CD133 Protein N-Glycosylation Processing Contributes to Cell Surface Recognition of the Primitive Cell Marker AC133 Epitope

Anthony B. Mak, Kim M. Blakely, Rashida A. Williams, Pier-Andrée Penttilä, Khan T. Osman, Dahlia Kasimer, Troy Ketela, and Jason Moffat

From the Donnelly Centre and Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario M5S 3E1, the Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, the Sickkids-University Health Network Flow Cytometry Facility, Toronto, Ontario M5G 1L7, and the Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario M5T 1P7, Canada

Background: Cell surface recognition of the AC133 epitope on CD133 marks many stem cell and cancer stem cell types. A large scale RNA interference screen identifies genes involved in N-glycosylation that regulate cell surface AC133 recognition.

Results: A large scale RNA interference screen identifies genes involved in N-glycosylation that regulate cell surface AC133 recognition.

Conclusion: CD133 N-glycosylation and its processing contribute to cell surface AC133 recognition.

Significance: Glycobiological differences between primitive and differentiated cells may be responsible for regulating cell surface AC133.

The AC133 epitope expressed on the CD133 glycoprotein has been widely used as a cell surface marker of numerous stem cell and cancer stem cell types. It has been recently proposed that posttranslational modification and regulation of CD133 may govern cell surface AC133 recognition. Therefore, we performed a large scale pooled RNA interference (RNAi) screen to identify genes involved in cell surface AC133 expression. Gene hits could be validated at a rate of 70.5% in a secondary assay using an orthogonal RNAi system, demonstrating that our primary RNAi screen served as a powerful genetic screening approach. Within the list of hits from the primary screen, genes involved in N-glycan biosynthesis were significantly enriched as determined by Ingenuity Canonical Pathway analyses. Indeed, inhibiting biosynthesis of the N-glycan precursor using the small molecule tunicamycin or inhibiting its transfer to CD133 by generating N-glycan-deficient CD133 mutants resulted in undetectable cell surface AC133. Among the screen hits involved in N-glycosylation were genes involved in complex N-glycan processing, including the poorly characterized Mgat4c, which we demonstrate to be a positive regulator of cell surface AC133 expression. Our study identifies a set of genes involved in CD133 N-glycosylation as a direct contributing factor to cell surface AC133 recognition and provides biochemical evidence for the function and structure of CD133 N-glycans.

The AC133 epitope was originally identified as a marker of CD34 human hematopoietic progenitors (1). Its further characterization found it to be specific to CD133 (prominin-1), a 120-kDa pentaspan transmembrane glycoprotein (2) localized primarily to plasma membrane protrusions (3, 4). Since its discovery, AC133 has been utilized as a marker of primitive cells, including both normal embryonic and adult stem cells (5). AC133 has also been used extensively to mark cancer stem cells (CSCs) from numerous tissues types, including brain (6), colon (7, 8), pancreas (9), liver (10), skin (11), prostate (12, 13), and ovarian (14). Furthermore, AC133 has marked CSCs that displayed resistance to cancer treatments, including chemotherapeutic drugs and radiation, suggesting that AC133-marked CSCs may be responsible for cancer relapse (15). Despite the widespread use of the AC133 marker to isolate, enrich, and characterize stem cells and CSCs, there is little known regarding the regulation of its expression. Caution has been reported on the utility of AC133 as a primitive cell surface marker because CD133 transcripts and protein have been detected in a variety of tissues and cell lines (3, 16, 17) that do not possess stem cell-like properties. Recently, Kemper et al. (18) demonstrated that the cell surface detection of AC133 is indeed specific to primitive cells and is lost during differentiation, whereas CD133 protein lacking the AC133 epitope can be maintained in differentiated cells. Given the above, this led to the hypothesis that posttranslational modifications and regulation of CD133 are responsible for the cell surface recognition of the AC133 epitope. Because the factors and biological processes involved in regulating cell surface AC133 recognition

3 The abbreviations used are: CSC, cancer stem cell; Dol-P, dolichol phosphate; ER, endoplasmic reticulum; esiRNA, endoribonuclease-prepared siRNA; KEGG, Kyoto Encyclopedia of Genes and Genomes; Endo Hf, endoglycosidase Hf; PNGase F, peptide N-glycosidase F; MAPLE, mammalian affinity purification lentiviral expression; PE, phycoerythrin; VA, versatile affinity tag.

4 The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables 1–3 and Figs. 1–3.

5 Supported by doctoral Canadian Graduate scholarships from the Natural Sciences and Engineering Research Council.

6 The AC133 epitope was originally identified as a marker of CD34 human hematopoietic progenitors (1). Its further characterization found it to be specific to CD133 (prominin-1), a 120-kDa pentaspan transmembrane glycoprotein (2) localized primarily to plasma membrane protrusions (3, 4). Since its discovery, AC133 has been utilized as a marker of primitive cells, including both normal embryonic and adult stem cells (5). AC133 has also been used extensively to mark cancer stem cells (CSCs) from numerous tissues types, including brain (6), colon (7, 8), pancreas (9), liver (10), skin (11), prostate (12, 13), and ovarian (14). Furthermore, AC133 has marked CSCs that displayed resistance to cancer treatments, including chemotherapeutic drugs and radiation, suggesting that AC133-marked CSCs may be responsible for cancer relapse (15). Despite the widespread use of the AC133 marker to isolate, enrich, and characterize stem cells and CSCs, there is little known regarding the regulation of its expression. Caution has been reported on the utility of AC133 as a primitive cell surface marker because CD133 transcripts and protein have been detected in a variety of tissues and cell lines (3, 16, 17) that do not possess stem cell-like properties. Recently, Kemper et al. (18) demonstrated that the cell surface detection of AC133 is indeed specific to primitive cells and is lost during differentiation, whereas CD133 protein lacking the AC133 epitope can be maintained in differentiated cells. Given the above, this led to the hypothesis that posttranslational modifications and regulation of CD133 are responsible for the cell surface recognition of the AC133 epitope. Because the factors and biological processes involved in regulating cell surface AC133 recognition

7 The abbreviations used are: CSC, cancer stem cell; Dol-P, dolichol phosphate; ER, endoplasmic reticulum; esiRNA, endoribonuclease-prepared siRNA; KEGG, Kyoto Encyclopedia of Genes and Genomes; Endo Hf, endoglycosidase Hf; PNGase F, peptide N-glycosidase F; MAPLE, mammalian affinity purification lentiviral expression; PE, phycoerythrin; VA, versatile affinity tag.
remain unclear, we took an unbiased genetic screening approach to identify genes that regulate cell surface AC133.

We present a large scale RNA interference (RNAi) screening approach targeting 11,248 genes to identify genes involved in cell surface AC133 expression. Genes that were identified as hits from our screen were well represented in the Kyoto Encyclopedia of Genes and Genomes (KEGG) curated list of N-glycan biosynthesis genes. We demonstrate that N-glycans are required for proper CD133 localization and cell surface AC133 expression. We also implicate the poorly characterized MGAT4C, a predicted glycosyltransferase, in CD133 N-glycosylation. Our data suggest that CD133 acts as an MGAT4C substrate for complex N-glycan processing, which contributes to cell surface AC133 recognition. Taken together, our study provides experimental evidence for the hypothesis that post-translational modification and regulation of CD133 is responsible for cell surface recognition of AC133.

**EXPERIMENTAL PROCEDURES**

**Cell Culture and Small Molecules**—Human embryonic kidney (HEK) 293 cells were cultured as described previously (19). SwaInsosine (S885000; Toronto Research Chemicals) and tunicamycin (T7765; Sigma-Aldrich) were suspended in 95% ethanol and were used to treat cells at 0.1% of the total volume for 24 h. Ethanol-only treatments were used as controls. CD133-VA was FLAG-immunoprecipitated as described previously (19) for PNGase F (P0704; New England Biolabs) and Endo Hf (P0703; New England Biolabs) assays, which were performed according to the manufacturer’s protocol.

**Lentivirus Production**—Lentivirus production and infection were performed as described previously (20).

**Western Blotting**—Cells were lysed, run on SDS-polyacrylamide gels, and transferred to PVDF membranes as described previously (19). Immunoblotting was performed as described previously (19) with the addition of AC133/1 pure (130-090-422; Miltenyi Biotec), MGAT4C (17841-1-P; Proteintech Group), and hypoxia-inducible factor 1α (3716, Cell Signaling Technology) antibodies.

**Site-directed Mutagenesis**—CD133 N-glycan-deficient mutants were generated using the QuikChange Lightning Multi Site-directed Mutagenesis kit (210518; Stratagene) according to the manufacturer’s protocol with an extension time of 13 min using the template pDONR223-CD133 (isoform 2) (described below) and appropriate primers for site-directed mutagenesis as listed in supplemental Table 3. All mutants were sequence-verified.

**Constructs**—The CD133 entry clone was generated by PCR amplification of CD133 (isoform 2) from MGC:20041 (IMAGE: 4644690) (Centre of Applied Genomics, Toronto, ON) with attB1-CD133 primer 5’-GGG GAC AAG TTT GTA CAA AAA AGC AGG CTT CAT GGC CCT GTG ACT CGG CTC C-3’ and attB2-CD133 primer 5’-GGG GAC CAC TTT GTA CAA GAA AGT GGG TCA TGT TGTG ATG GCC TGG TCA TAA CAG-3’ and introduced into pDONR-223 vector by a BP Clonase reaction (Invitrogen) according to the manufacturer’s protocol. CD133 was introduced into the mammalian affinity purification lentiviral expression (MAPLE) vector pLD-puro-CcVA as described previously (19) to generate a lentiviral expression vector. Individual CD133 N-glycosylation site mutants were introduced into pLD-puro-EcYFP using Gateway® cloning as described previously (19). pLD-puro-EcYFP was constructed by introducing a PCR-amplified YFP (sense, 5’-GAT CGA TCT CTA GA GTG AGC AAG GGC GAG GA-3’ and antisense, 5’-GAT CGA TCT TCG AAT TAC GAA GCT TGA GCT CGA GA-3’) into pLD-puro-EcF (19) using XbaI and BstBI restriction sites. All constructs were sequence-verified.

**Pooled Lentiviral shRNA Screen**—Prior to the screen, the highest 18.5% of HEK293/CD133-VA cells were sorted (as described below) to generate the cell population referred to as HEK293/AC133 and used solely for the screen. A 54K pooled shRNA library was obtained from Cheung et al. (21) (supplemental Table 1) (23), and 1 × 10^7 HEK293/AC133 cells were infected with the 54K pooled lentiviral shRNA library at a multiplicity of infection of 1, resulting in a 150-fold representation per individual shRNA (n = 3). More than 1 × 10^7 HEK293/AC133 cells treated with the lentiviral shRNA pooled library as described above were sorted into AC133low (lowest 2% of the population for AC133-PE staining) and AC133high fractions using a custom FACSArria (Becton Dickinson) 7 days after infection, and this was done for each of the three replicates.

**shRNA Probe Preparation and Microarray Analysis**—Extraction of genomic DNA and shRNA probe preparation for hybridization to microarray chips has been described previously (22, 23). Hits from our genetic screen were determined by first eliminating shRNAs that appeared above background (>274 intensity) in the AC133high population, in any of the replicates. Second, only shRNAs that appeared in the AC133low population with an intensity ≥1450 in each of the three replicates were considered hits. This resulted in a list of shRNAs that were ≥5-fold higher in the mean AC133low population than in the mean AC133high population (supplemental Table 1).

**Cell Sorting and Flow Cytometry**—AC133 cell surface expression of HEK293/CD133-VA cells was determined by staining cells with CD133/1(AC133)-PE (130-080-801; Miltenyi Biotec) or CD133/1(AC133)-APC (130-090-826; Miltenyi Biotec) antibodies according to the manufacturer’s protocol and analyzed using a FACSArria (Becton Dickinson). Cell sorting for the screen was performed by staining HEK293/CD133-VA cells infected with the pooled lentivirus-based shRNA library with the CD133/1(AC133)-PE according to the manufacturer’s protocol and counterstaining with propidium iodide at a final concentration of 1 μg/ml. Cells stained positive for propidium iodide were gated out. Cells were sorted using a FACSArria equipped with a 100-mW Coherent Sapphire 488-nm laser and a 20-mW JDS Uniphase 633-nm laser. Flow (version 7.6) was used to analyze data collected from the FACSArria. Cells sorted or analyzed by FACS were briefly treated with trypsin for dissociation.

**Endoribonuclease-prepared siRNA (esiRNA Interference)**—esiRNAs were produced as described previously (24) with primers containing T7 sequences (supplemental Table 2). An esiRNA against firefly luciferase (esiLuc) was used as a negative control as described previously (25). Cells transfected with 175 ng of esiRNA with 0.3 μl of Lipofectamine RNAiMAX (Invitrogen) and suspended in a total of 60 μl of OptiMEM were prepared according to the manufacturer’s protocol. The transfec-
CD133 N-Glycans Contribute to AC133 Recognition

tion mixtures were reverse transfected into 10,000 HEK293/CD133-VA or HEK293/CD133-YFP cells on tissue culture-treated 96-well glass bottom plates in a volume of 100 μl of DMEM + 10% FBS per well. Medium was changed after 24 h. Targeted gene knockdown by esiRNA was allowed to occur for an additional 48 h.

Cellomics High Content Screening—44 genes were validated using esiRNAs (supplemental Table 2) transfected into HEK293/CD133-VA or HEK 293/CD133-YFP cells, as described above under “esiRNA Interference,” in two independent replicates on 96-well plates (605182; Packard). 72 h after transfection, cells were fixed with 4% paraformaldehyde (electron microscopy grade; Electron Microscopy Sciences) in PBS (Wisent) for 10 min at room temperature. esiRNA-transfected HEK293/CD133-VA samples were incubated overnight at 4 °C with AC133/1 pure. Samples were washed twice with 0.25% BSA fraction V (Wisent) in PBS and incubated with goat anti-mouse Alexa Fluor 647-conjugated secondary antibodies (Molecular Probes) and Hoechst 33342 (Molecular Probes) for 1 h at room temperature away from light. Samples were washed twice with each 0.25% BSA fraction V and PBS. Cells were imaged using the ArrayScan VTi HCS reader automated fluorescent microscope (Cellomics, Thermos Scientific). Images were acquired using a 5× (0.25NA) objective lens and an ORCA-ER 1.00 digital camera. The microscope was set to autofocus in the Hoechst channel identifying individual nuclei based on an intensity value threshold of at least 61 (exposure time 0.008 s). Using the Target Activation Algorithm V3 (Thermo Scientific), an object mask was generated along the nuclei outer border based on the Hoechst intensity. The object mask from the Hoechst channel was copied to the GFP or Cy5 channel and expanded 10 pixels in diameter to cover the entire cell (supplemental Fig. 1). Cell surface AC133 intensity value per object was determined using the Cy5 channel (exposure time 0.8 s). CD133-YFP expression was determined using the GFP channel. 10,000 individual cells from 9 fields covering the entire well were imaged, and quantitative data of AC133 staining per cell or YFP expression per cell were acquired for each sample. Data were exported to Microsoft Excel and analyzed. Nonspecific AC133 staining or background YFP expression from the parental HEK293 cells was subtracted from values. Values were relative to the esiLuc control on the plate. Knockdowns that resulted in mean AC133 or YFP intensity per object of at least three S.D. away from the control esiLuc were considered as hits.

Quantitative PCR—RNA extraction and quantitative RT-PCR were performed as described previously (19) in technical triplicate and with the following modification. A standard curve for each primer set was generated by serial dilution of a mixture of equal amounts of each cDNAs sample of the sample set being analyzed. CD133 primers used were reported previously (26).

Immunofluorescence—Immunofluorescence of HEK293 cells stably expressing wild-type or mutant CD133 was described previously (19) with the addition of α-calnexin (1:1,000; Santa Cruz Biotechnology) at the 4 °C overnight incubation step when applicable. Z-stack images were captured at room temperature by the Leica DMi6000B confocal microscope with a Leica 20×/0.4 NA objective lens and a Hamamatsu EM-CCD digital camera (C9100-13), and imported into Volocity software. Extended focus was taken for images, and images were merged. The imaging medium was PBS.

Immunoprecipitation—FLAG immunoprecipitation was performed as described previously (19) with the exception of 0.5% Triton X-100 was used to lyse HEK293 cells, in replacement of Nonidet P-40.

RESULTS

Generation of a Human Cell Line Stably Expressing AC133—To discover factors involved in AC133 expression, we engineered the HEK293 cell line to express CD133 stably using MAPLE technology (19) (Fig. 1A). We elected to use HEK293 cells because they allow proper glycosylation of CD133 and cell surface AC133 expression (27). We chose to use CD133 isoform 2 because it has been demonstrated previously to be the predominantly expressed CD133 isoform in adult human stem cells and various cancers (27). Western blot analyses of protein lysates prepared from these cells demonstrated that a specific band at ~130 kDa corresponded to glycosylated CD133-VA (the VA tag contributes an additional ~10 kDa to CD133) as determined by immunoblotting with either FLAG and AC133 antibodies (Fig. 1B). Furthermore, FLAG and AC133 immunofluorescence of HEK293/CD133-VA cells displayed a plasma membrane staining pattern (Fig. 1C). This was confirmed further by the cell surface recognition of the AC133 epitope as determined by FACS analysis (Fig. 1D). To obtain a more homogeneously expressing AC133 population, we used FACS to sort the highest 18.5% of cell surface AC133-expressing HEK293/CD133-VA cells (hereafter referred to as HEK293/AC133).

Pooled RNAi Screen to Identify Cell Surface AC133 Regulatory Genes—We took a loss-of-function RNAi approach to identify genes that regulate cell surface AC133 expression (Fig. 2). To this end, a lentivirus-based pooled short hairpin RNAi (shRNA) library containing ~54,000 unique shRNAs designed to target 11,248 human genes (hereafter referred to as the 54K pool) was used (20–22) to infect 1 × 10^7 HEK293/AC133 cells at a multiplicity of infection of 1. This resulted in a 185-fold representation of each of the 54K shRNAs. Seven days after infection, pooled shRNA-treated HEK293/AC133 cells were stained with the AC133 antibody and counterstained with propidium iodide, a marker for dead or dying cells. The replicate cell populations were subsequently separated into AC133low (the lowest 2% of the cell population) and AC133high populations (i.e. the remaining 98% of cells) by cell sorting.

To quantify the shRNA barcodes from sorted fractions and to determine a target hit list, genomic DNA from each of the AC133low and AC133high sorts (n = 3) was made and used as a template for PCR amplification of shRNA barcodes. shRNA-amplified barcodes were then hybridized to a custom microarray and quantified as described previously (23) (supplemental Table 1). shRNAs were considered hits if their barcode intensity from microarray analysis was ≥5-fold higher in the AC133low fractions than in the AC133high fractions, which resulted in ~5% of the 54K shRNA library being identified as a hit from our screen. The pooled 54K shRNA library was designed to contain a set of negative control shRNAs targeting non-human genes,
including green fluorescent protein, red fluorescent protein, β-galactosidase, and firefly luciferase. None of the control shRNAs were identified as a hit in our screen. This supports the list of shRNAs or gene hits from the primary screen as putative regulators of cell surface AC133 expression.

**Secondary Screen to Validate Putative AC133 Regulators**

We selected 44 target genes that met the criteria described above as a hit in the primary screen for validation in a secondary screen. Among the hits to be validated were genes with functions in transcription, phosphorylation, receptor signaling, extracellular matrix organization, and glycosylation. To circumvent potential off-target effects specific to the shRNAs that were identified as a hits, we used an orthogonal RNAi approach, esiRNAs, to target gene knockdown (24). HEK293/CD133-VA cells were transfected with individual esiRNAs targeting each of 44 genes in an arrayed format. After 72 h, cells were stained for cell surface AC133 and imaged using a high content imaging platform to quantify AC133 staining per cell for each esiRNA knockdown (supplemental Fig. 1). Given the variable expression level of AC133 in our HEK293/CD133-VA cell line, we
took the mean cell surface AC133 staining per cell for each sample well.

Gene hits were considered validated if their esiRNA knockdown resulted in a mean cell surface AC133 intensity that was a minimum of 3 S.D. away from the negative control esiRNA targeting firefly luciferase (esiLuc) (supplemental Fig. 2A). 31 of the 44 genes tested could be validated by this criterion, resulting in a 70.5% validation rate for our secondary screen. This confirmed that a high percentage of hits from our primary screen were valid and could be subsequently analyzed to determine pathways involved in cell surface AC133 expression.

To investigate whether the selected target genes that resulted in a significant decrease in cell surface AC133 were due to decreased total CD133 protein levels, we screened the 44 selected gene hits using a HEK293 cell line expressing CD133 fused to the YFP in the same manner as described above (supplemental Fig. 2B). The CD133-YFP fusion allowed monitoring of total CD133 protein levels by measuring YFP expression levels. 10 of the 31 validated gene hits (32.3%) resulted in a significant decrease in total CD133 levels (supplemental Fig. 2B). Taken together, our primary screen identified factors that regulate CD133 protein stability, as well as factors that specifically regulate cell surface AC133.

The N-Glycan Biosynthesis Pathway Is Involved in Cell Surface AC133 Expression—To uncover pathways that may be enriched in our screening hits, we examined our hit list using Ingenuity Canonical Pathway analysis, which relies on the curated KEGG database. Interestingly, Ingenuity Canonical Pathway analysis of our gene hit list identified the KEGG N-glycan biosynthesis pathway to be significantly enriched (19/93 genes with a Fisher’s exact p value of 7.47 × 10^{-5}). In addition, several genes involved in N-glycosylation were validated from our secondary screen (Fig. 3), including ALG1, ALG5, ALG8, ALG10, GANC, MGAT3, and MGAT4C. This provides further evidence that the N-glycan biosynthesis pathway is involved in cell surface AC133 expression. Recent findings by Zhou et al. (28) have demonstrated that α2,3-sialylation regulates CD133 stability because neuraminidase treatment promotes its degradation via the lysosomal pathway. Consistent with their findings, we have identified the α2,3-sialyltransferases ST3GAL4, ST6GAL1, and ST6GAL2 as hits from our primary screen, further supporting the role of the N-glycosylation pathway in AC133 cell surface recognition.

Given that CD133 has been shown to be a heavily N-glycosylated protein (2) and that the N-glycan biosynthesis pathway fits the hypothesis that cell surface AC133 recognition relies on CD133 posttranslational regulation and/or modification (18, 29), we investigated further the role of N-glycan biosynthesis in cell surface AC133 expression.

Tunicamycin Results in CD133 Instability and Loss of Cell Surface AC133 Recognition—Prior to protein N-glycosylation, precursor glycans are synthesized on dolichol phosphate (Dol-P) along the endoplasmic reticulum (ER) membrane to generate Dol-P-P-Glc3Man9GlcNAc2. The ALG enzymes catalyze the biosynthesis of this N-glycan precursor and were among confirmed hits in our screen, including ALG1, ALG5, ALG8, and ALG10. To confirm further the importance of the biosynthesis of Dol-P-P-Glc3Man9GlcNAc2 for AC133 recognition, we treated HEK293/CD133-VA cells with the small molecule tunicamycin. Tunicamycin inhibits the first step of Dol-P-P-Glc3Man9GlcNAc2 biosynthesis catalyzed by ALG7, resulting in the inability to transfer the first GlcNAc residue to
CD133 N-Glycans Contribute to AC133 Recognition

These mutants contained a subpopulation of cells that were positive for CD133-YFP, but absent for the cell surface AC133 staining (Fig. 5C).

Given that none of the single-site N-glycan-deficient CD133 mutants resulted in a dramatic reduction of cell surface AC133, we mutated all eight N-glycosylation sites in CD133 (ΔN-glycans) and examined the consequences. Although the YFP tag on CD133 allowed us to monitor CD133 protein expression independently from cell surface AC133 expression, it may negatively impact CD133 expression because a subpopulation of cells that were both YFP- and AC133-negative could be observed (Fig. 5C). Therefore, we decided to introduce CD133ΔN-glycans into the MAPLE vector (19). We also used CD133(N405Q) as a control from the single-site N-glycan-deficient mutants to demonstrate that any changes in cell surface AC133 relative to wild-type CD133 were not due to the vector system. In contrast to single-site CD133 N-glycan-deficient mutants, the CD133ΔN-glycan mutants had no detectable cell surface AC133 expression (Fig. 6A). Contrary to the results obtained from tunicamycin treatment, loss of CD133 N-glycosylation does not affect its stability (Fig. 6B) or its transcription (Fig. 6C) as measured by the CD133ΔN-glycan mutants. Immunostaining of CD133 in HEK293/CD133ΔN-glycan mutants showed the loss of cell surface localization as determined by co-localization with the ER chaperone calnexin (CANX) and compared with wild-type CD133 or the CD133(N405Q) mutant (Fig. 6D).

Our experiments on CD133 N-glycan mutants demonstrate that individual N-glycan sites are not essential for proper cell surface AC133 expression, but are collectively required. A number of studies have shown an increase in cell surface AC133 expression under hypoxic conditions (31, 32). Although CD133 mRNA levels correlate with this increase (31, 32 and supplemental Fig. 3C), we examined whether N-glycosylation is also a contributing factor. Consistent with previous studies, we observed a significant increase in cell surface AC133 expression when HEK293/CD133-VA wild-type cells were grown under hypoxic conditions (supplemental Fig. 3A). However, hypoxia could not rescue the CD133ΔN-glycan mutants with regard to cell surface AC133 expression, and Western blot analysis showed no obvious differences in the molecular mass of wild-type CD133 (supplemental Fig. 3, A and B). Together, this suggests that CD133 N-glycosylation is not influenced by varying oxygen concentrations. Interestingly, CD133 protein and transcript levels are significantly increased in hypoxic conditions independently from N-glycosylation (supplemental Fig. 3, B and C), suggesting that the increase in cell surface AC133 levels during hypoxia is due to increased total CD133 levels.

Complex CD133 N-Glycans Contribute to Cell Surface AC133—We observed that two bands of different molecular mass are specific for CD133-VA when expressed in HEK293 cells (Fig. 1B). Given that glycoproteins that have passed quality control in the ER can be further processed in the Golgi for the creation of either hybrid or complex N-glycan structures by MGAT family of genes, we hypothesized that the different molecular mass bands were due to differential N-glycosylation. This hypothesis is consistent with the observation that treatment of HEK293/CD133-VA cells with either tunicamycin (Fig. 4A) or site-directed mutagenesis of all potential N-glycosa-
CD133 N-Glycans Contribute to AC133 Recognition

To explore our hypothesis further that CD133 undergoes complex N-glycosylation and that this complexity is required for cell surface AC133 expression, we performed a series of glycosidase treatments using purified CD133-VA to characterize each of the observed CD133 bands (Fig. 1B). Purified

tion sites (Fig. 6B) resulted in the collapse of both bands. Furthermore, this hypothesis is supported by the identification of genes involved in complex N-glycan processing from the primary screen including MGAT3, MGAT4B, MGAT4C, and MGAT5.

FIGURE 5. Single-site N-glycosylation CD133 mutants. A, lysates of single-site N-glycan-deficient CD133 mutants expressed in HEK293 cells were immunoblotted for AC133/1. β-Actin was used as a loading control. The experiment was performed in biological triplicate, and a representative replicate is shown. B, quantitative PCR was used to monitor CD133-YFP transcript of wild-type or single-site N-glycosylation CD133 mutant stables. Transcript levels were normalized to actin and are relative to wild-type CD133. Error bars represent S.D. of three independent replicates. C, representative FACS plot of wild-type CD133 or single-site N-glycosylation CD133 mutants tagged with YFP stably expressed in HEK293 cells stained with AC133-APC is shown (n = 3). Arrows indicate subpopulations that were noticeably different from WT.

FIGURE 6. CD133 N-glycans determine AC133 cell surface recognition. A, representative histogram from three biological replicates of HEK293 cells expressing either wild-type or mutated CD133 stained with AC133 is shown. B, CD133 levels were monitored by immunoblotting for wild-type or mutant CD133, and β-actin was used as a loading control. A representative set of blots is shown (n = 3). C, quantitative PCR was used to monitor CD133 transcript levels in HEK293 cells stably expressing wild-type or mutant CD133-FLAG. Transcript levels were normalized to actin, and error bars represent S.D. (n = 3). p values were determined by Student’s paired t test, with two-tailed distribution. D, FLAG immunofluorescence of HEK293 cells stably expressing wild-type or mutant versions of CD133-FLAG as indicated. Cells were also co-stained for the ER marker CANX. Scale bar, 50 μm.
CD133-VA was first treated with PNGase F, an endoglycosidase that cleaves GlcNAc linked to asparagine residues regardless of the N-glycan type (33). Treatment of CD133-VA with PNGase F resulted in both bands collapsing together and shifting to a molecular mass of ~100 kDa (Fig. 7A, lane 1 versus lane 3), which is the predicted mass of the nascent CD133-VA polypeptide. These changes in molecular mass of the CD133-specific bands suggest that both bands represent CD133 linked with N-glycans of different types. In parallel, we had treated purified CD133 with the endoglycosidase Endo Hf, which cleaves between the two innermost GlcNAc subunits linked to asparagines of only N-glycans that are of high mannose type and of some hybrid types, but is blocked by complex N-glycan biosynthesis, prior to FLAG immunoprecipitation of CD133. Consequently, this CD133 band is shifted to ~100 kDa (Fig. 7A, lane 1 versus lane 3), suggesting that this CD133 glycoform is not complex.

Prior to complex N-glycosylation, α-mannosidase II in the medial–Golgi cleaves the two terminal mannose sugars from GlcNAcMan$_3$GlcNAc$_2$ to generate the GlcNAcMan$_2$GlcNAc$_2$ intermediate. This intermediate is processed by proteins encoded by the MGAT family of genes to create complex N-glycan structures. Without the activity of α-mannosidase II, the resulting outcome is the induction of hybrid N-glycans, but not complex ones. Pretreatment of HEK293/AC133 cells with swainsonine, a small molecule inhibitor of α-mannosidase II (34) and complex N-glycan biosynthesis, prior to FLAG immunoprecipitation of CD133, resulted in the inability of CD133 to undergo complex N-glycan processing, but not its expression level (Fig. 7A, lane 6). Swainsonine treatment resulted in a significant reduction in cell surface AC133 expression as determined by FACS (Fig. 7B), which provides evidence that complex CD133 N-glycosylation is a contributing factor to cell surface AC133 detection.

**MGAT4C Contributes to Complex CD133 N-Glycan Processing**—Although swainsonine inhibits complex N-glycosylation, it does not provide details regarding the extent of N-glycan complexity. We focused on mannosyl (α-1,3-) glycoprotein β-1,4-N-acetylglucosaminyltransferase (isozyme C) (MGAT4C) because it validated as a hit in our screen (supplemental Fig. 2), and it is the most poorly characterized MGAT gene. In fact, no substrates for MGAT4C have been identified to...
CD133 N-Glycans Contribute to AC133 Recognition

date. MGAT4C is conserved in vertebrates and is predicted to be a glycosyltransferase that catalyzes a GlcNAc β–4 linkage to the core mannose residues of N-glycans, resulting in N-glycans with tri- or tetraantennary structures. Furthermore, MGAT4C is represented as one of the terminal genes in the KEGG N-glycan biosynthesis pathway (Fig. 3). Its activity on substrates would depend on upstream glycosylation-processing events. Therefore, demonstrating a role for MGAT4C in CD133 N-glycan processing and cell surface AC133 recognition would suggest the requirement of upstream N-glycosylation processing.

To validate that MGAT4C is involved in CD133 N-glycan processing in HEK293 cells, we knocked down MGAT4C in HEK293/CD133-VA cells using lentiviral-based RNA interference. Knockdown of MGAT4C resulted in a considerable reduction of cell surface AC133 (Fig. 7C) and was on-target as determined Western blot analysis (Fig. 7D). Moreover, the higher molecular mass band corresponding to complex CD133 N-glycans was decreased in molecular mass in the shMGAT4C-treated cells compared with control shRNA-treated cells, whereas the lower molecular mass band corresponding to CD133 N-glycans that are not complex did not change (Fig. 7D). In addition, MGAT4C can be immunoprecipitated with CD133-VA (Fig. 7E), suggesting that MGAT4C activity on CD133 is likely to be direct. Given the above, we suggest that CD133 is a substrate of MGAT4C and that MGAT4C is involved in processing complex N-glycan structures on CD133.

DISCUSSION

The utility of AC133 for marking specific stem cell populations as well as lineage-specific cancer stem cells is widespread. However, CD133 has been identified on differentiated cell types, and it has been observed that CD133 transcript and protein expression can be detected despite the absence of AC133 expression (3, 16, 17). The recent findings by Kemper et al. (18) have partially addressed this puzzle by demonstrating that AC133 is down-regulated during cancer cell differentiation whereas CD133 protein and mRNA levels remain unchanged. In our study, we provide experimental evidence to support their hypothesis directly and provide a detailed dissection of the N-glycosylation pathway components that are involved in cell surface AC133 expression.

We present a pooled large scale lentivirus-based shRNA screen in HEK293/AC133 cells. shRNAs that resulted in a dramatic reduction of cell surface AC133 (AC133low) were sorted by FACS and identified by microarray analysis (23). Pathway analysis of genes corresponding to shRNA hits showed enrichment for the N-glycan biosynthesis pathway in AC133 expression.

The experimental design and workflow of our genetic screen relied on the nature and the type of cells that were used. As AC133-expressing cells from tumor samples are described as a rare subpopulation (15), it would be impractical to use such limited cells for large scale genetic screens. And although a number of established cell lines have been reported to express cell surface AC133 endogenously, they exist as a heterogeneous population with varying AC133 expression levels (35). Our objective was solely to identify genes that regulated posttranslational CD133 for cell surface AC133 exposure, in an effort to address the current hypothesis of AC133 regulation in the field.

Although HEK293 cells do not possess stem cell-like properties found in cells that are typically marked by cell surface AC133, they do contain the biological pathways and processes for the posttranslational regulation and modifications required for cell surface AC133 recognition of exogenously expressed CD133 as shown by our study and previous reports (27). We used the predominant CD133 isoform identified in stem cells for our study (27); however, we cannot ignore the possibility of differences in the impact of N-glycosylation on other CD133 isoforms, which would require further investigation. Our MAPLE-generated HEK293 cell line expresses cell surface AC133 highly and provided us with a unique cell system to interrogate AC133 regulatory genes by genetic means.

The pooled RNAi genetic selection screen took advantage of a lentiviral shRNA library designed to target a total of 11,248 genes. HEK293/AC133 cells infected with this library were appropriately stained and subjected to FACS to positively select a subpopulation of shRNA-harboring cells that displayed decreased cell surface AC133 staining. The screen was restricted to viable cells as shRNAs that target essential genes are filtered by excluding cells in our FACS screen that stain positive for propidium iodide (see “Experimental Procedures”). Together, our genetic screening approach was a powerful means to identify systemically genes involved in cell surface AC133 expression in an unbiased manner from a large set of genes.

Although each shRNA was designed to specifically target only a single gene, we cannot eliminate off-target effects or insufficient knockdown as possible outcomes, which could result in false positive and false negatives, respectively. We elected to use esiRNAs to knock down a selected set of hits in a secondary validation screen because they provided an orthogonal approach to the shRNA library used in our primary screen. Genes that did not validate in the secondary screen may be false positives, or the esiRNAs may not be effective against the intended target. Nevertheless, we were able to show that hits could be validated at a rate of 70.5% (31/44). Further analysis of validated hits using cells expressing a CD133-YFP fusion revealed that a fraction of these factors also regulate CD133 protein stability, which ultimately affects cell surface AC133 expression. Further investigation into the biological pathways, processes, and mechanisms of regulation of these factors is required. Taken together, a high percentage of gene hits from the primary screen could be validated and dissected into cell surface AC133 or total CD133-regulating factors, suggesting that our pooled lentiviral shRNA screen is a powerful approach systematically to identify genes involved in regulating cell surface AC133 expression.

Miraglia et al. (2) have stated previously that AC133 relies on CD133 N-glycosylation. However, no data were directly shown to support this claim. They concluded that AC133 recognized a glycosylated structure of CD133 as determined by the inability of the AC133 antibody to immunoprecipitate CD133 from cells treated with tunicamycin (2). We found that HEK293/AC133 cells treated with tunicamycin resulted in a destabilization of the CD133 protein, which could potentially explain the inability
to immunoprecipitate CD133. In addition, in this study and in that of Kemper et al. (18), AC133 did not bind to a glycosylated epitope of CD133. Interestingly, mutation of N-glycosylation sites did not appear to alter drastically CD133 expression or stability, suggesting that the observations of CD133 instability from tunicamycin treatment are indirect. This may be due to the well known role of tunicamycin treatment in ER stress, which consequently results in perturbations in glycoprotein biosynthesis and in cell death (36).

Although single-site N-glycan-deficient CD133 mutants did not dramatically affect cell surface AC133 recognition, complete loss of CD133 N-glycosylation did. The CD133ΔN-glycan mutant was absent from the cell surface but was expressed and localized to the ER, as measured by co-localization with the CANX. CANX is an ER lectin that specifically interacts and with Glc,Man,GlcNAc2 glycans linked to polypeptides and functions as a chaperone by controlling the proper folding and assembly of glycoproteins for downstream N-glycosylation processing (37). It is conceivable that CD133ΔN-glycan is unable to bind to CANX, which results in improper protein folding and inability to pass the quality checkpoint in the ER. Intriguingly, these two distinct patterns of CD133 subcellular localization (extracellular and intracellular) have been reported in tumor samples from various cancer types (38–41). Given the above, our findings may speak to the glycosylation status of differentially localized CD133.

Although we demonstrate that CD133 N-glycosylation is required for cell surface AC133 expression, our genetic screen also identified a number of genes involved in complex N-glycan processing, including MGAT3, MGAT4B, MGAT4C, and MGAT5. However, inhibition of complex N-glycan processing of CD133 did not completely abolish cell surface AC133 (Fig. 7B), suggesting that it is a contributing, but not essential, factor. Given the above, it is conceivable that different CD133 glycoforms created from N-glycosylation of specific sites on CD133 and the extent of CD133 N-glycan processing result in varying degrees in cell surface AC133 recognition.

We further investigated the impact of complex CD133 N-glycan processing by focusing on MGAT4C. We found that MGAT4C processes complex CD133 N-glycans directly. Prior to our findings, MGAT4C was predicted to be a glycosyltransferase involved in catalyzing GlcNAc β1–4 linkage to the core mannose residues of N-glycans for tri- or tetraantennary N-glycan structures. Our demonstration that CD133 is a MGAT4C substrate provides experimental support for this prediction. Furthermore, given the utility of CD133 as a CSC marker, our report that CD133 is the first known substrate of MGAT4C suggests that MGAT4C may have an important role in certain cancers.

Kemper et al. (18) demonstrated that CD133 N-glycosylation in differentiated cancer cells was maintained to a lesser extent than in CSCs. Furthermore, as the loss of complex N-glycosylation does not appear to alter CD133 stability (Figs. 6B and 7B) or its plasma membrane localization (data not shown), our findings could support the proposal by Kemper et al. that differential N-glycan processing may contribute to a certain CD133 conformation or changes in CD133 tertiary structure to promote cell surface AC133 exposure.

We demonstrated that CD133 N-glycans are involved in cell surface AC133 expression and cell staining for cell surface AC133 may be a readout for certain N-glycosylation processing and activities. This is conceivable given that cell surface AC133 has already been used extensively to define distinct cell states (e.g. primitive versus differentiated cells). For example, complex N-glycosylation is enriched in CD133+ versus CD133− cells from cord blood (42). Our study suggests that further investigation of the glycobiochemical differences between primitive and differentiated cells may shed light on their distinct cellular behaviors.

Acknowledgments—We thank J. W. Dennis for critical reading of the manuscript and for insightful discussions, L. Pelletier for providing reagents, and J. Yam and P. Mero for technical assistance.

REFERENCES
1. Yin, A. H., Miraglia, S., Zanjani, E. D., Almeida-Porada, G., Ogawa, M., Leary, A. G., Olweus, J., Kearney, J., and Buck, D. W. (1997) Blood 90, 5002–5012
2. Miraglia, S., Godfrey, W., Yin, A. H., Atkins, K., Warnke, R., Holden, J. T., Bray, R. A., Waller, E. K., and Buck, D. W. (1997) Blood 90, 5013–5021
3. Corbeil, D., Röder, K., Hellwig, A., Tavian, M., Miraglia, S., Watt, S. M., Simmons, P. J., Peault, B., Buck, D. W., and Huttner, W. B. (2000) J. Biol. Chem. 275, 5512–5520
4. Weigmann, A., Corbeil, D., Hellwig, A., and Huttner, W. B. (1997) Proc. Natl. Acad. Sci. U.S.A. 94, 12425–12430
5. Bhatia, M. (2001) Leukemia 15, 1685–1688
6. Singh, S. K., Hawkins, C., Clarke, I. D., Squire, J. A., Bayani, J., Hide, T., Henkelman, R. M., Cusimano, M. D., and Dirks, P. B. (2004) Nature 432, 396–401
7. O’Brien, C. A., Pollett, A., Gallinger, S., and Dick, J. E. (2007) Nature 445, 106–110
8. Ricci-Vitiani, L., Lombardi, D. G., Pilozzi, E., Biffoni, M., Todaro, M., Peschle, C., and De Maria, R. (2007) Nature 445, 111–115
9. Olempska, M., Eisenach, P. A., Ammerpohl, O., Ungefroren, H., Fandrich, E., and Kalthoff, H. (2007) Hepatobiological Pancreas. Dis. Int. 6, 92–97
10. Yin, S., Li, J., Hu, C., Chen, X., Yao, M., Yan, M., Jiang, G., Ge, C., Xie, H., Wan, D., Yang, S., Zheng, S., and Gu, J. (2007) Int. J. Cancer 120, 1444–1450
11. Monzani, E., Facchetti, F., Galmozzi, E., Corsini, E., Benetti, A., Cazzavin, G., Gritti, A., Piccinni, A., Porro, D., Santinami, M., Invernici, G., Parati, E., Alessandri, G., and La Porta, C. A. (2007) Eur. J. Cancer 43, 935–946
12. Collins, A. T., Berry, P. A., Hyde, C., Stower, M. J., and Maitland, N. J. (2005) Cancer Res. 65, 10946–10951
13. Miki, J., Furuosato, B., Li, H., Gu, Y., Takahashi, H., Egawa, S., Sesterhenn, I. A., McLeod, D. G., Sraivasta, S., and Rhim, J. S. (2007) Cancer Res. 67, 3133–3161
14. Curley, M. D., Therrien, V. A., Cummings, C. L., Sergent, P. A., Koulouris, C. R., Friel, A. M., Roberts, D. J., Seiden, M. V., Scadden, D. T., Rueda, B. R., and Foster, R. (2009) Stem Cells 27, 2875–2883
15. Ferrandina, G., Petrillo, M., Bonanno, G., and Scambia, G. (2009) Expert Opin. Ther. Targets 13, 823–837
16. Floroc, M., Haase, M., Marzesco, A. M., Freund, D., Ehninger, G., Huttner, W. B., and Corbeil, D. (2005) Cell Tissue Res. 319, 15–26
17. Sgambato, A., Puglisi, M. A., Errico, F., Rafanelli, M., Boninsegna, A., Retino, A., Genovese, G., Coco, C., Gasbarrini, A., and Cittadini, A. (2010) J. Cell. Physiol. 224, 234–241
18. Kemper, K., Sprick, M. R., de Bree, M., Scoppelliti, A., Vermeulen, L., Hoek, M., Zeijstra, J., Pals, S. T., Mehmet, H., Stassi, G., and Medema, J. P. (2010) Cancer Res. 70, 719–729
19. Mak, A. B., Ni, Z., Hewel, J. A., Chen, G. I., Zhong, G., Karamboulas, K., Blakely, K., Smiley, S., Marcon, E., Rogovev, D., Li, J., Olsen, J. B., Wan, C., Punna, T., Isserlin, R., Chetyrkin, S., Gineras, A. C., Emili, A., Greenblatt, J. A. (2011) J. Cell. Physiol. 224, 234–241
CD133 N-Glycans Contribute to AC133 Recognition

J. and Moffat, J. (2010) Mol. Cell. Proteomics 9, 811–823

20. Moffat, J., Grueneberg, D. A., Yang, X., Kim, S. Y., Kloepfer, A. M., Hinkel, G., Piquani, B., Eisenhaure, T. M., Luo, B., Grenier, J. K., Carpenter, A. E., Foo, S. Y., Stewart, S. A., Stockwell, B. R., Hacohen, N., Hahn, W. C., Lander, E. S., Sabatini, D. M., and Root, D. E. (2006) Cell 124, 1283–1298

21. Cheung, H. W., Cowley, G. S., Weir, B. A., Boehm, J. S., Rusin, S., Scott, J. A., East, A., Ali, I. D., Lizotte, P. H., Wong, T. C., Jiang, G., Hsiao, J., Mermel, C. H., Getz, G., Barretina, J., Gopal, S., Tamayo, P., Gould, J., Tsherniak, A., Stransky, N., Luo, B., Ren, Y., Drapkin, R., Bhatia, S. N., Mesirow, J. P., Garraway, L. A., Meyerson, M., Lander, E. S., Sabatini, D. M., and Root, D. E. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 20380–20385

22. Luo, B., Cheung, H. W., Subramanian, A., Sharifnia, T., Okamoto, M., Yang, X., Hinkle, G., Boehm, J. S., Beroukhim, R., Weir, B. A., Mermel, C., Barbie, D. A., Awad, T., Zhou, X., Nguyen, T., Piquani, B., Li, C., Golub, T. R., Meyerson, M., Hacohen, N., Hahn, W. C., Lander, E. S., Sabatini, D. M., and Root, D. E. (2006) Proc. Natl. Acad. Sci. U.S.A. 103, 10388–10393

23. Ketela, T., Heisler, L. E., Brown, K. R., Ammar, R., Kasimer, D., Surendra, A., Ericson, E., Blakely, K., Karamboulas, D., Smith, A. M., Durbin, T., Arnoldo, A., Cheung-Ong, K., Koh, J. L., Gopal, S., Cowley, G. S., Yang, X., Grenier, J. K., Giaever, G., Root, D. E., Moffat, J., and Nislow, C. (2011) Genomics 12, 213

24. Kittler, R., Heninger, A. K., Franke, K., Habermann, B., and Buchholz, F. (2005) Nat. Methods 2, 779–784