

✂ Author's Choice

# The Mouse C2C12 Myoblast Cell Surface *N*-Linked Glycoproteome

IDENTIFICATION, GLYCOSITE OCCUPANCY, AND MEMBRANE ORIENTATION\*<sup>§</sup>

Rebekah L. Gundry<sup>‡§</sup>, Kimberly Raginski<sup>§¶</sup>, Yelena Tarasova<sup>‡§¶</sup>,  
Irina Tchernyshyov<sup>‡¶</sup>, Damaris Bausch-Fluck<sup>||</sup>, Steven T. Elliott<sup>¶‡</sup>,  
Kenneth R. Boheler<sup>§\*\*</sup>, Jennifer E. Van Eyk<sup>‡\*\*‡‡§§¶¶</sup>, and Bernd Wollscheid<sup>||\*\*</sup>

Endogenous regeneration and repair mechanisms are responsible for replacing dead and damaged cells to maintain or enhance tissue and organ function, and one of the best examples of endogenous repair mechanisms involves skeletal muscle. Although the molecular mechanisms that regulate the differentiation of satellite cells and myoblasts toward myofibers are not fully understood, cell surface proteins that sense and respond to their environment play an important role. The cell surface capturing technology was used here to uncover the cell surface *N*-linked glycoprotein subproteome of myoblasts and to identify potential markers of myoblast differentiation. 128 *bona fide* cell surface-exposed *N*-linked glycoproteins, including 117 transmembrane, four glycosylphosphatidylinositol-anchored, five extracellular matrix, and two membrane-associated proteins were identified from mouse C2C12 myoblasts. The data set revealed 36 cluster of differentiation-annotated proteins and confirmed the occupancy for 235 *N*-linked glycosylation sites. The identification of the *N*-glycosylation sites on the extracellular domain of the proteins allowed for the determination of the orientation of the identified proteins within the plasma membrane. One glycoprotein transmembrane orientation was found to be inconsistent with Swiss-Prot annotations, whereas ambiguous annotations for 14 other proteins were resolved. Several of the identified *N*-linked glycoproteins, including aquaporin-1 and  $\beta$ -sarcoglycan, were found in validation experiments to change in overall abundance as the myoblasts differentiate toward myotubes. Therefore, the strategy and data presented shed new light on the complexity of the myoblast cell surface subproteome and reveal new targets for the clinically important characterization of cell intermediates during myoblast differentiation into myotubes. *Molecular & Cellular Proteomics* 8:2555–2569, 2009.

Endogenous regeneration and repair mechanisms are responsible for replacing dead and damaged cells to maintain or enhance tissue and organ function. One of the best examples of endogenous repair mechanisms involves skeletal muscle, which has innate regenerative capacity (for reviews, see Refs. 1–4). Skeletal muscle repair begins with satellite cells, a heterogeneous population of mitotically quiescent cells located in the basal lamina that surrounds adult skeletal myofibers (5, 6), that, when activated, rapidly proliferate (7). The progeny of activated satellite cells, known as myogenic precursor cells or myoblasts, undergo several rounds of division prior to withdrawal from the cell cycle. This is followed by fusion to form terminally differentiated multinucleated myotubes and skeletal myofibers (7, 8). These cells effectively repair or replace damaged cells or contribute to an increase in skeletal muscle mass.

The molecular mechanisms that regulate differentiation of satellite cells and myoblasts toward myofibers are not fully understood, although it is known that the cell surface proteome plays an important biological role in skeletal muscle differentiation. Examples include how cell surface proteins modulate myoblast elongation, orientation, and fusion (for a review, see Ref. 8). The organization and fusion of myoblasts is mediated, in part, by cadherins (for reviews, see Refs. 9 and 10), which enhance skeletal muscle differentiation and are implicated in myoblast fusion (11). Neogenin, another cell surface protein, is also a likely regulator of myotube formation via the netrin ligand signal transduction pathway (12, 13), and the family of sphingosine 1-phosphate receptors (Edg receptors) are known key signal transduction molecules involved in regulating myogenic differentiation (14–17). Given the important role of these proteins, identifying and characterizing the cell surface proteins present on myoblasts in a more comprehensive approach could provide insights into the molecular mechanisms involved in skeletal muscle development and repair. The identification of naturally occurring cell surface proteins (*i.e.* markers) could also foster the enrichment and/or characterization of cell intermediates during differentiation that could be useful therapeutically.

Although it is possible to use techniques such as flow cytometry, antibody arrays, and microscopy to probe for known proteins on the cell surface in discrete populations,

From the Departments of <sup>‡</sup>Medicine, <sup>¶</sup>Biological Chemistry, and <sup>§§</sup>Biomedical Engineering, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21224, <sup>§</sup>NIA, National Institutes of Health, Baltimore, Maryland 21224, and <sup>||</sup>ETH Zurich, Institute of Molecular Systems Biology, NCCR Neuro Center for Proteomics, Zurich CH–8093, Switzerland

✂ Author's Choice—Final version full access.

Received, April 22, 2009, and in revised form, July 17, 2009

Published, MCP Papers in Press, August 4, 2009, DOI 10.1074/mcp.M900195-MCP200

these methods rely on a *priori* knowledge of the proteins present on the cell surface and the availability/specificity of an antibody. Proteomics approaches coupled with mass spectrometry offer an alternative approach that is antibody-independent and allows for the *de novo* discovery of proteins on the surface. One approach, which was used in the current study, exploits the fact that a majority of the cell surface proteins are glycosylated (18). The method uses hydrazide chemistry (19) to immobilize and enrich for glycoproteins/glycopeptides, and previous studies using this chemistry have successfully identified soluble glycoproteins (20–24) as well as cell surface glycoproteins (25–28). A recently optimized hydrazide chemistry strategy by Wollscheid *et al.* (29) termed cell surface capturing (CSC)<sup>1</sup> technology, reports the ability to identify cell surface (plasma membrane) proteins specifically with little (<15%) contamination from non-cell surface proteins. The specificity stems from the fact that the oligosaccharide structure is labeled using membrane-impermeable reagents while the cells are intact rather than after cell lysis. Consequently, only extracellular oligosaccharides are labeled and subsequently captured. Utilizing information regarding the glycosylation site then allows for a rapid elimination of nonspecifically captured proteins (*i.e.* non-cell surface proteins) during the data analysis process, a feature that makes this approach unique to methods where no label or tag is used. Additionally, the CSC technology provides information about glycosylation site occupancy (*i.e.* whether a potential *N*-linked glycosylation site is actually glycosylated), which is important for determining the protein orientation within the membrane and, therefore, antigen selection and antibody design.

To uncover information about the cell surface of myoblasts and to identify potential markers of myoblast differentiation, we used the CSC technology on the mouse myoblast C2C12 cell line model system (30, 31). This adherent cell line derived from satellite cells has routinely been used as a model for skeletal muscle development (*e.g.* Refs. 1, 32, and 33), skeletal muscle differentiation (*e.g.* Refs. 34–36), and studying muscular dystrophy (*e.g.* Refs. 37–39). Additionally, these cells have been used in cell-based therapies (*e.g.* Refs. 40–42). Using the CSC technology, 128 cell surface *N*-linked glycoproteins were identified, including several that were found to change in overall abundance as the myoblasts differentiate toward myotubes. The current data also confirmed the occupancy of 235 *N*-linked glycosites of which 226 were previously unconfirmed. The new information provided by the current study is expected to facilitate the development of

useful tools for studying the differentiation of myoblasts toward myotubes.

#### EXPERIMENTAL PROCEDURES

**Cell Culture**—Mouse myoblasts (C2C12 cell line) were cultured as described previously (43, 44). C2C12 cells were cultivated in growth medium (Dulbecco's modified Eagle's medium, L-Glu, penicillin/streptomycin, 20% fetal bovine serum (FBS), 4.5g/liter glucose) in 5% CO<sub>2</sub> and passaged at 70–80% confluency to maintain the undifferentiated myoblast population. Three biological replicates of undifferentiated C2C12 cells at ~70% confluency were used. For differentiation, cells were switched under confluent conditions (>70–80%) to low serum conditions (5% FBS).

**CSC Technology**—Approximately  $1 \times 10^8$  cells per biological replicate were taken through the CSC technology work flow as reported previously (28, 29) (Fig. 1). Cells were washed twice with labeling buffer (1× PBS (Quality Biological, Gaithersburg, MD), pH 6.5, 0.1% (v/v) FBS (Invitrogen)) followed by treatment for 15 min in 1.5 mM sodium *meta*-periodate (Pierce) in labeling buffer at 4 °C. Cells were washed with labeling buffer, collected, and centrifuged at  $225 \times g$  for 5 min at 25 °C. The pelleted cells were resuspended in 2.5 mg/ml biocytin hydrazide (Biotium, Hayward, CA) in labeling buffer for 1 h at 4 °C with gentle agitation, then washed with 1× PBS, and pelleted as above. Cells were resuspended in lysis buffer (10 mM Tris, pH 7.5, 0.5 mM MgCl<sub>2</sub>) and homogenized using a Dounce homogenizer. Cell lysate was centrifuged at  $2500 \times g$  for 10 min at 4 °C to remove the nucleus. The supernatant, containing the membranes, was centrifuged at  $210,000 \times g$  for 16 h at 4 °C. The membrane pellet was washed with 25 mM Na<sub>2</sub>CO<sub>3</sub>, resuspended in lysis buffer, and centrifuged at  $210,000 \times g$  for 30 min at 4 °C. The pellet was resuspended by sonication in 100 mM NH<sub>4</sub>HCO<sub>3</sub>, 5 mM tris(2-carboxyethyl)phosphine (Sigma), 0.1% (v/v) Rapigest (Waters). Proteins were then alkylated with 10 mM iodoacetamide for 30 min in the dark at 25 °C. The sample was then incubated with 1 μg of glycerol-free endoproteinase Lys-C (Calbiochem) at 37 °C for 4 h with end-over-end rotation and then with 20 μg of proteomics grade trypsin (Promega, Madison, WI) at 37 °C for 16 h with end-over-end rotation. The enzymes were inactivated by heating at 100 °C for 10 min followed by the addition of 10 μl of 1× Complete protease inhibitor mixture (Roche Applied Science). The peptide mixture was incubated with a 500-μl bead slurry of UltraLink Immobilized Streptavidin PLUS (Pierce) for 1 h at 25 °C. The beads were sequentially washed with the following: 5 M NaCl, 100 mM NH<sub>4</sub>HCO<sub>3</sub>, 5 M NaCl, 100 mM Na<sub>2</sub>CO<sub>3</sub>, 80% isopropanol, and 100 mM NH<sub>4</sub>HCO<sub>3</sub>. The beads were resuspended in 100 mM NH<sub>4</sub>HCO<sub>3</sub> and 500 units of glycerol-free endoproteinase peptide-*N*-glycosidase F (New England Biolabs, Ipswich, MA) and incubated at 37 °C for 16 h with end-over-end rotation to release the peptides from the beads. The collected peptides were desalted and concentrated using a C<sub>18</sub> UltraMicroSpin™ column (Nest Group, Southborough, MA) according to the manufacturer's instructions. In general,  $1 \times 10^8$  cells provided sufficient peptide quantity for two to three individual LC-MS/MS analyses.

**Mass Spectrometry**—For each biological replicate ( $n = 3$ ), two technical replicates were analyzed by LC-MS/MS using either an LTQ-Orbitrap (Thermo, Waltham, MA) or an LTQ-FT (Thermo). For the LTQ-Orbitrap, desalted peptides were resuspended in 12 μl of 0.1% (v/v) aqueous formic acid (FA). Two times 5 μl were injected and analyzed on an Agilent 1200 nano-LC system (Agilent, Santa Clara, CA) connected to an LTQ-Orbitrap mass spectrometer (Thermo) equipped with a nanoelectrospray ion source (Thermo). Peptides were separated on a BioBasic (New Objective, Woburn, MA) C<sub>18</sub> reversed phase HPLC column (75 μm × 10 cm) using a linear gradient from 5% B to 65% B in 60 min at a flow rate of 300 nl/min where mobile phase A was composed of 0.1% (v/v) aqueous FA and mobile

<sup>1</sup> The abbreviations used are: CSC, cell surface capturing; TM, transmembrane; FBS, fetal bovine serum; LTQ, linear trap quadrupole; FA, formic acid; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; CD, cluster of differentiation; GPI, glycosylphosphatidylinositol; ECM, extracellular matrix; NCAM, neural cell adhesion molecule.

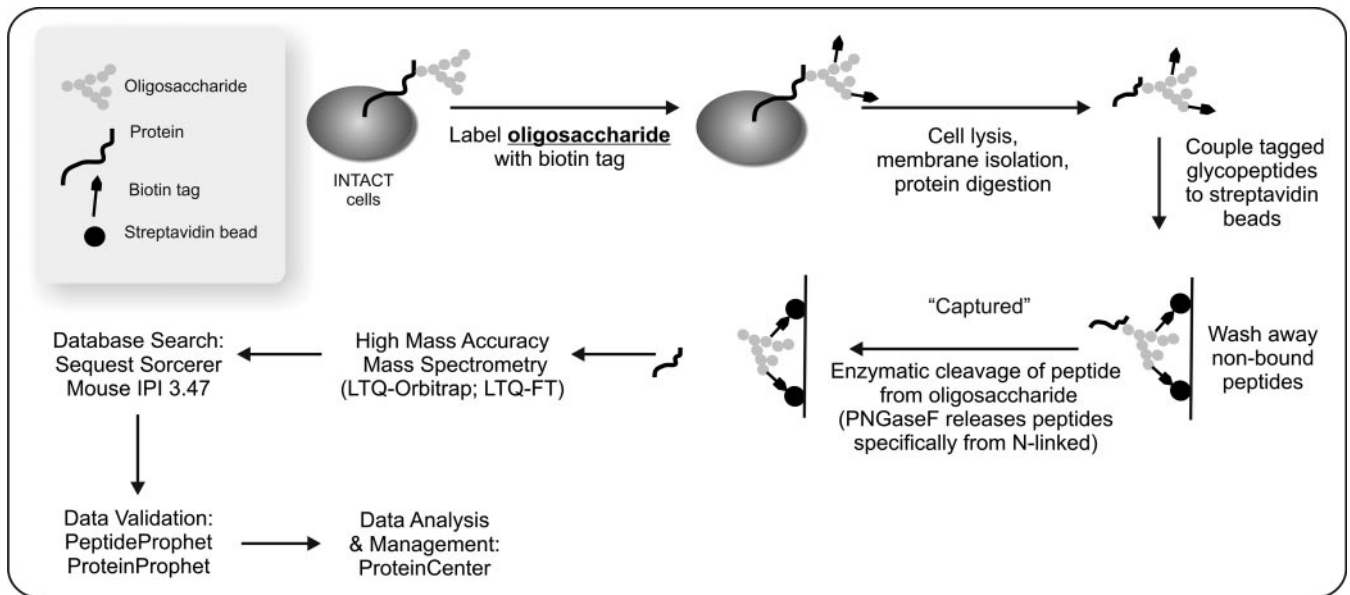


FIG. 1. **Schema of CSC technology for identifying N-linked glycoproteins.** Overview of the experimental work flow for enriching and identifying cell surface N-linked glycopeptides. *IPI*, International Protein Index; *PNGaseF*, peptide-N-glycosidase F.

phase B was 90% acetonitrile, 0.1% FA in water. Each MS1 scan was followed by CID (acquired in the LTQ part) of the five most abundant precursor ions with dynamic exclusion for 30 s. Only MS1 signals exceeding 10,000 counts triggered the MS2 scans. For MS1,  $2 \times 10^5$  ions were accumulated in the Orbitrap over a maximum time of 500 ms and scanned at a resolution of 60,000 full-width half-maximum (at 400 *m/z*). MS2 spectra (via CID) were acquired in normal scan mode in the LTQ using a target setting of  $10^4$  ions and an accumulation time of 30 ms. The normalized collision energy was set to 35%, and one microscan was acquired for each spectrum. For the LTQ-FT, desalted peptides were resuspended in 12  $\mu$ l of 0.1% (v/v) aqueous FA. Two times 4  $\mu$ l were injected and analyzed on a Tempo<sup>TM</sup> Nano 1D+ HPLC system (Applied Biosystems/MDS Sciex, Foster City, CA) connected to a 7-tesla Finnigan LTQ-FT-ICR instrument (Thermo) equipped with a nano-electrospray ion source (Thermo) using a C<sub>18</sub> reverse phase HPLC column (75  $\mu$ m  $\times$  15 cm) packed in house (Magic C<sub>18</sub> AQ 3  $\mu$ m; Michrom Bioresources, Auburn, CA) using a linear gradient from 4% B to 35% B in 60 min at a flow rate of 300 nl/min where mobile phase A was composed of 0.15% aqueous FA and mobile phase B was 98% (v/v) acetonitrile, 0.15% (v/v) FA in water. Each MS1 scan (acquired in the ICR cell) was followed by CID (acquired in the LTQ) of the five most abundant precursor ions with dynamic exclusion for 30 s. Only MS1 signals exceeding 150 counts were allowed to trigger MS2 scans with wideband activation disabled. For MS1,  $3 \times 10^6$  ions were accumulated in the ICR cell over a maximum time of 500 ms and scanned at a resolution of 100,000 full-width half-maximum (at 400 *m/z*). MS2 spectra were acquired in normal scan mode with a target setting of  $10^4$  ions and an accumulation time of 100 ms. Singly charged ions and ions with unassigned charge state were excluded from triggering MS2 events. The normalized collision energy was set to 32%, and one microscan was acquired for each spectrum.

**Mass Spectrometry Database Search**—Raw MS data were searched against the International Protein Index mouse v3.47 database (45) (55,298 entries; August 26, 2008) using Sorcerer 2<sup>TM</sup>-SEQUEST<sup>®</sup> (Sage-N Research, Milpitas, CA) with postsearch analysis performed using the Trans-Proteome Pipeline, implementing PeptideProphet (46) and ProteinProphet (47) algorithms. All raw data peak extraction was performed using Sorcerer 2-SEQUEST default

settings. Database search parameters were as follows: semienzyme digest using trypsin (after Lys or Arg) with up to two missed cleavages; monoisotopic precursor mass range of 400–4500 amu; and oxidation (Met), carbamidomethylation (Cys), and deamidation (Asn) allowed as differential modifications. Peptide mass tolerance was set to 50 ppm, fragment mass tolerance was set to 1 amu, fragment mass type was set to monoisotopic, and the maximum number of modifications was set to four per peptide. Advanced search options that were enabled included the following: XCorr score cutoff of 1.5, isotope check using a mass shift of 1.003355 amu, keep the top 2000 preliminary results for final scoring, display up to 200 peptide results in the result file, display up to five full protein descriptions in the result file, and display up to one duplicate protein reference in the result file. Error rates (false discovery rates) and protein probabilities (*p*) were calculated by ProteinProphet. Raw data from all three biological replicates were combined into a single database search.

**Protein Data Processing, Redundancy Removal, and Database Presentation**—The ProteinProphet interact-prot.xml result files were input into ProteinCenter (Proxeon Bioinformatics, Odense, Denmark) and filtered to contain only proteins with protein probability scores of *p* > 0.48. To prevent redundancy in protein identifications, proteins were grouped according to “indistinguishable proteins,” resulting in 128 protein groups. For the final database, isoform notation is provided only when a peptide that is unique to a specific protein isoform was identified. The protein list in supplemental Table S1 displays only those proteins identified by peptides containing an observed deamidation at the asparagine(s) within the conserved sequence motif for N-linked glycosylation: asparagine (N) followed by any amino acid except proline (X) followed by serine (S) or threonine (T), NX(S/T). Membrane topology predictions were obtained from three different prediction algorithms: publicly available versions of HMMTOP v2.0 (48, 49) and SOSUI v1.11 (50) and TMAP (51), which is integrated into ProteinCenter.

**Consideration of Single Peptide Identifications**—In this type of sample processing, the number of peptides identified for each protein is completely dependent upon the number of potential N-linked glycosylation sites and whether the glycosylation site is within a tryptic peptide of suitable *m/z* for MS analysis. For this reason, proteins identified by a single glycopeptide were not automatically excluded;



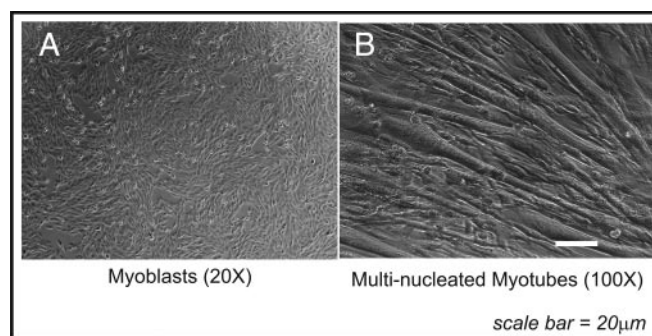
rather, great care was taken to appropriately evaluate and present the data for these identifications, which may fall into two categories. First, there are identifications for which a single peptide sequence was observed  $>2$  times (either multiple observations of the same charge state or as multiple charge states). This accounts for the majority of identifications based on a single peptide sequence. Second, there were several proteins for which a single peptide was observed  $\leq 2$  times. For identifications from the latter category, the annotated MS/MS spectra are provided in supplemental Table S8c. To ensure specificity for the proteins reported, the peptide sequence for any “single peptide identification” was searched (via BLAST) against NCBI nr to ensure that it mapped (with 100% homology) to only the protein reported. In supplemental Table S1 the protein identifications are sorted by false discovery rate (*i.e.* confidence), and the supplemental information contains all details regarding the peptides identified (supplemental Table S8), including spectra for single peptide identifications when appropriate.

**Comparison with Previous Studies of C2C12 Differentiation—**ProteinCenter was used to compare the experimentally derived data with data imported from the literature (current through December 2008). ProteinCenter clusters equivalent protein names into a single descriptor based on the amino acid sequence, allowing for an accurate comparison among multiple data sets regardless of the database searched. To compare the proteomics data, the accession numbers of the proteins reported by Tannu *et al.* (52), Kislinger *et al.* (53), and Capkovic *et al.* (54), as provided by the authors, were input directly into ProteinCenter. To compare the gene expression data, the list of detected genes as provided by Moran *et al.* (55) and Tomczak *et al.* (56) were converted to International Protein Index or Swiss-Prot protein accession numbers using either the Protein Identifier Cross-Reference Service (57) or IDconverter (58), and these protein accession numbers were imported into ProteinCenter. For Tomczak *et al.* (56) data, 103 of 2896 probes could not be assigned to protein accession numbers and are largely expressed sequence tags. For Moran *et al.* (55) data, only the 629 differentially regulated genes were compared as not all 11,000 that were probed were provided by the authors. Of these, four could not be assigned to protein accession numbers. In ProteinCenter, the protein accession lists for all data were clustered by 80% homology at the amino acid level. A summary of the previous studies with which the current data were compared can be found in supplemental Table S3.

**Western Blotting—**Antibodies were obtained for  $\beta$ -sarcoglycan (Novocastra; NCL-L-b-SARC), aquaporin 1 (Chemicon International, Temecula, CA; AB2219), and cadherin 2 (N-cadherin, CD325; BD Biosciences, San Jose, CA; 610920). For Western blot loading controls, topoisomerase I monoclonal antibody (BD Biosciences; 556597) was used. Cells were lysed in Laemmli buffer, and corresponding amounts of total protein (15–50  $\mu$ g, determined via BCA assay (Pierce); see Fig. 5) were separated on a 4–12% NuPAGE Bis-Tris gel (Invitrogen) according to the manufacturer’s standard protocol. The following antibody dilutions were used to detect the protein of interest: anti-aquaporin-1 rabbit polyclonal, 1:1000 dilution; anti- $\beta$ -sarcoglycan mouse monoclonal, 1:100 dilution; anti-cadherin 2 mouse monoclonal, 1:5000; and anti-topoisomerase I mouse monoclonal, 1:1000 dilution. Blots were developed with Amersham Biosciences ECL™ Western blotting detection reagent (GE Healthcare) according to the manufacturer’s protocol.

### RESULTS

**Cells—**The mouse myoblast C2C12 cell line was cultivated in medium containing high serum (20% FBS) and high glucose (4.5 g/liter). Under these conditions, cells maintained an undifferentiated fibroblast-like morphology with a single nucleus

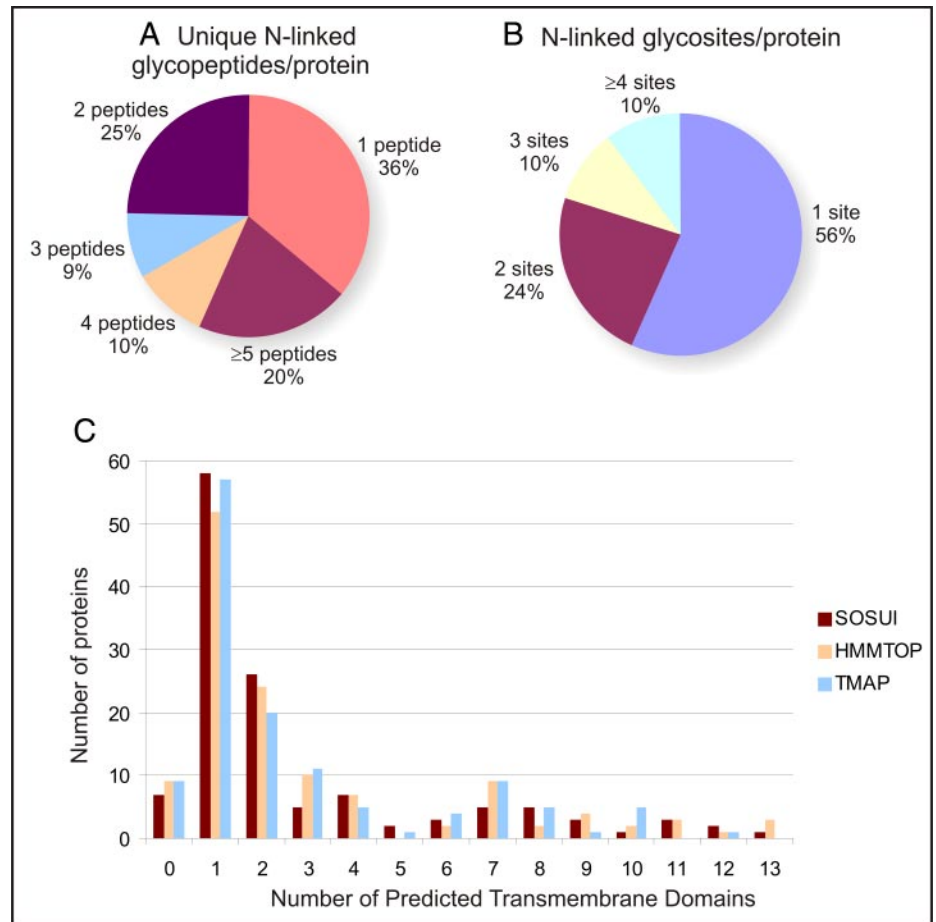


**FIG. 2. Images of myoblasts.** Bright field images of a monolayer of undifferentiated and mononuclear C2C12 myoblasts cultivated in high serum (20% FBS) (A) and a higher magnification of multinucleated C2C12 myotubes after differentiation for 5 days in low serum conditions (5% FBS) (B) are shown.

per cell, and no myotube formation was observed. Under confluent conditions ( $>80\%$ ) and after switching to low serum conditions (5% FBS), the cells spontaneously fused and formed myotubes, thus confirming their utility for studying the differentiation of non-muscle myoblast cells to skeletal muscle cells (Fig. 2).

**Database of Cell Surface N-Linked Glycoproteins on Undifferentiated C2C12 Cells—**The list of cell surface N-linked glycoproteins identified in the present study is presented in supplemental Table S1, and detailed information regarding all of the peptides attributed to each protein can be found in the supporting information (supplemental Table S8). A total of 128 N-linked cell surface glycoproteins were identified with probability scores of  $p > 0.48$ . Of these, 114 had scores of  $p > 0.9$ , corresponding to a false discovery rate of 1.1% as calculated by ProteinProphet (supplemental Table S4). Thirty-six proteins correspond to cluster of differentiation (CD) molecules (59), and all of the proteins listed in supplemental Table S1 were identified by peptides that met the following three criteria. 1) The peptide was captured by streptavidin beads, indicating that it was originally attached to a biotin-labeled oligosaccharide structure. 2) The captured peptide contains a deamidation (0.98-Da shift). 3) The deamidation occurs at asparagine within the N-linked glycosylation consensus amino acid sequence motif for N-linked glycosylation (NX(S/T)). As summarized in Fig. 3, 46 (36%) proteins were identified by a single unique glycopeptide, whereas 82 (64%) were identified by two or more unique glycopeptides. Of the 128 identified proteins, 117 have predicted transmembrane (TM) domains (based on the three prediction algorithms used), four are known GPI-anchored proteins, and five are known extracellular matrix (ECM) proteins. To provide an overview of the distribution of the number of TM domains per protein, the results from each topology prediction algorithm were averaged as not all algorithms predict the same number of TM domains for each protein (supplemental Table S1). This resulted in a total of 56, 23, and 41 proteins containing 1, 2, or  $\geq 3$  TM domains, respectively (Fig. 3C). Interestingly, unlike

**FIG. 3. Characterization of N-linked peptides and their corresponding proteins.** A, pie chart showing the distribution of the number of unique N-linked glycopeptides identified per protein, highlighting that 64% of the proteins were identified by two or more unique glycopeptides. Because the method captures only those peptides that are glycosylated, it is not expected to identify multiple peptides per protein, and this depends on whether the site of glycosylation lies within a tryptic peptide with a suitable *m/z* for MS analysis. B, pie chart showing the distribution of the number of N-linked glycosylation sites identified per protein, highlighting that two or more sites were identified for 44% of the proteins. C, bar graph showing the distribution of the number of transmembrane domains calculated using three different prediction algorithms, SOSUI, HMMTOP, and TMAP.



other methods for isolating membrane proteins, the CSC technology was able to identify proteins with as many as 13 transmembrane domains. Importantly, designation of the proteins in the current list as “cell surface proteins” is based only on the experimental data and not on database annotations. This allows for the inclusion of cell surface proteins that are not transmembrane proteins (*i.e.* GPI-anchored) and avoids mistakenly eliminating proteins that may have incomplete/ambiguous database or gene ontology term annotations regarding their subcellular localization.

**New Information Regarding Glycosylation Site Occupancy—**Identifying the site of glycosylation is useful for determining the orientation of a protein within the membrane as only the extracellular domains of plasma membrane proteins are N-linked glycosylated. As summarized in Fig. 3B, one site of glycosylation was identified for 72 (56%) proteins, whereas two or more glycosylation sites were identified for 56 (44%) proteins. For 10 proteins, the current data identified the only potential glycosylation site; whereas for eight other proteins, all predicted glycosylation sites ( $n = 2-3$ ) were observed (Table I). When determining whether the current data provided any new information regarding occupied glycosylation sites, occasionally the Swiss-Prot database predicted fewer sites than were observed in the current study. In other words, not

all NX(S/T) sites are listed in Swiss-Prot. For those proteins, EnsembleGly (60) was used to predict the number of NX(S/T) motifs in the extracellular domain, and the results from that prediction are included in Table I. In total, of the 235 N-linked glycosites identified here, 226 (96%) are not documented by experimental evidence in Swiss-Prot, meaning that the current data set adds considerable new information regarding N-linked glycosylation site occupancy for the proteins identified. Although most glycopeptide positions were consistent with predicted protein structures, the membrane orientation provided in Swiss-Prot was inconsistent with the data observed for zinc transporter ZIP14 where the glycosylation sites identified at residues 52 and 100 are annotated as in the cytoplasmic domain. In the case of 14 proteins where the membrane orientation listed in Swiss-Prot is ambiguous (*i.e.* only transmembrane domains are predicted, but no annotations are provided regarding extracellular *versus* cytoplasmic), the findings from the current study provide evidence for the correct orientation of these proteins (Table I). Five proteins identified as N-linked glycosylated in the current study are also predicted to be O-linked type glycosylated: glypican-1, thrombomodulin, neuropilin-1, basement membrane-specific heparan sulfate proteoglycan core protein, and chondroitin sulfate proteoglycan 4.

TABLE I  
N-Linked glycosite information for each protein identified via the CSC technology

The table lists the protein number (corresponds to supplemental Table S1), the protein name, the number of N-linked glycosylation sites confirmed in the current study, the number of potential N-linked glycosylation sites annotated in Swiss-Prot, whether these N-linked sites listed in Swiss-Prot are potential (i.e. the protein contains the NX(S/T) motif but no experimental evidence is available) or whether there is experimental (exp) evidence, and whether the glycopeptides identified in the current study are consistent with the extracellular domain (i.e. orientation) annotated in Swiss-Prot. Proteins are sorted by the number of potential N-linked glycosylation sites in increasing order. Y, observed glycopeptides map to predicted extracellular domain; N, observed glycopeptides map to predicted intracellular domain; NA, not applicable due to GPI or ECM; Amb, annotation regarding orientation is ambiguous. MHC, major histocompatibility complex.

Protein number	Protein name	No. identified sites	No. potential sites <sup>a</sup>	No. sites with exp evidence	Orientation consistent?	Protein number	Protein name	No. identified sites	No. potential sites <sup>a</sup>	No. sites with exp evidence	Orientation consistent?
45	Solute carrier family 2, facilitated glucose transporter member 1	1	1	0	Y	9	CD80 antigen	5	6	0	Y
63	Sphingosine 1-phosphate receptor 2 (Edg-5)	1	1	0	Y	15	Ectonucleotide pyrophosphatase/phosphodiesterase family member 1	3	6	0	Y
70	Ephrin-A5	1	1	1	NA	35	Neural cell adhesion molecule 1	1	6	1	Y
87	Lipid phosphate phosphohydrolase 2	1	1	0	Amb	37	Neuroplastin	3	6	1	Y
95	Solute carrier family 2, facilitated glucose transporter member 3	1	1	0	Y	48	Transmembrane protein 16F	3	6	0	Y
99	Mast cell antigen 32	1	1	0	Y	96	Tyrosine-protein kinase receptor UFO	1	6	0	Y
113	Sphingosine 1-phosphate receptor 5 or 8	1	1	0	Y	102	Cleft lip and palate transmembrane protein 1 homolog	1	6	0	Y
117	Ephrin-B1	1	1	0	Y	114	OX-2 membrane glycoprotein	1	6	0	Y
123	Aquaporin-1	1	1	0	Y	128	Calcitonin gene-related peptide type 1	1	6	0	Y
125	Myelin protein zero-like protein 1	1	1	0	Y	10	CD97 antigen	3	7	0	Y
3	Adipocyte adhesion molecule	1	2	0	Y	25	Hematopoietic progenitor cell antigen CD34	1	7	0	Y
18	Ephrin type-A receptor 2	2	2	0	Y	33	N-Acetylated $\alpha$ -linked acidic peptidase 2	3	7	0	Y
20	Excitatory amino acid transporter 1	2	2	0	Y	56	Poliovirus receptor-related protein 1	2	7	0	Y
21	Junctional adhesion molecule A	1	2	0	Y	73	Latrophilin 2	1	7	0	Y
44	Sodium/potassium-transporting ATPase subunit $\beta$ -3	1	2	0	Y	86	VPS10 domain-containing receptor SorCS2	1	7	0	Y
46	Translocon-associated protein $\alpha$	1	2	0	Y	103	Cadherin-2	1	7	0	Y
47	Transmembrane 4 L6 family member 1	1	2	0	Y	112	Plexin A1	1	7	0	Y
52	Zinc transporter ZIP10	2	2	0	Y	119	Ermilin-1	2	7	0	NA
59	Junctional adhesion molecule C	1	2	0	Y	1	4F2 cell surface antigen heavy chain	4	8	0	Y
62	Solute carrier family 12 member 2	1	2	0	Y	7	CD166 antigen	4	8	0	Y
67	Neutral amino acid transporter A	2	2	0	Y	23	Fibronectin	4	8	0	NA

TABLE I—continued

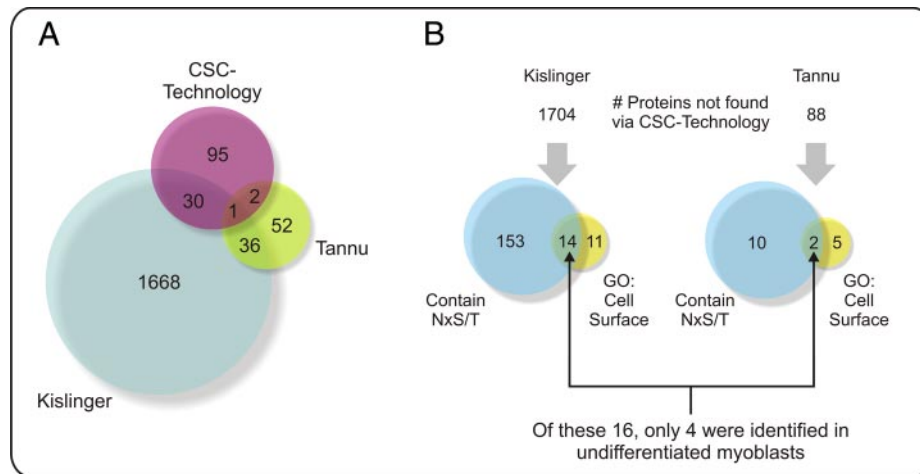
Protein number	Protein name	No. identified sites	No. potential sites <sup>a</sup>	No. sites with exp evidence	Orientation consistent?	Protein number	Protein name	No. identified sites	No. potential sites <sup>a</sup>	No. sites with exp evidence	Orientation consistent?
79	Transmembrane 9 superfamily member 3	1	2	0	Y	34	Neogenin	2	8	0	Y
91	Glypican-1	2	2	0	NA	42	Prostaglandin F <sub>2</sub> receptor negative regulator	2	8	0	Y
104	Trophoblast glycoprotein	1	2	0	Y	16	Embigin	8	9	0	Y
107	Major prion protein	1	2	0	NA	17	Endothelin-converting enzyme 1	3	10	0	Y
109	Tetraspanin-4	1	2	0	Y	60	Tyrosine-protein kinase-like 7	2	10	0	Y
120	Transmembrane protein 87A	2	2	0	Y	122	Epidermal growth factor receptor	1	10	3	Y
6	Basigin	2	3	0	Y	30	Integrin $\alpha V$	4	11	0	Y
12	Choline transporter-like protein 2	2	3	0	Y	39	$\beta$ -type platelet-derived growth factor receptor	3	11	0	Y
19	Ephrin type-B receptor 4	1	3	0	Y	97	Lysosome membrane protein 2	1	11	0	Y
24	H-1 class I histocompatibility antigen, D-K $\alpha$ chain	2	3	0	Y	5	Basement membrane-specific heparan sulfate proteoglycan core protein	4	12	0	NA
32	Macrophage mannose receptor 2	2	3	0	Y	31	Integrin $\beta 1$	2	12	0	Y
57	Protein ITFG3	3	3	0	Y	38	Oncostatin M-specific receptor subunit $\beta$	3	12	0	Y
80	Epithelial membrane protein 1	1	3	0	Amb	94	Receptor-type tyrosine-protein phosphatase $\mu$	1	12	0	Y
81	Synaptophysin-like protein 1	1	3	0	Y	4	Aminopeptidase N	5	13	0	Y
105	$\beta$ -Sarcoglycan	1	3	0	Y	27	Integrin $\alpha 3$	6	13	0	Y
106	LMBR1 domain-containing 1	1	3	0	Y	74	Lymphocyte antigen 75	1	13	0	Y
108	CMP-N-acetylneuraminatide- $\beta$ -galactosamide- $\alpha$ -2,3-sialyltransferase	1	3	0	Y	28	Integrin $\alpha 5$	5	14	0	Y
110	Immunoglobulin superfamily member 3	1	3	0	Y	13	Chondroitin sulfate proteoglycan 4	5	15	0	Y
115	Anthrax toxin receptor 1	1	3	0	Y	65	Integrin $\alpha 11$	1	16	0	Y
127	Cadherin-10	1	3	0	Y	88	Insulin-like growth factor 1 receptor	2	16	0	Y
11	Cell adhesion molecule 1	2	4	0	Y	69	Teneurin-3	2	17	0	Y
61	CD82 antigen	2	4	0	Y	98	Leucyl-cystinyl aminopeptidase	1	17	0	Y
64	Tissue factor	2	4	0	Y	68	Lysosomal membrane glycoprotein 1	1	18	2	Y
77	Solute carrier family 12 member 7	1	4	0	Y	72	Insulin receptor	1	18	0	Y
78	Solute carrier family 12, member 4	1	4	0	Y	92	Cation-independent mannose 6-phosphate receptor	1	20	0	Y
83	Ephrin type-B receptor 2	1	4	0	Y	126	Tenascin	1	20	1	NA
84	Thrombospondin	1	4	0	Y	82	Probable G-protein-coupled receptor 126	1	23	0	Y

TABLE I—continued

Protein number	Protein name	No. identified sites	No. potential sites <sup>a</sup>	No. sites with exp evidence	Orientation consistent?	Protein name	No. identified sites	No. potential sites <sup>a</sup>	No. sites with exp evidence	Orientation consistent?
85	Thrombospondin-1	1	4	0	NA	41	3	51	0	Y
100	Kin of IRRE-like protein 1	1	4	0	Y	22	3	0 (30) <sup>a</sup>	0	Amb
101	Tetraspanin-3	1	4	0	Y	76	1	0 (1) <sup>a</sup>	0	Amb
118	Transmembrane protein 87B	1	4	0	Y	14	4	0 (12) <sup>a</sup>	0	Y
2	Acid sphingomyelinase-like phosphodiesterase 3b	2	5	0	NA	40	6	0 (15) <sup>a</sup>	0	Amb
29	Integrin $\alpha 7$	1	5	0	Y	49	2	0 (15) <sup>a</sup>	0	Amb
36	Neuropilin-1	2	5	0	Y	116	1	0 (2) <sup>a</sup>	0	Amb
50	Vascular cell adhesion protein 1 (isoform 1)	2	5	0	Y	66	1	0 (3) <sup>a</sup>	0	Amb
51	Voltage-dependent calcium channel subunit $\alpha$ -2/ $\delta$ -1	3	5	0	Y	90	1	0 (3) <sup>a</sup>	0	Amb
54	Cadherin-15	2	5	0	Y	93	1	0 (3) <sup>a</sup>	0	Amb
55	Cation-dependent mannose 6-phosphate receptor	1	5	0	Y	26	2	0 (4) <sup>a</sup>	0	Amb
58	Golgi apparatus protein 1	2	5	0	Y	53	2	0 (5) <sup>a</sup>	0	N
71	Fibroblast growth factor receptor 4	1	5	0	Y	89	1	0 (5) <sup>a</sup>	0	Amb
75	P2X purinoceptor 7	1	5	0	Y	43	1	0 (8) <sup>a</sup>	0	Amb
111	Zinc transporter ZIP6	1	5	0	Y	121	3	0 (8) <sup>a</sup>	0	Amb
124	Semaphorin-7A	1	5	0	NA	8	2	1 (4) <sup>a</sup>	0	Y

<sup>a</sup> If no N-linked glycosylation sites are predicted in Swiss-Prot or if fewer are predicted than were observed, then the number of predicted N-linked glycosylation sites from EnsembleGly is provided. The first number is from Swiss-Prot; the number in parentheses is from EnsembleGly.





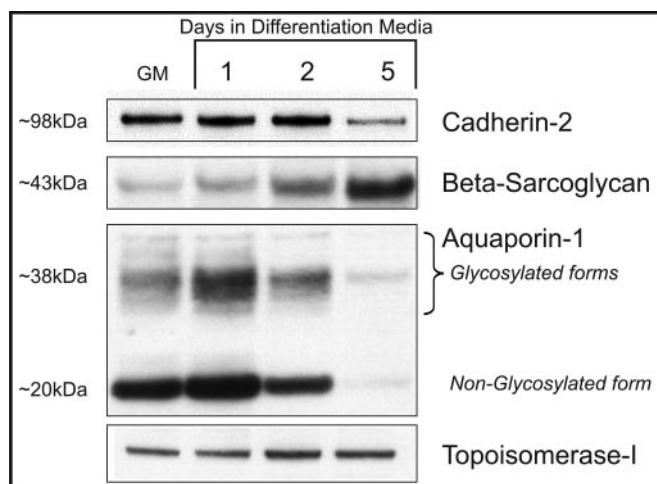
**FIG. 4. Comparison of the proteins identified by the CSC technology with those identified in other proteomics studies of C2C12 cells.** A, Venn diagram showing overlap of proteins identified in three proteomics studies each using different strategies to examine the mouse C2C12 myoblast proteome. In summary, 74% of proteins identified by the CSC technology were not identified in other, non-targeted studies. B, of the proteins identified by Kislinger *et al.* (53) and Tannu *et al.* (52) but not by the CSC technology, Venn diagrams show how many proteins are potentially N-linked (although no data exist to confirm their occupancy) and predicted to be cell surface proteins based on gene ontology (GO) term annotations, highlighting what the CSC technology may have missed. Of the 16 proteins that meet these criteria, only four were identified in undifferentiated myoblasts in the previous studies and thus could be expected to be observed in the current study. Refer to supplemental Tables S1 and S5 for proteins identified by the CSC technology but not by Kislinger *et al.* (53) and Tannu *et al.* (52).

**Comparison with Other Proteomics Studies of C2C12 Differentiation**—The list of N-linked glycoproteins identified here was compared with the data sets from two previously published global proteomics studies of undifferentiated and differentiated C2C12 cells (52, 53). Although cell culture conditions, sample handling, and mass spectrometry differed among the three studies, this comparison permitted us to assess whether the CSC technology was capable of discovering novel information. Most importantly, 74% of the proteins identified via the CSC technology were not present among the >1700 proteins identified in the previously published studies (Fig. 4). Additionally, 16 proteins containing the NX(S/T) sequence motif (although there are no data that predict these sites are actually occupied) and predicted to be localized to the cell surface were not identified by the CSC technology. Only four of these, however, were identified in the undifferentiated myoblasts in the previous proteomic studies, whereas the other 12 were only observed in latter stages of differentiation and thus are not expected to be found in the current study. In contrast, the CSC technology identified 18 proteins in the undifferentiated state that were only observed in differentiated cells in the other proteomics studies. All 18 were identified by  $\geq 2$  spectra in the current study (Table S6). For example, 97 spectra were observed for CD98 in the current study, although Kislinger *et al.* (53) only identified a single spectrum for CD98 after 10 days of differentiation. This exemplifies the ability of the CSC technology to identify proteins that may be less accessible to non-targeted methods.

**Biological Relevance of Identified Proteins in Skeletal Muscle Development and Repair**—To identify potential proteins for follow-up studies aimed at finding markers of differentia-

tion, a literature search of the potential biological relevance of each protein identified in the context of skeletal muscle development and repair was conducted (via PubMed), and the results are summarized in supplemental Table S7. Interesting proteins include M-cadherin, which is differentially expressed among satellite cells, proliferating myoblasts, and differentiated myotubes (61, 62) and is implicated in myoblast fusion (63). Other examples include thrombospondin-1, which may play a role in myoblast attachment (64), and glypican-1, which increases during myoblast differentiation and modulates myoblast proliferation and differentiation via the fibroblast growth factor 2 pathway (65–67). The information in the supplemental Table S7 is not intended to be an exhaustive list of the biological significance of all the proteins identified in the current study but rather to illustrate that a number of the proteins identified are known to have some relevance to skeletal muscle biology and, thus, could be targets of future follow-up studies focused on understanding the molecular events critical for skeletal muscle development and repair.

**Protein Abundance Changes in Cell Surface Glycoproteins with Differentiation**—To further refine the list of proteins that may be useful for future follow-up studies, previously published studies showing the change in protein abundance (52, 53) and mRNA expression (55, 56) of undifferentiated and differentiated C2C12 cells were analyzed to determine whether any of the proteins identified in the current study might show differential expression with differentiation. Seven mRNA transcripts and 33 proteins were reported to change in previous studies (supplemental Tables S2 and S3). Based on these comparisons, three proteins, aquaporin-1,  $\beta$ -sarcoglycan, and cadherin 2, were evaluated by Western blots to



**FIG. 5. Western blotting to probe for changes in protein abundance with differentiation.** Western blot images for cadherin 2 (30  $\mu\text{g}$  of total protein per lane),  $\beta$ -sarcoglycan (50  $\mu\text{g}$  of total protein per lane), and aquaporin-1 (15  $\mu\text{g}$  of total protein per lane) prepared from protein extracts of C2C12 cells grown in growth medium (GM) or in differentiation medium for 1, 2, or 5 days. Topoisomerase I loading control is representative for all blots. Molecular masses listed are approximate and are derived from the relationship to the molecular mass marker (not shown). For aquaporin-1, the Western blot shows both the glycosylated and non-glycosylated forms. The observed molecular mass for  $\beta$ -sarcoglycan is consistent with a glycosylated form, and the observed molecular mass for cadherin-2 is consistent with the non-glycosylated form.

determine whether the overall abundance of these proteins changes as the cells differentiate toward myotubes. Cells were cultivated as shown in Fig. 2, and samples were taken at 0, 1, 2, and 5 days after differentiation induction (low serum conditions). Western blotting confirmed the presence of aquaporin-1,  $\beta$ -sarcoglycan, and cadherin 2 on the undifferentiated C2C12 myoblasts; each of these proteins were identified by single peptide sequences via MS. As the myoblasts differentiated, Western blotting demonstrated a significant decrease in the overall abundance of aquaporin-1 (both glycosylated and non-glycosylated forms), a slight decrease in the overall abundance of cadherin 2, and a significant increase in the overall abundance of  $\beta$ -sarcoglycan (Fig. 5). These results are consistent with previous genomics and proteomics studies and, when combined with what is known about their biological function, highlight the potential role these proteins may play in myoblast differentiation as well as their possible use as cell surface markers.

Interestingly, the aquaporin-1 antibody used for Western blotting is reported to recognize both the glycosylated and non-glycosylated forms (see the product information), and this is consistent with what was observed in the current study where both forms were observed and appeared to decrease in abundance with differentiation. The molecular mass of  $\beta$ -sarcoglycan detected by the Western blot ( $\sim 43$  kDa) is consistent with a glycosylated form as the predicted molec-

ular mass of the native protein is  $\sim 35$  kDa. Finally, under the current conditions, only the non-glycosylated form of cadherin 2 was detected by the Western blot (both the observed and theoretical molecular mass,  $\sim 98$  kDa), although the MS data indicate that the protein is glycosylated. This may be a result of the gel and blotting experimental conditions (*i.e.* high molecular weight proteins are not transferred as efficiently), the glycosylated form may be much less abundant than the native form, or the antibody may preferentially recognize the non-glycosylated form. This exemplifies the utility of the MS-based CSC technology, which does not rely on the specificity or sensitivity of an antibody or optimization of gel conditions.

#### DISCUSSION

Using the CSC technology, the current study identified 128 cell surface *N*-linked glycoproteins on undifferentiated mouse C2C12 myoblasts; this is the largest library of cell surface *N*-linked glycoproteins described for this cell type to date. In addition to finding *N*-linked membrane glycoproteins not previously reported to be on the surface of C2C12 myoblasts, the current work adds new information about the occupancy of predicted glycosylation sites for 122 (95%) of the proteins identified as well as new information regarding protein orientation within the membrane for the proteins identified. Finally, the study provides examples of how potential markers of differentiation can be derived by starting from a characterization of the cell surface and then augmenting the data with comparisons with what is already known about protein biology as well as other proteome and transcriptome studies.

*Uncovering a Hidden Proteome*—Cell surface proteins (which include TM, GPI-anchored, and ECM proteins) are often undersampled in traditional proteomics approaches because of their relative insolubility/hydrophobicity (for TM proteins) and lower abundance compared with non-membrane proteins. To address this, a large number of studies have focused on identifying the proteins present in membrane (plasma and organelle membrane)-enriched fractions as opposed to whole cell lysates (for reviews, see Refs. 68–70). However, when using general biochemical membrane preparation techniques such as density gradients and ultracentrifugation alone, the identified membrane proteins cannot be distinguished as derived from the cell surface *versus* intracellular organelle membrane based upon experimental data. This is due, in part, to the fact that it is difficult to obtain purified plasma membrane proteins without contamination from membranes from other intracellular organelles such as the nucleus, mitochondria, endoplasmic reticulum, Golgi, and lysosome. In this case, researchers typically rely on available gene ontology or protein database annotations for classifying the subcellular location of identified proteins, although this information can be missing or incomplete in addition to the fact that a single protein may have several different locations annotated, and therefore, it may not be possible to unambiguously assign the localization of the protein in the particular cell type

examined for example. To overcome these challenges, a number of more targeted approaches have utilized creative solutions such as enzymatic “shaving” of extracellular domains on intact cells (71–77), fluorescent labeling (78, 79), lectin affinity (80–82), and biotinylation of cell surface proteins (25, 83–91). Each of these methods adds another level of specificity for plasma membrane proteins over intracellular membrane proteins. The approach used in the current study takes advantage of the fact that a majority of the cell surface proteins are glycosylated, thus allowing for their specific capture and ultimately allowing for the identification of cell surface proteins, which are less accessible to non-targeted methods. Most importantly, 74% of the proteins identified here have not been reported in previous global proteomics studies of the same cell type, highlighting the utility of this targeted approach, which effectively reduces sample complexity and allows for the identification of hydrophobic as well as lower abundance proteins.

**Importance of Confirming Glycosylation Site Occupancy—**Confirming glycosylation site occupancy is critical for determining the orientation of the protein within the membrane and antigen design for antibody development. Although publicly available protein databases (e.g. Swiss-Prot) contain information that may predict the presence of an *N*-linked glycosylation moiety due to the presence of an NX(S/T) sequence motif, they do not always provide conclusive experimental evidence that a potential glycosylation site is occupied (Table I). For 122 (95%) of the proteins identified here, none of the glycosylation sites listed in Swiss-Prot are documented by experimental evidence; thus the current data add new information regarding the occupation of these sites for most of the proteins identified. Importantly, not all occupied sites may be identified via the CSC technology as it is possible that a site may lie within a region of the protein that, after enzymatic digestion, does not result in a peptide with an appropriate *m/z* for detection. Thus, the absence of an identified site is not conclusive evidence that the site is not occupied. Although spontaneous or chemical deamidation at the asparagine within the NX(S/T) motif is possible and could lead to false positive assignments, the binding of biotin-labeled glycopeptides to the streptavidin beads enriches specifically for those peptides that are in fact glycosylated, reducing the likelihood that peptides identified in the captured fraction with deamidation at NX(S/T) were generated by chemical deamidation. It is further noted that there are several databases beyond Swiss-Prot that summarize experimental evidence regarding the occupancy of potential glycosylation sites (e.g. UniPep (92) and Human Protein Reference Database (93)). However, these resources are specific for human protein data and thus could not be used to determine whether the experimental data provided here (which uses murine cells) reports new confirmations of site occupancy. Of course, these resources are useful for predicting sites that are likely occupied in homologous proteins. In general, glycosylation of cell surface pro-

teins is critical for cell adhesion, motility, and cell-cell interactions. Specifically, previous studies have shown that glycosylation of a number of the proteins identified here affects their function. For example, glycosylation of calcitonin gene-related peptide type 1 is critical for its function as a receptor (94), glycosylation of Edg-1 affects lateralization and internalization of the receptor (95, 96), and glycosylation of NCAM has a role in attenuating myoblast fusion (97). Thus, the new knowledge regarding occupation of glycosylation sites may aid in further elucidating the biological implications of glycosylation for the proteins identified.

**Potential Markers of Myoblast Differentiation—**Aquaporin-1 was identified by the CSC technology in undifferentiated myoblasts but was not detected in the other proteomics studies of C2C12 differentiation (52, 53). Additional studies that have focused specifically on aquaporin-1 have shown the absence of aquaporin-1 on adult mouse muscle fibers (98), and similar results have been observed in rat (99), although its presence has been reported for human adult skeletal muscle (100, 101). The mRNA levels for aquaporin-1 were previously found to decrease significantly with myoblast differentiation (55, 56). Our results, which showed a decrease in the overall abundance of aquaporin-1 with differentiation, are therefore consistent with these previous studies. This change is intriguing as fluid transport, which is a function of aquaporin-1, and an increase in cell volume are important processes in muscle repair after injury (e.g. intense activity) (102–104).

$\beta$ -Sarcoglycan was identified by four spectra (one unique peptide) via the CSC technology in undifferentiated C2C12 myoblasts and three spectra in the Kislinger *et al.* (53) proteomics study. The Kislinger *et al.* (53) study shows that the number of spectra increased from one in undifferentiated myoblasts to three after 6 days of differentiation, a trend that is consistent with the observations by Western blotting in the current study. Also consistent with these observations are previous genomics studies that show an increase in  $\beta$ -sarcoglycan mRNA with differentiation (56). In skeletal and cardiac muscle,  $\beta$ -sarcoglycan is a member of the dystrophin-associated glycoprotein complex, a complex important for signaling and protecting muscle from contraction-induced injury (105) and required for maintenance of the sarcolemma (106, 107). Mutations in  $\beta$ -sarcoglycan, as well as other sarcoglycans, are associated with muscular dystrophy (108, 109).

Like  $\beta$ -sarcoglycan, cadherin 2 also has a well known biological role in skeletal muscle. Cadherin 2 (N-cadherin) was identified via the CSC technology, although it was not detected in the other proteomics studies of C2C12 differentiation (52, 53). Cadherin 2 is involved in calcium-dependent myoblast fusion during myogenesis (110–112), and the cadherin 2-catenin complex has been shown to be required for promoting differentiation in skeletal muscle (11, 110, 113–115).

Taken together, aquaporin-1,  $\beta$ -sarcoglycan, and cadherin 2 have both potential and known roles in muscle development

and repair, and thus, understanding their temporal patterns of protein expression on the cell surface, for example, may help in understanding the molecular mechanisms involved in skeletal muscle development and repair. One of the limitations currently faced is that the antibodies used in the current study were not developed against extracellular epitopes. However, if antibodies are developed that recognize the extracellular domain of the proteins, then they could be used in lineage tracing experiments, for example, because they would not require cell permeabilization. In this case, the information generated in the current study regarding orientation of the protein within the membrane as well as sites of glycosylation could aid in the development of suitable antibodies that could serve as truly valuable lineage markers.

In addition to the proteins that were shown to change in abundance with differentiation via Western blotting, several other proteins are of interest as they have also been identified in a proteomics analysis of the lipid rafts of satellite cells, which are developmental precursors of myoblasts (54). Five proteins were found in both studies (supplemental Table S1): CD56/NCAM, basigin/CD147, tyrosine-protein kinase-like 7, integrin  $\beta$ 1, and neurotrophin receptor-associated death domain. Of these, NCAM is particularly interesting as a potential marker of myoblast differentiation because, as in previous studies, it was either not found or rarely found in quiescent mouse satellite cells (62, 116) but rather was increasingly found in differentiating satellite cells and differentiating mouse myoblasts (54, 62).

**Summary and Conclusions**—In general, proteomics approaches to studying the cell surface are expected to add a welcomed complement to the data generated using flow cytometry, antibody arrays, and microscopy (117–119). Specifically, approaches that provide unambiguous information regarding the localization of the protein to the cell surface will be particularly useful. This is due to the fact that markers used for the selection and subsequent expansion of a particular cell type are, ideally, naturally occurring markers that are accessible to antibody binding without disruption of the cell. Therefore, advantages offered by the CSC technology compared with other proteomics methods are the ability 1) to identify *bona fide* cell surface proteins based upon the experimental data without relying on potentially incomplete database annotations and 2) provide confirmation of the membrane orientation of proteins and modifications that will be important for epitope selection and antibody design.

The current study provides a work flow for identifying potential cell surface markers of differentiation that includes 1) a targeted approach to efficiently access the plasma membrane proteome and 2) combining the results with what is known about mRNA expression, protein expression, and biological function. In summary, the work flow begins with a characterization of the cell surface and subsequently utilizes what is known about the genome and proteome to narrow the list of interesting proteins. The proteins selected via this process

were found to change in abundance with differentiation of the myoblasts toward myotubes and thus complement the collection of cell surface markers already known to characterize some of the cell types present during skeletal muscle development and repair (for reviews, see Refs. 2 and 120). Although there are currently a number of proteins described as cell surface markers (e.g. CD34, NCAM, and M-cadherin) for the differentiation of myoblasts, they are often present on heterogeneous subpopulations of cell types/stages (for reviews, see Refs. 2 and 120). Therefore, the need for refined panels of markers that can identify homogeneous populations of developmental intermediates is clear. Discovery-driven proteomics strategies, like the CSC technology, can now provide the rationale for the development of protein-specific antibodies against preselected differentiation marker candidates for subsequent single cell studies.

**Acknowledgments**—We thank Dr. Alexander Schmidt and Simon Sheng for technical assistance.

\* This work was supported, in whole or in part, by the National Institutes of Health through funding from the Intramural Research Program, NIA (to K. R. B.); Pathway to Independence Award K99-L094708-01 (to R. L. G.); NHLBI Proteomics Innovation Contract N01-HV-28180 (to J. E. V.); and Grant PH SCCOR-P50-HL-084946-01 (to J. E. V.). This work was also supported by NCCR Neural Plasticity and Repair of the Swiss National Science Foundation (to B. W.).

☒ The on-line version of this article (available at <http://www.mcponline.org>) contains supplemental Tables S1–S8.

¶ These authors contributed equally to this work.

\*\* These authors shared leadership.

¶¶ To whom correspondence should be addressed: Dept. of Medicine, Johns Hopkins University, Rm. 602, Mason F. Lord Bldg., Center tower, Baltimore, MD 21224. E-mail: [jvaneyk1@jhmi.edu](mailto:jvaneyk1@jhmi.edu).

## REFERENCES

1. Chargé, S. B., and Rudnicki, M. A. (2004) Cellular and molecular regulation of muscle regeneration. *Physiol. Rev.* **84**, 209–238
2. Hawke, T. J., and Garry, D. J. (2001) Myogenic satellite cells: physiology to molecular biology. *J. Appl. Physiol.* **91**, 534–551
3. Le Grand, F., and Rudnicki, M. A. (2007) Skeletal muscle satellite cells and adult myogenesis. *Curr. Opin. Cell Biol.* **19**, 628–633
4. Seale, P., and Rudnicki, M. A. (2000) A new look at the origin, function, and “stem-cell” status of muscle satellite cells. *Dev. Biol.* **218**, 115–124
5. Mauro, A. (1961) Satellite cell of skeletal muscle fibers. *J. Biophys. Biochem. Cytol.* **9**, 493–495
6. Bischoff, R. (1990) Interaction between satellite cells and skeletal muscle fibers. *Development* **109**, 943–952
7. Appell, H. J., Forsberg, S., and Hollmann, W. (1988) Satellite cell activation in human skeletal muscle after training: evidence for muscle fiber neof ormation. *Int. J. Sports Med.* **9**, 297–299
8. Horsley, V., and Pavlath, G. K. (2004) Forming a multinucleated cell: molecules that regulate myoblast fusion. *Cells Tissues Organs* **176**, 67–78
9. Geiger, B., and Ayalon, O. (1992) Cadherins. *Annu. Rev. Cell Biol.* **8**, 307–332
10. Kaufmann, U., Martin, B., Link, D., Witt, K., Zeitler, R., Reinhard, S., and Starzinski-Powitz, A. (1999) M-cadherin and its sisters in development of striated muscle. *Cell Tissue Res.* **296**, 191–198
11. Redfield, A., Nieman, M. T., and Knudsen, K. A. (1997) Cadherins promote skeletal muscle differentiation in three-dimensional cultures. *J. Cell Biol.* **138**, 1323–1331
12. Kang, J. S., Yi, M. J., Zhang, W., Feinleib, J. L., Cole, F., and Krauss, R. S.



- (2004) Netrins and neogenin promote myotube formation. *J. Cell Biol.* **167**, 493–504
13. Krauss, R. S., Cole, F., Gaio, U., Takaesu, G., Zhang, W., and Kang, J. S. (2005) Close encounters: regulation of vertebrate skeletal myogenesis by cell-cell contact. *J. Cell Sci.* **118**, 2355–2362
  14. Rapizzi, E., Donati, C., Cencetti, F., Nincheri, P., and Bruni, P. (2008) Sphingosine 1-phosphate differentially regulates proliferation of C2C12 reserve cells and myoblasts. *Mol. Cell. Biochem.* **314**, 193–199
  15. Meacci, E., Nuti, F., Donati, C., Cencetti, F., Farnararo, M., and Bruni, P. (2008) Sphingosine kinase activity is required for myogenic differentiation of C2C12 myoblasts. *J. Cell. Physiol.* **214**, 210–220
  16. Squecco, R., Sassoli, C., Nuti, F., Martinesi, M., Chellini, F., Nosi, D., Zecchi-Orlandini, S., Francini, F., Formigli, L., and Meacci, E. (2006) Sphingosine 1-phosphate induces myoblast differentiation through Cx43 protein expression: a role for a gap junction-dependent and -independent function. *Mol. Biol. Cell* **17**, 4896–4910
  17. Donati, C., Meacci, E., Nuti, F., Becciolini, L., Farnararo, M., and Bruni, P. (2005) Sphingosine 1-phosphate regulates myogenic differentiation: a major role for S1P2 receptor. *FASEB J.* **19**, 449–451
  18. 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
  19. Gemeiner, P., and Viskupic, E. (1981) Stepwise immobilization of proteins via their glycosylation. *J. Biochem. Biophys. Methods* **4**, 309–319
  20. 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
  21. Zhang, H., Liu, A. Y., Loriaux, P., Wollscheid, B., Zhou, Y., Watts, J. D., and Aebersold, R. (2007) Mass spectrometric detection of tissue proteins in plasma. *Mol. Cell. Proteomics* **6**, 64–71
  22. Liu, T., Qian, W. J., Gritsenko, M. A., Camp, D. G., 2nd, Monroe, M. E., Moore, R. J., and Smith, R. D. (2005) Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. *J. Proteome Res.* **4**, 2070–2080
  23. Liu, T., Qian, W. J., Gritsenko, M. A., Xiao, W., Moldawer, L. L., Kaushal, A., Monroe, M. E., Varnum, S. M., Moore, R. J., Purvine, S. O., Maier, R. V., Davis, R. W., Tompkins, R. G., Camp, D. G., 2nd, and Smith, R. D. (2006) High dynamic range characterization of the trauma patient plasma proteome. *Mol. Cell. Proteomics* **5**, 1899–1913
  24. Sun, B., Ranish, J. A., Utleg, A. G., White, J. T., Yan, X., Lin, B., and Hood, L. (2007) Shotgun glycopeptide capture approach coupled with mass spectrometry for comprehensive glycoproteomics. *Mol. Cell. Proteomics* **6**, 141–149
  25. Zhang, W., Zhou, G., Zhao, Y., White, M. A., and Zhao, Y. (2003) Affinity enrichment of plasma membrane for proteomics analysis. *Electrophoresis* **24**, 2855–2863
  26. Lewandrowski, U., Moebius, J., Walter, U., and Sickmann, A. (2006) Elucidation of N-glycosylation sites on human platelet proteins: a glycoproteomic approach. *Mol. Cell. Proteomics* **5**, 226–233
  27. 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
  28. Schiess, R., Mueller, L. N., Schmidt, A., Mueller, M., Wollscheid, B., and Aebersold, R. (2009) Analysis of cell surface proteome changes via label-free, quantitative mass spectrometry. *Mol. Cell Proteomics* **8**, 624–638
  29. 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
  30. Yaffe, D., and Saxel, O. (1977) Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature* **270**, 725–727
  31. Blau, H. M., Pavlath, G. K., Hardeman, E. C., Chiu, C. P., Silberstein, L., Webster, S. G., Miller, S. C., and Webster, C. (1985) Plasticity of the differentiated state. *Science* **230**, 758–766
  32. Burattini, S., Ferri, P., Battistelli, M., Curci, R., Luchetti, F., and Falcieri, E. (2004) C2C12 murine myoblasts as a model of skeletal muscle development: morpho-functional characterization. *Eur. J. Histochem.* **48**, 223–233
  33. Molkentin, J. D., and Olson, E. N. (1996) Defining the regulatory networks for muscle development. *Curr. Opin. Genet. Dev.* **6**, 445–453
  34. Capanni, C., Del Coco, R., Squarzonni, S., Columbaro, M., Mattioli, E., Camozzi, D., Rocchi, A., Scotlandi, K., Maraldi, N., Foisner, R., and Lattanzi, G. (2008) Prelamin A is involved in early steps of muscle differentiation. *Exp. Cell Res.* **314**, 3628–3637
  35. Murray, T. V., McMahon, J. M., Howley, B. A., Stanley, A., Ritter, T., Mohr, A., Zwacka, R., and Fearnhead, H. O. (2008) A non-apoptotic role for caspase-9 in muscle differentiation. *J. Cell Sci.* **121**, 3786–3793
  36. Buas, M. F., Kabak, S., and Kadesch, T. (2009) Inhibition of myogenesis by Notch: evidence for multiple pathways. *J. Cell Physiol.* **218**, 84–93
  37. Favreau, C., Delbarre, E., Courvalin, J. C., and Buendia, B. (2008) Differentiation of C2C12 myoblasts expressing lamin A mutated at a site responsible for Emery-Dreifuss muscular dystrophy is improved by inhibition of the MEK-ERK pathway and stimulation of the PI3-kinase pathway. *Exp. Cell Res.* **314**, 1392–1405
  38. Li, B., Lin, M., Tang, Y., Wang, B., and Wang, J. H. (2008) A novel functional assessment of the differentiation of micropatterned muscle cells. *J. Biomech.* **41**, 3349–3353
  39. Fanzani, A., Stoppani, E., Gualandi, L., Giuliani, R., Galbiati, F., Rossi, S., Fra, A., Preti, A., and Marchesini, S. (2007) Phenotypic behavior of C2C12 myoblasts upon expression of the dystrophy-related caveolin-3 P104L and TFT mutants. *FEBS Lett.* **581**, 5099–5104
  40. Bonacchi, M., Nistri, S., Nanni, C., Gelsomino, S., Pini, A., Cinci, L., Maiani, M., Zecchi-Orlandini, S., Lorusso, R., Fanti, S., Silvertown, J., and Bani, D. (2009) Functional and histopathological improvement of the post-infarcted rat heart upon myoblast cell grafting and relaxin therapy. *J. Cell Mol. Med.*, in press
  41. Formigli, L., Perna, A. M., Meacci, E., Cinci, L., Margheri, M., Nistri, S., Tani, A., Silvertown, J., Orlandini, G., Porciani, C., Zecchi-Orlandini, S., Medin, J., and Bani, D. (2007) Paracrine effects of transplanted myoblasts and relaxin on post-infarction heart remodelling. *J. Cell. Mol. Med.* **11**, 1087–1100
  42. Reinecke, H., and Murry, C. E. (2000) Transmural replacement of myocardium after skeletal myoblast grafting into the heart. Too much of a good thing? *Cardiovasc. Pathol.* **9**, 337–344
  43. Minty, A., Blau, H., and Kedes, L. (1986) Two-level regulation of cardiac actin gene transcription: muscle-specific modulating factors can accumulate before gene activation. *Mol. Cell. Biol.* **6**, 2137–2148
  44. Koban, M. U., Brugh, S. A., Riordon, D. R., Dellow, K. A., Yang, H. T., Tweedie, D., and Boheler, K. R. (2001) A distant upstream region of the rat multipartite Na(+)-Ca(2+) exchanger NCX1 gene promoter is sufficient to confer cardiac-specific expression. *Mech. Dev.* **109**, 267–279
  45. Kersey, P. J., Duarte, J., Williams, A., Karavidopoulou, Y., Birney, E., and Apweiler, R. (2004) The International Protein Index: an integrated database for proteomics experiments. *Proteomics* **4**, 1985–1988
  46. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392
  47. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646–4658
  48. Tusnády, G. E., and Simon, I. (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**, 849–850
  49. Tusnády, G. E., and Simon, I. (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J. Mol. Biol.* **283**, 489–506
  50. Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14**, 378–379
  51. Persson, B., and Argos, P. (1997) Prediction of membrane protein topology utilizing multiple sequence alignments. *J. Protein Chem.* **16**, 453–457
  52. Tannu, N. S., Rao, V. K., Chaudhary, R. M., Giorgianni, F., Saeed, A. E., Gao, Y., and Raghov, R. (2004) Comparative proteomes of the proliferating C2C12 myoblasts and fully differentiated myotubes reveal the complexity of the skeletal muscle differentiation program. *Mol. Cell. Proteomics* **3**, 1065–1082
  53. Kislinger, T., Gramolini, A. O., Pan, Y., Rahman, K., MacLennan, D. H., and

- Emili, A. (2005) Proteome dynamics during C2C12 myoblast differentiation. *Mol. Cell. Proteomics* **4**, 887–901
54. Capkovic, K. L., Stevenson, S., Johnson, M. C., Thelen, J. J., and Cornelison, D. D. (2008) Neural cell adhesion molecule (NCAM) marks adult myogenic cells committed to differentiation. *Exp. Cell Res.* **314**, 1553–1565
  55. Moran, J. L., Li, Y., Hill, A. A., Mounts, W. M., and Miller, C. P. (2002) Gene expression changes during mouse skeletal myoblast differentiation revealed by transcriptional profiling. *Physiol. Genomics* **10**, 103–111
  56. Tomczak, K. K., Marinescu, V. D., Ramoni, M. F., Sanoudou, D., Montanaro, F., Han, M., Kunkel, L. M., Kohane, I. S., and Beggs, A. H. (2004) Expression profiling and identification of novel genes involved in myogenic differentiation. *FASEB J.* **18**, 403–405
  57. Côté, R. G., Jones, P., Martens, L., Kerrier, S., Reisinger, F., Lin, Q., Leinonen, R., Apweiler, R., and Hermjakob, H. (2007) The Protein Identifier Cross-Referencing (PICR) service: reconciling protein identifiers across multiple source databases. *BMC Bioinformatics* **8**, 401
  58. Alibés, A., Yankilevich, P., Cañada, A., and Díaz-Uriarte, R. (2007) IDconverter and IDClight: conversion and annotation of gene and protein IDs. *BMC Bioinformatics* **8**, 9
  59. Zola, H., Swart, B., Banham, A., Barry, S., Beare, A., Bensussan, A., Boumsell, L., Buckley, C. D., Bühring, H. J., Clark, G., Engel, P., Fox, D., Jin, B. Q., Macardle, P. J., Malavasi, F., Mason, D., Stockinger, H., and Yang, X. (2007) CD molecules 2006—human cell differentiation molecules. *J. Immunol. Methods* **319**, 1–5
  60. Caragea, C., Sinapov, J., Silvescu, A., Dobbs, D., and Honavar, V. (2007) Glycosylation site prediction using ensembles of Support Vector Machine classifiers. *BMC Bioinformatics* **8**, 438
  61. Bornemann, A., and Schmalbruch, H. (1994) Immunocytochemistry of M-cadherin in mature and regenerating rat muscle. *Anat. Rec.* **239**, 119–125
  62. Irintchev, A., Zeschnig, M., Starzinski-Powitz, A., and Wernig, A. (1994) Expression pattern of M-cadherin in normal, denervated, and regenerating mouse muscles. *Dev. Dyn.* **199**, 326–337
  63. Wróbel, E., Brzóska, E., and Moraczewski, J. (2007) M-cadherin and beta-catenin participate in differentiation of rat satellite cells. *Eur. J. Cell Biol.* **86**, 99–109
  64. Adams, J. C., and Lawler, J. (1994) Cell-type specific adhesive interactions of skeletal myoblasts with thrombospondin-1. *Mol. Biol. Cell* **5**, 423–437
  65. Zhang, X., Liu, C., Nestor, K. E., McFarland, D. C., and Velleman, S. G. (2007) The effect of glypican-1 glycosaminoglycan chains on turkey myogenic satellite cell proliferation, differentiation, and fibroblast growth factor 2 responsiveness. *Poult. Sci.* **86**, 2020–2028
  66. Velleman, S. G., Coy, C. S., and McFarland, D. C. (2007) Effect of syndecan-1, syndecan-4, and glypican-1 on turkey muscle satellite cell proliferation, differentiation, and responsiveness to fibroblast growth factor 2. *Poult. Sci.* **86**, 1406–1413
  67. Velleman, S. G., Liu, C., Coy, C. S., and McFarland, D. C. (2006) Effects of glypican-1 on turkey skeletal muscle cell proliferation, differentiation and fibroblast growth factor 2 responsiveness. *Dev. Growth Differ.* **48**, 271–276
  68. Speers, A. E., and Wu, C. C. (2007) Proteomics of integral membrane proteins—theory and application. *Chem. Rev.* **107**, 3687–3714
  69. Macher, B. A., and Yen, T. Y. (2007) Proteins at membrane surfaces—a review of approaches. *Mol. Biosyst.* **3**, 705–713
  70. Josic, D., and Clifton, J. G. (2007) Mammalian plasma membrane proteomics. *Proteomics* **7**, 3010–3029
  71. Parnas, D., and Linal, M. (1997) Acceleration of neuronal maturation of P19 cells by increasing culture density. *Brain Res. Dev. Brain Res.* **101**, 115–124
  72. Tjalsma, H., Lambooy, L., Hermans, P. W., and Swinkels, D. W. (2008) Shedding & shaving: disclosure of proteomic expressions on a bacterial face. *Proteomics* **8**, 1415–1428
  73. Wu, C. C., MacCoss, M. J., Howell, K. E., and Yates, J. R., 3rd (2003) A method for the comprehensive proteomic analysis of membrane proteins. *Nat. Biotechnol.* **21**, 532–538
  74. Le Bihan, T., Goh, T., Stewart, I. I., Salter, A. M., Bukhman, Y. V., Dharsee, M., Ewing, R., and Wiñaniewski, J. R. (2006) Differential analysis of membrane proteins in mouse fore- and hindbrain using a label-free approach. *J. Proteome Res.* **5**, 2701–2710
  75. Rodríguez-Ortega, M. J., Norais, N., Bensi, G., Liberatori, S., Capo, S., Mora, M., Scarselli, M., Doro, F., Ferrari, G., Garaguso, I., Maggi, T., Neumann, A., Covre, A., Telford, J. L., and Grandi, G. (2006) Characterization and identification of vaccine candidate proteins through analysis of the group A Streptococcus surface proteome. *Nat. Biotechnol.* **24**, 191–197
  76. Wei, C., Yang, J., Zhu, J., Zhang, X., Leng, W., Wang, J., Xue, Y., Sun, L., Li, W., Wang, J., and Jin, Q. (2006) Comprehensive proteomic analysis of Shigella flexneri 2a membrane proteins. *J. Proteome Res.* **5**, 1860–1865
  77. Fischer, F., Wolters, D., Rögner, M., and Poetsch, A. (2006) Toward the complete membrane proteome: high coverage of integral membrane proteins through transmembrane peptide detection. *Mol. Cell. Proteomics* **5**, 444–453
  78. Khemiri, A., Galland, A., Vaudry, D., Chan Tchi Song, P., Vaudry, H., Jouenne, T., and Cosette, P. (2008) Outer-membrane proteomic maps and surface-exposed proteins of Legionella pneumophila using cellular fractionation and fluorescent labelling. *Anal. Bioanal. Chem.* **390**, 1861–1871
  79. Sidibe, A., Yin, X., Tarelli, E., Xiao, Q., Zampetaki, A., Xu, Q., and Mayr, M. (2007) Integrated membrane protein analysis of mature and embryonic stem cell-derived smooth muscle cells using a novel combination of CyDye/biotin labeling. *Mol. Cell. Proteomics* **6**, 1788–1797
  80. Fan, X., She, Y. M., Bagshaw, R. D., Callahan, J. W., Schachter, H., and Mahuran, D. J. (2004) A method for proteomic identification of membrane-bound proteins containing Asn-linked oligosaccharides. *Anal. Biochem.* **332**, 178–186
  81. Fan, X., She, Y. M., Bagshaw, R. D., Callahan, J. W., Schachter, H., and Mahuran, D. J. (2005) Identification of the hydrophobic glycoproteins of Caenorhabditis elegans. *Glycobiology* **15**, 952–964
  82. Ghosh, D., Krokhn, O., Antonovici, M., Ens, W., Standing, K. G., Beavis, R. C., and Wilkins, J. A. (2004) Lectin affinity as an approach to the proteomic analysis of membrane glycoproteins. *J. Proteome Res.* **3**, 841–850
  83. Goshe, M. B., Blonder, J., and Smith, R. D. (2003) Affinity labeling of highly hydrophobic integral membrane proteins for proteome-wide analysis. *J. Proteome Res.* **2**, 153–161
  84. Zhao, Y., Zhang, W., Kho, Y., and Zhao, Y. (2004) Proteomic analysis of integral plasma membrane proteins. *Anal. Chem.* **76**, 1817–1823
  85. Perosa, F., Luccarelli, G., Neri, M., De Pinto, V., Ferrone, S., and Dammacco, F. (1998) Evaluation of biotinylated cells as a source of antigens for characterization of their molecular profile. *Int. J. Clin. Lab. Res.* **28**, 246–251
  86. Castronovo, V., Waltregny, D., Kischel, P., Roesli, C., Elia, G., Rybak, J. N., and Neri, D. (2006) A chemical proteomics approach for the identification of accessible antigens expressed in human kidney cancer. *Mol. Cell. Proteomics* **5**, 2083–2091
  87. Scheurer, S. B., Rybak, J. N., Roesli, C., Brunisholz, R. A., Potthast, F., Schlapbach, R., Neri, D., and Elia, G. (2005) Identification and relative quantification of membrane proteins by surface biotinylation and two-dimensional peptide mapping. *Proteomics* **5**, 2718–2728
  88. Nunomura, K., Nagano, K., Itagaki, C., Taoka, M., Okamura, N., Yamai, Y., Sugano, S., Takahashi, N., Izumi, T., and Isobe, T. (2005) Cell surface labeling and mass spectrometry reveal diversity of cell surface markers and signaling molecules expressed in undifferentiated mouse embryonic stem cells. *Mol. Cell. Proteomics* **4**, 1968–1976
  89. Peirce, M. J., Wait, R., Begum, S., Saklatvala, J., and Cope, A. P. (2004) Expression profiling of lymphocyte plasma membrane proteins. *Mol. Cell. Proteomics* **3**, 56–65
  90. Yu, M. J., Pisitkun, T., Wang, G., Shen, R. F., and Knepper, M. A. (2006) LC-MS/MS analysis of apical and basolateral plasma membranes of rat renal collecting duct cells. *Mol. Cell. Proteomics* **5**, 2131–2145
  91. Sostaric, E., Georgiou, A. S., Wong, C. H., Watson, P. F., Holt, W. V., and Fazeli, A. (2006) Global profiling of surface plasma membrane proteome of oviductal epithelial cells. *J. Proteome Res.* **5**, 3029–3037
  92. Zhang, H., Loriaux, P., Eng, J., Campbell, D., Keller, A., Moss, P., Bonneau, R., Zhang, N., Zhou, Y., Wollscheid, B., Cooke, K., Yi, E. C., Lee, H., Peskind, E. R., Zhang, J., Smith, R. D., and Aebersold, R. (2006) UniPeP—a database for human N-linked glycosites: a resource for biomarker discovery. *Genome Biol.* **7**, R73
  93. Mishra, G. R., Suresh, M., Kumaran, K., Kannabiran, N., Suresh, S., Bala,

- P., Shivakumar, K., Anuradha, N., Reddy, R., Raghavan, T. M., Menon, S., Hanumanthu, G., Gupta, M., Upendran, S., Gupta, S., Mahesh, M., Jacob, B., Mathew, P., Chatterjee, P., Arun, K. S., Sharma, S., Chandrika, K. N., Deshpande, N., Palvankar, K., Raghavath, R., Krishnakanth, R., Karathia, H., Rekha, B., Nayak, R., Vishnupriya, G., Kumar, H. G., Nagini, M., Kumar, G. S., Jose, R., Deepthi, P., Mohan, S. S., Gandhi, T. K., Harsha, H. C., Deshpande, K. S., Sarker, M., Prasad, T. S., and Pandey, A. (2006) Human protein reference database—2006 update. *Nucleic Acids Res.* **34**, D411–D414
94. Kamitani, S., and Sakata, T. (2001) Glycosylation of human CRLR at Asn123 is required for ligand binding and signaling. *Biochim. Biophys. Acta* **1539**, 131–139
95. Kohno, T., and Igarashi, Y. (2004) Roles for N-glycosylation in the dynamics of Edg-1/S1P1 in sphingosine 1-phosphate-stimulated cells. *Glycoconj. J.* **21**, 497–501
96. Kohno, T., Wada, A., and Igarashi, Y. (2002) N-Glycans of sphingosine 1-phosphate receptor Edg-1 regulate ligand-induced receptor internalization. *FASEB J.* **16**, 983–992
97. Suzuki, M., Angata, K., Nakayama, J., and Fukuda, M. (2003) Polysialic acid and mucin type O-glycans on the neural cell adhesion molecule differentially regulate myoblast fusion. *J. Biol. Chem.* **278**, 49459–49468
98. Frigeri, A., Nicchia, G. P., Balena, R., Nico, B., and Svelto, M. (2004) Aquaporins in skeletal muscle: reassessment of the functional role of aquaporin-4. *FASEB J.* **18**, 905–907
99. Nielsen, S., Smith, B. L., Christensen, E. I., and Agre, P. (1993) Distribution of the aquaporin CHIP in secretory and resorptive epithelia and capillary endothelia. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7275–7279
100. Au, C. G., Cooper, S. T., Lo, H. P., Compton, A. G., Yang, N., Wintour, E. M., North, K. N., and Winlaw, D. S. (2004) Expression of aquaporin 1 in human cardiac and skeletal muscle. *J. Mol. Cell. Cardiol.* **36**, 655–662
101. Jimi, T., Wakayama, Y., Inoue, M., Kojima, H., Oniki, H., Matsuzaki, Y., Shibuya, S., Hara, H., and Takahashi, J. (2006) Aquaporin 1: examination of its expression and localization in normal human skeletal muscle tissue. *Cells Tissues Organs* **184**, 181–187
102. Taylor, S. R., Neering, I. R., Quesenberry, L. A., and Morris, V. A. (1992) Volume changes during contraction of isolated frog muscle fibers. *Adv. Exp. Med. Biol.* **311**, 91–101
103. Neering, I. R., Quesenberry, L. A., Morris, V. A., and Taylor, S. R. (1991) Nonuniform volume changes during muscle contraction. *Biophys. J.* **59**, 926–933
104. Sjogaard, G., Adams, R. P., and Saltin, B. (1985) Water and ion shifts in skeletal muscle of humans with intense dynamic knee extension. *Am. J. Physiol.* **248**, R190–R196
105. Li, D., Long, C., Yue, Y., and Duan, D. (2009) Sub-physiological sarcoglycan expression contributes to compensatory muscle protection in mdx mice. *Hum. Mol. Genet.* **18**, 1209–1220
106. Hack, A. A., Lam, M. Y., Cordier, L., Shoturma, D. I., Ly, C. T., Hadhazy, M. A., Hadhazy, M. R., Sweeney, H. L., and McNally, E. M. (2000) Differential requirement for individual sarcoglycans and dystrophin in the assembly and function of the dystrophin-glycoprotein complex. *J. Cell Sci.* **113**, 2535–2544
107. Ozawa, E., Mizuno, Y., Hagiwara, Y., Sasaoka, T., and Yoshida, M. (2005) Molecular and cell biology of the sarcoglycan complex. *Muscle Nerve* **32**, 563–576
108. Bönnemann, C. G., Modi, R., Noguchi, S., Mizuno, Y., Yoshida, M., Gussoni, E., McNally, E. M., Duggan, D. J., Angelini, C., and Hoffman, E. P. (1995) Beta-sarcoglycan (A3b) mutations cause autosomal recessive muscular dystrophy with loss of the sarcoglycan complex. *Nat. Genet.* **11**, 266–273
109. Lim, L. E., Duclos, F., Broux, O., Bourg, N., Sunada, Y., Allamand, V., Meyer, J., Richard, I., Moomaw, C., Slaughter, C., Tomé, F. M. S., Fardeau, M., Jackson, C. E., Beckmann, J. S., and Campbell, K. P. (1995) Beta-sarcoglycan: characterization and role in limb-girdle muscular dystrophy linked to 4q12. *Nat. Genet.* **11**, 257–265
110. Knudsen, K. A., Myers, L., and McElwee, S. A. (1990) A role for the Ca<sup>2+</sup>-dependent adhesion molecule, N-cadherin, in myoblast interaction during myogenesis. *Exp. Cell Res.* **188**, 175–184
111. Mege, R. M., Goudou, D., Diaz, C., Nicolet, M., Garcia, L., Geraud, G., and Rieger, F. (1992) N-cadherin and N-CAM in myoblast fusion: compared localisation and effect of blockade by peptides and antibodies. *J. Cell Sci.* **103**, 897–906
112. Cifuentes-Diaz, C., Nicolet, M., Goudou, D., Rieger, F., and Mège, R. M. (1993) N-cadherin and N-CAM-mediated adhesion in development and regeneration of skeletal muscle. *Neuromuscul. Disord.* **3**, 361–365
113. George-Weinstein, M., Gerhart, J., Blitz, J., Simak, E., and Knudsen, K. A. (1997) N-cadherin promotes the commitment and differentiation of skeletal muscle precursor cells. *Dev. Biol.* **185**, 14–24
114. Holt, C. E., Lemaire, P., and Gurdon, J. B. (1994) Cadherin-mediated cell interactions are necessary for the activation of MyoD in *Xenopus* mesoderm. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10844–10848
115. Knudsen, K. A. (1990) Cell adhesion molecules in myogenesis. *Curr. Opin. Cell Biol.* **2**, 902–906
116. Grounds, M. D. (1991) Towards understanding skeletal muscle regeneration. *Pathol. Res. Pract.* **187**, 1–22
117. Gundry, R. L., Boheler, K. R., Van Eyk, J. E., and Wollscheid, B. (2008) A novel role for proteomics in the discovery of cell-surface markers on stem cells: scratching the surface. *Proteomics Clin. Appl.* **2**, 892–903
118. Dormeyer, W., van Hoof, D., Braam, S. R., Heck, A. J., Mummery, C. L., and Krijgsvelde, J. (2008) Plasma membrane proteomics of human embryonic stem cells and human embryonal carcinoma cells. *J. Proteome Res.* **7**, 2936–2951
119. Spooncer, E., Brouard, N., Nilsson, S. K., Williams, B., Liu, M. C., Unwin, R. D., Blinco, D., Jaworska, E., Simmons, P. J., and Whetton, A. D. (2008) Developmental fate determination and marker discovery in hematopoietic stem cell biology using proteomic fingerprinting. *Mol. Cell. Proteomics* **7**, 573–581
120. Jankowski, R. J., Deasy, B. M., and Huard, J. (2002) Muscle-derived stem cells. *Gene Ther.* **9**, 642–647