Published in final edited form as:

Methods Mol Biol. 2018 ; 1722: 57–78. doi:10.1007/978-1-4939-7553-2_4.

Mass Spectrometry-Based Identification of Extracellular Domains of Cell Surface N-Glycoproteins: Defining the Accessible Surfaceome for Immunophenotyping Stem Cells and Their Derivatives

Chelsea M. Fujinaka, Matthew Waas, and Rebekah L. Gundry

Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI 53226

Abstract

Human stem cells and their progeny are valuable for a variety of research applications and have the potential to revolutionize approaches to regenerative medicine. However, we currently have limited tools to permit live isolation of homogeneous populations of cells apt for mechanistic studies or cellular therapies. While these challenges can be overcome through the use of immunophenotyping based on accessible cell surface markers, the success of this process depends on the availability of reliable antibodies and well-characterized markers, which are lacking for most stem cell lineages. This chapter outlines an iterative process for the development of new cell surface marker barcodes for identifying and selecting stem cell derived progeny of specific cell types, subtypes, and maturation stages, where antibody-independent identification of cell surface proteins is achieved using a modern chemoproteomic approach to specifically identify N-glycoproteins localized to the cell surface. By taking advantage of a large repository of available cell surfaceome data, proteins that are unlikely to confer cell type specificity can be rapidly eliminated from consideration. Subsequently, targeted quantitation by mass spectrometry can be used to refine candidates of interest, and a bioinformatic visualization tool is key to mapping experimental data to candidate protein sequences for the purpose of epitope selection during the antibody development phase. Overall, the process of developing cell surface barcodes for immunophenotyping is iterative and can include multiple rounds of discovery, refinement, and validation depending on the phenotypic resolution required.

Keywords

N-glycoproteins; Proteomics; Cell surface; Cell surface protein atlas; Immunophenotyping; Barcode; Mass spectrometry; Epitope selection; Extracellular domain

1 Introduction

Human stem cells and their progeny are valuable for a variety of research applications and have the potential to revolutionize therapeutic options for treating intractable diseases.

Specifically, stem cells and their progeny can be used to study molecular dynamics during very early stages of human development that are impossible to study in vivo [1] and to study potentially toxic effects of compounds in preclinical studies [2, 3]. It may also become possible to use stem cell progeny to repair damaged tissue and improve organ function by transplanting cells, tissue constructs, or secreted factors [4, 5]. Furthermore, patient-specific cells can be used to study cellular phenotypes and molecular mechanisms resulting from specific genetic mutations and may promote the advancement toward patient-specific medical therapy (i.e., personalized medicine). Despite promising preliminary results across a broad spectrum of lineages and applications, we currently have limited tools to permit the live isolation of homogeneous cell populations for the majority of stem cells or their progeny destined for mechanistic studies or cellular therapies.

Immunophenotyping is an approach where antibody-based detection of cell surface makers is used to identify, select, and quantify populations of cells with defined functional phenotypes. For example, immunophenotyping based on extracellular epitopes has enabled reproducible, nonmutagenic, and high-throughput live cell isolation and assessments of hematopoietic stem cell lineages [6, 7], but it has been underutilized in most other stem cell types. Once developed, immunophenotyping can be used both for live cell identification in quality control processes to assess cell culture purity and for live cell sorting of desired subpopulations. This strategy has tremendous potential to overcome current limitations and enable the isolation of homogeneous cell-type and maturation-stage specific populations of stem cells and their progeny. However, the success of this process depends on the availability of reliable antibodies and well-characterized markers, which are lacking for the majority of stem cell lineages. The lack of suitable markers can, in part, be addressed through antibody screening although there is limited availability of high quality antibodies to surface accessible epitopes. While approximately 370 antigens are targeted by monoclonal antibodies to cluster of differentiation (CD) molecules [8], this represents only a fraction of all molecules present on the cell surface (i.e., the surfaceome) and CD molecule screening has had limited success for nonhematopoietic stem cells [9–11]. Alternative approaches have been employed to evaluate the cell surface protein landscape based on transcriptomics [12], and physical and affinity enrichment strategies coupled with proteomics [13, 14], although each has limitations. Transcript abundance does not always correlate with protein abundance on the cell surface, transcriptomic data fail to inform subcellular localization, and the precise location of many membrane proteins are yet unannotated. This is further confounded by the fact that membrane proteins can be sequestered within the cell and only cycled to the cell surface in response to developmental events or cell signaling (e.g., aquaporins in kidney) [15]. Historically, direct identification and assessment of cell surface proteins using proteomic approaches has been challenging due to the relatively low abundance, hydrophobicity, and difficulty of purifying plasma membrane proteins away from intracellular membranes by biophysical methods. However, modern chemoproteomic strategies are overcoming these challenges and providing previously unprecedented views of the surfaceome of a variety of mammalian cell types [16–20]. When coupled with contemporary bioinformatics tools for comparing surfaceomes among cell types and mapping details of extracellular domains of surfaceome proteins, this approach can

efficiently identify new markers and facilitate antibody development for the purpose of live cell immunophenotyping.

This chapter details the steps involved in the discovery and refinement phases of an iterative process (Fig. 1) for the development of new cell surface marker barcodes for identifying and selecting stem cell derived progeny of specific cell types, subtypes, and maturation stages, where antibody-independent identification of cell surface proteins is achieved using a modern chemoproteomic approach to specifically identify N-glycoproteins localized to the cell surface. By taking advantage of a large repository of cell surfaceome data already available, proteins that are unlikely to provide any cell type specificity can be rapidly eliminated from consideration. Subsequently, targeted quantitation by mass spectrometry can be used to refine candidates of interest, and a bioinformatic visualization tool is key to mapping experimental data to candidate protein sequences for the purpose of epitope selection for antibody development. Overall, the process of developing cell surface barcodes is iterative and can include multiple rounds of discovery, refinement, and validation depending on the phenotypic resolution required.

In the first step of the process, the Cell Surface Capture (CSC)-Technology is used to identify extracellular epitopes of cell surface N-glycoproteins. Specifically, protein identity, transmembrane orientation and glycosylation site occupancy are simultaneously determined in a CSC-Technology experiment. The approach, originally developed by Wollscheid and colleagues [21], takes advantage of the prediction that >90% of cell surface proteins are glycosylated [22, 23]. Experimentally, extracellular oligosaccharides on live cells are oxidized and biotinylated using membrane-impermeable reagents. Proteins are enzymatically digested and the resulting biotinylated glycopeptides are enriched using streptavidin beads. Through the actions of the enzyme peptide-*N*-glycosidase F (PNGaseF), which cleaves between the innermost GlcNAc and asparagine residues, deglycosylated peptides are released and analyzed by high mass accuracy mass spectrometry. During the data analysis process, cell surface proteins are confidently identified by the signature deamidation (+0.984016 Da) observed on the asparagine residue within the conserved N-glycosylation sequence motif (Asn-x-Ser/Thr/Cys (NxS/T/C; where x is any amino acid except proline). Overall, the CSC-Technology generates a highly specific “snapshot” of the cell surfaceome at a particular time or stage. When performed in a comparative manner, it is possible to use the CSC-Technology to identify proteins that are cycled to the surface upon a specific stimulus or experimental condition. The CSC-Technology is thereby advantageous over predictive approaches or those that rely on more generic membrane protein enrichment strategies.

Extracellular epitopes from hundreds of cell surface N-glycoproteins can be identified in a single CSC-Technology experiment, and when compared with other datasets, the output facilitates the identification of novel and ubiquitous cell surface protein markers. Given the power of this approach, CSC-Technology data from over 80 mouse and human cell types (normal and diseased) have now been deposited into a publicly accessible repository, the Cell Surface Protein Atlas (CSPA; <http://wlab.ethz.ch/cspa/>) [24]. When these data are analyzed comparatively, classification of a particular protein as routinely vs. rarely observed among cell types can be made immediately. Although the presence of a protein among

disparate cell types may not preclude its utility for a specific context in vitro, this type of analysis can rapidly focus a candidate list from hundreds to tens in cases where a cell type specific marker is desired. In cases where proteins unique to a single cell type are rare, it is expected that panels of surface markers will be necessary for adequate barcoding and sorting well-defined cell populations [24, 25], and in this context the CSPA is also valuable for highlighting potential marker combinations that can be further examined in the next step (Fig. 1).

Once proteins of interest are identified, contemporary targeted mass spectrometry-based approaches for quantifying hundreds of targets in a single experiment [26] are critical to further refining newly discovered targets. Although antibody-based approaches such as immunofluorescence imaging and flow cytometry offer the ability to monitor protein abundance levels and can inform localization, they require validated and specific antibodies that are often not available for cell surface proteins. Moreover, determining optimal experimental conditions for each cell type of interest and validating antibody specificity requires significant time and effort. As with all antibody-based methods, even if highly specific antibodies are available, they require the epitope to be accessible, which may not be predictable among cell types, as unexpected conformational changes and post-translational modifications can mask epitope availability. For these reasons, we use targeted mass spectrometry-based approaches for simultaneously quantifying tens to hundreds of proteins of interest to further refine the candidates prior to downstream antibody-based studies. Ultimately, high quality antibodies (e.g., recombinant) that specifically recognize extracellular epitopes on live cells are necessary to advance immunophenotyping and sorting of stem cell subpopulations in a way that is analogous to the isolation of therapeutically relevant hematopoietic stem cell populations [6]. For this purpose, the CSC-Technology is especially advantageous as the experimental output confirms the occupancy of individual N-glycosylation sites of identified proteins and thereby confirms extracellular domains. In the final step of the process outlined in this chapter, data from the CSC-Technology are imported into the visualization tool Protter [27]. This approach can be used to confirm or correct predicted protein orientations and highlight occupied N-glycosites and other known modifications and sequence features that can be avoided during epitope selection. Additional bioinformatics analyses of identified glycoproteins can be accomplished using *Targets-search* as described in Chapter 19 of this book by Yan et al., In the validation phase, which is beyond the scope of this chapter and is also the most time-consuming, antibodies are validated for epitope and cell type specificity, and extensive functional analyses of the cells identified and selected by cell surface markers are performed. When successful, the end result is a barcode of cell surface proteins that can be correlated with specific cellular functions.

Overall, the iterative seven-step workflow described here is efficient for the identification of cell surface N-glycoproteins that can be used for live cell immunophenotyping. The steps described in this chapter include (Subheading 3.1) cell collection; (Subheading 3.2) oxidation and biotinylation; (Subheading 3.3) cell lysis and membrane enrichment; (Subheading 3.4) tryptic digestion; (Subheading 3.5) glycopeptide capture and elution; (Subheading 3.6) desalting and concentration of peptides; and (Subheading 3.7) mass spectrometry with data analysis. In Subheading 3.8, we show how to prioritize epitopes for

marker and antibody development. We have used this strategy to identify markers and to generate new antibodies for a variety of stem cells and their progeny, including: selecting pluripotent stem cell-derived hepatocytes [28] and cardiomyocytes ([29] and *in preparation*); isolating induced pluripotent stem cells reprogrammed from mouse fibroblasts [30]; and identifying proteins in human pluripotent stem cells that can be targeted by small molecules for the purpose of eliminating tumorigenic cells [25]. Although our previous work and this chapter have focused on stem cells, the approach is equally applicable to other mammalian cell types, as demonstrated by others [16, 20, 24, 31–33], and variations of this approach that rely on cysteine or lysine containing peptides have also been described [31].

2 Materials

Prepare all solutions with ultrapure water and analytical grade reagents, including liquid chromatography–mass spectrometry (LC-MS) and high-performance liquid chromatography (HPLC) grade water. Low-retention tubes should be used when possible.

2.1 Cell Collection

1. Phosphate buffered saline (PBS, 1×): 1× PBS pH 7.4. Add 900 mL water to 100 mL 10× PBS, pH 7.4 (Quality Biological, Gaithersburg, MD, USA). Store at 4 °C.
2. Liberase TH dissociation solution: Liberase TH (Roche Diagnostics, Indianapolis, IN, USA) (*see Note*¹).
3. PBS with 0.1% fetal bovine serum (FBS): 1× PBS pH 7.4, 0.1% (v/v) FBS. Add 1 mL FBS (Sigma-Aldrich Chemical Company, St. Louis, MO, USA) to 100 mL 10× PBS and fill to 1 L with water. Store at 4 °C.
4. 15 and 50 mL conical tubes.
5. Low speed centrifuge capable of holding 15 and 50 mL conical tubes.

2.2 Oxidation and Biotinylation of Cell Surface Oligosaccharides

1. Labeling buffer: 1× PBS, 0.1% (v/v) FBS, pH 6.5. Add 500 µL FBS to 50 mL 10× PBS, fill to 500 mL with water, and adjust the pH to 6.5 using 85% (w/v) H₃PO₄. Store at 4 °C (*see Notes*² and ³).
2. NaIO₄ stock solution: 195 mM NaIO₄ in labeling buffer. Weigh 0.01 g NaIO₄ (Sigma-Aldrich) in a microcentrifuge tube, add 240 µL labeling buffer and vortex until dissolved (*see Note*²).
3. Biotin hydrazide, 25 mg aliquots (Biotium, Hayward, CA, USA) (*see Note*⁴).

¹Liberase TH is an example of a trypsin-free cell dissociation solution for detaching adherent cells from plates and dishes. The solution and protocol used varies by cell type. It is important that the dissociation solution does not contain trypsin, as this will cleave extracellular protein domains. Acceptable solutions can include enzyme-free dissociation solutions, EDTA solutions, and/or those that contain collagenases. While it is possible to perform the oxidation and biotinylation on cells adhered to the culture dish, this strategy requires a significantly larger volume of the biotinylation solution and thus is more expensive.

²Prepare immediately before use.

³Typically ~5 drops of H₃PO₄ with a glass Pasteur pipette are necessary to reach the correct pH.

4. Laboratory rocker.

2.3 Cell Lysis and Membrane Enrichment

1. Hypotonic lysis buffer: 10 mM Tris-HCl pH 7.5, 0.5 mM MgCl₂. Combine 5 mL 1 M Tris-HCl pH 7.5 and 2.5 mL 100 mM MgCl₂. Fill to 500 mL and store at 4 °C.
2. GentleMACS M tubes and GentleMACS Dissociator (Miltenyi Biotec, San Diego, CA, USA).
3. Membrane prep buffer: 280 mM sucrose, 50 mM MES pH 6.5, 450 mM NaCl, 10 mM MgCl₂. Combine 47.92 g sucrose (Sigma-Aldrich), 25 mL 1M 2-(*N*-morpholino)-ethanesulfonic acid (MES hydrate)-NaOH pH 6.5 (Boston Bio-Products, Ashland, MA, USA), 45 mL 5 M NaCl (Sigma-Aldrich), 50 mL 100 mM MgCl₂. Fill to 500 mL with water and store at 4 °C.
4. Optima LE-80K ultracentrifuge, SW41 Ti rotor, and 9/16 × 3½ inch Ultra-clear tubes (Beckman Coulter, Brea, CA, USA) or equivalent.
5. Thermomixer R with block for 15 mL conical tubes (Eppendorf, Hauppauge, NY, USA).

2.4 Tryptic Digestion

1. 100 mM NH₄HCO₃ buffer: 100 mM NH₄HCO₃. Weigh 0.39 g NH₄HCO₃ (Sigma-Aldrich) and add water to 50 mL (*see Note* ²).
2. 1% AALS-I stock solution: 1% (w/v) Anionic acid labile surfactant-I (AALS-I) (Protea Biosciences, Inc., Morgantown, WV, USA). Dissolve 5 mg AALS-I in 500 µL LC-MS grade water. Store at 4 °C (*see Note* ⁵).
3. Tris(2-carboxyethyl) phosphine (TCEP) stock solution: 100 mM TCEP. Weigh 0.29 g TCEP (Sigma-Aldrich) and add water to 10 mL. Store 0.5 mL aliquots at −20 °C.
4. Iodoacetamide stock solution: 100 mM iodoacetamide. Weigh 0.184 g iodoacetamide (Sigma-Aldrich) and add water to 10 mL (*see Note* ²).
5. Sequencing grade-modified trypsin (Promega, Madison, WI, USA). Store at −80 °C.
6. Six inch Pasteur glass pipette.
7. Parafilm.

⁴While alternative quantities are available, the compound is difficult to transfer and weigh precisely due to static electricity. Therefore, we use preweighed single-use aliquots.

⁵It is important to use a mass spectrometry compatible surfactant to aid in the solubilization of membrane proteins. Alternative surfactants are available but should be tested for suitability.

2.5 Glycopeptide Capture and Elution

1. 1 mL Mobicol columns with 35 μ m pore, large filters (MoBi-Tec, Goettingen, Germany).
2. Vac-man laboratory vacuum manifold (Promega) and vacuum pump or in house vacuum line.
3. Streptavidin Plus UltraLink Resin (Pierce, Rockford, IL, USA). Store at 4 °C.
4. 20 mL BD Luer-Lok Disposable Syringes without needles (BD Biosciences, Franklin Lakes, NJ, USA).
5. Luer-lock caps (MoBiTec, Goettingen, Germany).
6. One-way luer-lock stopcocks (Promega).
7. Invitrosol stock solution: 0.1 $\times$ Invitrosol in 100 mM NH_4HCO_3 . Combine 500 μ L 5 $\times$ Invitrosol (ThermoFisher Scientific, Waltham, MA, USA) and 24.5 mL 100 mM NH_4HCO_3 . Store stock solution in 500 μ L aliquots at -20 °C.
8. 5 M NaCl solution (ThermoFisher Scientific).
9. 100 mM Na_2CO_3 solution: 100 mM Na_2CO_3 . Add 1 g Na_2CO_3 to 50 mL water and vortex until dissolved.
10. 80% isopropanol: 80% (v/v) 2-propanol. Mix 400 mL 2-propanol and 100 mL water.
11. 10 units/ μ L peptide *N*-deglycosylase F, glycerol-free (PNGase F) (Promega).
12. 0.2% TFA solution: 0.2% (v/v) trifluoroacetic acid (TFA). Combine 2 mL TFA in 1 L LC-MS grade water. Store in the dark at room temperature.
13. End over end Shaker.

2.6 Desalting and Concentration of Peptides

1. Hydrophobic packing material C18 Micro SpinColumns with 5–50 μ g capacity (Harvard Apparatus, Holliston, MA, USA).
2. 0.1% TFA solution: 0.1% (v/v). Combine 1 mL TFA in 1 L LC-MS grade water.
3. Elution solution: 70% (v/v) LC-MS grade acetonitrile with 0.1% formic acid, 30% (v/v) LC-MS grade water with 0.1% formic acid. For 100 mL of solution, combine 70 mL acetonitrile with 0.1% formic acid and 30 mL water with 0.1% formic acid.
4. Savant SpeedVac concentrator and UVS400 vacuum pump (ThermoFisher Scientific).
5. 0.5 and 1.5 mL low-binding microcentrifuge tube (ThermoFisher Scientific).

2.7 Mass Spectrometry Instrumentation and Analysis Software

1. Nanospray liquid chromatography-tandem mass spectrometry (nLC-MS/MS).

2. Reprosil-PUR C18-AQ, 3 μm , 120 $\text{\AA}$ (Dr. Maisch; Ammerbuch, Germany) or PicoChip (New Objective, Woburn, MA).
3. Eksigent nanoLC Ultra 2D system (Eksigent Technologies, Dublin, CA, USA) or comparable.
4. Q Exactive hybrid quadrupole orbitrap (ThermoFisher Scientific) or other high mass accuracy instrumentation.
5. Proteome Discoverer 2.1 (ThermoFisher Scientific) with SequestHT and MS Amanda search algorithms.
6. Percolator [34] and ptmRS [35] are publicly available bioinformatic tools for post-search peptide scoring and post-translational modification localization, respectively. These can be used as nodes within Proteome Discoverer.
7. SPSS Statistics (IBM) or other software package capable of hierarchical clustering.

3 Methods

3.1 Collect Cells and Oxidize Extracellular Oligosaccharides

1. Begin with a culture of approximately 10^8 cells (*see Note* ⁶). If working with adherent cells, wash cells with room temperature PBS to remove as many dead cells as possible. If working with cells in suspension, proceed to **step 3**.
2. Detach cells from the culture plate using a trypsin-free solution. The choice of this solution may vary by cell type, and can include collagenases, enzyme free, and/or EDTA solutions. Scraping may be possible, but this must be first tested with a small aliquot of cells, and cells should be stained with trypan blue to evaluate plasma membrane integrity (*see Notes* ¹ and ⁷).
3. Transfer cells and solution into a 50 mL conical tube. Centrifuge at $500 \times g$ at 4 °C for 3 min to collect cells.
4. Aspirate liquid and gently resuspend cells by first tapping against benchtop to loosen the pellet. Add 10 mL of cold PBS with 0.1% FBS and pipette up/down three times (pipette tip close to wall of tube to break up clumps). Remove a 50 μL aliquot for counting, then fill conical tube to 50 mL with PBS with 0.1% FBS. Centrifuge $500 \times g$ at 4 °C for 3 min to collect cells.
5. Aspirate liquid and gently resuspend cells by first tapping against benchtop to loosen the pellet. Add 20 mL cold labeling buffer. Keep tube on ice.
6. Add 100 μL NaIO_4 stock solution to cells/labeling buffer solution. Place cells on a very slow rocker in the dark for 15 min on ice or at 4 °C.

⁶It is possible to use fewer cells, such as $10\text{--}50 \times 10^6$, but the number of proteins identified will be fewer. We typically identify 300–500 proteins if starting with 10^8 cells, 200–250 proteins if starting with 50×10^6 cells.

⁷Plasma membrane integrity is critical to ensure that the oxidation/biotinylation only occurs on extracellular domains.

7. Add cold labeling buffer up to 50 mL to dilute solution. Centrifuge $500 \times g$ at 4°C for 3 min to collect cells.
8. Aspirate liquid. Gently resuspend cells by first tapping against benchtop to loosen the pellet. Add cold labeling buffer up to 50 mL. Centrifuge at $500 \times g$ at 4°C for 3 min to collect cells. Repeat for a total of two washes to remove NaIO_4 .

3.2 Biotinylate Extracellular Oligosaccharides

1. Aspirate liquid. Gently resuspend cells by first tapping against the benchtop to loosen the pellet. Add 4 mL cold labeling buffer.
2. Add 1 mL cold labeling buffer to one vial of biocytin hydrazide (25 mg). Vortex briefly until resuspended. Add all of this to the cell solution. Use 1 mL of cold labeling buffer to rinse the biocytin hydrazide vial and subsequently add this to the cell suspension. The final concentration will be approximately 10 mM biocytin hydrazide (5–10 mM is optimal).
3. Place tube in ice on a rocker to agitate slowly for 60 min.
4. Add PBS with 0.1% FBS up to 50 mL and invert several times to mix. Centrifuge $500 \times g$ at 4°C for 3 min to collect cells.
5. Aspirate liquid. Gently resuspend cells by first tapping against benchtop to loosen the pellet. Add PBS with 0.1% FBS up to 50 mL. Centrifuge at $500 \times g$ at 4°C for 3 min to collect cells. Repeat for a total of two washes.
6. At this step, either flash-freeze and store at -80°C for short term storage or proceed with next step directly.

3.3 Lyse Cells, Remove Nuclei, and Enrich Membranes

1. Aspirate PBS, and ensure that all the PBS is removed. Resuspend cells in 4 mL hypotonic lysis buffer. Set on ice for 10 min to swell the cells making it easier to lyse them.
2. Using a serological pipette, transfer cell solution to Orange Tube (M tube) for GentleMacs. Homogenize cells twice, using method D01 both times (*see Note 8*).
3. Centrifuge GentleMacs tube at $800 \times g$ at 4°C for 10 min to remove nuclei/debris. Transfer supernatant to a new 15 mL conical tube, and save on ice until **step 5**.
4. Using a serological pipette, add 2 mL hypotonic lysis buffer to the pellet remaining in the GentleMacs tube. Homogenize cell pellet twice, using method D01 both times (*see Note 8*).
5. Centrifuge both the GentleMacs tube and the 15 mL conical tube with the supernatant from **step 3** at $800 \times g$ at 4°C for 10 min to remove remaining

⁸The time and speed for disruption may need to be adjusted based on cell type. The goal is to obtain a white nuclear pellet with a slightly cloudy supernatant.

nuclei/debris. Combine the supernatant from the GentleMacs tube and the supernatant from the 15 mL conical into an ultracentrifuge tube. The total volume should be 6 mL.

6. Add an equal volume (6 mL) of membrane preparation buffer. Balance a second tube to within 0.01 g using membrane preparation buffer. Centrifuge in an ultracentrifuge at $210,000 \times g$ at 4 °C overnight (18 h) in an SW41 Ti rotor.
7. Aspirate liquid. Add 200 µL membrane wash buffer to the pellet. Vortex in thermomixer at 750 rpm at 4 °C for 30 min.
8. Add hypotonic lysis buffer until the ultracentrifuge tube is full and then balance a second tube to within 0.01 g.
9. Centrifuge in ultracentrifuge at $210,000 \times g$ at 4 °C for 30 min. Aspirate liquid.

3.4 Tryptic Digestion

1. Add 300 µL of 100 mM ammonium bicarbonate, 40 µL 1% AALS-I MS-compatible surfactant (to make final ~0.1% AALS-I solution), 25 µL 100 mM TCEP (5 mM final) to the ultra-centrifuge tube. Place tube on thermomixer at 37 °C at 750 rpm for 30 min to reduce proteins.
2. Add 45 µL 100 mM iodoacetamide (10 mM final). Place on thermomixer, covered with Parafilm for 30 min to alkylate proteins in the dark.
3. Add the contents of one vial (20 µg) of trypsin. Check pH and adjust to ~8.5 if necessary. Place on thermomixer overnight (see **Note** ⁹).
4. Lower the sample pH to 2 by adding one drop of 85% w/v phosphoric acid using a glass Pasteur pipette (~5 µL) to inactivate trypsin and precipitate lipids. The solution will become turbid.
5. Centrifuge at $16,000 \times g$ for 10 min to remove any nondigested material. Typically, an “oily” pellet will be observed. Use only supernatant to carry on to Subheading 3.5, **step 4**.

3.5 Glycopeptide Capture and Elution

1. Place an empty Mobicol column on the Vacman, using the stopcock as shown in Fig. 2a. Rinse Mobicol with 5×500 µL acetonitrile to remove polymer contamination, then rinse with 5×500 µL 100 mM ammonium bicarbonate.
2. Add 350 µL streptavidin bead slurry to Mobicol column. Wash beads with 3×500 µL 100 mM ammonium bicarbonate.
3. Remove Mobicol from Vacman. Insert plug into bottom of Mobicol and ensure it is snug.

⁹Depending on the amount of starting material, it may be beneficial to add another aliquot of trypsin if a relatively clear solution is not obtained after the first round of digestion.

4. Transfer peptide mixture to the Mobicol containing streptavidin beads. Do not add any of the pellet if there is one. Replace screw cap and ensure end plug is snug. Wrap entire tube in parafilm—wrap bottom to top to ensure plug remains tight. Incubate 1 h at room temperature with end-over-end rotation.
5. Prewash three Mobicol spin columns by following **step 1**. Add 200 μ L water to Mobicol and store temporarily to avoid drying out the membrane until use.
6. Assemble washing apparatus by attaching Mobicol to syringe via luer lock adaptor as shown in Fig. 2b. Transfer peptide/bead slurry to Mobicol column #1. Each time the beads are transferred to a new Mobicol, rinse the previous Mobicol with 3×300 μ L of the solution being used in the subsequent washing step to ensure all beads are transferred.
7. Wash beads with 10 mL 0.01 $\times$ Invitrosol in 100 mM ammonium bicarbonate.
8. Change Mobicol (*Mobicol #2*). Wash beads with 15 mL of 5 M NaCl, then 15 mL of 100 mM sodium carbonate, then 15 mL of 80% isopropanol.
9. Change Mobicol (*Mobicol #3*). Wash beads with 15 mL of 100 mM ammonium bicarbonate. Remove Mobicol from Vac-man. Insert plug into bottom of Mobicol and ensure it is snug.
10. Add 350 μ L of 100 mM ammonium bicarbonate to beads in Mobicol, followed by the addition of 3 μ L PNGaseF. Replace screw cap and ensure end plug is snug. Wrap entire tube in parafilm, wrapping from bottom to top to ensure plug remains tight. Incubate at 37 °C overnight with end-over-end rotation.
11. Unplug Mobicol and place into clean 1.5 mL microfuge tube. Centrifuge at 1000 $\times g$ for 30 s to collect liquid which contains cleaved peptides.
12. Add 350 μ L 0.2% TFA to beads and vortex gently to resuspend beads. Place Mobicol into clean 1.5 mL microfuge tube. Centrifuge at 1000 $\times g$ for 30 s to collect liquid.
13. Combine liquid from **steps 11** and **12** into a single 1.5 mL microfuge tube.

3.6 Desalting and Concentrating

1. Cut off approximately 25 mm of a P200 pipette tip and use as an adaptor to place the MicroSpin column on the vacuum manifold (Fig. 2c).
2. Rinse column with 5×200 μ L 100% HPLC grade acetonitrile then 10×200 μ L HPLC grade 0.1% TFA.
3. Place column into a clean 1.5 mL low-binding microcentrifuge tube.
4. Add 200 μ L of peptide sample to the top of the column. Do not add any pellet if there is one. Using a P200 pipette, slowly “push” sample through the column, from top to bottom. Repeat until all of the sample is through the column twice.
5. Place column on vacuum manifold. Rinse column with 10×200 μ L HPLC grade 0.1% TFA, then 10×200 μ L HPLC grade water.

6. Place column into a clean 0.5 mL low-binding microcentrifuge tube. Add 100 μ L of elution solution. Use a P200 pipette, slowly “push” sample through the column, from top to bottom. Centrifuge tube/column at $1000 \times g$ for 30 s to collect remaining eluent.
7. Place the sample in a SpeedVac or similar instrument until all of the liquid has evaporated.
8. Resuspend peptides in 12 μ L LC-MS grade 0.1% formic acid in water.

3.7 Mass Spectrometry and Data Analysis

1. Analyze peptides by nanospray liquid chromatography tandem mass spectrometry (LC-MS/MS). We typically use a 3 cm trap column followed by a 10 cm analytical column, both 75 μ m inner diameter packed with Reprosil-PUR C18-AQ, 3 μ m, 120 Å, either packed in house or in a prepacked PicoChip (New Objective, Woburn, MA). We typically perform two technical replicate injections of 5 μ L peptides using a gradient of mobile phase A (water, 0.1% formic acid) and mobile phase B (acetonitrile, 0.1% formic acid) of 3–20% B over 55 min, 20–30% B over 15 min, 30–50% B over 4 min, 50–98% B over 4 min at a flow rate of 300 nL/min (Eksigent nanoLC Ultra 2D system). Comparable chromatographic systems and settings will work equally well.
2. Data are best acquired using a high mass accuracy mass spectrometer with low limits of detection. In our case, this is a Q Exactive hybrid quadrupole orbitrap. The instrument is operated in data dependent mode to automatically switch between full scan MS and MS/MS acquisition. Full scan MS spectra are acquired in the Orbitrap with 70,000 resolution, accumulation of ions to a 1×10^6 target value based on predictive AGC, and dynamic exclusion is set to 30 s. The 15 most intense multiply charged ions ($2 \leq z \leq 8$) are sequentially isolated and fragmented in the octopole collision cell by higher energy collisional dissociation with a target value of 1×10^4 ions, maximum injection time of 120 ms and 35,000 resolution. Typical conditions are as follows: spray voltage of 2 kV, no sheath and auxiliary gas flow, heated capillary temperature 275 °C, normalized collision energy 25%, underfill ratio set to 0.1%, 2.0 m/z isolation window. Comparable instrumentation that offers high mass accuracy and low limits of detection will work equally well.
3. Raw data files from the MS instrument are searched via ProteomeDiscoverer using SequestHT and MS Amanda [36] search algorithms, Percolator [34] for post-search validation and ptmRS [35] for added confidence in localizing the site of deamidation (indicating the occupied N-glycosite), which can be especially helpful for peptides containing two asparagine residues (see **Note 10**). The search is performed against the SwissProt version of the appropriate organism database

¹⁰If more than 20% of the peptides identified are contaminants (i.e., do not contain a deamidation within the sequence motif), this is typically due to insufficient washing of the streptavidin beads after the glycopeptide capture. To further reduce nonspecific binding, streptavidin beads may be transferred to a new Mobicol repeatedly throughout the washing steps and volumes of washing solutions may be increased.

from UniProt (<http://www.uniprot.org/>). Search settings include the following: semitryptic digestion, MS1 tolerance of 10 ppm, MS2 tolerance of 0.02 Da, static modification of 57.021464 Da on cysteines representing carbamidomethylation from iodoacetamide treatment, dynamic modifications of 15.995 Da on methionine representing oxidation, and 0.984016 Da on asparagines representing deamidation resulting from cleavage of the N-linked oligosaccharide from the peptide backbone via PNGaseF treatment.

4. If the experiment is successful, 200–400 cell surface N-glycoproteins will be identified. Overall, more than 85% of the peptides identified should contain a deamidation at the asparagine within the sequence motif NxS/T/C (Fig. 3a) (*see Note*¹¹).
5. Export high confidence identifications from ProteomeDiscoverer or ProteinProphet as a .protXML file, which is the file format used in downstream analyses.

3.8 Prioritization of Epitopes for Marker and Antibody Development

1. To begin the selection of candidate makers that may be informative for a particular cell type, we utilize ProteinCenter to archive all data generated by the CSC-Technology. This work-flow enables comparisons of cell surface N-glycoproteins observed among various experimental conditions (e.g., cell types, organ types, and developmental stages). We maintain an in-house repository of data from three sources: data generated in-house, data from the Cell Surface Protein Atlas (Atlas [24]; downloaded from <http://wlab.ethz.ch/cspa/>), and data from other publications that may not be contained in the Atlas. While proteins of interest for distinguishing cellular phenotypes can be identified by directly including the desired cell types for comparison within a quantitative CSC-Technology experiment, the Atlas may still add value as a freely available repository populated with data from a variety of cell types.
2. To establish a custom database of cell surface N-glycoproteins, data from the Atlas and other publications are imported into ProteinCenter via .csv file format. Data generated in-house are imported as prot.xml files.
3. Once all data are imported into ProteinCenter, select the desired datasets and create a comparison. At this stage, contaminating proteins identified in the experiment can be eliminated by data filtering to keep only those proteins identified by peptides containing a deamidated asparagine within the conserved sequence motif (NxS/T/C) for N-glycosylation. To eliminate redundancy and enable comparisons among datasets that may have been generated using different protein databases, cluster data within the comparison by “indistinguishable proteins,” which will group proteins that cannot be distinguished based on the peptide evidence they share. This comparison provides a rapid strategy to

¹¹Notably, the Trans-Proteomic Pipeline, using Comet [38] and X!Tandem [39] algorithms followed by PeptideProphet [40] and ProteinProphet [41] also works well.

eliminate proteins widely observed across many cell types, and can highlight proteins observed in only a few cell types (Fig. 3b).

4. In addition to the binary comparisons described above, hierarchical clustering of the Atlas data can be used to identify combinations of proteins that may be informative for a particular cell type. An example is shown in Fig. 3c, where hierarchical clustering was performed using SPSS Statistics V22 using binary representation (1 = detected, 0 = not detected) of 1323 proteins from 15 human cell types. The furthest neighbor (i.e., complete linkage) cluster method was used in conjunction with the Dice similarity measure for determination of clusters. The Dice measure was used because of its exclusion of joint absences from consideration and the double weight it applies to matches. Together, these features are appropriate to address the limitations of working with discovery-based nonquantitative proteomics data (i.e., “not observed” does not confirm absence) and this strategy can reveal proteins potentially informative for distinguishing among protein groups. For example, the clustering result in Fig. 3c shows pluripotent stem cells cluster closely to their derivatives and to fibroblasts (the precursor from which the induced pluripotent stem cells were derived). Of the proteins of interest revealed by this strategy, we have previously used flow cytometry to confirm 12 as either positive or negative markers of pluripotency when compared to fibroblasts [25].
5. Proteins of interest identified in the data dependent discovery process described above can be quantified using targeted mass spectrometry approaches (e.g., multiple reaction monitoring, parallel reaction monitoring) to further refine the relative or absolute quantity among cell types, subtypes, and/or developmental stages in an antibody independent manner. This step can be especially important if it is necessary to analyze cell types that cannot be obtained in sufficient quantity to permit discovery-based analyses (e.g., <1E6 cells). An example of how parallel reaction monitoring was used to quantify integrin beta 1 and glucose transporter 3 simultaneously across three time points of human pluripotent stem cell-derived cardiomyocyte differentiation is shown in Fig. 4a. Similar approaches can be performed to quantify tens to hundreds of targets within a single experiment and can be performed using whole cell lysate, membrane enriched fractions, or N-glycopeptides captured via the CSC-Technology (see **Note** ¹²).
6. When selecting epitopes to be used for generating antibodies for the purpose of live cell analyses, it is important to select epitopes from the extracellular domain so that the antibodies can access the epitopes without permeabilization of the plasma membrane. In this case, the CSC-Technology data are valuable as the N-glycopeptides identified in the approach are derived from the extracellular

¹²The choice for sample type used for quantitation can depend on several factors. Whole cell lysate or biophysically enriched membrane fractions can be advantageous as the sample preparation is simpler and more peptides from across the protein can be used for quantitation (i.e., not only the formerly glycosylated N-glycopeptide). However, this type of preparation does not directly provide evidence of surface localization. The CSC-Technology can be used and will provide evidence of localization, but the peptides available for quantitation are restricted to those formerly glycosylated N-glycopeptides.

domain and these data can be used to verify the transmembrane orientation of the protein. Moreover, these data can reveal occupied N-glycosites that can be avoided during epitope selection. To visualize the N-glycopeptides identified in the experiment along with annotated protein features, upload the prot.xml file to Protter (<http://wlab.ethz.ch/protter/start/>) by choosing the option to “load a proteomics result file.” All peptides identified in the experiment will be mapped to their respective proteins and can be rapidly viewed in the context of known modifications and other sequence features (Fig. 4b).

Acknowledgments

This work was supported by National Institutes of Health grants R01HL126785 and R01HL134010 and the Paul G. Allen Family Foundation (Grant Award 11715).

References

1. Mummery CL, Zhang J, Ng ES, Elliott DA, Elefanty AG, Kamp TJ. Differentiation of human embryonic stem cells and induced pluripotent stem cells to cardiomyocytes: a methods overview. *Circ Res.* 2012; 111(3):344–358. [PubMed: 22821908]
2. Braam SR, Tertoolen L, van de Stolpe A, Meyer T, Passier R, Mummery CL. Prediction of drug-induced cardiotoxicity using human embryonic stem cell-derived cardiomyocytes. *Stem Cell Res.* 2010; 4(2):107–116. [PubMed: 20034863]
3. Liang P, Lan F, Lee AS, Gong T, Sanchez-Freire V, Wang Y, et al. Drug screening using a library of human induced pluripotent stem cell-derived cardiomyocytes reveals disease-specific patterns of cardiotoxicity. *Circulation.* 2013; 127(16):1677–1691. [PubMed: 23519760]
4. Chong JJ, Yang X, Don CW, Minami E, Liu YW, Weyers JJ, et al. Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. *Nature.* 2014; 510(7504):273–277. [PubMed: 24776797]
5. Schwartz SD, Hubschman JP, Heilwell G, Franco-Cardenas V, Pan CK, Ostrick RM, et al. Embryonic stem cell trials for macular degeneration: a preliminary report. *Lancet.* 2012; 379(9817):713–720. [PubMed: 22281388]
6. Bryder D, Rossi DJ, Weissman IL. Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. *Am J Pathol.* 2006; 169(2):338–346. [PubMed: 16877336]
7. Shizuru JA, Negrin RS, Weissman IL. Hematopoietic stem and progenitor cells: clinical and preclinical regeneration of the hematolymphoid system. *Annu Rev Med.* 2005; 56:509–538. [PubMed: 15660525]
8. Clark G, Stockinger H, Balderas R, van Zelm MC, Zola H, Hart D, Engel P. Nomenclature of CD molecules from the tenth human leucocyte differentiation antigen workshop. *Clin Transl Immunol.* 2016; 5(1):e57.
9. Turac G, Hindley CJ, Thomas R, Davis JA, Deleidi M, Gasser T, Karaoz E, Pruszk J. Combined flow cytometric analysis of surface and intracellular antigens reveals surface molecule markers of human neurogenesis. *PLoS One.* 2013; 8(6):e68519. [PubMed: 23826393]
10. Dubois NC, Craft AM, Sharma P, Elliott DA, Stanley EG, Elefanty AG, Gramolini A, Keller G. SIRPA is a specific cell-surface marker for isolating cardiomyocytes derived from human pluripotent stem cells. *Nat Biotechnol.* 2011; 29(11):1011–1018. [PubMed: 22020386]
11. Kelly OG, Chan MY, Martinson LA, Kadoya K, Ostertag TM, Ross KG, et al. Cell-surface markers for the isolation of pancreatic cell types derived from human embryonic stem cells. *Nat Biotechnol.* 2011; 29(8):750–756. [PubMed: 21804561]
12. Bye CR, Jonsson ME, Bjorklund A, Parish CL, Thompson LH. Transcriptome analysis reveals transmembrane targets on transplantable midbrain dopamine progenitors. *Proc Natl Acad Sci U S A.* 2015; 112(15):E1946–E1955. [PubMed: 25775569]

13. Cordwell SJ, Thingholm TE. Technologies for plasma membrane proteomics. *Proteomics*. 2010; 10(4):611–627. [PubMed: 19834916]
14. Elschenbroich S, Kim Y, Medin JA, Kislinger T. Isolation of cell surface proteins for mass spectrometry-based proteomics. *Expert Rev Proteomics*. 2010; 7(1):141–154. [PubMed: 20121483]
15. Kitchen P, Day RE, Salman MM, Conner MT, Bill RM, Conner AC. Beyond water homeostasis: diverse functional roles of mammalian aquaporins. *Biochim Biophys Acta*. 2015; 1850(12):2410–2421. [PubMed: 26365508]
16. Danzer C, Eckhardt K, Schmidt A, Fankhauser N, Ribrioux S, Wollscheid B, et al. Comprehensive description of the N-glycoproteome of mouse pancreatic beta-cells and human islets. *J Proteome Res*. 2012; 11(3):1598–1608. [PubMed: 22148984]
17. Tyleckova J, Valekova I, Zizkova M, Rakocyova M, Marsala S, Marsala M, Gadher SJ, Kovarova H. Surface N-glycoproteome patterns reveal key proteins of neuronal differentiation. *J Proteome*. 2016; 132:13–20.
18. Sun B, Ma L, Yan X, Lee D, Alexander V, Hohmann LJ, et al. N-glycoproteome of E14.Tg2a mouse embryonic stem cells. *PLoS One*. 2013; 8(2):e55722. [PubMed: 23405203]
19. Hofmann A, Thiesler T, Gerrits B, Behnke S, Sobotzki N, Omasits U, et al. Surfaceome of classical Hodgkin and non-Hodgkin lymphoma. *Proteomics Clin Appl*. 2015; 9(7–8):661–670. [PubMed: 26076441]
20. Ducret A, Kux van Geijtenbeek S, Roder D, Simon S, Chin D, Berrera M, et al. Identification of six cell surface proteins for specific liver targeting. *Proteomics Clin Appl*. 2015; 9(7–8):651–661. [PubMed: 26097162]
21. Wollscheid B, Bausch-Fluck D, Henderson C, O'Brien R, Bibel M, Schiess R, Aebersold R, Watts JD. Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat Biotechnol*. 2009; 27(4):378–386. [PubMed: 19349973]
22. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta*. 1999; 1473(1):4–8. [PubMed: 10580125]
23. Gahmberg CG, Tolvanen M. Why mammalian cell surface proteins are glycoproteins. *Trends Biochem Sci*. 1996; 21(8):308–311. [PubMed: 8772385]
24. Bausch-Fluck D, Hofmann A, Bock T, Frei AP, Cerciello F, Jacobs A, et al. A mass spectrometric-derived cell surface protein atlas. *PLoS One*. 2015; 10(3):e0121314. [PubMed: 25894527]
25. Boheler KR, Bhattacharya S, Kropp EM, Chuppa S, Riordon DR, Bausch-Fluck D, et al. A human pluripotent stem cell surface N-glycoproteome resource reveals markers, extracellular epitopes, and drug targets. *Stem Cell Rep*. 2014; 3(1):185–203.
26. Domon B, Gallien S. Recent advances in targeted proteomics for clinical applications. *Proteomics Clin Appl*. 2015; 9(3–4):423–431. [PubMed: 25504492]
27. Omasits U, Ahrens CH, Muller S, Wollscheid B. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics*. 2014; 30(6):884–886. [PubMed: 24162465]
28. Mallanna SK, Cayo MA, Twaroski K, Gundry RL, Duncan SA. Mapping the cell-surface N-glycoproteome of human hepatocytes reveals markers for selecting a homogeneous population of iPSC-derived hepatocytes. *Stem Cell Rep*. 2016; 7(3):543–556.
29. Kropp EM, Bhattacharya S, Waas M, Chuppa SL, Hadjantonakis AK, Boheler KR, Gundry RL. N-glycoprotein surfaceomes of four developmentally distinct mouse cell types. *Proteomics Clin Appl*. 2014; 8(7–8):603–609. [PubMed: 24920426]
30. Gundry RL, Riordon DR, Tarasova Y, Chuppa S, Bhattacharya S, Juhasz O, et al. A cell surfaceome map for immunophenotyping and sorting pluripotent stem cells. *Mol Cell Proteomics*. 2012; 11(8):303–316. [PubMed: 22493178]
31. Bausch-Fluck D, Hofmann A, Wollscheid B. Cell surface capturing technologies for the surfaceome discovery of hepatocytes. *Methods Mol Biol*. 2012; 909:1–16. [PubMed: 22903705]
32. Hofmann A, Gerrits B, Schmidt A, Bock T, Bausch-Fluck D, Aebersold R, Wollscheid B. Proteomic cell surface phenotyping of differentiating acute myeloid leukemia cells. *Blood*. 2010; 116(13):e26–e34. [PubMed: 20570859]

33. Schiess R, Mueller LN, Schmidt A, Mueller M, Wollscheid B, Aebersold R. Analysis of cell surface proteome changes via label-free, quantitative mass spectrometry. *Mol Cell Proteomics*. 2009; 8(4):624–638. [PubMed: 19036722]
34. Kall L, Canterbury JD, Weston J, Noble WS, MacCoss MJ. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat Methods*. 2007; 4(11):923–925. [PubMed: 17952086]
35. Taus T, Kocher T, Pichler P, Paschke C, Schmidt A, Henrich C, Mechtler K. Universal and confident phosphorylation site localization using phosphoRS. *J Proteome Res*. 2011; 10(12):5354–5362. [PubMed: 22073976]
36. Dorfer V, Pichler P, Stranzl T, Stadlmann J, Taus T, Winkler S, Mechtler K. MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. *J Proteome Res*. 2014; 13(8):3679–3684. [PubMed: 24909410]
37. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*. 2010; 26(7):966–968. [PubMed: 20147306]
38. Eng JK, Jahan TA, Hoopmann MR. Comet: an open-source MS/MS sequence database search tool. *Proteomics*. 2013; 13(1):22–24. [PubMed: 23148064]
39. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. 2004; 20(9):1466–1467. [PubMed: 14976030]
40. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem*. 2002; 74(20):5383–5392. [PubMed: 12403597]
41. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem*. 2003; 75(17):4646–4658. [PubMed: 14632076]

Fig. 1.

Iterative process for identifying extracellular epitopes and developing cell surface marker barcodes for live immunophenotyping. In the discovery phase, combining the CSC-Technology with the Cell Surface Protein Atlas provides a rapid method for identifying proteins that may be informative for identifying or selecting a specific cell type. Once putative candidates are selected, the refinement phase includes antibody-independent quantitation by targeted mass spectrometry to further refine those markers that are potentially most informative. Subsequently, Protter is used to visualize extracellular domains and guide epitope selection for antibody development. In the validation phase, monoclonal antibodies are validated and functional and molecular measurements can be performed on the cells identified by the antibodies in order to correlate the functional phenotype with the cell surface phenotype

Fig. 2. Vacuum manifold setup. Shown are the configurations for using the Mobicol (A, B) for glycopeptide capture and washing and the MicroSpin column (C) used for peptide desalting and concentration 3.6 Desalting and Concentrating

Fig. 3.

Exploiting the Cell Surface Protein Atlas to identify proteins of interest. (a) Annotated MS/MS spectrum of a representative peptide from Integrin Beta 1 that was identified in human pluripotent stem cell derived cardiomyocytes via the CSC-Technology. One identified site of N-glycosylation (N₁₁₅) within the sequence motif is indicated. Spectrum generated using Proteome Discoverer 2.1 (ThermoFisher Scientific). Inset table shows m/z values for fragment ions predicted to result from the proposed peptide sequence NPCTSEQNCTSPFSYK. Values in red or blue were observed in the MS/MS spectrum. (b)

Screenshot of ProteinCenter repository of CSC-Technology data for 59 human cell types. Color denotes that a protein (one protein/row) was identified in a particular cell type (one cell type/column). Data are shown for 20 cell surface N-glycoproteins, including those ubiquitously observed across most cell types, and those that are observed less commonly among those cell types represented in the Atlas. (c) Hierarchical clustering of 15 human cell types using 1323 cell surface N-glycoproteins identified via CSC-Technology can successfully cluster similar cell types

Fig. 4. CSC-Technology data can inform subsequent quantitative studies and antibody development. (a) Quantitative data for two cell surface proteins, Integrin Beta 1 and Glucose Transporter 3 across three time points of human pluripotent stem cell derived cardiomyocyte differentiation. Quantitation was achieved using parallel reaction monitoring to target both proteins in the same experiment and data were analyzed using Skyline [37]. Top row shows fragment ion chromatograms used for quantitation, and integrated data are summarized in the chart below. (b) Illustrations of the transmembrane topology of Glypican-1 and Integrin Beta 1 highlighting the extracellular peptides identified by CSC-Technology, including sites of N-glycosylation as well as additional annotated sequence features that may be considered during epitope selection for antibody development. Image generated using Protter