Elevated O-GlcNAc levels activate epigenetically repressed genes and delay mouse ES cell differentiation without affecting naive to primed cell transition

Speakman, Christopher M.; Domke, Tanja C. E.; Wongpaiboonwattana, Wikrom; Sanders, Kelly; Mudaliar, Manikhandan; van Aalten, Daan M. F.

Published in:
Stem Cells

DOI:
10.1002/stem.1761

Publication date:
2014

Document Version
Peer reviewed version

Link to publication in Discovery Research Portal

Citation for published version (APA):
Elevated O-GlcNAc levels activate epigenetically repressed genes and delay mouse ES cell differentiation without affecting naïve to primed cell transition

AUTHORS: Christopher M. Speakman¹, Tanja C.E. Domke¹, Wikrom Wongpaiboonwattana¹, Kelly Sanders², Manikhandan Mudaliar³,⁴, Daan M.F. van Aalten², Geoffrey J. Barton³, and Marios P. Stavridis¹

¹ Division of Cancer Research, Medical Research Institute, College of Medicine, Dentistry and Nursing, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK
² Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK
³ Division of Computational Biology, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK
⁴ Glasgow Polyomics, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G61 1QH, UK

AUTHORS CONTRIBUTIONS:

Christopher M. Speakman, Tanja C.E. Domke, Wikrom Wongpaiboonwattana: collection and assembly of data, data analysis;
Kelly Sanders: experimental design, collection and assembly of data, data analysis and interpretation;
Manikhandan Mudaliar, Geoffrey J. Barton: data analysis and interpretation;
Daan M.F. van Aalten: provision of reagents; experimental design, data analysis and interpretation
Marios P. Stavridis: conception and design, data collection and assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript.

CORRESPONDENCE:

Marios P. Stavridis, Medical Research Institute, Division of Cancer Research, Jacqui Wood Cancer Centre, University of Dundee, James Arrott Drive, Ninewells Hospital & Medical School, Dundee, DD1 9SY. Telephone: (0044) 1382 383388; E-mail: m.stavridis@dundee.ac.uk

KEY WORDS:

Embryonic Stem Cells, Cell Differentiation, O-GlcNAc, Post-translational Protein Modification, Signal Transduction, Oligonucleotide Microarrays

This is the accepted version of the following article: Speakman, C. M., Domke, T. C.E., Wongpaiboonwattana, W., Sanders, K., Mudaliar, M., van Aalten, D. M.F., Barton, G. J. and Stavridis, M. P. (2014), Elevated O-GlcNAc Levels Activate Epigenetically Repressed Genes and Delay Mouse ESC Differentiation Without Affecting Naïve to Primed Cell Transition. STEM CELLS, 32: 2605–2615, which has been published in final form at http://dx.doi.org/10.1002/stem.1761.
ABSTRACT

The differentiation of mouse embryonic stem (ES) cells is controlled by the interaction of multiple signaling pathways, typically mediated by post-translational protein modifications. The addition of O-linked N-acetylglucosamine (O-GlcNAc) to serine and threonine residues of nuclear and cytoplasmic proteins is one such modification (O-GlcNAcylation), whose function in ES cells is only now beginning to be elucidated. Here we demonstrate that the specific inhibition of O-GlcNAc hydrolase (Oga) causes increased levels of protein O-GlcNAcylation and impairs differentiation of mouse ES cells both in serum-free monolayer and in embryoid bodies (EBs). Use of reporter cell lines demonstrates that Oga inhibition leads to a reduction in the number of Sox1-expressing neural progenitors generated following induction of neural differentiation, as well as maintained expression of the ES cell marker Oct4 (Pou5f1). In EBs expression of mesodermal and endodermal markers is also delayed. However, the transition of naïve cells to primed pluripotency indicated by Rex1 (Zfp42), Nanog, Esrrb and Dppa3 downregulation and Fgf5 upregulation remains unchanged. Finally, we demonstrate that increased O-GlcNAcylation results in upregulation of genes normally epigenetically silenced in ES cells, supporting the emerging role for this protein modification in the regulation of histone modifications and DNA methylation.
INTRODUCTION

Over the last decade there have been major advances in our understanding of the mechanisms controlling embryonic stem (ES) cell behavior in response to the changing extracellular environment. These mechanisms largely involve engagement of signal transduction relays that operate by post-translational modifications of proteins. Under standard conditions ES cells grow as a mixture of “naïve” and “primed” cells. Signaling mediated by Erk1/2 target phosphorylation has recently been implicated in regulating the transition between these two states and initiation of differentiation. Reversible protein modification by addition of O-linked N-acetylglucosamine to serine or threonine residues (O-GlcNAcylation) was first described 30 years ago and occurs with similar time scales, dynamics and stoichiometry as protein phosphorylation, with which it sometimes competes. O-GlcNAcylation is found in all higher eukaryotes tested to date and has been implicated in development, epigenetic regulation and diseases such as diabetes and Alzheimer’s. Its addition and removal are catalyzed by one transferase (Ogt) and one hydrolase (Oga; also known as Mgea5) respectively.

To date there have been very few studies on O-GlcNAc function in embryonic stem cells, although accumulating evidence suggests a critical role for this modification. Elimination of Ogt in mouse ES cells or conditional deletion in somatic cells leads to death, consistent with an essential role in all cell types. Ogt has recently been identified as a protein partner of the essential ES cell transcription factor Oct4 (Pou5f1) by three independent studies. Furthermore, Oct4 has been shown to be modified by O-GlcNAc, and this was demonstrated to be important for regulation of a subset of its targets in ES cells and during reprogramming. A previous study has suggested that increased O-GlcNAcylation in ES cells prevents differentiation along the cardiac lineage in spontaneously differentiating embryoid bodies, however the mechanism for this or the stage at which differentiation
stalled were not determined. Ogt is a mammalian homologue of the *Drosophila super sex combs* (*sex*) gene, a member of the Polycomb group of transcriptional repressors\textsuperscript{13, 14} and has established roles in gene repression\textsuperscript{15}. Furthermore, recent studies implicate O-GlcNAc transferase in the regulation of the Tet epigenetic modifiers\textsuperscript{16-18}, suggesting that the O-GlcNAc modification and the enzymes controlling it regulate chromatin in pluripotent cells by multiple mechanisms.

In the present study we investigate in detail the effects of increased O-GlcNAcylation on neural differentiation of mouse ES cells. Our results show that Oga inhibition leads to delayed onset of neural differentiation although the transition of naïve cells to primed pluripotency proceeds unhindered. We also present a genome-wide gene expression analysis of ES cells and differentiating cells treated with a highly specific Oga inhibitor, revealing that increased O-GlcNAc levels control the expression of distinct groups of genes in ES cells associated with a recently described subpopulation resembling the 2-cell-stage embryo. Upregulation of these genesets is consistent with a disruption of the normal transcriptional repression programme operating in pluripotent cells.

**MATERIALS AND METHODS**

ES cell culture and differentiation

Mouse ES cells were cultured without feeders on 0.1% gelatin coated plastics in GMEM with 10% serum and LIF\textsuperscript{19} and differentiation was performed according to our previous protocol\textsuperscript{20}. The cell lines used were derivatives of E14Tg2aIV (46C\textsuperscript{21}, OCRG\textsuperscript{922}) or Oct4GiP\textsuperscript{23}. GlcNAcstatin C was obtained from GlycoBioChem (http://www.glycobiochem.com) and used at 1µM unless otherwise specified. For flow cytometry cells were dissociated using Accutase (Millipore) and resuspended in PBS/1%BSA containing 1µg/ml propidium iodide (PI; Sigma). Cellular debris and PI-positive cells were excluded from analysis. For clonal analysis
600/100 cells were plated per 10cm dish/well of 6-well plate and cultured for 6-8 days, then fixed and stained using an Alkaline phosphatase staining kit (Sigma).

siRNA induced knockdown of Oga and Ogt

Cells were seeded into 6-well dishes (2x10^5 cells /well). While still in suspension, cells were transfected with SmartPool siRNAs targeting Oga or Ogt or with a non-targeting siRNA pool (50 pmol per well, Thermo Scientific) using Lipofectamine RNAiMAX (Life Technologies) (5 µl per well). Cells were transfected again after 24 h.

Western blotting

Cells were cultured or differentiated in 9 cm dishes and treatments were for 24 h unless otherwise specified. Lysis was performed on ice in whole-cell lysis buffer (150 mM sodium chloride, 1.0% NP-40, 50 mM Tris, pH 8.0 with 1µM GlcNAcstatin C, Complete® protease inhibitor and PhosStop® phosphatase inhibitor cocktail tablets from Roche). For nuclear/cytoplasmic fractionation cells were lysed on ice in Buffer A (10 mM HEPES, 1.5 mM MgCl2, 10 mM KCl, 0.5 mM DTT, 0.05% NP40 pH 7.9 ) for 10 minutes and centrifuged for 10 minutes at 3000 rpm in a microfuge at 4ºC. The supernatant was removed as a cytoplasmic fraction and the pellet was resuspended in Buffer B (5 mM HEPES, 1.5 mM MgCl2, 0.2 mM EDTA, 0.5 mM DTT, 26% glycerol (v/v), pH 7.9) by micropestle homogenisation, NaCl added to 300 mM and incubated on ice for 30 minutes. A final spin at 12000 g was performed to separate insoluble material from the nuclear fraction. Protein concentration was determined by Coomassie protein assay (Pierce) and 10-30 µg of protein was loaded per lane of Life Technologies NuPage gels and transferred to PVDF membrane (Millipore). Membranes were blocked in either 5% Milk (Marvel) in TBST or 1-5% BSA (Merck) in TBST (all other chemicals from Sigma). Antibodies were incubated in blocking buffer overnight: anti-O-GlcNAc, CTD110.6 (Covance, 1/5000) or RL-2 (Santa Cruz, 1/1000); beta actin (Abcam, 1/2000); phosphoErk1/2 (Cell Signaling, 1/1000); Gsk3α and
phosphoGsk3α/β (Cell Signaling, 1/1000); phospho-Serine and phospho-Threonine (Cell Signaling, 1/1000); Oct4 (Abcam 1/1000); Sox2 (Abcam 1/1000). For the phospho-kinase antibody array (R&D Systems) 500 µg of total cell lysate was used according to manufacturers instructions and the arrays were developed using LumiGLO from Cell Signaling and imaged on a Fuji LAS3000 mini scanner. Densitometry was performed with ImageJ.

Immunoprecipitation
500µg of whole cell lysate was incubated overnight with a mixture of succinylated Wheat germ agglutinin-agarose and protein A sepharose beads in the presence of antibody RL-2 (anti-O-GlcNAc). After washing the bound proteins were eluted in 25µl of 2xLDS loading buffer (Life Technologies), boiled for 5 minutes and analysed by Western Blotting.

Immunocytochemistry
Cells were fixed in 4% PFA for 15 minutes, permeabilised in PBS, 0.1% Tween (PBST) and incubated in PBST +5% BSA (blocking buffer). Primary antibody (goat anti-Oct4, Santa Cruz) was diluted in blocking buffer at 1/50 and incubated overnight at 4ºC. Subsequently cells were incubated in fluorescent secondary antibody (AlexaFluor conjugate, Life Technologies, 1/1000) and 300 nM 4',6-diamidino-2-phenylindole (DAPI) for 1 h.

2D-PAGE
Nuclear samples extracted from ES cells treated with LIF, or 1 - 4 days treatment with N2B27 were enzymatically labelled using Click-IT O-GlcNAc labelling system (Life Technologies) according to the manufacturers instructions. Samples were chloroform-methanol precipitated and re-suspended in 7M urea, 2M thiourea, 4% CHAPS, 1% ASB-14 and 0.5% ampholytes. Passive rehydration was performed overnight followed by isoelectric focusing using pH 3-10 NL IPG strips for 4200 vh. Preceding equilibration, samples were separated by SDS-PAGE using 4-12% bis/tris gels and transferred to PVDF for detection by streptavidin-HRP. 10
prominent spots were excised from duplicate gels and analyzed by MALDI-TOF mass spectrometry

Quantitative RT-PCR

RNA was extracted from cells grown in 6-well plates using Nucleospin II RNA kit (Macherey-Nagel). DNase treatment was performed on-column during RNA extraction. One microgram of RNA was reverse transcribed with the qScript cDNA synthesis kit (Quanta Biosciences). PCR was performed on a BioRad iCycler or an AB 7500 using PerfeCTa® SYBR® Green FastMix® for iQ™ or PerfeCTa® SYBR® Green FastMix® Low Rox (both Quanta Biosciences) respectively. Primer sequences are shown in Supplementary Table 1.

Relative quantitation was performed using the method by Pfaffl and β-actin as a reference gene from technical duplicates or triplicates of at least 3 independent experiments.

Microarray analysis

RNA samples (from 4 independent experiments) were processed and hybridised to Affymetrix Gene 1.1ST arrays by Ark Genomics (http://www.ark-genomics.org) and the raw array expression data were obtained as CEL files. For quality control and probe sets annotations, the annotations files (Release 32, dated 23-06-2011) downloaded from Affymetrix were used. Background noise control (Detected Above Background), RMA normalisation and summarisation of probe set level data into transcript clusters were carried out using Affymetrix Power Tools. Quality analysis and differential expression analyses were performed in Partek GS 6.5 (version 6.11.0321) software and R (version 2.13.1) – Bioconductor using Limma and RankProd packages. The microarray data have been deposited in EBI ArrayExpress under accession number E-MEXP-3593. For gene set enrichment analysis (GSEA) GSEA v2.0.13 was used, along with gene sets from MSigDB, GenesigDB or custom made ones from the literature (Supplementary Table 5). Genes were
ranked by the Signal2Noise metric and the weighted2 enrichment statistic was used over 1000 geneset permutations. A false discovery rate q-value cut-off of 0.05 was applied.

RESULTS

Protein O-GlcNAcylation delays mouse embryonic stem cell differentiation

We decided to investigate the changes in O-GlcNAc signaling during ES cell neural differentiation. Global O-GlcNAc levels decline slightly during the first few days of differentiation, as measured by Western blotting with the O-GlcNAc specific antibody CTD110.6 (Figure 1A). However, this method may miss changes in the O-GlcNAcylation of lower abundance proteins. To test for more subtle changes in protein O-GlcNAcylation, we performed a 2D-PAGE-Western blot analysis of nuclear extracts of ES cells and cells during the first 4 days of monolayer differentiation using chemoenzymatic labeling of O-GlcNAcylated proteins (Figure 1B and Supplementary Figure 1A). The results revealed a dynamic pattern of nuclear O-GlcNAcylated proteins, with a dramatic change during the first 24 hours of differentiation. As the O-GlcNAcylation on the majority of the nuclear proteins disappeared following 24h of differentiation, we hypothesized that removal of the O-GlcNAc modification from some ES cell protein(s) may be required for differentiation onset. We also picked 10 spots showing dramatic regulation from the Day 1 gel and identified them by mass spectrometry. Our analysis identified 26 unique proteins. Although we were not able to positively identify O-GlcNAc modification on any of these proteins due to the instrumentation used, we note that 13 of the 26 had previously been identified as O-GlcNAc modified by other screens (Supplementary table 2).

We then sought to determine whether the O-GlcNAc levels in ES cells can be modulated by inhibiting the Oga enzyme activity that removes O-GlcNAc using GlcNAcstatin C (GNS), a highly potent and specific small molecule Oga inhibitor that exhibits 164-fold selectivity for
Oga when tested against the closely related lysosomal hexosaminidases HexA/B\textsuperscript{28} and a competitive inhibitor of the O-GlcNAc transferase Ogt (4Ac5SGlcNAc)\textsuperscript{29}. Treatment of ES cells with GNS resulted in a dose-dependent increase in cellular O-GlcNAc levels (Figure 1C) including O-GlcNAcylation of the transcription factor Sox2 on S248 (Figure 1D), with no discernable effects on cell viability or proliferation even after prolonged treatment (Figure 1E,F). Similar results were obtained by use of siRNA (Supplementary Figure 1B). Inhibition of Ogt on the other hand led to cell death within 4-5 days in culture (Supplementary Figure 1C), consistent with the previous observation that this protein is essential for ES cell viability\textsuperscript{5} (see also Figure 3A).

We then treated murine 46C ES cells (engineered to express the green fluorescent protein GFP from the endogenous neural–specific \textit{Sox1} locus\textsuperscript{21}) with GNS during monolayer differentiation into neural cells\textsuperscript{20}. Flow cytometry showed that a significantly smaller proportion of neural progenitors were generated during ES cell differentiation in the presence of GNS (24.6\% in GNS vs. 32.2\% in DMSO at day 3, a reduction of 23.6\%; n=4, p=0.0118) (Figure 2A, Supplementary Videos 1 and 2). Similar effects were seen using a structurally distinct Oga inhibitor, Thiamet G (21.4\% reduction at day 3). \textit{Sox1} mRNA levels were similarly reduced (Supplementary Figure 2A). This decrease in neural progenitors could reflect a bias of differentiation against neural fate, or be due to a more general effect on the onset of differentiation irrespective of lineage. To test this differentiating 46C cells were assessed for expression of the ES cell marker Oct4. GNS-treated ES cells remained Oct4 positive when vehicle-treated control cells have largely lost Oct4 immunoreactivity and instead expressed Sox1GFP (Figure 2B). We obtained the same result using the transgenic \textit{Oct4}-GFP reporter ES cell line Oct4GiP (Figure 2C), suggesting a delay in differentiation onset in the presence of GNS. Consistently, we also observed a decrease in endoderm (BMP2) and mesoderm markers Brachyury and Eomesodermin following embryoid body (EB)
differentiation (Figure 2D-F). Interestingly, primitive endoderm differentiation measured by qRT-PCR for Sox17 and Gata6 appears to be unaffected by GNS (Supplementary Figure 2B,C).

ES cell self-renewal can be maintained through the action of the transcription factor Stat3, operating downstream of the cytokine LIF. However, we could not detect any change in Stat3 phosphorylation in GNS treated cells compared to controls (Figure 3A), which suggests that the effect of raised O-GlcNAc levels on ES cell differentiation is independent of Stat3 activity. Furthermore, GNS is not able to substitute for LIF and promote the clonal expansion of ES cells in serum (Supplementary Figure 3A), nor does it affect cloning efficiency in the presence of LIF (Figure 3B), consistent with a delay rather than a complete block in differentiation.

**O-GlcNAc levels do not affect transition of naïve cells to a primed state**

Recent work has suggested that O-GlcNAcylation of Oct4 can maintain ES cells by regulation of downstream naïve state markers like Nanog, Rex1, Klf2 and 5 and others. When cultured in the presence of serum and LIF, ES cell populations consist of mixed naïve and primed cells, but in serum-free N2B27 monolayer differentiation conditions the transcripts for the naïve markers decline sharply after ~24h and the cells proceed to differentiate. To test whether this early transition is affected by O-GlcNAc levels we employed the OCRG9 line, expressing GFP under the control of the endogenous Rex1 locus. Under control conditions, OCRG9 cells become GFP negative at day 2 of differentiation (the delay between loss of Rex1 mRNA and loss of GFP is due to the stability of the latter). GNS-treated OCRG9 cells lost expression of Rex1-GFP at the same rate as the control (DMSO-treated) cells (Figure 3C). The transition from naïve to primed state is mediated by the actions of the Erk1/2 kinases downstream of fibroblast growth factor (Fgf) signaling. Recent work in
Drosophila has identified a requirement for O-GlcNAcylation for Erk1/2 signal transduction downstream of Fgf receptor \(^3\). However, consistent with a normal transition to the primed state, we did not detect any changes in the profile of Erk1/2 phosphorylation in GNS treated cells (Supplementary Figure 3B), suggesting that the signal regulating transition of naïve cells to primed pluripotency is unaltered by increased O-GlcNAc levels.

This is further confirmed by quantitative RT-PCR for other naïve cell markers Nanog, Rex1/Zfp42, Esrrb and Dppa3 (Figure 3D, Supplementary Figure 3C) as well as the transient upregulation of the epiblast marker Fgf5 in EB differentiation (Figure 3E). Taken together these results indicate that O-GlcNAc levels do not affect the initial events of ES cell differentiation.

In order to test this hypothesis further we performed a microarray experiment to measure GNS’s effect on global transcriptional changes during the transition of naïve pluripotent cells to a primed state. We treated ES cells with DMSO or GNS for 24h either in ES cell media or in neural differentiation conditions and compared gene expression between treatments. Principal component analysis (PCA) showed that the change from self-renewal media to monolayer differentiation media (including serum withdrawal) causes a major effect on global gene expression (37%). This is reflected in the clustering of the most significantly regulated genes that change their expression from ES cells to day 1 of differentiation irrespective of GNS (Figure 3F). We then compared the changes in gene expression during the first 24h of differentiation in the presence of GNS or DMSO compared to undifferentiated ES cells. This revealed that a great majority of the genes are common to the two conditions (Figure 3G), consistent with the hypothesis that naïve-to-primed transition is unaffected by O-GlcNAc levels. Many of the genes downregulated (fold change >1.2, \(p<0.05\)) in both vehicle and GNS samples belong to pathways regulating ES cell self-renewal (i.e. MAPK, Jak/Stat and Tgfβ superfamily; see Table 1) and include Bmp4, Tgfβ1 and downstream targets Id1, Id2, Id3;
Stat3 and downstream target Socs3 as well as transcription factors associated with naïve pluripotency (Klf2, 3, 4 and 5, Nanog, Rex1/Zfp42). Common upregulated genes include neural and mesodermal early differentiation regulators (Otx2, Tbx4, Neurogenin3, NeuroD1, NeuroD4) as well as primed pluripotency epiblast marker Fgf5, consistent with our RT-qPCR results. In total, 45 genes from our set also belong to the Plurinet network of pluripotency-related markers 31, 20 of which are upregulated in ES cells and 25 are upregulated in the day 1 samples (Supplementary Tables 3 and 4 respectively). This result confirms that raised O-GlcNAc levels do not affect the expression of genes controlling the onset of differentiation in serum-free monolayer.

Increased protein O-GlcNAcylation does not interfere with phosphorylation of major signaling pathways in ES cells

O-GlcNAc is frequently attached to serine and threonine residues that can also be phosphorylated and for many proteins this reciprocal relationship (often described as a “Yin-Yang”) regulates their activity. To investigate whether increased O-GlcNAc levels affect the global phosphorylation of proteins in ES cells we compared the global phospho-serine and phospho-threonine levels of ES cell protein extracts treated with various concentrations of GNS. Unlike global O-GlcNAc levels that showed a dose-dependent increase, protein phosphorylation levels remained relatively constant, suggesting that the majority of ES cell phosphosites are not occupied by O-GlcNAc (Supplementary Figure 4A). This result, however, could reflect a high abundance of constitutively phosphorylated proteins in these cells that conceal significant changes in less abundant and more dynamically regulated phosphoproteins. We therefore focused our attention on protein kinases as these proteins are often dynamically regulated by phosphorylation and mediate numerous important cellular responses. We analyzed kinase phosphorylation in whole-cell lysates from ES cells treated for
24h with either DMSO or GNS using a protein kinase array. However, none of the 46 kinase sites profiled showed a significant change in their basal level of phosphorylation by GNS (Supplementary Figure 4B) indicating that the ES cell kinome is not significantly affected by the “Yin-Yang” interplay between O-GlcNAc and phosphate.

**Raised O-GlcNAc levels affect transcription of repressed genes in ES cells**

We then analyzed ES cells treated for 24h with DMSO or GNS by microarrays, using Ranked Product analysis – a method better at detecting changes in gene expression from small number of replicates than the more commonly used SAM or ANOVA methods. Using a Ranked Product p-value < 0.05 we identified 971 differentially regulated genes. Of these, 516 were up-regulated following GNS treatment and 455 were down-regulated. Two of the most significantly regulated genes are *Ogt* and *Oga* (Figure 4A). Using RT-qPCR we found that *Ogt* expression is significantly reduced within 1h of treatment with GNS, whereas *Oga* up-regulation appears slower (Figure 4B,C). The regulation of Oga and Ogt by O-GlcNAc levels is also reflected at the protein level within 24h of treatment with GNS (Supplementary Figure 4C) or by knockdown of Ogt or Oga using siRNA (Figure 4D).

One other gene that stood out from this analysis is Zscan4 (Figure 4A). This gene has previously been associated with telomere maintenance in ES cells and is associated with a subpopulation of cells similar to that of the recently described “2C” state of privileged developmental plasticity, existing within ES cell cultures. 2C cells differ in gene expression from ES cells in that they express genes associated with zygotic genome activation and have been demonstrated to be totipotent (giving rise to extraembryonic as well as embryonic tissues in chimeras). Transcripts marking this subpopulation include retrotransposons normally repressed by epigenetic mechanisms, as well as chimeric transcripts of genes with junctions to murine endogenous retrovirus with leucine tRNA primer.
(MERVL) elements. Interestingly, the number of genes upregulated in the GNS treated samples were much larger than the DMSO samples, both for the genes enriched in ES cells and for genes enriched in Day 1 differentiating cells. We therefore performed pairwise comparisons between the ES cell and Day 1 samples in DMSO or GNS treatment for the most regulated genes (p<0.05, fold change >2). The majority of genes with expression higher in ES cells than Day 1 (~80%; 348/437) were similarly regulated both in DMSO and GNS. However, of the remaining 20% (those that were not common to GNS and DMSO), nearly three times more genes were higher in the GNS sample than in DMSO (65 vs. 24; Figure 5A). Similarly, the majority of the genes expressed at higher level in Day 1 samples compared to ES cells are common to GNS and DMSO (~75%; 160/214), but of those differentially regulated between the treatments, those regulated by GNS outnumbered those regulated by DMSO by a factor of 3.5 (42 vs. 12; Figure 5A). This result suggests that GNS treatment causes a general increase in gene expression both in ES cells and early differentiating cells (see also Figure 3G). We then turned to Gene Set Enrichment Analysis (GSEA) for further mining of our expression data. Search of the whole MSigDB and GeneSigDB databases using GSEA did not reveal a significant enrichment for any geneset so we created our own gene sets from publicly available microarray and ChIP-seq data as published in relevant papers. We focused our attention to targets of the ES cell core pluripotency network and genes regulated by the Polycomb group. Although there was no significant enrichment for polycomb complex 1 or 2 targets in our dataset (Supplementary Table 6), a specific subset of polycomb target genes, designated as Polycomb-repressed (PRCR), is highly significantly enriched in ES cells treated with GNS compared to the vehicle control (Figure 5B, Table 2). Strikingly, we also detected highly significant enrichment in the GNS treated samples for the gene sets corresponding to the 2C population (the “2C” sets as described previously) as well as the gene set “2C-MERVL+” which comprises the genes from the 2C set that contain
MERVL chimeric transcripts. The enrichment of 2C transcripts is highly significant both in undifferentiated ES cells and in Day 1 differentiation samples treated with GNS compared to their respective vehicle controls (Figure 5C,D, Table 2, Supplementary Table 6 and Supplementary Table 7). We validated the upregulation of three “2C” associated genes (Zfp352, Tdpoz3 and Zscan4) by Oga knockdown and GNS in ES cells using quantitative RT-PCR and show a consistent upregulation for all three (Figure 5C, Supplementary Figure 5A-C). These results therefore indicate that increased O-GlcNAc levels result in increased expression of epigenetically repressed genes, including those characteristic of a totipotent “2C” state.

**DISCUSSION**

Our results demonstrate that ES cell differentiation is reduced under conditions of increased O-GlcNAc signaling and pinpoint the deficit to a stage in differentiation after the loss of naive pluripotency but before the definitive differentiation as marked by the loss of Oct4 expression. Using a range of differentiation protocols we found a delay in the onset of differentiation towards embryonic lineages. We report that Oga inhibition does not affect the up- or down-regulation of the genes known to be regulated early during differentiation, but results in prolonged expression of the pluripotency master regulator Oct4 and a delay in the acquisition of stable neural differentiation marker Sox1. Therefore, our results clearly demonstrate that Oga inhibition does not affect the transition from the naïve to the primed state, indicating instead that elevated O-GlcNAc levels affect differentiation progression at a later stage. Taken together, our data show a clear effect of Oga inhibition on the initiation of ES cell differentiation that affects multiple somatic lineages.

We also report that Oga inhibition does not interfere with the steady-state phosphorylation levels of several key kinases, suggesting that the main mechanism of action operates
downstream of the signal transduction machinery. The fact that the phosphorylation status of kinases is unaffected by GNS in this context does not, however, preclude changes in kinase activity, localization or interaction with substrates. It also remains possible that the stimulated, maximal phosphorylation levels of some of these kinases will be affected by elevated O-GlcNAc despite no obvious effects at the basal phosphorylation levels.

An early study of Oga inhibition during ES cell differentiation identified a deficit in spontaneous cardiomyocyte generation in embryoid bodies (EBs) \(^{12}\). In that study a different, less specific Oga inhibitor was added to 5 day-old differentiating EBs, when ES cell markers are already lost and the cells have already committed to differentiating, highlighting a role for Oga in the commitment of mesodermal cells to the cardiomyocyte lineage. Our data extend and refine this previous observation by determining an earlier role in the exit from the pluripotent state. More recently, a report demonstrated a role for the O-GlcNAc modification of Oct4 in the onset of ES cell differentiation \(^{11}\). Our results are consistent with those findings, although we find that naïve ES cell markers decline normally under elevated O-GlcNAc conditions, demonstrating that the earliest events in ES cell differentiation are not O-GlcNAc dependent.

A number of recent papers have linked the O-GlcNAc transferase to the regulation of epigenetic regulators \(^{16-18,42}\). Our gene expression analysis has revealed that following Oga inhibition a number of repressed genes become activated. Gene set enrichment analysis showed a significant enrichment for a specific subset of polycomb repressor complex target genes, as well as genes associated with the early zygotic genome activation and retrovirus-like elements. This latter gene expression signature is thought to mark a developmentally privileged ES cell subpopulation capable of differentiation into extraembryonic lineages \(^{35,36}\). These genes are under the control of epigenetic regulators of gene expression and their expression can be induced by interfering with histone acetylation and methylation, either
pharmacologically or by genetic deletion of key enzymes. Our findings therefore suggest that the levels of O-GlcNAc in cells influence global gene expression regulated by major epigenetic mechanisms such as histone modifications and DNA methylation, and are consistent with a differentiation delay observed in ES cells treated with an inhibitor of histone deacetylases. Intriguingly, Oga has been shown to possess histone acetyltransferase (HAT) activity in vitro and this activity is distinct from its GlcNAcase function. It is therefore likely that although inhibition of Oga’s O-GlcNAcase activity by GNS raises O-GlcNAc levels and leads to feedback Oga upregulation, it does not affect the protein’s HAT activity (as has been shown for the less specific inhibitor Streptozotocin). This could possibly lead to increased histone acetylation and the derepression of transcription we observe, although the observation that 2C genes were upregulated even when Oga protein levels are reduced by siRNA (Figure 5E) indicates that this is not the main mechanism in this case. The negative feedback regulation from increased O-GlcNAc levels to a reduction in Ogt protein levels may also contribute to the phenotypes observed, by disrupting the formation of some of its protein complexes. O-GlcNAc can itself modify histones and promote transcriptional activation, so the precise mechanism by which gene expression is regulated by O-GlcNAc levels is likely to be complex and involve many, possibly redundant molecular players.

CONCLUSIONS

In conclusion, our results demonstrate that ES cell differentiation is delayed under conditions of increased O-GlcNAc signaling and pinpoint the delay to a stage in differentiation after the loss of naive pluripotency but before commitment to differentiation as marked by the loss of Oct4 expression. We utilized defined media and short timescales to minimize the possibility of multiple effects on several interacting cell types, and found a measurable, significant defect in the progression towards differentiation. We attribute this differentiation defect to a
disruption of the normal function of global activation and repression complexes, such as the histone deacetylases, polycomb group and Ten-eleven translocation proteins as well as direct effects of Oga and Ogt on chromatin.

ACKNOWLEDGEMENTS

This work was funded by Tenovus Tayside small project grants (MPS). We would like to thank Ian Chambers for hLIF-05, Osama Albarbarawi for mass spectrometry, Pauline Gammack, Claudio Guerri, Claire Higgins and Panagiotis Stamatiadis for technical assistance and Arno Müller, Steve Keyse and Kate Storey for helpful comments on the manuscript.

REFERENCES

FIGURE LEGENDS

Figure 1 Global protein O-GlcNAcylation in ES cells

A) Whole-cell lysates of ES cells and cells at various stages of neural differentiation were blotted for O-GlcNAc levels. No major changes are detectable in the most abundant bands B) 2D gels of nuclear extracts blotted for O-GlcNAcylation. The overall level of O-GlcNAc is reduced in differentiation conditions whereas a cluster of basic proteins is prominent in the Day 1 sample. C) Dose-dependent increase in protein O-GlcNAcylation after 24h treatment with 1µM GlcNAcstatin C. D) Undifferentiated and 3-day neural differentiated ES cells in DMSO or GNS were immunoprecipitated with antibody against O-GlcNAc (RL2) and succinylated WGA beads or antibody against Sox2 then blotted for Sox2 or Sox2 GlcNAcylated at S248. GNS increases O-GlcNAc levels on Sox2, especially on day 3 E) Measurement of apoptotic cells (pink) after 24h of vehicle (DMSO) or 1µM GlcNAcstatin C treatment shows no difference in apoptosis levels. F) ES cell proliferation in vehicle (DMSO), the Oga inhibitors Thiamet G and GlcNAcstatin C or its inactive stereoisomer, GalNacstatin C (all at 1µM).

Figure 2 GlcNAcstatin treatment impairs ES cell differentiation

A) The proportion of Sox1GFP expressing neural progenitors is decreased following GlcNAcstatin C (GNS) treatment during differentiation of 46C cells. B) 46C cells retain expression of Oct4 when differentiated in the presence of GNS while control cells differentiate into neural progenitors. Scale bar: 50µm. C) Oct4GiP cells retain Oct4-GFP expression at higher levels in differentiation conditions compared to vehicle only controls. D-F) ES cells were differentiated as embryoid bodies in the presence of DMSO or GNS for 3 or 5 days then assayed for expression of endodermal marker BMP2 (D) or mesodermal markers Brachyury (E) or Eomesodermin (F) by RT-qPCR. Mean +/- s.e.m. from n=3 experiments
**Figure 3 Increased O-GlcNAcylation does not affect naïve-primed ES cell transition**

A) 46C ES cells were cultured in the absence of LIF overnight and then stimulated with either LIF or the LIF receptor antagonist hLIF-05 in the presence of either DMSO or GlcNAcstatin C, then blotted for phospho-Stat3 (Y705). GlcNAcstatin C treatment does not interfere with the ability of LIF to stimulate Stat3 phosphorylation and does not cause Stat3 phosphorylation in the absence of LIF. B) GNS or Oga knockdown does not affect the ability of ES cells to form undifferentiated colonies, while inhibition or knockdown of Ogt results in cell death. 4Ac: 4-Acetyl-5S-GlcNAc (Ogt inhibitor). C) OCRG9 ES cells (Rex1EGFP reporter line) were differentiated in N2B27 for 2 days in vehicle or GlcNAcstatin C and analysed by flow cytometry for EGFP fluorescence. GlcNAcstatin treatment does not affect the loss of naïve marker Rex1. D) Quantitative RT-PCR analysis during monolayer differentiation shows that expression of the naïve marker Nanog is lost at equivalent rates in the presence of vehicle (DMSO; black bars) or GlcNAcstatin C (GNS; white bars). E) Upregulation of the primed pluripotency marker Fgf5 is unaffected by GNS treatment (D-E n=3, mean +/-s.e.m.) F) Hierarchical clustering (using Euclidean distance with complete linkage agglomeration method) of samples based on the top 31 differentially expressed genes (ranked by p-value) separates the samples according to the stage of differentiation and shows clustering of genes that contribute to differentiation. G) Pairwise comparison of genes regulated during the first 24h of differentiation in vehicle (DMSO) or GlcNAcstatin C (GNS) (fold change >1.2, p<0.05).

**Figure 4 O-GlcNAc levels control Ogt and Oga levels in ES cells**

A) Ranked Product analysis of genes regulated in ES cells by GNS treatment. Blue: significantly upregulated genes; yellow: significantly downregulated genes. B) Validation of
Ogt regulation by RT-qPCR (n=4, mean +/- s.e.m.) C) Validation of Oga expression by RT-qPCR (n=4, mean +/- s.e.m.). D) RNAi knockdown of Ogt and Oga results in a negative feedback regulation on the expression of Oga and Ogt respectively. * p<0.05, ** p<0.01 compared to t=0 (T-test).

Figure 5 Oga inhibition results in a broad genomic transcriptional de-repression

A) Bar charts showing the number of unique genes upregulated in ES cells or in Day 1 differentiation in the DMSO or GNS samples (p<0.05, fold change>2) B) Gene set enrichment analysis (GSEA) shows that GNS treatment results in an upregulation of the “Polycomb repressed” gene set. C) GSEA showing enrichment of the “2C” gene set in the GNS treated ES cells. D) The “2C” gene set is also enriched in the GNS treated Day 1 differentiating cells. E) RT-qPCR validation of the upregulation of 3 genes from the 2C cohort by knockdown of Oga (n=3, mean +/- s.e.m.) * p<0.05.
<table>
<thead>
<tr>
<th>Regulation</th>
<th>Term</th>
<th>Count</th>
<th>%</th>
<th>P-Value</th>
<th>Benjamini</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up in ES</td>
<td>Focal adhesion</td>
<td>57</td>
<td>3.00E-01</td>
<td>2.30E-12</td>
<td>3.90E-10</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Lysosome</td>
<td>39</td>
<td>2.00E-01</td>
<td>1.90E-10</td>
<td>1.60E-08</td>
</tr>
<tr>
<td>Up in ES</td>
<td>ECM-receptor interaction</td>
<td>28</td>
<td>2.00E-01</td>
<td>5.70E-08</td>
<td>3.10E-06</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Regulation of actin cytoskeleton</td>
<td>45</td>
<td>2.00E-01</td>
<td>1.70E-05</td>
<td>7.20E-04</td>
</tr>
<tr>
<td>Up in ES</td>
<td>TGF-beta signaling pathway</td>
<td>24</td>
<td>1.00E-01</td>
<td>2.80E-05</td>
<td>9.20E-04</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Pathways in cancer</td>
<td>59</td>
<td>3.00E-01</td>
<td>3.90E-05</td>
<td>1.10E-03</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Tight junction</td>
<td>31</td>
<td>2.00E-01</td>
<td>6.70E-05</td>
<td>1.60E-03</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Leukocyte transendothelial migration</td>
<td>28</td>
<td>2.00E-01</td>
<td>1.10E-04</td>
<td>2.20E-03</td>
</tr>
<tr>
<td>Up in ES</td>
<td>MAPK signaling pathway</td>
<td>49</td>
<td>3.00E-01</td>
<td>1.50E-04</td>
<td>2.70E-03</td>
</tr>
<tr>
<td></td>
<td>Arrhythmogenic right ventricular cardiomyopathy (ARVC)</td>
<td>20</td>
<td>1.00E-01</td>
<td>2.50E-04</td>
<td>4.20E-03</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Small cell lung cancer</td>
<td>21</td>
<td>1.00E-01</td>
<td>5.00E-04</td>
<td>7.50E-03</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Other glycan degradation</td>
<td>8</td>
<td>0.00E+00</td>
<td>7.50E-04</td>
<td>1.00E-02</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Glycosphingolipid biosynthesis</td>
<td>7</td>
<td>0.00E+00</td>
<td>2.10E-03</td>
<td>2.60E-02</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Endocytosis</td>
<td>36</td>
<td>2.00E-01</td>
<td>2.60E-03</td>
<td>3.00E-02</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Jak-STAT signaling pathway</td>
<td>29</td>
<td>2.00E-01</td>
<td>2.70E-03</td>
<td>3.00E-02</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Hypertrophic cardiomyopathy (HCM)</td>
<td>19</td>
<td>1.00E-01</td>
<td>2.90E-03</td>
<td>3.00E-02</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Glycosphingolipid biosynthesis</td>
<td>7</td>
<td>0.00E+00</td>
<td>3.20E-03</td>
<td>3.10E-02</td>
</tr>
<tr>
<td>Up in ES</td>
<td>ErbB signaling pathway</td>
<td>19</td>
<td>1.00E-01</td>
<td>4.30E-03</td>
<td>3.90E-02</td>
</tr>
<tr>
<td></td>
<td>Amino sugar and nucleotide sugar metabolism</td>
<td>12</td>
<td>1.00E-01</td>
<td>5.50E-03</td>
<td>4.70E-02</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Galactose metabolism</td>
<td>9</td>
<td>0.00E+00</td>
<td>5.80E-03</td>
<td>4.70E-02</td>
</tr>
<tr>
<td>Up in ES</td>
<td>Dorso-ventral axis formation</td>
<td>8</td>
<td>0.00E+00</td>
<td>6.30E-03</td>
<td>4.90E-02</td>
</tr>
<tr>
<td>Up in Day 1</td>
<td>Steroid biosynthesis</td>
<td>11</td>
<td>0.1</td>
<td>1.90E-08</td>
<td>3.40E-06</td>
</tr>
<tr>
<td>Up in Day 1</td>
<td>Terpenoid backbone biosynthesis</td>
<td>8</td>
<td>0.1</td>
<td>1.30E-05</td>
<td>1.10E-03</td>
</tr>
<tr>
<td>Up in Day 1</td>
<td>Biosynthesis of unsaturated fatty acids</td>
<td>9</td>
<td>0.1</td>
<td>2.60E-04</td>
<td>1.50E-02</td>
</tr>
<tr>
<td>Up in Day 1</td>
<td>Pathways in cancer</td>
<td>38</td>
<td>0.3</td>
<td>6.90E-04</td>
<td>3.00E-02</td>
</tr>
</tbody>
</table>
Table 2

Gene sets significantly enriched in GNS or DMSO treated ES cell or Day 1 differentiation samples.

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Enrichment Score</th>
<th>Normalized Enrichment Score</th>
<th>Nominal p-value</th>
<th>False Discovery Rate q-value</th>
<th>Enrichment in</th>
</tr>
</thead>
<tbody>
<tr>
<td>2C-MERVL+</td>
<td>45</td>
<td>0.85</td>
<td>1.85</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>ES GNS vs. ES DMSO</td>
</tr>
<tr>
<td>2C</td>
<td>466</td>
<td>0.66</td>
<td>1.78</td>
<td>&lt;0.001</td>
<td>0.001</td>
<td>ES GNS vs. ES DMSO</td>
</tr>
<tr>
<td>PRCR</td>
<td>1331</td>
<td>0.47</td>
<td>1.31</td>
<td>&lt;0.001</td>
<td>0.038</td>
<td>ES GNS vs. ES DMSO</td>
</tr>
<tr>
<td>2C</td>
<td>466</td>
<td>0.39</td>
<td>1.25</td>
<td>0.006</td>
<td>0.033</td>
<td>Day 1 GNS vs. Day 1 DMSO</td>
</tr>
</tbody>
</table>
**Stavridis Figure 2**

**A**

![Graph A: Sox1-EGFP expression](image)

**B**

Day 4 DMSO vs Day 4 GNS DAPI/Oct4/Sox1GFP images

**C**

![Graph C: Oct4-EGFP expression](image)

**D**

![Graph D: BMP2 mRNA expression](image)

**E**

![Graph E: Brachyury mRNA expression](image)

**F**

![Graph F: Eomesodermin mRNA expression](image)
Stavridis Figure 4
Stavridis Figure 5