Extensive Determination of Glycan Heterogeneity Reveals an Unusual Abundance of High Mannose Glycans in Enriched Plasma Membranes of Human Embryonic Stem Cells

Hyun Joo An, Phung Gip, Jae Han Kim, Shuai Wu, Kun Wook Park, Cheryl T. McVaugh, David V. Schaffer, Carolyn R. Bertozzi, and Carlito B. Lebrilla

Most cell membrane proteins are known or predicted to be glycosylated in eukaryotic organisms, where surface glycans are essential in many biological processes including cell development and differentiation. Nonetheless, the glycosylation on cell membranes remains not well characterized because of the lack of sensitive analytical methods. This study introduces a technique for the rapid profiling and quantitation of N- and O-glycans on cell membranes using membrane enrichment and nano-flow liquid chromatography/mass spectrometry of native structures. Using this new method, the glycome analysis of cell membranes isolated from human embryonic stem cells and somatic cell lines was performed. Human embryonic stem cells were found to have high levels of high mannose glycans, which contrasts with IMR-90 fibroblasts and a human normal breast cell line, where complex glycans are by far the most abundant and high mannose glycans are minor components. O-Glycosylation affects relatively minor components of cell surfaces. To verify the quantitation and localization of glycans on the human embryonic stem cell membranes, flow cytometry and immunocytochemistry were performed. Proteomics analyses were also performed and confirmed enrichment of plasma membrane proteins with some contamination from endoplasmic reticulum and other membranes. These findings suggest that high mannose glycans are the major component of cell surface glycosylation with even terminal glucoses. High mannose glycans are not commonly presented on the surfaces of mammalian cells or in serum yet may play important roles in stem cell biology. The results also mean that distinguishing stem cells from other mammalian cells may be facilitated by the major difference in the glycosylation of the cell membrane. The deep structural analysis enabled by this new method will enable future mechanistic studies on the biological significance of high mannose glycans on stem cell membranes and provide a general tool to examine cell surface glycosylation.

Glycosylation is the process by which oligosaccharides, termed glycans, are appended onto membrane and secreted proteins and lipids. It is the most common and complex form of post-translational modification, with ~50% of all eukaryotic proteins glycosylated (1, 2). A majority of glycans are found on the cell surface, where they are optimally poised to be the first cellular components encountered by approaching cells, pathogens, antibodies, or other molecules, as well as advertise information about the internal state and homeostasis of the cell (3, 4). Therefore, glycans play an essential role in many biological processes, including cell development and differentiation, cell-cell or cell-matrix communication, and pathogen-host recognition (3, 5–7). In fact, differences in glycan profiles between healthy and diseased states are utilized for clinical diagnosis (7), providing targets for many novel classes of therapeutics including cancer chemotherapy, diabetes treatment, and antibiotic and anti-viral medicine (5, 8). Glycans are highly heterogeneous in nature, varying in the composition of individual monosaccharide building blocks, the positions with which these building blocks link to each other, and the stereochemical disposition of the linkages (α or β). This complexity has presented a significant challenge for obtaining structural information about glycans at the molecular level.
level, particularly in contexts relevant to native cellular physiology. For this reason, the technology for elucidating their structures has lagged behind other major classes of biomolecules, such as protein, DNA, and RNA, in the molecular revolution currently underway in biology and medicine (9). Although several approaches including lectin binding (10–12), cell surface shaving (13), cell surface labeling (14, 15), and antibody-mediated membrane enrichment have been developed to identify surface glycans or glycoproteins, comprehensive and conclusive structural elucidation and identification remain laborious and difficult (16).

Human embryonic stem cells (hESCs) are of particular biomedical interest because they hold enormous potential for regenerative medicine and drug discovery. As a model system, they can also contribute to the understanding of human development and potentially help to guide cancer research because hESCs and cancer cells share similar characteristics (17–19). Therefore, structural elucidation of the components present on hESC membranes may provide a basis for understanding their role in hESC maintenance and differentiation. Recent studies suggest that glycans on the plasma membrane of hESCs change during differentiation; these changes can have profound effects on cellular function (20) and could be harnessed to meet the need to identify cell surface markers for isolating and purifying specific cell populations for therapeutic application. Indeed, one of the earliest pluripotent stem cell markers is SSEA-1 (stage-specific embryonic antigen-1), a glycan, otherwise known as Lexis X antigen, expressed on mouse embryonic stem cells (21, 22). Two other antigenic epitopes discovered were SSEA-3 and SSEA-4 (23), which are both glycolipids that have become the most common cell surface markers used to characterize hESCs (24).

Cell surface glycosylation may play an important role in development and may provide important new sources of markers for differentiation. Studies regarding the glycosylation of stem cell surfaces are limited. Wearne et al. (12) recently reported the use of fluorescence-labeled lectins to identify a number of specific structural motifs including mannose residues, α2,3- and α2,6-linked N-acetyllactosamine, α1,6-linked L-fucosyl, and β-D-galactosyl groups. In addition, they also found a number of common antigens including T, Tn, and sialyl-Tn. A more structurally intensive study of whole stem cell glycosylation was reported by Satomaa et al. (25). The cell surface studies were limited to lectins, which cannot be used quantitatively. Structural methods including nuclear magnetic resonance, mass spectrometry, and glycosidase digestion were used on whole cells where they showed that the N-glycan profile was rich in high mannose glycans as well as complex type structures that are terminated by both α2,3- and α2,6-sialylated oligosaccharides and fucosyl structures.

Here we characterize the N-glycan profile of human embryonic stem cell membrane glycans using a method that enables specific detection and quantitation and acquisition of structural information. Interestingly, hESCs have high levels of high mannose glycans on the cell surface, which is largely unprecedented in mammalian cells. Moreover, the hESCs were particularly rich in Man8 and Man9 structures, including Man9 with terminal glycoside still intact. This unusual glycomic signature might have functional implications, as well as practical utility in the characterization and isolation of hESCs.

**EXPERIMENTAL PROCEDURES**

**Human Embryonic Stem Cell Culture**—The National Institutes of Health-approved hESC lines, H1 and HSF-6, were maintained under feeder-free conditions using a chemically defined medium, X-VIVO 10 (Cambrex, Walkersville, MD), supplemented with human recombinant growth factors fibroblast growth factor and transforming growth factor-β1 (80 and 0.5 ng/ml, respectively; R & D Systems (10). The cells are propagated on hESC-qualified Matrigel-coated plates (BD Biosciences). Medium was exchanged daily after the first 48 h in culture, and the cells were passaged every 5–7 days using collagenase IV (200 units/ml; Invitrogen) and mechanically removed. For glycan analysis, the cells were collected after collagenase IV treatment, centrifuged, washed with PBS (Invitrogen) and then collected and collected at different passage numbers to obtain biological triplicates. Karyotype analysis was routinely performed and indicated that all samples were diploid and had no chromosomal abnormalities. The cells were routinely stained with pluripotent markers Oct4 and SSEA-4.

**Somatic Cell Line Culture**—IMR-90 human fibroblast cells (University of California Berkeley Tissue Culture Facility) were grown in Dulbecco’s modified Eagle’s high glucose medium (Invitrogen) supplemented with 10% fetal bovine serum (HyClone). The cells were passaged every 3 days using trypsin 0.25% and EDTA solution (Invitrogen). For glycan analysis, IMR-90s were collected using 0.5 mM EDTA, centrifuged, washed with PBS (Invitrogen), and then collected and collected at different passage numbers to obtain biological triplicates. Karyotype analysis was routinely performed and indicated that all samples were diploid and had no chromosomal abnormalities. The cells were routinely stained with pluripotent markers Oct4 and SSEA-4.

**Cell Membrane Extraction**—Membrane extraction was performed using ultracentrifugation. The pellets were thawed on ice with the addition of a homogenization buffer consisting of 0.25 M sucrose, 20 mM Hepes-KOH, pH 7.4, and protease inhibitor mixture (1:100; Calbiochem/EMD Chemicals). The cells were sonicated on ice, and cell lysates were centrifuged at 1,000 × g for 10 min to remove the nuclear fraction and debris. The supernatant was collected, and additional homogenization buffer was added for ultracentrifugation at 200,000 × g for 45 min at 4 °C to remove the cytoplasmic fraction.

1 The abbreviations used are: hESC, human embryonic stem cells; SPE, solid phase extraction; Hex, hexose; HexNAc, N-acetylhexosamine; NeuAc, sialic acid; Fuc, fucose; Man, mannose; GGC, graphitized carbon cartridge; ICR, ion cyclotron resonance; IRMPD, infrared multiphoton dissociation; Con A, Canavalia ensiformis; GNA, Galanthus nivalis; Man9, GlcNAc2Man9; Man8, GlcNAc2Man8; Man7, GlcNAc2Man7; JAC, Artocarpus integrifolia.

10.1074/mcp.M111.010660–2
The pellets were resuspended in 0.2 M Na2CO3 (pH 11) to break up the microsomes. The samples were spun twice more at 200,000 × g for 45 min to wash the samples of the cytoplasmic fraction. The supernatant was removed, and the membrane fractions were frozen at −20 °C.

**Western Blot Analysis**—All of the fractions (nuclear, cytoplasmic, and membranes) were analyzed by SDS-PAGE followed by Western blot using known organelle-specific markers for the nucleus (nuclear pore complex proteins; Covance), endoplasmic reticulum (Bip/GRP78; BD Biosciences), cytosol (α-tubulin; Sigma), and the plasma membrane (CD49b; BD Biosciences). Primary antibodies were probed with a horseradish peroxidase conjugated anti-mouse secondary antibody (IgG). Before Western blot analysis, membrane pellets were resuspended in 4% SDS buffer, and protein concentration was determined by the BCA assay (Pierce). The samples (20 μg) were separated by SDS/PAGE (4–12%; Bio-Rad).

**Glycan Enrichment**—For the analysis of N-glycans, 100 μl of 100 mM ammonium bicarbonate (NH4HCO3), 5 mM DTT (Promega) was added to the samples and heated to 100 °C for 2 min to denature the protein. After cooling at room temperature, 2 μl of peptide N-glycosidase F (New England Biolabs) was added to the mixture (pH 7.5) and incubated at 37 °C for 12 h in a water bath. 800 μl of chilled ethanol was added, and the mixture was frozen at −80 °C for 1 h and then centrifuged to separate glycans from deglycosylated proteins. The supernatant was completely dried down to remove the ethanol prior to solid phase extraction (SPE) using a graphitized carbon cartridge (GCC; Alttech).

For O-glycan analysis, alkaline borohydride solution (500 μl, mixture of 1.0 M sodium borohydride and 0.1 M sodium hydroxide) was added to the membrane fraction. The mixture was incubated at 42 °C for 12 h in a water bath. The addition of 1.0 M hydrochloric acid solution was slowly added in ice bath to stop the reaction and destroy excess sodium borohydride.

**Glycan Enrichment**—Released N- and O-glycans were purified and enriched by SPE-GCC. Prior to use, graphitized carbon cartridge (150 mg of bed weight, 4 ml of cartridge volume) was washed with nanopure water followed by 80% ACN in 0.05% TFA (v/v) and again with nanopure water. Glycan solutions were applied to the GCC cartridge and subsequently washed with several cartridge volumes of nanopure water at a flow rate of 1 ml/min to remove salts. Glycans were eluted stepwise with 10% ACN in H2O (v/v), 20% ACN in H2O (v/v), and 40% ACN in 0.05% TFA in H2O (v/v). Each fraction was collected and concentrated in vacuo prior to mass spectrometry analysis. Fractions were reconstituted in nanopure water prior to MS analysis.

**Mass Spectrometric Analysis**—Mass spectra were recorded on a Fourier transform ion cyclotron resonance (ICR) mass spectrometer with an external source HiResMALDI (IonSpec Corporation) equipped with a 7.0 Tesla magnet. The HiResMALDI was equipped with a pulsed Nd:YAG laser (355 nm), 2.5-Dihydroxy-benzoic acid was used as a matrix (5 mg/100 ml in 50% ACN:H2O) for both positive and negative modes. A saturated solution of NaCl in 50% ACN in H2O was used as a cation dopant to increase signal sensitivity. The glycan solution (0.7 μl) was applied to the MALDI probe followed by matrix solution (0.7 μl). The sample was dried under vacuum prior to mass spectrometric analysis.

**Structural Determination Using Infrared Multiphoton Dissociation (IRMPD)**—A desired ion was readily selected in the analyzer with the use of an arbitrary wave form generator and a frequency synthesizer. A continuous wave Parallax CO2 laser with 20-W maximum power and 10.6-μm wavelength was installed at the rear of the magnet and was used to provide the photons for IRMPD. The laser beam diameter is 6 mm as specified by the manufacturer. The laser beam was expanded to 12 mm by means of a 2× beam expander (Synrad) to ensure complete irradiation of the ion cloud through the course of the experiment. The laser was aligned and directed to the center of the ICR cell through a BaF2 window (Bicon Corporation). Photon irradiation time was optimized to produce the greatest number and abundance of fragment ions. The laser was operated at an output of ~13 W.

**NanoLC Mass Spectrometry**—GCC fractions were analyzed using a microfluidic HPLC-Chip-TOF MS (Agilent, CA). The microfluidic HPLC-Ch consists of an enrichment column, an LC separation column packed with porous graphitized carbon, and a nanoelectrospray tip. Separation was performed by a binary gradient A: 3% acetonitrile in 0.1% formic acid solution and B: 90% acetonitrile in 0.1% formic acid solution. The column was initially equilibrated and eluted with the flow rate at 0.3 μl/min for nanopump and 4 μl/min for capillary pump. The 65-min gradient was programmed as follows: 2.5–20 min, 0–16% B; 20–30 min, 16–44% B; 30–35 min, B increased to 100%, then continued 100% B to 45 min, finally 0% B for 20 min to equilibrate the chip column before next sample injection. Each possible composition of N-glycans was identified with the in-house program Glycox (26) and the N-glycan library (27) according to the mass tolerance with additional retention times and abundance information noted at the same time.

**Proteomic Analyses by LC-MS/MS Analyses**—The membrane proteins were dried and solubilized with 60 μl of 8 μL urea. The samples were then reduced with DTT and alkylated with iodoacetamide. After dilution in 180 μl of water, an overnight digestion with trypsin was performed. The peptides were then concentrated and desalted using C18 peptide trap (MichromeBioResources, Inc. Auburn, CA) before LC separation and online MS/MS. A NanoLC-2D system (Eksigent, Dublin, CA) coupled with an LTQ ion trap mass spectrometer (Thermo Finnigan) was used with a homemade fritless reverse phase microcapillary column (75 μm × 180 mm; packed with Magic C18AQ, 3 μm 100 Å; Michrom Bio Resources) and vented column configuration. Digested samples were transferred from the autosampler to the online trap column (0.15 mm × 20 mm; packed with Magic C18AQ, 3 μm, 100 Å) and desalted. The peptides were eluted from the trap and separated on the capillary column using a reverse phase gradient at a flow rate of 300 nl/min and directly electrosprayed into the mass spectrometer. A cycle of one MS survey scan followed by 10 MS/MS scans was repeatedly acquired over the LC gradient. Dynamic exclusion for 1-min duration was utilized. Buffers were 0.1% formic acid in water (buffer A) and 0.1% formic acid in acetonitrile (buffer B). A 107-min gradient (2–40% B for 95 min, followed by 40–80% B for 12 min) was used. Protein identification based on LC-MS/MS was performed using X! Tandem with a fragment ion mass tolerance of 0.040 Da and a parent ion tolerance of 1.8 Da. Iodoacetamide derivatization of cysteine was specified as a fixed modification. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two identified peptides.

**Immunofluorescence**—hESCs were fixed with 2% paraformaldehyde and rinsed three times in PBS. The cells were blocked with staining buffer (2% fetal bovine serum in PBS) and then stained with pluripotential marker SSEA-4 (2.5 μg/500 μl/well; Millipore) for 30 min at room temperature. The wells were rinsed in PBS before adding FITC-conjugated lectins (20 μg/ml; EY Labs) and Alexa Fluor 594-conjugated goat anti-mouse secondary antibody to SSEA-4 (1:400; Invitrogen). The wells were rinsed in PBS and stained with a solution of 1× Hoechst 23187 (Sigma) as a nuclear stain and analyzed using an Olympus IX71 fluorescent microscope. Control wells were stained with mouse IgG3 isotype (Invitrogen), and the lectins were incubated with their respective inhibitory sugar.

**Flow Cytometry**—hESCs were collected after incubation with collagen IV (200 units/ml; Invitrogen) and mechanically removed. The
colonies were dissociated into single-cell suspensions in 0.5 mM EDTA, then filtered through a 40-micron cell strainer, and counted. The cells were blocked with staining buffer (2% fetal bovine serum in PBS) and then stained with pluripotent marker SSEA-4 (2.5 μg/500 μL/500,000 cells; Millipore) for 30 min on ice. The cells were washed and stained with APC-conjugated goat anti-mouse secondary antibody to SSEA-4 and 5, 10, 20, or 40 μg/ml of the following FITC-conjugated lectins: Canavalienformis or Galanthus nivalis (EV Labs). To validate binding specificity, hESCs were also stained with lectins preincubated with sugar haptens: methyl-α-mannoside and yeast mannan, respectively (Sigma). After 30 min on ice, the cells were washed and resuspended in staining buffer with propidium iodide to distinguish dead cells from live cells. Flow cytometry (BD FACs Calibur from BD Biosciences) was performed, and the data were analyzed using FlowJo software (TreeStarInc). At least three independent assays were carried out for each lectin. The final quantitation represents bodies to SSEA-4 and 5, 10, 20, or 40 μg/ml of the following FITC-conjugated lectins. hESCs were also stained with mouse IgG3 isotype (Invitrogen), as a control for SSEA-4 labeling.

RESULTS

The experimental strategy, including: (i) the purification of cell membrane fractions from whole cell lysates by ultracentrifugation, (ii) release and enrichment of surface glycans by SPE using a graphitized carbon, (iii) glycan profiling by high performance mass spectrometry, and (iv) isomer separation and quantitation by nanoLC, is outlined in supplemental Fig. S1. These methods were developed and streamlined to profile cell membrane glycans effectively and selectively by mass spectrometry. As further described below, the quantitation and localization of glycans on the hESCs membrane were also validated by flow cytometry and immunocytochemistry.

Isolation of Membrane Fractions by Ultracentrifugation in hESCs—Selective isolation of membrane fractions from whole cells with a compatible buffer that allows MS detection of glycans was a crucial component of the methodology development. Although plasma membrane purification would have been a more rigorous approach, the sensitivity of the method at this point is still insufficient for the analysis of glycans from less than 50 million cells.

Ultracentrifugation was employed as the technique to isolate membrane fractions from cells. However, although mass spectrometry can be a powerful method for analyzing biomolecules because of its intrinsic speed and sensitivity, coupling the two techniques has proven challenging in part because of such incompatibilities in sample processing (16). For example, the typical buffer solution used in ultracentrifugation contains high concentration of detergents and solvents such as SDS, Triton X-100, EDTA, sucrose, sodium, and protease inhibitors, and several of these components are known to be deleterious to the mass spectrometry analyses.

The buffer solutions used in ultracentrifugation were thus optimized to render them more compatible with mass spectrometry. Homogenization buffer consisting of 0.25 M sucrose, 20 mM Hapes-KOH, 1 mM EDTA, 1% SDS, and a protease inhibitor mixture was initially used. However, in the initial MS analysis, only polymeric material with a regular mass spacing and non-carbon isotopic distribution were observed in supplemental Fig. S2, indicative of chemical background material only and an absence or complete suppression of the signal from glycans. To resolve this issue, the buffer conditions were reformulated. One key change was the removal of EDTA from the homogenization buffer, a change that did not affect the quality of the cellular fractionation step, but more importantly allowed observation of masses correlating to glycans in the mass spectra. Other changes included eliminating SDS and adding rigorous washes with nanopure water. The resulting purity of the membrane fractions was validated by SDS-PAGE gel electrophoresis followed by Western blotting using organelle-specific antibodies (supplemental Fig. S3).

Proteomic analyses were performed to determine the gross protein and membrane protein content in H1 stem cell. LC-MS/MS analyses identified a total of 335 proteins present in H1 membrane fraction. Two algorithms using HMMTOP2.0 and TMHMM2.0 were used to predict integral membrane proteins and transmembrane helices. Of the 335 proteins identified in the membrane fractions, 224 (59.1%) proteins contain at least one sequence predicted to be transmembrane helices by the HMMTOP (Hidden Markov Model for Topology Prediction) method (28), whereas 161 (42.5%) proteins contain at least one by the TMHMM2.0 (Tied Mixture Hidden Markov Model) method (29). We further performed Gene Ontology term analysis of the H1 stem cell membrane fraction using the Database for Annotation, Visualization and Integrated Discovery (DAVID) V6.7 (30) to determine the membrane portion relative to other cell parts. In the process, seven proteins were excluded because their gene symbols were not assigned. Of the remaining 321 proteins used in the Gene Ontology analysis, 217 proteins (67.6%) are categorized as cell membrane (supplemental Fig. S4). We examined these proteins further and found plasma membrane proteins that have been confirmed by earlier publications (31–33). These proteins are listed in supplemental Table S1 along with their potential sites of glycosylation based on known consensus sequences. The entire list of proteins identified in the membrane fraction is provided in the supplemental materials (Excel format). Included in the list are proteins categorized as belonging to membranes other than plasma membranes such as ribosome and hence endoplasmic reticulum, suggesting that there are contributions from other membranes in the fraction. However, because membranes sometimes share the same proteins, and the complete identification of cell membrane proteins is still ongoing, it may be difficult to provide conclusive determination as to what is and is not a cell membrane protein.

Mass Profiling of Glycans from hESC Cell Membranes—Global release methods were used to separate glycan from protein, including chemical and enzymatic to access O- and N-glycans, respectively. The exact glycan compositions of hexoses (Hex), N-acetylhexosamines (HexNAc), sialic acids

10.1074/mcp.M111.010660–4 Molecular & Cellular Proteomics 11.4
in both hES cell lines. In this analysis, the base peak, and high mannose glycans are in abundance of hES cell membranes of the two different hES cell lines HSF-6 and H1 are shown in Fig. 1. The putative structures of only the abundant species were assigned based on known glycobiology (or N-glycan biosynthesis) and tandem mass spectrometry; however, zooming in on the low abundance signals resulted in the identification of even more glycans. Overall, glycan mass profiling from H1 and HSF-6 cells is very similar. The m/z 1905.643 ([M + Na]⁺ corresponding to GlcNAc₂Man₉ (Man9) is the base peak, and high mannose glycans are in abundance in both hES cell lines. In this analysis, ~42 N-glycan compositions were identified on the hES cell membranes (supplemental Table S2).

To confirm glycan compositions, as well as obtain detailed structural information such as the putative glycan structures shown in Fig. 1, selected ions were subjected to tandem mass spectrometry using IRMPD. Tandem MS allows for the observance of ions arising from sequential loss of individual monosaccharides from the native glycan structure, thus supplementing the composition data obtained above via MALDI-FT-ICR MS with information on the connections of the monosaccharides to each other. One representative IRMPD analysis is shown for the ion at m/z 2067.687 ([M + Na]⁺), corresponding to 2HexNAc and 10Hex (9Man + 1Hex), for which tandem MS afforded extensive fragments in a single MS/MS event (supplemental Fig. S5). The glycosidic bond cleavages corresponding to the loss of the two core GlcNAc residues present in all N-glycans (m/z 1847 and 1643, respectively) were readily observed, along with fragments corresponding to subsequent monosaccharide losses (i.e. m/z 1847, 1685, 1523, 1361, and 1198 and m/z 1643, 1481, 1319, 1157, 995, and 833). The ion at m/z 833 corresponds to the trimannosyl core with two extra mannoses ([Man5-H₂O+Na⁺], confirming that this glycan is a high mannose type N-glycan. By contrast, the IRMPD spectrum of a complex type N-glycan (m/z 2012.719) is shown in supplemental Fig. S5.

O-Glycans were released from the membrane glycoproteins of hESCs by reductive β-elimination and enriched by SPE. The results indicate that O-glycans are present on hESC membranes, but at significantly lower abundances than N-glycans. The O-glycans are primarily neutral oligosaccharides comprised of mucin core type structures containing HexNAc and Hex residues. The dominant mass in three biological replicates, at relatively high abundances, corresponded to m/z 772 (2HexNAc:2Hex). All O-glycans found on the cell membrane commonly have a 2HexNAc:2Hex motif, indicating the presence of the core 1 structure (34). However, as shown in supplemental Fig. S6, the relatively high abundances of residual N-glycans even after treatment with peptide N-glycosidase F suppressed the signals from the O-glycans. With the somatic cell membranes, O-glycosylation was even less likely to be undetectable with this method. As such, this report focused on N-glycan profiling of hESC membranes.

**Mass Profiling of Glycans from Somatic Cell Membranes**—To determine whether the MS-based method employed in this study can be applied to other cell lines and whether somatic cells present high mannose glycans at the same abundance as hESCs, N-glycans were released from membrane fractions of IMR-90 fibroblast and MCF-10A human breast cell line and profiled by mass spectrometry. These somatic cells lines are commonly used, in part because they can go through several passages before cellular senescence (35, 36). The method provides sufficient cellular material necessary for MS analysis of membrane N-glycans but appears to provide few or no O-glycans, suggesting that O-glycans are significantly lower abundant than N-glycans, as previously mentioned. Glycans found in membranes of the two somatic cell lines were also summarized in supplemental Table S2. In general, the abundant species correspond to complex type structures for both cell lines, although high mannose glycans are more abundant in IMR-90 fibroblasts than in breast cell lines. For example, the complex type composition with m/z 1809.651 (5Hex:4HexNAc:1Fuc) is the base peak in IMR-90 fibroblast and the composition with m/z 2539.910 (7Hex:6HexNAc:1Fuc) is the base peak for MCF-10A breast cells. Both compositions are minor components in the stem cells. Therefore, complex type glycans containing Hex:HexNAc:Fuc:NeuAc are the major species found in human somatic cell lines, which contrast with hESCs.

**Isomer Separation and Quantitation of Cell Surface Glycans**—Glycan composition can be readily obtained based solely on MALDI-FT-ICR MS, and the abundances can also be determined with reasonable precision using peak intensities if the homogeneity of the sample surface can be guaranteed with the proper matrix (37, 38). However, this analysis does not provide information on the number of isomers associated with the compositions. Therefore, to analyze the isomers, glycans were further examined with nanoLC. Amicrochip packed with graphitized carbon was used to chromatographically separate the glycans in this study (39). The microchip was interfaced with TOF mass analyzer that routinely provides a mass measurement accuracy of less than 5 ppm. This technique achieves both isomer separation and direct quantitation with greater precision than MALDI MS.

Aliquots of each SPE fraction were combined, and the mixtures were analyzed by nanoLC MS in triplicate. Representative base peak chromatograms of N-glycans for four different cell membranes are shown in Fig. 2, with the putative structures for the most abundant glycans assigned to specific peaks, keeping in mind that each peak may still be composed of small mixtures. Both stem cells, H1 and HSF-6, show very similar base peak chromatograms and contrasts to the
FIG. 1. Representative MALDI-FT-ICR mass spectra of N-glycans found in hESC membranes in the positive detection ion mode. Glycans eluted from SPE fractions (10, 20, and 40% ACN) were combined prior to MS analysis. Glycan mass profiles of HSF-6 hESC (A) and H1 hESC (B) membranes are shown. The structures are putative and are based on accurate masses and tandem mass spectrometry.
IMR-90 and MCF-10A cell lines. In both human embryonic stem cell membranes, high mannose structures are the most abundant glycans and eluted earlier (between 14 and 20 min). IMR-90 and MCF-10A contain more complex type glycans, which elute later (between 20 and 28 min). In general, the mass profiles performed by MALDI-FT-ICR MS are consistent with the abundances from nanoLC/MS (Fig. 2, A and B, and supplemental Table S2). In addition, differentiation between IMR-90 and MCF-10A is also possible, because the profile of IMR-90 appears to be in intermediate between the stem cells and MCF10A.

We identified an average of 170 distinct features, which includes anomers and other isomers, arising from an average of 69 glycan compositions in H1 and HSF-6 hESC membranes (supplemental Table S3). Representative extracted ion chromatograms of high mannose glycans found on HSF-6 hESC membrane, GlcNAc₂Man₈ (Man8) and GlcNAc₃Man₇ (Man7), are shown in Fig. 3 to illustrate the separation and the number of isomers that may comprise a specific composition. The anomeric contribution caused by the separation on porous graphitized carbon (40) is small for these species so that the peaks correspond in these cases to linkage isomers. The extracted ion chromatogram of Man8 shows two structural isomers for m/z 861.302 ([M/H]⁺) with retention times of 15.66 and 16.66 min. Similarly, Man7 shows three structural isomers with retention times of 15.54, 16.44, and 16.56 min. The H1 hESC membrane also shows the same number of isomers from Man7 and Man8 but with different peak abundances (supplemental Fig. S7 and Table S3).

Chromatographic peak areas were used for overall glycan quantitation. N-Glycans are classified into two groups, namely high mannose and complex/hybrid type (non-high mannose glycan). The relative abundances of N-glycan types are depicted in Fig. 4A. Indeed, high mannose glycans are the most abundant in both H1 and HSF-6 hESCs, accounting for 74% in H1 and 85% in HSF-6 of the total N-glycans identified.

Fig. 2. Quantitation and isomer separation of N-glycans on cell membranes. The combined SPE fraction (10, 20, and 40%) was analyzed by nanoLC-Chip/TOF. Shown are representative base peak chromatograms of N-glycans found on four different cell membranes: HSF-6 hESC (A), H1 hESC (B), IMR-90 fibroblast (C), and normal breast cell line, MCF10A (D). Proposed structures corresponding to the most abundant glycan at that retention time are provided. An asterisk represents non-glycan peaks.
Cell Surface Glycome using Mass Spectrometry

Fig. 3. Extracted ion chromatogram of high mannose type glycans found in HSF-6 hESC. The left panel shows Man7 isomers, and the right panel shows Man8 isomers.

whereas complex/hybrid type N-glycans comprise the remaining 26% in H1 and 15% in HSF-6, respectively. Man8 and Man9 are the major glycans present in two hESC membranes, accounting for more than 50% of total high mannose glycans (Fig. 4B). However, human IMR-90 fibroblasts and MCF-10A breast cells had significantly lower levels of high mannose glycans (45 and 28%, respectively) compared with hESCs. For the MCF-10A cells, the complex type glycans dominate, accounting for more than 70% (Fig. 4B), whereas for IMR-90, it is 55%. Finally, the complex/hybrid type glycans are generally fucosylated in both hESC lines, with more than 70% of such glycans containing at least one fucose residue (Fig. 4B). For IMR-90 and MCF-10A, the degree of fucosylation in the complex type glycans is 76 and 68%, respectively (data not shown).

The reproducibility between biological replicates and analytical replicates was examined. Biological replicates consisted of hESCs collected on different days from different passage numbers, whereas analytical replicates were from the same biological sample analyzed multiple times. For the analytical replicates, normalized absolute peak intensities measured by three MS experiments were plotted against one other (supplemental Fig. S8). High correlation coefficients ($r = 0.96–0.97$) were obtained from three analytical replicates. The variation among biological experiments was further examined. The correlation coefficients ($r$) between the normalized intensity of glycan peaks were in the range of 0.86–0.92, showing good correlation between biological replicates (supplemental Fig. S8B). The results clearly show that profiling cell membrane glycans by MS is highly reproducible.

Validation of Glycan Expression Using Lectins—The majority of glycans found on the surface of hESCs via the MS approach were high mannose N-glycans. To determine whether these types of glycans, observed in the general cellular membrane fractions, are specifically found on hESC surfaces, hESCs were labeled with lectins: plant and animal cell surface glycosylation and provide extensive heterogeneity of membrane proteins followed by high performance mass spectrometry.

In this study, we developed a method to measure cell surface glycosylation and provide extensive heterogeneity of the hESC glycome from membrane proteins. Enrichment of membrane proteins followed by high performance mass spectrometry.
Cell Surface Glycome using Mass Spectrometry

spectrometry and chromatography makes this approach both highly specific and sensitive. The resulting approach provides comprehensive and highly quantitative structural information including isomer separation.

Stem cell surface glycosylation is dominated by high mannose glycans (Fig. 4A). These results are consistent with both the lectin cell surface studies by Wearne et al. (12) and the glycomic profile of whole hESC by Satomaa et al. (25). The major difference between this study and the previous two is that we enrich plasma membrane and determine the glycosylation specifically from the membrane. We confirm that the products we obtain are primarily from the membrane. We then confirm the results with lectins and flow cytometry analyses. Furthermore, we quantitate the relative contribution of each glycan types as well as specific isomers on the cell surface. These experiments differ from those by Wearne et al. (12) and Satomaa et al. (25) in that those were performed on whole cell lysates. The cell surface glycosylation were inferred indirectly using lectins, which are not quantitative when comparing different glycan types. Therefore, although high mannose may show on surfaces with lectins, so would complex and hybrid types. There is no method using lectins to determine whether high mannose is more abundant compared with the other glycan types. Glycans from whole cells were indeed examined by MS and by NMR, but there were no efforts to determine the relative concentrations of each glycan types on the cell surface.

The amount of high mannose is unprecedented, consisting of ~75% of the total N-glycan species, which is the most abundant type of glycosylation. Moreover, hESCs are rich in Man8 and Man9 glycans (Fig. 4B) and contain a large amount of Man9 + Glc, suggesting that hESC membranes have an unprecedented amount of terminal glucose. These results contrast with fully differentiated cell lines that are richer in complex type glycans, further suggesting new types of glycan markers for determining hESC cells.

The high abundance of high mannose glycans on hESCs further contrasts with the levels, for example, found in blood and human somatic cell lines, where complex types are by far the most abundant, and high mannose glycans are very minor components (Fig. 4) (40). More generally, high mannose glycans are not commonly observed on the surfaces of mammalian cells (41), with the exception of macrophages (42). High mannose glycans are involved in initial steps in N-glycan processing and control of protein folding in the endoplasmic reticulum. The initiating event is the processing of N-glycan precursor, Man9, and three glucose by glucosidases I and II to yield to Man9 + glucose, which is subsequently trimmed back in the process of generating complex type glycans (43). If the membrane fractions were primarily plasma membrane, the abundance of Man9 suggests that the trimming process is incomplete in hESCs, perhaps yielding “immature” glycoproteins. Indeed, the ion at m/z 2067,68, corresponding to Man9 with one extra hexose, potentially a glucose, was readily identified by tandem MS (supplemental Fig. S5). Because the membrane extraction included all types of membranes from the cell, glycans derived from the endoplasmic reticulum were included in the analyses. However, human somatic cell lines such as IMR-90 fibroblasts and MCF-10A breast cells had significantly lower levels of high mannose glycans compared with hESCs, and in particular MCF-10A cells had predominantly complex type glycans. Interestingly, IMR-90 fibroblasts, which are a cell type derived from human fetal lung (35), also have high mannose glycans on the cell surface, assessed by flow cytometry (data not shown). It is unclear whether this represents an “embryonic” characteristic of IMR-90 fibroblasts that is shared with hESCs or possibly whether high mannose glycans are serving a similar function in both type of cells. In contrast to IMR-90s, MCF-10As are a cell type derived from adult breast epithelium (36).

Contrary to mammalian cells, yeast and viruses have high mannose glycans on the cell surface. Cells in the human immune system such as macrophages express mannose receptors that can bind to terminal/high mannose glycans on the surfaces of non-self entities (e.g. bacteria, yeast, and viruses) to detect and initiate pathogen phagocytosis. Thus, investigations of the role of high mannose glycans have been mostly directed to their involvement in initiating the innate immune system or in protein folding in the endoplasmic reticulum. A few studies, however, have suggested that high mannose glycans mediate cell-cell fusion (42, 44), mediating sperm-egg fusion, myoblast fusion, and osteoclast formation. Thus, high mannose glycans on hESCs may play a role in cellular binding and recognition. More recently, our group and others have shown that high mannose glycans are prevalent on the cell surface of various tumor cells compared with normal cells (45, 46). Understanding the role of high mannose glycans on hESCs could be applicable to cancer biology because similar characteristics have been observed between tumor and stem cells, such as the ability to self-renew, the expression of cell surface markers, and the activation of signaling pathways (47).

10.1074/mcp.M111.010660–10

Molecular & Cellular Proteomics 11.4
Further analyses will be performed to identify the proteins that high mannose glycans are modifying and the sites of attachment, to discern whether high mannose glycans are functionally modifying cell surface proteins and the functional significance on both hESCs compared with somatic cell lines. Also, future studies include investigating changes in high mannose glycans on the cell surface during differentiation, as well as confirming that the glycans come only from surface proteins and not from contamination by the endoplasmic reticulum, although the same contamination should also be present in the other cell lines.

If indeed the glycan structures come primarily from the surface proteins, these findings suggest that the abundance and presence of high mannose glycans specifically on the cell surface may have biological significance. However, it has immediate consequences. It suggests that stem cells have...
less sialylation than differentiated cells, because sialylation accompanies complex and hybrid but are not found in high mannose glycans. Furthermore, the presence of Man9 + glucose in high abundances also suggests the unprecedented presence of terminal glucose residues on the surfaces. Thus, the deep structural analysis enabled by this methodology motivates future mechanistic studies on the biological significance of cell surface glycosylation.

Acknowledgments—We thank Dr. Jong Shin Yoo for assistance in proteomics data analysis.

H. J. A., P. G., and C. B. L. designed the overall research. P. G., D. V. S., and C. R. B. designed the hESCs cells culture, membrane isolation, lectin assays, and immunocytochemistry of hESCs. H. J. A. did all of the MS analyses. S. W. performed the nanoLC-Chip/TOF analyses. J. H. K. did data analysis for peak normalization and reproducibility. G. W. P. did proteomics data analysis. H. J. A., P. G., C. T. M., and C. B. L. wrote the paper.

* This work was supported by National Institutes of Health Grants RO1GM049077 (to C. B. L.) and GM6604? (to C. R. B.), California Institute for Regenerative Medicine Grant RS1-00365 (to C. R. B.), and the Converging Research Center Program through Ministry of Education, Science and Technology Grant 2011K000968 (to H. J. A.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

This article contains supplemental material.

These authors contributed equally to this work as co-first authors.

Supported by the California Institute for Regenerative Medicine and the National Science Foundation and by California Institute for Regenerative Medicine predoctoral fellowships.

These authors contributed equally to this work.

To whom correspondence should be addressed. Tel.: 530-752-0504; Fax: 530-752-8995; E-mail: cblebrilla@ucdavis.edu.

REFERENCES

1. 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

2. Lebrilla, C. B., and Mahal, L. K. (2009) Post-translation modifications. Curr. Opin. Chem. Biol. 13, 373–374

3. Ohitsu, K., and Marth, J. D. (2006) Glycosylation in cellular mechanisms of health and disease. Cell 126, 855–867

4. Paulson, J. C., Blixt, O., and Collins, B. E. (2006) Sweet spots in functional glycomics. Nat. Chem. Biol. 2, 238–248

5. Dube, D. H., and Bertozzi, C. R. (2005) Glycans in cancer and inflammation: Potential for therapeutics and diagnostics. Nat. Rev. Drug Discov. 4, 477–488

6. Cooke, C. L., An, H. J., Kim, J., Canfield, D. R., Torres, J., Lebrilla, C. B., and Solnick, J. V. (2009) Modification of gastric mucin oligosaccharide expression in rhesus macaques after infection with Helicobacter pylori. Gastroenterology 137, 1061–1071

7. An, H. J., Kronewitter, S. R., de Leoz, M. L., and Lebrilla, C. B. (2009) Glycomics and disease markers. Curr. Opin. Chem. Biol. 13, 601–607

8. Fuster, M. M., and Esko, J. D. (2005) The sweet and sour of cancer: Glycans and cancer progression. Nat. Chem. Biol. 1, 238–248

9. Wearne, K. A., Winter, H. C., O’Shea, K., and Goldstein, I. J. (2006) Use of lectins for probing differentiated human embryonic stem cells for carbohydrates. Glycobiology 16, 981–990

10. Wollscheid, B., Bausch-Fluck, D., Henderson, C., O’Brien, B., Bibel, M., Schiess, R., Aebersold, R., and Watts, J. D. (2009) Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. Nat. Biotechnol. 27, 378–386

11. Ito, H., Kuno, A., Sawaki, H., Mogabe, M., Ozaki, H., Tanaka, Y., Mizokami, M., Shoda, J., Angata, T., Sato, T., Hirabayashi, J., Ikehara, Y., and Narimatsu, H. (2009) Strategy for glycoproteomics: Identification of glyco-alteration using multiple glycan profiling tools. J. Proteome Res. 8, 1358–1367

12. McNeil, J. (2004) Embryonic stem cells in drug discovery. Nat. Rev. Drug Discov. 3, 70–80

13. Lanctot, P. M., Gage, F. H., and Varli, A. P. (2007) The glycans of stem cells. Curr. Opin. Chem. Biol. 11, 373–380

14. Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. W., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., and Jones, J. M. (1998) Embryonic stem cell lines derived from human blastocysts. Science 282, 1145–1147

15. Northwestern, D. C., and Lebrilla, C. B. (2006) A new computer program (GlycoX) to determine simultaneously the glycosylation sites and oligosaccharide heterogeneity of glycoproteins. J. Proteome Res. 5, 2800–2808

16. Kornwitter, 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

17. Sasame, T., and Simon, I. (1998) Principles governing amino acid composition of integral membrane proteins: Application to topology prediction. J. Mol. Biol. 283, 489–506

18. Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 175–182

19. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57

20. Feller, R., Rossier, G., Spindler, B., Meier, C., Kühn, L., and Verrey, F. (1999) Amino acid transport of γ-L-type by heterodimers of 4F2hc/CD98 family. EMBO J. 18, 49–57

21. Zhang, H., Li, X. J., Martin, D. B., and Aebersold, R. (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat. Biotechnol. 21, 660–666

22. Kayano, T., Fukushima, H., Eddy, R. L., Fan, Y. S., Byers, M. G., Shows, T. B., and Bell, G. I. (1988) Evidence for a family of human glucose
transporter-like proteins: Sequence and gene localization of a protein expressed in fetal skeletal muscle and other tissues. J. Biol. Chem. 263, 15245–15248
34. Brockhausen, I. (1997) Biosynthesis and functions of O-glycans and regulation of mucin antigen expression in cancer. Biochem. Soc Trans 25, 871–874
35. Nichols, W. W., Murphy, D. G., Cristofalo, V. J., Toji, L. H., Greene, A. E., and Dwight, S. A. (1977) Characterization of a new human diploid cell strain, IMR-90. Science 196, 60–63
36. Soule, H. D., Maloney, T. M., Wolman, S. R., Peterson, W. D., Jr., Brenz, R., McGrath, C. M., Russo, J., Pauley, R. J., Jones, R. F., and Brooks, S. C. (1990) Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. Cancer Res. 50, 6075–6086
37. Harvey, D. J. (1993) Quantitative aspects of the matrix-assisted laser-desorption mass-spectrometry of complex oligosaccharides. Rapid Commun. Mass Spectrom. 7, 614–619
38. Grey, C., Edebrink, P., Krook, M., and Jacobsson, S. P. (2009) Development of a high performance anion exchange chromatography analysis for mapping of oligosaccharides. J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 877, 1827–1832
39. Ninñou, M., An, H., Yin, H., Killeen, K., Miyamoto, S., Grimm, R., and Lebrilla, C. B. (2009) Nanoliquid chromatography-mass spectrometry of oligosaccharides employing graphitized carbon chromatography on microchip with a high-accuracy mass analyzer. Electrophoresis 26, 3641–3649
40. Chu, C. S., Ninñou, M. R., Clowers, B. H., Perkins, P. D., An, H. J., Yin, H., Killeen, K., Miyamoto, S., Grimm, R., and Lebrilla, C. B. (2009) Profile of native N-linked glycan structures from human serum using high performance liquid chromatography on a microfluidic chip and time-of-flight mass spectrometry. Proteomics 9, 1939–1951
41. 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
42. Morishima, S., Morita, I., Tokushima, T., Kawashima, H., Miyasaka, M., Omura, K., and Murota, S. (2003) Expression and role of mannos receptor/terminal high-mannose type oligosaccharide on osteoclast precursors during osteoclast formation. J. Endocrinol. 176, 285–292
43. Molinari, M. (2007) N-Glycan structure dictates extension of protein folding or onset of disposal. Nat. Chem. Biol. 3, 313–320
44. Kurachi, T., Morita, I., Oki, T., Ueki, T., Sakaguchi, K., Enomoto, S., and Murota, S. (1994) Expression on outer membranes of mannos residues, which are involved in osteoclast formation via cellular fusion events. J. Biol. Chem. 269, 17572–17576
45. Johns, T., G., Mellman, I., Cartwright, G. A., Ritter, G., Old, L. J., Burgess, A. W., and Scott, A. M. (2005) The antitumor monoclonal antibody 806 recognizes a high-mannose form of the EGF receptor that reaches the cell surface when cells over-express the receptor. FASEB J. 19, 780–782
46. de Leoz, M. L., Young, L. J., An, H. J., Kronewitter, S. R., Kim, J., Miyamoto, S., Borowsky, A. D., Chew, H. K., and Lebrilla, C. B. (2011) High-mannose glycans are elevated during breast cancer progression. Mol. Cell. Proteomics 10, M110.002717
47. Reya, T., Morrison, S. J., Clarke, M. F., and Weissman, I. L. (2001) Stem cells, cancer, and cancer stem cells. Nature 414, 105–111

In order to cite this article properly, please include all of the following information: 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(4):M111.010660. DOI: 10.1074/mcp.M111.010660.