

# Characteristic Changes in Cell Surface Glycosylation Accompany Intestinal Epithelial Cell (IEC) Differentiation: High Mannose Structures Dominate the Cell Surface Glycome of Undifferentiated Enterocytes\*s

Dayoung Park‡, Kristin A. Brune‡, Anupam Mitra§, Alina I. Marusina§, © Emanual Maverakis§, and Carlito B. Lebrilla‡¶

Changes in cell surface glycosylation occur during the development and differentiation of cells and have been widely correlated with the progression of several diseases. Because of their structural diversity and sensitivity to intra- and extracellular conditions, glycans are an indispensable tool for analyzing cellular transformations. Glycans present on the surface of intestinal epithelial cells (IEC) mediate interactions with billions of native microorganisms, which continuously populate the mammalian gut. A distinct feature of IECs is that they differentiate as they migrate upwards from the crypt base to the villus tip. In this study, nano-LC/ESI QTOF MS profiling was used to characterize the changes in glycosylation that correspond to Caco-2 cell differentiation. As Caco-2 cells differentiate to form a brush border membrane, a decrease in high mannose type glycans and a concurrent increase in fucosylated and sialylated complex/hybrid type glycans were observed. At day 21, when cells appear to be completely differentiated, remodeling of the cell surface glycome ceases. Differential expression of glycans during IEC maturation appears to play a key functional role in regulating the membrane-associated hydrolases and contributes to the mucosal surface innate defense mechanisms. Developing methodologies to rapidly identify changes in IEC surface glycans may lead to a rapid screening approach for a variety of disease states affecting the GI Molecular & Cellular Proteomics 14: 10.1074/mcp. tract. M115.053983, 2910-2921, 2015.

Proliferative stem cells located in the base of intestinal crypts form specialized differentiated cell types as they migrate up the villi. A continuous cell turnover occurs every four to eight days as newly differentiated cells eventually replace older cells at the tip of the villus. Self-renewing intestinal epithelial cells (IECs)<sup>1</sup> are highly susceptible to malignant growths, which arise from imbalances in cellular proliferation, differentiation, and apoptosis. If the number of developing cells outbalances the number of mature cells undergoing apoptosis, an abnormal growth of tissue can form which, in some cases, may lead to malignant tumors. Thus, a greater understanding of the molecular details of IEC differentiation may lead to novel insight into the pathophysiology of a variety of GI diseases, including cancer.

IECs are known to have highly glycosylated surfaces (1-3). The distribution of these displayed glycans across the cell surface is sensitive to the internal state as well as the external environment of the cell (4-6). In particular, variations in glycosylation patterns have been reported to occur as a function of cellular development and cancer progression (7-10). For example, global changes in cell surface sialylation have been observed during kidney endothelial, uterine epithelial, and lymphoid cell maturation (11-15). Similar variations in glycosylation have also been observed during malignant tumor progression, where the cells undergo rapid proliferation and resistance to apoptosis. In this context, recent studies in colorectal cancer tissues have demonstrated the expression of higher levels of high mannose type glycans and bisecting N-acetylglucosamine (GlcNAc) compared with normal colorectal tissues (16, 17). Although the mechanisms that control these specific changes are not well understood, it is evident

From the ‡Department of Chemistry, University of California, Davis, California 95616; §Department of Dermatology, University of California, Davis School of Medicine, Sacramento, California 95816

Received July 23, 2015, and in revised form, August 21, 2015 Published, MCP Papers in Press, September 9, 2015, DOI 10.1074/ mcp.M115.053983

Author contributions: D.P., E.M., and C.B.L. designed research; D.P., K.A.B., A.M., and A.I.M. performed research; D.P., E.M., and C.B.L. wrote the paper; D.P. analyzed glycomics and proteomics data; A.M. and A.I.M. analyzed transcriptomics data.

<sup>&</sup>lt;sup>1</sup> The abbreviations used are: IEC, Intestinal Epithelial cell; HM, High Mannose; C, Complex; H, Hybrid; Fuc, Fucose; NeuAc, *N*-Acetyl-neuraminic Acid; Man, Mannose; GlcNAc, *N*-Acetylglucosamine; ALP, Alkaline Phosphatase; *p*-NPP, *p*-Nitrophenyl Phosphate; TCC, Total Compound Chromatogram; ECC, Extracted Compound Chromatogram; SEM, Scanning Electron Microscope.

that glycosylation is critical for these cellular transformations. Determining the surface glycosylation changes of IECs in relation to their differentiation may lead to a better understanding of this malignancy.

Among the various intestinal epithelial cell types, absorptive enterocytes are the most abundant (18). Differentiation of enterocytes results in the formation of a brush border on their apical surface, a structure that controls membrane-associated hydrolases and nutrient transport. The human cell line Caco-2 has been widely accepted and used as an in vitro model for absorptive intestinal epithelial cells since its establishment in 1974 (19, 20). A characteristic feature of Caco-2 cells is their spontaneous enterocyte-like differentiation in culture after cells reach confluence (21). Although proliferation and differentiation of Caco-2 has been studied extensively, including quantitative proteomic and transcriptomic analyses (22-26), the associated changes in glycosylation that accompany Caco-2 cell differentiation have yet to be comprehensively characterized. More specifically, there is only modest knowledge about its cell surface glycome, despite the importance of the plasma membrane in many key biological functions. Earlier glycosylation targeted studies have focused on changes in glycosyltransferase activity and mRNA levels. Briefly, upon differentiation, increased activity was observed for GlcNAc transferase II and V, which are involved in N-glycosylation (27), and for  $\beta$ -3-galactosyltransferase,  $\alpha$ -2-fucosyltransferase, sialyltransferase, and  $\beta$ -6-GlcNAc transferase, which are relevant to O-glycan biosynthesis (28). Additionally, differentiation-dependent changes in mRNA expression were observed for α-2,6-sialyltransferase (29). A more global glycan analysis was performed by lectin array profiling of the surfaces of Caco-2 cells, demonstrating that lectins which recognize branched fucose and  $\alpha$ -2,6-sialic acid were effective at Caco-2 cell binding (30, 31). Although these studies provide qualitative indication of the presence of carbohydrate motifs on the cell surface, the complete composition or the relative amounts of individual structures cannot be distinguished. Furthermore, these methods do not provide information about the underlying protein scaffold. Precise identification of glycan compositions with structural detail and additional glycoproteomic analysis is necessary to adequately monitor changes in glycosylation patterns associated with cell differentiation.

Recent advancements in mass spectrometry have overcome the limitations inherent to earlier glycan profiling methodologies (32–44). Herein we employ an MS-based analytical approach utilizing nano-LC separation with high resolution TOF MS for accurate detection of compounds, enabling rapid identification and quantitation of N-glycan alterations during Caco-2 cell differentiation. With microchip nano-LC separation, isomeric forms of glycans were distinguished by their retention times. High resolution TOF MS analysis provides accurate mass measurements and consequently, detailed and selective assignment of over 200 glycan compounds from a single injection (45). Using membrane enrichment methods compatible with mass spectrometry (46), we have targeted our analysis to the cell membrane compartment to identify the specific glycan features that accompany Caco-2 cell differentiation. Additionally, the corresponding membrane-localized proteins, from which glycans were released, were identified. Finally, select glycosylation-related mRNA expression levels were quantified during differentiation. Monitoring changes in specific structures is important to identify disease-associated deviations from normal processes of intestinal cell growth and differentiation, which will allow for therapeutic interventions to be initiated earlier.

### EXPERIMENTAL PROCEDURES

*Cell Culture*—Caco-2 cells were obtained from American Type Culture Collection (ATCC, Manassas, VA) and grown in Eagle's Minimum Essential Medium (EMEM) supplemented with nonessential amino acids, 2 mM L-glutamine, 10% (v/v) fetal bovine serum (Life Technologies, Grand Island, NY), 1 mM sodium pyruvate, 1.5 g/l sodium bicarbonate, 100 U/ml penicillin, and 100  $\mu$ g/ml streptomycin. Cells were subcultured at 80% confluency and maintained at 37 °C in a humidified incubator with 5% CO<sub>2</sub>. On days 5, 7, 14, 21, and 24, cells were collected in biological triplicates by scraping. For consistency, ~2 × 10<sup>6</sup> cells were used per replicate for analysis.

*Electron Microscopy*—Cells were grown on 13 mm round coverslips and fixed in 2% paraformaldehyde and 2.5% glutaraldehyde (pH 7.2) in 0.1 M phosphate buffer at 4 °C on days 5, 7, 10, 14, 18, 21, and 24. Samples were dehydrated in a critical point dryer (Tousimis, Rockville, MD) and coated with gold using a sputter coater (Ted Pella Inc., Redding, CA). Images were taken on a FEI XL30 electron microscope (FEI, Hillsboro, OR).

Alkaline Phosphatase Activity—Alkaline phosphatase (ALP) activity was measured in cell homogenates using a colorimetric assay as per the manufacturer's instructions (Abcam, Cambridge, MA). Cells lysed at different growth times were incubated with 50  $\mu$ M *p*-nitrophenyl phosphate (*p*-NPP) for 60 min. When the substrate is dephosphorylated, *p*-nitrophenyl is produced, which turns the solution to a visible yellow color. Absorbance readings were taken at 405 nm with an electronic microplate reader (Fisher Scientific, Pittsburgh, PA). Enzyme activities were determined with a standard curve generated using intestinal ALP enzyme (Abcam). ALP activity was measured in mU, where a Unit is defined as the amount of enzyme to hydrolyze 1  $\mu$ mol of *p*-NPP per minute at 37 °C.

*Membrane Extraction*—Details of the isolation of the cell membrane fraction have been described previously (47). In brief, after harvesting cell pellets were resuspended in homogenization buffer containing 0.25 M sucrose, 20 mM HEPES-KOH (pH 7.4), and 1:100 protease inhibitor mixture (EMD Milipore, Billerica, MA). Cells were lysed on ice using a probe sonicator (Qsonica, Newtown, CT) and lysates were pelleted by centrifugation at 2000  $\times$  *g* for 10 min to remove the nuclear fraction and unlysed cells followed by a series of ultracentrifugation steps at 200,000  $\times$  *g* for 45 min to remove other nonmembrane subcellular fractions.

*N-Glycan Release and Enrichment*—Membrane pellets were suspended with 100  $\mu$ l of 100 mM NH<sub>4</sub>HCO<sub>3</sub> in 5 mM dithiothreitol and heated for 10 s at 100 °C to thermally denature the proteins. To release the glycans, 2  $\mu$ l of peptide N-glycosidase F (New England Biolabs, MA) was added to the samples, which were then incubated at 60 °C in a microwave reactor (CEM Corporation, Matthews, NC) for 10 min at 20 watts. After addition of 400  $\mu$ l of ice-cold ethanol, samples were frozen for 1 h at -80 °C to precipitate deglycosylated

proteins and centrifuged for 20 min at 15,000 rpm. The supernatant containing N-glycans was collected and dried.

Released N-glycans were purified by solid-phased extraction using porous graphitized carbon packed cartridges (Grace, Deerfield, IL). Cartridges were first equilibrated with alternating washes of nanopure water and a solution of 80% (v/v) acetonitrile and 0.05% (v/v) trifluoroacetic acid in water. Samples were loaded onto the cartridge and washed with nanopure water at a flow rate of 1 ml/min to remove salts and buffer. N-Glycans were eluted with a solution of 40% (v/v) acetonitrile and 0.05% (v/v) trifluoroacetic acid in water and dried.

Nano-LC/MS and LC/MS/MS Analysis—Glycan samples were reconstituted in nanopure water and analyzed using an Agilent nano-LC/ESI QTOF MS system (Agilent, Santa Clara, CA). Samples are introduced to the MS with a microfluidic chip, which consists of enrichment and analytical columns packed with porous graphitized carbon and a nanoelectrospray tip. The programmed 32-min run applies a binary gradient to separate and elute glycans at a flow rate of 0.4 µl/min: (A) 3% (v/v) acetonitrile and 0.1% (v/v) formic acid in water and (B) 90% (v/v) acetonitrile in 1% (v/v) formic acid in water. MS spectra were acquired at 1.5 s per spectrum over a mass range of m/z 500–2000 in positive ionization mode. Mass inaccuracies were corrected with reference masses m/z 622.029, 922.010, 1221.991, and 1521.971.

Collision-induced dissociation (CID) was performed with nitrogen gas using a series of collision energies ( $V_{collision}$ ) dependent on the m/z values of the N-glycans, based on the equation:

$$V_{\text{collision}} = \text{slope}(m/z) + \text{offset}$$

where the slope and offset were set at (1.8/100 Da) V and -2.4 V, respectively.

N-Glycan compositions were identified with an in-house retrosynthetic library according to accurate mass (48). This library was constructed based on knowledge of the mammalian N-glycan biosynthetic pathway. Quantitative reproducibility and tandem MS confirmation of library searches were previously validated, enabling rapid and accurate assignment of glycan compositions (36, 45, 48, 49). Signals above a signal-to-noise ratio of 5.0 were filtered and deconvoluted using MassHunter Qualitative Analysis B.03.01 software (Agilent). Deconvoluted masses were compared with theoretical masses using a mass tolerance of 20 ppm and a false discovery rate of 0.6%. Relative abundances were determined by integrating peak areas for observed glycan masses and normalizing to the summed peak areas of all glycans detected. Statistical evaluation of glycan abundances were performed using a two-tailed, unpaired Student's t test.

Protein Enrichment—Extracted membrane proteins were resolubilized in 8  $\,$ M urea. After reduction of proteins with dithiothreitol and alkylation with iodoacetamide, samples were diluted to 1  $\,$ M urea and incubated with trypsin at 37 °C overnight. The resulting peptides were enriched using C18 packed pipette tips (Agilent). Columns were washed and conditioned with acetonitrile followed by 0.1% (v/v) trifluoroacetic acid in water. Samples were introduced to the column, washed with 0.1% (v/v) trifluoroacetic acid in water, and eluted with a solution of 80% (v/v) acetonitrile and 0.1% (v/v) trifluoroacetic acid in water. Peptides were concentrated *in vacuo* prior to mass spectrometric analysis.

Proteomic LC-MS/MS Analysis—Membrane peptides were reconstituted in 2% (v/v) acetonitrile and 0.1% (v/v) trifluoroacetic acid in water and separated using a reverse-phase Michrom Magic C18AQ column (200  $\mu$ m, 150 mm) coupled with a Q Exactive Plus mass spectrometer through a Proxeon nano-spray source (Thermo Scientific, Fremont, CA). A 90-min binary gradient was applied at a flow rate of 2  $\mu$ l/min with (A) 0.1% (v/v) formic acid in water and (B) 100% acetonitrile: 0–75 min, 5–35% (B); 75–82 min, 35–80% (B); 82–84 min, 80% (B); 84–85 min, 80–5% (B). Per acquisition, the instrument was run in a data-dependent mode as follows: spray voltage, 2.2 kV; ion transfer capillary temperature, 200 °C; full scan mass range, *m/z* 350–1600; MS automatic gain control,  $1 \times 10^6$ ; MS maximum injection time, 30 ms; MS/MS automatic gain control,  $5 \times 10^4$ ; MS/MS maximum injection time, 50 ms; precursor resolution, 70,000; product ion resolution, 17,500; precursor ion isolation width of *m/z* 1.6; normalized collision energy, 27.

Raw data were exported using xCalibur, version 2.0 (Thermo Scientific). Proteins were identified from the tandem mass spectra using X! Tandem (2015.04.01) against the Ensembl v76 human protein database (86137 entries) (50) with the following parameters: mass tolerances of 10 ppm for the precursor and 20 ppm for fragment ions; peptide probability >0.95 using PeptideProphet (51); carbamidomethylation of cysteine as a fixed modification; oxidation of methionine and tryptophan and deamidation of asparagine and glutamine as variable modifications; two missed cleavage sites; at least two unique peptides. Protein identifications were filtered at a 1% false discovery rate and were accepted if E-values were less than or equal to 0.02. Additional protein information was retrieved through functional annotation analysis using the PANTHER database (52). N-glycosylation sites were predicted using the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc).

Real-Time PCR for Relevant Genes of Glycosylation Pathway-Undifferentiated, partially differentiated, and fully differentiated Caco-2 cells (2  $\times$  10<sup>6</sup> cells per preparation) were scraped on day 5, 14, and 21 respectively and washed twice with PBS  $(1 \times)$ , followed by resuspension in RNAlater (Life Technologies). Briefly, total RNA were extracted using RNeasy plus mini kit (Qiagen, Valencia, CA) and the quantity and quality of RNA were determined by using Qubit Fluorometer (Life Technologies) and TapeStation 2200 (Agilent) following manufacturer's protocol. Total RNA were reverse transcribed to cDNA using iScript Reverse Transcription Supermix (Bio-Rad, Hercules, CA) following manufacturer's instructions. The predesigned human glycosylation PrimePCR plates (Bio-Rad) were used for real-time PCR using the CFX96 Touch Real-Time PCR detection system (Bio-Rad) and the analysis was performed using CFX Manager 3.1 software (Bio-Rad). Each 96-well plate has 88 target genes relevant to glycosylation pathway and 3 reference genes (GAPDH, TBP, and HPRT1). The gene expression was normalized to reference genes and presented as fold changes.

#### RESULTS

Caco-2 cells were grown in standard culture conditions and analyzed at different growth time points during the process of their spontaneous enterocyte-like differentiation. Undifferentiated cells proliferated to confluency at day 7, after which cells accelerate their differentiation. Five time points were chosen and are referred to as follows: day 5, undifferentiated; day 7, confluent; day 14, partially differentiated; day 21, fully differentiated; day 24, post-differentiated.

*Brush Border Formation*—Differentiation of Caco-2 cells was evaluated colorimetrically for alkaline phosphatase (ALP) activity, a marker indicative of the presence of an established brush border. ALP activity levels were barely detectable in the early phases of proliferation (day 5) but continued to increase after cells reached confluence (day 7) (Fig. 1*A*). By day 18, the enzyme activity increased more than 11-fold relative to the measured activity at day 5.

In addition to molecular validation of Caco-2 differentiation, electron micrographs were taken of the apical surface to FIG. 1. Biochemical and morphological validation of Caco-2 differentiation. *A*, Alkaline phosphatase (ALP) activity during Caco-2 cell growth. Standard deviation is indicated by error bars (n = 3). *B*, SEM images of the apical side of the Caco-2 cell monolayer at different stages of growth. Microvilli were fully formed by day 21.



observe the growth-dependent formation of a brush border membrane. Scanning electron microscope (SEM) images of Caco-2 cells fixed at different growth times confirm the presence of microvilli, which become more numerous and elongated after cells reach confluence (Fig. 1B). By day 21, the entire cell layer was covered with microvilli. Thus for our analysis, cells are considered undifferentiated until day 7 and fully differentiated at day 21.

Progressive Cell Surface Glycomic Changes during Enterocytic Differentiation—To characterize the progressive modifications in glycosylation during enterocytic differentiation, Caco-2 cell surface glycans were analyzed at days 5, 7, 14, 21, and 24 after seeding. Three replicate samples from each day were isolated from the same passage. LC-MS spectral searching using an in-house library identified an average of more than 200 glycan compounds in the cell membrane preparations on the specified days. Validation of library searches were performed previously (36, 45, 48, 49), enabling rapid assignment based on accurate mass. Relative intensities of each compound were determined using the area under each peak, as described under Experimental Procedures.

From the sum of all signals, glycans were grouped into subtypes, which include high mannose (HM), nondecorated complex/hybrid (C/H), fucosylated C/H, and sialylated C/H. Their distribution on the Caco-2 cell surface underwent significant changes as cells differentiated. The most abundant glycans on undifferentiated cell surfaces were high mannose type structures, collectively summing to about 30% of the total signal. In contrast, differentiated cells were dominated by fucosylated and sialylated glycans, comprising respectively around 83 and 75% of the total number of glycans found on the cell surface. Along with the increases in the abundances of complex/hybrid type glycans, the percent of total high mannose type glycans on differentiated cells decreased to ~7%. The transition from high mannose to complex/hybrid rich cell surfaces is shown as a function of time in Fig. 2. Notably, the rate of change in the abundances was most dramatic before cells reached confluence and in the early phases of proliferation (days 5–7). After cells reached confluence, levels of high mannose and decorated complex/hybrid type glycans changed to a lesser degree (days 14–21). Upon completion of differentiation at day 21, the distribution of glycans remained relatively constant at day 24. Throughout the course of differentiation, nondecorated complex/hybrid type glycans that lacked fucose and sialic acid residues summed to less than 3% of the total cell surface glycans.

Structural Characterization of Differentially Expressed Glycans in Undifferentiated versus Fully Differentiated Cells-In addition to observing global trends, individual glycan compositions were monitored at the same time points as an initial search for key differentially expressed glycans. Specifically, we compared the glycans found in undifferentiated cells (day 5) and in fully differentiated cells (day 21), which are representative of the states before and after differentiation, respectively. For compositional analysis, isomeric contributions were grouped together and thus linkage positions are not specified. Compositions were determined by accurate mass and further confirmed by tandem MS (supplemental Fig. S1). On average, 119 nonredundant glycan compositions were identified in undifferentiated cells and 113 compositions in fully differentiated cells. Of these glycans, the changes in the abundances of 71 compositions were significant after differentiation (p < 0.05), as determined by a two-tailed, unpaired Student's t test. These changes are shown in Fig. 3 from day 5 to day 24, where colors represent log intensities of high abundant (red) and low abundant (green) glycans. Putative structures of glycans that increased in abundances after differentiation are shown in the upper box whereas those that decreased are



Fig. 2. Global glycan profiling of all detected glycans based on the sum of their abundances as a function of growth time. Summed abundances have been grouped by glycan subtype: high mannose, neutral complex/hybrid (C/H), fucosylated C/H, and sialylated C/H. Error bars indicate standard deviation between three biological replicates.

shown in the lower box. Data for identified glycans including composition, mass, glycan subtype, p values, and relative abundance are presented in supplemental Table S1.

Among the glycans found to be differentially expressed between undifferentiated and differentiated cells, residuespecific structural features were evaluated. To obtain structural information, select ions were fragmented by CID. The observance of fragment ions m/z 790.3 (Hex1HexNAc3) and m/z 1155.4 (Hex<sub>2</sub>HexNAc<sub>4</sub>) were used to identify the presence of a bisecting GlcNAc. Fragmentation spectra of sialylated species exhibited fragments at m/z 292.1 (NeuAc) and m/z 657.2 (Hex, HexNAc, NeuAc,). Nano-LC PGC-separation enabled evaluation of individual isomers. For example, Fig. 4A shows the extracted compound chromatograms (ECCs) of sialylated N-glycans where isomers are resolved at different retention times. Two isomers with composition Hex<sub>6</sub>HexNAc<sub>5</sub> Fuc<sub>1</sub>NeuAc<sub>2</sub> eluting at a) 11.4 min and b) 12.0 min showed inverse trends. Although isomer a decreased in abundance, isomer b increased after differentiation. According to the MS/MS spectra, isomer a contained a core fucose residue and isomer b contained an antennae fucose residue, as evidenced by fragment ions m/z 512.2 (Hex<sub>1</sub>HexNAc<sub>1</sub>Fuc<sub>1</sub>) and m/z 803.3 (Hex1HexNAc1Fuc1NeuAc1) (Fig. 4B).

Comparative changes of specific glycan features on undifferentiated and differentiated cell surfaces are shown in Fig. 5. High mannose type structures represented the most abundant glycans in undifferentiated cells. In particular, an accumulation of Man 8 and Man 9 was observed before differentiation. In differentiated cells, regardless of size, each individual high mannose form (Man 5 - Man 9) decreased in abundances. Truncated paucimannosidic glycans also showed significant decreases. In contrast, complex type glycans containing a bisecting GlcNAc increased in differentiated (45%, relative abundance) compared with undifferentiated (16%) cells. Additionally, increases in sialylation were readily observed in fully differentiated cells.

Direct quantitative analysis using peak areas of chromatograms required reproducible ion counts between samples. The reproducibility of the method is shown with triplicate experiments in supplemental Fig. S2. The total compound chromatograms (TCCs) for released glycans from three biological replicates of differentiated Caco-2 cells are overlaid in supplemental Fig. S2A. Accounting for biological variability, our method produces reproducible separation and detection of glycan structures. With cell cultures, reproducibility between passages is also an important factor to consider. Quantitative distribution of glycans identified from cells at the same stage of differentiation analyzed at different passages are shown in supplemental Fig. S2B. High reproducibility between biological replicates and between passages demonstrates that identifying glycosylation patterns is a robust way to distinguish between the stages of differentiation.

Comparative Cell Membrane Proteome Analysis in Undifferentiated versus Differentiated Cells—Proteomic analysis was



Fig. 3. Heatmap display of statistically significant glycan compositions from the undifferentiated state (day 5) to the postdifferentiated state (day 24). Pair-wise comparisons were conducted between undifferentiated and differentiated cell samples. Log values of abundances range from low (green) to high (red). Compositions that increased in abundance are arranged from lowest to highest abundance. Glycans that showed a decrease in abundance are ordered from highest to lowest abundance. Each glycan composition is labeled as high mannose (HM), neutral (N), fucosylated (F), sialylated (S), or fucosylated and sialylated (FS) glycans.



FIG. 4. **MS/MS identification of PGC-resolved isomers.** *A*, Extracted compound chromatograms of two isomeric forms (*a* and *b*) of  $Hex_6HexNAc_5Fuc_1NeuAc_2$  at different times points during Caco-2 differentiation. *B*, MS/MS spectra of isomers *a* and *b* showing the distinct monosaccharide fragmentation patterns.

conducted for undifferentiated (day 5) and fully differentiated (day 21) cells to identify the plasma membrane proteins that may be associated with the glycans. Following membrane enrichment as outlined above, proteins were treated with trypsin and analyzed by LC-CID-MS/MS. Database searching identified 211 membrane proteins for undifferentiated cells and 247 proteins for differentiated cells, with a false discovery rate of <1%. Of these, 173 (82%) and 201 (81%) were identified respectively as glycoproteins, with at least one potential N-glycosylation site. Identifications were confirmed as plasma membrane-localized proteins by gene ontology annotation. Known Caco-2 cell membrane hydrolases such as dipeptidyl peptidase IV, sucrase isomaltase, alkaline phosphatase, and aminopeptidase were found in the prepared cell membrane fractions as previously reported (23). Other essential transmembrane proteins, including receptors, cadherins, integrins, solute carriers, and tight junction proteins were identified with high confidence. Protein identifications are summarized in supplemental Table S2. Out of 296 nonredundant proteins, a total of 163 (55%) overlapping proteins were found between undifferentiated and differentiated cells, as depicted by the

Venn diagram in Fig. 6A. Additionally, undifferentiated and differentiated cells share  $\sim$ 55% (133 glycoproteins) of the same glycoproteins on their cell surfaces (Fig. 6B).

Contingent on the state of differentiation, several membrane glycoproteins were found exclusively on certain days. These unique proteins were classified according to their molecular functions using the PANTHER database (Fig. 6*C*). Biological functions of membrane glycoproteins that are unique to differentiated cells are similar to those uniquely identified on undifferentiated cell surfaces. These include cytoskeletal, transporter, transferase, receptor, and signaling molecule proteins. In comparison to undifferentiated cells, differentiated cells have a higher proportion of cell membrane hydrolases, proteases, cell junction proteins, and immune proteins. In contrast, nucleic acid binding, extracellular matrix, and chaperone proteins were more abundant in undifferentiated than in differentiated cells.

*Real-Time PCR Analysis* – Total RNA extracted from undifferentiated (day 5), partially differentiated (day 14), and fully differentiated (day 21) Caco-2 cells were subjected to human glycosylation PrimePCR plate profiling to determine the







expression of genes that encode enzymes involved in posttranslational oligosaccharide biosynthesis. A minimum threshold of a fourfold change was established to distinguish biologically relevant changes. Comparative differential expression analysis demonstrated that 14 out of 88 target genes were up-regulated more than fourfold in partially differentiated cells compared with undifferentiated cells, as listed in Fig. 7A. During N-glycan biosynthesis, high mannose glycans are first cleaved by mannosidases to generate hybrid and complex type glycans. An increase in the expression of mannosidases in partially differentiated cells is consistent with the observed decreases in high mannose type glycans and concurrent increases in complex type glycans observed in the glycomic profiles (Fig. 2). Among the up-regulated genes in partially differentiated Caco-2 cells, the gene MGAT3 encodes for a protein that produces bisecting GlcNAc glycans and the gene B4GALT5 encodes for a galactosyltransferase. Accordingly, glycans composed of bisecting GlcNAc and galactose residues were observed in increased levels in differentiated samples (Fig. 5). The expression of a similar set of genes increased by more than fourfold in fully differentiated cells

when compared with expression levels in undifferentiated cells (Fig. 7*B*).

Partially differentiated and fully differentiated cells differed by only three target genes that increased more than fourfold (Fig. 7*C*). The mannosidase-coding gene, MAN1C1, was upregulated in fully differentiated cells. In partially differentiated cells, two genes involved in O-linked oligosaccharide biosynthesis were up-regulated. As cells continued to differentiate post-confluency, expression of genes associated with glycan biosynthesis remains relatively the same, which agrees with the corresponding glycomic profiles.

## DISCUSSION

Differentiation of Caco-2 cells grown under the conditions described here begins between days 5 and 7, with the emergence of extended microvilli structures and detectable levels of intestinal alkaline phosphatase activity. Mass spectrometric structural analysis revealed that after day 7, the cell surface glycomic profile begins to show marked preference toward complex type glycans. As the major changes in the cell surface during IEC differentiation is because of microvilli



FIG. 6. Membrane proteins identified in undifferentiated and fully differentiated cells. Venn diagrams illustrating the overlap of Caco-2 cell membrane *A*, proteins and *B*, glycoproteins. *C*, PANTHER molecular function analysis of unique glycoproteins.

formation, we propose that the glycan changes are largely microvilli-associated. Indeed, hydrolases such as alkaline phosphatase, which are present on the apical border, are highly glycosylated (53, 54), and their function and stability have been shown to be associated with the presented glycans (55, 56). Improper glycosylation would result in rapid degra-



Fig. 7. Comparative normalized expression levels of relevant target genes involved in the glycosylation pathway. Numbered dots indicate genes that increased expression more than fourfold (threshold) as described in the tables. *A*, undifferentiated *versus* partially differentiated, *B*, undifferentiated *versus* fully differentiated, and *C*, partially *versus* fully differentiated Caco-2 cells.

dation of these brush border enzymes. As seen in the micrographs, the percent of microvilli covering the cell surface continue to increase until the surface is entirely covered on day 21. The presentation of complex type glycan structures on cell surfaces coincides with the development and dominance of microvilli.

Discrimination between undifferentiated and differentiated states was observable in the abundances of high mannose type structures present on the cell surface. Higher levels of this subtype have been reported in the membrane proteins of moderately and poorly differentiated colorectal cancer cells (16) and in the whole cell lysate of undifferentiated Caco-2 cells (57). We observe an accumulation of Man 8 and Man 9 selectively on the cell surface of undifferentiated Caco-2. This observation is proposed to be because of the trafficking of proteins before additional processing of larger high mannose type structures. In the N-glycan biosynthetic pathway, mannosidases are responsible for cleaving oligomannose chains before the transfer of GlcNAc residues to form hybrid and

complex type glycans. Defects in the regulation of this pathway would result in accumulation of these structures. In comparison, as Caco-2 cells matured, a redistribution of the relative abundances of Man 5 - Man 9 on the cell surface was observed. Significant decreases in high mannose type glycans were accompanied by increases in decorated complex type glycans.

Glycosylation and differentiation appear highly correlated with sialylation being the key modification. Particular attention is given to sialic acid as it resides on the terminal ends of N-glycans. Increases in sialylation of differentiated cell membrane glycans can be explained by the up-regulated expression of the ST6GAL1 gene, which encodes for a sialyltransferase that adds sialic acid to a galactose in an  $\alpha$ -2,6-linkage. The preference for the cell to regulate the addition of  $\alpha$ -2,6-sialic acids on differentiated cells suggests a correlation between sialic acid linkage and differentiation. The prominence of sialic acid residues on epithelia has functional significance on how the cell interacts with the external environment (58,

59). *In vivo*, differentiated cells are positioned on the tips of the villi and are therefore exposed to luminal pathogens. It is important for the host to regulate the types of glycans expressed on differentiated cells as pathogens recognize certain sialic acid moieties with high specificity (2).

Another distinguishing feature of the mature glycosylation profile is elevated levels of the bisecting GlcNAc containing oligosaccharides. Increases were also observed in the MGAT3 transcript from differentiated cells, which encodes for  $\beta$ -1,4-*N*-acetylglucosaminyltransferase III (GnT-III), responsible for the introduction of a bisecting GlcNAc. The bisecting GlcNAc modification of complex and hybrid N-glycans has important influences on membrane proteins. For example, the presence of a bisecting GlcNAc has implications for proteins involved in cell-cell adhesion, such as E-cadherin, which is known to be a target protein of GnT-III (60, 61). Thus, we speculate that the presence of bisecting GlcNAc on differentiated cells is involved in the regulation of cell-cell adhesion.

The consequences of gene expression changes are large as the gene-encoded enzymes are involved in multiple steps during the synthesis of glycoconjugates, which include N-glycoproteins, O-glycoproteins, and glycolipids. Further studies would be needed to directly correlate transcript changes to the glycan patterns. Although not all structural features were investigated here, other changes are notable based on the gene expression changes, such as disialyl units, heparin sulfates, and O-glycans. Complete structural analyses of the surface glycans are ongoing and will be a subject of future reports. At the protein level,  $\sim$ 50% of proteins remain unchanged on the cell surface as Caco-2 cells differentiate. Differentiation-related changes in proteins involve a variety of new molecular functions. Glycan-specific changes of these proteins may serve as distinguishing markers for understanding the regulation of differentiation and has implications in understanding the irregularities that lead to malignant transformations.

The ability to extensively characterize and quantitate glycan structures using mass spectrometric profiling is a valuable tool for understanding biological processes. Our results demonstrate that IEC cell differentiation is accompanied by characteristic changes in the cell surface glycome, which may be specific and sensitive markers of differentiation stages. Changes in glycan modifications on the cell membrane may be a general mechanism of differentiation. Differentiated Caco-2 cells are an important model cell line that exhibits a characteristic glycan profile on their surfaces. It is important to note that careful attention should be given when harvesting Caco-2 cells as changes in glycosylation expression patterns are growth-dependent.

*Acknowledgments*—We thank Patricia Kysar and Fred Hayes for their assistance with scanning electron microscopy and Anthony Herren for Orbitrap performance.

\* This work was supported by the National Institutes of Health under award number GM049077 (to C.B.L.). EM was supported by career awards from the Howard Hughes Medical Institute and the Burroughs Wellcome fund and by a NIH New Innovator Award (DP2OD008752).

S This article contains supplemental Figs. S1 and S2 and Tables S1 and S2.

¶ To whom correspondence should be addressed: Department of Chemistry, University of California, Davis, One Shields Avenue, Davis, CA 95616. Tel.: 530-752-0504; Fax: 530-752-8995; E-mail: cblebrilla@ucdavis.edu.

#### REFERENCES

- Moran, A. P., Gupta, A., and Joshi, L. (2011) Sweet-talk: role of host glycosylation in bacterial pathogenesis of the gastrointestinal tract. *Gut* 60, 1412–1425
- Marcobal, A., Southwick, A. M., Earle, K. A., and Sonnenburg, J. L. (2013) A refined palate: bacterial consumption of host glycans in the gut. *Glycobiology* 23, 1038–1046
- Karlsson, K. A. (1999) Bacterium-host protein-carbohydrate interactions and pathogenicity. *Biochem. Soc. Trans.* 27, 471–474
- van Kooyk, Y., and Rabinovich, G. A. (2008) Protein-glycan interactions in the control of innate and adaptive immune responses. *Nat. Immunol.* 9, 593–601
- Ohtsubo, K., and Marth, J. D. (2006) Glycosylation in cellular mechanisms of health and disease. *Cell* **126**, 855–867
- Marth, J. D., and Grewal, P. K. (2008) Mammalian glycosylation in immunity. Nat. Rev. Immunol. 8, 874–887
- Reis, C. A., Osorio, H., Silva, L., Gomes, C., and David, L. (2010) Alterations in glycosylation as biomarkers for cancer detection. *J. Clin. Pathol.* 63, 322–329
- Adamczyk, B., Tharmalingam, T., and Rudd, P. M. (2012) Glycans as cancer biomarkers. *Biochim. Biophys. Acta* 1820, 1347–1353
- Haltiwanger, R. S., and Lowe, J. B. (2004) Role of glycosylation in development. Annu. Rev. Biochem. 73, 491–537
- Dennis, J. W., Granovsky, M., and Warren, C. E. (1999) Protein glycosylation in development and disease. *Bioessays* 21, 412–421
- Laitinen, L., Virtanen, I., and Saxén, L. (1987) Changes in the glycosylation pattern during embryonic development of mouse kidney as revealed with lectin conjugates. J. Histochem. Cytochem. 35, 55–65
- Kimber, S. J., Stones, R. E., and Sidhu, S. S. (2001) Glycosylation changes during differentiation of the murine uterine epithelium. *Biochem. Soc. Trans.* 29, 156–162
- Despont, J. P., Abel, C. A., and Grey, H. M. (1975) Sialic acids and sialyltransferases in murine lymphoid cells: indicators of T cell maturation. *Cell. Immunol.* **17**, 487–494
- Reisner, Y., Linker-Israeli, M., and Sharon, N. (1976) Separation of mouse thymocytes into two subpopulations by the use of peanut agglutinin. *Cell. Immunol.* 25, 129–134
- Toporowicz, A., and Reisner, Y. (1986) Changes in sialyltransferase activity during murine T cell differentiation. *Cell. Immunol.* **100**, 10–19
- Sethi, M. K., Thaysen-Andersen, M., Smith, J. T., Baker, M. S., Packer, N. H., Hancock, W. S., and Fanayan, S. (2014) Comparative N-glycan profiling of colorectal cancer cell lines reveals unique bisecting GlcNAc and alpha-2,3-linked sialic acid determinants are associated with membrane proteins of the more metastatic/aggressive cell lines. *J. Proteome Res.* **13**, 277–288
- Balog, C. I., Stavenhagen, K., Fung, W. L., Koeleman, C. A., McDonnell, L. A., Verhoeven, A., Mesker, W. E., Tollenaar, R. A., Deelder, A. M., and Wuhrer, M. (2012) N-glycosylation of colorectal cancer tissues: a liquid chromatography and mass spectrometry-based investigation. *Mol. Cell. Proteomics* **11**, 571–585
- Crosnier, C., Stamataki, D., and Lewis, J. (2006) Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nat. Rev. Genet.* 7, 349–359
- Fogh, J., Wright, W. C., and Loveless, J. D. (1977) Absence of HeLa cell contamination in 169 cell lines derived from human tumors. *J. Natl. Cancer Inst.* 58, 209–214
- Rousset, M. (1986) The human colon carcinoma cell lines HT-29 and Caco-2: two in vitro models for the study of intestinal differentiation.

Biochimie 68, 1035-1040

- Pinto, M., Robineleon, S., Appay, M. D., Kedinger, M., Triadou, N., Dussaulx, E., Lacroix, B., Simonassmann, P., Haffen, K., Fogh, J., and Zweibaum, A. (1983) Enterocyte-like differentiation and polarization of the human-colon carcinoma cell-line Caco-2 in culture. *Biol. Cell* 47, 323–330
- Stierum, R., Gaspari, M., Dommels, Y., Ouatas, T., Pluk, H., Jespersen, S., Vogels, J., Verhoeckx, K., Groten, J., and van Ommen, B. (2003) Proteome analysis reveals novel proteins associated with proliferation and differentiation of the colorectal cancer cell line Caco-2. *Biochim. Biophys. Acta* 1650, 73–91
- Pshezhetsky, A. V., Fedjaev, M., Ashmarina, L., Mazur, A., Budman, L., Sinnett, D., Labuda, D., Beaulieu, J. F., Menard, D., Nifant'ev, I., and Levy, E. (2007) Subcellular proteomics of cell differentiation: quantitative analysis of the plasma membrane proteome of Caco-2 cells. *Proteomics* 7, 2201–2215
- Tadjali, M., Seidelin, J. B., Olsen, J., and Troelsen, J. T. (2002) Transcriptome changes during intestinal cell differentiation. *Biochim. Biophys. Acta* 1589, 160–167
- Turck, N., Richert, S., Gendry, P., Stutzmann, J., Kedinger, M., Leize, E., Simon-Assmann, P., Van Dorsselaer, A., and Launay, J. F. (2004) Proteomic analysis of nuclear proteins from proliferative and differentiated human colonic intestinal epithelial cells. *Proteomics* 4, 93–105
- Tremblay, E., Auclair, J., Delvin, E., Levy, E., Ménard, D., Pshezhetsky, A. V., Rivard, N., Seidman, E. G., Sinnett, D., Vachon, P. H., and Beaulieu, J. F. (2006) Gene expression profiles of normal proliferating and differentiating human intestinal epithelial cells: a comparison with the Caco-2 cell model. *J. Cell. Biochem.* **99**, 1175–1186
- Brockhausen, I., Romero, P. A., and Herscovics, A. (1991) Glycosyltransferase changes upon differentiation of CaCo-2 human colonic adenocarcinoma cells. *Cancer Res.* 51, 3136–3142
- Amano, J., Kobayashi, K., and Oshima, M. (2001) Comparative study of glycosyltransferase activities in Caco-2 cells before and after enterocytic differentiation using lectin-affinity high-performance liquid chromatography. Arch. Biochem. Biophys. **395**, 191–198
- Dall'Olio, F., Malagolini, N., Guerrini, S., Lau, J. T., and Serafini-Cessi, F. (1996) Differentiation -dependent expression of human beta-galactoside alpha 2,6-sialyltransferase mRNA in colon carcinoma CaCo-2 cells. *Glycoconj. J.* **13**, 115–121
- Zheng, T., Peelen, D., and Smith, L. M. (2005) Lectin arrays for profiling cell surface carbohydrate expression. J. Am. Chem. Soc. 127, 9982–9983
- Tao, S. C., Li, Y., Zhou, J., Qian, J., Schnaar, R. L., Zhang, Y., Goldstein, I. J., Zhu, H., and Schneck, J. P. (2008) Lectin microarrays identify cell-specific and functionally significant cell surface glycan markers. *Glycobiology* 18, 761–769
- Zaia, J. (2008) Mass spectrometry and the emerging field of glycomics. Chem. Biol. 15, 881–892
- Leymarie, N., and Zaia, J. (2012) Effective use of mass spectrometry for glycan and glycopeptide structural analysis. *Anal. Chem.* 84, 3040–3048
- North, S. J., Hitchen, P. G., Haslam, S. M., and Dell, A. (2009) Mass spectrometry in the analysis of N-linked and O-linked glycans. *Curr. Opin. Struct. Biol.* **19**, 498–506
- Wuhrer, M. (2013) Glycomics using mass spectrometry. *Glycoconj. J.* 30, 11–22
- Hua, S., An, H. J., Ozcan, S., Ro, G. S., Soares, S., DeVere-White, R., and Lebrilla, C. B. (2011) Comprehensive native glycan profiling with isomer separation and quantitation for the discovery of cancer biomarkers. *Analyst* 136, 3663–3671
- Hong, Q., Lebrilla, C. B., Miyamoto, S., and Ruhaak, L. R. (2013) Absolute quantitation of immunoglobulin G and its glycoforms using multiple reaction monitoring. *Anal. Chem.* 85, 8585–8593
- Ruhaak, L. R., Miyamoto, S., and Lebrilla, C. B. (2013) Developments in the identification of glycan biomarkers for the detection of cancer. *Mol. Cell. Proteomics* 12, 846–855
- Morelle, W., and Michalski, J. C. (2007) Analysis of protein glycosylation by mass spectrometry. *Nat. Protoc.* 2, 1585–1602
- Han, L., and Costello, C. E. (2013) Mass spectrometry of glycans. *Biochemistry* 78, 710–720
- 41. Harvey, D. J., Scarff, C. A., Crispin, M., Scanlan, C. N., Bonomelli, C., and

Scrivens, J. H. (2012) MALDI-MS/MS with traveling wave ion mobility for the structural analysis of N-linked glycans. *J. Am. Soc. Mass Spectrom.* **23**, 1955–1966

- Pabst, M., and Altmann, F. (2011) Glycan analysis by modern instrumental methods. *Proteomics* 11, 631–643
- Reinhold, V., Zhang, H., Hanneman, A., and Ashline, D. (2013) Toward a platform for comprehensive glycan sequencing. *Mol. Cell. Proteomics* 12, 866–873
- Zauner, G., Deelder, A. M., and Wuhrer, M. (2011) Recent advances in hydrophilic interaction liquid chromatography (HILIC) for structural glycomics. *Electrophoresis* 32, 3456–3466
- Hua, S., Williams, C. C., Dimapasoc, L. M., Ro, G. S., Ozcan, S., Miyamoto, S., Lebrilla, C. B., An, H. J., and Leiserowitz, G. S. (2013) Isomer-specific chromatographic profiling yields highly sensitive and specific potential N-glycan biomarkers for epithelial ovarian cancer. *J. Chromatogr. A* 1279, 58–67
- 46. An, H. J., Gip, P., Kim, J., Wu, S., Park, K. W., McVaugh, C. T., Schaffer, D. V., Bertozzi, C. R., and Lebrilla, C. B. (2012) Extensive Determination of Glycan Heterogeneity Reveals an Unusual Abundance of High Mannose Glycans in Enriched Plasma Membranes of Human Embryonic Stem Cells. *Mol. Cell. Proteomics* **11**, M111.010660
- An, H. J., Gip, P., Kim, J., Wu, S., Park, K. W., McVaugh, C. T., Schaffer, D. V., Bertozzi, C. R., and Lebrilla, C. B. (2012) Extensive determination of glycan heterogeneity reveals an unusual abundance of high mannose glycans in enriched plasma membranes of human embryonic stem cells. *Mol. Cell. Proteomics* **11**, M111.010660
- Kronewitter, S. R., An, H. J., de Leoz, M. L., Lebrilla, C. B., Miyamoto, S., and Leiserowitz, G. S. (2009) The development of retrosynthetic glycan libraries to profile and classify the human serum N-linked glycome. *Proteomics* 9, 2986–2994
- Ruhaak, L. R., Miyamoto, S., Kelly, K., and Lebrilla, C. B. (2012) N-Glycan profiling of dried blood spots. *Anal. Chem.* 84, 396–402
- Craig, R., and Beavis, R. C. (2004) TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 20, 1466–1467
- Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392
- 52. Thomas, P. D., Kejariwal, A., Campbell, M. J., Mi, H., Diemer, K., Guo, N., Ladunga, I., Ulitsky-Lazareva, B., Muruganujan, A., Rabkin, S., Vandergriff, J. A., and Doremieux, O. (2003) PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res.* **31**, 334–341
- Kenny, A. J., and Maroux, S. (1982) Topology of microvillar membrance hydrolases of kidney and intestine. *Physiol. Rev.* 62, 91–128
- Semenza, G. (1986) Anchoring and biosynthesis of stalked brush border membrane proteins: glycosidases and peptidases of enterocytes and renal tubuli. *Annu. Rev. Cell Biol.* 2, 255–313
- Wetzel, G., Heine, M., Rohwedder, A., and Naim, H. Y. (2009) Impact of glycosylation and detergent-resistant membranes on the function of intestinal sucrase-isomaltase. *Biol. Chem.* **390**, 545–549
- Trugnan, G., Rousset, M., Chantret, I., Barbat, A., and Zweibaum, A. (1987) The posttranslational processing of sucrase-isomaltase in HT-29 cells is a function of their state of enterocytic differentiation. *J. Cell Biol.* **104**, 1199–1205
- 57. Fujitani, N., Furukawa, J., Araki, K., Fujioka, T., Takegawa, Y., Piao, J., Nishioka, T., Tamura, T., Nikaido, T., Ito, M., Nakamura, Y., and Shinohara, Y. (2013) Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2105–2110
- Schauer, R. (2009) Sialic acids as regulators of molecular and cellular interactions. *Curr. Opin. Struct. Biol.* **19**, 507–514
- Varki, A. (2008) Sialic acids in human health and disease. *Trends Mol. Med.* 14, 351–360
- Gumbiner, B. M. (2005) Regulation of cadherin-mediated adhesion in morphogenesis. Nat. Rev. Mol. Cell Biol. 6, 622–634
- Takahashi, M., Kuroki, Y., Ohtsubo, K., and Taniguchi, N. (2009) Core fucose and bisecting GlcNAc, the direct modifiers of the N-glycan core: their functions and target proteins. *Carbohydr. Res.* **344**, 1387–1390