MCP Papers in Press. Published on June 27, 2012 as Manuscript M112.018119

Endo-β-N-acetylglucosaminidases from infant-gut associated bifidobacteria release complex N-glycans from human milk glycoproteins.

Daniel Garrido[a,d], Charles Nwosu[b,d], Santiago Ruiz-Moyano[a,d], Danielle Aldredge[b,d], J. Bruce German[c,d], Carlito B. Lebrilla[b,d] and David A. Mills[a,c,d].

Departments of Viticulture & Enology[a], Chemistry[b], Food Science & Technology[c] and Foods for Health Institute[d], University of California Davis. 1 Shields Ave, Davis CA 95616.

Corresponding author: David A. Mills. 1 Shields Ave., Davis, CA 95616, USA. damills@ucdavis.edu. Phone: 530-754-7821; fax: 530-752-0382.

Running title: Bifidobacterial endoglycosidases active on complex N-glycans.

Abbreviations: GlcNAc: N-acetylglucosamine; Man: mannose; Fuc: fucose; GH: glycosyl hydrolase; hLF: human lactoferrin; bLF: bovine lactoferrin; IgA: immunoglobulin A; IgG: immunoglobulin G; RNaseB: ribonuclease B.

Keywords: *Bifidobacterium*, Endo-β-N-acetylglucosaminidase, N-linked glycan, human milk, glycoprotein, lactoferrin.

Summary

Breastfeeding is one of the main factors guiding the composition of the infant gut microbiota in the first months of life. This process is shaped in part by the high amounts of human milk oligosaccharides that serve as a carbon source for saccharolytic bacteria such as *Bifidobacterium* species. Infant-borne bifidobacteria have developed various molecular strategies for utilizing these oligosaccharides as a carbon source. Here we hypothesized that these species can also interact with structurally similar N-glycans found in host glycoproteins as those found in human milk and the intestinal epithelium. Endo-β-N-acetylglucosaminidases were identified in certain isolates of *Bifidobacterium longum* subsp. *longum*, *Bifidobacterium longum* subsp. *infantis* and *Bifidobacterium breve*, and their presence correlated with the ability of these strains to deglycosylate glycoproteins. An endoglycosidase from *B. infantis* ATCC 15697, EndoBI-1, was active towards all major types of N-linked glycans found in glycosylated proteins. Its activity was not affected by core fucosylation or extensive fucosylation, antenna number or sialylation, releasing several N-glycans from human lactoferrin and immunoglobulins A and G. Extensive N-deglycosylation of whole breast milk was also observed after coincubation with this enzyme. Mutation of the active site of EndoBI-1 did not abolish binding to N-glycosylated proteins, and this mutant specifically recognized $Man_3GlcNAc_2$(α1-6Fuc), the core structure of human N-glycans. EndoBI-1 is constitutively expressed in *B. infantis,* and incubation of the bacterium with human or bovine lactoferrin led to the induction of genes associated to import and consumption of human milk oligosaccharides, suggesting linked regulatory mechanisms among these glycans. This work reveals an unprecedented interaction of bifidobacteria with host N-glycans, and describes a novel endoglycosidase with broad specificity on diverse N-glycan types, potentially a useful tool for glycoproteomics studies.

Introduction

Breast milk is an intriguing and complex fluid that supports the growth, development and protection of the newborn. Its composition includes essential nutrients such as lactose, fatty acids and proteins (1), as well as a myriad of bioactive compounds critical for the protection and correct development of the infant in the first months of life (2). Breastfeeding is one of the main factors in the establishment of the intestinal microbiota in infants (3). The presence of certain species of *Bifidobacterium* is a major feature of breast-fed infants (4), and the dominance of these microorganisms is thought to be associated with beneficial health effects (5, 6). This enrichment has been explained by the ability of bifidobacteria to degrade and utilize human milk oligosaccharides (HMO) as a carbon source (7).

Proteins represent an important fraction of breast milk. A great variability exists among different proteins types and concentrations across different mothers and stages of lactation (8). Milk proteins are readily utilized by the infant (9), and are also critical in the protection of the newborn. For example, human lactoferrin (hLF) is one of the most abundant proteins in human milk, and hLF or its derived peptides display broad antimicrobial and anti-inflammatory effects, among several other biological activities (10, 11).

Virtually all secreted proteins in eukaryotes, including those in human milk, are glycosylated (12). While some milk caseins are O-linked glycosylated, lactoferrin and immunoglobulins contain N-linked glycans (12, 13). Asparagine-linked glycosylation is one of the most common post-translational modifications of eukaryotic proteins (14). In general, the role of N-linked glycosylation in folding, secretion and resistance to proteolysis is understood for several proteins (15, 16), and several examples have exemplified the crucial role of N-glycans in

protein function, such as bacterial recognition (17), intracellular signaling (18) and antigen binding and presentation (19).

Interestingly, certain microorganisms, mostly pathogens, have acquired the ability to cleave N-glycans from glycoproteins. This trait is associated with the use of these oligosaccharides as a carbon source (20), or altering the biological function of certain glycoproteins such as immunoglobulins (21). Bacterial Endo-β-N-acetylglucosaminidases (EC 3.2.1.96; endoglycosidases) are widespread enzymes that cleave the *N-N'*-diacetyl chitobiose moiety characteristic of the pentasaccharide $Man_3GlcNAc_2$ found in all N-glycans (22). These enzymes belong to glycosyl hydrolase families GH18 or GH85. Prominent examples of GH18 enzymes are EndoH from *Streptomyces plicatus* (23), EndoE from *Enterococcus faecalis* (24) and EndoS from *Streptococcus pyogenes* (25). EndoD from *Streptococcus pneumoniae* (26) is a member of GH85. Their substrate specificities are usually limited to either high mannose or complex N-glycans and some require additional exoglycosidases for glycan release. EndoH is commonly used for glycoprotein deglycosylation and determination the type of N-linked glycosylation, however EndoH substrate limitations hamper its use in comprehensive glycoproteomic analysis.

The niche that infant-gut associated bifidobacteria colonize is characterized by high amounts of milk oligosaccharides, as well as proteins or peptides arriving from breast milk or from the developing infant gut. While some bifidobacteria apparently can use mucin O-linked oligosaccharides as a carbon source (27, 28), whether these microorganisms can interact with N-glycosylated proteins has yet to be addressed. In this work we explored the ability of infant gut isolates of bifidobacteria to release N-glycans from host glycoproteins, and we also describe

some of the properties of a bifidobacterial endoglycosidase that releases N-glycans from diverse host glycoproteins.

Experimental procedures.

Bacteria and media. *Bifidobacterium* strains used in this study (Table S1) were obtained from the Japanese Collection of Microorganisms (Riken Biosource Center Japan), the American Type Culture Collection (Manassas, VA), and the University of California Davis Viticulture and Enology Culture Collection (Davis, CA). For routine experiments, bifidobacteria were grown on de Mann-Rogose-Sharp (MRS) broth supplemented with 0.05 % w/v L-cysteine (Sigma-Aldrich, St. Louis, MO). Chemically defined Zhang-Mills-Block-1 (ZMB-1) media (29) was used for evaluation of bacterial growth on glycoproteins or transcriptional analyses. Cells were anaerobically grown in a vinyl chamber (Coy Laboratory Products, Grass Lake, MI) at 37° C for 24 h, in an atmosphere consisting of 5% carbon dioxide, 5% hydrogen, and 90% nitrogen.

Chemicals. Cyanogen bromide (CNBr) activated sepharose 4B (S4B) beads, ribonuclease B from bovine pancreas (RNaseB), immunoglobulin G from human serum (IgG), immunoglobulin A from human colostrum (30), lactoferrin from human milk (hLF), lactoferrin from bovine milk (bLF) and 2,5- dihydroxylbenzoic acid (DHB) were all obtained from Sigma Aldrich (St. Louis, MO). Graphitized carbon cartridges were purchased from Grace Davison Discovery Sciences (Deerfield, IL). All chemicals used were either of analytical grade or better. Claristar yeast mannoprotein was a gift from DSM Food Specialties (Parsippany, NJ).

Incubations and growth of bifidobacteria on glycoproteins. Bifidobacterial isolates were grown on 2 ml of MRS with no carbon source (mMRS), supplemented with 2% lactose to mid-late exponential phase. Two hundred µl of culture were centrifuged for 1 min at 12000 x *g*, and resuspended in 200 µl of mMRS supplemented with 5 mg/ml of RNaseB. Incubations were run for 18 hours, and supernatants were recovered after centrifugation 1 min at 12000 x *g*. A 1:10 dilution of each supernatant was denatured in glycoprotein denaturing buffer (0.5% SDS and 40

mM DTT), analyzed on 4-15% precast SDS-PAGE gels (Bio-Rad, Carlsbad CA) and stained with Bio-safe coomassie blue (Bio-Rad). Growth of specific bacteria was also analyzed on 96 well plates containing 200 µl of ZMB-1 media supplemented with 10 mg/ml of hLF or bLF, or 50 mg/ml of Claristar yeast mannoprotein. Cultures were inoculated at 2% and grown for 72 h in a PowerWave microplate reader (BioTek Instruments, Inc., Winoosky, VT), under anaerobic conditions at 37° C. Growth was monitored using the Gen5 1.10 software (BioTek). Cultures were grown in triplicate, and controls with no glycoprotein and no bacteria were run and subtracted from OD_{600} values.

Endoglycosidase sequence determination. Protein coding sequences belonging to GH18 found in the genomes in *B. infantis* ATCC 15697 (Blon_2468), *B. infantis* 157F (BLIF_1310) and *Enterococcus faecalis* OG1RF (EndoEα) were aligned using MUSCLE. Conserved regions were selected and converted to DNA to design degenerate primers (Table S2). A similar approach was used with sequences encoding GH85 enzymes, found in the published genome sequences of *B. longum* DJO10A (BLD_0197), *B. longum* NCC2703 (BL1335) and *B. breve* UCC2003. Genomic DNA was prepared from overnight cultures on MRS for each strain used in this study using the DNeasy Blood & Tissue Kit (Qiagen, Valencia CA). Fifty µl PCR reactions contained 1 U of Phusion DNA polymerase (Finnzymes, Vantaa, Finland), 1 ng of DNA, 0.2 mM of dNTPs and 2.5 µM of each degenerate primer (Table S2), and were run in a PTC200 Thermo Cycler (MJ Research, Ramsey, MN). The PCR program included an initial denaturation at 98° C for 30 s, 30 cycles of denaturation at 98° C 10 s, annealing at 55° C for 30 s, extension at 72° C 1 min, and a final extension at 72° C for 7 min. PCR products were purified using the Qiaquick PCR product purification kit (Qiagen), and sequenced at the UC Davis DNA sequencing facility. Sequences encoding GH18 enzymes were analyzed using BioEdit 7.1.3, and later expanded and

fully determined using the DNA Walking SpeedUp Premix Kit (Seegene, Rockville MD), and the TSP142 primers listed in Table S2. GH85-encoding gene sequences were directly determined using primers GH85cF and GH85cR.

Bioinformatic analyses. The Integrated Microbial Genomes (31) database was used to find GH18 and GH85 protein sequences in *Bifidobacterium* genomes and to determine genetic landscapes for GH18-type and GH85-type genes found in the genomes of *B. infantis* ATCC 15697, *B. infantis* 157F and *B. longum* DJO10A. Multiple sequence alignments were performed using MUSCLE, using the Maximum Likelihood algorithm in MEGA v 5.0.

Gene cloning and expression. Genomic DNA from *B. infantis* ATCC 15697, *B. infantis* SC142 and *B. longum* DJO10A was amplified using the cloning primers indicated in Table S2, targeting GH18 or GH85 sequences. Signal peptides and transmembrane domains were not amplified to facilitate protein expression and purification from *E. coli*. Gene amplification by PCR, cloning, protein expression and purification were performed as (32). Induction was performed with 0.5 mM IPTG at 28° C (EndoBI-1, EndoBI-2 and EndoBI-1mut), or with 1 mM IPTG at 37° C (EndoBB). Proteins were concentrated using Amicon Ultra 30 kDa 4 ml columns, and buffer was exchanged for saline sodium citrate 1X using Bio-Gel P-30 in SSC Buffer columns (Bio-Rad).

Glycoprotein digestion by bifidobacterial endoglycosidases. Optimal enzymatic conditions for endoglycosidases EndoBI-1, EndoBI-2 and EndoBB were determined by incubation with RNaseB. Reactions were performed in a 10 µl volume and included 4 µg of RNaseB, 1 µg of each enzyme and 4 µl of 0.2 M Na_2HPO_4 with pH values between 5.0 and 7.0 at 37° C. Reactions were run for 1 h, stopped with 1 M Na_2CO_3, treated with the denaturing buffer described above and loaded into 4-15% precast polyacrylamide SDS gels. Optimal temperature for each reaction was determined at each optimal pH, and reactions were performed at 4°, 30°, 37°, 45°, 55° and

65° C for 1 h. Heat resistance was evaluated by incubating each glycosidase at 95° C for 1, 5 and 30 min, and enzyme reactions were then carried out under optimal conditions. Digestions of hLF and bLF (Sigma) were performed under optimal conditions using 4 µg of each glycoprotein and incubated for 18 h with 1 µg of each endoglycosidase, or 1 µl of glycerol-free peptide:N-glycosidase F (PNGaseF 500U/µl; New England Biolabs, Ipswich, MA). Finally 20 µl of a fresh breast milk sample (kindly provided by a lactation study directed by Jennifer Smilowitz at UC Davis) was incubated for 18 h at 37° C with 10 µg of EndoBI-1, 10 µg of EndoBI-1 D184N or 1 µl of PNGaseF in 20 mM Na_2HPO_4 pH 5.0. Lactoferrin and human milk digestions were evaluated in 7.5% precast SDS-PAGE gels under denaturing conditions. All experiments were performed at least in duplicates.

EndoBI-1 immobilization to sepharose beads. To eliminate unwanted interferences and contamination from the enzyme solution, EndoBI-1 was immobilized to sepharose beads activated with CNBr. This also allowed for multiple usage of the enzyme on different samples over a few weeks. CNBr-activated sepharose beads of 40-165 µm diameter were covalently coupled to EndoBI-1 via the well-established coupling chemistry (33). The actual immobilization of EndoBI-1 to the sepharose beads was achieved using a slightly modified version of the protocol reported earlier (34, 35). In this study 150 mg of the lyophilized S4B beads were coupled to 300 µg of EndoBI-1 prior to the glycoprotein digestion.

Glycan release by EndoBI-1. Model glycoproteins used in this study were RNaseB, bLF, hLF, IgA and IgG. Each glycoprotein was individually digested with bead-immobilized EndoBI-1 while thoroughly rinsing the beads after each digestion to eliminate cross-contamination. Glycoproteins were prepared in 0.2 M Na_2HPO_4 pH 5.0 at 1 mg/ml in a final volume of 300 µL in 1.5 ml tubes. The digestion mixture including the beads was incubated at 37° C overnight with

gentle agitation. The resultant digest mixture was then carefully drawn out following centrifugation. Purification of the resultant glycans was then achieved via solid phase extraction (36) using C18 and graphitized carbon cartridges as earlier described by our group (37). In this study, a clean mixture of the resultant glycans was then eluted with 9 ml of 0.05% trifluoroacetic acid in 40% acetonitrile (ACN) in water (v/v) followed by vacuum drying using a speed vac prior to MS analysis.

Instrumentation. Glycans purified by SPE were completely dried in a speed vac, reconstituted in 50 μL of deionized water and were directly analyzed by mass spectrometry. Glycan stock solutions (0.75 μL) were individually spotted on a stainless steel MALDI target with each spot mixed with an equal volume of DHB matrix solution made up of 0.05 mg/ml DHB in 50% ACN: 50% water. The glycan-DHB spots were then allowed to dry prior to the actual MS analyses. In this study, an IonSpec HiRes MALDI FT-ICR mass spectrometer (Lake Forest, CA) equipped with an external ion source based on a third harmonic Nd:YAG laser (355 nm) and a 7.0 Tesla actively shielded superconducting magnet, served as the platform for all the experiments described herein. Glycans were analyzed with the MALDI FT-ICR MS instrument in both the positive- and negative ion modes. Once released by the EndoBI-1 enzyme and detected via MS analysis, glycan assignment was achieved by a combination of accurate mass measurement and the use of an in-house tool (oligosaccharide calculator). The oligosaccharide calculator takes into account the various combinations of monosaccharides that match up in mass to the observed glycan mass within a predetermined error window using maltooligosaccharides. (38). This information is integrated with biological knowledge of N-linked glycosylation, such as connectivity and position of each monosaccharide. Cartoons of N-glycans represent putative structures.

Site directed mutagenesis. A plasmid containing the EndoBI-1 coding sequence (described above) was resynthesized with mutagenic primers AmpR and 2468mutF (Table S2) using the Change-IT multiple mutation site directed mutagenesis kit (USB Corporation, Santa Clara CA) and following manufacturer instructions. Mutated plasmids were cloned into Top10 competent cells (Invitrogen), and after verifying the proper mutation were transformed into BL21 competent cells. EndoBI-1 D184N was purified as described in the previous section, with induction carried on with 0.5 mM IPTG at 28° C for 6 h.

Glycan array analysis. Purified EndoBI-1 D184N (100 μg/ml, 200 μl), was analyzed for glycan binding to the Mammalian Printed Array v5.0 (provided by the Consortium for Functional Glycomics). Protocols are available at http://www.functionalglycomics.org. Detection was performed using an Anti-His-FITC antibody (Invitrogen).

***B. infantis* gene expression.** *B. infantis* cells were grown on ZMB-1 media with 2% lactose or glucose as describe above. Six ml of an exponential culture (OD_{600} 0.8-1) were centrifuged for 1 min at 12000 x *g*, and immediately resuspended in 5 ml of prewarmed ZMB-1 supplemented with either human lactoferrin or bovine lactoferrin (5 mg/ml). Cultures were rapidly returned to anaerobic conditions, and 1 ml of each culture was taken anaerobically every hour. One ml of the original culture grown on lactose or glucose (t=0), and hourly time points of incubations with bLF or hLF (t=1-3 h), were centrifuged at 12000 x *g* for 1 min, and the pellet was resuspended in 1 ml of RNAlater (Ambion, Austin, TX). The experiment was done in duplicate. Cell suspensions were stored overnight at 4° C and then at -80° C until use. RNA extraction, quality check and cDNA conversion were performed as previously described (39). Relative quantification for genes listed in table S2 was performed in a 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, CA) and using the Fast Sybr Green Master Mix (Applied

Biosystems). Reaction conditions were as recommended by the manufacturer using 0.5 μM of each primer. Primers for qPCR were designed using the NCBI primer design tool, checking for specificity across the *B. infantis* ATCC 15697 genome (Table S2).

Fluorescence assays. Binding of EndoBI-1 D184N to different glycoproteins was determined after overnight coating in microtiter 96 well plates of 20 μmoles of RNaseB, bLF, hLF or BSA in PBS buffer at room temperature. The experiment was performed in triplicates. Wells were washed with PBS three times, and blocked after incubation with BSA 3% at RT for 1h. Twenty μmol of EndoBI-1 D184N or BSA were added to the wells and incubated for 2 h at 37° C in PBS buffer adjusted to pH 5.0. Wells were washed three times with PBS-Tween 20 0.05%, and incubated for 1 h with a 1:500 dilution of FITC-Anti-His (C-term) antibody (Invitrogen). After 4 washes with PBS-Tween, fluorescence was monitored in a Synergy2 Microplate reader (Biotek), at 485/530 nm emission/excitation. In another set of experiments, breast milk samples incubated overnight with EndoBI-1, EndoBI-1 D184N or PNGaseF as performed above were coated overnight in a microtiter 96-well plate at room temperature. After washing three times with PBS buffer, wells were incubated with a 1:500 dilution of 5 mg/ml of fluorescein labeled Concavalin A (Vector labs, Burlingame CA) for 1 h at 37° C. Wells were washed four times with PBS-Tween 20 0.05%, and fluorescence was read as described above. The experiment with three replicates was repeated twice. Statistical analysis of the data was carried out by one-way analysis of variance, and the means were separated by Tukey's honest significant differences test using the SPSS software package version 10.0.6. (SPSS Inc., Chicago, Illinois, USA).

Results

Infant gut isolates of bifidobacteria display endo-N-acetylglucosaminidase activity. Bovine ribonuclease B (RNaseB) is a 17 kDa model glycoprotein that contains one glycosylation site, composed of high mannose N-linked glycans. Cleavage by endoglycosidases results in a molecule of 14 kDa (40). Overnight incubations of a panel of 76 bifidobacterial isolates (Table S1) with RNaseB suggested that endoglycosidase activity is present in only certain isolates. Incubation of *B. infantis* ATCC 15697 with 5 mg/ml of RNaseB led to a gradual deglycosylation of this glycoprotein over time (Figure 1A). In general other *B. infantis* strains degraded RNaseB weakly (Figure 1C). Only a few *B. longum* subsp. *longum (B. longum)* isolates and none of the *B. bifidum* strains examined displayed this phenotype (Figure 1B). Interestingly, certain isolates of *B. breve* such as KA179 and JCM7019, completely deglycosylated RNaseB (Figure 1D).

Distribution of endo-N-acetylglucosaminidase gene sequences in bifidobacteria. Based on known endo-N-acetylglucosaminidase sequences found in bifidobacteria, degenerate primers were used to search for GH18 or GH85 sequences in the strains used in this study. Certain isolates were found to contain a gene encoding either a GH18 or GH85 enzyme. All strains containing one of these sequences cleaved RNaseB in vitro, and conversely strains lacking these genes did not show endoglycosidase activity (Table S1). This suggested a correlation between the presence of a GH18 or GH85 enzyme and the observed RNaseB deglycosylation phenotype.

A phylogenetic tree (Figure 1E) classified these protein sequences in three types. One group was exclusively found in *B. infantis* strains including the type strain ATCC 15697 (termed GH18a), and distantly related to EndoE from *E. faecalis*. Another set of sequences found in strains of *B. infantis*, *B. breve* and *B. longum* shared 60% aminoacid identity with GH18a, and it was termed

GH18b. Sequences belonging to GH85 were almost exclusively found in fecal *B. breve* isolates. Multiple alignments revealed a high degree of conservation of the proposed active site for each glycosidase family (Figure S1; (24, 41)). The genomic landscape for these genes supported their association with carbohydrate metabolism. Blon_2468 in *B. infantis* ATCC 15697 is located in a gene cluster that in addition contains a phosphotransferase (PTS) system specific for N-acetylglucosamine (Figure S2). BLIF_1310 in *B. infantis* 157F (GH18b), and BLD_0197 in *B. longum* DJO10A (GH85) are located near ABC transporters predicted to import oligosaccharides and two or three putative α-mannosidases (Figure S2).

Enzymatic properties of bifidobacterial endo-N-acetylglucosaminidases. Based on the sequence alignments obtained (Figure 1E), a representative gene of each glycosyl hydrolase type was cloned, expressed and purified in *E. coli* (Figure S3). The recombinant endo-β-N-acetylglucosaminidases from *B. infantis* ATCC 15697 (EndoBI-1; Blon_2468), *B. infantis* SC142 (EndoBI-2), and *B. longum* DJO10A (EndoBB; BLD_0197), all exhibited maximum glycolytic activity at pH 5.0 and their optimal temperatures ranged from 30° C to 45° C (Figure S4). Interestingly, the enzymatic activity of EndoBI-1 and EndoBI-2 was not significantly impaired by incubation at 95° C for 1 or 5 minutes, implying that they are heat resistant enzymes (Figure 2A). This property was not observed for EndoBB. Other properties of these enzymes are listed in Table 1.

Human lactoferrin (hLF) contains core fucosylated complex N-glycans, predominantly at two glycosylation sites (42). Bovine lactoferrin (bLF) represents a minor fraction of bovine milk, and it contains mainly oligomannose N-linked glycans at five glycosylation sites (35). Overnight incubations of bLF and hLF with the three recombinant *Bifidobacterium* endoglycosidases indicated that EndoBI-1 and EndoBI-2 were able to cleave bLF and also hLF, as observed by

discrete changes in MW on SDS-PAGE gels (Figures 2B). EndoBB did not display glycolytic activity against bLF or hLF (Figure 2B).

EndoBI-1 cleaves the chitobiose core of high-mannose and complex N-glycans. In order to better determine the enzymatic properties of EndoBI-1, the enzyme was immobilized using sepharose beads and incubated with several glycoproteins with varying glycosylation types. Glycans released were analyzed using MALDI FT-ICR mass spectrometry. Similarly to related endoglycosidases, EndoBI-1 acted on the chitobiose core of N-glycans, probably leaving a GlcNAc residue (and an α1-6 fucose in core fucosylated glycans) attached to the glycosylated asparagine residue. Activity on bLF was detected and predominantly oligomannose N-glycans were released, with a minor amount of complex/hybrid glycans (Figure 3A; (43)), consistent with the general N-glycosylation of the protein. N-glycans released from RNaseB contained between 5 and 9 mannose residues (Figure S5; (40)). EndoBI-1 was also shown to deglycosylate glycoproteins such as hLF, IgA and IgG (Figures 3B, 3C and Figure S5). These proteins are characterized by complex core fucosylated glycans, and are essentially resistant to several commercial endoglycosidases under native conditions. Glycans released from hLF and IgA were bi and triantennary, and contained up to two sialic acid residues and up to three fucoses attached to the lactosamine chains. The profile of the glycans released by EndoBI-1 was similar to PNGaseF (10, 42). IgG deglycosylation by EndoBI-1 revealed biantennary complex N-glycans with lesser fucosylation and sialylation compared to IgA, and similar to those observed after PNGaseF cleavage (Figure S5, (44, 45)).

A mutant of EndoBI-1 binds the core of N-linked glycans. The conserved active site of GH18 enzymes includes the motif D-X-E where both Asp and Glu are crucial for activity (46). Asp184 in EndoBI-1 was replaced by site-directed mutagenesis to Asn184 (EndoBI-1 D184N). The

mutated enzyme specifically bound to the core of N-glycans, $Man_3GlcNAc_2$, across 600 glycans in a mammalian glycan array (Figure 4A). Interestingly, EndoBI-1 D184N also showed significant binding to the α1-6 fucosylated pentasaccharide, characteristic of human N-linked glycoproteins. When equimolar amounts of RNaseB, bLF and hLF were coated to microwell plates, EndoBI-1 D184N showed a significant binding to these proteins compared to a non-glycosylated control (Figure 4B).

EndoBI-1 has glycosidase activity on human milk glycoproteins. Breast milk is a complex fluid, characterized by diverse amounts of N-linked, O-linked and non-glycosylated proteins. Overnight incubation of a fresh human milk sample with EndoBI-1 or PNGaseF produced a shift in the molecular weight of mainly one protein, probably lactoferrin as deduced from its MW (Figure 5A). No change was observed when a breast milk sample was incubated with EndoBI-1 D184N. In a parallel experiment, the total amount of N-linked glycans, estimated as the amount of α-mannose detected by the lectin Concavalin A conjugated to FITC (ConA-FITC), was determined in digested milk samples. EndoBI-1 and PNGaseF, but not EndoBI D184N, significantly decreased the amount of α-mannose in breast milk (Figure 5B), suggesting an extensive removal of N-linked glycans.

Growth of bifidobacteria on N-glycosylated proteins. Vigorous growth of bifidobacteria was observed when 5% of yeast mannoprotein, heavily N-glycosylated cell wall proteins purified from *Saccharomyces cerevisiae*, were used as the sole carbon source (Figure S6). A higher OD_{600} was obtained under these conditions for *B. breve* SC139 and *B. breve* KA179, strains that showed high endoglycosidase activity against RNaseB (Figure 1). Weak growth was observed on hLF or bLF as the sole carbon source was tested *in vitro* (data not shown).

Impact of hLF and bLF on *B. infantis* gene expression. The molecular response of *B. infantis* ATCC 15697 to bLF and hLF was tested incubating the microorganism with 5 mg/ml of bLF and hLF in a resting cell assay. This revealed an increased expression of Blon_2468 (EndoBI-1) by comparison to cells grown on glucose, and the level of expression of this gene was similar to that from cells grown on lactose (Figure S7). Incubations with bLF or hLF resulted in higher expression of other genes adjacent to Blon_2468 including Blon_2470 and Blon_2471, encoding part of a PTS system specific for GlcNAc, (Figure 6A). A similar trend was observed for Blon_0177 and Blon_0178, genes also associated to PTS systems in *B. infantis*. Other genes induced by these glycoproteins were Blon_0881 and to a lesser extent Blon_0882, key enzymes that participate in metabolism of GlcNAc and sialic acid. Putative genes in *B. infantis* associated to mannose metabolism (Blon_2380, solute binding protein for manno-oligosaccharides, and Blon_0868 and Blon_0869, α-mannosidases) were not affected by the presence of bLF or hLF. Conversely, several genes associated to the import and consumption of human milk oligosaccharides in *B. infantis* were significantly induced by hLF, and to a lesser extent bLF (Figure 6B). In general the highest induction was observed after 1 hour of incubation. These genes included Blon_2344, Blon_2347, Blon_0883 and Blon_2177, solute-binding proteins that bind different classes of HMO associated to ABC transporters, as well as Blon_2335 and Blon_2336, two key fucosidases in the *B. infantis* genome (32, 39).

Discussion

Human milk contains a constellation of factors that shape the microbial content of the infant gastrointestinal tract (47). They include a range of bioactive glycans (either as free HMOs or as conjugates bound to proteins or lipids) that simultaneously enrich beneficial commensals within, and deflect pathogens from, the intestine. The mechanism by which free HMOs enrich infant-borne bifidobacteria has been characterized at the molecular level (7) (48), however relatively little work has explored the impact of human milk glycoproteins on the developing infant gut microbiota. A prebiotic character has been previously suggested for N-linked human milk lactoferrin (49, 50), as well as O-linked substrates such as bovine glycomacropeptide (51), a peptide derived from κ-casein. While bifidobacterial endoglycosidases that release O-linked glycans from mucins have been described (28, 52), it is not clear if this is valid for O-linked glycans found in human milk.

To examine if N-linked glycoproteins serve as potential growth substrates endoglycosidase activity was tested across a large panel of bifidobacterial isolates obtained from infant feces. Only certain isolates of *B. infantis* (40%), *B. longum* (21%) and *B. breve* (36%) exhibited this activity, with the latter species showing the highest glycolytic activity on RNaseB. Genes encoding endo-β-N-acetylglucosaminidases belonging to GH18a, GH18b or GH85 were found only in isolates with endoglycosidase activity, suggesting that the presence of these genes explain the observed phenotype. In this study we determined that *Bifidobacterium* strains do not extensively use host glycoproteins as the sole carbon source (by comparison to certain pathogens described below) under the conditions assayed. However, several strains possessing an endoglycosidase showed a significant growth on pure yeast mannoproteins (Figure S6), which consists of high-mannose heavily N-glycosylated cell wall proteins from yeast. This suggested

that bifidobacteria could utilize N-glycans as the sole carbon source under certain conditions. Moreover, the broad range of activity on complex N-linked glycans exhibited by EndoBI-1 or EndoBI-2 may endow bifidobacteria with a competitive advantage by enabling cleavage and consumption or a diverse range of N-linked glycans.

The ability to release N-glycans from host proteins has been mainly associated to GH18 enzymes found in bacterial pathogens such as *E. faecalis* (53), *S. pyogenes (54)* and *Capnocytophaga canimorsus* (20). These bacteria can extensively grow on different glycoproteins as a carbon source, and as in the case of EndoS from *S. pyogenes*, IgG-specific deglycosylation severely impairs its recognition by immune effectors, increasing bacterial survival in blood (25).

It is possible that bifidobacterial endoglycosidases could also modulate of the activity of host glycoproteins. An increasing amount of evidence suggests a crucial role for N-glycans in the function of several host proteins (19, 55, 56). For example, recognition of Gram-positive bacteria by IgA is dependent on its glycosylation (17), and intracellular signaling and NF-kB activation of the toll-like receptor 3 (18) is modulated by N-glycans. C-type lectins, galectins and sialic-acid-binding Ig-like lectins are immune and cell response mediators that specifically recognize different epitopes in N-glycans (22). While lactoferrin N-linked glycosylation is variable during lactation (10, 12), studies about the impact of glycosylation of this protein with respect to its resistance to proteolysis and iron binding (57-60) are contradictory. If certain *Bifidobacterium* isolates have the ability to remove N-glycans from lactoferrin, destabilization of the protein could favor to the production of antimicrobial peptides such as lactoferricin B or lactoferrampin. Interestingly, some studies have suggested that lactoferrin has a bifidogenic effect (49, 61). While in the present work specific enzymes in bifidobacteria have been determined to

deglycosylate human lactoferrin, more studies are needed to address closer the impact of lactoferrin or its derived peptides on these microorganisms.

GH18 and GH85 endoglycosidases specifically cleave the *N*-*N*'-diacetylchitobiose core of N-linked glycans. Here we studied some of the enzymatic properties of EndoBI-1 and EndoBI-2, representatives of two clades of GH18 sequences found in bifidobacteria (Figure 1E). While their aminoacid sequences were only 60% identical and possessed different gene contexts (Figure S2), their active sites were conserved, and both acted on bLF and hLF, containing oligomannose and complex N-glycans, respectively. Further description of the N-glycans released from EndoBI-1 by mass spectrometry indicated that it was also active on IgA and IgG. The enzyme did not recognize O-linked glycans or HMO (data not shown). In general the specificity of most known endoglycosidases is limited to high mannose glycans (for example EndoH (23)). EndoS from *S. pyogenes* acts solely on IgG (25), and the affinities of EndoEα from *E. faecalis* for other proteins than RNaseB have not been further studied (24). Endoglycosidases F1, F2 and F3 from *Elizabethkingia miricola* show a preference for either high mannose or complex oligosaccharides (62). By comparison, several features suggested that EndoBI-1 might prove useful as a novel tool for diverse applications such as proteomics and glycoproteomics. Firstly, EndoBI-1 is heat stable as incubation at 95° C for 5 minutes did not severely impact its activity. Moreover, the glycolytic activity found on a varied range of target glycoproteins revealed that this enzyme cleaved the chitobiose core of high mannose N-glycans (RNaseB and bLF), and core α1-6 fucosylated, bi or triantennary complex N-glycans with up to two sialic acid and up to three fucose residues decorating the lactosamine chains (Figure 3 and S4). More precise kinetic studies are required for determining the impact of these modifications on enzyme activity. These analyses were done using native glycoproteins, and a much greater

deglycosylation rate after denaturation is expected. Moreover, EndoBI-1 was active directly on human milk—a complex matrix of lipids, oligosaccharides and proteins with disparate glycosylation types—successfully removing a significant proportion of the total amount of N-glycans (Figure 5). Finally, EndoBI-1 will be useful in applications such as detection and characterization of glycosylated regions of proteins under non-denaturing conditions, and also an improved determination of glycosylation sites of core and non-core fucosylated N-glycans, after chitobiose cleavage and tripsinization.

Further evidence for the affinity of EndoBI-1 for N-linked glycans was determined in a parallel set of experiments. An active site mutant of the enzyme (EndoBI-1 D184N) lacked catalytic activity but it retained the ability to bind to $Man_3GlcNAc_2$(α1-6Fuc) in a glycan array. While this suggests that this mutation does not significantly alter the binding pocket of this enzyme, further structural studies are required to understand the mechanism of action of EndoBI-1 and the wide range of N-linked glycans it recognizes.

We also partially characterized EndoBB (BLD_0197) from *B. longum* DJO10A, representative of GH85 sequences found in infant gut bifidobacteria. The activity of this enzyme was much more limited, cleaving RNaseB but not bLF or hLF. Bifidobacterial GH85 enzymes are distantly related to EndoD from *S. pneumoniae* (26). EndoD acts on complex core fucosylated N-glycans, but only when lactosamine chains have been trimmed by exoglycosidases (26). It is possible that *B. breve* GH85 endoglycosidases collaborates with additional glycosyl hydrolases. The presence of α-mannosidases and an ABC importer for oligosaccharides adjacent to these genes suggests that their function is related. It is also possible that these clusters are also active on mannose-based oligosaccharides from plant origin.

Strains of *B. infantis* have been studied by their remarkable ability to use HMO as the only carbon source (63). Genes induced by HMO in *B. infantis*, such as solute binding proteins and α-fucosidases (32, 39), were also up-regulated by hLF and bLF (Figure 6). These results suggested that, while not extensively using these glycoproteins as a carbon source (at least under the conditions tested in this study), *B. infantis* was still able to respond to these substrates in a similar fashion as to HMO.

In conclusion, in this work we described the interaction of infant-gut associated bifidobacteria with N-linked glycans found in host glycoproteins such as those found in breast milk, and we determined the discrete molecular determinants associated with this interaction. Finally, we have characterized the enzymatic properties of EndoBI-1 from *B. infantis,* which showed a remarkable activity on a wide range of host N-linked glycans.

Acknowledgments

The authors would like to thank Dr. Jennifer Smilowitz and Lan Guan for providing reagents for this study, Dr. Alberto Martin Gonzalez for his support in statistical analysis and finally Dr. David F. Smith and Dr. Jamie Heimburg-Molinaro from the Consortium for Functional Glycomics Core H for their glycan array analysis. We thank DSM for the gift of mannoproteins used in this study. Daniel Garrido was sponsored in part through a Fulbright-Conicyt Chile scholarship, and Santiago Ruiz-Moyano was supported by the Ministry of Education and Science of Spain and University of Extremadura, Spain. This work was supported by the University of California Discovery Grant Program, the California Dairy Research Foundation and National Institute of Health Award R01HD061923.

References

1. Neville, M. C. (2009) Classic studies of mammary development and milk secretion: 1945 - 1980. *J Mammary Gland Biol Neoplasia* 14, 193-197.
2. German, J. B., Dillard, C. J., and Ward, R. E. (2002) Bioactive components in milk. *Curr Opin Clin Nutr Metab Care* 5, 653-658.
3. Adlerberth, I., and Wold, A. E. (2009) Establishment of the gut microbiota in Western infants. *Acta Paediatr* 98, 229-238.
4. Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R. N., Anokhin, A. P., Heath, A. C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J. G., Lozupone, C. A., Lauber, C., Clemente, J. C., Knights, D., Knight, R., and Gordon, J. I. (2012) Human gut microbiome viewed across age and geography. *Nature* 486, 222-227.
5. Le Huerou-Luron, I., Blat, S., and Boudry, G. (2010) Breast- v. formula-feeding: impacts on the digestive tract and immediate and long-term health effects. *Nutr Res Rev* 23, 23-36.
6. Conroy, M. E., Shi, H. N., and Walker, W. A. (2009) The long-term health effects of neonatal microbial flora. *Curr Opin Allergy Clin Immunol* 9, 197-201.
7. Garrido, D., Barile, D., and Mills, D. A. (2012) A molecular basis for bifidobacterial enrichment in the infant gastrointestinal tract. *Adv Nutr* 3, 415S-421S.
8. Mitoulas, L. R., Kent, J. C., Cox, D. B., Owens, R. A., Sherriff, J. L., and Hartmann, P. E. (2002) Variation in fat, lactose and protein in human milk over 24 h and throughout the first year of lactation. *Br J Nutr* 88, 29-37.
9. Prentice, A., Ewing, G., Roberts, S. B., Lucas, A., MacCarthy, A., Jarjou, L. M., and Whitehead, R. G. (1987) The nutritional role of breast-milk IgA and lactoferrin. *Acta Paediatr Scand* 76, 592-598.
10. Barboza, M., Pinzon, J., Wickramasinghe, S., Froehlich, J. W., Moeller, I., Smilowitz, J. T., Ruhaak, L. R., Huang, J., Lonnerdal, B., German, J. B., Medrano, J. F., Weimer, B. C., and Lebrilla, C. B. (2012) Glycosylation of human milk lactoferrin exhibits dynamic changes during early lactation enhancing its role in pathogenic bacteria-host interactions. *Mol Cell Proteomics* 11, M111 015248.
11. Gonzalez-Chavez, S. A., Arevalo-Gallegos, S., and Rascon-Cruz, Q. (2009) Lactoferrin: structure, function and applications. *Int J Antimicrob Agents* 33, 301 e301-308.
12. Froehlich, J. W., Dodds, E. D., Barboza, M., McJimpsey, E. L., Seipert, R. R., Francis, J., An, H. J., Freeman, S., German, J. B., and Lebrilla, C. B. (2010) Glycoprotein expression in human milk during lactation. *J Agric Food Chem* 58, 6440-6448.
13. Picariello, G., Ferranti, P., Mamone, G., Roepstorff, P., and Addeo, F. (2008) Identification of N-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass spectrometry. *Proteomics* 8, 3833-3847.
14. Apweiler, R., Hermjakob, H., and Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* 1473, 4-8.
15. Weber, A. N., Morse, M. A., and Gay, N. J. (2004) Four N-linked glycosylation sites in human toll-like receptor 2 cooperate to direct efficient biosynthesis and secretion. *J Biol Chem* 279, 34589-34594.
16. Roth, J., Zuber, C., Park, S., Jang, I., Lee, Y., Kysela, K. G., Le Fourn, V., Santimaria, R., Guhl, B., and Cho, J. W. (2010) Protein N-glycosylation, protein folding, and protein quality control. *Mol Cells* 30, 497-506.
17. Mathias, A., and Corthesy, B. (2011) Recognition of gram-positive intestinal bacteria by hybridoma- and colostrum-derived secretory immunoglobulin A is mediated by carbohydrates. *J Biol Chem* 286, 17239-17247.
18. Sun, J., Duffy, K. E., Ranjith-Kumar, C. T., Xiong, J., Lamb, R. J., Santos, J., Masarapu, H., Cunningham, M., Holzenburg, A., Sarisky, R. T., Mbow, M. L., and Kao, C. (2006) Structural and functional analyses of the human Toll-like receptor 3. Role of glycosylation. *J Biol Chem* 281, 11144-11151.
19. Ryan, S. O., Bonomo, J. A., Zhao, F., and Cobb, B. A. (2011) MHCII glycosylation modulates Bacteroides fragilis carbohydrate antigen presentation. *J Exp Med* 208, 1041-1053.

20. Renzi, F., Manfredi, P., Mally, M., Moes, S., Jeno, P., and Cornelis, G. R. (2011) The N-glycan glycoprotein deglycosylation complex (Gpd) from Capnocytophaga canimorsus deglycosylates human IgG. *PLoS Pathog* 7, e1002118.
21. Collin, M., Shannon, O., and Bjorck, L. (2008) IgG glycan hydrolysis by a bacterial enzyme as a therapy against autoimmune conditions. *Proc Natl Acad Sci U S A* 105, 4265-4270.
22. Varki, A. (2009) *Essentials of glycobiology*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
23. Trimble, R. B., and Maley, F. (1977) The use of endo-beta-N-acetylglucosaminidase H in characterizing the structure and function of glycoproteins. *Biochem Biophys Res Commun* 78, 935-944.
24. Collin, M., and Fischetti, V. A. (2004) A novel secreted endoglycosidase from Enterococcus faecalis with activity on human immunoglobulin G and ribonuclease B. *J Biol Chem* 279, 22558-22570.
25. Allhorn, M., Olin, A. I., Nimmerjahn, F., and Collin, M. (2008) Human IgG/Fc gamma R interactions are modulated by streptococcal IgG glycan hydrolysis. *PLoS One* 3, e1413.
26. Muramatsu, H., Tachikui, H., Ushida, H., Song, X., Qiu, Y., Yamamoto, S., and Muramatsu, T. (2001) Molecular cloning and expression of endo-beta-N-acetylglucosaminidase D, which acts on the core structure of complex type asparagine-linked oligosaccharides. *J Biochem* 129, 923-928.
27. Ruas-Madiedo, P., Gueimonde, M., Fernandez-Garcia, M., de los Reyes-Gavilan, C. G., and Margolles, A. (2008) Mucin degradation by Bifidobacterium strains isolated from the human intestinal microbiota. *Appl Environ Microbiol* 74, 1936-1940.
28. Fujita, K., Oura, F., Nagamine, N., Katayama, T., Hiratake, J., Sakata, K., Kumagai, H., and Yamamoto, K. (2005) Identification and molecular cloning of a novel glycoside hydrolase family of core 1 type O-glycan-specific endo-alpha-N-acetylgalactosaminidase from Bifidobacterium longum. *J Biol Chem* 280, 37415-37422.
29. Zhang, G., Mills, D. A., and Block, D. E. (2009) Development of chemically defined media supporting high-cell-density growth of lactococci, enterococci, and streptococci. *Appl Environ Microbiol* 75, 1080-1087.
30. Vincent, F., Davies, G. J., and Brannigan, J. A. (2005) Structure and kinetics of a monomeric glucosamine 6-phosphate deaminase: missing link of the NagB superfamily? *J Biol Chem* 280, 19649-19655.
31. Markowitz, V. M., Korzeniewski, F., Palaniappan, K., Szeto, E., Werner, G., Padki, A., Zhao, X., Dubchak, I., Hugenholtz, P., Anderson, I., Lykidis, A., Mavromatis, K., Ivanova, N., and Kyrpides, N. C. (2006) The integrated microbial genomes (IMG) system. *Nucleic Acids Res* 34, D344-348.
32. Sela, D. A., Garrido, D., Lerno, L., Wu, S., Tan, K., Eom, H. J., Joachimiak, A., Lebrilla, C. B., and Mills, D. A. (2012) Bifidobacterium longum subsp. infantis ATCC 15697 alpha-Fucosidases Are Active on Fucosylated Human Milk Oligosaccharides. *Appl Environ Microbiol* 78, 795-803.
33. Deguchi, K., Ito, H., Baba, T., Hirabayashi, A., Nakagawa, H., Fumoto, M., Hinou, H., and Nishimura, S. (2007) Structural analysis of O-glycopeptides employing negative- and positive-ion multi-stage mass spectra obtained by collision-induced and electron-capture dissociations in linear ion trap time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 21, 691-698.
34. Nwosu, C. C., Strum, J. S., An, H. J., and Lebrilla, C. B. (2010) Enhanced detection and identification of glycopeptides in negative ion mode mass spectrometry. *Anal Chem* 82, 9654-9662.
35. Nwosu, C. C., Seipert, R. R., Strum, J. S., Hua, S. S., An, H. J., Zivkovic, A. M., German, B. J., and Lebrilla, C. B. (2011) Simultaneous and extensive site-specific N- and O-glycosylation analysis in protein mixtures. *J Proteome Res* 10, 2612-2624.
36. Rentz, D. M., Locascio, J. J., Becker, J. A., Moran, E. K., Eng, E., Buckner, R. L., Sperling, R. A., and Johnson, K. A. (2010) Cognition, reserve, and amyloid deposition in normal aging. *Ann Neurol* 67, 353-364.
37. An, H. J., Peavy, T. R., Hedrick, J. L., and Lebrilla, C. B. (2003) Determination of N-glycosylation sites and site heterogeneity in glycoproteins. *Anal Chem* 75, 5628-5637.

38. Clowers, B. H., Dodds, E. D., Seipert, R. R., and Lebrilla, C. B. (2008) Dual polarity accurate mass calibration for electrospray ionization and matrix-assisted laser desorption/ionization mass spectrometry using maltooligosaccharides. *Anal Biochem* 381, 205-213.
39. Garrido, D., Kim, J. H., German, J. B., Raybould, H. E., and Mills, D. A. (2011) Oligosaccharide binding proteins from Bifidobacterium longum subsp. infantis reveal a preference for host glycans. *PLoS One* 6, e17315.
40. Roberts, G., Tarelli, E., Homer, K. A., Philpott-Howard, J., and Beighton, D. (2000) Production of an endo-beta-N-acetylglucosaminidase activity mediates growth of Enterococcus faecalis on a high-mannose-type glycoprotein. *J Bacteriol* 182, 882-890.
41. Yamamoto, S., Muramatsu, H., and Muramatsu, T. (2005) Mutational studies on endo-β-N-acetylglucosaminidase D which hydrolyzes core portion of asparagine-linked complex type oligosaccharides. *Glycoconjugate Journal* 22, 35-42.
42. Yu, T., Guo, C., Wang, J., Hao, P., Sui, S., Chen, X., Zhang, R., Wang, P., Yu, G., Zhang, L., Dai, Y., and Li, N. (2011) Comprehensive characterization of the site-specific N-glycosylation of wild-type and recombinant human lactoferrin expressed in the milk of transgenic cloned cattle. *Glycobiology* 21, 206-224.
43. Hua, S., Nwosu, C. C., Strum, J. S., Seipert, R. R., An, H. J., Zivkovic, A. M., German, J. B., and Lebrilla, C. B. (2011) Site-specific protein glycosylation analysis with glycan isomer differentiation. *Anal Bioanal Chem*.
44. Raju, T. S., Briggs, J. B., Borge, S. M., and Jones, A. J. (2000) Species-specific variation in glycosylation of IgG: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics. *Glycobiology* 10, 477-486.
45. Wuhrer, M., Stam, J. C., van de Geijn, F. E., Koeleman, C. A., Verrips, C. T., Dolhain, R. J., Hokke, C. H., and Deelder, A. M. (2007) Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum. *Proteomics* 7, 4070-4081.
46. Synstad, B., Gaseidnes, S., Van Aalten, D. M., Vriend, G., Nielsen, J. E., and Eijsink, V. G. (2004) Mutational and computational analysis of the role of conserved residues in the active site of a family 18 chitinase. *Eur J Biochem* 271, 253-262.
47. Newburg, D. S. (2009) Neonatal protection by an innate immune system of human milk consisting of oligosaccharides and glycans. *J Anim Sci* 87, 26-34.
48. Garrido, D., Ruiz-Moyano, S., and Mills, D. A. (2012) Release and utilization of N-acetyl-d-glucosamine from human milk oligosaccharides by Bifidobacterium longum subsp. infantis. *Anaerobe*.
49. Rahman, M. M., Kim, W. S., Ito, T., Kumura, H., and Shimazaki, K. (2009) Growth promotion and cell binding ability of bovine lactoferrin to Bifidobacterium longum. *Anaerobe* 15, 133-137.
50. Kim, W. S., Ohashi, M., Tanaka, T., Kumura, H., Kim, G. Y., Kwon, I. K., Goh, J. S., and Shimazaki, K. (2004) Growth-promoting effects of lactoferrin on L. acidophilus and Bifidobacterium spp. *Biometals* 17, 279-283.
51. Janer, C., Pel$\sqrt{}^{\circ}$ ez, C., and Requena, T. (2004) Caseinomacropeptide and whey protein concentrate enhance Bifidobacterium lactis growth in milk. *Food Chemistry* 86, 263-267.
52. Kiyohara, M., Nakatomi, T., Kurihara, S., Fushinobu, S., Suzuki, H., Tanaka, T., Shoda, S., Kitaoka, M., Katayama, T., Yamamoto, K., and Ashida, H. (2012) alpha-N-acetylgalactosaminidase from infant-associated bifidobacteria belonging to novel glycoside hydrolase family 129 is implicated in alternative mucin degradation pathway. *J Biol Chem* 287, 693-700.
53. Roberts, G., Homer, K. A., Tarelli, E., Philpott-Howard, J., Devriese, L. A., and Beighton, D. (2001) Distribution of endo-beta-N-acetylglucosaminidase amongst enterococci. *J Med Microbiol* 50, 620-626.
54. Collin, M., and Olsen, A. (2001) EndoS, a novel secreted protein from Streptococcus pyogenes with endoglycosidase activity on human IgG. *EMBO J* 20, 3046-3055.
55. Almogren, A., Bonner, A., Perkins, S. J., and Kerr, M. A. (2009) Functional and structural characterisation of human colostrum free secretory component. *Mol Immunol* 46, 1534-1541.

56. Vagin, O., Tokhtaeva, E., Yakubov, I., Shevchenko, E., and Sachs, G. (2008) Inverse correlation between the extent of N-glycan branching and intercellular adhesion in epithelia. Contribution of the Na,K-ATPase beta1 subunit. *J Biol Chem* 283, 2192-2202.
57. Davidson, L. A., and Lonnerdal, B. (1988) Specific binding of lactoferrin to brush-border membrane: ontogeny and effect of glycan chain. *Am J Physiol* 254, G580-585.
58. van Berkel, P. H., Geerts, M. E., van Veen, H. A., Kooiman, P. M., Pieper, F. R., de Boer, H. A., and Nuijens, J. H. (1995) Glycosylated and unglycosylated human lactoferrins both bind iron and show identical affinities towards human lysozyme and bacterial lipopolysaccharide, but differ in their susceptibilities towards tryptic proteolysis. *Biochem J* 312 (Pt 1), 107-114.
59. van Veen, H. A., Geerts, M. E., van Berkel, P. H., and Nuijens, J. H. (2004) The role of N-linked glycosylation in the protection of human and bovine lactoferrin against tryptic proteolysis. *Eur J Biochem* 271, 678-684.
60. Brines, R. D., and Brock, J. H. (1983) The effect of trypsin and chymotrypsin on the in vitro antimicrobial and iron-binding properties of lactoferrin in human milk and bovine colostrum. Unusual resistance of human apolactoferrin to proteolytic digestion. *Biochim Biophys Acta* 759, 229-235.
61. Petschow, B. W., Talbott, R. D., and Batema, R. P. (1999) Ability of lactoferrin to promote the growth of Bifidobacterium spp. in vitro is independent of receptor binding capacity and iron saturation level. *J Med Microbiol* 48, 541-549.
62. Trimble, R. B., and Tarentino, A. L. (1991) Identification of distinct endoglycosidase (endo) activities in Flavobacterium meningosepticum: endo F1, endo F2, and endo F3. Endo F1 and endo H hydrolyze only high mannose and hybrid glycans. *J Biol Chem* 266, 1646-1651.
63. Locascio, R. G., Ninonuevo, M. R., Kronewitter, S. R., Freeman, S. L., German, J. B., Lebrilla, C. B., and Mills, D. A. (2009) A versatile and scalable strategy for glycoprofiling bifidobacterial consumption of human milk oligosaccharides. *Microb Biotechnol* 2, 333-342.

Figure Legends

Figure 1: Endoglycosidase activity found in *Bifidobacterium* isolates. A: Deglycosylation of RNaseB by *B. infantis* ATCC 15697 over time. Overnight incubation with RNaseB was evaluated for other isolates of *B. longum* (B), *B. infantis* (C) or *B. breve* (D). E: Phylogenetic representation of endoglycosidase sequences found in bifidobacterial isolates.

Figure 2: Properties of the recombinant endoglycosidases in bifidobacteria. A: Heat tolerance of EndoBI-1, EndoBI-2 or EndoBB evaluated in SDS-PAGE gels, as evaluated by RNaseB deglycosylation with each enzyme incubated at 95° C for the times indicated. B: Coincubations of bLF (upper panel) and hLF (lower panel) with EndoBI-1 (1), EndoBI-2 (2), EndoBB (3) or PNGaseF (4). Control (C) non-digested reactions were included in both experiments.

Figure 3: MALDI-FT-ICR MS analysis of glycans released after coincubation of EndoBI-1 with A: bLF; B: hLF (positive mode); C: IgA (positive mode). Green circles: mannose; blue squares: GlcNAc; red triangles: fucose; yellow circles: galactose; pink diamonds: NeuAc.

Figure 4: Properties of EndoBI-1 D184N. A: Glycan array analysis of the mutant enzyme binding to mammalian glycans (x-axis). Bars represent SD of sextuplicates. Structures bound are provided, and legend is as in Figure 3. B: Binding of EndoBI-1 D184N to coated glycoproteins, as detected by a FITC-Anti His antibody. Error bars represent SD from triplicate experiments, and values of BSA binding to glycoproteins were withdrawn from EndoBI-1 D184N. Asterisks represent samples with $p < 0.05$ compared to BSA.

Figure 5: EndoBI-1 activity in breast milk. A: SDS-PAGE gel of overnight incubation of human milk (lane 1, control) with EndoBI-1 (lane 2), EndoBI-1 D184N (lane 3) or PNGaseF (lane 4). Protein identities were deduced from (12). B: Amount of N-glycosylation (proportional to α-

mannose) in samples from A. Error bars represent SD from triplicate experiments. Asterisks represent samples with $p < 0.05$ compared to control.

Figure 6: Fold changes in gene expression for *B. infantis* ATCC 15697 genes during time coincubation with bLF or hLF, as indicated in the figure legend. Locus tags are described in the text. Error bars represent SD from three biological replicates. A: Genes associated to GlcNAc metabolism and located close to EndoBI-1; B: Genes previously described to be associated or induced by HMO.

Table 1: General properties of the endo-β-N-acetylglucosaminidases described in this study

	EndoBI-1	**EndoBI-2**	**EndoBB**
Family	GH18	GH18	GH85
Calculated MW (recombinant protein)	47 kDa	53 kDa	98 kDa
Transmembrane domains	2	2	1
Optimum pH	5.0	5.0	5.0
Optimum temperature	30-45° C	30-45° C	30-45° C
Heat resistance	Yes	Yes	No

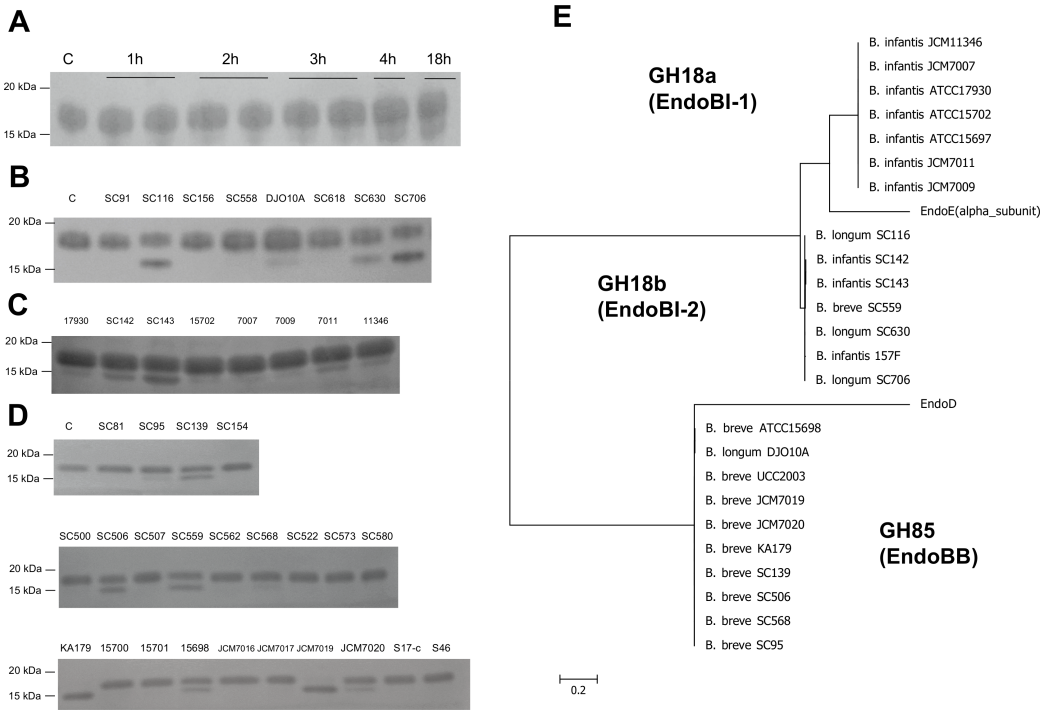
A
C
1h
2h
3h
4h
18h
20 kDa
15 kDa
B
C
SC91
SC116
SC156
SC558
DJO10A
SC618
SC630
SC706
20 kDa
15 kDa
C
17930
SC142
SC143
15702
7007
7009
7011
11346
20 kDa
15 kDa
D
C
SC81
SC95
SC139
SC154
20 kDa
15 kDa
SC500
SC506
SC507
SC559
SC562
SC568
SC522
SC573
SC580
20 kDa
15 kDa
KA179
15700
15701
15698
JCM7016
JCM7017
JCM7019
JCM7020
S17-c
S46
20 kDa
15 kDa
E
GH18a
(EndoBI-1)
B. infantis JCM11346
B. infantis JCM7007
B. infantis ATCC17930
B. infantis ATCC15702
B. infantis ATCC15697
B. infantis JCM7011
B. infantis JCM7009
EndoE(alpha_subunit)
GH18b
(EndoBI-2)
B. longum SC116
B. infantis SC142
B. infantis SC143
B. breve SC559
B. longum SC630
B. infantis 157F
B. longum SC706
EndoD
B. breve ATCC15698
B. longum DJO10A
B. breve UCC2003
B. breve JCM7019
B. breve JCM7020
B. breve KA179
B. breve SC139
B. breve SC506
B. breve SC568
B. breve SC95
GH85
(EndoBB)
0.2

FIGURE 1

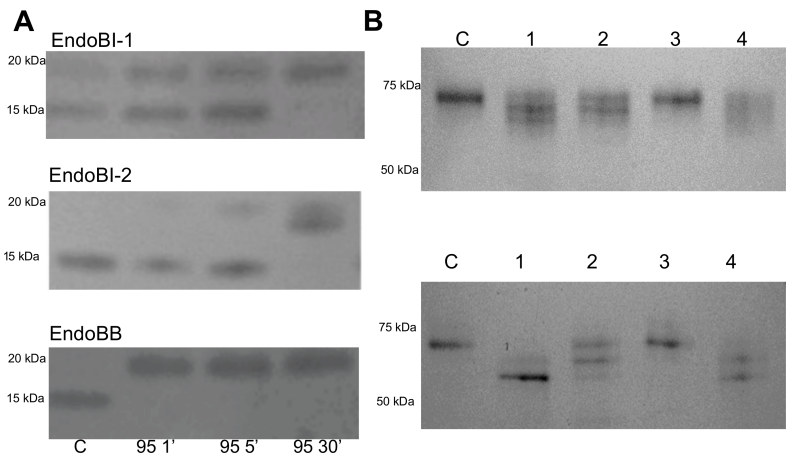
A
EndoBI-1
20 kDa
15 kDa
EndoBI-2
20 kDa
15 kDa
EndoBB
20 kDa
15 kDa
C
95 1'
95 5'
95 30'
B
C
1
2
3
4
75 kDa
50 kDa
C
1
2
3
4
75 kDa
50 kDa

FIGURE 2

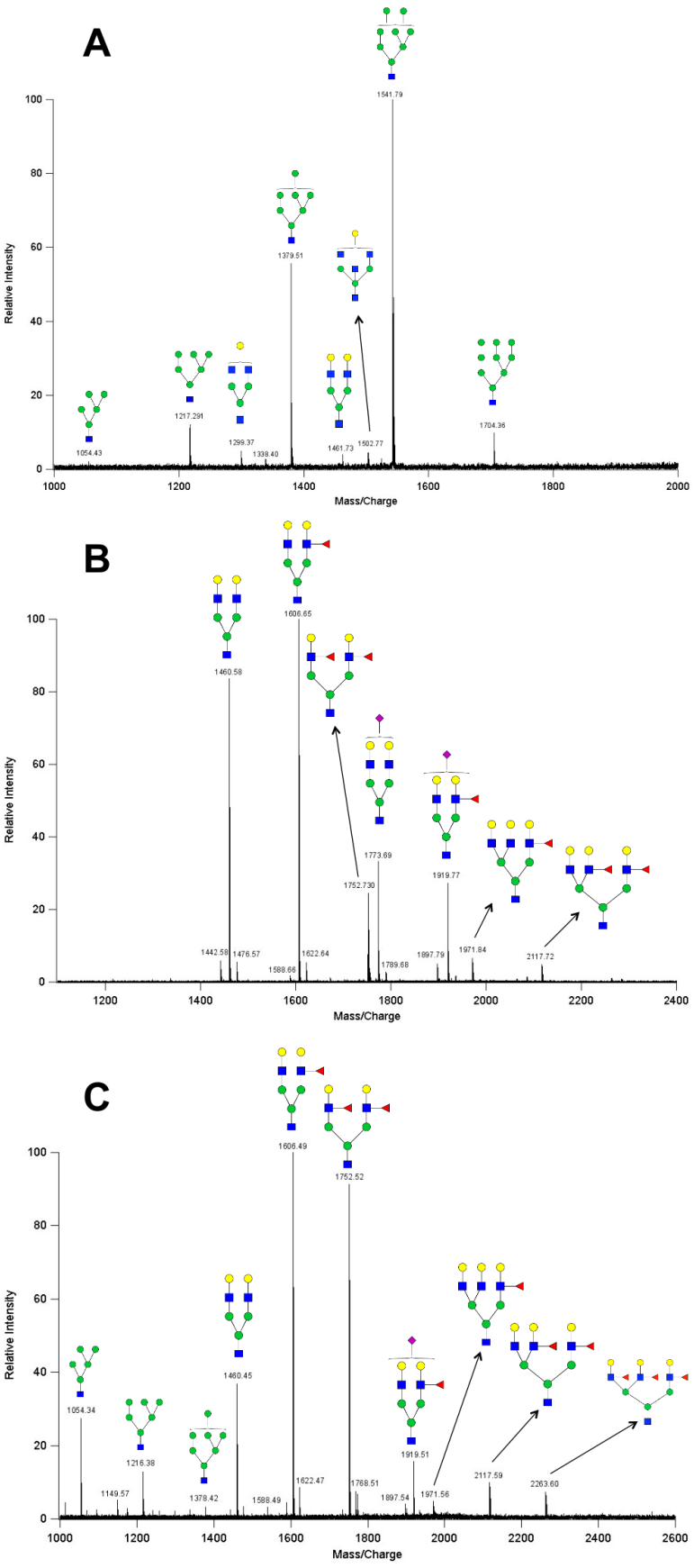
A
Relative Intensity
Mass/Charge
1541.79
1379.51
1217.291
1299.37
1338.40
1461.73
1502.77
1704.36
1054.43
B
Relative Intensity
Mass/Charge
1606.65
1460.58
1773.69
1752.730
1919.77
1971.84
2117.72
1442.58
1476.57
1622.64
1588.66
1789.68
1897.79
C
Relative Intensity
Mass/Charge
1606.49
1752.52
1460.45
1054.34
1216.38
1149.57
1378.42
1588.49
1622.47
1768.51
1897.54
1919.51
1971.56
2117.59
2263.60

FIGURE 3

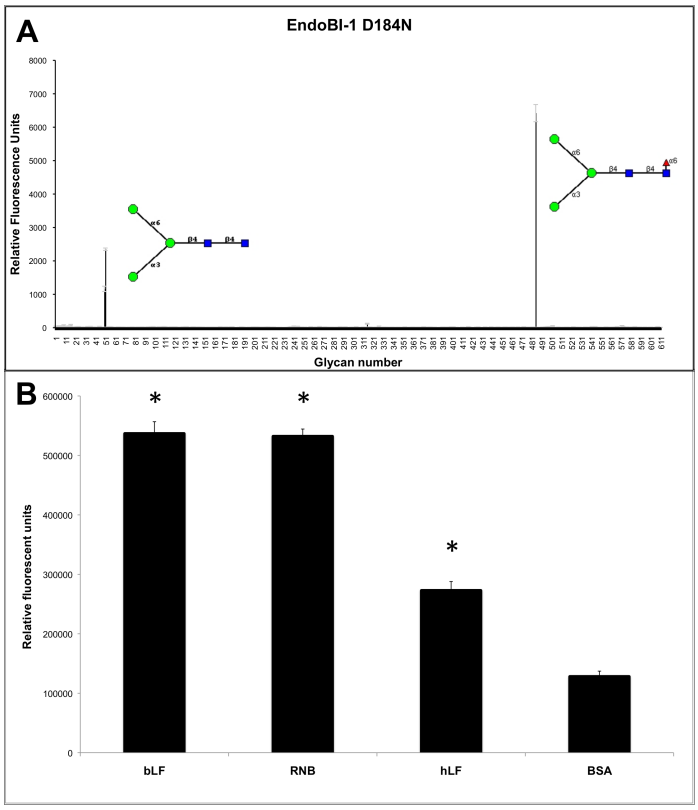
A
EndoBI-1 D184N
Relative Fluorescence Units
Glycan number
B
Relative fluorescent units
bLF
RNB
hLF
BSA

FIGURE 4

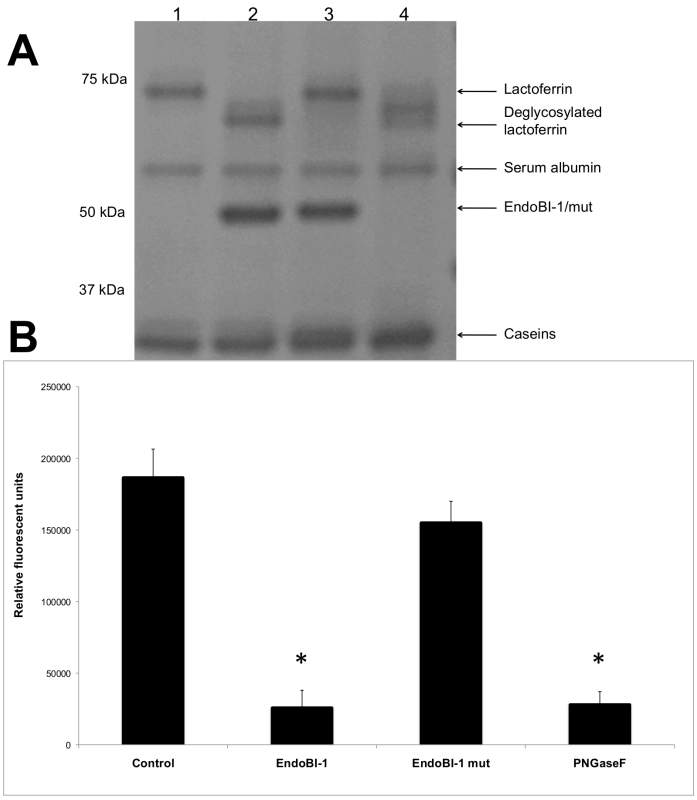
A
1
2
3
4
75 kDa
50 kDa
37 kDa
Lactoferrin
Deglycosylated lactoferrin
Serum albumin
EndoBI-1/mut
Caseins
B
Relative fluorescent units
250000
200000
150000
100000
50000
0
*
*
Control
EndoBI-1
EndoBI-1 mut
PNGaseF

FIGURE 5

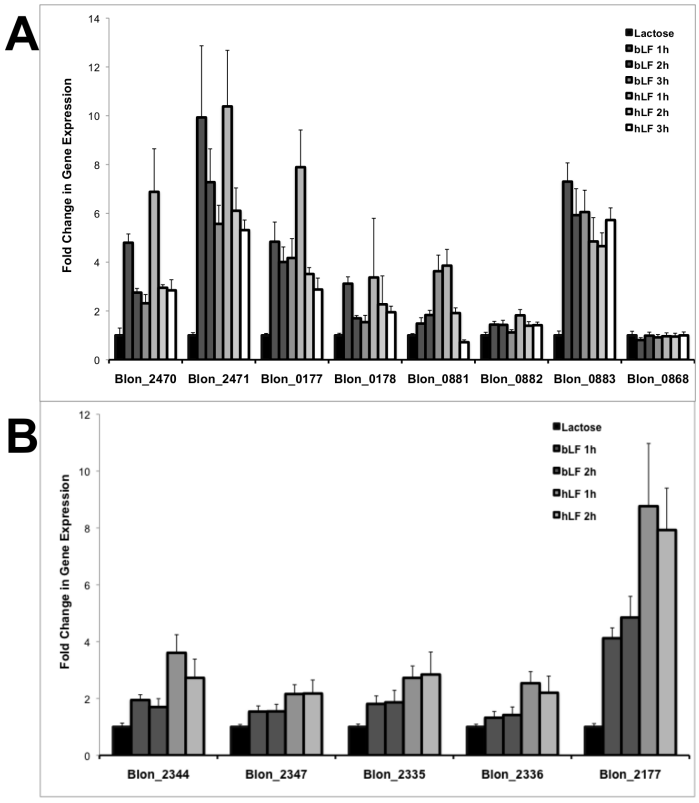
A
Fold Change in Gene Expression
14
12
10
8
6
4
2
0
Lactose
bLF 1h
bLF 2h
bLF 3h
hLF 1h
hLF 2h
hLF 3h
Blon_2470
Blon_2471
Blon_0177
Blon_0178
Blon_0881
Blon_0882
Blon_0883
Blon_0868
B
Fold Change in Gene Expression
12
10
8
6
4
2
0
Lactose
bLF 1h
bLF 2h
hLF 1h
hLF 2h
Blon_2344
Blon_2347
Blon_2335
Blon_2336
Blon_2177

FIGURE 6