⌘ Author's Choice

© 2011 by The American Society for Biochemistry and Molecular Biology, Inc.
This paper is available on line at http://www.mcponline.org

Targeted Identification of Metastasis-associated Cell-surface Sialoglycoproteins in Prostate Cancer*[S]

Lifang Yang‡, Julius O. Nyalwidhe‡, Siqi Guo§, Richard R. Drake‡, and O. John Semmes‡¶

Covalent attachment of carbohydrates to proteins is one of the most common post-translational modifications. At the cell surface, sugar moieties of glycoproteins contribute to molecular recognition events involved in cancer metastasis. We have combined glycan metabolic labeling with mass spectrometry analysis to identify and characterize metastasis-associated cell surface sialoglycoproteins. Our model system used syngeneic prostate cancer cell lines derived from PC3 (N2, nonmetastatic, and ML2, highly metastatic). The metabolic incorporation of AC_4ManNAz and subsequent specific labeling of cell surface sialylation was confirmed by flow cytometry and confocal microscopy. Affinity isolation of the modified sialic-acid containing cell surface proteins via click chemistry was followed by SDS-PAGE separation and liquid chromatography-tandem MS analysis. We identified 324 proteins from N2 and 372 proteins of ML2. Using conservative annotation, 64 proteins (26%) from N2 and 72 proteins (29%) from ML2 were classified as extracellular or membrane-associated glycoproteins. A selective enrichment of sialoglycoproteins was confirmed. When compared with global proteomic analysis of the same cells, the proportion of identified glycoprotein and cell-surface proteins were on average threefold higher using the selective capture approach. Functional clustering of differentially expressed proteins by Ingenuity Pathway Analysis revealed that the vast majority of glycoproteins overexpressed in the metastatic ML2 subline were involved in cell motility, migration, and invasion. Our approach effectively targeted surface sialoglycoproteins and efficiently identified proteins that underlie the metastatic potential of the ML2 cells. ***Molecular & Cellular Proteomics 10: 10.1074/mcp.M110.007294, 1–16, 2011.***

From the ‡Leroy T. Canoles Cancer Research Center, Department of Microbiology and Molecular Cell Biology, Eastern Virginia Medical School, Norfolk, Virginia VA 23507; §Frank Reidy Research Center for Bioelectrics, Old Dominion University, Norfolk, VA 23508

Received December 22, 2010, and in revised form, February 18, 2011

⌘ *Author's Choice*—Final version full access.

Published, MCP Papers in Press, March 29, 2011, DOI 10.1074/mcp.M110.007294

Covalent attachment of carbohydrates to proteins is one of the most common post-translational modifications with more than 50% of eukaryotic proteins thought to be glycosylated (1). As would be expected, protein glycosylation is prevalent in membrane and secreted proteins where they play a decisive role in cellular recognition events involved in cell adhesion, cell-to-cell communication, and receptor-ligand interactions (2–4). Glycans also critically influence the physiochemical properties of proteins that impact protein folding, solubility, and turnover (5, 6). As a direct result of its importance in a variety of biological and pathological processes, changes in glycosylation of proteins are strongly correlated with cancer prognosis and the malignant properties of tumor cells. This includes changes in the levels of sialylation and fucosylation, polylactosaminylation, higher-ordered *N*-linked branching, and truncated O-linked glycans (7–10).

Cell-surface glycoproteins commonly contain sialic acid (SA)[1] as the monosaccharide located on the nonreducing terminus of glycans. Evidence from both patient histochemical analysis and experimental tumor models demonstrate that altered sialylation of tumor cell surfaces is associated with a metastatic tumor phenotype (11, 12). These surface sialylation changes have been reported reflecting the amount, type, distribution, and bonding of sialic acids to adjacent molecules. For instance, a positive correlation can be established between the levels of cell-surface sialylation and metastatic ability of various experimental tumors (13, 14). In addition, the distribution of sialic acid on specific *N*- or *O*-linked oligosaccharides has been demonstrated to alter the metastatic potential of cancer cell lines (15). In some cases tumor development has been associated with over expression of β1–6 branching that results in increased number of Galβ1–4GlcNAc structures available for sialylation (16–18). Therefore, exploring cell-surface sialylation changes during tumor development and disease progression likely affords excellent

[1] The abbreviations used are: SA, sialic acid; LC, liquid chromatography; MS/MS, tandem mass spectrometry; DMEM, Dulbecco's modified Eagle's medium; PBS, phosphate-buffered saline; AC_4ManNAz, tetraacetylated N-azidoacetyl-D-mannosamine; ManNAc, N-acetyl-D-mannosamine; SAv, streptavidin; FITC, fluorescein isothiocyanate; IPA, Ingenuity Pathway Analysis.

opportunities to identify sensitive and specific cancer biomarkers.

Elucidation of structural details of cell-surface glycosylation by mass spectrometry is hampered by the limited relative abundance of surface proteins compared with cytosolic components, the complex and microheterogeneous nature of glycans, and the inherent complexities of deciphering low energy carbohydrate fragmentation ions *versus* higher energy peptide fragments in complex mixtures (19). The type of mass spectrometer and ionization energies to be used, and the complexity of the sample, are critical parameters for successful glycoprotein analysis (19). In recent years, lectin- and antibody-based affinity selection has been used with some success to purify glycoproteins and glycopeptides with specific structures (20–24). Other approaches for glycopeptide characterization arise from the exploitation of glycan chemical reactivity. Larsen *et al.* took advantage of the high affinity of titanium dioxide microcolumns toward SA residues to isolate SA-containing peptides from serum under highly acidic conditions (25). Two similar approaches, involving hydrazide and boronic acid chemistry, capitalize on the *cis*-diols present in monosaccharides. Zhang *et al.* described the use of hydrazide chemistry for purification by directly coupling of glycoproteins to a solid support (26, 27). Similarly, Sparbier *et al.* used boronic acid -functionalized beads to covalently capture glycoproteins followed by elution with acid (28). Although these methods are effective at the enrichment and identification of broad classes of glycoproteins/glycopeptides, they still lack the specificity and selectivity required for analysis of specific cell surface glycoproteins that could serve as potential cancer biomarkers.

In this study, we describe a glycoproteomic identification strategy for the selective detection, isolation and identification of cell-surface sialoglycoproteins from cultured cell lines. The method utilizes the sialic acid biosynthetic pathway for the incorporation of monosaccharide bearing bioorthogonal functional handles (tetraacetylated N-azidoacetyl-D- mannosamine) into cellular sialic acid (29–31). These reagents have previously been used to label and visualize cell surface expression of glycoproteins via microscopy. To illustrate the potential of using this cell labeling procedure in biomarker discovery, we combined it with an MS-based proteomics approach as applied to a syngeneic metastatic prostate cancer cell line model.

EXPERIMENTAL PROCEDURES

Materials—Complete™ protease inhibitors were purchased from Roche Applied Sciences (Indianapolis, IN), sequencing grade trypsin was from Promega (Madison, WI), and Immobilon-FL PDVF membrane was from Millipore (Billerica, MA). Protein-free blocking buffer and high capacity streptavidin agarose resin was from Thermo Scientific (Rockford, IL). Dulbecco's modified Eagle media (DMEM) and fetal bovine serum (FBS) were from Invitrogen (Carlsbad, CA). Antibiotics-antimycotic, Click-iT™ ManNAz metabolic glycoprotein reagent, Click-iT Biotin Protein Analysis Detection Kit, propidium iodide, fluorescein conjugated streptavidin (streptavidin-FITC), 4–12% NuPAGE® Bis-Tris gels, and lithium dodecyl sulfate buffer were from Invitrogen (Carlsbad, CA). 2× Laemmli buffer, 7.5% Criterion® Tris-HCl Gel, and Bio-Safe Coomassie blue were from Bio-Rad (Hercules, CA). The streptavidin-IR 800 and species-specific secondary antibodies conjugated to IR 800 or IR700 were from Li-COR Biosciences (Lincoln, PA).

Cell Lines—PC3-N2 and PC3-ML2, two sublines of PC3 prostate cancer cells, were kindly provided by Dr. Mark Stearns (Drexel University). The PC3 cell line was originally derived from a skeletal metastasis in a patient with primary prostate adenocarcinoma. The N2 and ML2 cell lines have been developed on the basis of their invasiveness *in vitro* and metastatic potential *in vivo*. Both cells were tumorigenic when injected subcutaneously in Severe combined immunodeficiency mice. However, N2 cells were unable to migrate through a Matrigel-coated membrane *in vitro* as well as induce metastases in severe combined immunodeficiency mice, whereas ML2 cells were highly invasive *in vitro* and induced skeletal metastases in more than 80% of mice (32).

Cell Culture and Metabolic Labeling—N2 and ML2 cells were cultured in DMEM medium supplemented with 10% FBS and 1% antibiotics at 37 °C with 5% CO_2. For metabolic labeling, growth medium was replaced at 70% cell confluence with complete DMEM medium containing the indicated concentration of an azido-modified sugar, tetraacetylated N-azidoacetyl-D- mannosamine (AC_4ManNAz), or a control sugar, *N*-acetyl-D-mannosamine (ManNAc), and cells were incubated for 1–3 days. Cells then were dissociated from the plastic surface by nonenzyme dissociation buffer.

Flow Cytometric Analysis of Cell-surface Sialic Acid Labeling—After metabolic labeling, N2 and ML2 cells were harvested, washed with 0.1% FBS/phosphate-buffered saline (PBS), and resuspended (10^6 cells) in 100 μl click reaction solution with the indicated amount of each component. The reaction was incubated at room temperature for 30 min, and then cells were washed three times with 0.1% FBS/PBS. Cells were subsequently stained with streptavidin-FITC (1 μg/sample in 100 μl 2% FBS/PBS) for 30 min at 4 °C, and washed with 2% FBS/PBS three times. Prior to flow cytometric analysis, cells were incubated with 1 μg/ml propidium iodide in 500 μl 2% FBS/PBS at 4 °C for 20 min. Data was acquired by FACScalibur (BD Biosciences, San Jose, CA) and analyzed by Flowjo software (Tree Star Inc., Ashland, OR).

Confocal Microscopy Analysis of Cell-surface Sialic Acid Labeling—N2 and ML2 cells were seeded onto glass coverslips in 6-well plates containing 10% FBS/DMEM. Growth medium was supplemented with 40 μM AC_4ManNAz or ManNAc for 3 days. Cells were washed with ice-cold PBS, fixed with 2% paraformaldehyde, and then subjected to click reaction solution (25 μl biotin-alkyne, 2.5 μl CuSO4, 2.5 μl and 5 μl for component D and E, 65 μl PBS). Subsequently, the fixed and labeled cells were washed with PBS and stained with streptavidin-FITC and the nuclei were counterstained with propidium iodide. The coverslips were inverted onto glass slides and mounted with VectorShield medium (Vector Labs, Burlingame, CA), and sealed with nail polish. Fluorescent images were examined and captured with a Zeiss confocal microscope.

Cell-surface Sialoglycoprotein Labeling and Detection in Cell Extracts—N2 and ML2 cells were seeded in 15-cm dishes and treated with the optimized labeling conditions for AC_4ManNAz and ManNAc (20 μM for 1 day) in growth medium. The cells were then harvested with nonenzymatic dissociation buffer, 5×10^7 cells were collected and washed with 0.1% FBS/PBS. Harvested cells were subjected to click reaction solution as described above. After chemical conjugation, cell pellets were washed with ice-cold PBS twice to remove unreacted reagents and then lysed in lysis buffer I (1% Nonidet P-40, 150 mM NaCl, protease inhibitor, 100 mM sodium phosphate, pH7.5)

using a Dounce homogenizer. The total cell lysate was further cleared by methanol-chloroform precipitation and resolved in lysis buffer II (1% SDS, protease inhibitor, 50 mM Tris-HCl, pH 8.0). Protein concentrations were measured using the BCA protein assay (Pierce). To detect biotin-labeled sialoglycoproteins in cell extracts, 20 μg of labeled protein lysate was resolved by SDS-PAGE. Electrophoresed proteins were transferred onto PVDF membrane, blocked with Odyssey Blocking Buffer (Rockland Immunochemicals, Gilbertsville, PA), probed with streptavidin-IR 800, and visualized and quantified using an Odyssey infrared imaging system.

Sialoglycoprotein Capture—Streptavidin beads were pretreated with protein-free blocking buffer overnight at 4 °C and washed five times with PBS. Cell lysate (2 mg) was incubated with 100 μl beads in 0.3% Nonidet P-40/PBS overnight at 4 °C with a rotating shaker. The captured glycoproteins were washed intensively with modified RIPA buffer (150 mM NaCl, 2% SDS, 1% Nonidet P-40, 1% Na Deoxycholate, 50 mM Tris-HCl, pH 7.5). Bound material was eluted by boiling for 10 min in 100 μl 2× Laemmli sample buffer. For assessment of capture efficiency, 10 μl of the eluent, along with the input and flow-through fractions were separated by SDS-PAGE and probed with streptavidin-IR 800 by Western blot.

One-dimensional Gel Electrophoresis and In-gel Digestion—Captured glycoproteins (45 μl) were separated on a 1.0 mm 7.5% Tris-HCl polyacrylamide gels. Gels were stained with Coomassie blue and imaged on a Typhoon 9410 (GE Healthcare, Piscataway, NJ). Twenty-seven equally spaced gel pieces were excised from each lane, spanning the full height of the gel (30–300 kDa). Individual gel pieces were destained with 25 mM NH_4HCO_3 in 50% acetonitrile (ACN), reduced with 20 mM dithiothreitol in 25 mM NH_4HCO_3 for 45 min at 56 °C, and alkylated with 55 mM iodoacetamide in 25 mM NH_4HCO_3 in the dark for 1 h at room temperature. After washing, the gel pieces were dehydrated with ACN and dried using a speed vac. Trypsin (12.5 ng/μl in 25 mM NH_4HCO_3, pH 8.0) was added to each gel piece, and the gel pieces were allowed to swell on ice for 1 h. Excess trypsin was removed, replaced with 25 mM NH_4HCO_3, 10% ACN, pH 8.0, and the mixture was incubated overnight at 37 °C. The digest was collected and peptides were extracted 2X with 50% ACN/0.5% formic acid. The combined extracts were then dried using a SpeedVac for subsequent LC-MS/MS analysis.

Whole-cell Extract and In-solution Digestion—For N2 and ML2 cell lysate experiments; both cell lines were cultured in complete growth medium to 80% confluence. Cells were then washed with PBS and the pellet resuspended in 160 μl dissolution buffer containing 100 mM NH_4HCO_3 and TFE (1:1 v/v). The samples were sonicated 3× for 20 s and incubated at 60 °C for 1 h. The lysates were centrifuged to remove cell debris and unbroken cells and the supernatant was collected. The protein concentration of the samples was determined by BCA assay and normalized for each sample and before reduction and alkylation with TCEP and iodoacetamide respectively. The concentration of TFE was reduced to 5% by the addition of 1.4 ml 100 mM NH_4HCO_3. Trypsin was added at a ratio of 1:50 protease to protein and the digestion proceeded at 37 °C for 18 h with constant mixing. After digestion the sample was dried down in a SpeedVac, desalted and lyophilized before liquid chromatography-tandem MS (LC-MS/MS) analysis.

LC-MS/MS Analysis—Digests were resuspended in 20 μl Buffer A (5% ACN, 0.1% formic acid, 0.005% heptafluorobutyric acid) and 15 μl loaded onto a 12-cm × 0.075 mm fused silica capillary column packed with 5 μM diameter C-18 beads (The Nest Group, Southborough, MA) using a programmed automatic injection. Peptides were eluted over 80 min, by applying a 0–80% linear gradient of Buffer B (95% acetonitrile, 0.1% formic acid, 0.005% HFBA) at a flow rate of 200 μl/min with a precolumn flow splitter resulting in a final flow rate of ~300 nl/min directly into the source. In some cases, the gradient was extended to 150 min to acquire more MS/MS spectra. An LTQ™ Linear Ion Trap (ThermoFinnigan, San Jose, CA) was run in an automated collection mode with an instrument method composed of a single segment and five data-dependent scan events with a full MS scan followed by four MS/MS scans of the highest intensity ions. Normalized collision energy was set at 35, activation Q was 0.250 with minimum full scan signal intensity at 1×10^5 with no minimum MS^2 intensity specified. Dynamic exclusion was turned on utilizing a three-minute repeat count of 2 with the mass width set at 1.0 *m/z*. Peak lists were generated using Xcaliber (version 2.1). Sequence analysis was performed with MASCOT (version 2.2.03) using the SwissProt 57.1 database with a human taxonomy filter enabled that contained 20,405 entires. The database searches were performed with fixed modification as carbamidomethyl (Cys) and variable modifications as oxidation (Met) and deamidation (Asn, Gln). Enzyme specificity was selected to trypsin with 1 missed cleavage. The mass tolerance was set at 1 for precursor ions and 0.8 for fragment ions. Threshold score for acceptance of individual spectra was set at 0.05. All the MS/MS spectra were manually inspected to verify the validity of the database search results. False discovery rates were estimated to be 0.25% on the protein level by searching a decoy version of the SwissProt protein database. The relative abundance of peptides was estimated by spectral counting.

Deglycosylation Assays—Enymatic deglycosylation of samples was performed prior to analysis by western. For PNGase F digestions, cell lysates (20 μg) were denatured and reduced for 10 min at 100 °C in 0.5% SDS and 1% 2-mercaptoethanol. The samples were then adjusted to 1% Nonidet P 40 and 50 mM sodium phosphate, pH 7.5, and incubated with PNGase F (1,000 NEB units) overnight at 37 °C. The experiments using neuraminidase reactions were adjusted to 50 mM sodium acetate, pH 6.0 prior to incubation overnight at 37 °C with 50 NEB units neuraminidase per 20 μg of protein. Deglycosylation digestion of Fetuin was used to optimize conditions (data not shown).

Western Blot Analysis—Whole-cell lysates were collected in M-PER lysis buffer containing 1x protease inhibitor mixture. The protein concentration was determined by the BCA protein assay. Cell lysates, streptavidin pull-down fractions and deglycosylated protein samples were separated by electrophoresis through 4–12% or 7.5% SDS-PAGE and then transferred to Immobilon-FL PDVF membranes. Membranes were blocked in LiCor blocking buffer (Rockland Immunochemicals, Gilbertsville, PA) diluted with PBS (1:1), then incubated with anti-CDCP1 polyclonal (#4115, 1:1000, Cell Signaling Technology, Danvers, MA), or anti-integrin β1 mouse monoclonal (#610467, 0.1 μg/ml, BD PharMingen, San Diego, CA) primary antibodies overnight at 4 °C. After 5× washes in PBS, membranes were incubated with species-specific (goat anti-rabbit, 1:15,000; goat anti-mouse, 1:15,000) IRDye700- or IRDye800-conjugated secondary antibodies for 1 h at room temperature, and visualized with a LiCor Odyssey infrared imager (LiCor, Lincoln, NE). Consistent protein loading was determined by reprobing membranes stripped in Restore Western blot stripping buffer with anti-actin antibody (#612656, 0.1 μg/ml, BD PharMingen, San Diego, CA) or anti-GAPDH antibody (#25778, 1:2000, Santa Cruz Biotechnology, Santa Cruz, CA).

Bioinformatic Analysis—Initial basic information on the structure of the identified proteins was obtained through literature reports. The ProteinID Finder (Proteome Solutions) program was employed to extract such available information from the UniProt database. Because of the limited annotation of protein glycosylation for most human proteins and the shortage of subcellular location information for hypothetical proteins and functionally uncharacterized proteins, we also subjected each identified protein to prediction algorithms based on protein sequence analysis. For protein glycosylation we used NetNGlyc (http://www.cbs.dtu.dk/services/NetNGlyc/) to predict the possible presence of the consensus NXS/T glycosylation

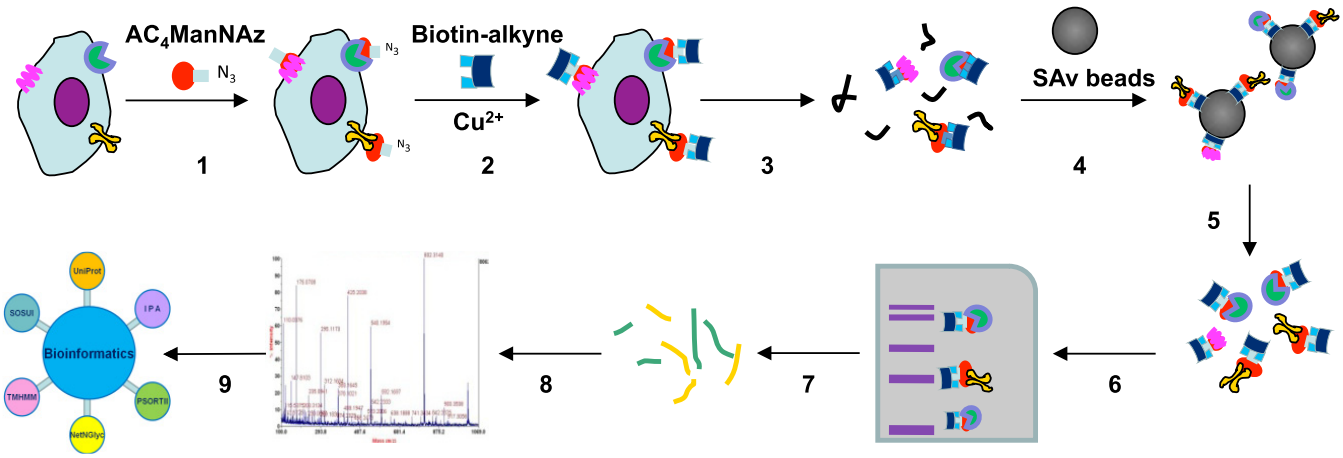

FIG. 1. **Experimental workflow**. (*1*) Metabolic labeling of cells with the mannose analog, peracetylated azido-mannose (AC_4ManNAz). (*2*) Chemoselective conjugation of azido sugars with a biotinylated alkyne capture reagent via Cu (I) catalyzed click chemistry. (*3*) Lysis of labeled cells. (*4*) Affinity purification using streptavidin (SAv) resins. (*5*) Elution of captured sialoglycoproteins. (*6*) SDS-PAGE separation of sialoglycoproteins. (*7*) Isolation of gel slice and subsequent digestion and release of peptides. (*8*) Analysis of peptides by LC-MS/MS. (*9*) Bioinformatic analysis.

motif. For subcellular location, all identified proteins were analyzed with two transmembrane prediction algorithms SOSUI (http://bp.nuap.nagoya-u.ac.jp/sosui/sosui) and TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/) indicating hydrophobic protein sequence regions.

To generate the lists of cell-surface glycoproteins uniquely expressed in N2 or ML2 cells, identified proteins were required to meet the following 3 criteria: (a) identified from one cell line; (b) unique peptide ≥ 1, and NXS/T motif ≥ 1 or UniProt indicated *N*-linked or O-linked glycosylation; and (c) transmembrane region ≥ 1, or UniProt membrane subcellular location.

The list of identified proteins that were unique to each experimental group was subjected to the commercially available curator database software Ingenuity Pathways Analysis (IPA) to determine their molecular function and interacting networks.

RESULTS

Methodology Overview—Our strategy for interrogation of cell surface sialoglycoproteins using selective chemical tagging followed by high-affinity enrichment and GeLC-MS/MS analysis is summarized in Fig. 1. The method consists of several steps: (1) metabolic labeling of N2 and ML2 cells with the azide-containing mannose analog, peracetylated azido-mannose (AC_4ManNAz); (2) chemoselective conjugation of azide sugars with a biotinylated alkyne capture reagent via Cu (I) catalyzed click chemistry in live cells; (3) affinity enrichment of the labeled cell-surface sialylated proteins by streptavidin capture; and (4) separation by one-dimensional gel electrophoresis and identification by LC-MS/MS. When compared with other published approaches, the theoretical advantage of our approach is the targeted selectivity for sialyl glycosylated proteins on the cell surface.

Efficient Cell Surface Expression of AC_4ManAz-Sialoglycoproteins—A critical parameter in the proposed strategy is the efficient labeling and surface expression of the azide-containing sialoglycoproteins. N2 and ML2 cells were subjected to metabolic incorporation of AC_4ManNAz, conjugated with biotin with the click reaction, and subsequently stained with streptavidin-FITC to visualize the glycans harboring the azide group. We conducted fluorescence-assisted flow cytometric analysis of nonpermeabilized cells for determining the efficiency of streptavidin-FITC labeling. As quantified by flow cytometry (Fig. 2*A*), the AC_4ManNAz-treated N2 and ML2 cells displayed 80–100 fold greater FITC-specific fluorescence intensity than the control ManNAc-treated and untreated cells. This result demonstrates that the azide moieties were efficiently incorporated into cellular glycans. We next conducted confocal microscopy analysis of streptavidin labeled cells to specifically assess the cellular location of the azide-modified glycans. As seen in Fig. 2*B*, prominent FITC staining primarily at the cell membrane was observed when the cells were subjected to metabolic incorporation of ManNAz. Consistent with flow cytometric data, confocal microscopic analysis of cells treated with control ManNAc displayed much reduced staining, confirming the labeling via azide-containing glycans is specific. Furthermore, there was no evidence of cell surface labeling for the control groups. We further optimized the azide labeling and click reaction to maximize cell viability/vigor (see supplemental Fig. S1), and determined that 24 h incubation of 20 μM AC_4ManNAz and 25% (v/v) biotin-alkyne was optimal.

Targeted Isolation of Cell-surface Sialoglycoproteins—Having demonstrated efficient labeling of cell surface glycoconjugates, we continued with a general characterization of whole-cell lysates of Biotin-labeled glycoconjugates via SDS-PAGE separation and visualization by reaction with streptavidin-IR800. Expression of β-actin was used for normalization. Overall sialylation, as judged by streptavidin staining, was restricted to extracts from the ManNAz labeled cells (Fig. 3*A*). In addition, there appeared to be greater sialylated glycoproteins in the ML2 cell line when compared with the N2 cell line.

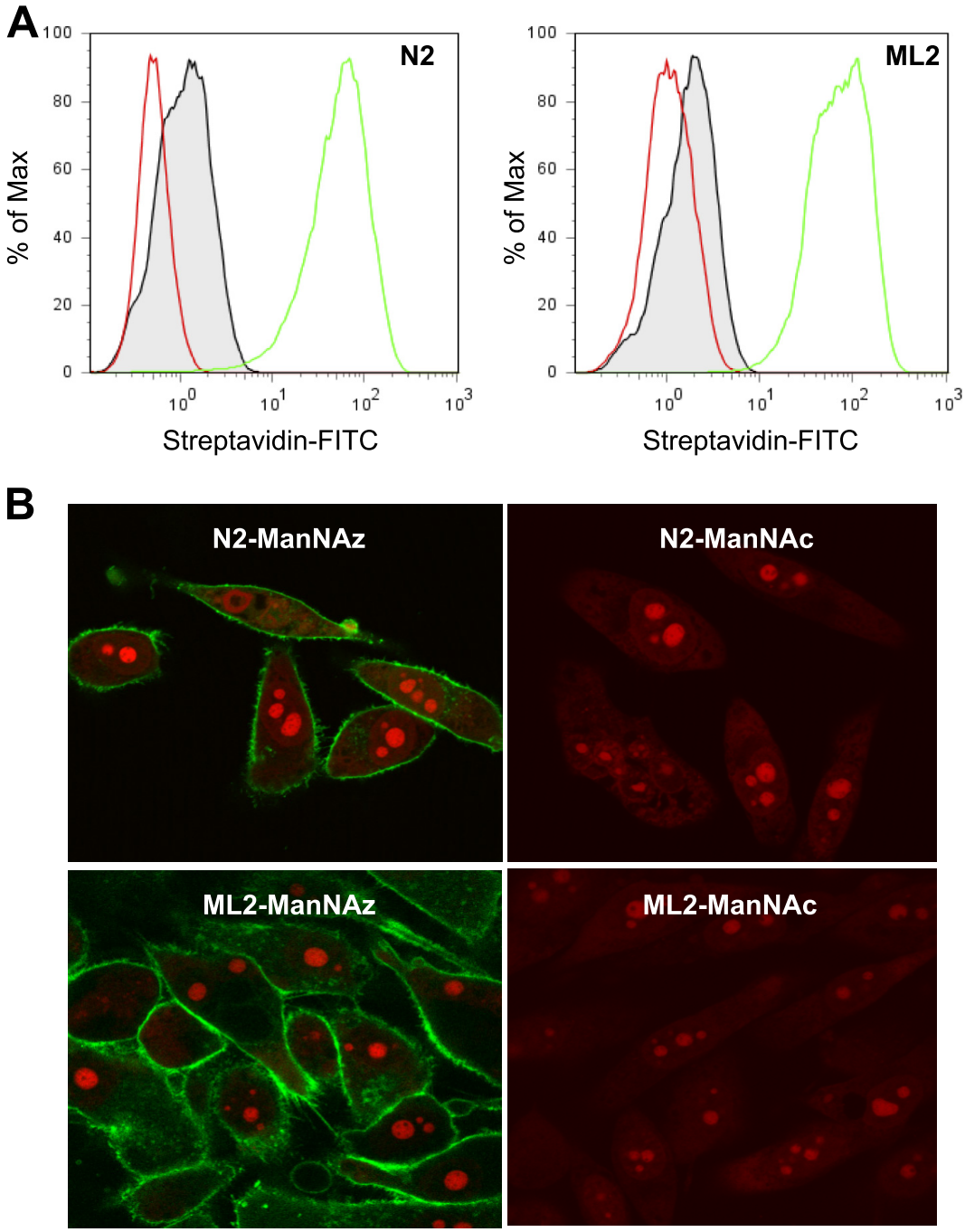

FIG. 2. **Labeling of cell surface sialoglycoproteins.** *A*, Flow cytometry analysis of N2 and ML2 cells treated with AC_4ManNAz or ManNAc. Intact cells were biotin-tagged through click reaction, and visualized with streptavidin-FITC. PI negative gated FACS histogram showing three populations of cells: control cells (red), cells treated with ManNAc (gray filled), and cells treated with 40 μM AC_4ManNAz (green) for 3 days. *B*, Confocal microscopy of N2 and ML2 cells for visualization of the fluorescently labeled cell-surface sialoglycoproteins. Cells were treated with 40 μM AC_4ManNAz or ManNAc for 3 days and then fixed, conjugated with biotin-alkyne via click reaction and "stained" with streptavidin-FITC (green) and PI (red).

We also observed a major protein band that migrated just under 150 kDa and reacted with streptavidin. These observations were consistent with the prior immunofluorescence analyses in demonstrating a specific uptake and labeling of the ManNAz treated cells.

We next examined the efficiency of our system for the enrichment of biotin-labeled sialoglycoproteins. The above whole-cell lysates were incubated with streptavidin-derivatized beads as described. The bound sialoglycoproteins were eluted by boiling and subjected to SDS-PAGE analysis. In

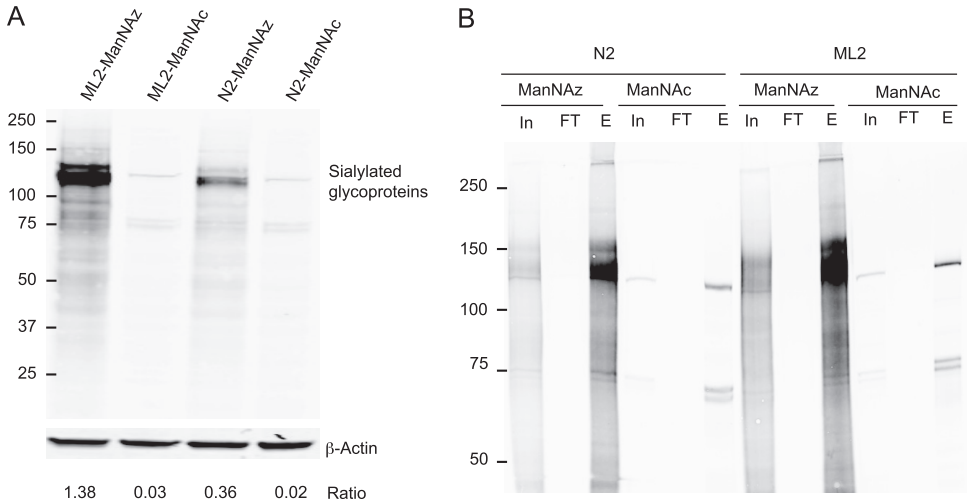

FIG. 3. **Capture and enrichment of azide-tagged sialoglycoproteins.** *A*, Detection of sialylated glycoprotein in cell lysates. N2 and ML2 cells were metabolically labeled with 20 μM ManNAz for 24 h and conjugated with 25% (v/v) biotin-alkyne. Cells were lysed and total protein was extracted as described under "Experimental Procedures." *Top panel*: Protein was resolved by SDS-PAGE and visualized by incubation with streptavidin-IR 800. Reactive bands indicate sialylated glycoproteins. The same blot was probed with anti-β-actin to verify equal protein loading. The relative expression of sialylated glycoproteins after normalization with the β-actin is shown at the bottom of each lane. *B*, Enrichment of azide-tagged sialoglycoproteins by affinity chromatography. Azide-tagged and biotin-conjugated sialoglycoproteins from the total lysate of N2 and ML2 cells were captured by streptavidin beads, separated by SDS-PAGE and visualized by reaction with streptavidin-IR 800. Shown are the results from 20 μg of postclick cell lysate (Input), 5% of the flow through material that did not bind to the beads (Flow-Through), 5% of the eluted material from 2 mg of protein that bound to the beads (Eluent). In, input; FT, flow through; E, eluent. Shown is a representative of four experimental replicates.

both cell types we observed a dramatic enrichment for biotin-conjugated proteins in the bound and eluted fractions from the ManNAz-treated cells (Fig. 3*B*). In contrast there was dramatically reduced overall staining of biotin-labeled protein in the ManNAc-labeled control cells. We again observed an intense signal at sub-150 Kda in eluted fractions from cells treated with ManNAz, and absent or minimal signal in flow-through fractions and eluent from the control group. These results demonstrate efficient capture and enrichment of the azide-labeled proteins.

Identification of Surface Glycoproteins by Mass Spectrometry—Because we were able to establish both efficient cell surface labeling and effective isolation of sialoglycoproteins, we next performed a proof-of-principle experiment for the application of this approach to the identification of sialoglycoproteins differentially expressed between ML2 and N2 cell lines. The cells were processed for metabolic labeling as described above, the tagged proteins were enriched with streptavidin (SAv) resins, the eluted proteins separated by SDS-PAGE, and analyzed by LC-MS/MS. In total, 538 nonredundant proteins were identified. This total derived from 324 proteins from the N2 line and 372 proteins from the ML2 line when metabolically labeled with ManNAz. In contrast the control cells lacking ManNaz yielded 132 proteins from the N2 line and 75 proteins from the ML2 line. The total number of proteins identified using a sliding scale for unique peptides is shown in Table I. Approximately 60% of proteins identified contained two or more unique peptide hits. Supplemental Table 1, A–D contains detailed information on all of the proteins identified for N2, ML2, and their controls. Another purpose of setting up control groups was to evaluate the extent of nonspecific interaction between the SAv resin and proteins in the N2 and ML2 cell lysates. Although a preblocking step and high detergent resin wash was carried out, several proteins were still identified in the control groups. Most of these proteins were high abundant housekeeping proteins, such as actin, tubulin, plectin-1, pyruvate carboxylase, pyruvate kinase, and heat shock proteins. These proteins were subtracted from the list of total proteins identified in the treated groups, further narrowing the pool of proteins identified as cell surface molecules (216 proteins for N2 and 245 for ML2 group).

TABLE I
Total number of proteins identified for each experimental group

No. of peptides identified	N2-ManNAz	N2-ManNAc	ML2-ManNAz	ML2-ManNAc
1	128	64	153	44
2	62	24	72	10
3	29	13	47	6
4	23	11	22	8
≥5	82	20	78	7
Total	324	132	372	75

The enrichment efficiency for sialoglycoproteins is a key factor in characterization of the utility of this methodology. As was shown in Fig. 3, the presence of sialoglycoproteins was significantly enhanced after affinity chromatography by SAv beads when compared with raw cell lysate or processed control extracts (ManNAc). The specific enrichment of glyco-

TABLE II
*Most abundant glycoproteins identified in the N2 and ML2 cell lines. *p = 0.0015 (N2); < 0.0001 (ML2)*

A. N2 Cells

	Acc #	Protein	N2-ManNAz		N2-ManNAc	
			Spect-counts	Uniq-pep	Spect-counts	Uniq-pep
1	O75694	Nuclear pore complex protein Nup155	21	17	4	3
2	P05556	Integrin β-1	64	14	0	0
3	P14625	Endoplasmin	17	10	6	3
4	P26006	Integrin α-3	7	5	0	0
5	P17301	Integrin α-2	7	5	0	0
6	P15144	Amin opeptidase N	8	4	0	0
7	P02786	Transferrin receptor protein 1	6	4	3	2
8	P16144	Integrin β-4	5	4	0	0
9	P16070	CD44 antigen	6	3	0	0
10	O00592	Podocalyxin-like protein 1	6	3	0	0

B. ML2 Cells

	Acc #	prot_desc	ML2-ManNAz		ML2-ManNAc	
			Spect-counts	Uniq-pep	Spect-counts	Uniq-pep
1	P05556	Integrin β-1	75	14	0	0
2	P15144	Amino peptidase N	21	14	0	0
3	O75694	Nuclear pore complex protein Nup155	20	11	2	2
4	P17301	Integrin α-2	17	10	0	0
5	P16144	Integrin β-4	10	8	0	0
6	P14625	Endoplasmin	16	7	2	1
7	P02786	Transferrin receptor protein 1	10	7	1	1
8	Q02413	Desmoglein-2	7	5	0	0
9	P35613	Basigin	11	4	0	0
10	P26006	Integrin α-3	9	4	0	0

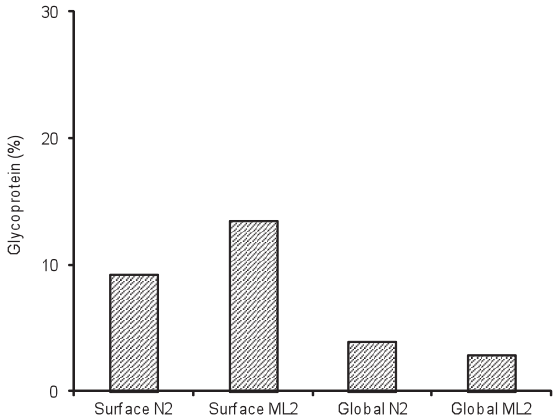

FIG. 4. **Improved capture of known glycoproteins.** Identified proteins were cross-referenced to the UniProt Knowledgebase for known glycoproteins. The percent of total proteins that were matched to known glycoproteins is shown, in comparison to that typically determined by global proteomic approaches.

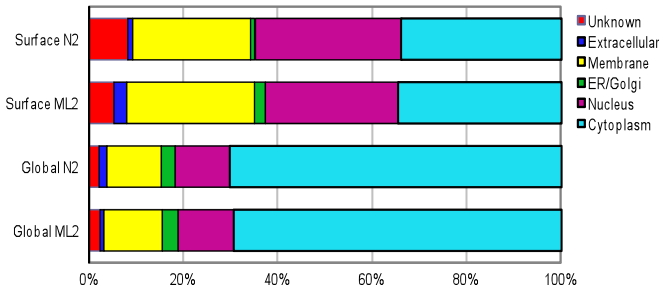

FIG. 5. **Subcellular location of identified proteins.** The cellular location was determined by cross-referencing the annotations from UniProtKB. Each protein was designated with only one subcellular location. The protein classes are shown as percent of total protein on the *x* axis. This analysis was conducted for both the surface glycoprotein and global methodologies.

proteins was further examined using the results from the mass spectrometry characterization. As an initial analysis we determined the top 10 most abundant glycoproteins from each cell line. The expression of these abundant proteins were estimated and ranked by spectral counts and unique peptides as shown in Table IIA and Table IIB. We then compared the relative abundance of these top 10 glycoproteins, as derived from the sialoglycoprotein-targeted enrichment/isolation, between the ManNaz and ManNAc labeled cells. Based upon this comparison, the top most abundant glycoproteins were present in far higher incidence from the ManNAz-treated group. We observed a similar enrichment by the ManNAz group when we employed emPIA score to rank the glycoproteins (see supplemental Table 3). These results clearly demonstrate a substantial increase in the selective capture of the most abundant known cellular glycoproteins via this methodology.

As a final assessment of the extent of enrichment of glycoproteins on the cell surface and in the extracellular compartment, we compared our targeted method to a nonselective global proteomic analysis of whole-cell lysates of the same

TABLE III
Cell-surface N-linked glycoproteins overexpressed in N2 cells

Spect-counts	Uniq-pep	Acc#	Gene Name	Protein Name	Predicted MW (Da)	Observed MW (Da)	Functions
4	4	P28290	SSFA2	Sperm-specific antigen 2	139,554	170,000	Actin binding
1	1	Q92667	AKAP1	Akinase anchor protein 1, mitochondrial	98,250	120,000	Protein binding
1	1	Q00341	HDLBP	Vigilin	141,995	140,000	Sterolmetabolism
1	1	Q8NFJ5	GPRC5A	Retinoic acid-induced protein 3	40,624	300,000	Signal transduction
4	3	Q9Y678	COPG	Coatomer subunit γ	98,967	90,000	ER-Golgi vesicle transport
1	1	P50416	CPT1A	Carnitine O-palmitoyltransferase 1, liver isoform	88,995	300,000	Lipid metabolism
2	2	P35222	CTNNB1	Catenin β-1	86,069	90,000	Cell adhesion; signal transduction
1	1	O60716	CTNND1	Catenin Δ-1	108,674	100,000	Cell adhesion; signal transduction
1	1	P08174	CD55	Complement decay-accelerating factor	42,400	75,000	Complement activation
3	3	Q7L2E3	DHX30	Putative ATP-dependent RNA helicase DHX30	134,938	130,000	ATP binding
1	1	O60610	DIAPH1	Protein diaphanoushomolog 1	141,942	140,000	Receptor binding
1	1	Q8TDM6	DLG5	Disks large homolog 5	215,446	240,000	Cell adhesion; signal transduction
3	2	P50570	DNM2	Dynamin-2	98,345	95,000	Vesicular trafficking
1	1	Q9BW60	ELOVL1	Elongation of very long chain fatty acids protein 1	32,755	300,000	Fatty acid biosynthesis
1	1	O95864	FADS2	Fatty acid desaturase 2	52,340	300,000	Lipid metabolism
1	1	Q05397	PTK2	Focal adhesion kinase 1	119,956	120,000	Integrin-mediated signaling pathways
11	7	Q92616	GCN1L1	Translational activator GCN1	294,953	300,000	Translation activator
2	1	Q8TCT9	HM13	Minor histocompatibility antigen H13	41,747	300,000	Proteolysis; protein binding
1	1	Q86YT6	MIB1	E3 ubiquitin-protein ligase MIB1	112,461	110,000	Notch signal pathway, ubiquitination regulation
1	1	Q9UM54	MYO6	Myosin-VI	150,965	150,000	Protein transport; endocytosis
12	4	Q5SRE5	NUP188	Nucleoporin NUP188 homolog	198,369	220,000	Component of the nuclear pore complex (NPC)
1	1	O75781	PALM	Paralemmin	42,221	60,000	Cell shape
1	1	O00408	PDE2A	cGMP-dependent 3′,5′-cyclic phosphodiesterase	107,360	50,000	Signal transduction
1	1	P13796	LCP1	Plastin-2	70,815	70,000	T-cell activation; intracellular protein transport
2	1	Q8TEM1	NUP210	Nuclear pore membrane glycoprotein 210	205,895	300,000	Protein transport
6	3	O75400	PRPF40A	Pre-mRNA-processing factor 40 homolog A	109,022	120,000	mRNA processing
1	1	Q63HN8	RNF213	RING finger protein 213	378,474	300,000	Nucleoside-triphosphatase activity
4	3	O15027	SEC16A	Protein transport protein Sec 16A	234,855	250,000	ER-Golgi transport
1	1	Q15437	SEC23B	Protein transport protein Sec23B	87,393	80,000	ER-Golgi transport
2	2	P53992	SEC24C	Protein transport protein Sec24C	119,779	120,000	ER-Golgi transport
1	1	P46977	STT3A	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit STT3A	81,104	300,000	N-linked glycosylation
1	1	Q92797	SYMPK	Symplekin	141,915	140,000	Cell adhesion; mRNA processing

TABLE III—*continued*

Spect-counts	Uniq-pep	Acc#	Gene Name	Protein Name	Predicted MW (Da)	Observed MW (Da)	Functions
4	3	Q9NXF1	TEX10	Testis-expressed sequence 10 protein	106,349	90,000	Unknown
1	1	O94901	UNC84A	Protein unc-84 homolog A	90,806	300,000	Protein binding, cytoskeletal anchoring
1	1	Q9H1C4	UNC93B1	Protein unc-93 homolog B1	66,930	300,000	Immunity; intracellular protein transport
1	1	Q969M3	YIPF5	Protein YIPF5	28,370	40,000	ER-Golgi transport

cell model. Whole-cell lysates were prepared from N2 and ML2 cells, directly digested by trypsin and subjected to LC-MS/MS as described in materials and methods. Three nano-LC-MS/MS runs of tryptic digests showed that a significant number of total proteins (310 for N2 and 280 for ML2) could be detected using this simple nonfractionation approach (see supplemental Table 4). When analyzed for protein content, we found that 3.8% of the N2 proteome and 2.9% of the ML2 proteome identified from this global approach were annotated as glycoproteins using the UniProt database. In comparison, our targeted approach yielded 9.3% and 13.5% respectively (Fig. 4). Although the use of the UniProt database to determine if a protein is glycosylated is notoriously under-representative, a fourfold enrichment was revealed. Most striking was that 10 of 12 internal glycoproteins, found to be abundant and identified in the targeted approach, were not identified in the whole-cell proteome approach.

We next conducted a comparative analysis of the two approaches with specific emphasis on which cellular compartment the proteins were derived. Fig. 5 displays the UniProt-designated subcellular distribution of proteins identified in the global proteomic approach and the surface sialoglycoprotein targeted approach. We observed that about 13.2% of proteins from the whole-cell analysis of N2 and ML2 were classified as being extracellular and membrane-bound. This result is consistent with studies using whole-cell approaches to examine breast cancer cell (33) and ovarian cancer cell proteome (34). In contrast, our cell-surface glycoprotein approach resulted in a proteome for N2 and ML2 with 25.9% and 29.8% of proteins being secreted or cell-surface proteins, respectively. These cell-surface proteins include cluster differentiation markers, cell surface receptors, membrane transport, and cell adhesion proteins, such as CD44, CD166, integrins, anexins, calnexin, etc. In contrast, a small portion of proteins identified using our targeted approach were classified as having endoplasmic reticulum/Golgi location (~2%). A large portion of proteins were classified as intracellular (>50%), either cytoplasm or nucleus, whereas 18 proteins for N2 and 13 proteins for ML2 remain unclassified. Recently, Sardana *et al.* conducted a cellular "secretome" analysis in which the conditioned medium of three prostate cancer cell lines (including PC3) was examined. They identified ~2000 proteins of which 12% were classified as known extracellular proteins (35). Our surface glycoprotein approach observed 32 proteins that were in common with this "secretome" study and that 85% of these were classified as membrane proteins. These findings underscore the value of our targeted strategy to capture and isolate cell-surface glycoproteins.

Selective Identification of Differentially Expressed Cell-surface Sialoglycoproteins—Currently less than 5% of proteins are annotated as glycoproteins in the UniProt database (36). This is a striking number when current estimates suggest that more than 50% of eukaryotic proteins are glycosylated. Therefore, we assembled the identified proteins that could be attributed to membrane-bound and extracellular compartments by UniProt database or predicted to contain transmembrane domains (SOSUI, TMHMM). These were further analyzed for the prediction of sites for *N*-linked glycosylation (NetNGlyc). Using these criteria, we compared the surface sialoglycoprotein expression profiles of the metastatic variants. Ultimately, 36 and 44 cell-surface glycoproteins were uniquely identified in N2 and ML2 respectively. A list of these proteins, their UniProt accession number, and molecular function is presented in Table III and IV (The annotated spectra for protein identifications relying on a single peptide are available in supplemental Tables 5 and 6). Known functions of these proteins include cell-cell or cell-matrix adhesion, protein transporter, signaling in response to cytokines and growth factors, induction of coagulation, and proteolysis.

Functional Mapping of Differential Expressed Sialoglycoproteins—We next subjected the selected group of differentially expressed cell-surface glycoproteins to analysis via IPA. Among the 36 cell-surface glycoproteins unique to the nonmetastatic N2 cells (Table III), many proteins, such as, complement decay-accelerating factor (CD55), Catenin β-1 (CTNNB1), focal adhesion kinase 1 (PTK2), Myosin-VI (MYO6), A kinase anchor protein 1 (AKAP1), and symplekin (SYMPK), can positively or negatively regulate apoptosis pathway (Fig. 6) and cell adhesion (catenin Δ-1; CTNND1/disks large homolog 5;DLG5/symplekin; SYMPK). Other proteins overexpressed reflect basic cell functions involving vesicular transport, lipid metabolism and mRNA processing and splicing.

Conversely, the 44 cell-surface glycoproteins unique to the metastatic ML2 cells (Table IV), a large number of proteins are

TABLE IV
Cell-surface N-linked glycoproteins overexpressed in ML2 cells

Spect-counts	Uniq-pep	Acc#	Gene Name	Protein Name	Predicted MW (Da)	Observed MW (Da)	Functions
1	1	P08195	SLC3A2	4F2 cell-surface antigen heavy chain (CD98)	58,023	80,000	Cell growth; amino-acid transport; carbohydrate metabolism
1	1	P49748	ACADVL	Very long-chain specific acyl-CoA dehydrogenase	70,745	70,000	Fatty acid metabolism
3	2	O95573	ACSL3	Long-chain-fatty-acid-CoAligase 3	81,338	85,000	Fatty acid metabolism
3	2	Q5T9A4	ATAD3B	ATPase family AAA domain-containing protein 3B	73,098	70,000	ATP binding
11	4	P35613	BSG	Basigin	42,573	50,000–70,000	Cell surface receptor
5	3	P27797	CALR	Calreticulin	48,283	60,000	Molecular calcium binding chaperone
1	1	Q01518	CAP1	Adenylyl cyclase-associated protein 1	52,222	60,000	Cell polarity maintenance; signal trasduction
1	1	Q96A33	CCDC47	Coiled-coil domain-containing protein 47	56,123	65,000	Protein binding
4	2	Q9H5V8	CDCP1	CUB domain-containing protein 1	94,241	130,000	Cell adhesion; tumor progression and metastasis
2	2	P12109	COL6A1	Collagen α-1 (VI) chain	109,602	130,000	Cell adhesion
2	1	A8TX70	COL29A1	Collagen α-5(VI) chain	291,796	120,000	Cell adhesion
1	1	Q15046	KARS	Lysyl-tRNA synthetase	68,461	70,000	Protein biosynthesis
1	1	Q9BW27	NUP85	Nuclear pore complex protein Nup85	75,826	75,000	Protein transport; mRNA transport
1	1	Q9NZ01	TECR	Synaptic glycoprotein	36,410	300,000	Lipid/steroid biosynthesis
1	1	Q02413	DSG1	Desmoglein-1	114,670	70,000	Cell adhesion
1	1	P06744	GPI	Glucose-6-phosphate isomerase	63,335	65,000	Angiogenesis; gluconeogenesis
1	1	A6NKF9	GPR89C	Putative protein GPR89C	36,760	300,000	Unknown
1	1	P43304	GPD2	Glycerol-3-phosphate dehydrogenase	81,296	80,000	Calcium ion binding
2	1	A1L0T0	ILVBL	Acetolactate synthase-like protein	68,452	70,000	Transferase activity; magnesium ion binding
1	1	Q14721	KCNB1	Potassium voltage-gated channel subfamily B member 1	96,672	60,000	Ion transport
1	1	Q8TF66	LRRC15	Leucine-rich repeat-containing protein 15	65,268	75,000	Protein binding
1	1	P29966	MARCKS	Myristoylated alanine-rich C-kinase substrate	31,707	80,000	Substrate for protein kinase C; binds calmodulin, actin, and synapsin
2	1	Q7Z434	MAVS	Mitochondrial antiviral-signaling protein	57,063	75,000	Innate immune defense against viruses
1	1	Q8NHP6	MOSPD2	Motile sperm domain-containing protein 2	60,051	60,000	Unknown
1	1	Q6PIF6	MYO7B	Myosin-VIIb	243,612	250,000	Actin binding motor activity
2	1	O15118	NPC1	Niemann-Pick C1 protein	144,868	250,000	Intracellular trafficking of cholesterol
7	4	Q9UKX7	NUP50	Nucleoporin 50 kDa	50,512	60,000	Protein transport; mRNA transport

TABLE IV—*continued*
Cell-surface N-linked glycoproteins overexpressed in ML2 cells

Spect-counts	Uniq-pep	Acc#	Gene Name	Protein Name	Predicted MW (Da)	Observed MW (Da)	Function
1	1	Q6V1P9	DCHS2	Protocadherin-23	323,744	55,000	Cell adhesion
1	1	Q9H307	PNN	Pinin	81,679	130,000	Cell adhesion; mRNA processing
8	6	P14923	JUP	Junction plakoglobin	82,434	85,000	Junctional plaque formation
1	1	Q15063	POSTN	Periostin	93,883	70,000	Cell adhesion
4	3	P18031	PTPN1	Tyrosine-protein phosphatase non-receptor type 1	50,505	55,000	Regulation of insulin signal pathways
8	7	P46060	RANGAP1	RanGTPase-activating protein 1	63,958	80,000	Signal transduction
1	1	Q9GZS1	POLR1E	DNA-directed RNA polymerase I subunit RPA49	54,441	55,000	rRNA transcription
3	3	P04843	RPN1	Dolichyl-diphosphooligo-saccharide-protein glycosyltransferase subunit 1	68,641	70,000	N-linked glycosylation
1	1	Q9Y265	RUVBL1	RuvB-like 1	50,538	60,000	Cell cycle; cell division, growth regulation
1	1	Q96QD8	SLC38A2	Sodium-coupled neutral amino acid transporter 2	56,332	300,000	Amino-acid transport
1	1	Q9Y512	SAMM50	Sorting and assembly machinery component 50 homolog	52,342	60,000	Mitochondrial outermembrane translocase complex assembly
3	2	P35498	SCN1A	Sodium channel protein type 1 subunit α	230,876	300,000	Ion transport
2	1	O15269	SPTLC1	Serine palmitoyltransferase 1	53,281	50,000	Lipid metabolism
1	1	O00186	STXBP3	Syntaxin-binding protein 3	68,633	70,000	Vesicle docking in exocytosis
7	4	O94826	TOMM70A	Mitochondrial import receptor subunit TOM70	68,096	75,000	Receptor for protein import into mitochondria
3	1	Q15849	SLC14A2	Urea transporter, kidney	102,213	70,000	Urea transporter
1	1	P38606	ATP6V1A	V-type proton ATPase catalytic subunit A	68,660	70,000	ATP synthesis coupled proton transport

involved in cell movement, migration, and cell invasion. There are 11 proteins involved in cell movement (Fig. 7). Notably, six proteins, Basigin (BSG), Periostin (POSTN), Glucose-6-phosphate isomerase (GPI), Calreticulin (CALR), Leucine-rich repeat-containing protein 15 (LRRC15), and Myristoylated alanine-rich protein kinase C substrate (MARCKS) are all involved in the invasion of tumor cell lines. An additional five proteins, Junction plakoglobin (JUP), Tyrosine-protein phosphatase nonreceptor type 1 (PTPN1), Niemann-Pick C1 protein (NPC1), Sodium-coupled neutral amino acid transporter 2 (SLC38A2), and Adenylyl cyclase-associated protein 1 (CAP1), are involved in cellular migration and/or movement. Six of the aforementioned proteins (SLC38A2, JUP, PTPN1, POSTN, CALR, and BSG) and Pinin (PNN), CUB domain-containing protein 1 (CDCP1), DNA-directed RNA polymerase I subunit RPA49 (POLR1E), and Collagen α-1(VI) chain (COL6A1), are involved in growth/colony formation. Cumulatively, the types of glycoproteins that were differentially expressed in the N2 and ML2 cells functionally reflect their respective nonmetastatic and metastatic phenotypes.

Verification of the Expression of Integrin β1 and CDCP1—We next employed orthoganol approaches to examine two candidate proteins as a verification of the ability of the described methodology to target sialoglycoproteins that are differentially expressed across our cell model. We targeted Integrin β1 and CDCP1 for further analysis and would propose that such verification be a routine component of our workflow. Integrin β1 (CD29) was identified in both N2 and ML2 cells, where as CDCP1 was uniquely identified in the metastatic ML2 cell line. We first examined the cell surface glyco-

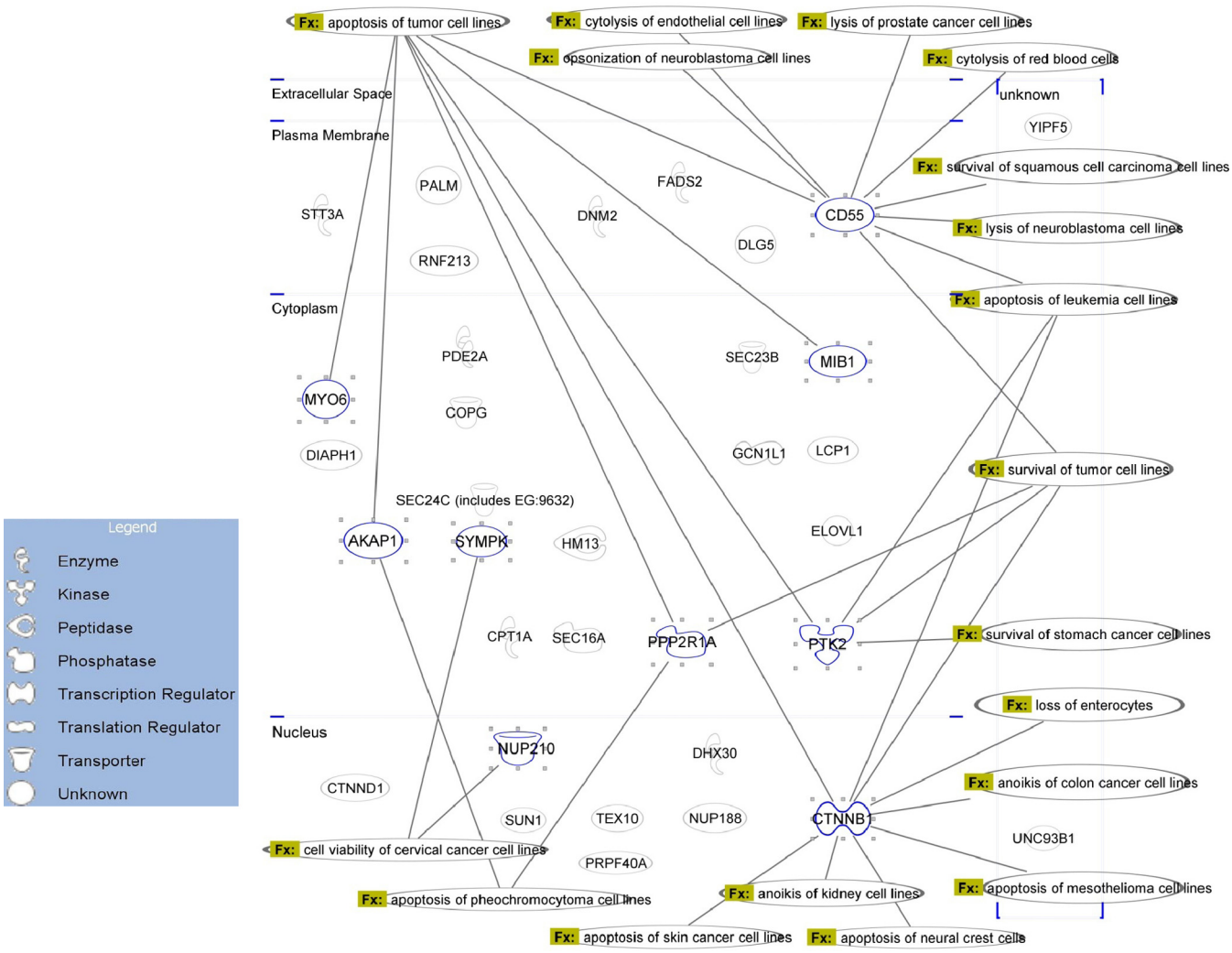

FIG. 6. **Functional pathway analysis of glycoproteins overexpressed in PC3-N2.** Ingenuity Pathway Analysis (IPA) was employed to functionally map the cell-surface glycoproteins overexpressed in PC3-N2 cells. The network shows the biological functions (Fx) that have been associated with these glycoproteins in the context of disease. "Apoptosis of tumor cell lines" is the top listed function, as determined by the number of overexpressed glycoproteins (highlighted in blue) associated with this activity. Individual proteins with known biochemical activities are highlighted with class-specific shapes (indicated in the legend).

protein enriched eluents derived from the streptavidin-based purification shown in Fig. 3*B*. An equal volume of surface enriched fractions from N2 and ML2 cells and their respective controls was resolved by SDS-PAGE and probed with antibodies to the indicated proteins. When analyzed for Integrin β1, a band with an apparent molecular weight of 130 kDa was found in N2 and ML2 ManNAz-treated cells at similar abundance (Fig. 8*A*, *top panel*), consistent with the observed LC-MS/MS results. Immunoblotting for CDCP1 revealed the 135 kDa full-length form as well as a 70 kDa truncated species (truncated species not shown). The CDCP1 levels were ~fourfold greater on the surface of ML2 compared with N2 cells (Fig. 8*A*, *middle panel*). The ManNAc-labeled control cells exposed to the same enrichment procedure had undetectable levels of the selected proteins, underscoring the efficiency of our surface glycoproteomic strategy. The consistent amounts of SAv across all samples served as a control for normalization of the data. This result provides verification of the relative expression of these proteins.

We then determined the expression of the target proteins directly from whole-cell lysates derived from N2 and ML2. Western blot analysis of these extracts confirmed a uniform expression of Integrin β1 between N2 and ML2 (Fig. 8*B*, *top panel*). However, the analysis of CDCP1 (135 kDa) showed that expression in ML2 total cell lysate was higher (1.6-fold) than N2 lysate (Fig. 8*B*, *middel panel*). The expression results were normalized to the endogenous levels of GAPDH. It is interesting to note that although the expression of total CDCP1 is higher in ML2 than N2 it seems that the surface expression of CDCP1 is the most striking difference in this cell model.

We next addressed the glycosylation status of the selected proteins to determine the effectiveness of our approach to

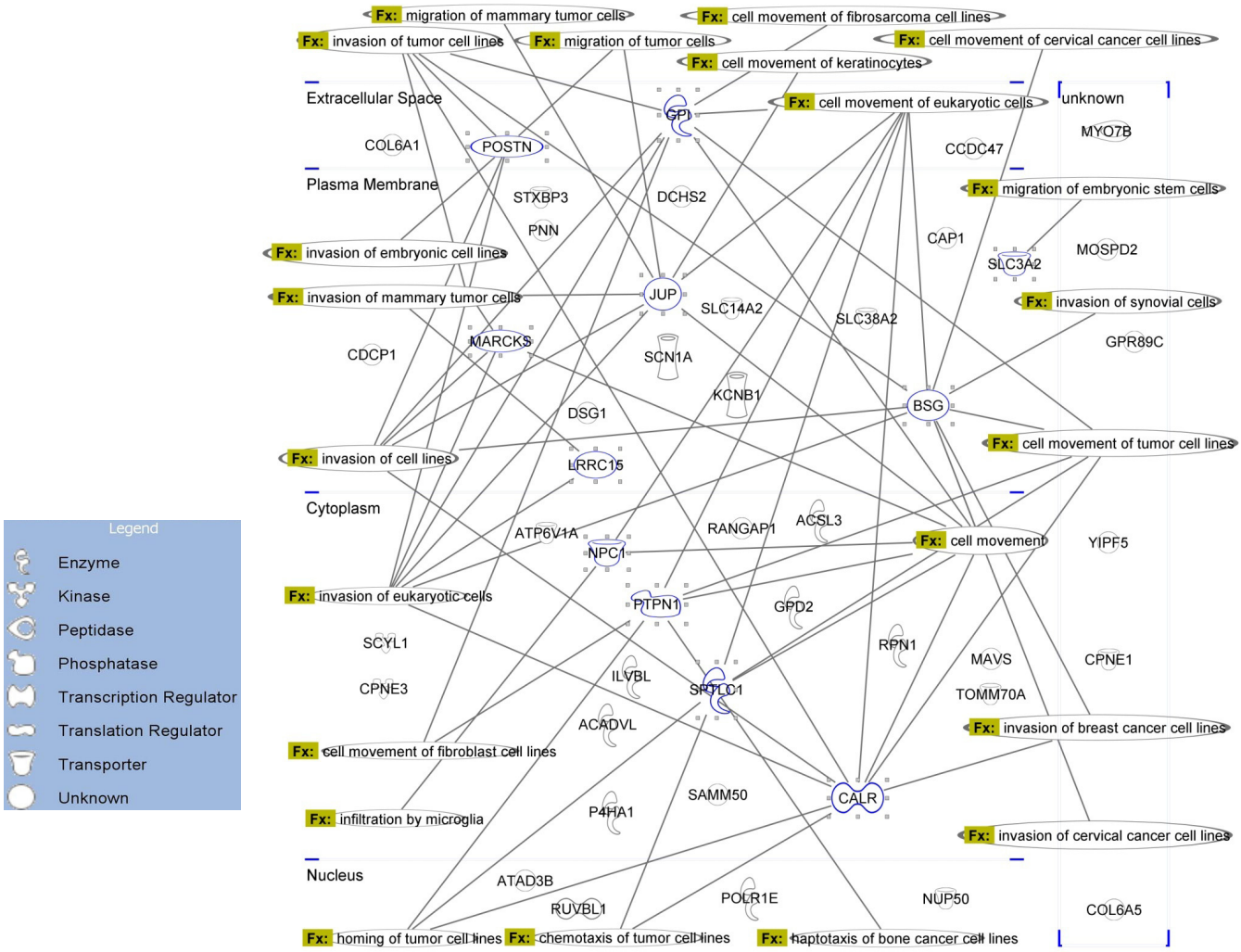

FIG. 7. **Functional pathway analysis of glycoproteins overexpressed in PC3-ML2.** Ingenuity Pathway Analysis (IPA) was employed to functionally map the cell-surface glycoproteins overexpressed in PC3-ML2 cells. The network shows the biological functions (Fx) that have been associated with these glycoproteins in the context of disease. "Cell movement" and "invasion" are the top listed functions, as determined by the number of glycoproteins (highlighted in blue) associated with these activities. Individual proteins with known biochemical activities are highlighted with class-specific shapes (indicated in the legend).

target sialoglycoproteins. Whole-cell extracts were treated with Protein N-glycosidase F (PNGase F), which efficiently cleaves *N*-linked glycans, and neuraminidase, which hydrolyzes only the terminal sialic acid residues. The digested and undigested extracts were separated by SDS-PAGE, and subjected to Western blot analysis. The analysis of integrin β1 showed that PNGase F treatment resulted in the disappearance of the 130 kDa band and coincident generation of a lower molecular weight band of ~90 kDa (Fig. 8*C*, *upper panel*). Likewise, deglycosylation with PNGase F led to a mass shift of about 40 kDa for CDCP1 (Fig. 8*C*, *middle panel*). These results clearly demonstrate that integrin β1 and CDCP1 are heavily N-glycosylated. When examined for specific sialylation we observed that neuraminidase (Neura) treatment resulted in a reduction in the apparent mass for integrin β1 and CDCP1 (Fig. 8*D*), indicating a higher sialic acid content in N-glycans of both protein species. This result is consistent with our approach being able to successfully target sialoglycoproteins and identify those that are differentially expressed in our model.

DISCUSSION

Metastatic spread of cancer involves tumor cell extravasation and subsequent invasion of surrounding tissue, all processes that are tightly regulated by cell surface mechanisms (37, 38). The identification of surface-accessible glycoproteins, which are altered in the metastatic process, will help to unfold the underlying biological events that support cancer metastasis. Developing experimental approaches to selectively interrogate surface glycoproteins should facilitate the discovery of new biomarkers for cancer diagnosis, prognosis,

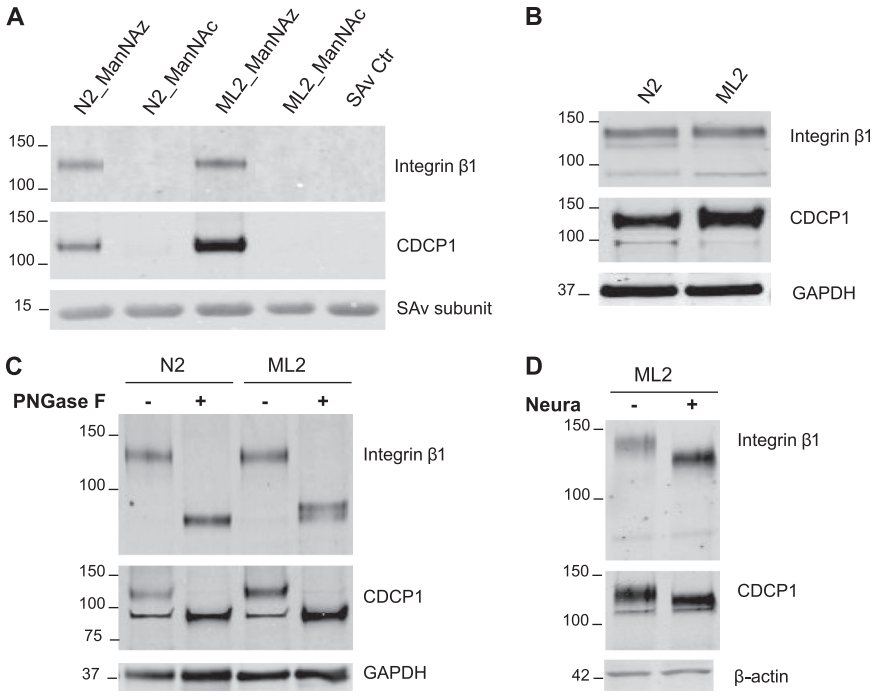

FIG. 8. **Characterization of Integrin β1 and CDCP1 Expression.** *A*, Western blot analysis of cell-surface glycoprotein enriched fractions. N2 and ML2 cells were metabolically labeled with ManNAz or ManNAc, conjugated with biotin-alkyne and isolated via streptavidin. The purified eluates from each group were separated by SDS-PAGE, and immunoblotted for selected proteins as described. Streptavidin (SAv) was analyzed to normalize for loading. *B*, Western blot analysis of expression of the indicated proteins in total lysates of N2 and ML2 cells. Anti-GAPDH was included as a loading control. (*C*, *D*) Characterization of the glycosylation status of integrin β1 and CDCP1. Total cell lysates (20 μg) of N2 and ML2 cells were subjected to digestions with PNGase F (*C*) or neuraminidase (*D*) as described and then subjected to SDS-PAGE. The separated proteins were analyzed by Western blot analysis as described. Analysis of endogenous GAPDH or β actin was included as a loading control.

and help define new therapeutic targets. In this study, a novel surface glycoprotein targeted strategy, based on metabolic labeling and click chemistry, was developed for purification and characterization of sialoglycoproteins. Metabolic labeling of glycans with sugar analogues and subsequent probe capture via click chemistry has been applied to cellular microscopy (39, 40). However, to our knowledge, our data is the first to demonstrate its application in targeted proteomics.

Click chemistry refers to a bioorthogonal ligation event between azide and alkyne (31). Because of its small size, metabolic stability, and lack of reactivity with natural biofunctional structures, azides can be metabolically introduced into cellular substrates with minimal physiological perturbation. The Cu-catalyzed azide-alkyne cycloaddition click reaction offers high-sensitivity and high-specificity for azide detection. Our flow cytometry results and confocal images confirmed the high selectivity and specificity of click chemistry for labeling cell-surface sialoglycoconjugates. We further demonstrated that click conjugation of biotin allowed for selective and efficient isolation of sialoglycoproteins. The ready availability of metabolic probes targeting alternative glycans makes this an attractive strategy for targeting other classes of glycoproteins. Currently there are commercially available Click-friendly products for labeling of *O*-linked, *O*-linked *N*-acetylglucosamine modified, sialic acid modified and fucosylated glycoproteins.

GeLC-MS/MS analysis of the targeted proteins revealed a proteome that was enriched to ~29% for cell surface or excreted proteins. This compares favorably with other proteomic approaches using a nonselective surface biotinylation and MudPIT analysis (34, 41), in which cell-surface protein enrichment of ~30% was reported for human ovarian cancer cell lines. More recently, generic glycan capture followed by the specific release of *N*-linked glycosylated peptides via peptide-N-glycosidase F and LC-MS/MS has been reported to successfully identify *N*-linked glycopeptides (26). Because the authors assumed that all proteins identified by the approach were glycoproteins there was no report of target efficiency. Aebersold and coworkers adapted this system to target surface membrane glycoproteins (42) and McDonald *et al.* demonstrated improved glycoprotein capture using lectin affinity based enrichment (43). Although both glycan capture strategies were successful, the technology is still challenging because of a potential for false positive identification that arises from the dependence on single peptide analysis, 1 Da glycan shift and competition with natural Asn to Asp conversion. To eliminate the false positive identification of the *N*-linked peptides, MS data acquisition from more intense isotopic peaks and manual inspection of mass spectra are required. Because *N*-linked glycosylation predominates in secreted and cell-surface proteins, we would suggest that combining a selective click chemistry approach as we have described with this gel-free system for the analysis of the cell-surface glycoproteome would be a successful alternative strategy.

The significant number of proteins annotated as nuclear or cytoplasmic that were present among our cell-surface sialoglycoproteome is noteworthy. To some extent, occurrence of these predominantly intracellular proteins may be because of their high abundance. Despite our efforts to optimize the labeling protocol and increase wash stringency, some proteins classified as intracellular were still precipitated by the SAv resin. High-salt (1.5 M NaCl) and high-pH (0.1 M Na_2CO_3)

buffers have been reported to reduce such contamination and may reduce this background (44, 45). Another possible reason for the contamination is limited cell death during the click reaction. The reaction requires copper as a catalyst, which is toxic to the cell (46). Fortunately, a new generation click chemistry is available and has been shown to maximally preserve cell viability through bypassing the requirement of copper (47). Bertozzi's group applied copper-free click chemistry for dynamic *in vivo* imaging, claiming comparable kinetics to the copper-catalyzed reaction and speed with no apparent toxicity (47). We have similarly observed identical labeling performance and selectivity with much less cytotoxicity when cells were labeled with these copper-free reagents (data not shown). Improving cell viability could improve proteome selectivity and would also allow for kinetic experiments post-reaction via click. Finally, one must consider existing evidence for the occurrence of cytoplasmic proteins bound to the extracellular membrane or exposed on the cell surface, particularly in aggressive cancer cells (48–50). Thus, a portion of our "cytoplasmic" proteins that we detected may be associated with glycosylated surface binding partners. For this reason it is critical to employ robust orthoganol verification of the identified proteins similar to the strategy we described for verification of integrin β1 and CDCP1 in our system.

In addition to the enriched recovery of surface glycoproteins the identification of proteins differentially expressed between the nonmetastatic (N2) and metastatic (ML2) prostate cancer cell lines was supportive of our model system. Our findings revealed substantial changes in the abundance of proteins that have previously been associated with metastatic spread of cancer and the discovery of some new cell-surface sialoglycoproteins with potential roles in metastasis. The enrichment for surface glycoproteins demonstrated by this method provides proof of concept of the utility of this approach in the discovery of sialic acid-containing glycoproteins of disease significance. Although many cellular proteins are glycosylated, they are rarely identified by global proteomics approaches. We demonstrated that enrichment for surface sialoglycoproteins facilitates identification of biologically relevant surface proteins including a substantial enrichment of cell-surface glycoproteins of low abundance. These included adhesion proteins implicated in cancer development such as protocadherin-23 and CUB-domain containing protein 1, protein binding and signal transduction molecules such as leucine-rich repeat-containing protein 15 and focal adhesion kinase 1. The success at identification of low abundance surface glycoproteins is likely a result of the selective simplification of the complexity of a cellular proteome.

In summary, we describe a glycoproteomic strategy for the analysis of specific subclasses of cell-surface glycoproteins. The approach combines selective metabolic labeling of cell-surface glycans, affinity purification of sialylated proteins, SDS-PAGE separation, and subsequent LC-MS/MS for protein identification. A total of 80 proteins differentially expressed among cell lines were identified, including many integral plasma membrane glycoproteins that were not reported by previous proteomic analyses. Given the importance of the cell-surface glycoproteins for drug design, biomarker discovery, and the regulation of cellular behaviors, the described protocol is a useful addition to currently available tools to expedite the discovery of new cancer biomarker and therapeutic targets in diseases and delineation of signal transduction pathways.

Acknowledgments—We thank Dr. Mark E. Stearns at Drexel University for kindly providing us PC3-N2 and PC3-ML2 cells, and Michael D. Ward at our group for his excellent technical assistance with LC-MS/MS analysis.

* This work was funded in part by National Institutes of Health National Cancer Institute Grants R01 CA135087 (R. R. D.) and U01 CA085067 (O. J. S.).

[S] This article contains supplemental Figs. S1 to S3 and Tables S1 to S4.

¶ To whom correspondence should be addressed: Department of Microbiology and Molecular Cell Biology, Eastern Virginia Medical School, Norfolk, Virginia 23507. Tel.: 757-446-5904; Fax: 757-446-5766; Email: semmesoj@evms.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. Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3,** 97–130
3. Taniguchi, N., Miyoshi, E., Gu, J., Jianguo, G., Honke, K., and Matsumoto, A. (2006) Decoding sugar functions by identifying target glycoproteins. *Curr. Opin. Struct. Biol.* **16,** 561–566
4. Rudd, P. M., Elliott, T., Cresswell, P., Wilson, I. A., and Dwek, R. A. (2001) Glycosylation and the immune system. *Science* **291,** 2370–2376
5. Roth, J. (2002) Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. *Chem. Rev.* **102,** 285–303
6. Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., and Hart, G. W. (2009) *Essentials of glycobiology*. 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
7. Drake, P. M., Cho, W., Li, B., Prakobphol, A., Johansen, E., Anderson, N. L., Regnier, F. E., Gibson, B. W., and Fisher, S. J. (2010) Sweetening the pot: adding glycosylation to the biomarker discovery equation. *Clin. Chem.* **56,** 223–236
8. Dwek, M. V., and Brooks, S. A. (2004) Harnessing changes in cellular glycosylation in new cancer treatment strategies. *Curr. Cancer Drug Targets* **4,** 425–442
9. Lau, K. S., and Dennis, J. W. (2008) N-Glycans in cancer progression. *Glycobiology* **18,** 750–760
10. Ono, M., and Hakomori, S. (2004) Glycosylation defining cancer cell motility and invasiveness. *Glycoconj. J.* **20,** 71–78
11. Miyagi, T., Wada, T., Yamaguchi, K., and Hata, K. (2004) Sialidase and malignancy: a minireview. *Glycoconj. J.* **20,** 189–198
12. Varki, N. M., and Varki, A. (2007) Diversity in cell surface sialic acid presentations: implications for biology and disease. *Lab. Invest.* **87,** 851–857
13. Fogel, M., Altevogt, P., and Schirrmacher, V. (1983) Metastatic potential severely altered by changes in tumor cell adhesiveness and cell-surface sialylation. *J. Exp. Med.* **157,** 371–376
14. Babál, P., Janega, P., Cerná, A., Kholová, I., and Brabencová, E. (2006) Neoplastic transformation of the thyroid gland is accompanied by changes in cellular sialylation. *Acta Histochem.* **108,** 133–140
15. Passaniti, A., and Hart, G. W. (1988) Cell surface sialylation and tumor metastasis. Metastatic potential of B16 melanoma variants correlates with their relative numbers of specific penultimate oligosaccharide struc-

tures. *J. Biol. Chem.* **263,** 7591–7603
16. Brooks, S. A., and Leathem, A. J. (1998) Expression of N-acetyl galactosaminylated and sialylated glycans by metastases arising from primary breast cancer. *Invasion Metastasis* **18,** 115–121
17. Dennis, J. W., Laferté, S., Fukuda, M., Dell, A., and Carver, J. P. (1986) Asn-linked oligosaccharides in lectin-resistant tumor-cell mutants with varying metastatic potential. *Eur. J. Biochem.* **161,** 359–373
18. Abbott, K. L., Aoki, K., Lim, J. M., Porterfield, M., Johnson, R., O'Regan, R. M., Wells, L., Tiemeyer, M., and Pierce, M. (2008) Targeted glycoproteomic identification of biomarkers for human breast carcinoma. *J. Proteome Res.* **7,** 1470–1480
19. Wuhrer, M., Catalina, M. I., Deelder, A. M., and Hokke, C. H. (2007) Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **849,** 115–128
20. Drake, R. R., Schwegler, E. E., Malik, G., Diaz, J., Block, T., Mehta, A., and Semmes, O. J. (2006) Lectin capture strategies combined with mass spectrometry for the discovery of serum glycoprotein biomarkers. *Mol. Cell Proteomics* **5,** 1957–1967
21. Madera, M., Mechref, Y., and Novotny, M. V. (2005) Combining lectin microcolumns with high-resolution separation techniques for enrichment of glycoproteins and glycopeptides. *Anal. Chem.* **77,** 4081–4090
22. Qiu, R., and Regnier, F. E. (2005) Use of multidimensional lectin affinity chromatography in differential glycoproteomics. *Anal. Chem.* **77,** 2802–2809
23. Yang, Z., and Hancock, W. S. (2005) Monitoring glycosylation pattern changes of glycoproteins using multi-lectin affinity chromatography. *J. Chromatogr. A* **1070,** 57–64
24. Zhao, J., Simeone, D. M., Heidt, D., Anderson, M. A., and Lubman, D. M. (2006) Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum. *J. Proteome Res.* **5,** 1792–1802
25. Larsen, M. R., Jensen, S. S., Jakobsen, L. A., and Heegaard, N. H. (2007) Exploring the sialiome using titanium dioxide chromatography and mass spectrometry. *Mol. Cell Proteomics* **6,** 1778–1787
26. 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
27. Zhang, H., Yi, E. C., Li, X. J., Mallick, P., Kelly-Spratt, K. S., Masselon, C. D., Camp, D. G., 2nd, Smith, R. D., Kemp, C. J., and Aebersold, R. (2005) High throughput quantitative analysis of serum proteins using glycopeptide capture and liquid chromatography mass spectrometry. *Mol. Cell Proteomics* **4,** 144–155
28. Sparbier, K., Koch, S., Kessler, I., Wenzel, T., and Kostrzewa, M. (2005) Selective isolation of glycoproteins and glycopeptides for MALDI-TOF MS detection supported by magnetic particles. *J. Biomol. Tech.* **16,** 407–413
29. Laughlin, S. T., Agard, N. J., Baskin, J. M., Carrico, I. S., Chang, P. V., Ganguli, A. S., Hangauer, M. J., Lo, A., Prescher, J. A., and Bertozzi, C. R. (2006) Metabolic labeling of glycans with azido sugars for visualization and glycoproteomics. *Methods Enzymol.* **415,** 230–250
30. Laughlin, S. T., and Bertozzi, C. R. (2009) Imaging the glycome. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 12–17
31. Prescher, J. A., and Bertozzi, C. R. (2005) Chemistry in living systems. *Nat. Chem. Biol.* **1,** 13–21
32. Wang, M., and Stearns, M. E. (1991) Isolation and characterization of PC-3 human prostatic tumor sublines which preferentially metastasize to select organs in S.C.I.D. mice. *Differentiation* **48,** 115–125
33. Kulasingam, V., and Diamandis, E. P. (2007) Proteomics analysis of conditioned media from three breast cancer cell lines: a mine for biomarkers and therapeutic targets. *Mol. Cell Proteomics* **6,** 1997–2011
34. Faça, V. M., Ventura, A. P., Fitzgibbon, M. P., Pereira-Faça, S. R., Pitteri, S. J., Green, A. E., Ireton, R. C., Zhang, Q., Wang, H., O'Briant, K. C., Drescher, C. W., Schummer, M., McIntosh, M. W., Knudsen, B. S., and Hanash, S. M. (2008) Proteomic analysis of ovarian cancer cells reveals dynamic processes of protein secretion and shedding of extra-cellular domains. *PLoS One* **3,** e2425
35. Sardana, G., Jung, K., Stephan, C., and Diamandis, E. P. (2008) Proteomic analysis of conditioned media from the PC3, LNCaP, and 22Rv1 prostate cancer cell lines: discovery and validation of candidate prostate cancer biomarkers. *J. Proteome Res.* **7,** 3329–3338
36. Gupta, R., and Brunak, S. (2002) Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac. Symp. Biocomput.* **7,** 310–322
37. Fuster, M. M., and Esko, J. D. (2005) The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat. Rev. Cancer* **5,** 526–542
38. Psaila, B., and Lyden, D. (2009) The metastatic niche: adapting the foreign soil. *Nat. Rev. Cancer* **9,** 285–293
39. Hsu, T. L., Hanson, S. R., Kishikawa, K., Wang, S. K., Sawa, M., and Wong, C. H. (2007) Alkynyl sugar analogs for the labeling and visualization of glycoconjugates in cells. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 2614–2619
40. Sawa, M., Hsu, T. L., Itoh, T., Sugiyama, M., Hanson, S. R., Vogt, P. K., and Wong, C. H. (2006) Glycoproteomic probes for fluorescent imaging of fucosylated glycans in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 12371–12376
41. Conn, E. M., Madsen, M. A., Cravatt, B. F., Ruf, W., Deryugina, E. I., and Quigley, J. P. (2008) Cell surface proteomics identifies molecules functionally linked to tumor cell intravasation. *J. Biol. Chem.* **283,** 26518–26527
42. Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R., 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
43. McDonald, C. A., Yang, J. Y., Marathe, V., Yen, T. Y., and Macher, B. A. (2009) Combining results from lectin affinity chromatography and glycocapture approaches substantially improves the coverage of the glycoproteome. *Mol. Cell Proteomics* **8,** 287–301
44. Zhao, Y., Zhang, W., Kho, Y., and Zhao, Y. (2004) Proteomic analysis of integral plasma membrane proteins. *Anal. Chem.* **76,** 1817–1823
45. Zhao, Y., Zhang, W., White, M. A., and Zhao, Y. (2003) Capillary high-performance liquid chromatography/mass spectrometric analysis of proteins from affinity-purified plasma membrane. *Anal Chem* **75,** 3751–3757
46. Baskin, J. M., Prescher, J. A., Laughlin, S. T., Agard, N. J., Chang, P. V., Miller, I. A., Lo, A., Codelli, J. A., and Bertozzi, C. R. (2007) Copper-free click chemistry for dynamic in vivo imaging. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 16793–16797
47. Laughlin, S. T., Baskin, J. M., Amacher, S. L., and Bertozzi, C. R. (2008) In vivo imaging of membrane-associated glycans in developing zebrafish. *Science* **320,** 664–667
48. Kaiser, B. K., Yim, D., Chow, I. T., Gonzalez, S., Dai, Z., Mann, H. H., Strong, R. K., Groh, V., and Spies, T. (2007) Disulphide-isomerase-enabled shedding of tumour-associated NKG2D ligands. *Nature* **447,** 482–486
49. Shin, B. K., Wang, H., Yim, A. M., Le Naour, F., Brichory, F., Jang, J. H., Zhao, R., Puravs, E., Tra, J., Michael, C. W., Misek, D. E., and Hanash, S. M. (2003) Global profiling of the cell surface proteome of cancer cells uncovers an abundance of proteins with chaperone function. *J. Biol. Chem.* **278,** 7607–7616
50. Roesli, C., Borgia, B., Schliemann, C., Gunthert, M., Wunderli-Allenspach, H., Giavazzi, R., and Neri, D. (2009) Comparative analysis of the membrane proteome of closely related metastatic and nonmetastatic tumor cells. *Cancer Res.* **69,** 5406–5414