Characterization of Outer Membrane Proteome of *Akkermansia muciniphila* Reveals Sets of Novel Proteins Exposed to the Human Intestine

Noora Ottman[1,2], Laura Huuskonen[3], Justus Reunanen[3,4], Sjef Boeren[5], Judith Klievink[6], Hauke Smidt[1], Clara Belzer[1] and Willem M. de Vos[1,3,6]*

[1] Laboratory of Microbiology, Wageningen University, Wageningen, Netherlands, [2] Metapopulation Research Centre, University of Helsinki, Helsinki, Finland, [3] Department of Veterinary Biosciences, University of Helsinki, Helsinki, Finland, [4] Microbiology and Biotechnology, Department of Food and Environmental Sciences, University of Helsinki, Helsinki, Finland, [5] Laboratory of Biochemistry, Wageningen University, Wageningen, Netherlands, [6] Immunobiology, Department of Bacteriology and Immunology, and Research Programs Unit, University of Helsinki, Helsinki, Finland

Akkermansia muciniphila is a common member of the human gut microbiota and belongs to the Planctomycetes-Verrucomicrobia-Chlamydiae superphylum. Decreased levels of *A. muciniphila* have been associated with many diseases, and thus it is considered to be a beneficial resident of the intestinal mucus layer. Surface-exposed molecules produced by this organism likely play important roles in colonization and communication with other microbes and the host, but the protein composition of the outer membrane (OM) has not been characterized thus far. Herein we set out to identify and characterize *A. muciniphila* proteins using an integrated approach of proteomics and computational analysis. Sarkosyl extraction and sucrose density-gradient centrifugation methods were used to enrich and fractionate the OM proteome of *A. muciniphila*. Proteins from these fractions were identified by LC-MS/MS and candidates for OM proteins derived from the experimental approach were subjected to computational screening to verify their location in the cell. In total we identified 79 putative OM and membrane-associated extracellular proteins, and 23 of those were found to differ in abundance between cells of *A. muciniphila* grown on the natural substrate, mucin, and those grown on the non-mucus sugar, glucose. The identified OM proteins included highly abundant proteins involved in secretion and transport, as well as proteins predicted to take part in formation of the pili-like structures observed in *A. muciniphila*. The most abundant OM protein was a 95-kD protein, termed PilQ, annotated as a type IV pili secretin and predicted to be involved in the production of pili in *A. muciniphila*. To verify its location we purified the His-Tag labeled N-terminal domain of PilQ and generated rabbit polyclonal antibodies. Immunoelectron microscopy of thin sections immunolabeled with these antibodies demonstrated the OM localization of PilQ, testifying for its predicted function as a type IV pili secretin in *A. muciniphila*. As pili structures are known to be involved in the modulation of host immune responses,

this provides support for the involvement of OM proteins in the host interaction of *A. muciniphila*. In conclusion, the characterization of *A. muciniphila* OM proteome provides valuable information that can be used for further functional and immunological studies.

INTRODUCTION

Akkermansia muciniphila is a Gram-negative, anaerobic bacterium, which colonizes the mucus layer of the human gastrointestinal (GI) tract (Derrien et al., 2004). *A. muciniphila* is considered to be an important member of the GI microbiota, because of the inverse correlation between its abundance and several intestinal disorders, including inflammatory bowel diseases and obesity (Png et al., 2010; Karlsson et al., 2012; Rajilic-Stojanovic et al., 2013). Moreover, experiments with germ-free mice mono-associated with *A. muciniphila*, or conventional mice on a high fat diet that are fed *A. muciniphila*, have shown that *A. muciniphila* plays a role in host immune response, restoration of mucus layer thickness and mucus production (Derrien et al., 2011; Everard et al., 2013; Shin et al., 2014). In addition, extracellular vesicles from *A. muciniphila* were shown to have protective effects on the development of dextran sulfate sodium (DSS) induced colitis in mice (Kang et al., 2013). Finally, *A. muciniphila* has also been shown to adhere to intestinal epithelium and improve enterocyte monolayer integrity of Caco-2 cells (Reunanen et al., 2015). These findings suggest important host-bacteria interactions, the mechanisms of which are yet to be discovered (see Derrien et al., 2016 for a recent review).

Bacterial outer membrane (OM) proteins play important roles in communication with other microbes and the host, as well as in colonization and substrate transport (Tseng et al., 2009; Galdiero et al., 2012). Subcellular fractionation techniques combined with mass spectrometry-based proteomic analysis are powerful tools for identifying proteins in different bacterial compartments. These techniques have been successfully used for studying the protein composition of intestinal bacteria, such as the OM of *Bacteroides fragilis* and *Bacteroides thetaiotaomicron* (Elhenawy et al., 2014; Wilson et al., 2015), surface proteins of *Propionibacterium freudenreichii* (Le Marechal et al., 2015) and OM vesicles of *Escherichia coli* Nissle 1917 (Aguilera et al., 2014).

Akkermansia muciniphila is a member of the Planctomycetes-Verrucomicrobia-Chlamydiae (PVC) superphylum, which contains bacteria from several groups and various environments with different lifestyles (Wagner and Horn, 2006; Gupta et al., 2012; Kamneva et al., 2012). Bacteria from this superphylum were previously suggested to have a compartmentalized cell plan with a cytoplasmic membrane as the outermost membrane, and an intracytoplasmic membrane containing a condensed nucleoid and ribosomes (Lee et al., 2009). However, these observations have been challenged by more recent data suggesting that the PVC cell plan is actually a variation, not an exception, of the Gram-negative cell plan, and that the bacteria have an outer and an inner membrane (IM) with possible invaginations of the IM inside the cytoplasm (Devos, 2014).

There is limited information available on the membrane structure and composition of *A. muciniphila,* and most reports have focused on *in silico* analysis of Verrucomicrobia membranes, instead of experimental approaches (Santarella-Mellwig et al., 2010; Kamneva et al., 2012; Speth et al., 2012). Recently, the proteome of a termite hindgut representative of the Verrucomicrobia, *Diplosphaera colotermitum* TAV2, was experimentally studied, but this report did not focus on membrane proteins (Isanapong et al., 2013). The presence of OM biomarkers, including genes involved in lipopolysaccharide (LPS) insertion, in the genome of *A. muciniphila* was confirmed computationally (Speth et al., 2012). We have experimentally verified the presence of LPS in *A. muciniphila* (Ottman, 2015). No genes coding for membrane coat-like proteins were found in *A. muciniphila*, unlike in some other Verrucomicrobia (Santarella-Mellwig et al., 2010).

In the present study, we set out to identify and characterize *A. muciniphila* proteins using an integrated approach of proteomics and computational analysis. Successful extraction of OM proteins was established, and the proteins were identified with liquid chromatography-tandem mass spectrometry (LC-MS/MS). The abundance of *A. muciniphila* proteins in the OM fraction was compared to the whole proteome of *A. muciniphila* and a fraction enriched for intracellular proteins. Candidates for OM proteins derived from the proteomics analysis were subjected to computational screening to verify their location in the cell. The OM location of the most abundant membrane protein, termed PilQ, in *A. muciniphila* was confirmed by immunoelectron microscopy. PilQ is predicted to be a secretin involved in the production of type IV pilins, long surface exposed filaments involved in a variety of functions, including motility and adherence to host cells, and found in virtually all prokaryotes, including the PVC superphylum (Berry and Pelicic, 2015). We also explored the presence of these proteins as a function of the growth on mucus, resulting in the identification of several OM proteins related to mucosal colonization. The results indicate that *A. muciniphila* produces OM proteins involved in secretion and transport in high abundance, and these, including PilQ, may be involved in its capacity to interact with the host.

MATERIALS AND METHODS

Bacterial Growth Conditions

Akkermansia muciniphila Muc[T] (ATTC BAA-835) was grown in a basal medium as described previously, except without the

addition of rumen fluid (Derrien et al., 2004). The medium was supplemented with either hog gastric mucin (0.5 %, Type III; Sigma-Aldrich, St. Louis, MO, USA) purified by ethanol precipitation as described previously (Miller and Hoskins, 1981), or D-glucose (10 mM, Sigma–Aldrich). The medium with glucose was also supplemented with Bacto^{TM} casitone (BD, Sparks, MD, USA), BBL^{TM} yeast extract (BD), tryptone (Oxoid Ltd, Basingstoke, Hampshire, England), peptone (Oxoid ltd) (2 g/l each) and L-threonine (Sigma–Aldrich) (2 mM). Incubations were performed in serum bottles sealed with butyl-rubber stoppers at 37°C under anaerobic conditions provided by a gas phase of 182 kPa (1.5 atm) N_2/CO_2. Growth was measured by following optical density at 600 nm (OD600) using a spectrophotometer.

Bacterial Fractionation Methods

Membrane proteins of *A. muciniphila* were isolated from liquid cultures with two different methods, using either *N*-lauroyl-sarcosine (sarkosyl) or sucrose density-gradient centrifugation, as described previously, with minor modifications (Hobb et al., 2009).

Briefly, for sarkosyl treatments, 250 ml cultures of *A. muciniphila* were grown on mucin for 16 h (OD$_{600}$ = 1.56) or on glucose for 40 h (OD$_{600}$ = 0.34). Cells were harvested, washed twice with phosphate buffered saline (PBS) and resuspended in 9 ml 10 mM HEPES, pH 7.4, and lysed by passing the culture three times through a French press (Aminco, American Instrument Co., Inc., Silver Spring, Maryland, USA) at 1000 psi (40 K cell). The lysed cell preparation was centrifuged at 10 000 *g* for 10 min at 4°C to remove cell debris and unlysed cells. The membranes were collected by ultracentrifugation of the supernatant at 100 000 *g* for 1 h at 4°C. The supernatant was collected and stored at −20°C. This sample was later analyzed by LC-MS/MS as the intracellular fraction. The pellet was resuspended in 2 ml 10 mM HEPES, pH 7.4, washed in a total volume of 10 ml 10 mM HEPES, pH 7.4, and spun again in the ultracentrifuge (using the conditions described above). The pellet was resuspended in 5 ml 1 % (w/v) *N*-lauroylsarcosine (sarkosyl) (Sigma–Aldrich) in 10 mM HEPES, pH 7.4, and incubated at 37°C for 30 min with shaking to solubilize cytoplasmic membranes. The sarkosyl-treated membranes were spun at 100 000 *g* for 1 h at 4°C and the pellet was washed with 7 ml 10 mM HEPES, pH 7.4. Following the final ultracentrifugation, the pellet containing the OM fraction was resuspended in 1 ml 10 mM HEPES, pH 7.4 and stored at −20°C.

For the sucrose-density gradient centrifugation treatments, 250 ml cultures of *A. muciniphila* were grown on mucin for 16 h (OD600 = 1.51) or on glucose for 40 h (OD600 = 0.35). Cells were harvested, washed twice with phosphate buffered saline (PBS) and resuspended in 7 ml 10 mM HEPES, pH 7.4, and lysed by passing the culture three times through a French press at 1000 psi (40K cell). The lysed cells preparation was centrifuged at 10 000 *g* for 10 min at 4°C. The supernatant was ultracentrifuged at 100 000 *g* for 60 min at 4°C to pellet the total membranes. The membrane pellet was washed in 10 ml 10 mM HEPES, 0.05 M EDTA pH 7.5 (HE buffer) and ultracentrifuged again. The final membranes were homogenized in 2 ml HE buffer.

Continuous sucrose gradients were prepared by layering sucrose solutions (prepared in HE buffer) into 14 × 95 mm polyallomer ultracentrifuge tubes (Seton Scientific, Petaluma, CA, USA) in the following order: 0.4 ml 60% (w/v), 0.9 ml 55%, 2.2 ml 50%, 2.2 ml 45%, 2.2 ml 40%, 1.3 ml 35% and 0.4 ml 30%. Total membranes were layered on top of each gradient, with no more than 2.5 ml per gradient. Sucrose gradients were centrifuged in a TST 41.14-41000 RPM swinging-bucket rotor (Kontron) at 250 000 *g* for 16 h at 4°C. The sucrose gradient tubes were then removed from the rotor buckets and 500 µl fractions (24 fractions for each gradient) were collected from the bottom of each gradient by puncturing the tube with a needle and allowing the sample to drip out by gravity. The samples were stored in 2 ml low binding tubes (Eppendorf, Hamburg, Germany) at −20°C.

For comparison, the whole proteome fraction was obtained from *A. muciniphila* cultures grown with mucin or glucose as the carbon source. Bacterial cells from an overnight 2 ml culture were spun down, washed with PBS and suspended into 500 µl of PBS. Cells were lysed by sonication, using a Branson sonifier equipped with a 3 mm tip (four pulses of 20 s with 30 s rest on ice in-between each pulse, strength of the pulse was 4). The samples were stored in 2 ml low binding tubes (Eppendorf, Hamburg, Germany) at −20°C.

Protein Identification by Mass Spectrometry

To determine the protein content of cell extracts, the Qubit® Protein Assay Kit (Life technologies, Eugene, OR, USA) was used according to the manufacturer's instructions. Samples were loaded on a 10% acrylamide separation gel (25201, Precise^{TM} Protein Gels, Thermo Scientific, Rockford, IL, USA) using the mini-PROTEAN 3 cell (Bio-Rad Laboratories, Hercules, CA, USA). The electrophoresis procedure was according to the manufacturer's instructions. Gels were stained using Coomassie Brilliant Blue (CBB) R-250 as indicated in the protocol of the mini-PROTEAN 3 system.

In-gel digestion of proteins and purification of peptides were done following a modified version of a previously described protocol (Rupakula et al., 2013). Disulphide bridges in proteins were reduced by covering whole gels with reducing solution (10 mM dithiothreitol, pH 7.6, in 50 mM NH_4HCO_3), and the gels were incubated at 60°C for 1 h. Alkylation was performed for 1 h by adding 25 ml of iodoacetamide solution (10 mM iodoacetamide in 100 mM Tris-HCl, pH 8.0). Gels were thoroughly rinsed with demineralized water in between steps. Each of the used gel lanes was cut into either five slices (sarkosyl-extracted OM, intracellular proteins and whole proteome) or one slice (sucrose density-gradient centrifugation samples), and the slices were cut into approximately 1 mm × 1 mm × 1 mm cubes and transferred to separate 0.5 ml protein LoBind tubes (Eppendorf, Hamburg, Germany). Enzymatic digestion was done by adding 50 µl of trypsin solution (5 ng/µl trypsin in 50 mM NH_4HCO_3) to each tube, and by incubating at room temperature overnight with gentle shaking. Extraction of peptides was performed with manual sonication in an ultrasonic water bath for 1 s before the supernatant was transferred to

a clean protein LoBind tube. Trifluoroacetic acid (10%) was added to the supernatant to reach a pH between 2 and 4. The supernatant was used for LC-MS/MS analysis. Samples were measured by nLC–MS/MS with a Proxeon EASY nLC and a LTQ-Orbitrap XL mass spectrometer as previously described (Lu et al., 2011).

LC–MS data analysis was performed as described previously (Rupakula et al., 2013) with false discovery rates (FDRs) set to 0.01 on peptide and protein level, and additional result filtering (minimally 2 peptides necessary for protein identification of which at least one is unique and at least one is unmodified). To analyze the abundance of proteins in the fractions, their label-free quantification (LFQ) intensities was compared (Cox et al., 2014). Non-existing LFQ intensity values due to not enough quantified peptides were substituted with a value lower than the LFQ intensity value for the least abundant, quantified protein. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository with the dataset identifier PXD004400.

Enzyme Activity Assay

Enzyme activity was tested in the OM and intracellular fractions by incubating them at 37°C with colorimetric or fluorimetric substrates, as described previously (Rosendale et al., 2012). The following substrate/enzyme combinations were used: 4-nitrophenyl *N*-acetyl-β-D-glucosaminide/*N*-acetyl-β-glucosaminidase, *p*-nitrophenyl-alpha-L-fucoside/fucosidase, PNP *N*-acetyl-glucosamine-sulfate/GlcNAc-sulfatase, and 2′-(4-methylumbelliferyl)-α-D-*N*-acetylneuraminic acid/sialidase. Substrates were purchased from Carbosynth (Compton, Berkshire, UK). Reaction volume was 20 μl and substrate concentration was 1 mM in each reaction. All reactions were performed in duplicate.

Bioinformatics Analysis

Proteins were categorized based on results of SignalP v.4.0 (set for Gram-negative bacteria) (Petersen et al., 2011), LipoP v.1.0 (Juncker et al., 2003), SecretomeP v.2.0 (set for Gram-negative bacteria) (Bendtsen et al., 2005), TMHMM v.2.0 (Krogh et al., 2001) and BOMP (Berven et al., 2004). Subcellular localization was determined using CELLO v.2.5 (set for Gram-negative bacteria) (Yu et al., 2004) and PSORTb v3.0 (set for Gram-negative bacteria) (Yu et al., 2010). BLAST searches were run against the non-redundant (nr) database at http://blast.ncbi.nlm.nih.gov/ using default settings. TIGRFAM and Pfam were used for screening proteins with PEP-CTERM domains (Selengut et al., 2007; Finn et al., 2014).

Overproduction and Purification of Amuc_1098A Protein

The DNA coding for the N-terminal half of PilQ, the product of the gene with locus tag Amuc_1098, was amplified by PCR from the chromosomal DNA of *A. muciniphila* using a primer pair containing NdeI site (5′ATA*CATATG*GATGGCGGCGCCGTCGGAACCTC; restriction site in italics) and XhoI

site (5′ATA*CTCGAG*GGGGCTTAAGGCCGGAGGAGCTTT; restriction site in italics) (Sigma–Aldrich). The obtained PCR product was digested with NdeI and XhoI and cloned into similarly digested pET26b vector (Novagen) and transformed into *E. coli* TOP10 (Invitrogen). Agarose gel electrophoresis and sequencing were used to verify the right insert and the obtained plasmid (termed pAmuc_1098A) was subsequently transformed into the expression strain *E. coli* BL21 (DE3) (Invitrogen) containing plasmid pSJS1240 (Kim et al., 1998) and transformants were selected for resistance to kanamycin and spectinomycin (both at 50 μg/ml). Protein production was induced in an exponentially growing aerobic culture at optical density at 600 nm of approximately 0.5 by adding isopropyl β-d-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After a further overnight incubation at 37°C, cells were collected and lysed by sonication and the protein, referred to as P-Amuc_1098A, was purified under native conditions using Ni-NTA His Bind Resin (Novagen) according to manufacturer's instructions. The purity was tested with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Generation of Polyclonal Antibody against P-Amuc_1098A Protein

Polyclonal rabbit antibodies were produced against P-Amuc_1098A protein with an immunization protocol as described previously (Johnston et al., 1991), at the Laboratory Animal Centre of the University of Helsinki. In short, a rabbit was immunized four times with 3 weeks interval for each injection. Pre-immunization sera were collected before initiating the immunization. In the first immunization, purified P-Amuc_1098A protein (192 μg) was mixed with an equal amount of Freund's complete adjuvant. For the subsequent immunizations (booster injections), P-Amuc_1098A protein (96 μg) was diluted 1:1 volume in Freund's incomplete adjuvant and injected into rabbit. Blood was collected from the rabbit 10 days after the last booster injection. The blood was allowed to clot for 1-2 h at 37°C after which the blood clot was separated from the serum by centrifugation at 3 000 *g* for 10 min. Then the serum was divided into aliquots and stored at −80°C prior to use.

Immunogold Labeling and Electron Microscopy of Thin Sections

Immunogold labeling of *A. muciniphila* thin sections was performed as described previously with minor changes (Reunanen et al., 2012). In short, *A. muciniphila* cells were grown in mucin medium for two overnights and washed once with phosphate buffer (0.1 M Na-phosphate, pH 7.4) before fixation. The fixed cells were embedded in Lowicryl HM20 resin, 50 nm thin sections were cut from polymerized Lowicryl and mounted on nickel grids. Grids containing the *A. muciniphila* thin sections were treated with anti-P-Amuc_1098A rabbit immune serum (1:5 dilution), and subsequently with protein A conjugated to 10 nm diameter gold particles. Finally, the grids were post-stained with uranyl acetate and lead citrate using an UltroStain apparatus (Leica). The grids were examined and

images obtained with a JEOL JEM-1400 transmission electron microscope (Jeol Ltd., Tokyo, Japan).

Statistical Analysis

Statistical analysis of the results from the enzyme activity assay was performed by one-way analysis of variance (ANOVA) using IBM SPSS software (IBM SPSS Statistics 22); p values <0.05 were considered significant.

RESULTS

Akkermansia muciniphila OM proteins were isolated using two different methods, based on either selective sarkosyl extraction or sedimentation gradient centrifugation of membrane proteins, derived from cells grown with mucin or glucose as major carbon and energy sources (**Figure 1**). Label-free mass spectrometry of the trypsin-treated fractions was performed to identify the extracted proteins in comparison with those recovered from total cells (whole proteome). In addition, the location of the identified proteins was predicted by using the CELLO software. To determine the abundance of proteins within the fractions from different cellular locations, we compared their LFQ intensities as these are commonly used as a proxy for absolute protein abundance in mass spectrometry data analysis (Cox et al., 2014).

An over 20-fold enrichment of predicted OM proteins was achieved in the sarkosyl-treated membranes in comparison with the whole proteome and intracellular fraction of both mucin and glucose-grown cells of *A. muciniphila* (**Figures 1A,B**). Moreover, the predicted periplasmic and extracellular proteins were slightly enriched in the OM fraction at the expense of cytoplasmic proteins that were enriched in the intracellular fraction.

Complementary to the sarkosyl treatment procedure, we applied sucrose density gradient centrifugation of *A. muciniphila* membrane fractions obtained from cells grown on mucin or glucose. Analysis of the recovered proteins from each fraction revealed different protein abundance in the mucin versus glucose-grown cell membranes, likely reflecting the adaptation to the different environmental conditions (Supplementary Figure S1). The abundance of OM proteins was the lowest in the fraction with the lowest sucrose density, and increased along the gradient for proteins derived from both growth conditions (**Figures 1C,D**). However, the enrichment of predicted OM proteins was not more than a factor of 2–5 by the sucrose gradient centrifugation and hence it was evident that the sarkosyl treatment procedure was superior (**Figures 1C,D**).

The sarkosyl-treated OM fraction was compared with the intracellular fraction derived from the same *A. muciniphila* cells (**Figures 1A,B**) for the activity of four relevant mucin-degrading
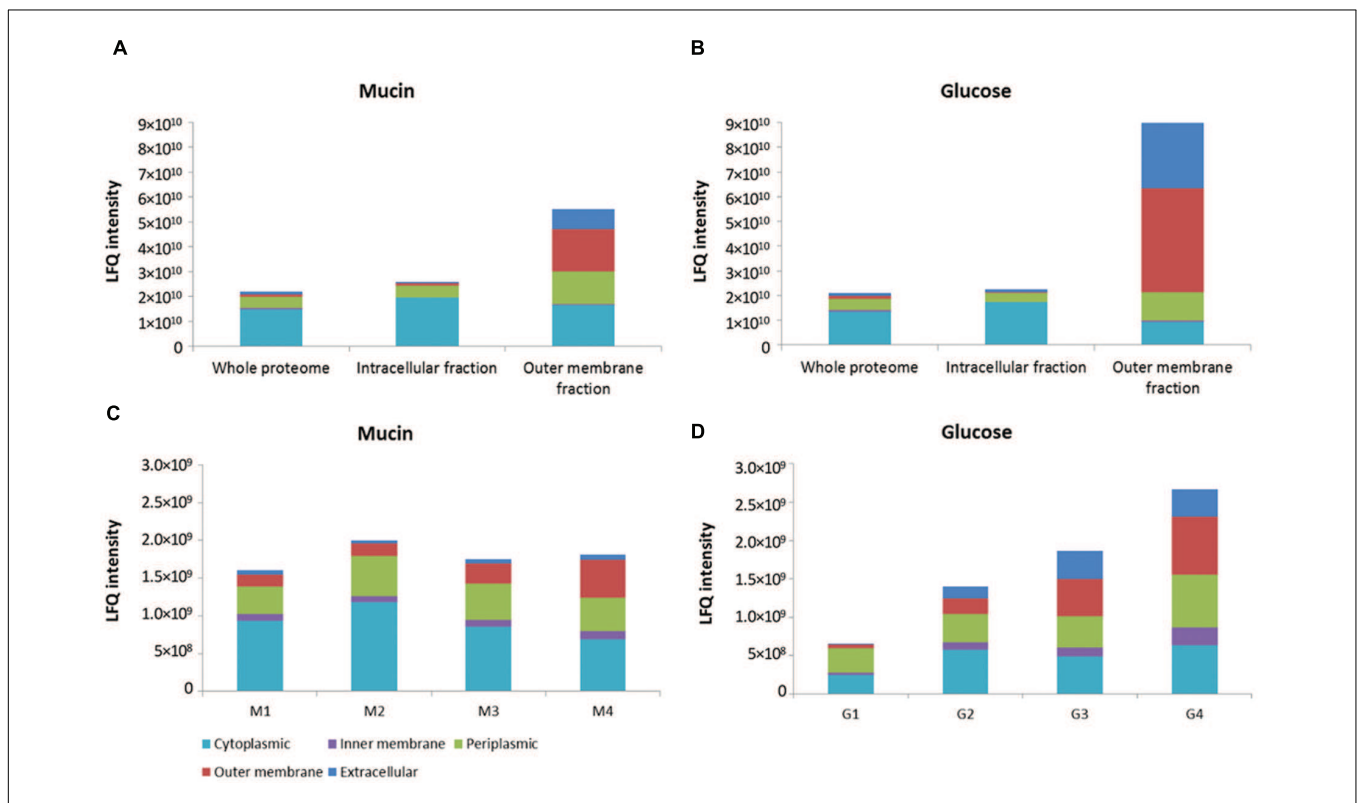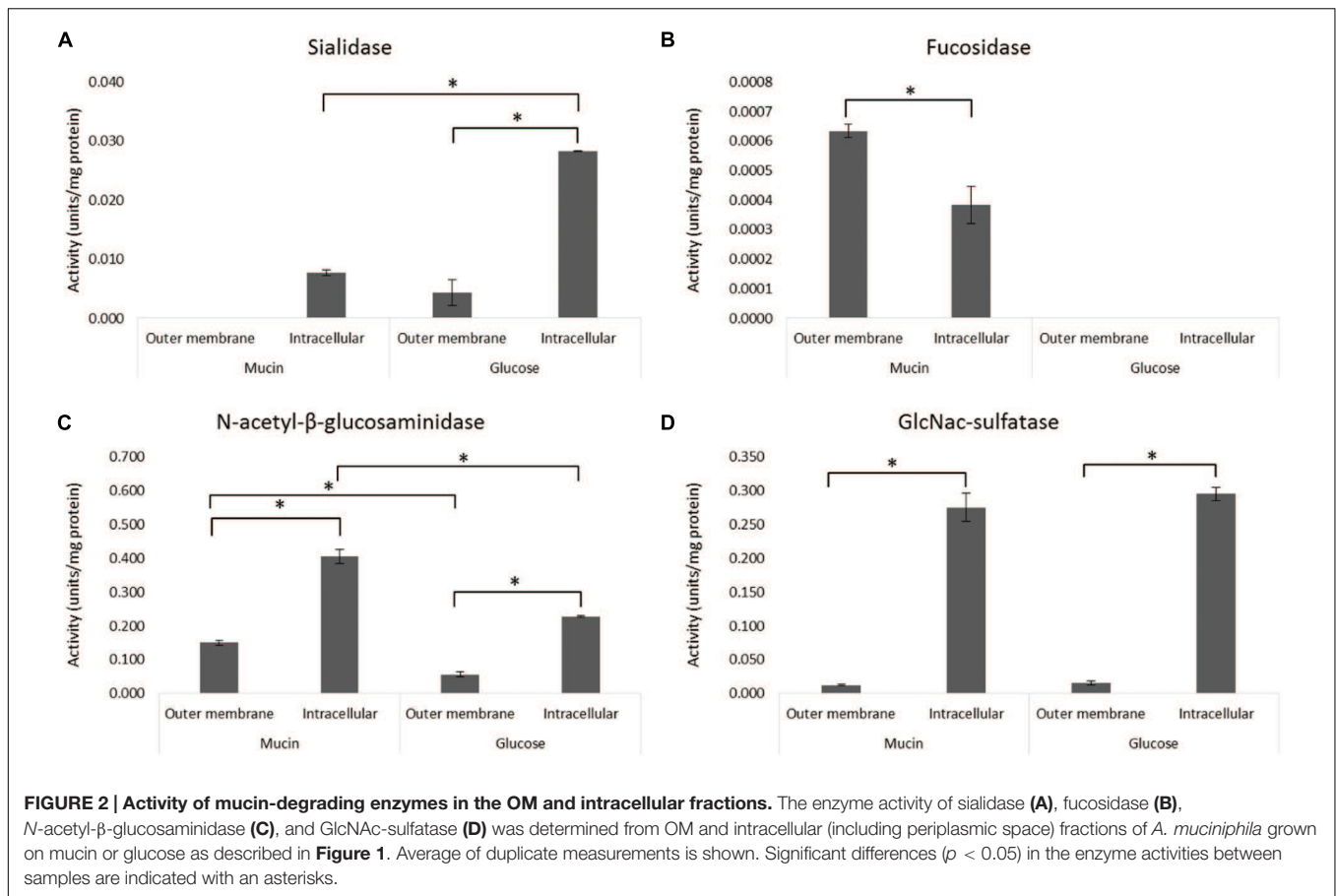

FIGURE 1 | Distribution of proteins from different bacterial compartments in isolated fractions. OM fraction isolated from sarkosyl-extracted membranes in comparison to whole proteome and intracellular fraction of both mucin and glucose-grown cells of *A. muciniphila* **(A,B)**. Fractions of sucrose density-gradient centrifugation (see Supplementary Figure S1) of membranes of *A. muciniphila* grown on mucin or glucose **(C,D)**. Label-free quantification (LFQ) intensity was used as a proxy for absolute protein abundance. CELLO software was used to predict protein localization.

FIGURE 2 | Activity of mucin-degrading enzymes in the OM and intracellular fractions. The enzyme activity of sialidase **(A)**, fucosidase **(B)**, *N*-acetyl-β-glucosaminidase **(C)**, and GlcNAc-sulfatase **(D)** was determined from OM and intracellular (including periplasmic space) fractions of *A. muciniphila* grown on mucin or glucose as described in **Figure 1**. Average of duplicate measurements is shown. Significant differences ($p < 0.05$) in the enzyme activities between samples are indicated with an asterisks.

enzymes (sialidase, fucosidase, *N*-acetyl-β-glucosaminidase, GlcNAc-sulfatase) (**Figure 2**). Except for the fucosidase activity, it appeared that the sialidase, *N*-acetyl-β-glucosaminidase and GlcNAc-sulfatase activity was significantly higher in the intracellular than the OM fractions, for both glucose and mucin-grown cells (**Figures 2A,C,D**). As none of the candidate enzymes for these activities was predicted to be located in the OM, this supported the purification of OM proteins by the sarkosyl treatment. In contrast, most candidate proteins were predicted to carry a signal sequence and hence their presence in the intracellular fraction indicated that they are at least partly among the periplasmic space proteins that are also found in this fraction (**Figures 1A,B**). Different results were obtained for the fucosidase activity (**Figure 2B**). This enzymatic activity was enriched in the OM fraction, which is compatible with the experimentally verified OM location of one of the fucosidases (encoded by the gene with locus tag Amuc_0146; see below). Remarkably, this fucosidase activity was strongly induced in the extracts of *A. muciniphila* grown on mucin, which is in line with the production of 1,2-propanediol from fucose during growth on mucin but not on glucose (Ottman, 2015).
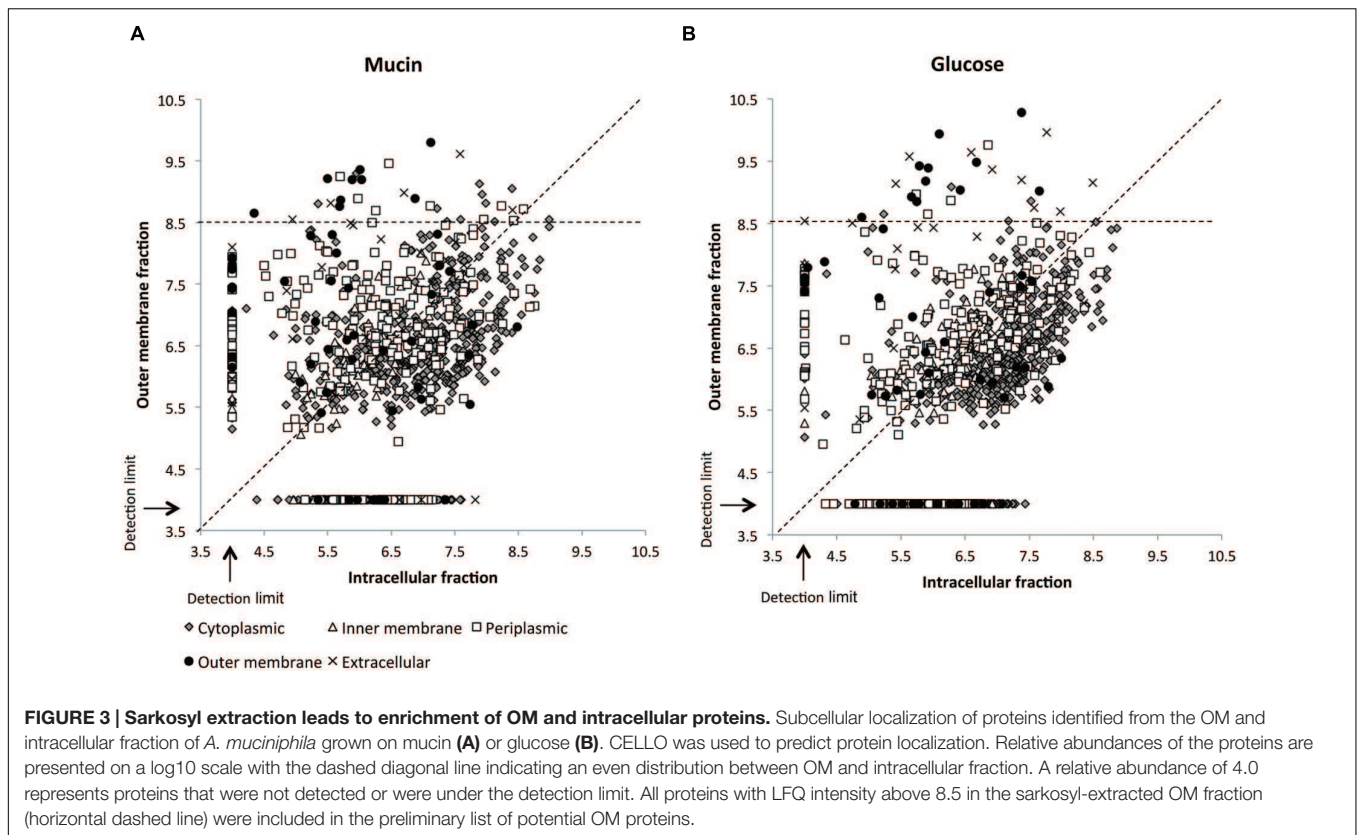
Analysis of the *A. muciniphila* Proteome

While over 1000 different proteins could be identified in the total *A. muciniphila* proteome, the number of sarkosyl-treated OM proteins amounted to slightly lower amounts (812 and 719 proteins from the mucin and glucose grown cells, respectively) but exceeded those found in the sucrose gradient fractions (up to ∼500 proteins), indicative of the efficiency of the detergent treatment (Supplementary Table S1). Of further interest was the comparison of the protein abundancies of these sarkosyl-treated OM proteins and the intracellular proteins obtained from the same *A. muciniphila* cells, grown on either mucin or glucose (**Figure 3**). The resulting comparative proteome plots showed that many of the predicted OM proteins were highly abundant in the OM fraction with a relative LFQ intensity above 8.5. Remarkably, several predicted extracellular, periplasmic and cytoplasmic proteins (12, 12, and 14, respectively) were also similarly abundant and enriched in the OM fraction in *A. muciniphila* cells grown on either mucin or glucose. Hence, all the proteins with an LFQ intensity above 8.5 and proteins predicted as OM proteins by CELLO or PSORTb (a total of 145) were further analyzed as described below.

Identification, Annotation and Classification of *A. muciniphila* OM Proteins

Next, we confirmed the predicted localization of the 145 proteins that were experimentally found to be enriched in the OM fraction by applying several criteria, supported by advanced

FIGURE 3 | Sarkosyl extraction leads to enrichment of OM and intracellular proteins. Subcellular localization of proteins identified from the OM and intracellular fraction of *A. muciniphila* grown on mucin **(A)** or glucose **(B)**. CELLO was used to predict protein localization. Relative abundances of the proteins are presented on a log10 scale with the dashed diagonal line indicating an even distribution between OM and intracellular fraction. A relative abundance of 4.0 represents proteins that were not detected or were under the detection limit. All proteins with LFQ intensity above 8.5 in the sarkosyl-extracted OM fraction (horizontal dashed line) were included in the preliminary list of potential OM proteins.

bioinformatic predictions: a protein was considered as a potential OM protein if (1) both CELLO and PSORTb programs predicted the protein to be localized in the OM, (2) the LipoP software predicted the protein to be a lipoprotein, (3) BOMP software predicted the presence of beta-barrel(s), or (4) the protein had no signal protein, but SecretomeP predicted unclassical secretion (Supplementary Figure S2). Proteins that had a signal peptide but did not fulfill any of the above criteria were assumed to be periplasmic. After applying these criteria, 73 proteins from mucin-grown cells and 76 proteins from glucose-grown cells in the isolated OM fractions were predicted to be true OM proteins. Moreover, 5 out of the 14 cytoplasmic and 9 out of the 12 periplasmic proteins included in the preliminary list were identified as true OM proteins. In total, 79 truly predicted OM and membrane-associated extracellular proteins were identified (Supplementary Table S2). It is of interest to note that the majority of these were also present in the sucrose density-gradient samples: 62 proteins for the mucin-grown and 58 proteins for the glucose-grown condition (data not shown).

Additionally, we searched for potential OM proteins, which were not present in either of the sarkosyl-extracted protein fractions. The *A. muciniphila* genome was found to encode additional 17 proteins, which were predicted as OM proteins according to the applied criteria (see above), and 12 (71%) of these were uncharacterized proteins (Supplementary Table S3). Two of the proteins (products of genes tagged as Amuc_0356 and Amuc_0480) were found to be present in the sucrose density-gradient fractions. Three of the proteins (products of

genes tagged as Amuc_0983, Amuc_1011 and Amuc_1945) were present in low amounts when the whole proteome of *A. muciniphila* was analyzed. However, all these proteins are likely to be produced, since RNA-seq analysis of *A. muciniphila* grown on mucin or glucose confirmed the production of transcripts for the corresponding 17 genes (Ottman, 2015).

The proteins identified as true OM proteins had multiple and different functions (Supplementary Figure S3). Nine proteins obtained from cells grown at both conditions were OM proteins containing an autotransporter barrel domain. OM proteins enriched during growth on mucin (total of four) or on glucose (total of three) were found to be involved in RND (Resistance-Nodulation-Division) efflux systems. Mucin-grown cells contained three and glucose-grown cells two YD repeat proteins. The function of many of the predicted OM proteins was unknown (26 proteins for mucin-grown and 29 for glucose-grown cells). BLAST searches were performed for all uncharacterized proteins with LFQ intensity above 8.5, but this did not reveal any closely related proteins.

Influence of Environmental Growth Conditions on OM Protein Abundancies

From the identified OM proteins, 23 out of the 79 showed more than a 10-fold change in abundancy between *A. muciniphila* grown on either mucin or glucose (Supplementary Table S2). Six out of the nine proteins containing an autotransporter barrel domain were present at a higher level in glucose-grown as

compared to mucin-grown cells. In addition, a two-component regulator propeller domain protein and a peptidase were more abundant in mucin-grown cells, whilst a glycoside hydrolase family 16 protein and peptidyl-prolyl *cis-trans* isomerase were present in higher abundance in glucose-grown cells (products of genes with locus tags Amuc_0371 and Amuc_0576 as well as Amuc_2108 and Amuc_1026, respectively). Many uncharacterized proteins were among the ones with highest fold-change differences between the two conditions. There were four uncharacterized proteins (products of genes with locus tags Amuc_0789, Amuc_0837, Amuc_1333, Amuc_2165) exclusively found in the OM fraction extracted from glucose-grown cells and one uncharacterized protein (encoded by the gene with locus tag Amuc_0006), exclusively found in mucin-grown cells.

BLAST searches and subsequent TIGRFAM and PFAM analyses revealed that three of the glucose-exclusive proteins in the OM fraction contained a PEP-CTERM domain (Supplementary Table S4). This domain is thought to be involved in protein sorting and cell surface localization (Haft et al., 2006). The domain includes the motif Pro-Glu-Pro (PEP), which is considered a potential recognition or processing site, followed by a predicted transmembrane helix and a cluster rich in basic amino acids. These target proteins are generally destined to transit cellular membranes during their biosynthesis and undergo further posttranslational modifications, such as glycosylation. PEP-CTERM domains have been first identified in the predicted proteome of *Verrucomicrobium spinosum* and have only been found in the genomes of several Gram-negative bacteria that encode EpsH (exopolysaccharide locus protein H), including members of the PVC cluster (Haft et al., 2012). An EpsH-like gene is also found in the genome of *A. muciniphila* and upon further screening for the presence of PEP-CTERM proteins encoded by the genome of *A. muciniphila*, we found 23 predicted proteins containing this domain (Supplementary Table S4). The synthesis of 9 of these was confirmed by mass spectrometry, as they were present in at least one of the analyzed bacterial fractions. However, these appeared to be differentially regulated as some were only found in *A. muciniphila* grown on either glucose or mucin.

Abundant OM Proteins and OM Location of the PilQ Protein in *A. muciniphila*

Three OM proteins (encoded by the genes with locus tags Amuc_0336, Amuc_1310 and Amuc_1098) were the most highly abundant in *A. muciniphila* grown on mucin or glucose, altogether making up to 20–40% of the total OM protein (Supplementary Table S2). BLAST analysis predicted the 17-kD protein encoded by the gene with locus tag Amuc_1310 to belong to a protein superfamily that includes several *Rickettsia* genus-specific 17 kD-surface antigen proteins, suggesting its potential interaction with the environment. The 82-kD protein encoded by the gene with locus tag Amuc_0336 could be annotated as a TonB-dependent receptor and showed low but significant similarity (sequence identity ∼25%) to TonB-dependent receptors from *Pseudomonas* spp, which are OM proteins involved in heme and iron transport (Takase et al., 2000).

By far the most prominent protein, which was found to be over twofold more abundant than any other OM protein, is the product of gene with locus tag Amuc_1098, further termed PilQ. The 907-residue protein PilQ was initially annotated as a types II and III secretion system protein (see Supplementary Table S2) and contains a highly conserved pilus secretin domain at position 650-876, while its N-terminal part has a conserved domain at position 71-239 of a beta-barrel assembly machinery (BAM) containing three tetratricopeptide repeats. Highly similar homologues (25–40% identity) of the *A. muciniphila* PilQ are predicted to be produced by other Verrucomicrobia, such as *V. spinosum*, *Haloferula* sp. and *Lentisphaera araneosa*, the homologue of the latter being annotated as the type IV fimbrial biogenesis protein PilQ. These results strongly suggests that Verrucomicrobia, and specifically *A. muciniphila,* may produce type IV pili that are secreted in a process that involves the OM located secretin PilQ (Berry and Pelicic, 2015).

Previous electron microscope (EM) images from *A. muciniphila* have revealed pili-like structures (Derrien et al., 2004). Therefore, we used immunogold labeling of *A. muciniphila* thin sections to determine the location of PilQ. The N-terminal domain of PilQ (termed P-Amuc_1098A) labeled with a C-terminal His-Tag was overproduced and purified from *E. coli* since the gene for the entire PilQ could not be stably maintained in the *E. coli* host Top10 (data not shown). Polyclonal rabbit antibodies against the purified termed P-Amuc_1098A were generated, labeled with protein A-conjugated gold particles, and used to treat thin sections of *A. muciniphila* cells that were subsequently examined by transmission electron microscopy (**Figure 4**). The results show that a major part of the protein A–gold particles were located in the OM, confirming the location of PilQ in the OM of *A. muciniphila*. Moreover, each thin section has various PilQ molecules, as is to be expected by the high level of abundance of the OM protein PilQ in *A. muciniphila.*

DISCUSSION

Using a combination of experimental mass-spectrometry based identification and advanced bioinformatics, we identified 79 putative OM proteins in *A. muciniphila*, which comprises 3.6% of the 2176 predicted protein-coding sequences in this mucus-degrading intestinal bacterium. This fraction is slightly higher than that found in other Gram-negative bacteria where it was estimated that 2–3% of the genome codes for OM proteins (Molloy et al., 2000; Wimley, 2003). As we included membrane-associated extracellular proteins in the analysis, the actual number of OM proteins in *A. muciniphila* is presumably somewhat lower. Extracellular proteins were included as they may be involved in communication with the host and are therefore of great interest. Seven complementary bioinformatic algorithms were applied to characterize *A. muciniphila* OM and extracellular proteins based on the predicted secretion signals, occurrence of transmembrane strands, presence of beta-barrel structures and attachment to the cell wall by lipid motifs. Moreover, the enrichment of the OM fractions was characterized by analysis of the activity of mucin-degrading
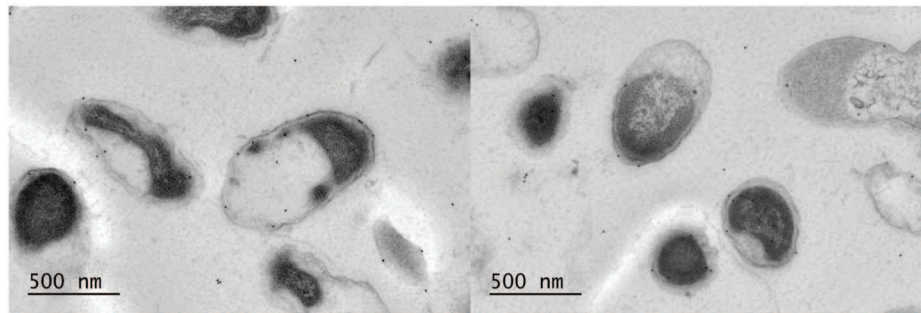
FIGURE 4 | Immunoelectron microscopy labeling of *A. muciniphila* thin sections visualizes PilQ to be localized on the OM. The cells were labeled with anti-P-Amuc_1098A antibodies and detection was done by using 10-nm protein A-conjugated gold particles. Two representative electronmicrographs are shown.

enzymes in *A. muciniphila*. This included the sialidase activity, encoded by two secreted enzymes that recently were purified and characterized (Tailford et al., 2015).

Since it was discovered that sarkosyl can selectively solubilize cytoplasmic and inner membranes while conserving the integrity of the OM, this ionic detergent has been widely used in the purification of OM proteins in Gram-negative bacteria (Rhomberg et al., 2004; Brown et al., 2010; Moumene et al., 2015). In comparison to other methods, sarkosyl extraction leads to higher purity and better reproducibility of the OM extracts, while the basis for OM resistance to this detergent is not known (Filip et al., 1973; Frankel et al., 1991). In our study, sarkosyl-treatment was superior to the sucrose gradient centrifugation of *A. muciniphila* membranes and led to enrichment of OM proteins (**Figure 1**). Several proteins from other cellular locations were also present and the high sensitivity of LC-MS/MS might be one of the reasons why so many proteins were identified from the fractions. Hence, by optimizing the protocol, the abundance of non-OM proteins could be further reduced.

More insight into the OM proteome biosynthesis of *A. muciniphila* was generated through isolation and identification of proteins covering the most important pathways of OM biogenesis. The YaeT protein (the product of the gene with locus tag Amuc_1053), which is an essential protein for OM protein biogenesis due to its role in insertion of beta-barrel proteins into the OM (Werner and Misra, 2005; Tokuda, 2009), was among the 10 most highly abundant proteins in the extracted OM fractions. Furthermore, the organic solvent tolerance protein (encoded by the gene with the locus tag Amuc_1439), known also as increased membrane permeability (Imp) protein, needed for LPS transport to the bacterial cell surface (Bos et al., 2004), was found in high abundance. These results support the typical Gram-negative membrane architecture of *A. muciniphila*. However, we obtained no support for the presence of an intracytoplasmic membrane as is found in some PVC species (Lee et al., 2009). A recent study on the abundance of *A. muciniphila* in antibiotic-treated subjects has reported partial compartmentalization of some *A. muciniphila* cells (Dubourg et al., 2013). However, *A. muciniphila* does not contain the highly conserved genetic module, termed DUF1501, preferentially present in compartmentalized PVC species, nor

the Planctomycetes and Verrumicrobia metabolosome genes (Kamneva et al., 2012).

The *A. muciniphila* protein PilQ was found to be the most highly produced OM protein in *A. muciniphila*, regardless of the growth conditions. The primary structure of the *A. muciniphila* PilQ includes a N-terminal BAM domain with three tetratricopeptide repeats, resembling the pilotin proteins that are essential for the folding and insertion of proteins in the OM of Gram-negative bacteria and involved in the biogenesis of type IV pili (Koo et al., 2013). Moreover, the *A. muciniphila* PilQ includes a C-terminal secretin domain, which is highly conserved in all secretins that allow passage through the OM of type IV pilins, long surface exposed filaments involved in a variety of functions, including motility and adherence to host cells (Berry and Pelicic, 2015). No atomic structures of secretins have been solved yet but advanced EM analysis has recently provided insight in the *Thermus thermophilus* and *Myxococcus xhantus* pilin machineries, showing that the PilQ of these bacteria form the OM channels as well as part of the periplasmic ring where type IV pili are being extruded (Burkhardt et al., 2011; Chang et al., 2016). The structure of the *A. muciniphila* PilQ is unique but conserved in proteins predicted to be produced by other Verrucomicrobia, including *V. spinosum*, *Haloferula* sp. and *L. araneosa*. We found the *A. muciniphila* PilQ to be enriched in the sarkosyl-resistant OM fraction, which is in line with the observation that the well-characterized *T. thermophilus* PilQ is resistant to sarkosyl treatment (Burkhardt et al., 2011). The immunogold labeling results presented here confirm the localization of PilQ in the OM of *A. muciniphila*. Based on this observation, the demonstrated OM location and the predicted structural homology of PilQ, we propose that the *A. muciniphila* PilQ is involved in the formation of type IV pili. The earlier observed protruding filaments in *A. muciniphila* could represent such type IV pili but further studies are needed to confirm this (Derrien et al., 2004). Hence, PilQ is an interesting candidate for immune signaling in *A. muciniphila*, as the type IV pili are known to be involved in a variety of interactions with the host (Ehara et al., 1993; Craig and Li, 2008; von Ossowski et al., 2013).

Other groups of OM proteins found to be produced by *A. muciniphila* were autotransporters, RND multidrug efflux pumps and YD proteins. Autotransporters are an extensive family

of proteins, which can be either secreted or cell-surface-exposed, and function for example as enzymes, adhesins, cytotoxins or mediate bacterial motility (Grijpstra et al., 2013). They have mainly been studied for their virulence-related properties in pathogens but are also commonly found in non-pathogenic bacteria. RND multidrug efflux pumps exist in a tripartite form traversing both the OM and IM, and thus they are able to efficiently pump out drug molecules directly into extracellular space (Nikaido, 2011). The function of YD proteins is not very well studied but they have been suggested to be involved in carbohydrate binding and bacterial interactions with eukaryotic host cells (Jackson et al., 2009; Koskiniemi et al., 2013).

In addition to the OM proteins produced by *A. muciniphila*, we identified another 12 OM proteins from the genome using *in silico* analysis. These were not present in any of the studied fractions. Based on RNA-seq data all the genes encoding these proteins are transcriptionally active (Ottman, 2015), suggesting that their absence from the proteomics dataset could be due to a methodological reason, such as their low level production, high hydrophobicity, or post-translational modification. While we can presently not distinguish between these possibilities, we have been able to detect proteins that are post-translationally modified. These include the PEP-CTERM proteins that are involved in a specific sortase system and may be heavily post-translationally modified, including removal of N-terminal and C-terminal transmembrane domains and extensive glycosylation (Haft et al., 2012). *A. muciniphila* was predicted to encode 23 PEP-CTERM proteins and we detected nine of these by mass spectrometry. The proteins were present in relatively low abundance in the whole proteome samples, which may explain the difficulty in detecting them. Higher levels of PEP-CTERM proteins were detected in the intracellular fractions compared to OM fractions but we cannot exclude the possibility that this is related to different levels of glycosylation. While originally detected in *V. spinosum* but widely spread in the Gram-negatives, including the PVC cluster, peptides of the PEP-CTERM proteins have so far been experimentally identified in several Archaea (Haft et al., 2012). Hence, these findings suggest that *A. muciniphila* may be a well-suited model organism for studying the bacterial PEP-CTERM/exosortase system that so far has been little characterized.

Among the highly abundant proteins in the OM fractions were several proteins annotated as cytoplasmic, including six ribosomal proteins. It is possible that these proteins were so abundant in the cultures that they ended up in high concentrations also in the OM fraction, but it may also be that they have secondary roles in the bacteria. Proteins that are able to perform two or more functions are called moonlighting proteins (Copley, 2012). Several glycolytic, housekeeping and ribosomal proteins are often found on the surface of bacteria where they develop other functions (Henderson and Martin, 2011; Sherry et al., 2011).

The presence of specific OM proteins in Gram-negative bacteria is modulated by the available carbon and energy sources (Papasotiriou et al., 2008; Haussmann et al., 2009; Yang et al., 2011). The abundance of one third of the putative *A. muciniphila* OM proteins was altered during growth on glucose in comparison to mucin. However, for most of these proteins, the presence of other proteins with the same function was found in both conditions. This suggests that the specific function is not missing, but is taken over by a different protein. Unfortunately, many of the proteins that were most affected by the difference in carbon sources remain uncharacterized, despite efforts to search for related proteins. This is a common issue when handling proteomics data, especially as *A. muciniphila* belongs to the PVC superphylum, which has not been studied extensively yet. Further research into the function of these proteins could reveal important characteristics of *A. muciniphila*.

This study provides the first proteomic characterization of *A. muciniphila* OM proteins. We have identified highly abundant proteins involved in secretion and transport, as well as many uncharacterized proteins. In particular, we confirmed by immunoelectron microscopy the localization of the most abundant OM protein PilQ that is predicted to be involved in type IV pili formation by *A. muciniphila*.

AUTHOR CONTRIBUTIONS

NO, HS, CB, WdV conceived and designed the experiments. NO, LH, JR, SB, JK performed the experiments. NO, CB, WdV analyzed the data. NO and WdV drafted the manuscript and all authors contributed to the final article.

FUNDING

ACKNOWLEDGMENT

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.01157

REFERENCES

Aguilera, L., Toloza, L., Gimenez, R., Odena, A., Oliveira, E., Aguilar, J., et al. (2014). Proteomic analysis of outer membrane vesicles from the probiotic strain *Escherichia coli* Nissle 1917. *Proteomics* 14, 222–229. doi: 10.1002/pmic.201300328

Bendtsen, J. D., Kiemer, L., Fausboll, A., and Brunak, S. (2005). Non-classical protein secretion in bacteria. *BMC Microbiol.* 5:58. doi: 10.1186/1471-2180-5-58

Berry, J. L., and Pelicic, V. (2015). Exceptionally widespread nanomachines composed of type IV pilins: the prokaryotic Swiss Army knives. *FEMS Microbiol. Rev.* 39, 134–154. doi: 10.1093/femsre/fuu001

Berven, F. S., Flikka, K., Jensen, H. B., and Eidhammer, I. (2004). BOMP: a program to predict integral beta-barrel outer membrane proteins encoded within genomes of Gram-negative bacteria. *Nucleic Acids Res.* 32, W394–W399. doi: 10.1093/nar/gkh351

Bos, M. P., Tefsen, B., Geurtsen, J., and Tommassen, J. (2004). Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9417–9422. doi: 10.1073/pnas.0402340101

Brown, R. N., Romine, M. F., Schepmoes, A. A., Smith, R. D., and Lipton, M. S. (2010). Mapping the subcellular proteome of *Shewanella oneidensis* MR-1 using sarkosyl-based fractionation and LC-MS/MS protein identification. *J. Proteome Res.* 9, 4454–4463. doi: 10.1021/pr100215h

Burkhardt, J., Vonck, J., and Averhoff, B. (2011). Structure and function of PilQ, a secretin of the DNA transporter from the thermophilic bacterium *Thermus thermophilus* HB27. *J. Biol. Chem.* 286, 9977–9984. doi: 10.1074/jbc.M110.212688

Chang, Y. W., Rettberg, L. A., Treuner-Lange, A., Iwasa, J., Sogaard-Andersen, L., and Jensen, G. J. (2016). Architecture of the type IVa pilus machine. *Science* 351:aad2001. doi: 10.1126/science.aad2001

Copley, S. D. (2012). Moonlighting is mainstream: paradigm adjustment required. *Bioessays* 34, 578–588. doi: 10.1002/bies.201100191

Cox, J., Hein, M. Y., Luber, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* 13, 2513–2526. doi: 10.1074/mcp.M113.031591

Craig, L., and Li, J. (2008). Type IV pili: paradoxes in form and function. *Curr. Opin. Struct. Biol.* 18, 267–277. doi: 10.1016/j.sbi.2007.12.009

Derrien, M., Belzer, C., and de Vos, W. M. (2016). *Akkermansia muciniphila* and its role in regulating host functions. *Microb. Pathog.* doi: 10.1016/j.micpath.2016.02.005 [Epub ahead of print].

Derrien, M., Van Baarlen, P., Hooiveld, G., Norin, E., Muller, M., and de Vos, W. M. (2011). Modulation of mucosal immune response, tolerance, and proliferation in mice colonized by the mucin-degrader *Akkermansia muciniphila*. *Front. Microbiol.* 2:166. doi: 10.3389/fmicb.2011.00166

Derrien, M., Vaughan, E. E., Plugge, C. M., and de Vos, W. M. (2004). *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int. J. Syst. Evol. Microbiol.* 54, 1469–1476. doi: 10.1099/ijs.0.02873-0

Devos, D. P. (2014). PVC bacteria: variation of, but not exception to, the Gram-negative cell plan. *Trends Microbiol.* 22, 14–20. doi: 10.1016/j.tim.2013.10.008

Dubourg, G., Lagier, J. C., Armougom, F., Robert, C., Audoly, G., Papazian, L., et al. (2013). High-level colonisation of the human gut by Verrucomicrobia following broad-spectrum antibiotic treatment. *Int. J. Antimicrob. Agents* 41, 149–155. doi: 10.1016/j.ijantimicag.2012.10.012

Ehara, M., Ichinose, Y., Iwami, M., Utsunomiya, A., Shimodori, S., Kangethe, S. K., et al. (1993). Immunogenicity of *Vibrio cholerae* O1 fimbriae in animal and human cholera. *Microbiol. Immunol.* 37, 679–688. doi: 10.1111/j.1348-0421.1993.tb01692.x

Elhenawy, W., Debelyy, M. O., and Feldman, M. F. (2014). Preferential packing of acidic glycosidases and proteases into *Bacteroides* outer membrane vesicles. *MBio* 5:e00909-14. doi: 10.1128/mBio.00909-14

Everard, A., Belzer, C., Geurts, L., Ouwerkerk, J. P., Druart, C., Bindels, L. B., et al. (2013). Cross-talk between *Akkermansia muciniphila* and intestinal epithelium controls diet-induced obesity. *Proc. Natl. Acad. Sci. U.S.A.* 110, 9066–9071. doi: 10.1073/pnas.1219451110

Filip, C., Fletcher, G., Wulff, J. L., and Earhart, C. F. (1973). Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent sodium-lauryl sarcosinate. *J. Bacteriol.* 115, 717–722.

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: the protein families database. *Nucleic Acids Res.* 42, D222–D230. doi: 10.1093/nar/gkt1223

Frankel, S., Sohn, R., and Leinwand, L. (1991). The use of sarkosyl in generating soluble protein after bacterial expression. *Proc. Natl. Acad. Sci. U.S.A.* 88, 1192–1196. doi: 10.1073/pnas.88.4.1192

Galdiero, S., Falanga, A., Cantisani, M., Tarallo, R., Della Pepa, M. E., D'Oriano, V., et al. (2012). Microbe-host interactions: structure and role of Gram-negative bacterial porins. *Curr. Protein Pept. Sci.* 13, 843–854. doi: 10.2174/138920312804871120

Grijpstra, J., Arenas, J., Rutten, L., and Tommassen, J. (2013). Autotransporter secretion: varying on a theme. *Res. Microbiol.* 164, 562–582. doi: 10.1016/j.resmic.2013.03.010

Gupta, R. S., Bhandari, V., and Naushad, H. S. (2012). Molecular signatures for the PVC clade (Planctomycetes, Verrucomicrobia, Chlamydiae, and Lentisphaerae) of bacteria provide insights into their evolutionary relationships. *Front. Microbiol.* 3:327. doi: 10.3389/fmicb.2012.00327

Haft, D. H., Paulsen, I. T., Ward, N., and Selengut, J. D. (2006). Exopolysaccharide-associated protein sorting in environmental organisms: the PEP-CTERM/EpsH system. Application of a novel phylogenetic profiling heuristic. *BMC Biol.* 4:29. doi: 10.1186/1741-7007-4-29

Haft, D. H., Payne, S. H., and Selengut, J. D. (2012). Archaeosortases and exosortases are widely distributed systems linking membrane transit with posttranslational modification. *J. Bacteriol.* 194, 36–48. doi: 10.1128/JB.06026-11

Haussmann, U., Qi, S. W., Wolters, D., Rogner, M., Liu, S. J., and Poetsch, A. (2009). Physiological adaptation of Corynebacterium glutamicum to benzoate as alternative carbon source – a membrane proteome-centric view. *Proteomics* 9, 3635–3651. doi: 10.1002/pmic.200900025

Henderson, B., and Martin, A. (2011). Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease. *Infect. Immun.* 79, 3476–3491. doi: 10.1128/IAI.00179-11

Hobb, R. I., Fields, J. A., Burns, C. M., and Thompson, S. A. (2009). Evaluation of procedures for outer membrane isolation from *Campylobacter jejuni*. *Microbiology* 155, 979–988. doi: 10.1099/mic.0.024539-0

Isanapong, J., Sealy Hambright, W., Willis, A. G., Boonmee, A., Callister, S. J., Burnum, K. E., et al. (2013). Development of an ecophysiological model for *Diplosphaera* colotermitum TAV2, a termite hindgut Verrucomicrobium. *ISME J.* 7, 1803–1813. doi: 10.1038/ismej.2013.74

Jackson, A. P., Thomas, G. H., Parkhill, J., and Thomson, N. R. (2009). Evolutionary diversification of an ancient gene family (rhs) through C-terminal displacement. *BMC Genomics* 10:584. doi: 10.1186/1471-2164-10-584

Johnston, B. A., Eisen, H., and Fry, D. (1991). An evaluation of several adjuvant emulsion regimens for the production of polyclonal antisera in rabbits. *Lab. Anim. Sci.* 41, 15–21.

Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H., and Krogh, A. (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci.* 12, 1652–1662. doi: 10.1110/ps.0303703

Kamneva, O. K., Knight, S. J., Liberles, D. A., and Ward, N. L. (2012). Analysis of genome content evolution in pvc bacterial super-phylum: assessment of candidate genes associated with cellular organization and lifestyle. *Genome Biol. Evol.* 4, 1375–1390. doi: 10.1093/gbe/evs113

Kang, C. S., Ban, M., Choi, E. J., Moon, H. G., Jeon, J. S., Kim, D. K., et al. (2013). Extracellular vesicles derived from gut microbiota, especially *Akkermansia muciniphila*, protect the progression of dextran sulfate sodium-induced colitis. *PLoS ONE* 8:e76520. doi: 10.1371/journal.pone.0076520

Karlsson, C. L., Onnerfalt, J., Xu, J., Molin, G., Ahrne, S., and Thorngren-Jerneck, K. (2012). The microbiota of the gut in preschool children with normal and excessive body weight. *Obesity (Silver Spring)* 20, 2257–2261. doi: 10.1038/oby.2012.110

Kim, R., Sandler, S. J., Goldman, S., Yokota, H., Clark, A. J., and Kim, S. H. (1998). Overexpression of archaeal proteins in *Escherichia coli*. *Biotechnol. Lett.* 20, 207–210. doi: 10.1023/A:1005305330517

Koo, J., Tang, T., Harvey, H., Tammam, S., Sampaleanu, L., Burrows, L. L., et al. (2013). Functional mapping of PilF and PilQ in the *Pseudomonas aeruginosa* type IV pilus system. *Biochemistry* 52, 2914–2923. doi: 10.1021/bi3015345

Koskiniemi, S., Lamoureux, J. G., Nikolakakis, K. C., t'Kint de Roodenbeke, C., Kaplan, M. D., Low, D. A., et al. (2013). Rhs proteins from diverse bacteria mediate intercellular competition. *Proc. Natl. Acad. Sci. U.S.A.* 110, 7032–7037. doi: 10.1073/pnas.1300627110

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315

Le Marechal, C., Peton, V., Ple, C., Vroland, C., Jardin, J., Briard-Bion, V., et al. (2015). Surface proteins of *Propionibacterium freudenreichii* are involved in its anti-inflammatory properties. *J. Proteomics* 113, 447–461. doi: 10.1016/j.jprot.2014.07.018

Lee, K. C., Webb, R. I., Janssen, P. H., Sangwan, P., Romeo, T., Staley, J. T., et al. (2009). Phylum Verrucomicrobia representatives share a compartmentalized cell plan with members of bacterial phylum Planctomycetes. *BMC Microbiol.* 9:5. doi: 10.1186/1471-2180-9-5

Lu, J., Boeren, S., de Vries, S. C., van Valenberg, H. J., Vervoort, J., and Hettinga, K. (2011). Filter-aided sample preparation with dimethyl labeling to identify and quantify milk fat globule membrane proteins. *J. Proteomics* 75, 34–43. doi: 10.1016/j.jprot.2011.07.031

Miller, R. S., and Hoskins, L. C. (1981). Mucin degradation in human colon ecosystems. Fecal population densities of mucin-degrading bacteria estimated by a "most probable number" method. *Gastroenterology* 81, 759–765.

Molloy, M. P., Herbert, B. R., Slade, M. B., Rabilloud, T., Nouwens, A. S., Williams, K. L., et al. (2000). Proteomic analysis of the *Escherichia coli* outer membrane. *Eur. J. Biochem.* 267, 2871–2881. doi: 10.1046/j.1432-1327.2000.01296.x

Moumene, A., Marcelino, I., Ventosa, M., Gros, O., Lefrancois, T., Vachiery, N., et al. (2015). Proteomic profiling of the outer membrane fraction of the obligate intracellular bacterial pathogen *Ehrlichia ruminantium*. *PLoS ONE* 10:e0116758. doi: 10.1371/journal.pone.0116758

Nikaido, H. (2011). Structure and mechanism of RND-type multidrug efflux pumps. *Adv. Enzymol. Relat. Areas Mol. Biol.* 77, 1–60.

Ottman, N. (2015). *Host Immunostimulation and Substrate Utilization of the Gut Symbiont Akkermansia muciniphila.* Ph.D. dissertation, Wageningen University, Wageningen.

Papasotiriou, D. G., Markoutsa, S., Meyer, B., Papadioti, A., Karas, M., and Tsiotis, G. (2008). Comparison of the membrane subproteomes during growth of a new *Pseudomonas* strain on lysogeny broth medium, glucose, and phenol. *J. Proteome Res.* 7, 4278–4288. doi: 10.1021/pr800192n

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8, 785–786. doi: 10.1038/nmeth.1701

Png, C. W., Linden, S. K., Gilshenan, K. S., Zoetendal, E. G., McSweeney, C. S., Sly, L. I., et al. (2010). Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *Am. J. Gastroenterol.* 105, 2420–2428. doi: 10.1038/ajg.2010.281

Rajilic-Stojanovic, M., Shanahan, F., Guarner, F., and de Vos, W. M. (2013). Phylogenetic analysis of dysbiosis in ulcerative colitis during remission. *Inflamm. Bowel Dis.* 19, 481–488. doi: 10.1097/MIB.0b013e31827fec6d

Reunanen, J., Kainulainen, V., Huuskonen, L., Ottman, N., Belzer, C., Huhtinen, H., et al. (2015). *Akkermansia muciniphila* adheres to enterocytes and strengthens the integrity of epithelial cell layer. *Appl. Environ. Microbiol.* 81, 3655–3662. doi: 10.1128/AEM.04050-14

Reunanen, J., von Ossowski, I., Hendrickx, A. P. A., Palva, A., and de Vos, W. M. (2012). Characterization of the SpaCBA pilus fibers in the probiotic *Lactobacillus rhamnosus* GG. *Appl. Environ. Microbiol.* 78, 2337–2344. doi: 10.1128/Aem.07047-11

Rhomberg, T. A., Karlberg, O., Mini, T., Zimny-Arndt, U., Wickenberg, U., Rottgen, M., et al. (2004). Proteomic analysis of the sarcosine-insoluble outer membrane fraction of the bacterial pathogen *Bartonella henselae*. *Proteomics* 4, 3021–3033. doi: 10.1002/pmic.200400933

Rosendale, D. I., Blatchford, P. A., Sims, I. M., Parkar, S. G., Carnachan, S. M., Hedderley, D., et al. (2012). Characterizing kiwifruit carbohydrate utilization in vitro and its consequences for human faecal microbiota. *J. Proteome Res.* 11, 5863–5875. doi: 10.1021/pr300646m

Rupakula, A., Kruse, T., Boeren, S., Holliger, C., Smidt, H., and Maillard, J. (2013). The restricted metabolism of the obligate organohalide respiring bacterium *Dehalobacter restrictus*: lessons from tiered functional genomics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20120325. doi: 10.1098/rstb.2012.0325

Santarella-Mellwig, R., Franke, J., Jaedicke, A., Gorjanacz, M., Bauer, U., Budd, A., et al. (2010). The compartmentalized bacteria of the planctomycetes-verrucomicrobia-chlamydiae superphylum have membrane coat-like proteins. *PLoS Biol.* 8:e1000281. doi: 10.1371/journal.pbio.1000281

Selengut, J. D., Haft, D. H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W. C., et al. (2007). TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Res.* 35, D260–D264. doi: 10.1093/nar/gkl1043

Sherry, A. E., Inglis, N. F., Stevenson, A., Fraser-Pitt, D., Everest, P., Smith, D. G., et al. (2011). Characterisation of proteins extracted from the surface of *Salmonella* Typhimurium grown under SPI-2-inducing conditions by LC-ESI/MS/MS sequencing. *Proteomics* 11, 361–370. doi: 10.1002/pmic.200900802

Shin, N. R., Lee, J. C., Lee, H. Y., Kim, M. S., Whon, T. W., Lee, M. S., et al. (2014). An increase in the *Akkermansia* spp. population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. *Gut* 63, 727–735. doi: 10.1136/gutjnl-2012-303839

Speth, D. R., van Teeseling, M. C., and Jetten, M. S. (2012). Genomic analysis indicates the presence of an asymmetric bilayer outer membrane in planctomycetes and verrucomicrobia. *Front. Microbiol.* 3:304. doi: 10.3389/fmicb.2012.00304

Tailford, L. E., Owen, C. D., Walshaw, J., Crost, E. H., Hardy-Goddard, J., Le Gall, G., et al. (2015). Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. *Nat. Commun.* 6:7624. doi: 10.1038/ncomms8624

Takase, H., Nitanai, H., Hoshino, K., and Otani, T. (2000). Requirement of the *Pseudomonas aeruginosa* tonB gene for high-affinity iron acquisition and infection. *Infect. Immun.* 68, 4498–4504. doi: 10.1128/Iai.68.8.4498-4504.2000

Tokuda, H. (2009). Biogenesis of outer membranes in Gram-negative bacteria. *Biosci. Biotechnol. Biochem.* 73, 465–473. doi: 10.1271/bbb.80778

Tseng, T. T., Tyler, B. M., and Setubal, J. C. (2009). Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC Microbiol.* 9(Suppl. 1):S2. doi: 10.1186/1471-2180-9-S1-S2

Vizcaino, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., et al. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 44, D447–D456. doi: 10.1093/nar/gkv1145

von Ossowski, I., Pietila, T. E., Rintahaka, J., Nummenmaa, E., Makinen, V. M., Reunanen, J., et al. (2013). Using recombinant Lactococci as an approach to dissect the immunomodulating capacity of surface piliation in probiotic *Lactobacillus rhamnosus* GG. *PLoS ONE* 8:e64416. doi: 10.1371/journal.pone.0064416

Wagner, M., and Horn, M. (2006). The Planctomycetes, Verrucomicrobia, Chlamydiae and sister phyla comprise a superphylum with biotechnological and medical relevance. *Curr. Opin. Biotechnol.* 17, 241–249. doi: 10.1016/j.copbio.2006.05.005

Werner, J., and Misra, R. (2005). YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of *Escherichia coli*. *Mol. Microbiol.* 57, 1450–1459. doi: 10.1111/j.1365-2958.2005.04775.x

Wilson, M. M., Anderson, D. E., and Bernstein, H. D. (2015). Analysis of the outer membrane proteome and secretome of *Bacteroides fragilis* reveals a multiplicity of secretion mechanisms. *PLoS ONE* 10:e0117732. doi: 10.1371/journal.pone.0117732

Wimley, W. C. (2003). The versatile beta-barrel membrane protein. *Curr. Opin. Struct. Biol.* 13, 404–411. doi: 10.1016/S0959-440X(03)00099-X

Yang, J. N., Wang, C., Guo, C., Peng, X. X., and Li, H. (2011). Outer membrane proteome and its regulation networks in response to glucose

concentration changes in *Escherichia coli*. *Mol. Biosyst.* 7, 3087–3093. doi: 10.1039/c1mb05193h

Yu, C. S., Lin, C. J., and Hwang, J. K. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci.* 13, 1402–1406. doi: 10.1110/ps. 03479604

Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., et al. (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26, 1608–1615. doi: 10.1093/bioinformatics/btq249