and incubated on ice for 20 min. Trypsinolysis was carried on overnight at 37°C and stopped by addition of 2 μl of 1% (v/v) formic acid. Tryptic peptides were analyzed using an LC-MS/MS workflow as already described [Mastroleo et al., 2009]. Briefly, peptides were separated by reverse phase chromatography using a 40-min ACN gradient

(4–45%) and analyzed on an HCTultra Ion Trap (Bruker, Goes, The Netherlands) in MS/MS mode. MS spectra were acquired across the 300–1,500 m/z mass range, and the four ions with the highest intensity were selected for MS/MS and excluded after one spectrum for 30 s. Acquired data were analyzed with the Mascot search engine against the NCBI database (NCBInr_20130409) restricted to *L. rhamnosus* GG.

Data Analysis and Glycosylation Score

Proteins were first filtered through the identification p value: only proteins with $p < 0.05$ were selected. Peptides with a Mascot ion score <30 were also discarded. Finally, only proteins identified with at least 2 unique peptides were further analyzed. A glycosylation score was attributed to each identified protein as follows: 1 point was given to a protein identified as a top hit in a spot with no other putative glycoproteins present (a putative glycoprotein is a protein with a glycosylation score >2, top hit protein is defined as the protein with the highest Mascot score for a spot); 2 points were given each time a protein was identified as the sole protein in a reactive spot; 3 points were attributed if the protein has been detected as a putative glycoprotein in the repetition of the experiment. These weights were assigned based on the importance of the experimental evidence leading to the identification of the hits as being true glycoproteins. Retrieval of proteins in independent repetitions of the experiment and as sole protein in a reactive spot was more highly valued in the assessment of results compared to it being merely a top hit amongst other proteins present in a 2D gel spot.

Mining of the Results for General Trends

All 41 putative glycoproteins were checked for their functional role in the KEGG database (genome.jp/kegg/), which enabled the delineation of general functional classes among the results. The classes were described in accordance to GO terms [Ashburner et al., 2000].

Purification of GAPDH (LGG_00933)

Protein fractions of *L. rhamnosus* GG were prepared as described above and GAPDH was purified as described by Saad et al. [2009]. All fractions, except for SP, were dissolved in 25 mM NH_4HCO_3, pH 7.5, 5 mM EDTA, 1 mM PMSF (buffer A) after TCA precipitation. The SP fraction was concentrated and resuspended in buffer A using a tangential flow filtration system (KrosFlo; Spectrum Labs, Breda, The Netherlands). All samples were dialyzed overnight (4°C) against buffer A. The dialyzed fractions were applied to a 5-ml fast-flow blue Sepharose CL 6B column (GE Healthcare) pre-equilibrated with buffer A at a flow rate of 2 ml/min. Elution was performed using a linear gradient of 0–10 mM nicotine amide adenine dinucleotide in buffer A, pH 8.5 (buffer B) [Saad et al., 2009].

Assessment of the Glycosylation State of GAPDH (LGG_00933)

Glycosylation of GAPDH in the different proteome fractions of *L. rhamnosus* GG was tested using the Pro-Q® Emerald 488 glycoprotein gel stain kit (Molecular Probes®) on 1D gels, as described earlier [Lebeer et al., 2012]. Gels were imaged using a Typhoon 9400 laser scanner (GE Healthcare). Gels were poststained with Sypro® to visualize the proteins. Confirmation that the bands corresponded to GAPDH was obtained by Western blots developed with GAPDH yeast antibody (Nordic Immunology, Susteren, The Netherlands). The high level of conservation among GAPDH proteins justified usage of this antibody in general, and binding of this antibody to purified GAPDH of *L. plantarum* CMPG5300 ([Malik et al., 2014; Malik, pers. commun.] further confirmed this.

Acknowledgments

H.L.P.T. acknowledges the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT Vlaanderen) for her PhD scholarship. S.L. previously held a postdoc grant from Fonds Wetenschappelijk Onderzoek-Vlaanderen (FWO) to study protein glycosylation in lactobacilli and now holds a 'Krediet aan Navorsers' (28960). J.V., S.L., H.L.P.T., and G.S. are also grateful for the financial support of the University of Leuven via PF3M100234 and UAntwerp via BOF and IOF-SBO funding. H.L.P.T. would like to thank Pieter Vanbosseghem for his help with the figures.

References

Anzengruber J, Pabst M, Neumann L, Sekot G, Heinl S, Grabherr R, Altmann F, Messner P, Schaffer C: Protein *O*-glucosylation in *Lactobacillus buchneri*. Glycoconj J 2014;31:117–131.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25–29.

Balonova L, Hernychova L, Mann BF, Link M, Bilkova Z, Novotny MV, Stulik J: Multimethodological approach to identification of glycoproteins from the proteome of *Francisella tularensis*, an intracellular microorganism. J Proteome Res 2010;9:1995–2005.

Bauerl C, Perez-Martinez G, Yan F, Polk DB, Monedero V: Functional analysis of the p40 and p75 proteins from *Lactobacillus casei* BL23. J Mol Microbiol Biotechnol 2010;19:231–241.

Boel G, Jin H, Pancholi V: Inhibition of cell surface export of group A streptococcal anchorless surface dehydrogenase affects bacterial adherence and antiphagocytic properties. Infect Immun 2005;73:6237–6248.

Castaldo C, Vastano V, Siciliano RA, Candela M, Vici M, Muscariello L, Marasco R, Sacco M: Surface displaced alfa-enolase of *Lactobacillus plantarum* is a fibronectin binding protein. Microb Cell Fact 2009;8:14.

Champasa K, Longwell SA, Eldridge AM, Stemmler EA, Dube DH: Targeted identification of glycosylated proteins in the gastric pathogen *Helicobacter pylori* (Hp). Mol Cell Proteomics 2013;12:2568–2586.

Claes IJ, Schoofs G, Regulski K, Courtin P, Chapot-Chartier MP, Rolain T, Hols P, von Ossowski I, Reunanen J, de Vos WM, Palva A, Vanderleyden J, De Keersmaecker SC, Lebeer S: Genetic and biochemical characterization of the cell wall hydrolase activity of the major secreted protein of *Lactobacillus rhamnosus* GG. PLoS One 2012;7:e31588.

Coyne MJ, Fletcher CM, Chatzidaki-Livanis M, Posch G, Schaffer C, Comstock LE: Phylum-wide general protein *O*-glycosylation system of the Bacteroidetes. Mol Microbiol 2013;88:772–783.

Deepika G, Karunakaran E, Hurley CR, Biggs CA, Charalampopoulos D: Influence of fermentation conditions on the surface properties and adhesion of *Lactobacillus rhamnosus* GG. Microb Cell Fact 2012;11:116.

Ding W, Nothaft H, Szymanski CM, Kelly J: Identification and quantification of glycoproteins using ion-pairing normal-phase liquid chromatography and mass spectrometry. Mol Cell Proteomics 2009;8:2170–2185.

Doron S, Snydman DR, Gorbach SL: *Lactobacillus* GG: bacteriology and clinical applications. Gastroenterol Clin North Am 2005;34:483–498, ix.

Fairbanks G, Steck TL, Wallach DF: Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. Biochemistry 1971;10:2606–2617.

FAO/WHO: Probiotics in Food. Health and Nutritional Properties and Guidelines for Evaluation. Report of a Joint FAO/WHO Expert Consultation on Evaluation of Health and Nutritional Properties of Probiotics in Food including Powder Milk with Live Lactic Acid Bacteria. FAO Food and Nutrition Paper 85. Rome, WHO/Food and Agriculture Organization of the United Nations, 2001, ftp://ftp.fao.org/docrep/fao/009/a0512e/a0512e00.pdf.

Fischer W, Haas R: The RecA protein of *Helicobacter pylori* requires a posttranslational modification for full activity. J Bacteriol 2004;186:777–784.

Fletcher CM, Coyne MJ, Comstock LE: Theoretical and experimental characterization of the scope of protein *O*-glycosylation in *Bacteroides fragilis*. J Biol Chem 2011;286:3219–3226.

Fletcher CM, Coyne MJ, Villa OF, Chatzidaki-Livanis M, Comstock LE: A general *O*-glycosylation system important to the physiology of a major human intestinal symbiont. Cell 2009;137:321–331.

Fredriksen L, Mathiesen G, Moen A, Bron PA, Kleerebezem M, Eijsink VG, Egge-Jacobsen W: The major autolysin Acm2 from *Lactobacillus plantarum* undergoes cytoplasmic *O*-glycosylation. J Bacteriol 2012;194:325–333.

Fredriksen L, Moen A, Adzhubei AA, Mathiesen G, Eijsink VG, Egge-Jacobsen W: *Lactobacillus plantarum* WCFS1 *O*-linked protein glycosylation: an extended spectrum of target proteins and modification sites detected by mass spectrometry. Glycobiology 2013;23:1439–1451.

Granato D, Perotti F, Masserey I, Rouvet M, Golliard M, Servin A, Brassart D: Cell surface-associated lipoteichoic acid acts as an adhesion factor for attachment of *Lactobacillus johnsonii* La1 to human enterocyte-like Caco-2 cells. Appl Environ Microbiol 1999;65:1071–1077.

Hart C, Schulenberg B, Steinberg TH, Leung WY, Patton WF: Detection of glycoproteins in polyacrylamide gels and on electroblots using Pro-Q Emerald 488 dye, a fluorescent periodate Schiff-base stain. Electrophoresis 2003;24:588–598.

Henderson B: An overview of protein moonlighting in bacterial infection. Biochem Soc Trans 2014;42:1720–1727.

Henderson B, Martin AC: Protein moonlighting: a new factor in biology and medicine. Biochem Soc Trans 2014;42:1671–1678.

Huard C, Miranda G, Wessner F, Bolotin A, Hansen J, Foster SJ, Chapot-Chartier MP: Characterization of AcmB, an N-acetylglucosaminidase autolysin from *Lactococcus lactis*. Microbiology 2003;149:695–705.

Izquierdo E, Horvatovich P, Marchioni E, Aoude-Werner D, Sanz Y, Ennahar S: 2-DE and MS analysis of key proteins in the adhesion of *Lactobacillus plantarum*, a first step toward early selection of probiotics based on bacterial biomarkers. Electrophoresis 2009;30:949–956.

Jeffery CJ: Moonlighting proteins: old proteins learning new tricks. Trends Genet 2003;19:415–417.

Kankainen M, Paulin L, Tynkkynen S, von Ossowski I, Reunanen J, Partanen P, Satokari R, Vesterlund S, Hendrickx AP, Lebeer S, De Keersmaecker SC, Vanderleyden J, Hamalainen T, Laukkanen S, Salovuori N, Ritari J, Alatalo E, Korpela R, Mattila-Sandholm T, Lassig A, Hatakka K, Kinnunen KT, Karjalainen H, Saxelin M, Laakso K, Surakka A, Palva A, Salusjarvi T, Auvinen P, de Vos WM: Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human-mucus binding protein. Proc Natl Acad Sci USA 2009;106:17193–17198.

Kinoshita H, Uchida H, Kawai Y, Kawasaki T, Wakahara N, Matsuo H, Watanabe M, Kitazawa H, Ohnuma S, Miura K, Horii A, Saito T: Cell surface *Lactobacillus plantarum* LA 318 glyceraldehyde-3-phosphate dehydrogenase (GAPDH) adheres to human colonic mucin. J Appl Microbiol 2008a;104:1667–1674.

Kinoshita H, Wakahara N, Watanabe M, Kawasaki T, Matsuo H, Kawai Y, Kitazawa H, Ohnuma S, Miura K, Horii A, Saito T: Cell surface glyceraldehyde-3-phosphate dehydrogenase (GAPDH) of *Lactobacillus plantarum* LA 318 recognizes human A and B blood group antigens. Res Microbiol 2008b;159:685–691.

Lebeer S, Claes IJ, Balog CI, Schoofs G, Verhoeven TL, Nys K, von Ossowski I, de Vos WM, Tytgat HL, Agostinis P, Palva A, Van Damme EJ, Deelder AM, De Keersmaecker SC, Wuhrer M, Vanderleyden J: The major secreted protein Msp1/p75 is *O*-glycosylated in *Lactobacillus rhamnosus* GG. Microb Cell Fact 2012;11:15.

Lebeer S, Verhoeven TL, Francius G, Schoofs G, Lambrichts I, Dufrene Y, Vanderleyden J, De Keersmaecker SC: Identification of a gene cluster for the biosynthesis of a long, galactose-rich exopolysaccharide in *Lactobacillus rhamnosus* GG and functional analysis of the priming glycosyltransferase. Appl Environ Microbiol 2009;75:3554–3563.

Malik S, Siezen RJ, Renckens B, Vaneechoutte M, Vanderleyden J, Lebeer S: Draft genome sequence of *Lactobacillus plantarum* CMPG5300, a human vaginal isolate. Genome Announc 2014;2:e01149-14.

Mastroleo F, Leroy B, Van Houdt R, s' Heeren C, Mergeay M, Hendrickx L, Wattiez R: Shotgun proteome analysis of *Rhodospirillum rubrum* S1H: integrating data from gel-free and gel-based peptides fractionation methods. J Proteome Res 2009;8:2530–2541.

Ramiah K, van Reenen CA, Dicks LM: Surface-bound proteins of *Lactobacillus plantarum* 423 that contribute to adhesion of Caco-2 cells and their role in competitive exclusion and displacement of *Clostridium sporogenes* and *Enterococcus faecalis*. Res Microbiol 2008;159:470–475.

Rolain T, Bernard E, Beaussart A, Degand H, Courtin P, Egge-Jacobsen W, Bron PA, Morsomme P, Kleerebezem M, Chapot-Chartier MP, Dufrene YF, Hols P: *O*-glycosylation as a novel control mechanism of peptidoglycan hydrolase activity. J Biol Chem 2013;288:22233–22247.

Saad N, Urdaci M, Vignoles C, Chaignepain S, Tallon R, Schmitter JM, Bressollier P: *Lactobacillus plantarum* 299v surface-bound GAPDH: a new insight into enzyme cell walls location. J Microbiol Biotechnol 2009;19:1635–1643.

Sanchez B, Bressollier P, Chaignepain S, Schmitter JM, Urdaci MC: Identification of surface-associated proteins in the probiotic bacterium *Lactobacillus rhamnosus* GG. Int Dairy J 2009a;19:85–88.

Sanchez B, Schmitter JM, Urdaci MC: Identification of novel proteins secreted by *Lactobacillus rhamnosus* GG grown in de Mann-Rogosa-Sharpe broth. Lett Appl Microbiol 2009b;48:618–622.

Sanchez-Rodriguez A, Tytgat HL, Winderickx J, Vanderleyden J, Lebeer S, Marchal K: A network-based approach to identify substrate classes of bacterial glycosyltransferases. BMC Genomics 2014;15:349.

Scott AE, Twine SM, Fulton KM, Titball RW, Essex-Lopresti AE, Atkins TP, Prior JL: Flagellar glycosylation in *Burkholderia pseudomallei* and *Burkholderia thailandensis*. J Bacteriol 2011a;193:3577–3587.

Scott NE, Bogema DR, Connolly AM, Falconer L, Djordjevic SP, Cordwell SJ: Mass spectrometric characterization of the surface-associated 42 kDa lipoprotein JlpA as a glycosylated antigen in strains of *Campylobacter jejuni*. J Proteome Res 2009;8:4654–4664.

Scott NE, Parker BL, Connolly AM, Paulech J, Edwards AV, Crossett B, Falconer L, Kolarich D, Djordjevic SP, Hojrup P, Packer NH, Larsen MR, Cordwell SJ: Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of *Campylobacter jejuni*. Mol Cell Proteomics 2011b;10:M000031-MCP000201.

Sonck KA, Kint G, Schoofs G, Vander Wauven C, Vanderleyden J, De Keersmaecker SC: The proteome of *Salmonella* Typhimurium grown under in vivo-mimicking conditions. Proteomics 2009;9:565–579.

Thompson R, Creavin A, O'Connell M, O'Connor B, Clarke P: Optimization of the enzyme-linked lectin assay for enhanced glycoprotein and glycoconjugate analysis. Anal Biochem 2011;413:114–122.

Tytgat HL, Lebeer S: The sweet tooth of bacteria: common themes in bacterial glycoconjugates. Microbiol Mol Biol Rev 2014;78:372–417.

Tytgat HL, Schoofs G, Driesen M, Proost P, Van Damme EJ, Vanderleyden J, Lebeer S: Endogenous biotin-binding proteins: an overlooked factor causing false positives in streptavidin-based protein detection. Microb Biotechnol 2015;8:164–168..

Van Damme EJ, Peumans WJ: Biosynthesis of the snowdrop *(Galanthus nivalis)* lectin in ripening ovaries. Plant Physiol 1988;86:922–926.

Van Damme EJ, Smith DF, Cummings R, Peumans WJ: Glycan arrays to decipher the specificity of plant lectins. Adv Exp Med Biol 2011; 705:757–767.

van Sorge NM, Bleumink NM, van Vliet SJ, Saeland E, van der Pol WL, van Kooyk Y, van Putten JP: N-glycosylated proteins and distinct lipooligosaccharide glycoforms of *Campylobacter jejuni* target the human C-type lectin receptor MGL. Cell Microbiol 2009;11: 1768–1781.

Vik A, Aas FE, Anonsen JH, Bilsborough S, Schneider A, Egge-Jacobsen W, Koomey M: Broad spectrum *O*-linked protein glycosylation in the human pathogen *Neisseria gonorrhoeae*. Proc Natl Acad Sci USA 2009;106: 4447–4452.

Yan F, Cao H, Cover TL, Whitehead R, Washington MK, Polk DB: Soluble proteins produced by probiotic bacteria regulate intestinal epithelial cell survival and growth. Gastroenterology 2007;132:562–575.

Young NM, Brisson JR, Kelly J, Watson DC, Tessier L, Lanthier PH, Jarrell HC, Cadotte N, St Michael F, Aberg E, Szymanski CM: Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. J Biol Chem 2002; 277:42530–42539.