Extracellular O-Linked N-Acetylglucosamine Is Enriched in Stem Cells Derived from Human Umbilical Cord Blood

Heli Suila, Tia Hirvonen, Ilja Ritamo, Suvi Natunen, Jarno Tuimala, Saara Laitinen, Heidi Anderson, Johanna Nystedt, Jarkko Räbinä, and Leena Valmu

Abstract

Stem cells have a unique ability to self-renew and differentiate into diverse cell types. Currently, stem cells from various sources are being explored as a promising new treatment for a variety of human diseases. A diverse set of functional and phenotypical markers are used in the characterization of specific therapeutic stem cell populations. The glycans on the stem cell surface respond rapidly to alterations in cellular state and signaling and are therefore ideal for identifying even minor changes in cell populations. Many stem cell markers are based on cell surface glycan epitopes including the widely used markers SSEA-3, SSEA-4, Tra 1-60, and Tra 1-81. We have now discovered by mRNA analysis that a novel glycosyltransferase, epidermal growth factor (EGF) domain-specific O-linked GlcNAc transferase (EOGT), is highly expressed in stem cells. EOGT is responsible for adding O-linked N-acetylglucosamine (O-GlcNAc) to folded EGF domains on extracellular proteins, such as those on the Notch receptors. We were able to show by immunological assays that human umbilical cord blood–derived mesenchymal stromal cells display O-GlcNAc, the product of EOGT, and that O-GlcNAc is further elongated with galactose to form O-linked N-acetyllactosamine. We suggest that these novel glycans are involved in the fine tuning of Notch receptor signaling pathways in stem cells.

Key words: biomarkers; cellular biology; stem cells

Introduction

The outmost layer of cell surface is a dense network of carbohydrates that create cell type–specific signatures, serve as receptors, mediate signals, as well as diversify and fine tune the functions of cell surface proteins and lipids. Cell surface glycan epitopes are ideal for identifying and isolating specific cell types, and are widely used as stem cell markers, for example. We, among others, have previously characterized the glycomes of stem cells from various sources and reported that stem cells display characteristic glycosylation features that change upon differentiation.

O-linked N-acetylglucosamine (O-GlcNAc) on proteins has traditionally been considered as an intracellular modification that cooperates with phosphorylation to regulate a variety of cellular processes including intracellular signaling, cytokinesis, and transcription. It has also been shown to play a role in various human diseases such as type 2 diabetes and neurodegeneration, as well as in cancer etiology. Intracellular O-GlcNAc is a dynamic modification, the turnover of O-GlcNAc is catalyzed by two enzymes: O-GlcNAc transferase (OGT) and O-linked β-N-acetylglucosaminidase (OGA).

Recently, it has become apparent that O-GlcNAc can also be found extracellularly. The glycosyltransferase that catalyzes the transfer of GlcNAc from UDP-GlcNAc to epidermal growth factor (EGF) repeats of extracellular proteins, EGF domain-specific O-linked GlcNAc transferase (EOGT), has also been identified and characterized in Drosophila and mouse. Moreover, it has been observed that extracellular O-GlcNAc can be further modified by galactose to generate an O-linked N-acetyllactosamine (O-LacNAc) structure. EOGT acts independently of the previously known OGT and O-GlcNAcylates serine or threonine residues within the consensus sequence C5XXXXS/TG located between the fifth and sixth conserved cysteines of folded EGF-like domains. EOGT localizes in the lumen of the endoplasmic reticulum and is highly conserved. EOGT has been shown to act on the EGF domains of Notch, Dumpy, Delta, and Serrate in Drosophila, and thrombospondin-1 and EGF domain containing membrane proteins in mouse brain. Recently, it was demonstrated that mutations in human
EOGT induce Adams–Oliver syndrome, a multiple congenital anomaly syndrome characterized by vertex scalp defects and terminal transverse limb defects.\textsuperscript{20,21}

In this report we present evidence that human EOGT is highly expressed in stem cells. As a result of this novel type of enzyme activity, we have detected O-GlcNAc and O-LacNac on the human stem cell surface for the first time. This unique modification may play a crucial role in the fine tuning of Notch signaling in stem cells.

Materials and Methods

Cells

Umbilical cord blood (UCB) units were obtained via the Finnish Cord Blood Bank, Finnish Red Cross Blood Service, Helsinki, Finland. Voluntary donors gave informed consent, and the study protocol was accepted by the ethical review boards of the Helsinki University Central Hospital and the Finnish Red Cross Blood Service. Cord blood was collected after normal vaginal delivery. The preparation of mesenchymal stromal/stem cell (MSC) lines and isolation of CD34\(^+\)/CD38\(^-\) cells from UCB was performed as previously described.\textsuperscript{5,22}

Expression analysis

Total RNA from CD34\(^+\)/CD38\(^-\) cells derived from UCB was purified with RNeasy Mini Kit (Qiagen GmbH, Hilden, Germany). cDNA synthesis was performed using Superscript III kit and oligo dtprimers (Invitrogen by Life Technologies, Carlsbad, CA). PCR was performed according to standard protocols using Phusion polymerase (Finnzymes, Thermo Fisher Scientific, Inc., Waltham, MA) and the following primers: for EOGT nucleotides 10–31 for forward and 1563–1584 for reverse primer (GenBank accession number AJ868234), and for β-2-M nucleotides 88–110 for forward and 398–422 for reverse primer (GenBank accession number NM_004048).

The construction of the in-house mRNA microarray data set has been described in detail.\textsuperscript{9} The public domain microarray expression data set was described by Kilpinen et al.\textsuperscript{23} and is located at www.genesapiens.org.

Dissociation enhanced lanthanide fluorescence immunoassay

For dissociation-enhanced lanthanide fluorescence immunoassay (DELFIA), 2500 UCB-MSCs per well were grown on 96-well plates overnight and fixed with 4% paraformaldehyde for 10 min at room temperature. The fixed cells were blocked with 2% bovine serum albumin (BSA) for 1 h and incubated with 1 μL of β-(1–4)-galactosidase from Streptococcus pneumoniae (Prozyme Inc., Hayward, CA) or with 1 μL of recombinant OGA from Bacteroides thetaiotaomicron (from R&D Systems, Abingdon, United Kingdom) or with 1 μL β-(1–4)-galactosidase and increasing amounts of OGA as indicated or buffer only (50 mM sodium phosphate, 0.1% BSA, 100 mM NaCl, pH 5.8) overnight at 37°C. Cells were washed with DELFIA wash buffer (Perkin Elmer, Turku, Finland) and incubated sequentially at room temperature with 1:50 dilution of O-GlcNAc antibody (CTD110.6) (Cell Signaling Technology Inc., Danvers, MA) and 1 μg/mL Eu-labeled secondary antibody (N1-anti mouse antibody, Perkin Elmer) in assay buffer (50 mM Tris, 0.9% NaCl, 0.3% BSA). After thorough washes DELFIA enhancement solution (Perkin Elmer) was added to each well and the signal was quantitated using VICTOR2 1420 multilabel counter (Perkin Elmer). The experiment was performed with three technical replicates and repeated twice.

Cell surface protein analysis by mass spectrometry

For cell surface protein analysis, UCB-MSCs were detached by mild and fast trypsinization (TrypLE Express, Gibco, Gaithersburg, MD) and washed with phosphate-buffered saline. The cell surface protein fraction was enriched using biotinylation of intact cells and further captured using streptavidin-coupled magnetic beads, as previously described.\textsuperscript{24} In liquid reduction, alkylation, and digestion of proteins were performed as described earlier.\textsuperscript{25} The sample was vacuum dried and dissolved in 0.1% formic acid for mass spectrometric analysis.

Protein digests were analysed with liquid chromatography (LC)–mass spectrometry (MS). Peptides were loaded to a reversed-phase precolumn (ProteoCol Guard-C18, 150 μm × 10 mm, SGE, Austin, TX) with 0.1% formic acid and separated in reversed-phase analytical column (PepMap 100, 75 μm × 150 mm, Thermo Fisher Scientific Inc.) with linear gradient of acetonitrile. Ultimate 3000 LC instrument (Thermo Fisher Scientific Inc.) was operated in nanoscale with a flow rate of 0.3 μL/min. Eluted peptides were introduced to LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific Inc.) via ESI Chip interface (Advion BioSciences Inc., Ithaca, NY) in positive-ion mode.

Data files from mass spectrometer were processed with Mascot Distiller (Matrix Science Ltd., London, UK, version 2.3). The processed data was searched with Mascot Server (Matrix Science Ltd., version 2.2) against human proteins in UniProtKB database (release 2011_09). The search criteria were as follows: enzyme trypsin; maximum missed cleavages 1; variable modifications: lysine 3-(carbamidomethyl)propanoylation, protein N-terminal 3-(carbamidomethyl)propanoylation, cysteine carboxamidomethylation, methionine oxidation; peptide mass tolerance ± 10 ppm; fragment mass tolerance ± 0.8 Da and instrument type ESI-TRAP.

Results

As part of our studies on UCB-derived stem and progenitor cell glycomes we searched for glycosyltransferase-encoding genes that would be enriched in stem cells. We performed reverse-transcription (RT)-PCR analysis in UCB-derived CD34\(^+\) progenitor cells and detected the expression of the gene AER61, now known as EO GT (also known as C3orf64 and EOGT1) (Fig. 1A). Using a public domain mRNA microarray expression database,\textsuperscript{23} we discovered that EO GT is distinctively expressed in different stem and progenitor cell types (Fig. 1B).

It was recently found out that EO G T O-GlcNAcylates EGF-domains on Drosophila Notch receptors,\textsuperscript{14} and that Drosophila Notch has 17 EGF repeats with an EOGT consensus site.\textsuperscript{15} We studied the mRNA expression of Notch receptors in human UCB-derived stem and progenitor cells using our in-house mRNA microarray databases generated from UCB-derived cells.\textsuperscript{9} In UCB-MSCs, Notch 2 and
3 displayed the highest relative mRNA expression levels among the Notch receptors (Fig. 2A). In order to verify the presence of Notch receptors on UCB-MSCs, we extracted the cell surface proteins by biotinylating the UCB-MSC surface and capturing the biotinylated proteins on magnetic streptavidin beads. The cell surface proteins were then digested with trypsin, separated by nanoscale liquid chromatography, and subsequently analyzed by high-resolution MS. Altogether eight tryptic peptides were identified from human Notch 3 with a good Mascot score indicating good quality of the MS/MS fragmentation spectra (Fig. 2B). However, we were unable to detect Notch glycopeptides. This may be due to either the calcium-sensitive shedding of the extracellular part of the Notch receptor or the MS detection limit (many glycopeptides do not pass the detection limit in MS analysis due to ion suppression in the presence of nonglycosylated peptides and heterogeneity of the glycan portion). Human Notch 3 has 21 EGF repeats with a perfect match to the consensus site for EOGT modification. A schematic diagram of human Notch 3 is shown in Fig. 2C.

To detect the presence of O-GlcNAc on the surface of intact UCB-MSCs, we took advantage of a well-characterized O-GlcNAc antibody. Of all the available O-GlcNAc antibodies, CTD110.6 is the least dependent on protein structure and recognizes the widest range of O-GlcNAcylated proteins. As expected, O-GlcNAc was detected on the cell surface (Fig. 3B). However, when the cells were treated with β(1-4)-galactosidase, the signal increased several fold, suggesting that O-GlcNAc is further modified by galactose to generate O-linked LacNAc. Treatment with OGA quenched most of the signal. Treatment with both enzymes using a constant amount of β(1-4)-galactosidase and increasing amounts of OGA resulted in a dose-dependent decrease in signal intensity, indicating specific removal of O-LacNAc and O-GlcNAc, respectively (Fig. 3C). A part of the signal remained after treatment with β(1-4)-galactosidase and...
OGA. Although no cross-reactivity toward common N-linked oligosaccharides have been found for the monoclonal antibody CTD110.6, the background signal may arise from a large amount of exposed GlcNAc terminals on common N- and O-glycans. It can be concluded that while the decrease of signal reflects O-GlcNAc detected by monoclonal antibody CTD110.6 and released by OGA, the remaining signal level is likely to reflect background arising from other GlcNAc terminals revealed by the galactosidase treatment.

**Discussion**

Stem cells are considered to offer future cure for various human diseases. However, several aspects of stem cell biology are still undiscovered. To produce therapeutic cell populations, more knowledge should be gained on the biochemistry of stem cells from different origins. In this study we show by mRNA expression analysis that a novel enzyme responsible for extracellular O-GlcNAcylation of protein EGF domains, EOGT, is distinctively expressed in stem and progenitor cells. Further, we detected O-GlcNAc for the first time on human UCB-MSCs by specific immunochemical detection and showed that the O-GlcNAc moiety undergoes additional modification resulting in the presence of O-LacNAc on the cell surface.

Adams–Oliver syndrome is characterized by cytoskeletal dysregulation with perturbed Notch signaling. It was recently shown that frame shift deletion mutations, which impair O-GlcNAcylation of extracellular protein domains in human EOGT, can induce this congenital disorder. In *Drosophila*, O-GlcNAcylation of extracellular protein domains appears to be important for cell–matrix interactions and the correct targeting of Dumpy. Loss of EOGT is larval lethal, and knock-down of EOGT in wings induces wing blistering. Originally, extracellular O-GlcNAc modification was first identified from *Drosophila* Notch EGF repeat.
The Notch pathway is a highly conserved cell signaling system that is involved in numerous stages of development. It regulates cell fate specification, differentiation, left–right asymmetry, apoptosis, somitogenesis, angiogenesis, and plays a key role in stem cell proliferation (reviewed by Perdigoto and Bardin29 and Jafar-Nejad et al.30). Mammals possess four different Notch receptors. The extracellular domain of Notch receptors can be modified by regular N-glycans and several unusual glycans, which modulate Notch signaling pathways, namely O-glucose,31 O-fucose (reviewed by Stanley and Okajima32), O-xylene,33 and O-GlcNAC.19 The extracellular domain of human Notch 3 is mainly composed of tandem repeats of EGF domains, containing 34 extracellular EGF domains, respectively, and 21 of these contain a perfect match to EOGT modification consensus site.

We speculate that in human stem cells the novel glycans, extracellular O-GlcNAc and extracellular O-LacNAc, play a crucial role in the fine tuning of Notch signaling, even though the exact nature of this regulation needs to be further examined. Extracellular O-GlcNAc modification in stem cells is a novel and exceptional type of glycosylation that was earlier believed to be restricted to the cytoplasm and nucleus. It participates in creating a stem cell–specific surface signature and reflects the unique characteristics of therapeutic stem cells.

Acknowledgments

The authors would like to thank the Finnish Red Cross Blood Service Cord Blood Bank, and acknowledge Teija Kupari (present affiliation: Thermo Fisher Scientific, Finland), Lotta Andersson, Birgitta Rantala, and Sirkka Hirschovits-Gerz for excellent technical assistance. This work was supported by the Finnish Funding Agency for Technology and Innovation (TEKES); the EVO Medical Research Fund of Finnish Red Cross Blood Service, and the Finnish Glycoscience Graduate School.

Author Disclosure Statement

No competing financial interests exist.

References

1. Cummings RD. The repertoire of glycan determinants in the human glycome. Mol Biosyst. 2009;5:1087–1104.
2. Lancetot PM, Gage FH, Varki AP. The glycans of stem cells. Curr Opin Chem Biol. 2007;11:373–380.
3. Heiskanen A, Hirvon T, Salo H, et al. Glycomics of bone marrow–derived mesenchymal stem cells can be used to evaluate their cellular differentiation stage. Glycoconjug J. 2009;26:367–384.
4. Heiskanen A, Satomaa T, Tiitinen S, et al. N-glycolynneuraminic acid xenoantigen contamination of human embryonic and mesenchymal stem cells is substantially reversible. Stem Cells. 2007;25:197–202.
5. Hemmoranta H, Hautaniem S, Niemi J, et al. Transcription-proliferating profiles shared and unique markers for CD34+ and CD133+ cells. Stem Cells Dev. 2006;15:839–851.
6. Hemmoranta H, Satomaa T, Blomqvist M, et al. N-glycan structures and associated gene expression reflect the characteristic N-glycosylation pattern of human hematopoietic stem and progenitor cells. Exp Hematol. 2007;35:1279–1292.
7. Jaatinen T, Hemmoranta H, Hautaniem S, et al. Global gene expression profile of human cord blood-derived CD133+ cells. Stem Cells. 2006;24:631–641.
8. Satomaa T, Heiskanen A, Mikkola M, et al. The N-glycome of human embryonic stem cells. BMC Cell Biol. 2009;10:42.
9. Suila H, Pitkanen V, Hirvon T, et al. Are globoseries glycosphingolipids SSEA-3 and −4 markers for stem cells derived from human umbilical cord blood? J Mol Cell Biol. 2011;3:99–107.
10. Hirvon T, Suila H, Kotouvoori A, et al. The i blood group antigen as a marker for umbilical cord blood–derived mesenchymal stem cells. Stem Cells Dev. 2012;21:455–464.
11. Zachara NE, Hart GW. Cell signaling, the essential role of O-GlcNAC! Biochim Biophys Acta. 2006;1761:599–617.
12. Hart GW, Slawson C, Ramirez-Correa G, et al. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annu Rev Biochem. 2011;80:825–858.
13. Vocadlo DJ. O-GlcNAc processing enzymes: catalytic mechanisms, substrate specificity, and enzyme regulation. Curr Opin Chem Biol. 2012;16:488–497.
14. Matsuura A, Ito M, Sakaiedani Y, et al. Linked N-acetylglucosamine is present on the extracellular domain of notch receptors. J Biol Chem. 2008;283:35486–35495.
15. Sakaiedani Y, Nomura T, Matsuura A, et al. Linked-N-acetylglucosamine on extracellular protein domains mediates epithelial cell-matrix interactions. Nat Commun. 2011;2:583.
16. Sakaiedani Y, Ichiyanagi N, Saito C, et al. Linked-N-acetylglucosamine modification of mammalian Notch receptors by an atypical O-GlcNAc transferase Eogt1. Biochem Biophys Res Commun. 2012;419:14–19.
17. Alfaro JF, Gong CX, Monroe ME, et al. Tandem mass spectrometry identifies many mouse brain O-GlcNAcylated proteins including EGF domain–specific O-GlcNAc transferase targets. Proc Natl Acad Sci U S A. 2012;109:7280–7285.
18. Hoffmann BR, Liu Y, Mosher DF. Modification of EGF-like module 1 of thrombospondin-1, an animal extracellular protein, by O-linked N-acetylglucosamine. PLoS One. 2012;7:e32762.
19. Muller R, Jenny A, Stanley P. The EGF repeat–specific O-GlcNAc-transferase Eogt interacts with notch signaling and pyrimidine metabolism pathways in Drosophila. PLoS One. 2013;8:e62835.
20. Shaheen R, Aghan M, Keppler-Noreuil K, et al. Mutations in EOGT confirm the genetic heterogeneity of autosomal-recessive Adams-Oliver syndrome. Am J Hum Genet. 2013;92:599–604.
21. Cohen I, Silberstein E, Perez Y, et al. Autosomal recessive Adams-Oliver syndrome caused by homozygous mutation in EOGT, encoding an EGF domain–specific O-GlcNAc transferase. Eur J Hum Genet 2014;22:374–378.
22. Lahtinen T, Nystedt J, Laitinen S. The isolation and culture of human cord blood–derived mesenchymal stem cells under hypoxic conditions. In: Mesenchymal Stem Cell Assays and Applications. Vemuri M, Chase LG, Rao MS (eds.) Humana Press: New York, NY; pp. 63–73; 2011.
23. Kilpinen S, Autio R, Ojala K, et al. Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. Genome Biol. 2008;9:R139.
24. Scheurer SB, Roesli C, Neri D, et al. A comparison of different biotinylation reagents, tryptic digestion procedures,
and mass spectrometric techniques for 2-D peptide mapping of membrane proteins. Proteomics. 2005;5:3035–3039.

25. Kinter M, Sherman NE. Protein sequencing and identification using tandem mass spectrometry. In: Protein Sequencing and Identification Using Tandem Mass Spectrometry. Kinter M, Sherman NE, eds. New York: John Wiley & Sons: New York, pp. 160–163; 2000.

26. Rand MD, Grimm LM, Artavanis-Tsakonas S, et al. Calcium depletion dissociates and activates heterodimeric notch receptors. Mol Cell Biol. 2000;20:1825–1835.

27. Comer FI, Vosseller K, Wells L, et al. Characterization of a mouse monoclonal antibody specific for O-linked N-acetylglucosamine. Anal Biochem. 2001;293:169–177.

28. Whelan SA, Hart GW. Identification of O-GlcNAc sites on proteins. Methods Enzymol. 2006;415:113–133.

29. Perdigoto CN, Bardin AJ. Sending the right signal: Notch and stem cells. Biochim Biophys Acta. 2013;1830:2307–2322.

30. Jafar-Nejad H, Leonardi J, Fernandez-Valdivia R. Role of glycans and glycosyltransferases in the regulation of Notch signaling. Glycobiology. 2010;20:931–949.

31. Fernandez-Valdivia R, Takeuchi H, Samarghandi A, et al. Regulation of mammalian Notch signaling and embryonic development by the protein O-glucosyltransferase Rumi. Development. 2011;138:1925–1934.

32. Stanley P, Okajima T. Roles of glycosylation in Notch signaling. Curr Top Dev Biol. 2010;92:131–164.

33. Lee TV, Sethi MK, Leonardi J, et al. Negative regulation of notch signaling by xylose. PLoS Genet. 2013;9:e1003547.

Address correspondence to:

Heli Suila, MSc
Finnish Red Cross Blood Service
Kivihaantie 7
FI-00310 Helsinki
Finland

E-mail: heli.suila@helsinki.fi

Abbreviations Used

- β-2-M = β-2-microglobulin
- BSA = bovine serum albumin
- DELFIA = dissociation-enhanced lanthanide fluorescence immunoassay
- EGF = epidermal growth factor
- EOGT = EGF domain-specific O-linked GlcNAc transferase
- LC = liquid chromatography
- MS = mass spectrometry
- MSC = mesenchymal stromal/stem cell
- MS/MS = tandem mass spectrometry
- OGA = O-linked β-N-acetylglucosaminidase
- O-GlcNAc = O-linked N-acetylglucosamine
- OGT = O-GlcNAc transferase
- O-LacNAc = O-linked N-acetyllactosamine
- RT-PCR = reverse-transcription PCR
- UCB = umbilical cord blood