O-GlcNAc Regulates Erythroid Gene Transcription

By

Zhen Zhang
Master of Science, The Fourth Military Medical University, China, 2008
Bachelor of Forensic Medicine, Shanxi Medical University, China, 2005

Submitted to the graduate degree program in Biochemistry and Molecular Biology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Co-Chair: Dr. Chad Slawson

Co-Chair: Dr. Kenneth R. Peterson

Dr. Joan Conaway

Dr. Joseph Fontes

Dr. Soument Paul

Dr. Luciano DiTacchio

Date Defended: April 12, 2017

The dissertation committee for Zhen Zhang certifies that this is the approved version of the following dissertation:

O-GlcNAc Regulates Erythroid Gene Transcription

Co-Chair: Dr. Chad Slawson

Co-Chair: Dr. Kenneth R. Peterson

Date Approved: April 12, 2017

Abstract

O-GlcNAc is a post-translational modification on serine or threonine residues by β-N-acetylglucosamine. O-GlcNAc is added and removed from serine and threonine residues by the O-GlcNAc processing enzymes, O-GlcNAc-transferase (OGT) and O-GlcNAcase (OGA), respectively. The levels of O-GlcNAc can rapidly change in response to fluctuations in the extracellular environment and return to a baseline level quickly after stimulus removal. This process termed O-GlcNAc homeostasis is critical to the regulation of many cellular functions. However, the relationship between O-GlcNAc homeostasis, OGT, OGA, and gene transcription is unknown. In this dissertation, I seek to address some of the fundamental mechanisms into the control of transcription by O-GlcNAcylation. First, I explored how changes in O-GlcNAc homeostasis affect the transcription of OGT and OGA. We treated several human cell lines with Thiamet-G (TMG, an OGA inhibitor) to increase overall O-GlcNAc levels resulting in decreased OGT and increased OGA protein expression. OGT transcription level slightly declined, but OGA significantly increased with TMG treatment. Pretreating cells with protein translation inhibitor cycloheximide did not stabilize OGT or OGA protein in the presence of TMG; nor did TMG stabilize OGT and OGA transcription when cells were treated with RNA transcription inhibitor actinomycin D. RNA Pol II chromatin immunoprecipitation at the OGA promoter showed that RNA Pol II occupancy at the transcription start site was lower after prolonged TMG treatment. Together, these data suggest that OGA transcription was sensitive to changes in O-GlcNAc homeostasis and was potentially regulated by O-GlcNAc.

Next, we investigated how O-GlcNAcylation regulates $^{A}\gamma$-globin transcription. Erythropoiesis is the process of generating erythrocytes from erythroid progenitor cells. During this process, numerous genes are up-regulated or down-regulated. Transcription factor GATA-1

is the master regulator for erythropoiesis. However, the mechanism underling how GATA-1 regulate gene activation or repression is not well understood. Here, we utilized different cell and animal models with OGA inhibitor TMG to address this question. We first studied the $^{A}\gamma$-globin gene repression at adult stage mediated by GATA-1/FOG-1/Mi2β repressor complex at the -566 GATA site at the $^{A}\gamma$-globin gene promoter. We demonstrated that OGT and OGA interact with the $^{A}\gamma$-globin promoter at the -566 GATA repressor site; however, mutation of the GATA site to GAGA significantly reduced OGT and OGA promoter interactions in β-YAC bone marrow cells (BMCs). When WT β-YAC BMCs are treated with an OGA inhibitor Thiamet-G (TMG), the occupancy of OGT, OGA, and Mi2β at the $^{A}\gamma$-globin promoter was increased. In addition, OGT and Mi2β recruitment was increased at the $^{A}\gamma$-globin promoter when γ-globin becomes repressed in post-conception day E18 human β-YAC transgenic mouse fetal liver. Furthermore, we showed that Mi2β is modified with O-GlcNAc and both OGT and OGA interacts with Mi2β, GATA-1, and FOG-1. Taken together, our data suggested that O-GlcNAcylation is a novel mechanism of γ-globin gene regulation, mediated by modulating the assembly of the GATA-1/FOG-1/Mi2β repressor complex at the -566 GATA motif within the promoter.

Then, we asked if O-GlcNAc regulates GATA-1 target gene transcription during erythropoiesis. In this study, we utilized a well-established cell model of erythropoiesis, G1E-ER4 cells, a murine GATA-1 null erythroblast line that undergoes erythroid differentiation when GATA-1 activity is restored by the addition of β-estradiol (E2). Interestingly, overall O-GlcNAc levels dramatically decreased after GATA-1 activation in G1E-ER4 cells, with a slight increase in OGA and a decrease in OGT protein levels. GATA-1 interacts with OGT/OGA, and this interaction increased during erythropoiesis. Transcriptome analysis of G1E-ER4 cells treated with E2 and Thiamet-G (TMG, an OGA inhibitor) revealed that 1,173 genes changed expression

patterns compared to E2 treatment only, including 433 GATA-1 target genes. These data suggest a subset of GATA-1 target genes are regulated through O-GlcNAcylation. Next, we demonstrated that the occupancy of GATA-1 and OGT/OGA, and the overall O-GlcNAc level at *Laptm5* GATA binding site decreased when OGA was inhibited by TMG during erythropoiesis. Our data suggests that O-GlcNAcylation plays a role in regulating a subset of GATA-1 targeted erythroid genes and suggest O-GlcNAcylation as a mechanism regulating GATA-1 function at specific GATA-1 targeted genes.

Acknowledgments

I would like to thank my mentor Dr. Chad Slawson for his advice and support during my graduate training. I would also like to thank my collaborator Dr. Kenneth R. Peterson and colleagues in his lab for their support on these projects. Thank my committee members Drs. Joan Conaway, Joseph Fontes, Luciano DiTacchio, and Soumen Paul for their kindly suggestion and guidance to my projects.

I would like to thank Ee Phie Tan and Miranda Machacek in Chad Slawson Lab for technical support and helpful discussion on the projects, Dr. Zhuan Li in Steven Weinman Lab and Diana Kalinowska in Luciano DiTacchio Lab for ChIP technical support, Dr. Devin Koestler and Stefan Graw in Department of Biostatistics for processing the RNA-Seq data, Genome Sequencing Facility for RNA-sequencing, Drs. Antonio Artigues and Maria T. Villar in Mass Spectrometry/Proteomic Core Laboratory for identifying the Mi2β O-GlcNAc site, and all the students and faculty in Department of Biochemistry and Molecular Biology for providing a good research environment.

Thank KUMC Biomedical Research Training Program, National Institute of Diabetes and Digestive and Kidney Diseases R01DK100595 to C. Slawson and K. R. Peterson, and National Institute of General Medical Sciences (P20GM103549 & P30GM118247) of the National Institutes of Healtfor funding support.

Thank Dr. Gerald Hart from Department of Biological Chemistry at the Johns Hopkins University for providing O-GlcNAc antibodies.

At last, I would like thank my parents and my wife for their yearly support. Nothing will be done without them.

Table of Contents

List of Figures

Chapter 1: Introduction

1.1 Chromatin and Transcription Initiation

Each human diploid cell contains around 2 meters of chromosomal DNA, which is compacted in a tiny nucleus. How is the long DNA compacted into the tiny space? It is because of the help of histones, which are positive-charged proteins and bind tightly to negative-charged DNA. This DNA-histone complex is referred to as chromatin. A nucleosome is the basic structural and functional unit of the chromatin. The nucleosomes fold, loop, and coil, to form the chromosome (1). Each nucleosome has the following structure: two copies of histone H2A, H2B, H3, and H4 form an octamer, which is wrapped by approximately 1.7 turns of DNA (146 base pairs) (2). The addition of one histone H1 and another 20 base pairs of DNA link the adjacent nucleosome. The N-terminus of histones project out of the nucleosome core and form histone tails. Many sites on the histone tail can be targeted for post-translational modification (PTM), such as acetylation, phosphorylation and methylation, which are referred to as the histone code. However, histone PTMs have also been found in the globular domain of core histones (3).

The modifications on histones are epigenetic changes. Epigenetics refers to functionally relevant changes on chromatin or RNA without involving changes in the nucleotide sequence. Epigenetic regulation is dynamically regulated by epigenetic writers, readers, and erasers. Epigenetic writers such as histone acetyltransferases (HATs) add the modification to the targets; epigenetic readers such as proteins containing bromodomains bind to these epigenetic marks; and epigenetic erasers such as histone deacetylases (HDACs) remove the epigenetic marks. The integration of signaling at the epigenetic level can change many cellular processes including gene transcription (4-6).

Eukaryotic transcription is an elaborate process that eukaryotic cells use genomic DNA as template to syntheses pre-mRNA and later form mature mRNA. Transcription includes three different stages: initiation, elongation, and termination. Transcription of pre-mRNA is catalyzed by RNA polymerase II (Pol II), which is conserved in eukaryotes and contains 12 subunits. The largest subunit in RNA Pol II RPB1 has a CTD (carboxyl-terminal domain) consisting of multiple repeats of a heptamer YSPTSPS. The second and fifth Serine of the heptamer can be phosphorylated during transcription elongation (7).

Transcription initiation by Pol II requires 5 general transcription factors (GTFs), which are termed as TFIID (TATA binding protein [TBP]), TFIIB, TFIIF, TFIIE, and TFIIH. In the first initiation step, TFIID recognizes and binds to a specific DNA sequence (TATA box) in the promoter. Then, TFIIB binds to both Pol II and TFIID to facilitate binding of Pol II to TFIID. In the third step, TFIIF binds and stabilizes the complex by recruiting TFIIE and TFIIH to the complex. At this point, a fully assembled pre-initiation complex (PIC) forms (8,9).

Following formation of PIC, TFIIH DNA helicase subunit unwinds the DNA template at the transcription start site (TSS) in an ATP-dependent manner and creates an open complex. Recruitment of TFIIH also results in the phosphorylation of Pol II at Serine 5 of the CTD heptamer repeats. Then the first phosphodiester bond of nascent transcripts is synthesized by Pol II, and transcription initiation occurs. However, the Pol II pauses and stops transcribing after producing a short transcript. Subsequently, pTEF-b kinase is recruited and phosphorylates CTD heptamer repeats at Serine 2 allowing transcription elongation (8,9).

1.2 What is O-GlcNAc

O-linked N-acetylglucosamine (O-GlcNAc) is a post-translational modification (PTM) first discovered by Gerald W. Hart and Carmen-Rosa Torres in 1984 (10). O-GlcNAc is a

reversible modification on serine or threonine residues by β-N-acetylglucosamine. O-GlcNAc is present in all higher eukaryotes studied including plants. Unlike other glycosylation, O-GlcNAc is not elongated to more complex structures and localized to the cytoplasm, nucleus, and mitochondria. Since Ser/Thr residues are also modified by phosphorylation, O-GlcNAc is often reciprocal with phosphorylation.

O-GlcNAc transferase (OGT) is the enzyme that adds the O-GlcNAc modification, whereas O-GlcNAcase (OGA) removes it (11). The removal and addition of O-GlcNAc termed O-GlcNAc cycling is highly dynamic. Uridine diphosphate-N-acetyl-glucosamine (UDP-GlcNAc), the end-point of the hexosamine biosynthetic pathway (HBP), is the high-energy donor substrate for OGT. Since the HBP is heavily influenced by various metabolic events including the metabolism of glucose, amino and fatty acids, and nucleotides, O-GlcNAcylation is sensitive to nutrient availability (11). O-GlcNAc regulates many cellular processes, for example, cellular stress response, protein translation, signal transduction, and transcription. An imbalance in the homeostasis of O-GlcNAc does contribute to the development of diseases including cancer, diabetes, and Alzheimer's (12-14).

Human OGT gene is localized near centromere on X-chromosome. N-terminus of OGT contains 11.5 tetratricopeptide repeat (TPR) domains and C-terminus contains unique catalytic domain and PIP3 (phosphatidylinositol 3,4,5-trisphosphate) binding domain. OGT is found in all tissue and highly expressed in brain. There is no known consensus site for OGT, several proteins target OGT to cellular substrates, for example, OIP106, p38MAPK (11); Human OGA (MGEA5) is mapped to chromosome 10 and localized mostly in the cytoplasm. The N-terminus of OGA contains unique hyaluronidase domain and C-terminus contains putative GCN5 like histone-acetyltransferase. OGA can be cleaved by caspase-3 upon apoptosis but still maintains

activity. Both OGT and OGA are conserved across species, and knockout of OGT or OGA is lethal in dividing cells (11).

The transcriptional regulation of OGT and OGA is unclear. It is suggested that the substantial length of introns can affect gene transcription in Drosophila (15), hence the transcription of OGT may also be affected since it has one or more long introns. Another study also proposed that PRC2 (Polycomb Repressor Complex 2) is required for maintaining normal Ogt protein and O-GlcNAc levels (16). However, the mechanism underlying how OGT and OGA are transcriptionally regulated is unknown.

1.3 O-GlcNAcylation Regulates Transcription

In 1989, Kelly and Hart reported that Drosophila polytene chromosomes are enriched with O-GlcNAc, with condensed chromatin brightly stained and transcriptionally active "puff" regions less stained with FITC-WGA (a lectin binds to O-GlcNAc). Furthermore, they found that O-GlcNAc moieties are presented on many chromatin proteins associated with other nuclear components (17). Later, Sakabe, *et al*. demonstrated that histones are modified by O-GlcNAc (18), and some of the O-GlcNAc sites of H2A, H2B, and H4 were identified by mass spectrometry (18,19). The O-GlcNAc levels of histones increase during heat shock (18) and H3 O-GlcNAc level decreases during mitosis (20), suggesting that histone O-GlcNAc level changes in response to different cellular condition. O-GlcNAcylation of H2B at S112 recruits H2B ubiquitin ligase (BRE1A/1B) to promote H2B K120 monoubiquitination, which is a histone marker of transcription activation (21).

RNA Pol II CTD is modified by O-GlcNAc at Thr 4 and Ser 5 (22), and O-GlcNAc and O-Phosphate of CTD are mutually exclusive (23). OGT interacts with Pol II and GTFs, such as TFIIA, TFIID, and TFIIF. The O-GlcNAc cycling on the CTD at gene promoter is important for

gene transcription (24). In this cycle, unphosphorylated RNA Pol II (Pol IIA) CTD is first modified with O-GlcNAc (Pol IIγ) by OGT, promoting the association of the general transcription factors with Pol IIγ. Next, O-GlcNAc is removed from Pol IIγ by OGA converting Pol IIγ back to Pol IIA. Subsequently, Pol IIA is phosphorylated (Pol IIO) by TFIIH and p-TEFb and transcription elongation occurs (24).

OGT interacts with numerous transcription factors and mediates gene activation and silencing (11,19,25,26). For example, OGT interacts with mSin3A recruiting OGT to promoters to repress transcription (27). Furthermore, inactivation of transcription factors and RNA polymerase II by O-GlcNAc modifications contribute to gene repression (27); *Drosophila* polycomb group (PcG) gene (*sxc*) encodes for OGT, and null mutations in OGT lead to polycomb defects suggesting OGT is critical for PcG medicated gene silencing (28); Tet2 interacts with OGT and recruits OGT to O-GlcNAcylated H2B S112 during gene transcription (29). In summary, O-GlcNAc can regulate gene transcription not only by modulating RNA Pol II and histones, but also by modifying transcription regulators such as co-activators or repressors through various mechanisms.

1.4 Transcription Regulation of Erythropoiesis

Erythropoiesis is the process in which erythroid progenitors proliferate and differentiate into reticulocytes. This process is regulated by numerous transcription factors. At adult stage, erythropoiesis is characterized by fetal γ-globin repression and activation of master transcription factor GATA-1.

There are several mechanisms of fetal γ-globin silencing at adult stage. One of the well-known γ-globin silencers is B-cell lymphoma-leukemia A (BCL11A), which forms a protein complex with nucleosome remodeling and deacetylase complex (NuRD), GATA-1 and SOX6 to

repress γ-globin expression (30). Another one is Krüppel-like factor 1 (KLF1), required for γ-globin gene activation, which also stimulates BCL11A expression to represses γ-globin expression (30-32). Both BCL11A and KLF1 are positively regulated by binding of Mi2β (CHD4, Chromodomain Helicase DNA Binding Protein 4) a component of the NuRD complex (33). We previously demonstrated that ^{A}γ-globin gene is silenced during development following sequential binding of GATA-1, FOG-1 (Friend of GATA-1), and Mi2β proteins at the -566 GATA site relative to the ^{A}γ-globin gene CAP site in the fetal liver, between embryonic day E16-18 (34,35). GATA-1 is the DNA binding moiety in this complex, but GATA-1 can recruit both co-activators and co-repressors in both FOG-1-dependent and independent pathways (36-39). However, the determinants of repressor versus activator complex assembly during development are not well understood.

Another characteristic of erythroid differentiation is the activation of master transcription factor GATA-1, resulting in up- or down-regulation of its target gene transcription. What are the mechanisms underlying how GATA-1 mediates gene transcription activation versus repression? In some instances, co-localization of GATA-1 with Scl/TAL1 at chromatin sites assures the recruitment of activator or repressor (40-42); to a limited degree the WGATAR adjacent motifs (40,43,44) and histone modifications (45,46) also play a role. However, these mechanisms do not take account of the protein PTMs, which can affect protein-protein interaction and chromatin structure.

In this dissertation, we postulate that O-GlcNAcylation plays a role in modulating GATA-1/FOG-1/Mi2β repressor complex activity and regulating target gene transcription upon activating of GATA-1.

Chapter 2: O-GlcNAcase Expression is Critical in Maintaining O-GlcNAc Cellular Homeostasis

"This research was originally published in Frontiers in Endocrinology. Zhen Zhang, Ee Phie Tan, Nicole J. VandenHull, Kenneth R. Peterson, and Chad Slawson. O-GlcNAcase expression is sensitive to changes in O-GlcNAc homeostasis. *Front. Endocrinol*. 2014, 5:206. doi: 10.3389/fendo.2014.00206. © 2014 Zhang, Tan, VandenHull, Peterson and Slawson."

O-linked N-acetylglucosamine (O-GlcNAc) is a post-translational modification (PTM) first discovered by Gerald W. Hart and Carmen-Rosa Torres in 1984 (10). They initially used bovine milk galactosyltransferase (GalT1) to probe for terminal N-acetylglucosamine glycoconjugates on T-cells and unexpectedly discovered the existence of single β-N-acetylglucosamine conjugated proteins inside the cell (10). O-GlcNAc is a reversible modification that is ubiquitiously expressed in higher eukaryotes. O-GlcNAc transferase (OGT) is the enzyme that adds the O-GlcNAc modification, whereas O-GlcNAcase (OGA) removes it (47,48). Because uridine diphosphate-N-acetyl-glucosamine (UDP-GlcNAc), the end point of the hexosamine biosynthetic pathway, is the high-energy donor substrate for OGT, O-GlcNAcylation is sensitive to nutrient availability (11). Furthermore, the removal and addition of O-GlcNAc termed O-GlcNAc cycling is highly dynamic. Changes in hormones, nutrients, or the environment cause within minutes to several hours changes to the total level of O-GlcNAc on proteins (49-51). Importantly, O-GlcNAc cycling rates affect transcription regulatory pathways, cell cycle progression, and respiration (17,24,52-54).

Since O-GlcNAcylation plays a significant role in regulating a wide panel of cellular processes, and aberrant O-GlcNAcylation contributes to the development of diseases, understanding the regulation of OGT and OGA is, therefore, important. Several studies report

that the expression of OGT and OGA is sensitive to changes in total cellular O-GlcNAc levels (55,56). Elevation of O-GlcNAc levels via pharmacological inhibition of OGA causes OGT protein expression to decrease and OGA protein expression to increase (55). A rapid decrease in OGA protein expression occurs in mice embryonic fibroblasts when OGT is knocked out (56). Cells appear to actively keep a specific level of O-GlcNAcylation suggesting a certain homeostatic level of O-GlcNAc must be maintained for optimal cellular function. Although alterations of cellular O-GlcNAc levels affect OGT and OGA expression, the exact mechanism as to explain this phenomenon is still unclear. An imbalance in the homeostasis of O-GlcNAc does contribute to the development of diseases including cancer, diabetes, and Alzheimer's (12-14,57).

To further address how cells adjust OGT and OGA protein expression in response to alterations in O-GlcNAc levels, we measured in different cell lines OGT and OGA protein and mRNA expression and stability after pharmacologically inhibition of OGA by Thiamet-G (TMG, an OGA inhibitor). In these experiments, we were able to show that the OGA mRNA levels were more sensitive compared to OGT to alterations in O-GlcNAc, and RNA Pol II occupancy at the OGA transcription start site (TSS) was decreased after prolonged TMG treatments. Altogether, our data show that the protein expression of OGT and OGA is sensitive to changes in O-GlcNAc levels, and OGA transcription is sensitive to alterations in O-GlcNAc homeostasis.

2.1 Material and Methods

Antibodies and Reagents

All primary and secondary antibodies used for immunoblotting were used at a 1:1,000 and 1:10,000 dilution, respectively. Anti-O-linked *N*-acetylglucosamine antibody [RL2] (ab2739) was purchased from Abcam. Antibodies for OGT (AL-34) and OGA (345) were

gracious gifts from the Laboratory of Gerald Hart in the Department of Biological Chemistry at the Johns Hopkins University School of Medicine. Actin (A2066) antibody and anti-chicken IgY HRP (A9046) were purchased from Sigma. Chromatin immunoprecipitation (ChIP) grade mouse (G3A1) mAb IgG1 isotype control (5415) and RNA polymerase II antibody, clone CTD4H8 (05-623) were purchased from Cell Signaling Technologies and Millipore, respectively. Anti-rabbit HRP (170-6515) and anti-mouse HRP (170-6516) were purchased from Bio-Rad.

All reagents were purchased form Sigma unless otherwise noted. Cycloheximide (CHX, C7698, Sigma) was used at 50 μg/ml for HeLa cells and 25 μg/ml for K562 cells (58,59). Actinomycin D (AMD, A1410, Sigma) was used at 0.5 μg/ml for HeLa cells and 5 μg/ml for K562 cells (59,60).

Cell Culture

HeLa cells and SH-SY5Y neuroblastoma cells were cultured in DMEM (Invitrogen) supplemented with 10% fetal bovine serum (FBS, Gemini) and 1% penicillin/streptomycin (Invitrogen). K562 cells were cultured in RPMI-1640 (Invitrogen) supplemented with 10% fetal bovine serum, 1× MEM non-essential amino acids solution (Invitrogen), 1 mM sodium pyruvate (Invitrogen), 2.5 mM HEPES, and 1% penicillin/streptomycin. Cells were treated with 10 μM Thiamet-G (TMG, S.D. Specialty Chemicals) for 6, 8, 24, or 48 h with fresh TMG supplied daily. Cells were also pretreated with CHX for 4 h, followed by TMG treatment for 8 h or AMD for 0.5 h, followed by TMG treatment for 6 h. Cells were infected with OGT, OGA, or green fluorescent protein (GFP) virus at a multiplicity of infection (MOI) of 75 for 24 h. After different treatments, cells were harvested for western blot, quantitative PCR (qPCR), or ChIP assay.

Immunoblotting

Cells were lysed on ice for 30 min in Non-idet P-40 (NP-40) Lysis Buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 40 mM GlcNAc, and 1% Nonidet P-40) with 1 mM PMSF, 1 mM sodium fluoride, 1 mM β-glycerol phosphate, and 1× protease inhibitor cocktail I (leupeptin 1 mg/ml, antipain 1 mg/ml, benzamidine 10 mg/ml, and 0.1% aprotinin). Cell lysates were mixed with 4× protein solubility mixture (100 mM Tris-HCl, pH 6.8, 10 mM EDTA, 8% SDS, 50% sucrose, 5% β-mercaptoethanol, 0.08% Pyronin-Y). All electrophoresis was performed with 4–15% gradient polyacrylamide gels (Criterion Gels, Bio-Rad) and separated at 120 V, followed by transfer to PVDF membrane (Immobilon, Millipore) at 0.4 A. Blots were blocked by 3% BSA in TBST (25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 0.05% Tween-20) at room temperature for 20 min, then incubated with primary antibody at 4°C overnight. After washing with TBST for 5 × 5 min, blots were incubated with HRP-conjugated secondary for 1 h at room temperature, then washed with TBST again and developed using chemiluminescent substrate (HyGlo E2400; Denville Scientific). Blots were stripped in 200 mM glycine, pH 2.5 at room temperature for 1 h and probed with different primary antibodies. All immunoblotting results were repeated in three independent experiments (53). OGA and OGT relative protein levels were measured by analyzing the bands density using ImageJ 1.48 (http://imagej.nih.gov/ij/download.html) then normalized to the density of actin. All experiments were repeated three times, and average relative fold changes were calculated.

Total RNA Isolation and RT-PCR

Total RNA was isolated by TRI reagent solution (AM9738, Ambion) according to manufacture's instruction. Briefly, 2×10^6 cells were resuspended by 1 mL TRI reagent solution. Then, 200 µl of chloroform was added to extract RNA. After spinning down, upper phase containing total RNA was collected and incubated with equal amount isopropanol. RNA pellets

were then precipitated by centrifugation, washed once with 70% ethanol, air-dried, and dissolved in nuclease free water (Invitrogen).

RNA concentration was measured by Nanodrop 2000c (Thermo) and 1 μg of total RNA was used for reverse transcription (RT) using iScript Reverse Transcription Supermix (170-8841, Bio-Rad) following manufacturer's instruction. In all, 10 μl of each completed reaction mix was incubated in a thermal cycler (Model 2720, Applied Biosystems) using the following protocol: priming 5 min at 25°C, RT 30 min at 42°C, and RT inactivation 5 min at 85°C. cDNA products were diluted with 90 μl nuclease free water and analyzed by qPCR. All qPCR results were repeated in three independent experiments (34).

ChIP Assay

K562 cells were harvested by centrifugation at 1,000 g for 5 min and washed twice with 1× PBS. Cells were then crosslinked by 2 mM EGS (21565, Pierce) in PBS at room temperature for 30 min, followed by 1% formaldehyde (BP531-25, Fisher) for another 10 min. Crosslinking reaction was terminated by 125 mM glycine. Cell pellets were collected and lysed on ice for 30 min by cell lysis buffer (10 mM Tris-HCl, pH 8.1, 10 mM NaCl, 0.5% NP-40) with protease inhibitors. Chromatin was collected by spinning down, and the pellets were resuspended in cold nuclear lysis buffer (50 mM Tris-HCl, pH 8.1, 10 mM EDTA, 1% SDS, 25% glycerol) with protease inhibitors. In total, 300 μl of nuclear lysis buffer was used to resuspend chromatin from 2×10^6 cells.

Chromatin DNA was sheared to the size of 100-300 bp by sonication (Model Q800R, Active Motif) with the following protocol: amplification 75%, pulse on 15 s, pulse off 45 s, temperature 3°C. 200 μl of sheared chromatin was diluted by adding 1 ml of ChIP buffer (20 mM Tris-HCl pH8.1, 1.2% Triton X-100, 1.2 mM EDTA, 20 mM NaCl) with protease

inhibitors. 2 μg of control IgG and specific antibody were added to diluted chromatin respectively, followed by end to end rotation at 4°C overnight. At the same time, 12 μl of diluted chromatin was saved as input and processed later. Next day, 30 μl of PBS washed M-280 Sheep Anti-Mouse IgG Dynabeads (11204D, Invitrogen) was added to the chromatin, followed by rotating at 4°C for 4 h. Dynabeads were separated by DynaMag-2 Magnet (12321D, Invitrogen) and subsequently washed with 1 ml of the following buffer for 5 min at 4°C: wash buffer 1 (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 150 mM NaCl), wash buffer 2 (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 300 mM NaCl), wash buffer 3 (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 500 mM NaCl), wash buffer 4 (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.0), and TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA).

Complexes were eluted from beads with 500 μl elution buffer (1% SDS, 0.1 M $NaHCO_3$, 40 mM Tris-HCl, pH 8.0, 10 mM EDTA) and added with 200 mM NaCl. Eluates and inputs were treated at the same time with RNase A (EN0531, Thermo) at 65°C overnight, followed by proteinase K (25530-031, Invitrogen) treatment for 2 h. DNA was extracted by phenol/chloroform/isoamyl alcohol (AC327111000, Fisher) and precipitated by glycogen (10814-010, Invitrogen) and ethanol (61). DNA pellets were air-dried, dissolved in 50 μl nuclease free water, and analyzed by qPCR.

qPCR Assay

cDNA or ChIP DNA samples were analyzed by qPCR using SsoAdvanced Universal SYBR Green Supermix (172-5271, Bio-Rad) according to manufacturer's instruction. Briefly, 2 μl of cDNA or 5 μl of ChIP DNA samples, SYBR green supermix, nuclease free water, and primers for the target gene were mixed with a total reaction volume of 20 μl. The primer

sequences are listed as following: OGT Fwd CATCGAGAATATCAGGCAGGAG, OGT Rev CCTTCGACACTGGAAGTGTATAG; OGA Fwd TTCACTGAAGGCTAATGGCTCCCG, OGA Rev ATGTCACAGGCTCCGACCAAGT; HPRT Fwd ATTGGTGGAGATGATCTCTCAACTTT, HPRT Rev GCCAGTGTCAATTATATCTTCCACAA; −1000 OGA TSS Fwd TTGGGTCTCCTTGCTGTATG, −1000 OGA TSS Rev ACCTCACAGGTTGAGATAGATTT; OGA TSS Fwd GGGCTAGCCTATTAAGCTTCTTTA, OGA TSS Rev AGGGTGGCAAGCAGAAAT; +2700 OGA TSS Fwd TCCTTTCAGAGTTGCTCCAATA, +2700 OGA TSS Rev CAGTCAACCGAAACCATGAAC. A 96-well PCR plate (AVRT-LP, Midsci) with the mixture was loaded to CFX96 Touch Real-Time PCR Detection System (185-5195, Bio-Rad) with the following protocol: polymerase activation and DNA denaturation 30 s at 95°C, amplification denaturation 5 s at 95°C and annealing 30 s at 60 or 62°C with 40 cycles, and melt curve 65–95°C with 0.5°C increment 5 s/step.

Data Analysis

Quantification cycle (Cq) value was recorded by CFX Manager™ software. For cDNA qPCR data, dynamic range of RT and amplification efficiency was evaluated before using ΔΔCq method to calculate relative gene expression change. For ChIP DNA qPCR data, Cq value was normalized to percentage of input. Data generated in three independent experiments was presented as means ± standard error and analyzed using two-tailed Student's *t*-test with $P < 0.05$ as significant difference.

2.2 Results

Alteration in *O*-GlcNAc Levels Changes the Protein Expression of OGT and OGA

Previous reports demonstrated that different pharmacological inhibitors of OGA, PUGNAc and GlcNAc-thiazoline, rapidly increase the protein expression of OGA (55,62). We explored this phenomenon using another highly selective inhibitor of OGA Thiamet-G (TMG) (63). We altered the *O*-GlcNAc levels of SH-SY5Y neuroblastoma, HeLa cervical carcinoma, and K562 leukemia cells with TMG and measured *O*-GlcNAc, OGT, and OGA levels at various time points up to 48 h of TMG treatment. The *O*-GlcNAc levels were increased in the TMG treated samples while the pattern of *O*-GlcNAcylation was unique to each of the three cells lines used. OGA protein expression increased while OGT protein expression decreased gradually in the prolonged TMG treatment time points (**Figures 1a-c**). Additionally, in SY5Y cells, we used adenoviral-mediated OGT or OGA infection to alter *O*-GlcNAc levels. GFP was used as a control for the adenoviral infection. Cells overexpressing OGT showed an elevation in *O*-GlcNAc levels and a slight increase in OGA protein expression compared to control, while cells overexpressing OGA showed a decrease in *O*-GlcNAc levels and a slight increase in OGT protein expression compared to control (**Figure 1d**).

TMG Does Not Stabilize OGA Protein Expression

In order to explore the reason why TMG increases OGA protein level, we asked the question does TMG increase OGA protein stability. We pretreated cells with CHX to inhibit protein translation (64). We treated HeLa (**Figure 2a**) and K562 (**Figure 2b**) cells with TMG and observed a robust increase in OGA protein level (**Figures 2a and b**, Lane 2) compared to control cells without any treatment (**Figures 2a and b** , Lane 1). When HeLa cells were treated with CHX, OGT protein levels dramatically decreased compared to control, and we did not observe much of a decrease in OGA protein expression (**Figure 2a**, Lane 3). However, both OGT and OGA protein levels were dramatically decreased after CHX treatment in K562 cells

(**Figure 2b**, Lane 3) compared to control. Combination of CHX and TMG did not change the OGA or OGT protein levels (**Figures 2a and b**, Lane 4) compared to CHX treatment only suggesting that the TMG mediated increase in OGA protein expression was not due to increased stability of the protein.

OGA Transcript Level is Increased after TMG Treatment

Next, we investigated if OGT or OGA transcript level was altered after TMG treatment. We analyzed OGA mRNA level in SH-SY5Y (**Figure 3a**), HeLa (**Figure 3b**), and K562 (**Figure 3c**) cells. We found OGA mRNA level increased from 6 h TMG treatment in all three cell lines and was still elevated above control after 48 h TMG treatment (**Figures 3a and c**). The OGA mRNA level corresponded with the increase in protein level in **Figure 1**. However, the OGT mRNA level did not significantly change (**Figures 3d-f**). We also demonstrated that the corresponding OGA and OGT mRNA levels increased slightly but not significantly when OGT or OGA were overexpressed in SH-SY5Y cells (**Figures 3g and h**).

TMG Does Not Stabilize OGA mRNA

Next, we asked the question whether increased OGA mRNA level after TMG treatment was due to stabilized OGA mRNA. AMD, a RNA synthesis inhibitor, was used to test OGA mRNA stability. TMG treated HeLa cells showed an increase of OGA mRNA level compared to control cells without any treatment (**Figure** 4**a**). When cells were treated with AMD, both OGA and OGT mRNA levels were dramatically decreased compared to control (**Figures 4a and b**). Combination of AMD and TMG did not change the OGA and OGT mRNA levels compared to AMD treatment only (**Figures 4a and b**). The same results were observed when using K562 cells (**Figures 4c and d**).

RNA Pol II Occupancy is Decreased at OGA TSS after 48 h TMG Treatment

We next investigated RNA Pol II occupancy at the OGA TSS via RNA Pol II ChIP. In control K562 cells, RNA Pol II was bound to OGA TSS with little binding upstream (−1000) or downstream (+2700) of the TSS. However, after 48 h TMG treatment, RNA Pol II occupancy was decreased at the OGA TSS compared to control cells (**Figure 5a**). Normal mouse IgG ChIP was used as a negative control (**Figure 5b**).

2.3 Discussion

The production of UDP-GlcNAc, the substrate for OGT, integrates various metabolic substrates allowing the O-GlcNAc modification to act as a nutrient sensor (11,65). Consequently, cells are sensitive to changes in O-GlcNAc levels due to nutritional and metabolic flux and will adjust cellular functions accordingly. Prolonged alterations in homeostatic levels of O-GlcNAc will cause the protein expression of OGT and OGA to change in an effort to restore O-GlcNAc homeostasis (11). Exactly how cells sense alterations to homeostatic levels of O-GlcNAc and adjust OGT and OGA expression to compensate for the changes in O-GlcNAcylation is unclear. For example, pharmacological inhibition of OGA rapidly increases cellular O-GlcNAc levels; however, the protein expression of OGA will also increase in response to the elevation of O-GlcNAc (55,62). We sought to explore the mechanism as to how OGT and OGA protein expression changes in response to alterations in cellular O-GlcNAc levels. In agreement with previous reports, we found an increase in OGA protein expression as quickly as 8 h in HeLa cells and 24 h in K562 and SY5Y cells after treatment with TMG. OGT protein expression also decreased in these later time points (**Figure 1**). Due to the fact that increased levels of O-GlcNAc can increase the stability of proteins, such as p53 (66) and TET (ten-eleven translocation) (67), we postulated that increased O-GlcNAc could stabilize OGA. K562 or HeLa cells exposed to CHX in the presence of TMG showed no difference in the stability of either

OGT or OGA (**Figure 2**) suggesting that the increase in OGA protein expression was not due to increased stability and more likely to an increase of OGA transcripts.

Importantly, decreased O-GlcNAc levels do not necessarily increase OGT levels in all cell types; for example, blocking GFAT (glutamine fructose-6-phosphate amidotransferase) activity with 6-diazo-5-oxo-l-norleucine (DON) in HeLa cells lowered O-GlcNAc levels but did not increase OGT protein expression (55). On the other hand, OGA protein levels quickly decreased after Cre-mediated knockout of OGT in mouse embryonic fibroblasts (56), but OGA knockdown in colon cancer cells did not significantly decrease OGT protein expression (68). Changes in OGA protein expression appear more sensitive to changes in O-GlcNAc than OGT in HeLa cells, while both OGT and OGA expression significantly changed in SY5Y cells (**Figure 1**). Overexpression of OGA did not substantially influence OGT protein expression (**Figure 1**), and OGT overexpression did not change OGA expression (**Figure 1**). Recently, the development of a selective OGT inhibitor allowed for a dramatic reduction in cellular O-GlcNAcylation (69), which in turn caused OGA protein expression to rapidly decrease with only a minimal increase in OGT protein expression (69). The dynamic change in OGA protein expression was seen in the development of disease as well. In red blood cells of prediabetic individuals, OGA expression was significantly increased (70), and OGA protein levels correlated with increased blood glucose in these prediabetic patients. These data suggest that higher blood glucose levels promote increased flux through the hexosamine biosynthetic pathway leading to elevated OGT activity, followed by OGA protein levels increasing to restore cellular O-GlcNAc homeostasis in erythrocyte precursor cells. Together, these data support the proposed hypothesis that if OGT acts as a nutrient sensor allowing for rapid changes in O-GlcNAcylation due to

alterations, the cellular concentration of UDP-GlcNAc (71), then OGA should be less sensitive to nutrient changes and more sensitive to changes in O-GlcNAcylation.

In order to respond to changes in O-GlcNAc levels, cells rapidly and dramatically alter the expression of OGA mRNA (**Figure 3**). In the case of OGT, we did not detect a significant change in OGT mRNA levels after TMG treatment. The rapid increase in OGA mRNA levels after TMG treatment would argue that either OGA transcripts were more stable or transcriptional activity at the OGA promoter was increasing. We tested transcript stability by inhibiting RNA polymerase II with AMD (60). Interestingly, OGA and OGT transcript levels were not more stable after TMG treatment in the presence of AMD (**Figure 4**) suggesting that the increase in OGA mRNA levels with TMG was due to an increase in OGA gene transcription.

Next, we performed ChIP at the OGA promoter with an antibody that recognized all forms of RNA Pol II (phosphorylated and non-phosphorylated forms). After 48 h of prolonged TMG treatment in K562 cells, total RNA Pol II at the promoter was decreased compared to the control samples (**Figure 5**). Potentially, an antibody directed against the actively transcribing phosphorylated C-terminal domain (CTD) of RNA Pol II might have demonstrated an increase in enrichment of the phosphorylated forms of RNA Pol II at the promoter while non-phosphorylated forms of RNA Pol II would be less associated with the promoter. Interestingly, RNA Pol II is O-GlcNAcylated on the CTD at the fourth position of the CTD repeat, which is between the two activating phosphorylations at serine two and serine five on the CTD, which is needed for transcriptional elongation (22). Both O-GlcNAcylation and phosphorylation appeared to be mutually exclusive suggesting a cycle of O-GlcNAcylation and phosphorylation on the CTD repeats (23). Several groups have suggested that OGT and OGA work together to promote gene transcription by organizing the RNA Pol II preinitiation complex (PIC) (24,72). O-

GlcNAcylation was shown to promote the formation of the PIC in an in vitro transcription assay system; however, OGA activity was required for full transcriptional activation suggesting that OGT modified RNA Pol II, which initiated the formation of the PIC, while OGA was then required to remove the O-GlcNAc on the stalled RNA Pol II allowing for phosphorylation and transcription elongation (24). We have yet to explore RNA Pol II occupancy at the OGA promoter after a short TMG treatment (for example 6 h), which might yield a different result and needs to be studied further. The mRNA levels of OGA in K562 cells did begin to decrease at the 48 h TMG treatment suggesting that the OGA promoter might become inactive after prolonged TMG treatment. Reciprocal binding of OGT and OGA at active gene promoters provides several interesting future questions into the nature of transcriptional regulation, and the control of both the OGT and OGA promoter might be regulated in this manner.

Many transcription factors are modified by O-GlcNAc (13) and likely alteration of the O-GlcNAcylation level of a transcription factor could mediate the change in OGA transcription. We used the predictive software TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html) to identify potential transcription factor-binding sites in the first 1000 base pairs upstream of the OGA TSS (73). Among the transcription factor-binding sites in this sequence, GATA and MZF were the most predicted transcription factors. Due to the essential and ubiquitous expression of OGA (11), we anticipated that several housekeeping transcription factors might bind to this region, but we found only few of these. Interestingly, both GATA and MZF are important transcription factors regulating hematopoietic development (34,74). Perhaps the increased in OGA expression in prediabetic red blood cells (70) was partially due to changes in either of these two proteins. O-GlcNAcylation changes might lead to alteration of GATA or MZF occupancy at the OGA promoter. Some GATA family

members are modified by O-GlcNAc (16); thus, this presents an interesting avenue to explore in more detail.

Together, our data demonstrate that OGA protein and mRNA expression is sensitive to cellular levels of O-GlcNAc. Some disease states have OGA expression uncoupled from O-GlcNAc levels (75). In many different cancers, O-GlcNAc homeostasis appears to be disrupted with increased OGT protein expression and O-GlcNAc levels (76). Several pancreatic cancer cell lines have increased O-GlcNAc levels when compared to an immortalized control cell line; importantly, OGT protein expression was increased while OGA protein expression was decreased (75). The uncoupling of OGA expression to O-GlcNAc homeostasis could be an indicator of cancer progression and suggest that an increase of OGA protein expression would be beneficial therapeutically. Determining how O-GlcNAc regulates OGA expression and transcription will be crucial for understanding the biology of O-GlcNAcylation and how O-GlcNAc homeostasis is disrupted in disease.

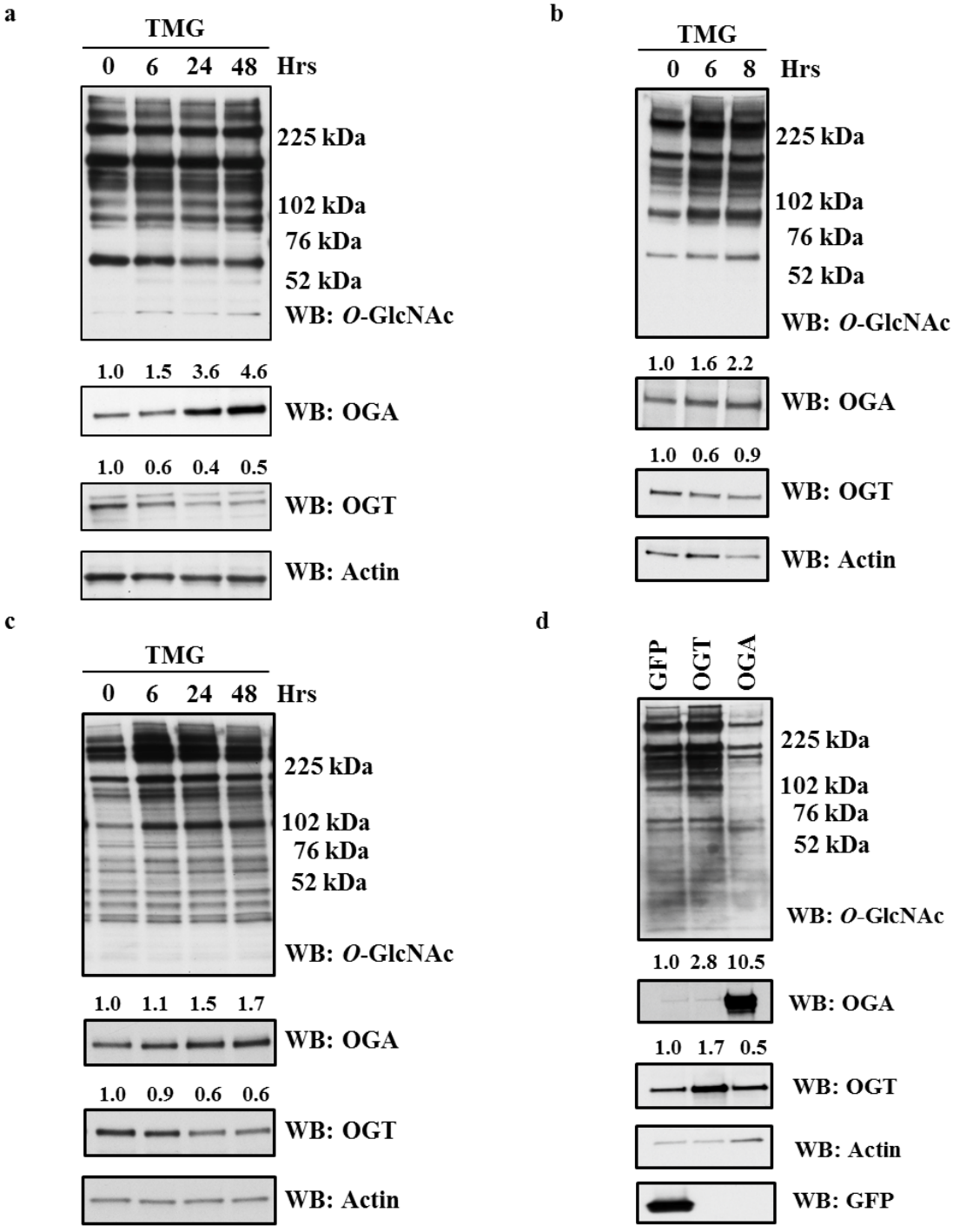

Figure 1: OGA protein level was increased after TMG treatment. **a,** SH-SY5Y neuroblastoma cells. **b,** HeLa cervical cells. **c,** K562 leukemia cells were treated with 10 μM TMG for indicated time. **d,** SH-SY5Y cells were infected with GFP, OGT, and OGA adenovirus at 75 MOI for 24 h. Cells were lysed, overall *O*-GlcNAc level, OGA and OGT protein level were analyzed by western blot with actin as a loading control. Average fold change for OGT and OGA is listed on the blots.

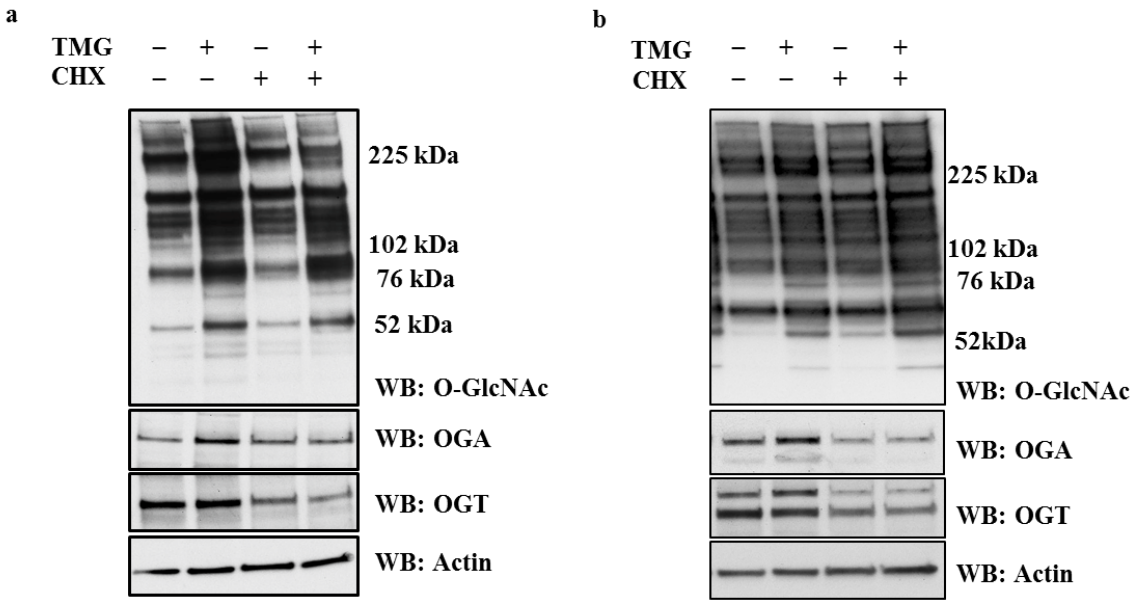

Figure 2:TMG does not stabilize OGA protein. **a**, HeLa cells and **(b)** K562 cells were treated with TMG, CHX (protein translation inhibitor), and CHX + TMG. Cells were lysed, overall O-GlcNAc level, OGA and OGT protein level were analyzed by western blot, with actin as loading control.

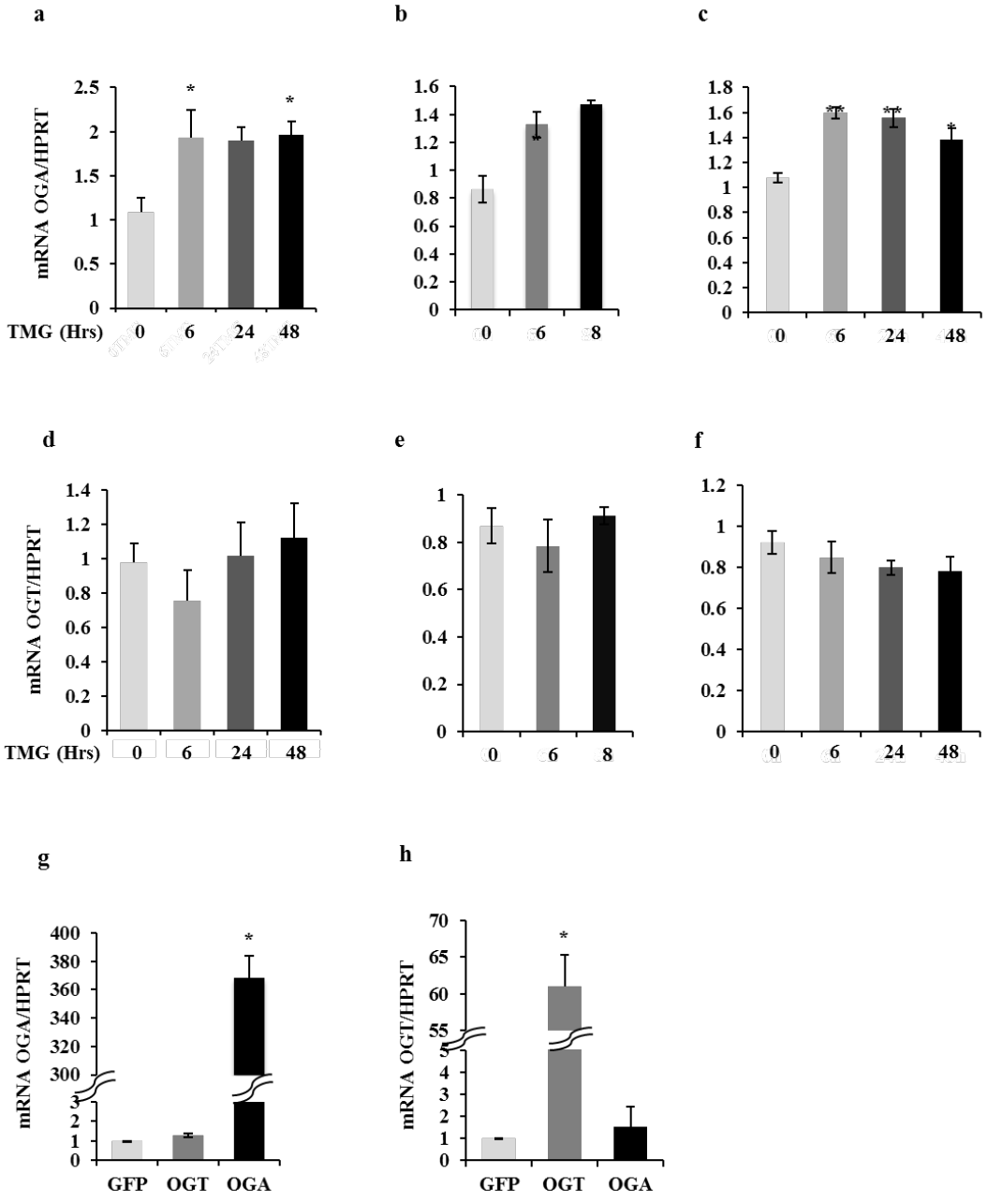

Figure 3:OGA mRNA level was increased after TMG treatment. After TMG treatment, relative OGA mRNA level in (**a**) SH-SY5Y, (**b**) HeLa, and (**c**) K562 cells, as well as OGT mRNA level in (**d**) SH-SY5Y, (**e**) HeLa, and (**f**) K562 cells was analyzed by qPCR. (**g**) OGA mRNA level and (**h**) OGT mRNA level in SH-SY5Y cells infected with GFP, OGT, and OGA adenovirus at 75 MOI for 24 h, respectively, were also analyzed by qPCR. Hypoxanthine-guanine phosphoribosyltransferase (HPRT) was served as internal control. *P < 0.05. **P < 0.01, compared with control (TMG 0 h or GFP), n = 3, Student's t-test.

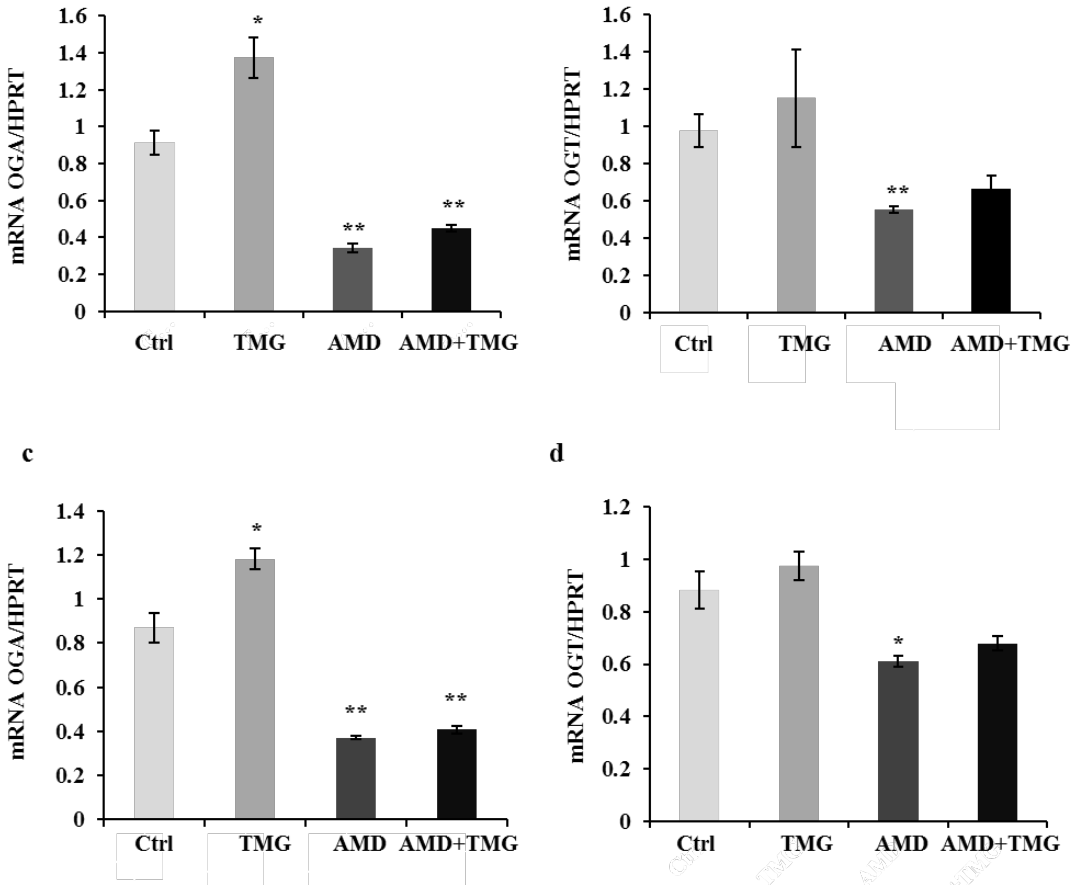

Figure 4:TMG does not stabilize OGA mRNA. HeLa cells **(a** and **b**) and K562 cells (**c** and **d**) were treated with TMG, AMD (RNA transcription inhibitor), and AMD + TMG, respectively. OGA (**a** and **c**) and OGT (**b** and **d**) mRNA level were analyzed by qPCR, with HPRT as internal control. *P < 0.05. **P < 0.01, compared with control, n = 3, Student's t-test.

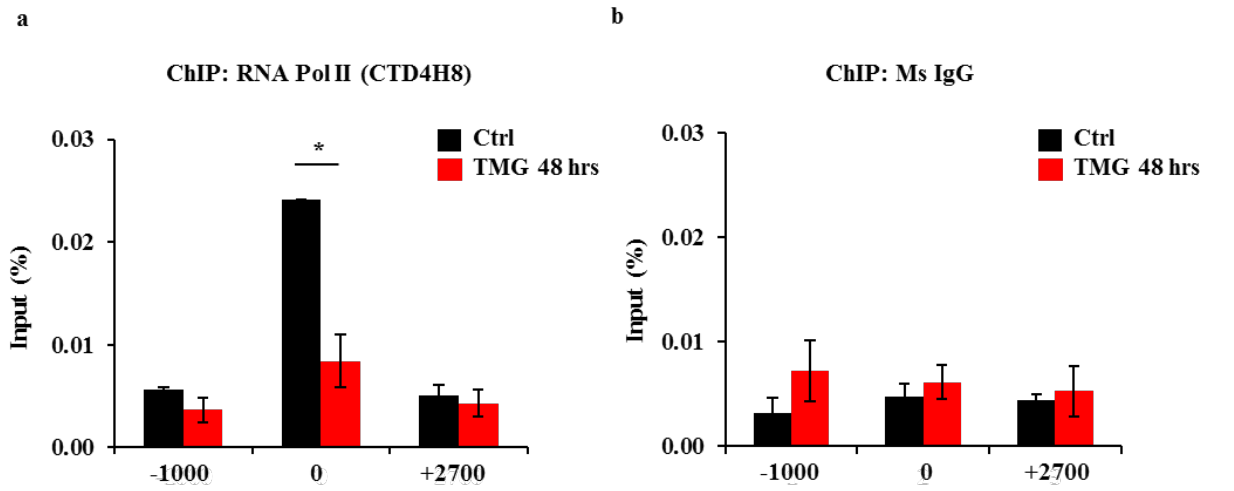

Figure 5:RNA Pol II occupancy at OGA TSS was decreased after 48 h TMG treatment in K562 cells. (**a**) RNA Pol II ChIP assay was performed on control and 48 h TMG treated cells. ChIP DNA was analyzed by qPCR using a set of primer targeting 1000 bp upstream of OGA TSS (−1000), OGA TSS (0), and +700 bp downstream of OGA TSS (+2700). *P < 0.05, n = 3, Student's t-test. (**b**) Normal mouse IgG ChIP served as a negative control.

Chapter 3: OGT and OGA Interact with Mi2β Protein at the $^{A}\gamma$-Globin Promoter

"This research was originally published in The Journal of Biological Chemistry. Zhen Zhang, Flávia C. Costa, Ee Phie Tan, Nathan Bushue, Luciano DiTacchio, Catherine E. Costello‖, Mark E. McComb, Stephen A. Whelan‖, Kenneth R. Peterson, and Chad Slawson. O-Linked N-Acetylglucosamine (O-GlcNAc) Transferase and O-GlcNAcase Interact with Mi2β Protein at the $^{A}\gamma$-Globin Promoter. *The Journal of Biological Chemistry*. 2016, 291:15628-40. © the American Society for Biochemistry and Molecular Biology."

The human β-globin gene cluster is comprised of five functional β-like globin genes: an embryonic gene (ε), duplicated fetal γ genes ($^{G}\gamma$ and $^{A}\gamma$), a minor adult gene (δ) and the adult β gene, which are expressed during development in the same order as they are arrayed, 5'-ε-$^{G}\gamma$-$^{A}\gamma$-δ-β-3'. The expression of these genes is not only controlled by an upstream locus control region (LCR), but also regulated by epigenetic signals, as well as transcription factors and their DNA binding motifs (32,77-79). There are two major switches of human β-like globin gene expression during development. The first switch is from embryonic to fetal globin, characterized by silencing of the embryonic ε-globin gene in the yolk sac and activation of the fetal γ-globin genes in the liver. The second switch is from the fetal to adult globins, characterized by the progressive silencing of γ-globin gene expression in the liver, and activation of adult globins (δ and β) in bone marrow (30).

Patients with sickle cell disease (SCD), caused by a point mutation in the β-globin gene, suffer chronic damage of multiple organs, increased risk of stroke, as well as cardiovascular abnormalities and dysfunction (80,81). However, SCD pathophysiology is ameliorated if the patients carry compensatory mutations that result in continued expression of the fetal γ-globin genes (fetal hemoglobin, HbF), a condition called hereditary persistence of fetal hemoglobin

(HPFH) (31). Thus, a logical clinical goal for treatment of this β-hemoglobinopathy is to up-regulate γ-globin synthesis (31). There are several mechanisms of γ-globin silencing. One of the well-known γ-globin silencers is B-cell lymphoma-leukemia A (BCL11A), which forms a protein complex with nucleosome remodeling and deacetylase complex (NuRD), GATA-1 and SOX6 to repress γ-globin expression (30). Krüppel-like factor 1 (KLF1), required for γ-globin gene activation, also stimulates BCL11A expression, which in turn represses γ-globin expression (30-32). Both BCL11A and KLF1 are positively regulated by binding of Mi2β (CHD4, Chromodomain Helicase DNA Binding Protein 4) a component of the NuRD complex (33). We previously demonstrated another modality of fetal globin repression using human γ-globin locus yeast artificial chromosome (β-YAC) transgenic mice. The ^{A}γ-globin gene is silenced during development following sequential binding of GATA-1, FOG-1 (Friend of GATA-1), and Mi2β proteins at the -566 GATA site relative to the ^{A}γ-globin gene CAP site in the fetal liver, between post-conception days E16-18 (34,35). GATA-1 is the DNA binding moiety in this complex, but GATA-1 can recruit both coactivators and corepressors in both FOG-1-dependent and independent pathways (36-39). The determinants of repressor versus activator complex assembly during development are not well understood. We postulate that one potential mechanism is through post-translational modifications (PTM) such as O-GlcNAcylation.

O-GlcNAcylation is the attachment of a single N-acetyl-glucosamine moiety to serine or threonine residues in mitochondrial, nuclear and cytoplasmic proteins. The enzyme O-GlcNAc transferase (OGT) adds the modification to proteins while O-GlcNAcase (OGA) removes the modification (11). O-GlcNAcylation integrates signals from a variety of nutrients in order to regulate cell signaling pathways, transcription, and cellular metabolism (82). Furthermore, the rate of O-GlcNAc cycling (the addition and then removal of O-GlcNAc) affects cell cycle

progression, mitotic signaling and spindle formation, and cellular respiration (11,13,14,53,54,83). Both OGT and OGA are essential for cellular function. Knockout mutations of OGT or OGA are embryonic lethal (84,85). Disruption of O-GlcNAcylation contributes to the development of diseases including cancer, diabetes, and Alzheimer's (12-14,57,86).

Growing evidence demonstrates that O-GlcNAcylation plays an important role in regulating transcription. O-GlcNAcylation is part of the histone code, and histones are O-GlcNAcylated under diverse cellular conditions (18,20,21). A recent study suggests that O-GlcNAc cycling also plays an important role in maintaining the epigenetic machinery in *Drosophila* (87). The RNA Pol II carboxy terminal domain (CTD) is modified by O-GlcNAc (22) facilitating pre-initiation complex (PIC) formation (24). Furthermore, numerous transcription factors are modified by O-GlcNAc and interact with OGT (11,25,27,29,88-90). Thus, O-GlcNAcylation could potentially regulate γ-globin gene transcription.

In this study, we asked the question whether O-GlcNAcylation plays a role in organizing the ^{A}γ-globin promoter. Herein, we demonstrate that both OGT and OGA interact with the ^{A}γ-globin promoter in both immortalized β-YAC bone marrow cells (BMCs) and in β-YAC transgenic mouse fetal liver. The occupancy of OGT, OGA, and Mi2β at the ^{A}γ-globin promoter is increased after TMG (an OGA inhibitor) treatment. Furthermore, we show that Mi2β is modified by O-GlcNAc and both OGT and OGA interact with GATA-1/FOG-1/Mi2β repressor complex. Taken together, our data suggest that O-GlcNAcylation is a novel mechanism that regulates ^{A}γ-globin gene expression by modulating GATA-1/FOG-1/Mi2β repressor complex recruitment to the -566 GATA site in the promoter during development.

3.1 Material and Methods

Antibodies

Primary antibodies and secondary antibodies for immunoblotting were used at a 1:1,000 and 1:10,000 dilution respectively. Antibodies for chromatin immunoprecipitation (ChIP) assay were used at 2 μg per reaction.

Antibodies for immunoblotting: O-GlcNAc (CTD 110.6), OGT (AL-34) and OGA (345) were gracious gifts from the Laboratory of Gerald Hart in the Department of Biological Chemistry at the Johns Hopkins University School of Medicine. Anti-O-linked N-acetylglucosamine antibody [RL2] (ab2739), anti-GAPDH antibody (ab9484), and anti-CHD4 antibody [3F2/4] (ab70469) were purchased from Abcam. Actin antibody (A2066), anti-chicken IgY HRP (A9046), and anti-mouse IgM HRP (A8786) were purchased from Sigma. Goat anti-rabbit IgG HRP (170-6515) and goat anti-mouse IgG HRP (170-6516) were purchased from Bio-Rad. Rabbit anti-goat IgG HRP (31402) was purchased from Thermo Scientific. Goat anti-rat IgG HRP (NA935V) was purchased from GE Healthcare. Mi2β Antibody (H-242) (sc-11378 X, also used for ChIP assay), GATA-1 antibody (sc-265 X), FOG antibody (sc-9361 X), normal rabbit IgG (sc-2027), normal chicken IgY (sc-2718), normal mouse IgG (sc-2025) and IgM (sc-3881), rat IgG (sc-2026), and normal goat IgG (sc-2028) were purchased from Santa Cruz Biotechnology. Antibodies for chromatin immunoprecipitation (ChIP): Rabbit control IgG (ab46540) was purchased from Abcam. OGT antibody (61355) was purchased from Active Motif. Anti-OGA antibody (SAB4200267) was purchased from Sigma.

Cell Culture

K562 cells were cultured in RPMI-1640 (R8758, Sigma) supplemented with 10% heat-inactivated fetal bovine serum (FBS, 100-106, Gemini), 1X MEM non-essential amino acids solution (Sigma), 1 mM sodium pyruvate (Sigma), 2.5 mM HEPES (Sigma), and 1X

penicillin/streptomycin (Sigma). K562 cells were treated with 10 μM Thiamet-G (TMG, S.D. Specialty Chemicals) for 6 hrs, followed by 0.75 mM sodium butyrate (NaB, B5887, Sigma) for 4 days or 10 μM hemin (51280, Sigma) and 3 mM N,N'-Hexamethylenebisacetamide (HMBA, 208320500, Acros Organics) for 3 days (34).

Wild-type (WT) and -566 $^{A}\gamma$-globin mutant murine chemical inducer of dimerization (CID)-dependent human β-YAC BMCs were generated as previously described (34,35). WT CID-dependent β-YAC BMCs have the normal GATA sequence at -566 of the $^{A}\gamma$-globin promoter, while -566 mutant CID-dependent β-YAC BMCs have a T>G point mutation at nucleotide -566. Cells were cultured in HyClone Iscove's Modified Dulbecco's Medium (IMDM, SH30005.02, Thermo Scientific) supplemented with 10% heat-inactivated FBS (Gemini), 1X penicillin/streptomycin (Sigma) and the CID, 5 μM CL-COB-II-293 (synthesized by the University of Kansas COBRE CCET Core C Synthesis Lab, commonly called AP20187). The cells were passaged at 2 x 10^5 cells/ml every three days in the presence of CL-COB-II-293 (91). WT β-YAC BMCs were treated with 10 μM TMG for 4 days, with fresh TMG and culture medium supplied daily.

Murine erythroleukemia (MEL) cell line stably expressing bacteria biotin ligase *birA* is a kindly gift from the Laboratory of Alan B. Cantor at Harvard Medical School, and cultured as previously described (42). All cells were incubated at 37°C, 5% CO_2 in a 95% humidified incubator.

Immunoprecipitation

Cells were lysed on ice for 30 min in Nonidet P-40 (NP-40) Lysis Buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 40 mM GlcNAc, and 1% NP-40) supplemented with 1 mM PMSF, 1 mM sodium fluoride, 1 mM β-glycerol phosphate, and 1X

protease inhibitor cocktail I (1 μg/ml leupeptin, 1 μg/ml antipain, 10 μg/ml benzamidine and 0.1% aprotinin). Two mg of cell lysates were incubated with 2 μg of antibody in a 1 ml reaction volume overnight with end-to-end rotation at 4°C. Next day, 20 μl of anti-Mouse IgM (μ-chain specific)-Agarose (A4540, Sigma), Protein G Sepharose 4 Fast Flow (17-0618-01, GE), or anti-Chicken IgY Agarose (DAIgY-AGA-2, Gallus Immunotech) slurry were added into the mixture, followed by end-to-end rotation at 4°C for 2 hrs. Agarose beads were washed 3x with 1 ml NP-40 Lysis Buffer and mixed with 20 μl 2X protein solubility mixture (50 mM Tris-HCl, pH 6.8, 5 mM EDTA, 4% SDS, 25% sucrose, 2.5% β-mercaptoethanol, 0.04% Pyronin-Y). Forty μg of cell lysates were mixed with 4X protein solubility mixture as input. The beads and input were heated at 95°C for 2 min and were subjected to immunoblotting.

Immunoblotting

All electrophoresis was performed with 4-15% gradient Criterion TGX Precast Gels (567-1084, Bio-Rad) at 140 volts, followed by transferring protein to PVDF membrane (IPVH00010, Millipore) at 0.4 Amps. Blots were blocked by 3% BSA in TBST (25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 0.05% Tween-20) at room temperature for 20 min, then incubated with primary antibody at 4°C overnight. After washing with TBST 5x for 5 min each time, blots were incubated with HRP-conjugated secondary antibody for 1 hr at room temperature, followed by TBST washes again and developed using chemiluminescent substrate (E2400, Denville Scientific). Blots were stripped in 200 mM glycine, pH 2.5 at room temperature for 1 hr and probed with other primary antibodies. All immunoblotting results were repeated in three independent experiments and representative images are shown.

ChIP Assay

WT and -566 $^{A}\gamma$-globin mutant CID-dependent b-YAC BMCs were used or WT β-YAC transgenic mouse fetal liver single cell suspensions were prepared as previously described (34,35). Briefly, fetal liver from WT β-YAC transgenic mice at 12 and 18 days post-conception (E12 or E18) were isolated and kept on ice in 1X PBS. The liver tissue was cut into small pieces with a scissor and a single cell suspension was prepared by repeatedly passing the liver pieces through a 1 ml syringe with sequentially smaller needle sizes (16-, 18- and 20-gauge, BD Biosciences) and then subjected to ChIP assay.

The ChIP assay was performed using a previously described method (92) with slight modifications. Briefly, cells were cross-linked by 2 mM EGS (21565, Pierce) in PBS for 30 min, followed by 1% formaldehyde (BP531-25, Fisher) for 10 min at room temperature. The cross-linking reaction was terminated by the addition of 2.5 M glycine to a final concentration of 125 mM. Cells were lysed and chromatin was collected by centrifugation at 1,000x *g*. Three hundred μl of Nuclear Lysis Buffer (50 mM Tris-HCl, pH 8.1, 10 mM EDTA, 1% SDS, 40 mM GlcNAc, and 25% glycerol) was used to resuspend chromatin pellets from 5 x 10^6 cells. Chromatin DNA was sheared to the size of 100~300 bp and immuoprecipitated with 2 μg of control IgG or specific antibody, respectively, at 4°C overnight. Next day, 30 μl of PBS-washed M-280 Sheep Anti-Rabbit IgG Dynabeads (11204D, Invitrogen) were added to the chromatin, followed by rotation at 4°C for 4 hrs. Dynabeads were separated using a DynaMag-2 Magnet (12321D, Invitrogen) and subsequently washed and eluted as previously described (92). The eluate was subjected to reverse cross-linking by sequential treatments with RNAse A and proteinase K (92). ChIP DNA was extracted as previously described (92), and dissolved in 50 μl nuclease free water (AM9906, Ambion), followed by quantitative PCR (qPCR).

Total RNA Isolation and RT-PCR

Total RNA was isolated as previously described (92). Briefly, RNA was isolated from 2 x 10^6 cells using TRI reagent solution (T9424, Sigma) according to the manufacturer's instructions. For reverse transcription (RT), 0.5 μg of total RNA was used in iScript Reverse Transcription Supermix (170-8841, Bio-Rad) following the manufacturer's instructions. Ten μl of each completed reaction mix was incubated in a thermal cycler (Model 2720, Applied Biosystems) using the following protocol: priming 5 min at 25°C, RT 30 min at 42°C, and RT inactivation 5 min at 85°C. cDNA products were diluted with 90 μl nuclease free water and analyzed by qPCR.

qPCR Assay

cDNA and ChIP DNA were analyzed by qPCR as described previously (92) and according to the manufacturer's instructions. Two μl of cDNA or 5 μl of ChIP DNA sample, 10 μl of SsoAdvanced Universal SYBR Green Supermix (172-5271, Bio-Rad), nuclease free water, and primers for the target gene were mixed together in a total reaction volume of 20 μl. The primer sequences are listed as following: Human γ-globin Fwd GTATTGCTTGCAGAATAAAGCC, Rev GACCGTTTTGGCAATCCATTTC; Mouse HPRT Fwd GGCCAGACTTTGTTGGATTTG, Rev CGCTCATCTTAGGCTTTGTATTTG; ^{A}γ-globin Promoter Fwd CTAATCTATTACTGCGCTGA, Rev GTTTCTAAGGAAAAAGTGCT; Human OGT Fwd CATCGAGAATATCAGGCAGGAG, Rev CCTTCGACACTGGAAGTGTATAG; Human OGA Fwd TTCACTGAAGGCTAATGGCTCCCG, Rev ATGTCACAGGCTCCGACCAAGT; Human HPRT Fwd ATTGGTGGAGATGATCTCTCAACTTT, Rev GCCAGTGTCAATTATATCTTCCACAA. The reactions were run in a CFX96 Touch Real-Time PCR Detection System (185-5195, Bio-Rad) using the following conditions: polymerase activation and DNA denaturation 30 s at 95°C,

amplification denaturation 5 s at 95°C, and amplification annealing and extension 30 s at 60°C or 62°C for 40 cycles.

qPCR Data Analysis

Quantification cycle (Cq) values were calculated by CFX Manager™ software. For cDNA qPCR data, the dynamic range of RT and amplification efficiency was evaluated before applying the ΔΔCq method to calculate relative gene expression change. The transcription level of the target gene was normalized to the internal control as fold change. For ChIP DNA qPCR data, the Cq value was normalized as percent of input. Data generated in three independent experiments are presented as mean ± standard error; the two-tailed Student *t*-test statistic was applied with $P < 0.05$ considered to be a significant difference.

3.2 Results

OGT and OGA interact with the ^{A}γ-globin gene promoter

To determine whether OGT and OGA interact with the ^{A}γ-globin gene promoter, we utilized WT and -566 mutant murine CID-dependent β-YAC BMCs (34,35). At the -566 region of the ^{A}γ-globin promoter relative to the CAP site, WT β-YAC BMCs have a GATA motif that is bound by a GATA-1/FOG-1/Mi2β repressor complex when γ-globin is silenced during the adult stage of hematopoiesis. In the -566 mutant β-YAC BMCs, a T>G point mutation at nucleotide -566 (**Figure 6a**) results in an HPFH phenotype. This mutation changes the GATA site to a GAGA site preventing binding of the GATA-1/FOG-1/Mi2β repressor complex, and results in continued γ-globin gene transcription in the adult stage of erythropoiesis (**Figure 6b**). Both WT and -566 mutant β-YAC BMCs express OGT and OGA, and have similar overall O-GlcNAc levels (**Figure 6c**). Next, we performed ChIP assays in both of these BMC populations to test for OGT and OGA occupancy at the ^{A}γ-globin promoter. In WT β-YAC BMCs both OGT and

OGA associated with the -566 region of the $^{A}\gamma$-globin promoter; however, in -566 mutant β-YAC BMCs, this association was dramatically decreased (**Figure 6d**). These data suggest that OGT and OGA interact with component(s) of the GATA-1/FOG-1/Mi2β repressor complex at -566 GATA site of $^{A}\gamma$-globin promoter.

OGT, OGA and Mi2β increase at the $^{A}\gamma$-globin gene promoter after TMG treatment

Our data suggest that OGT and OGA play a role in regulating γ-globin transcription by modulating GATA-1/FOG-1/Mi2β repressor complex activity at the $^{A}\gamma$-globin promoter. To test this hypothesis, we treated the WT β-YAC BMCs for 4 days with Thiamet-G (TMG, S.D. Specialty Chemicals), a competitive inhibitor of OGA, to increase overall O-GlcNAc levels. After TMG treatment for 4 days, the overall O-GlcNAc levels were increased, with increased OGA and decreased OGT protein level (**Figure 7a**), a similar effects that we observed previously (92). Interestingly, the γ-globin transcription level was decreased (**Figure 7b**). In order to seek the mechanism of deceased γ-globin transcription after TMG treatment, we performed ChIP assays on WT β-YAC BMCs. After TMG treatment for 4 days, the occupancy of both OGT and OGA was increased at the $^{A}\gamma$-globin promoter compared to the cells without TMG treatment (**Figure 7c**). We also observed an increase of Mi2β occupancy at $^{A}\gamma$-globin promoter after TMG treatment (**Figure 7d**). These data demonstrate that OGT, OGA, and Mi2β increase at the $^{A}\gamma$-globin promoter after TMG treatment.

OGT and OGA interact with the $^{A}\gamma$-globin promoter during fetal liver development

In β-YAC transgenic mice, the γ-globin gene is expressed during early fetal development; however, prior to birth γ-globin expression is silenced in hematopoietic cells of the fetal liver, partly through the gradual recruitment of the GATA-1/FOG-1/Mi2β repressor complex (34,35). To determine if OGT and OGA associate with the $^{A}\gamma$-globin promoter during different fetal

stages of hematopoietic development as gene silencing progresses, we performed ChIP assays in β-YAC transgenic mouse fetal liver single cell suspensions to test for recruitment of Mi2β, OGT and OGA. At post-conception day 12 (E12), γ-globin is expressed in the murine fetal liver. However, γ-globin is repressed at post-conception day 18 (E18), when the GATA-1/FOG-1/Mi2β repressor complex is recruited to the -566 ^{A}γ-globin GATA silencer site (34). As a positive control for repressor complex binding, we confirmed our previous data (34) that Mi2β binds to the ^{A}γ-globin promoter at day E18 when γ-globin is repressed, but not in day E12 fetal liver when γ-globin is expressed (**Figure 8a**). Interestingly, OGT occupancy was slightly increased at the ^{A}γ-globin promoter at day E18 compared to day E12 (**Figure 8b**). We did not observe an increase of OGA occupancy at the ^{A}γ-globin promoter at day E18 compared to day E12, although OGA interacts with the promoter on both days (**Figure 8c**). These data suggest that recruitment of OGT and OGA to the ^{A}γ-globin promoter may contribute to ^{A}γ-globin gene silencing.

Mi2β is modified by O-GlcNAc

Since OGT and OGA interact with the ^{A}γ-globin promoter where the GATA-1/FOG-1/Mi2β repressor complex binds, we determined whether any component(s) of the repressor complex is modified by O-GlcNAc. For these experiments, we used human erythroleukemia K562 cells, a well-established model system for globin studies (93). The addition of sodium butyrate (NaB) to K562 cells increases γ-globin expression (93). Our experiments confirmed this phenotype; the γ-globin mRNA level was increased by NaB as expected (**Figure 9a**). After 4 days treatment, overall O-GlcNAc levels were decreased, with an increase of OGA protein level and a decrease of OGT protein level (**Figure 9b**), which corresponded to an increased OGA mRNA level (**Figure 9c**) and a decreased OGT mRNA level (**Figure 9d**). We also used hemin

and HMBA (H-H), another well-established inducer of differentiation in K562 cells (34) that also relieves γ-globin repression (data not shown). We used these two different methods to alleviate ^{A}γ-globin promoter repression in order to eliminate any artifacts from these chemical inducers that could skew the data.

Next, we assessed if any proteins of the GATA-1/FOG-1/Mi2β complex were modified by O-GlcNAc pre- or post-γ-globin promoter activation. We used a specific antibody to O-GlcNAc to immunoprecipitate potential O-GlcNAc modified proteins from K562 lysates before and after NaB treatment. TMG was used to increase the overall O-GlcNAc levels by competitively inhibiting OGA. Blots were probed individually with Mi2β, FOG-1, and GATA-1 antibodies, respectively. We found that CTD 110.6 immunoprecipitated Mi2β before NaB treatment (Day 0), and TMG treatment increased the amount of Mi2β immunoprecipitated by CTD 110.6. However, CTD 110.6 immunoprecipitated Mi2β protein level was decreased after 4 days of NaB treatment (**Figure 10a,** top panel). Similar results were observed when using H-H as inducer instead of NaB (**Figure 10a,** bottom panel).

We also performed the converse immunoprecipitation using Mi2β antibody and probed the blot for O-GlcNAc. Mi2β was O-GlcNAc modified prior to NaB induction, and this modification dramatically decreased after 4 days of NaB treatment (**Figure 10b,** top panel). The same results were obtained when H-H were used as inducer (**Figure 10b,** bottom panel). These data suggest that O-GlcNAcylation of Mi2β changes during γ-globin induction.

We also performed two control experiments to further validate O-GlcNAcylation of Mi2β. Mi2β was immunoprecipitated from duplicate sets of lysates. Both sets were transferred to a PVDF membrane, which was cut in half. One set was probed with CTD 110.6 antibody detecting O-GlcNAcylated Mi2β (**Figure 10c**). However, detection of O-GlcNAcylated Mi2β

on the other set was abolished when CTD 110.6 antibody was pre-incubated with 500 mM free GlcNAc (**Figure 10d**). In another experiment, Mi2β immunoprecipitates were treated with *Cp*NagJ hexosaminidase (gift from Daan van Aaltan, University of Dundee), which removes oxygen-linked hexoses from proteins (94). In the absence of *Cp*NagJ treatment, Mi2β immunoprecipitates were O-GlcNAcylated, while detection of the Mi2β O-GlcNAcylation was completely lost after *Cp*NagJ treatment (**Figure 10e**).

Mi2β interacts with OGT

Since Mi2β is modified by O-GlcNAc, we next explored if Mi2β interacts with OGT. First, we performed IP experiments using OGT antibody on lysates prepared from K562 cells that had been treated with or without NaB. OGT interacts with Mi2β prior to NaB induction but not after 4 days of NaB induction (**Figure 11a**). The lost interaction of OGT and Mi2β after NaB treatment corresponds to the decreased O-GlcNAcylation of Mi2β after NaB induction (**Figure 11a and b,** top panel). Similar IP results were observed when cells were differentiated by H-H for three days (**Figure 11b**).

We also performed IPs against Mi2β in K562 cell lysates prior to and following NaB induction. Mi2β antibody was able to co-IP OGT prior to NaB treatment, but unable to immunoprecipitate OGT after 4 days of NaB treatment (**Figure 11c**). The same results were seen when we used H-H to differentiate cells (**Figure 11d**). These data demonstrate that Mi2β interacts with OGT before induction of γ-globin in K562 cells or prior to their terminal differentiation, but not after treatments that induce HbF or terminal erythroid differentiation.

Mi2β interacts with OGA

Since our data demonstrated OGA also interacts at the -566 region of the $^{A}\gamma$-globin promoter, we next tested whether Mi2β interacts with OGA as well. We performed co-IP

experiments as described in the previous section. OGA antibody was able to co-IP Mi2β prior to and after NaB induction (**Figure 12a**). Similar co-IP results were observed when cells were differentiated by H-H (**Figure 12b**). As described above for the Mi2β-OGT interaction experiments, we performed the converse co-IP using Mi2β antibody and looked for interactions with OGA. Mi2β antibody was able to co-IP OGA prior to and after NaB treatment (**Figure 12c**). The same results were obtained when we used H-H to differentiate cells (**Figure 12d**). These data indicate that Mi2β interacts with OGA before and after induction of γ-globin in K562 cells.

OGT and OGA interact with GATA-1 and FOG-1

Our data demonstrated that both OGT and OGA interact with Mi2β, a component of GATA-1/FOG-1/Mi2β repressor complex. Thus, OGT and OGA might also interact with GATA-1 and FOG. In order to confirm the interaction of OGT/OGA with GATA-1/FOG-1, we performed IPs in MEL *birA* cells (42) since they express relative higher GATA-1 and FOG-1 protein compared to K562 and β-YAC BMCs (data not shown). GATA-1 antibody was able to pull down both OGT and OGA (**Figure 13a**); FOG-1 antibody can also pull down OGT and Mi2β (**Figure 13b**); and OGA antibody was able to IP FOG-1 (**Figure 13c**). These data suggest that OGT and OGA can also interact with GATA-1 and FOG-1 besides Mi2β.

3.3 Discussion

Using WT and -566 mutant CID-dependent β-YAC BMCs, we demonstrated that both OGT and OGA interact with the ^{A}γ-globin promoter (**Figure 6**) at the -566 GATA site occupied by the GATA-1/FOG-1/Mi2β repressor complex when γ-globin is silenced. In addition, the occupancy of OGT, OGA and Mi2β at the ^{A}γ-globin promoter is increased after TMG treatment (**Figure 7**). In β-YAC transgenic mouse fetal liver, OGT and Mi2β recruitment was increased at

the -566 GATA site of the $^{A}\gamma$-globin promoter at post-conception day E18 when γ-globin is repressed compared to day E12 when γ-globin is expressed (**Figure 8**). Furthermore, we determined that Mi2β is modified by O-GlcNAc (**Figure 10**), and OGT/OGA interacts with GATA-1/FOG-1/Mi2β repressor complex (**Figure 11-13**). Taken together our data suggest that O-GlcNAcylation plays a role in regulating $^{A}\gamma$-globin gene expression via recruitment of the GATA-1/FOG-1/Mi2β repressor complex at the -566 GATA site of the $^{A}\gamma$-globin promoter.

O-GlcNAcylation plays an important role in controlling transcription. O-GlcNAc regulates gene transcription by directly modifying transcription factors/co-factors, RNA Pol II CTD (24), or by altering how the activity of these factors is modulated by other PTMs (25). A previous study by Ranuncolo et al. suggested that an O-GlcNAcylation cycle on RNA Pol II C-terminal domain (CTD) occurs at gene promoters. In this cycle, unphosphorylated RNA Pol II (Pol IIA) CTD is first modified with O-GlcNAc (Pol IIγ) by OGT, promoting the association of the general transcription factors with Pol IIγ. Next, O-GlcNAc is removed from Pol IIγ by OGA converting Pol IIγ back to Pol IIA. Subsequently, Pol IIA is phosphorylated (Pol IIO) by TFIIH and P-TEFb and transcription elongation occurs (24). Correspondingly, we demonstrated that both OGT and OGA interact with the $^{A}\gamma$-globin promoter in β-YAC BMCs (**Figure 6d**) and β-YAC transgenic mouse fetal liver (**Figure 8b and c**). The addition of TMG to the β-YAC BMCs decreased OGA activity and reduced O-GlcNAc cycling of Mi2β at the $^{A}\gamma$-globin promoter leading to the recruitment of more OGA to the $^{A}\gamma$-globin promoter to maintain the normal O-GlcNAc cycling (**Figure 7c**). In addition, TMG can also increase the Mi2β O-GlcNAc levels resulting in the extended occupancy of Mi2β (**Figure 7d**), other components of the GATA-1/FOG-1/Mi2β repressor complex, and OGT (**Figure 7c**) at the $^{A}\gamma$-globin promoter. All the above effects caused by TMG treatment could stabilize GATA-1/FOG-1/Mi2β repressor

complex and reduce the γ-globin transcription level (**Figure 7b**). Furthermore, these data suggest that both OGT and OGA interactions at the ^{A}γ-globin promoter are important for the recruitment and organization of the GATA-1/FOG-1/Mi2β repressor complex at the -566 GATA site. Similar to the assembly of the PIC at gene promoters described by Ranuncolo et al., the assembly of the GATA-1/FOG-1/Mi2β repressor complex at ^{A}γ-globin promoter might also require the O-GlcNAcylation of Mi2β first to recruit other factors, followed by OGA removal of the O-GlcNAc.

Although the GATA-1/FOG-1/Mi2β repressor complex is assembled sequentially during development in a murine model (35), the mechanisms controlling this assembly are unknown. Our data show that Mi2β is modified by O-GlcNAc (**Figure 10**). O-GlcNAcylation of Mi2β may facilitate the sequential recruitment of Mi2β to the GATA-1/FOG-1 proteins already bound at the ^{A}γ-globin promoter. Mi2β/NuRD also mediates ^{A}γ-globin gene silencing independent of the GATA-1/FOG-1/Mi2β repressor complex at the -566 GATA site of the ^{A}γ-globin promoter. Mi2β binds directly to BCL11A and KLF1, and positively regulates BLC11A and KLF1 expression, which in turn leads to γ-globin silencing (33). Thus, O-GlcNAcylation of Mi2β likely plays a role in mediating γ-globin silencing by affecting the action of several previously identified γ-globin repressor complexes that interact with different *cis*-regulatory motifs in the promoters of the γ-globin genes.

The GATA-1/FOG-1/Mi2β repressor complex mediates one mode of ^{A}γ-globin gene silencing; the components of which are sequentially recruited to the -566 GATA silencer of the ^{A}γ-globin promoter (34,35). How this repressor complex is assembled and organized during development is unclear. Our data show that OGT and OGA interact with Mi2β (a component of NuRD), GATA-1 and FOG-1. (**Figure 11-13**). These data suggested that OGT and OGA are

part of the GATA-1/FOG-1/Mi2β repressor complex. Mi2β or other components of the GATA-1/FOG-1/Mi2β repressor complex may mediate the recruitment of OGT and OGA to the $^{A}\gamma$-globin promoter. Interestingly, we show that besides OGT, OGA also interacts with Mi2β, suggesting that O-GlcNAc cycling on Mi2β may be a potentially important mechanism for repressor complex assembly and transcriptional repression. Of note, OGT interacts with numerous transcription factors and mediates gene activation and silencing. For example, OGT interacts with mSin3A recruiting OGT to promoters to repress transcription. Furthermore, inactivation of transcription factors and RNA polymerase II by O-GlcNAc modifications contribute to gene repression (27). *Drosophila* polycomb group (PcG) gene (*sxc*) encodes for OGT, and null mutations in OGT lead to polycomb defects suggesting OGT is critical for PcG medicated gene silencing (28). Potentially, OGT is mediating Mi2β repression of the $^{A}\gamma$-globin promoter.

The transcription factor GATA-1 is essential for erythroid development. This DNA-binding protein functions in either transcriptional activator or repressor protein complexes as the DNA-docking moiety, in either a FOG-1-dependent or independent fashion (36-39). A fundamental question remains as to what are the mechanisms that determine whether GATA-1 functions in an activation complex or a repressor complex. In some instances, binding by partner protein TAL1 at a neighboring recognition site assures the recruitment of an activator complex (95), and to a limited degree the WGATAR binding site context plays a role (96). However, these mechanisms do not explain the outcome for the majority of protein complex binding events. The O-GlcNAcylation status of Mi2β during erythroid development may provide a mechanism by which different co-repressors or co-activators are recruited to interact with GATA-1. Moreover, GATA-1 itself may be a target of O-GlcNAcylation, influencing how it

interacts with co-activators or co-repressors. In addition, FOG-1 also may be a target for O-GlcNAcylation, and the differential post-translational modification of FOG-1 might determine the recruitment of different co-activators or co-repressors from those recruited by GATA-1 alone.

O-GlcNAcylation and phosphorylation can compete for specific Ser/Thr sites on a protein, or influence the PTM state of nearby Ser/Thr sites. For example, threonine 58 on c-Myc is reciprocally modified; phosphorylation promotes its degradation, but O-GlcNAc promotes its stability (97). O-GlcNAcylation of C/EBPβ (CCAAT enhancer-binding protein β) influences adjacent phosphorylation sites. O-GlcNAcylation of C/EBPβ Ser180 and Ser181 prevents the phosphorylation of Thr188, Ser184 and Thr179, which is required for C/EBPβ DNA binding activity (98). Many phosphorylation sites have been mapped on Mi2β by mass spectrometry (99-104), and some of these sites could potentially be modified by O-GlcNAc. The reciprocal occupancy of GlcNAc or phosphate at these sites could fine tune Mi2β activity or interactions. Moreover, O-GlcNAcylation of Mi2β may also influence the phosphorylation state of specific sites on Mi2β and therefore its function in the formation of and/or within the GATA-1/FOG-1/Mi2β repressor complex. In future studies, mapping Mi2β O-GlcNAc sites and exploring their biological function will be critical to understanding how O-GlcNAc cycling regulates the NuRD complex.

In summary, we demonstrated that OGT and OGA associate with the $^{A}\gamma$-globin promoter (**Figure 6 and 8**); Inhibition of OGA by TMG leads to increased occupancy of OGT, OGA and Mi2β at the $^{A}\gamma$-globin promoter and decreased γ-globin transcription level (**Figure 7**); both OGT and OGA can interact with Mi2β, GATA-1 and FOG-1 (**Figure 11-13**); and Mi2β is O-

GlcNAcylated (**Figure 10**). Our data indicate that Mi2β O-GlcNAcylation, OGT, and OGA facilitate assembly of the GATA-1/FOG-1/Mi2β (NuRD) repressor complex.

Based on our findings, we propose that OGT/OGA contributes to the regulation of GATA-1/FOG-1/Mi2β mediated ^{A}γ-globin gene repression. O-GlcNAc cycling on Mi2β is mediated by interactions with OGT and OGA (**Figure 14a**). Potentially, O-GlcNAcylated Mi2β is recruited to -566 of ^{A}γ-globin promoter with OGT and OGA by GATA-1 and FOG-1. Subsequently, Mi2β recruits other potential co-factors (**Figure 14a and b**) to form a stable repressor complex. At this point, the ^{A}γ-globin transcription is repressed (**Figure 14b**). When ^{A}γ-globin transcription is activated, Mi2β interactions with OGT or other co-factors, is reduced, OGA removes O-GlcNAc from Mi2β, and the repressor complex is dissembled (**Figure 14c**). However, several questions still need to be addressed. How is GATA-1/FOG-1/Mi2β assembly affected when the OGT is altered? Although Mi2β is modified by O-GlcNAc, are GATA-1, FOG-1 or other components of the NuRD complex modified by O-GlcNAc? Once the mechanistic role of O-GlcNAcylation in the assembly of the GATA-1/FOG-1/Mi2β repressor complex is understood, we may be able to target this complex to reverse γ-globin gene silencing and effectively treat SCD patients by increasing their HbF to beneficial therapeutic levels.

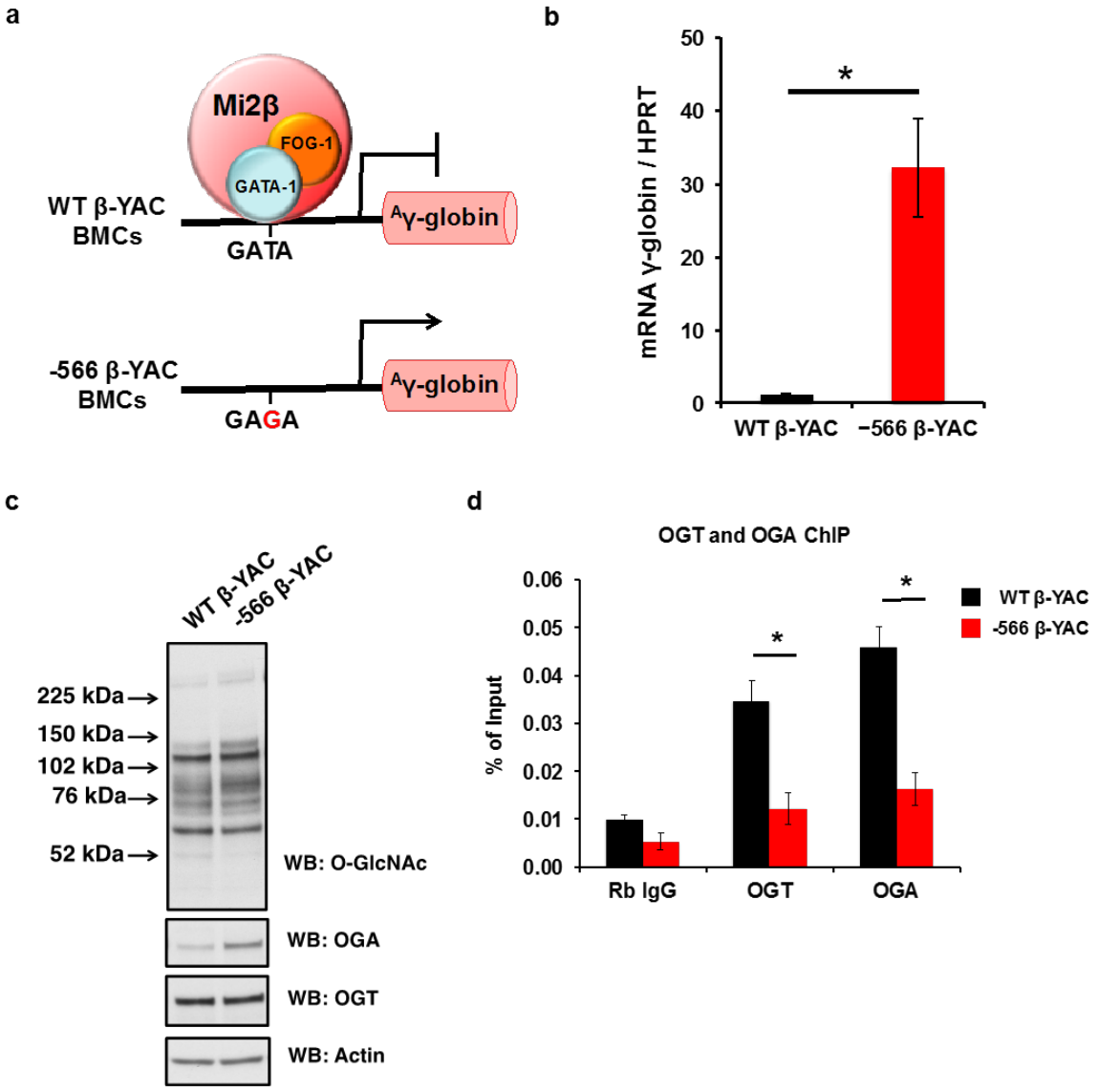

Figure 6:OGT and OGA interact with the $^{A}\gamma$-globin promoter in CID-dependent β-YAC BMCs. **a**, WT CID-dependent β-YAC BMCs have a normal GATA motif that is bound by a GATA-1·FOG-1·Mi2β repressor complex, mediating γ-globin repression; whereas −566 mutant CID-dependent β-YAC BMCs have a T→G point mutation at nucleotide −566, which alleviates γ-globin repression. **b**, γ-globin mRNA level in WT and −566 mutant β-YAC BMCs was analyzed by qPCR. Mouse HPRT was used as an internal control. **c**, Overall O-GlcNAc levels and OGA and OGT protein expression in total cell lysates from WT and −566 mutant CID-dependent β-YAC BMCs were analyzed by immunoblotting. Actin was used as a loading control. **d**, OGT and OGA ChIP assays were performed on WT and −566 mutant β-YAC BMCs, respectively. ChIP DNA was analyzed by qPCR using a set of primers targeting the −566 GATA site of the $^{A}\gamma$-globin promoter. Normal rabbit (Rb) IgG served as a negative control. All experiments were performed with at least three biological replicates (* indicates $p < 0.05$, Student's t test). Error bars represent S.E. WB, Western blotting.

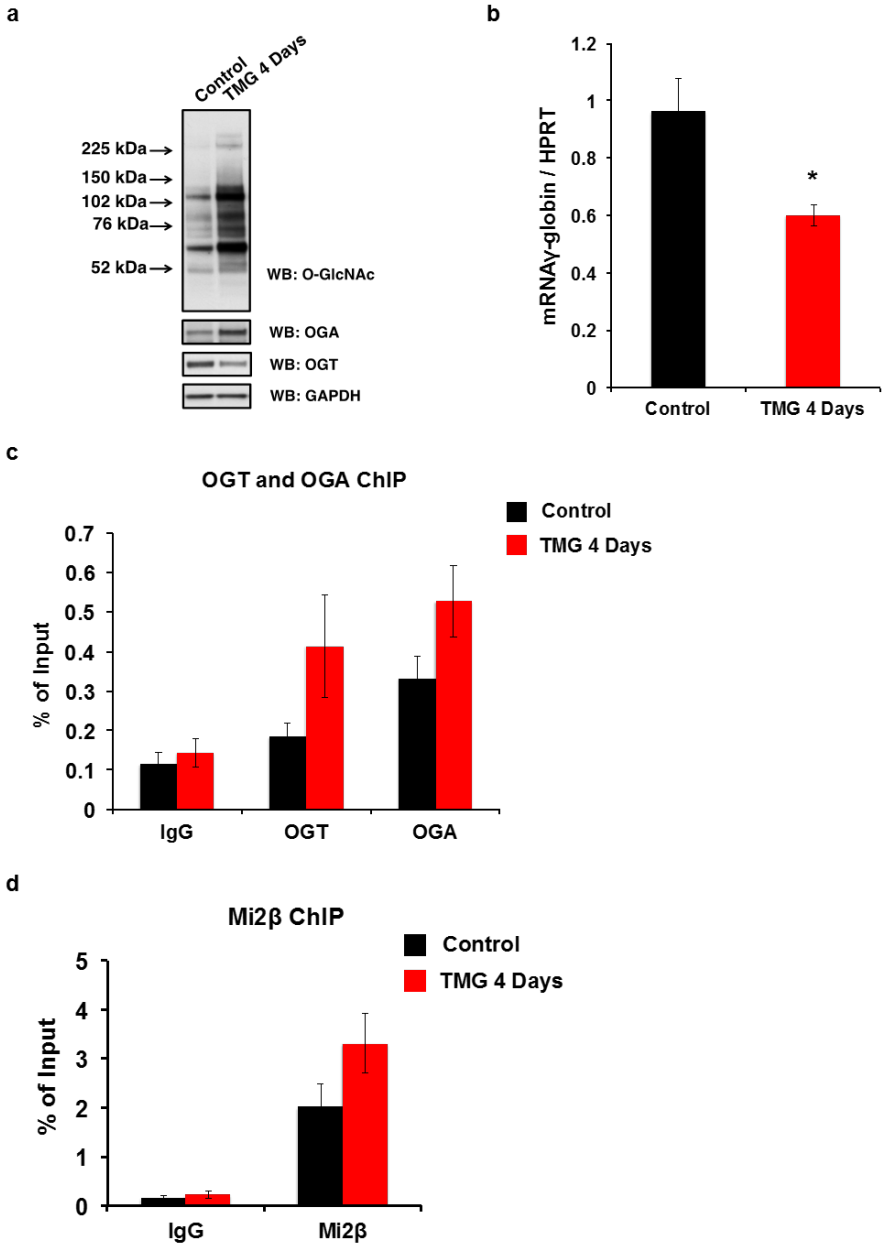

Figure 7:OGT, OGA, and Mi2β increase at the ^{A}γ-globin gene promoter after TMG treatment in WT β-YAC BMCs. **a**, overall O-GlcNAc levels and OGA and OGT protein expression in total cell lysates from control and 4-day TMG-treated WT β-YAC BMCs cells were analyzed by immunoblotting. GAPDH was used as a loading control. **b**, γ-globin mRNA level in control and 4-day TMG-treated WT β-YAC BMCs cells was analyzed by qPCR. Mouse HPRT was used as an internal control. OGT/OGA (**c**) and Mi2β (**d**) ChIP assays were performed on control and 4-day TMG-treated WT β-YAC BMCs cells, respectively. ChIP DNA was analyzed by qPCR using a set of primers targeting the −566 GATA site of the ^{A}γ-globin promoter. Normal rabbit IgG served as a negative control. All experiments were performed with four biological replicates (* indicates $p < 0.05$, Student's t test). Error bars represent S.E. WB, Western blotting.

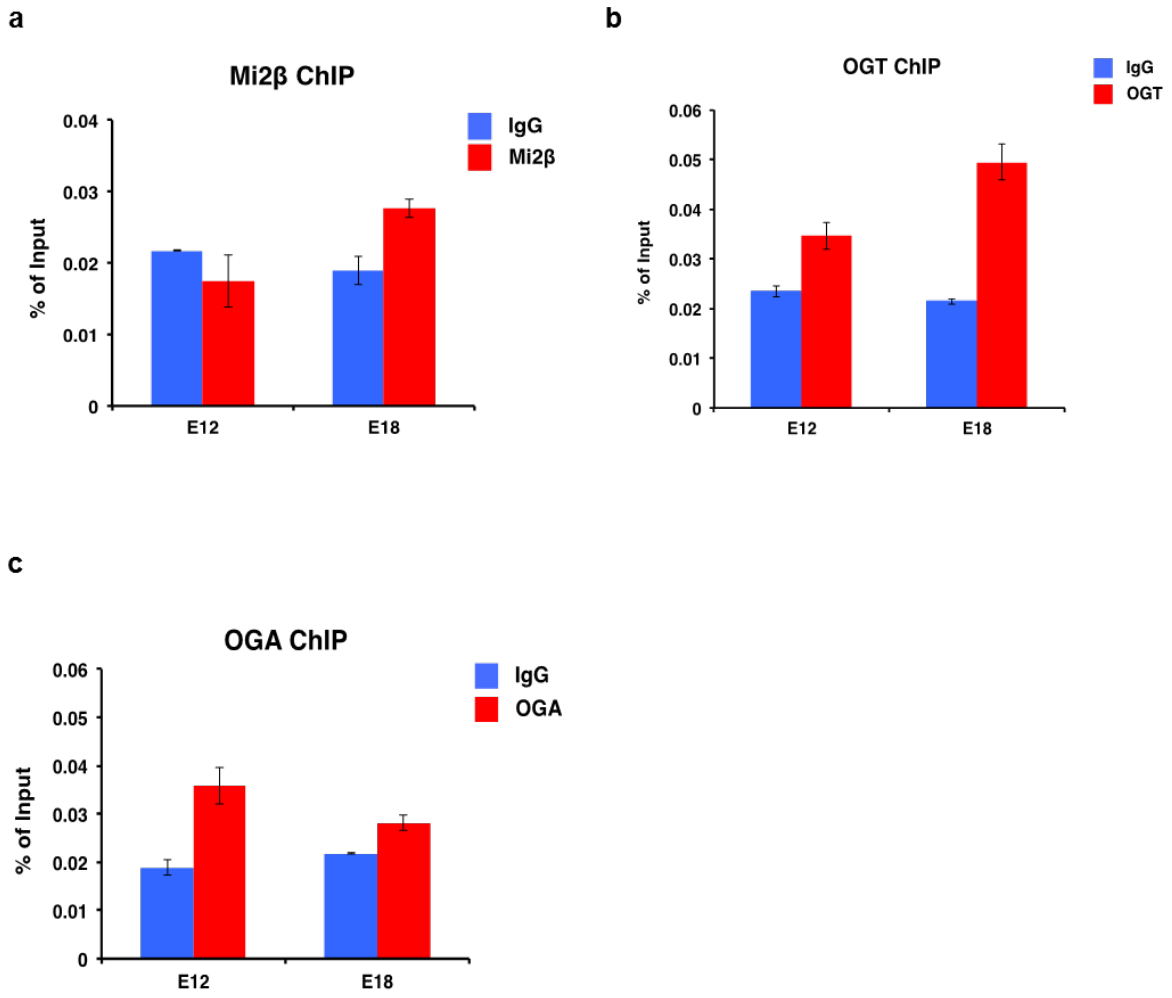

Figure 8:OGT and OGA interact with the $^{A}\gamma$-globin promoter during fetal liver development in β-YAC transgenic mice. Mi2β (**a**), OGT (**b**), and OGA (**c**) ChIP assays were performed using postconception day E12 (γ-globin is expressed) and E18 (γ-globin is repressed) fetal liver single cell suspension prepared from β-YAC transgenic mouse conceptuses. ChIP DNA was analyzed by qPCR using a set of primers targeting the −566 GATA site of the $^{A}\gamma$-globin promoter. Normal rabbit IgG served as a negative control. All experiments were performed with at least three biological replicates. Error bars represent S.E.

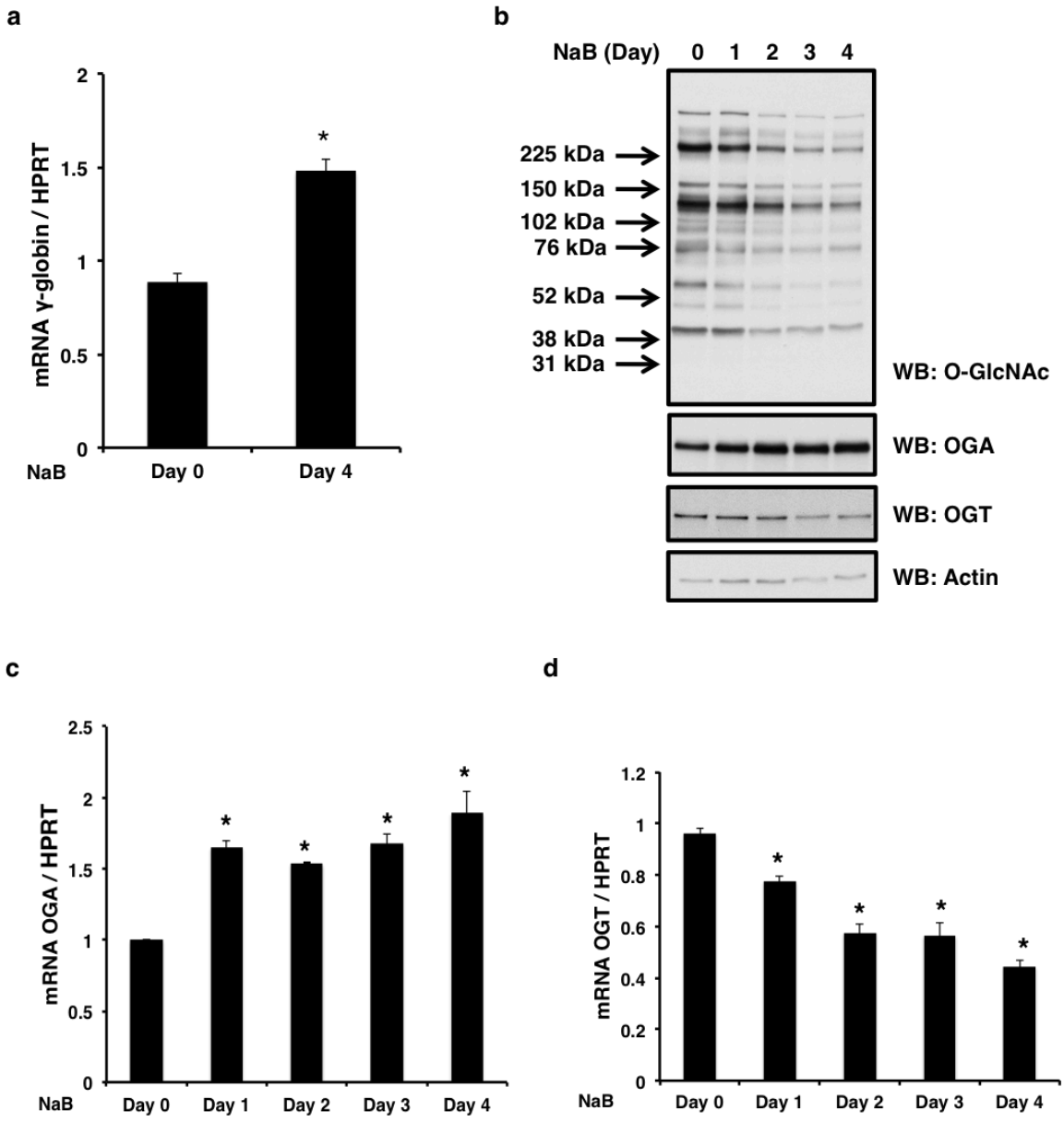

Figure 9:K562 γ-globin expression is increased after NaB induction. K562 cells were treated with NaB for 4 days. Cells were collected daily. **a**, γ-globin mRNA levels were measured before and after NaB induction by qPCR. **b**, overall O-GlcNAc levels and OGA and OGT protein expression from total cell lysates were analyzed by immunoblotting. Actin was used as a loading control. OGA (**c**) and OGT (**d**) mRNA levels were measured before and after NaB induction by qPCR. Human HPRT was used as an internal control in all qPCR analyses. All experiments were performed with at least three biological replicates (* indicates $p < 0.05$, Student's t test). Error bars represent S.E. WB, Western blotting.

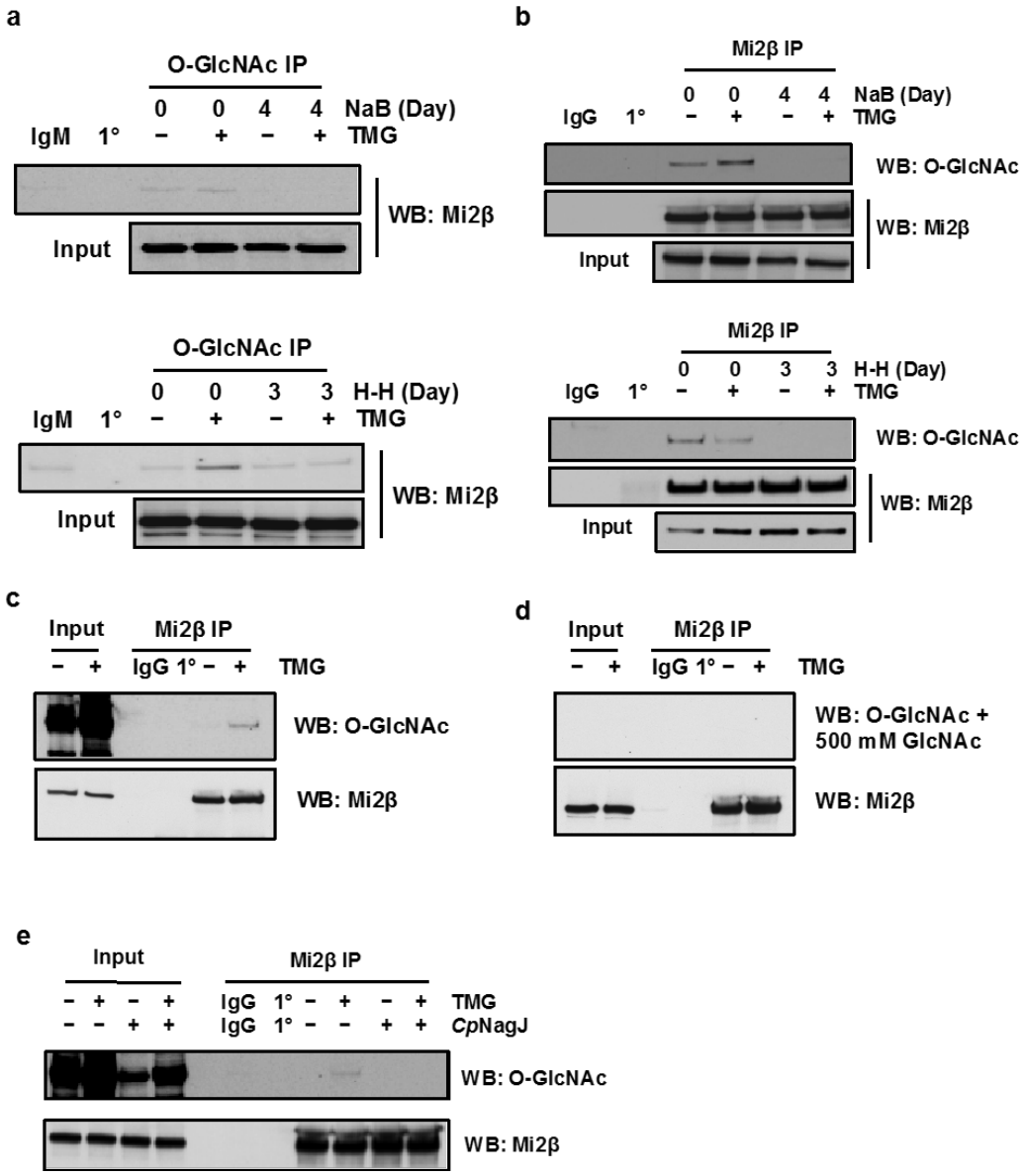

Figure 10:Mi2β is modified by O-GlcNAc. **a**, O-GlcNAcylated protein was immunoprecipitated by an O-GlcNAc-specific antibody (CTD 110.6) in K562 cells before and after NaB induction (top panel) or H-H induction (bottom panel); the blot was probed with Mi2β antibody. **b**, Mi2β protein was immunoprecipitated from K562 cells before and after NaB induction (top panel) or H-H induction (bottom panel); the blot was probed with CTD 110.6. c and d, Mi2β protein was immunoprecipitated from K562 cells, and the blot was probed with CTD 110.6 (**c**) or CTD 110.6 preincubated with 500 mM GlcNAc (**d**). **e**, Mi2β protein was immunoprecipitated from K562 cells and treated with hexosaminidase CpNagJ at 37 °C for 3 h; the blot was probed with CTD 110.6. The OGA inhibitor TMG was used to increase overall O-GlcNAc levels. Isotype immunoprecipitation (normal mouse IgM or rabbit IgG) and antibody-alone immunoprecipitation (1°) were used as negative controls. All experiments were performed with at least three biological replicates. WB, Western blotting.

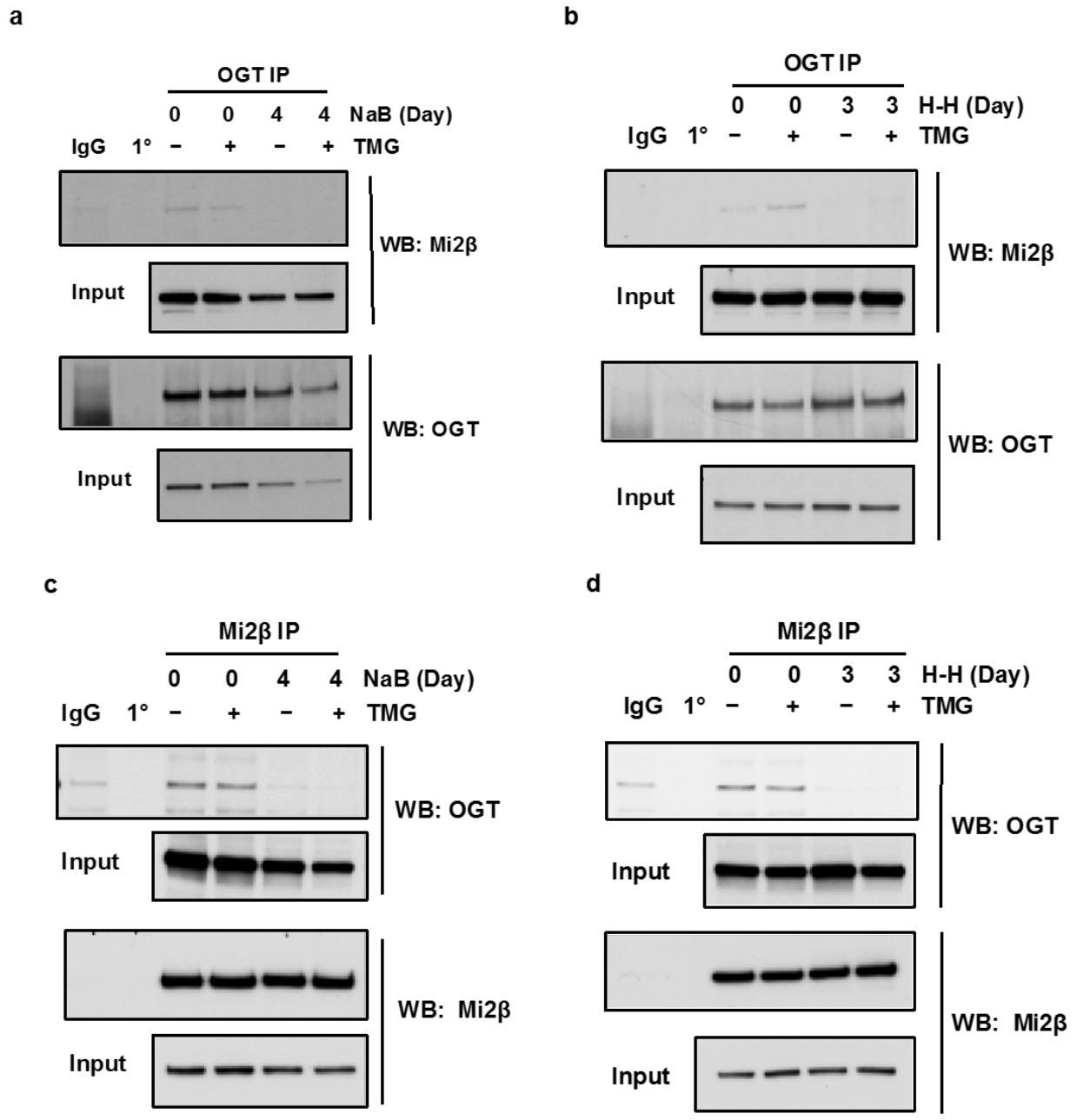

Figure 11:Mi2β interacts with OGT. a and b, OGT was immunoprecipitated with an OGT-specific antibody (AL-34) in K562 cells before and after NaB induction (**a**) or hemin and HMBA induction (**b**), and the blot was probed with Mi2β antibody. c and d, Mi2β protein was immunoprecipitated from K562 cells before and after NaB induction (**c**) or hemin and HMBA induction (**d**), and the blot was probed with OGT antibody. The OGA inhibitor TMG was used to increase the overall O-GlcNAc levels. Isotype immunoprecipitation (normal rabbit or mouse IgG) and antibody-alone precipitation (1°) were used as negative controls. All experiments were performed with at least three biological replicates. WB, Western blotting.

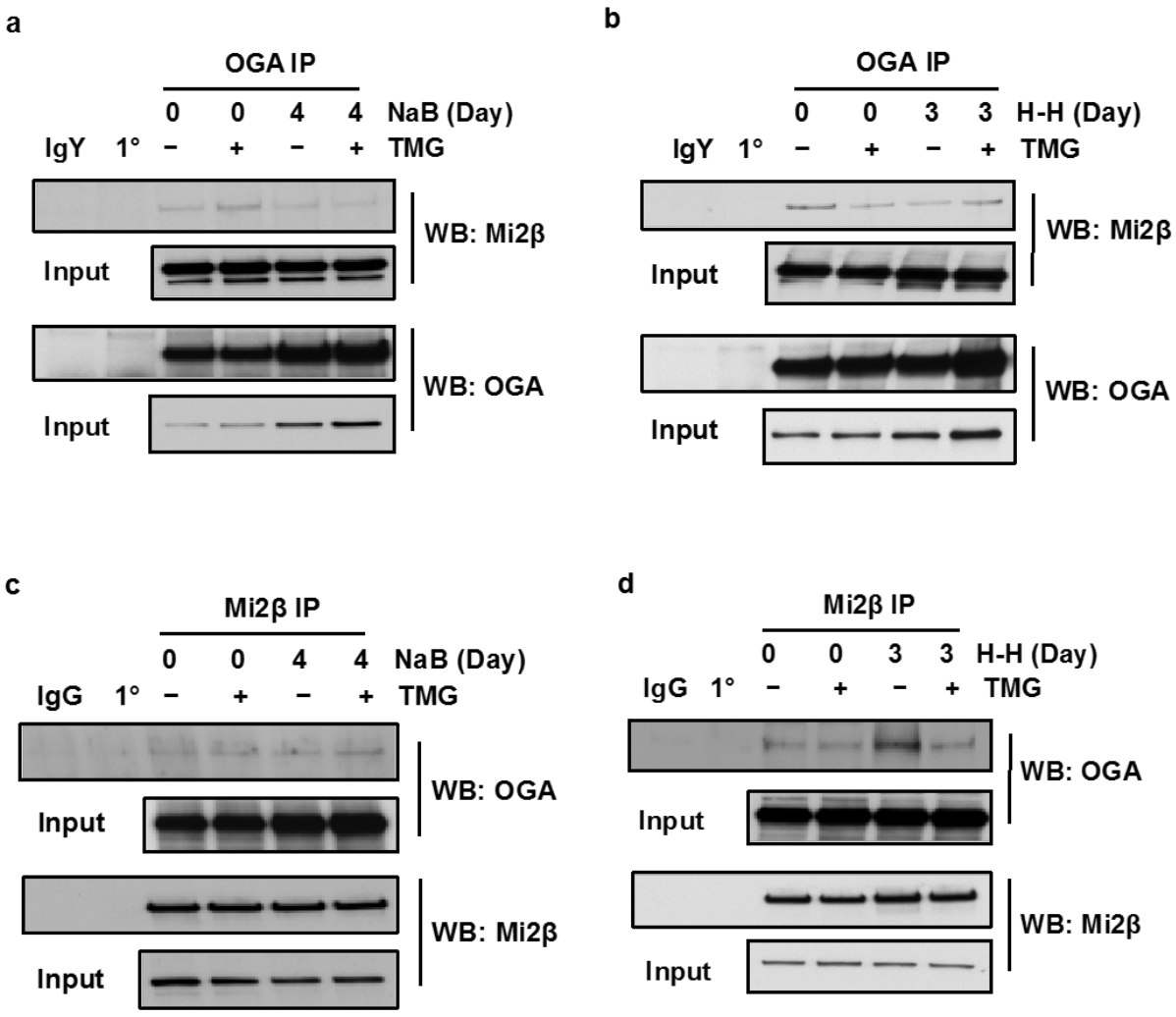

Figure 12:Mi2β interacts with OGA. **a** and **b**, OGA was immunoprecipitated using an OGA-specific antibody in K562 cells before and after NaB induction (**a**) or hemin and HMBA induction (**b**), and the blot was probed with Mi2β antibody. c and d, Mi2β protein was immunoprecipitated from K562 cells before and after NaB induction (**c**) or hemin and HMBA induction (**d**), and the blot was probed with OGA antibody. The OGA inhibitor TMG was used to increase overall O-GlcNAc levels. Isotype immunoprecipitation (normal chicken IgY or rabbit IgG) and antibody-alone immunoprecipitation (1°) were used as negative controls. All experiments were performed with at least three biological replicates. WB, Western blotting.

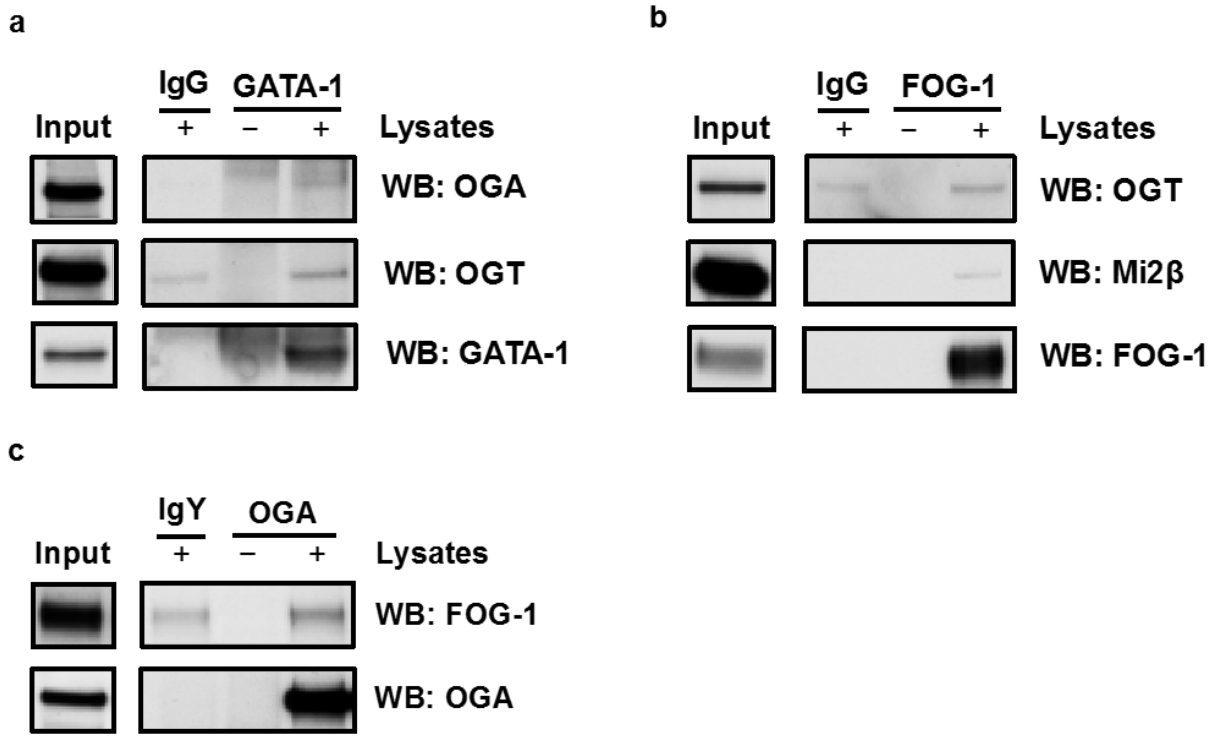

Figure 13:OGT and OGA interact with GATA-1 and FOG-1 in MEL birA cells. **a**, GATA-1 was immunoprecipitated using a GATA-1-specific antibody in MEL birA cells, and the blot was probed with OGA and OGT antibodies. **b**, FOG-1 was immunoprecipitated from MEL birA cells, and the blot was probed with OGT and Mi2β antibodies. **c**, OGA was immunoprecipitated from MEL birA cells, and the blot was probed with FOG-1 antibody. Isotype (IgY or IgG) immunoprecipitation and antibody-alone immunoprecipitation (without lysates) were used as negative controls. All experiments were performed with three biological replicates. WB, Western blotting.

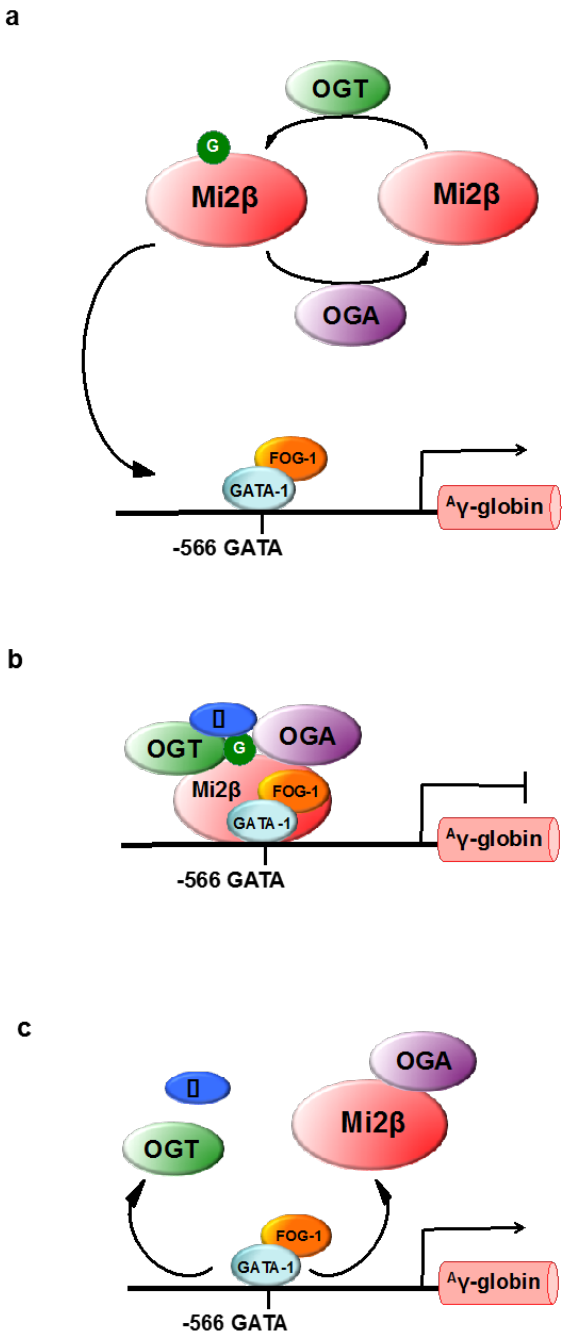

Figure 14:Proposed mechanism of OGT/OGA regulation of GATA-1·FOG-1·Mi2β mediated $^{A}\gamma$-globin repression. **a**, O-GlcNAc (G) cycling on Mi2β is processed by OGT and OGA. Potentially, O-GlcNAcylated Mi2β is recruited to −566 of $^{A}\gamma$-globin promoter with OGT and OGA by GATA-1 and FOG-1. **b**, once O-GlcNAcylated Mi2β is recruited to the promoter, other potential cofactors (blue oval with "□") are recruited to the promoter as well to form a stable repressor complex, repressing $^{A}\gamma$-globin transcription. **c**, when $^{A}\gamma$-globin transcription is activated, Mi2β does not interact with OGT or other cofactors, and OGA removes O-GlcNAc from Mi2β, dissembling the repressor complex.

Chapter 4: O-GlcNAc Regulates GATA-1 Targeted Erythroid Gene Transcription

Erythropoiesis is the process in which erythroid progenitors proliferate and differentiate into reticulocytes. Although erythropoiesis is regulated by numerous transcription factors, GATA-1, a transcription factor bearing a dual-zinc finger and WGATAR DNA binding motif, is the master regulator of erythroid differentiation (105-108). During erythroid differentiation, many GATA-1 target genes are activated or repressed (45,46,109,110). How GATA-1 mediates gene transcription activation versus repression is not well understood. For example, the co-occupancy of GATA-1, Scl/TAL1, and other erythroid transcription factors, and the recruitments of transcription co-activator or repressor at specific gene loci, determine the GATA-1 target gene transcription status to some extend (41,42,45,111). Epigenetic changes of histones and variation of WGATAR adjacent motif also play a role (40,43,45,46,112). These mechanisms provide a detailed view of DNA occupancy by GATA-1 with other transcription factors, and show that changes to chromatin structure during erythropoiesis influence GATA-1 function, however, these mechanisms do not explain the full breadth of GATA-1 regulation. Potentially, post-translational modifications (PTM) of transcription complexes during erythropoiesis could regulate GATA-1 function such as changing protein-protein interactions of transcription co-activators or repressors or chromatin structure. We contend that O-GlcNAcylation is one of the PTMs contributing to GATA-1 regulation.

O-linked β-N-acetylglucosamine (O-GlcNAc) is the modification of serine and threonine residues by β-D-N-acetylglucosamine. The sugar moiety is added and removed from serine and threonine residues by the O-GlcNAc processing enzymes, O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA), respectively (82). O-GlcNAc is critical for transcriptional regulation including both transcriptional activation and repression. For example, TET (ten-eleven

translocation) proteins and HCF-1 (host cell factor 1) co-localize with OGT at gene promoters associated with activating histone marks to promote target gene transcription (19,29,113,114). OGT also interacts with mSin3A and components of polycomb repressive complex (PRC) 1 and 2 to mediate transcription repression by forming stable repressor complexes (27,115,116). OGA can localize to transcription start sites (TSS) overlapping with RNA polymerase II (117). In our previous study, we also demonstrated that OGT and OGA interact with Mi2β, a component of GATA-1·FOG-1·Mi2β repressor complex, at the $^{A}\gamma$-globin promoter. O-GlcNAcylation of Mi2β might modulate GATA-1·FOG-1·Mi2β repressor complex formation to regulate $^{A}\gamma$-globin transcription (118).

In this study, we asked the question whether O-GlcNAc regulates GATA-1 targeted gene transcription during erythropoiesis. We utilized a well-established cell model of erythropoiesis, G1E-ER4 cells, a murine GATA-1 null erythroblast line that undergoes erythroid differentiation when GATA-1 activity is restored by the addition of β-estradiol (E2) (109,119). Interestingly, overall O-GlcNAc levels dramatically decreased upon GATA-1 activation with a slight increase in OGA and a decrease in OGT protein expression, and the interactions between GATA-1-ER and OGT/OGA were enhanced. We performed Next Generation RNA-Sequencing analysis on G1E-ER4 cells treated with E2 and/or Thiamet-G (TMG, an OGA inhibitor) and identified transcriptional alterations of 433 GATA-1 target genes when cells were treated with E2+TMG compared to E2 treatment only. We selected the GATA-1 target gene *Laptm5* to investigate GATA-1 and OGT/OGA occupancy at the GATA-1 binding site during erythropoiesis. LAPTM5 (lysosomal protein transmembrane 5) is specifically expressed on hematopoietic cells (120) and positively modulates inflammatory signaling pathways and cytokine secretion in macrophages (121). Our ChIP data demonstrated that the occupancy of GATA-1 and OGT/OGA, and the

overall O-GlcNAc level at *Laptm5* GATA binding site decreased when OGA is inhibited by TMG during erythropoiesis. Our data suggests that O-GlcNAc regulates specific GATA-1 target genes during erythropoiesis, OGA inhibition leads to an amplification of GATA-1 activation or repression, and suggest O-GlcNAcylation is a mechanism to control GATA-1 target gene transcription.

4.1 Material and Methods

Antibodies

Primary antibodies and secondary antibodies for immunoblotting were used at 0.5 μg/ml and 1:10,000 dilution respectively. Antibodies for immunoprecipitation (IP) and chromatin immunoprecipitation (ChIP) assay were used at 2 μg per reaction. Antibodies for immunoblotting: Anti-OGT (AL-34) and OGA (345) were gracious gifts from the Laboratory of Gerald Hart in the Department of Biological Chemistry at the Johns Hopkins University School of Medicine. Anti-O-linked N-acetylglucosamine antibody (RL2) (12440061), also used for IP and ChIP, and rabbit anti-goat IgG HRP (31402) were purchased from ThermoFisher Scientific. Anti-GAPDH antibody (ab9484) and anti-GATA2 antibody (ab109241) were purchased from Abcam. Anti-chicken IgY HRP (A9046) was purchased from Sigma. Goat anti-rabbit IgG HRP (170-6515) and goat anti-mouse IgG HRP (170-6516) were purchased from Bio-Rad. Goat anti-rat IgG HRP (NA935V) was purchased from GE Healthcare. GATA-1 antibody (sc-265 X), also used for IP, and FOG antibody (sc-9361 X) were purchased from Santa Cruz Biotechnology. Antibodies for IP: Normal rabbit IgG (sc-2027), normal mouse IgG (sc-2025), and rat IgG (sc-2026) were purchased form Santa Cruz Biotechnology. Anti-OGT (DM-17) antibody (O6264) and anti-OGA antibody (SAB4200267) were purchased from Sigma. Antibodies for ChIP: Rabbit control IgG (ab46540) and anti-GATA1 antibody (ab11852) were purchased from

Abcam. ChIP grade mouse (G3A1) mAb IgG1 isotype control (5415) was purchased from Cell Signaling Technologies. OGT antibody (61355) was purchased from Active Motif. Anti-OGA antibody (SAB4200267) was purchased from Sigma.

Cell Culture

G1E (GATA-1^{-} Erythroid) and G1E-ER4 cells are kind gifts from Dr. Soumen Paul (The University of Kansas Medical Center). G1E is a GATA-1 null hematopoietic cell line derived from murine ES cells (119). G1E-ER4 cell line is a G1E designated subclone stably expressing fusion product of GATA-1 and human estrogen receptor ligand binding domain (GATA-1-ER) (109). G1E and G1E-ER4 cells were cultured in IMDM (12440061, ThermoFisher Scientific) supplemented with 15% heat-inactivated FBS (Gemini), 100 U/ml penicillin and 100 μg/ml streptomycin (P4333, Sigma), 120 nM 1-thioglycerol (M6145, Sigma), 2.5 U/ml recombinant mouse erythropoietin (rm EPO, 587606, Biolegend), and 50 ng/ml recombinant mouse stem cell factor (rm SCF, 78064.2, Stemcell Technology). Puromycin was added to the G1E-ER4 medium at 1 μg/ml to maintain selection pressure. Cells were incubated at 37°C, 5% CO_2 in a 95% humidified incubator. Cells were seeded at 4 x 10^4 cells/ml in T25 flasks with 10 ml culture medium overnight. Cells were then treated with 10 μM Thiamet-G (TMG, S.D. Specialty Chemicals) or/and 1 μM β-estradiol (E2, E2758, Sigma) (122). After 30 hrs, cells were harvested for subsequent analysis.

Western Blot

Cells were lysed on ice as described previously (118). Fifty μg of total cell lysates were used for electrophoresis and subjected to Western blot as previously described (118). All Western blot results were repeated in three independent experiments and representative images are shown.

Immuniprecipitation (IP)

IP was performed as described previously (118). Two mg of cell lysates and 2 μg of antibody were used in each reaction. Forty μg of cell lysates were used as input.

Benzidine Staining and May-Grünwald/Giemsa Stain

Benzidine Staining was performed to access the hemoglobin content as described previously (109). Benzidine reagent was prepared by dissolving 60 mg of o-diansidine (D9143, Sigma) in 29.7 ml H2O and 0.5 ml glacial acetic acid. Prior to staining, one part of hydrogen peroxide (216763, Sigma) was added into 10 parts benzidine reagent.

A hundred thousand cells were collected and resuspended with 100 μl of 1X PBS, followed by incubating with 10 μl of fresh benzidine preparation for 2 min. Cells were then loaded onto cytospin chambers (10-356, Fisher Scientific) and spun by a Shandon™ CytoSpin™ Centrifuge for 3 minutes at 700 rpm. Slides were air-dried, counterstained in May-Grünwald (MG1L, Sigma) for 5 seconds, and quick-washed twice in distilled water. Positive cells will be black, while negative will be light purple.

May-Grünwald/Giemsa Stain was performed to access the morphology changes (109). Slides were prepared as described above, stained in May-Grünwald for 2 min, in fresh Giemsa (48900, Sigma, 1:20 dilution from stock) for 10 min, and quick-washed twice in distilled water. Cells staining images were visualized and taken under Nikon Eclipse 80i digital microscopy (Nikon Instruments INC, Melville, NY).

Total RNA Isolation and RT-PCR

Total RNA was isolated from 5 x 10^6 cells using TRI Reagent (T9424, Sigma) according to the manufacturer's instructions. For reverse transcription (RT), 0.5 μg of total RNA was reverse-transcribed to cDNA by iScript Reverse Transcription Supermix (170-8841, Bio-Rad)

following the manufacturer's instructions. Twenty µl of each completed reaction mix was incubated in a thermal cycler (Model 2720, Applied Biosystems) using the following protocol: priming 5 min at 25°C, RT 20 min at 46°C, RT inactivation 1 min at 95°C, and hold at 4°C.

RNA-Sequencing

RNA was extracted by Tri Reagent from three independent experiments. RNA sequencing library was prepared using Illumina TruSeq Stranded mRNA Library Preparation Kit Set A (RS-122-2101, Illumina) according to manufacturer's instruction. Eight hundred ng of total RNA was used for the library preparation. Total RNA and sequencing library were analyzed by Agilent 2100 Bioanalyzer for quality control at KUMC Genome Sequencing Facility prior to sequencing. RNA library was sequenced at two lane flow cell Rapid Run mode using Illumina HiSeq 2500 Sequencing System composed of the HiSeq2500 sequencer, cBOT Clustering Station and the IlluminaCompute (iCompute) Server.

RNA-Sequencing Data Analysis

The FastQC (0.11.2) and RSEM (1.2.22) softwares were used to assess the quality of the RNA sequencing results, and align the reads to the mouse genome reference GRCm38/mm10 and calculate the gene expression values, respectively. R (3.2.2) and EdgeR (3.14.0) softwares were used to normalize the expression values using the TMM-method (weighted trimmed mean of M-values), followed by differential expression analysis. First, we maximized the negative binomial conditional common likelihood to estimate a common dispersion value across all genes (estimateCommonDisp). Next, we estimated tagwise dispersion values by an empirical Bayes method based on weighted conditional maximum likelihood (estimateTagwiseDisp). Finally, we calculated the differentially gene expression by computing the genewise exact tests for differences in the means between two groups of negative-binomially distributed counts. To

reduce the burden of multiple testing in differential gene expression analyses, a filter was initially applied to reduce the number of genes. Genes were removed if they did not present a meaningful gene expression across all samples; only genes with cpm (counts per million) of >10 in least two samples were considered in differential expression analyses. The Benjamini and Hochberg procedure was used to control the false discovery rate (FDR). The following R-packages were utilized for calculations and visualizations: edgeR and Gplot.

Chromatin Immuniprecipitation (ChIP) Assay

The ChIP assay was performed using a previously described method (118,123) with slight modifications. Briefly, chromatin was prepared and sheared as previously described (118). For each ChIP reaction, one hundred µl of sheared chromatin was diluted with 500 µl IP buffer (150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1% Triton X-100, 40 mM GlcNAc, and 50 mM Tris-HCl, pH 7.5), supplemented with protease inhibitor, and incubated with 2 µg of control IgG or specific antibody, respectively, at 4°C overnight. Next day, 20 µl of PBS-washed Dynabeads Protein G (10004D, ThermoFisher Scientific) were added to the mixture, followed by rotation at 4°C for 2 hrs. Dynabeads were washed by 1 ml cold IP buffer 6 x for 5 min each wash, then mixed with 100 µl 10% Chelex 100 (1421253, Bio-rad) slurry; for input samples, 90 µl of 10% Chelex 100 was mixed with 10 µl of sheared chromatin. Briefly vortex ChIP samples and input samples and boil at 95 °C for 10 min, followed by treatment of 1 ul Rnase A (10 mg/ml) for 15 min and 1 µl proteinase K (20 mg/ml) for 30 min at 55 °C. Samples were boiled again for 10 min to inactivate proteinase K and centrifuged at 12,000 g for 1 min at 4 °C. Seventy µl of supernatant was transferred to a new tube then 130 µl of DEPC-treated water (AM9906, ThermoFisher Scientific) was added to the beads and mixed well, followed by centrifugation.

One hundred and thirty μl of supernatant was collected and pooled with the previous supernatant for a total of 200 μl. DNA samples were stored at -20°C for subsequent qPCR analysis.

Quantitative Polymerase Chain Reaction (qPCR)

qPCR was performed as described previously (118). Two μl of template for RT-qPCR or 9.6 μl of template for ChIP-qPCR, 10 μl of SsoAdvanced Universal SYBR Green Supermix (172-5271, Bio-Rad), and 0.2 μl of each forward and reverse primers (100 μM) were mixed together in a total reaction volume of 20 μl. The primer sequences for measuring target genes transcription levels are listed as following: *Actb* FWD TTTCCAGCCTTCCTTCTTGG, REV GGCATAGAG GTCTTTACGGATG; *Hbb-b1* FWD GTGAGCTCCACTGTGACAAGCT, REV GGTGGC CCAGCACAATCACGATC (35); *Fndc5* FWD CTCTGTCCCATTTCCTACCTTAC, REV GATGTTCACAGAGGGCTTCTT; *Laptm5* FWD GATGGCAAGCTGTCCTTATCT, REV GTGTGATCACCAGACCCTAAAT; *Mvb12a* FWD CCCAAGATTTGAGGGCAAGA, REV CCACCACGAAGCCATAGTTATAC; *Parp14* FWD CAGAGCATGGCCTTGAGATAA, REV CCAGTGCCAGAAAGGAGAAA. The primer sequences for measuring ChIP DNA quantity are listed as following: +8 kb *Laptm5* TSS (GATA-1 binding site) FWD GGTGTTGTGGTGGCATATAGA, REV CCGCTTCCTGTCACAGTTTA; +58 kb *Laptm5* TSS FWD CCCAGTGCTTTGTCTGTTACT, REV GCTCTGAAGTGATTCGAGGAAA. The reactions were run in a CFX96 Touch Real-Time PCR Detection System (185-5195, Bio-Rad) as described previously (118).

qPCR Data Analysis

Quantification cycle (Cq) values were calculated by CFX Manager™ software. For cDNA qPCR data, the dynamic range of RT and amplification efficiency were evaluated before applying the ΔΔCq method to calculate relative gene expression change. The transcription level

of the target gene was normalized to the internal control as fold change. For ChIP DNA qPCR data, the Cq value was normalized as percent of input. Data generated in at least three independent experiments are presented as mean ± standard error; the two-tailed Student *t*-test statistic was applied with $P < 0.05$ considered to be a significant difference.

Gene Set Enrichment Analysis (GSEA)

GSEA was performed according to the "GSEA USER GUIDE" (124,125). For gene ontology (GO) analysis, biological processes gene sets from Molecular Signature Database (MsigDB) were used. Pathways with NOM (Nominal) p-value < 5% and FDR q-value < 25% were included into the analysis.

4.2 Results

O-GlcNAc Levels Decrease upon GATA-1 Restoration

In order to determine the role of O-GlcNAc during erythropoiesis, we first used a well-established erythropoiesis model system. G1E (GATA-1^{-} erythroid) is a GATA-1 null hematopoietic cell line derived from murine ES cells (119). The G1E-ER4 cell line is a G1E designated subclone stably expressing GATA-1-ER, a fusion product of GATA-1 and human estrogen receptor ligand binding domain (109). Upon restoring GATA-1 activity by adding β-estradiol (E2) for 30 hrs, G1E-ER4 cells undergo red blood cell differentiation and recapitulate a stage from late BFU-E (burst forming unit-erythroid) to basophilic erythroblasts (109,122), characterized by decreased cell size, hemoglobin accumulation, nuclear condensation, and cell membrane organization. After GATA-1 restoration by E2 treatment, GATA-1-ER expression was increased as reported previously (**Figure 15a**) (126,127); GATA-2 transcription is repressed by GATA-1, resulting in the decrease of GATA-2 protein level (43) (**Figure 15a**); GATA-1 target gene FOG-1 protein level (**Figure 15a**) and β-major globin gene transcription level

(**Figure 15e**) were increased. At the same time, the morphologic maturation (**Figure 15c**) and hemoglobin induction (**Figure 15d**) were observed in G1E-ER4 cells treated with E2. In order to determine how changes in O-GlcNAcylation could affect GATA-1 induced differentiation, we treated cells with both E2 and TMG, an OGA inhibitor. TMG+E2 did not change the cell morphology (**Figure 15c**), hemoglobin induction (**Figure 15d**), and β-major globin gene transcription level (**Figure 15e**) compared to E2 treatment only. Interestingly, the O-GlcNAc levels were dramatically decreased after E2 treatment, with an increase of OGA and a decrease of OGT protein expression. When G1E-ER4 cells were treated with E2 following a time course, we found that the O-GlcNAc levels at 30 hrs were the lowest compared to other time points (**Figure 15b**). However, the O-GlcNAc levels, and OGT/OGA expression in G1E cells did not change at 30 hrs compared to 0 hr (**Figure 15a and b**). These data indicate that O-GlcNAc levels decrease upon GATA-1 restoration and erythroid differentiation.

GATA-1 Interacts with OGT and OGA in G1E-ER4 Cells

Our previous study demonstrated that OGT and OGA interact with GATA-1·FOG-1·Mi2β repressor complex, and Mi2β-OGT/OGA interaction pattern changes when γ-globin transcription level is altered (118). To address how OGT/OGA interact with GATA-1 during G1E-ER4 differentiation, we performed immunoprecipitation (IP) using antibodies against O-GlcNAc (RL2), OGT, OGA, and GATA-1 from G1E-ER4 cell lysates treated with TMG or/and E2 (**Figure 16**). We found that little GATA-1-ER could be pulled down by O-GlcNAc, OGT, or OGA antibody before E2 treatment, however, O-GlcNAc, OGT, and OGA antibody all can robustly IP GATA-1-ER after E2 treatment (**Figure 16a-c**); these antibodies can also pull down FOG-1 before and after E2 treatment, and E2 treatment increased the amount of immunoprecipitated FOG-1 with a molecular weight shift suggesting an increase in FOG-1

phosphorylation (128,129) (**Figure 16a-c**). GATA-1 antibody can IP OGT/OGA before and after E2 treatment (**Figure 16d**). We also performed co-IP to test if GATA-1 or FOG-1 could be modified by O-GlcNAc, however, we could not detect the O-GlcNAcylation of GATA-1 or FOG-1 by Western blot (**data not shown**). These data indicate that the interaction between GATA-1 and OGT/OGA, and the O-GlcNAcylation of GATA-1 interacting proteins increase during G1E-ER4 cell differentiation.

Inhibition of OGA Changes Erythroid Gene Transcription Network

Since O-GlcNAc plays an important role in transcription regulation and OGT/OGA interact with GATA-1, we tested if the erythroid gene transcription network was altered by disruption of O-GlcNAcylation during G1E-ER4 differentiation. We performed Next Generation RNA-seq analysis on G1E-ER4 cells without any treatment, and with TMG and/or E2 treatment for 30 hrs. 8,271 transcripts were included into the RNA-seq analysis from the different treatments, and the shared gene numbers of down-regulated (**Figure 17a**) and up-regulated genes (**Figure 17b**) are showed in Venn diagram. Our RNA-seq data suggest that inhibition of OGA by TMG changes transcription levels of numerous genes during erythropoiesis. In order to characterize the biological processes altered by inhibition of OGA, GSEA (Gene Set Enrichment Analysis) analysis of the 8,271 transcripts was performed (124,125). Top 10 up- (top panel) and down-regulated (bottom panel) biological processes (**Figure 17c**) and examples of enrichment plots (**Figure 17d**) were listed respectively, comparing TMG+E2 with E2 treatment only. Interestingly, biological processes of myeloid leukocyte activation and inflammatory response were up-regulated, which indicated that inhibition of OGA during erythropoiesis might activate some leukocyte gene programs usually repressed during erythroid differentiation. These results indicate that many biological processes were changed when OGA was inhibited during

erythropoiesis, and these changes might enable erythroid cells to gain some leukocyte characteristics during differentiation.

We compared the transcription profile of G1E-ER4 cells treated with E2+TMG with E2 only (**Figure 18a).** 1,173 transcripts were differentially expressed in E2+TMG compared to E2 treatment only, and 530 genes were up-regulated while 643 down-regulated (**Figure 18b-c**). These data demonstrated that the G1E-ER4 transcriptome was altered with inhibition of OGA during differentiation.

O-GlcNAc Regulates GATA-1 Target Gene Transcription during Erythropoiesis

GATA-1 is a master transcription factor during erythropoiesis and is important for erythroid maturation (106,130). Our data demonstrated that inhibition of OGA can alter gene transcription profiles during erythropoiesis; therefore, O-GlcNAc could regulate GATA-1 target gene transcription. We used an online database Harmonizome (131) to search for GATA-1 target genes in G1E-ER4 after E2 treatment. Totally 4,072 GATA-1 target genes were found from ENCODE transcription factor binding site profiles (45,132), and 433 of them are differentially expressed in TMG+E2 compared with E2 only (**Figure 18d**). Of these 433 genes, 243 genes were up-regulated and 190 genes down-regulated (**Figure 18d**). Similar GSEA analysis was performed on these 433 genes. Interestingly, immune response and defense response biological processes, which are characteristics of leukocytes, were also up-regulated in TMG+E2 treated cells compared to E2 only (**data not shown**). Our data demonstrated that a subset of GATA-1 target genes is regulated by O-GlcNAc during erythropoiesis.

In order to assess the quality of our RNA-Seq data analysis and confirm the differential expression of candidate genes, we selected 4 genes to perform qPCR analysis. *Laptm5*, *Fndc5*, *Parp14*, and *Mvb12a* are all GATA-1 target genes, which also have a robust transcriptional

change with TMG+E2 treatment compared to E2 only. Total RNA was extracted from cells with control, TMG or/and E2 treatment, and qPCR analysis was performed. After E2 treatment for 30 hrs, *Laptm5*, *Fndc5*, and *Parp14* transcription level was increased compared to control (**Figure 19a-c**), while *Mvb12a* decreased (**Figure 19d**). When cells were treated with TMG+E2, *Laptm5*, *Fndc5*, and *Parp14* transcription level was increased compared to E2 only (**Figure 19a-c**), while *Mvb12a* decreased (**Figure 19d**). These results were consistent with the fold change from RNA-Seq data.

Inhibition of OGA Decreases the Occupancy of GATA-1, OGT/OGA, and O-GlcNAc level at the *Laptm5* GATA Binding Site

Since inhibition of OGA by TMG altered a subset of GATA-1 target genes during erythropoiesis, we explored the association of O-GlcNAc, OGT, and OGA at GATA-1 target gene *Laptm5*. LAPTM5 is specifically expressed in hematopoietic cells (120) and positively modulates inflammatory signaling pathways and cytokine secretion in macrophages (121). The *Laptm5* gene has a GATA binding site located at the first intron, approximately 8 kb downstream of TSS (45). We performed ChIP assays using antibody against GATA-1, OGT, OGA, and O-GlcNAc to assess the occupancy of GATA-1, OGT, OGA, and overall O-GlcNAc level at the GATA binding site during erythropoiesis.

In control G1E-ER4 cells, we saw the occupancy of OGT/OGA near *Laptm5* GATA binding site, which was also modified by O-GlcNAc (**Figure 20b-d)**, suggesting that O-GlcNAcylation near *Laptm5* GATA binding site might be important to maintain the *Laptm5* repression state. Upon GATA-1 restoration by E2, GATA-1 occupancy dramatically increased at the GATA binding site compared to control, however, its occupancy decreased at the same site when cells were treated with E2+TMG (**Figure 20a, left panel**); OGT occupancy did not change

and OGA occupancy increased at the GATA binding site after E2 treatment compared to control (**Figure 20b and c, left panel**), which might contribute to the decrease of the overall O-GlcNAc levels at the GATA binding site (**Figure 20d, left panel**). Interestingly, the occupancy of OGT, OGA, and overall O-GlcNAc level at GATA binding site decreased after TMG+E2 compared to E2 (**Figure 20b-d, left panel**), and the similar pattern was observed when comparing TMG treatment only with control (**Figure 20b-d, left panel**). In order to show that these changes were specific to the GATA binding site, we used a DNA region 50 kb downstream of the GATA binding site as a negative control. As expected, we did not find any changes comparing TMG+E2 with E2 only (**Figure 20a-d, right panel**). Our data demonstrated that the occupancy of GATA-1 and OGA, and overall O-GlcNAc level at *Laptm5* GATA binding site was changed upon GATA-1 restoration, and these changes could be disrupted by inhibition of OGA.

4.3 Discussion

We utilized a well-established erythropoiesis model (G1E-ER4 cells) to investigate the role of O-GlcNAc during GATA-1 mediated erythroid differentiation. After restoration of GATA-1 activity for 30 hrs, interestingly, overall O-GlcNAc levels dramatically decreased, with an increase of OGA and a decrease of OGT protein levels **(Figure 15a and b)**. Restoration of GATA-1 also promoted the interaction of GATA-1-ER with OGT and OGA **(Figure 16b-c)**. Next, we performed RNA-Seq analysis of G1E-ER4 cells treated with E2 and TMG. 1,173 genes changed transcription level compared to cells treated with E2 only **(Figure 18b)**, including 433 GATA-1 target genes **(Figure 18d)**. The transcription level changes were further confirmed by qPCR **(Figure 19)**. Our ChIP assay demonstrated that the occupancy of GATA-1 and OGA, and overall O-GlcNAc level at *Laptm5* GATA binding site were changed upon GATA-1 activation,

and these changes were diminished by inhibition of OGA **(Figure 20)**. Our data suggests that O-GlcNAc plays a role in regulating a subset of GATA-1 controlled erythroid genes.

O-GlcNAc regulates cell differentiation and development (133,134). The overall O-GlcNAc level decreases during embryonic stem cell (ESC) neural differentiation (135), epidermal keratinocytes (136), cardiomyocytes (137), myoblasts (138), and leukemia K562 cell differentiation (118); while overall O-GlcNAc levels increased during differentiation of mouse chondrocyte (139) and osteoblast cells (140). Here, we observed a dramatic decrease of overall O-GlcNAc levels in G1E-ER4 cells after E2 induction for 30 hrs, and even more reduction at 48 hrs (**data not shown**), with decreased OGT and increased OGA protein level (**Figure 15a and b**); however, the overall O-GlcNAc and OGT/OGA protein levels did not change in G1E cells at 30 hrs compared to control (**Figure 15a and b**). We do not know whether the decrease of overall O-GlcNAc levels is a characteristic of all hematopoietic linage differentiation or only specific for erythroid differentiation and how OGT and OGA are regulated to maintain the O-GlcNAc levels during erythropoiesis. Nor do we know whether this decrease is a driving force for differentiation or a result of differentiation. However, our data does recapitulate the data from differentiated cells lines showing a decrease in overall O-GlcNAcylation and suggesting that in most cells, differentiation requires a reduction in O-GlcNAc levels.

Numerous studies show that OGT, OGA, and O-GlcNAc are critical for transcription regulation (19,25,26,141). O-GlcNAc can modulate transcription activator or repressor complexes by the interactions of OGT/OGA with the components of protein complexes (27,118). O-GlcNAc is part of the histone code (18,19) and modulates other epigenetic changes on DNA or histones, such as DNA methylation (29,114), and histone acetylation (27) or methylation (114,115) affecting chromatin structure and leading to transcription activation or repression.

OGT and OGA regulates preinitiation complex (PIC) formation and RNA Pol II activity (24,117,142). Since GATA-1 expression increases with E2 treatment only, more GATA-1-ER was immunoprecipitated after E2 treatment by an O-GlcNAc, OGT, and OGA antibody respectively compared to control (**Figure 16a-c**). We did not find GATA-1 itself could be modified with O-GlcNAc by IP (**data not shown**); however, GATA-1 interacting proteins are likely modified by O-GlcNAc since upon differentiation, the O-GlcNAc antibody can pull down GATA-1-ER (**Figure 16a)** despite the decrease of overall O-GlcNAc levels in the total cell lysates compared to control (**Figure 15a**). The changes of OGT/OGA protein level, their interaction with GATA-1-ER, and O-GlcNAc levels of GATA-1-ER interacting proteins during erythropoiesis, could change protein-protein interaction of GATA-1 mediated activator or repressor complexes, chromatin structure, and RNA Pol II activity of GATA-1 target genes. All these events during erythropoiesis could contribute to GATA-1 mediated gene activation or repression.

In order to explore the role of O-GlcNAc during erythropoiesis, we pharmacologically inhibited OGA by TMG in G1E-ER4 with or without E2 treatment, and then performed RNA-seq analysis. Inhibition of OGA by TMG during G1E-ER4 differentiation leads to many erythroid genes being more activated or repressed compared to E2 treatment only. When we ran GSEA using our data set, interestingly, we found that biological process of myeloid leukocyte activation is significantly enriched, and tRNA metabolic process attenuated in TMG+E2 compared to E2 treatment only (**Figure 17c**). These data suggest inhibition of OGA by TMG changes gene transcription and protein synthesis profiles during erythropoiesis, and these changes might reprogram the cell and result in myeloid leukocyte-like cell morphology or behavior. However, it would be interesting to explore if O-GlcNAc plays a role in determining

hematopoietic cell fate at an earlier stage, for example, inhibition of OGA by TMG in hematopoietic stem cell stage before lineage commitment occurs and then determine if changes in O-GlcNAc skew hematopoietic stem cells differentiation toward a particular lineage.

We selected 1,173 differentially expressed genes (TMG+E2 compared to E2 only) and determined which of those genes were GATA-1 targeted using an online database (131). We found 433 differentially expressed genes under control of GATA-1 with 243 genes up-regulated and 190 down-regulated (**Figure 18d**). Interestingly, 84% of activated GATA-1 target genes became more activated in the presence of TMG; while 79% repressed GATA-1 target genes became more repressed in the presence of TMG, indicating that inhibition of OGA amplifies GATA-1 mediated transcription. Recently, we showed that in WT β-YAC bone marrow cells, transcription of $^{A}\gamma$-globin was repressed by GATA-1·FOG-1·Mi2β repressor complex, and TMG treatment in these cells led to even more repression of $^{A}\gamma$-globin transcription (118). These data point to a compelling phenomenon in which O-GlcNAc cycling at GATA-1 binding sites controls the strength of the transcriptional response. If O-GlcNAc cycling randomly promotes the formation of GATA-1 mediated activator or repressor complexes, we would anticipate TMG would amplify activation or repression of GATA-1 target genes randomly regardless of whether GATA-1 activates or represses these genes during erythropoiesis. Instead, we found that the transcription levels of most GATA-1 target genes were amplified when O-GlcNAc cycling was disrupted. These data argue that O-GlcNAc cycling does not dictate the activation or repression of GATA-1 target genes but finely tunes the actions of GATA-1 mediated activator or repressor complexes at GATA binding sites. Importantly, this might be a general phenomenon of O-GlcNAc cycling at gene promoters or other protein complexes and warrants further investigation.

One of the GATA-1 target genes is *Laptm5*, which is specifically expressed in

hematopoietic cells (120). LAPTM5 positively modulates inflammatory signaling pathways and cytokine secretion in macrophages (121). The transcription of *Laptm5* was activated upon GATA-1 restoration and was enhanced in the presence of TMG (**Figure 19a**). Strikingly, the occupancy of GATA-1, OGT/OGA, and O-GlcNAc level at *Laptm5* GATA binding site decreased after inhibition of OGA upon GATA-1 restoration (**Figure 20**). These changes could lead to the recruitment of other transcription co-activators or open chromatin structure at *Laptm5* GATA binding site resulting in an increase of *Laptm5* transcription compared to E2 treatment only. Even though GATA-1 occupancy at *Laptm5* GATA binding site was decreased in TMG+E2 treatment (**Figure 20a**), TMG amplified the increase of *Laptm5* transcription compared to E2 treatment only (**Figure 19a**). This suggests that TMG treatment could stabilize the interaction of GATA-1 with other transcription factors at the GATA binding site reducing the recruitment or turnover of these transcription factors. Hence a reduced occupancy of GATA-1 at the GATA binding site, but GATA-1 is either more stable or more active in promoting transcription. However, it is unknown what are these transcription factors at the *Laptm5* GATA binding site.

Based on our data, we propose that OGT and OGA are important in regulating GATA-1 target gene *Laptm5* transcription during erythropoiesis. In normal G1E-ER4 cells, *Laptm5* transcription is repressed by an O-GlcNAcylated repressor complex, which interacts with OGT/OGA near the GATA binding site (**Figure 21a**). Upon GATA-1 restoration by E2, GATA-1 is recruited to GATA site located in the first intron of *Laptm5*, followed by the recruitment of a transcription activator complex activating *Laptm5* transcription (**Figure 21a**). OGT interacts and modifies the component(s) of this transcription activator complex promoting the recruitment of other transcription co-activators. OGA then removes the O-GlcNAc to form a stable activator

complex. This process is similar to the preinitiation complex formation regulated by OGT and OGA (142). When OGA is inhibited by TMG during erythropoiesis, the O-GlcNAc cycling on the activator complex is disrupted, leading to the decreased occupancy of OGT, OGA, and GATA-1 at the GATA binding site. The components of the activator complex might also change (**Figure 21b**). However, this new activator complex could be more stable and active due to the changes in protein-protein interactions within the complex or chromatin structure resulting in even more active transcription of *Laptm5* (**Figure 21c**). These data suggest that inhibition of OGA by TMG could affect a subset of GATA-1 target gene transcription by enhancing the gene activation or repression.

In summary, we induced G1E-ER4 cells with E2 to investigate the role of O-GlcNAc during erythropoiesis. Our data suggests that O-GlcNAc plays an important role in regulating GATA-1 target genes transcription during erythropoiesis. Finally, we demonstrate that reduced O-GlcNAc cycling enhanced GATA-1 mediated gene activation or repression providing a novel mechanism as to how O-GlcNAcylation can promote GATA-1 targeted gene transcription activation or repression.

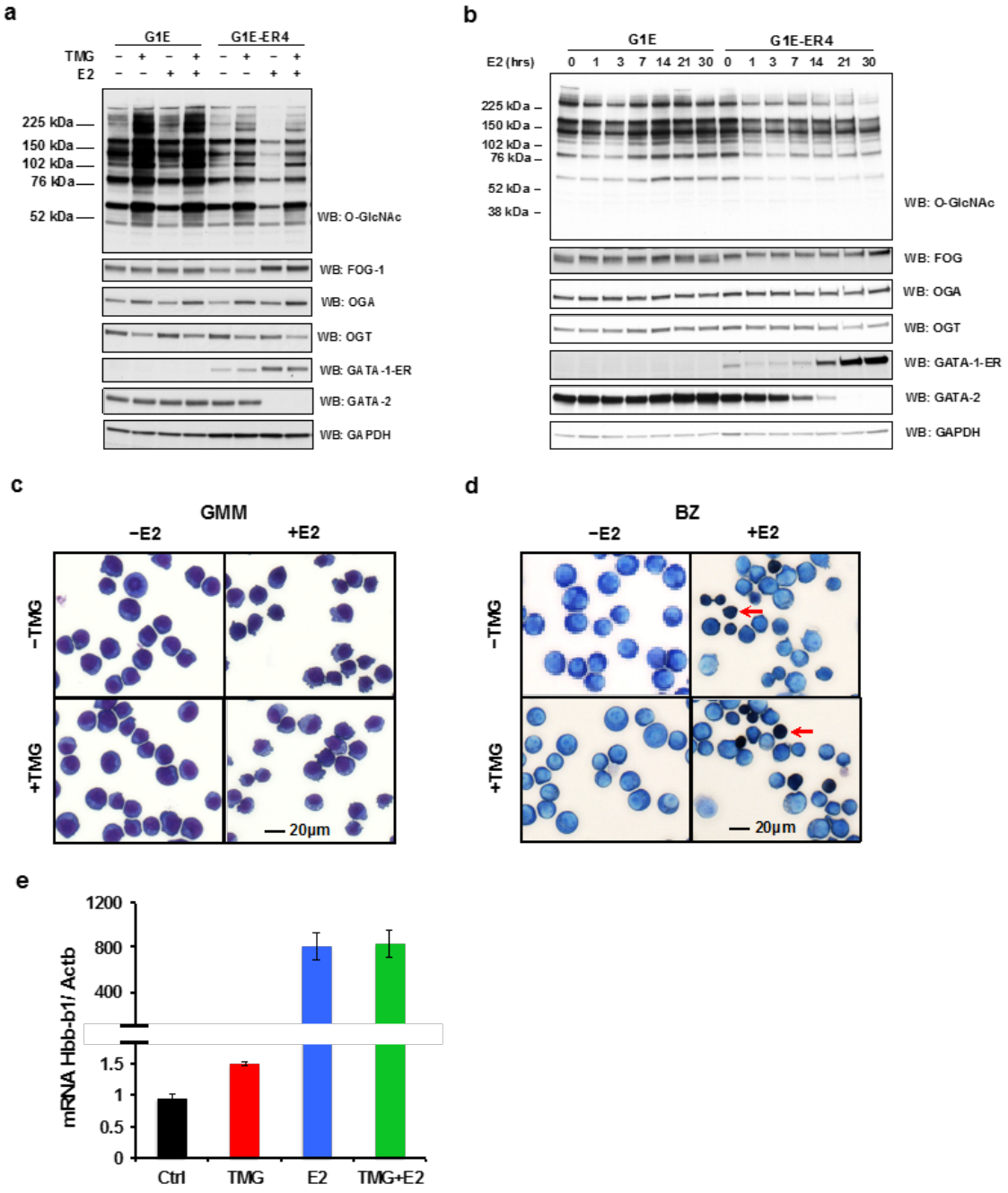

Figure 15: O-GlcNAc levels decrease upon GATA-1 restoration. G1E and G1E-ER4 cells were treated with TMG (OGA inhibitor) or/and E2 for 30 hrs. Cells were harvested at (**a)** 30 hrs or (**b**) following a time course, and total cell lysates were subjected to immunoblotting. G1E-ER4 cells were subjected to (**c**) May-Grunwald Giemsa (GMM) staining for cell morphology and (**d**) benzidine staining (BZ) for hemoglobin content. Red arrow shows hemoglobin positive cells. **e**. Transcription level of β-major globin gene (Hbb-b1) was measured by qPCR. and β-actin (Actb) was used as internal control.

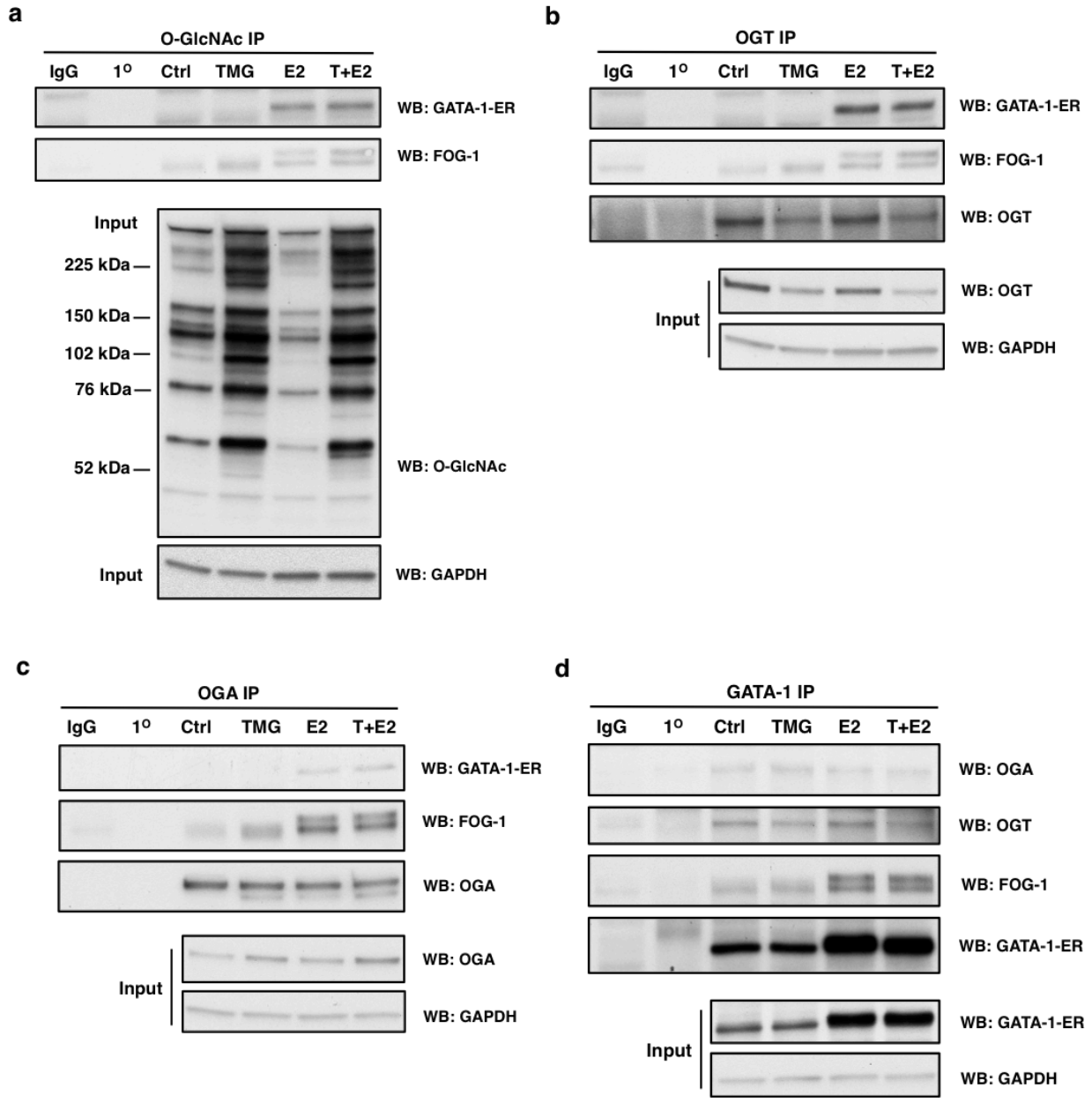

Figure 16:GATA-1 interacts with OGT and OGA in G1E-ER4 cells. G1E-ER4 cells were treated with TMG (T) or/and E2 for 30 hrs. Cells were harvested and subjected to immunoprecipitation (IP). O-GlcNAcylated protein, OGT, OGA, and GATA-1-ER were immunoprecipitated with a specific **(a)** O-GlcNAc antibody (RL2), **(b)** OGT, **(c)** OGA, and **(d)** GATA-1 antibody, respectively. Blots were probed with RL2, GATA-1, FOG-1, OGT, OGA, and GAPDH respectively. Isotype (normal rabbit or mouse IgG) IP and antibody alone precipitation (1°) were used as negative controls. All experiments were performed with at least 3 biological replicates.

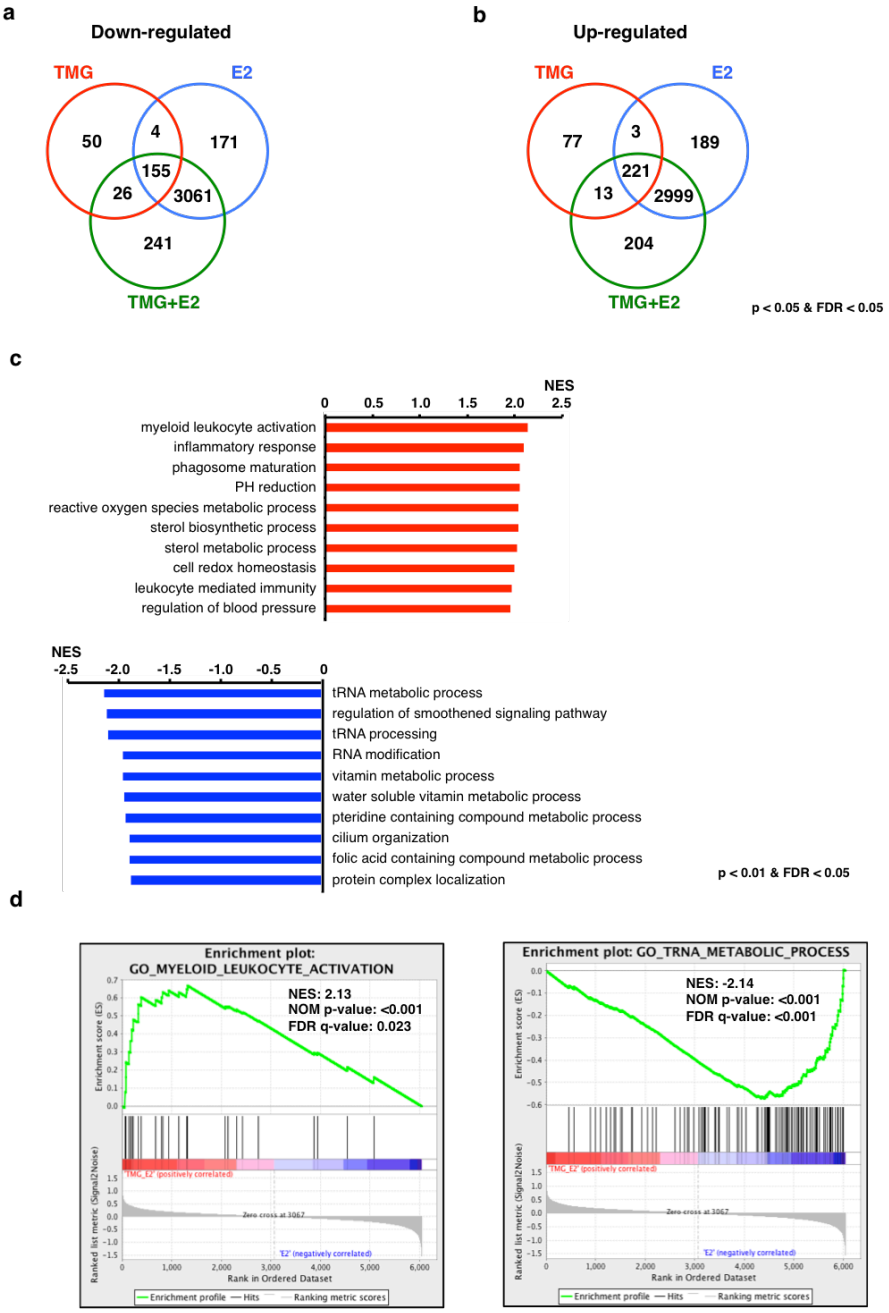

Figure 17: Inhibition of OGA changes gene transcription network during erythropoiesis. Venn diagram of (**a**) down- and (**b**) up-regulated transcripts in TMG, E2, and TMG+E2 treatment compared to control, respectively. Transcripts with p value<0.05 and FDR (False Discover Rate) <0.05 were included in the analysis. **c**. Top 10 up- (top) and down-regulated (bottom) biological processes analyzed by GSEA are ranked by NES (Normalized Enrichment Score). **d**. Enrichment plot of an up-regulated biological process myeloid leukocyte activation (left) and down-regulated biological process tRNA metabolic process (right). NOM (Nominal) p-value < 5% and FDR q-value < 25% were used for biological processes analysis.

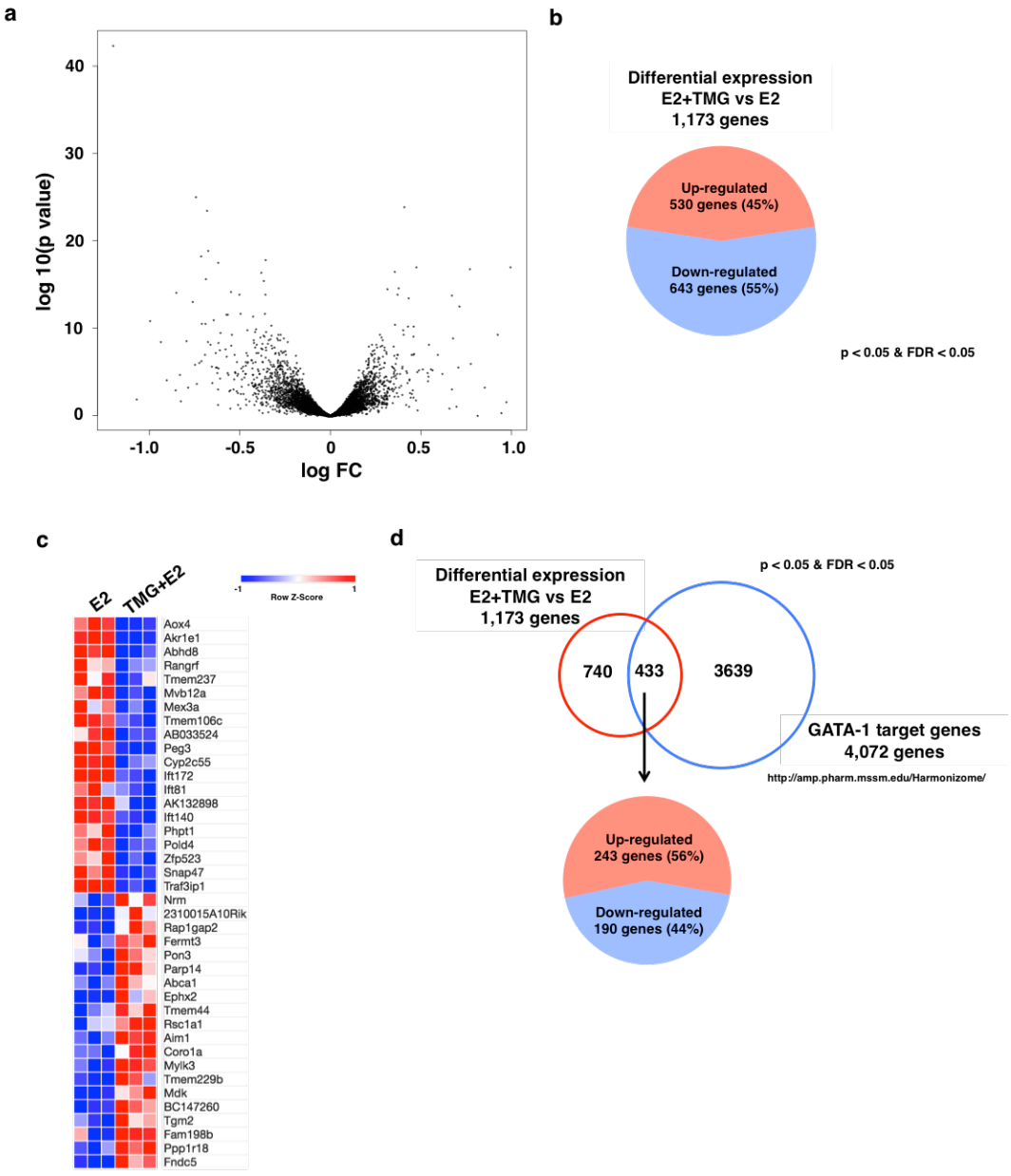

Figure 18: O-GlcNAc regulates GATA-1 target Gene transcription during erythropoiesis. **a**. The Volcano Plot of differentially expressed genes in E2+TMG compared to E2 treatment only**.** The negative log of P value is plotted on the Y-axis, and the log of the fold change (FC) is plotted on the X-axis. **b,** 1,173 genes were differentially expressed in E2+TMG compared to E2 treatment only. Pie chart shows the distribution of up- (45%) and down-regulated (55%) genes differentially expressed. Transcripts with p value < 0.05 and FDR < 0.05 were included in the analysis. **c**, The heat map of top 20 differentially expressed genes. **d**, Venn diagram of differential expression of 1,173 genes (TMG+E2 versus E2) and 4,072 GATA-1 targeted genes from database. Of 433 differentially expressed GATA-1 target genes, 243 genes were up-regulated and 190 genes down-regulated.

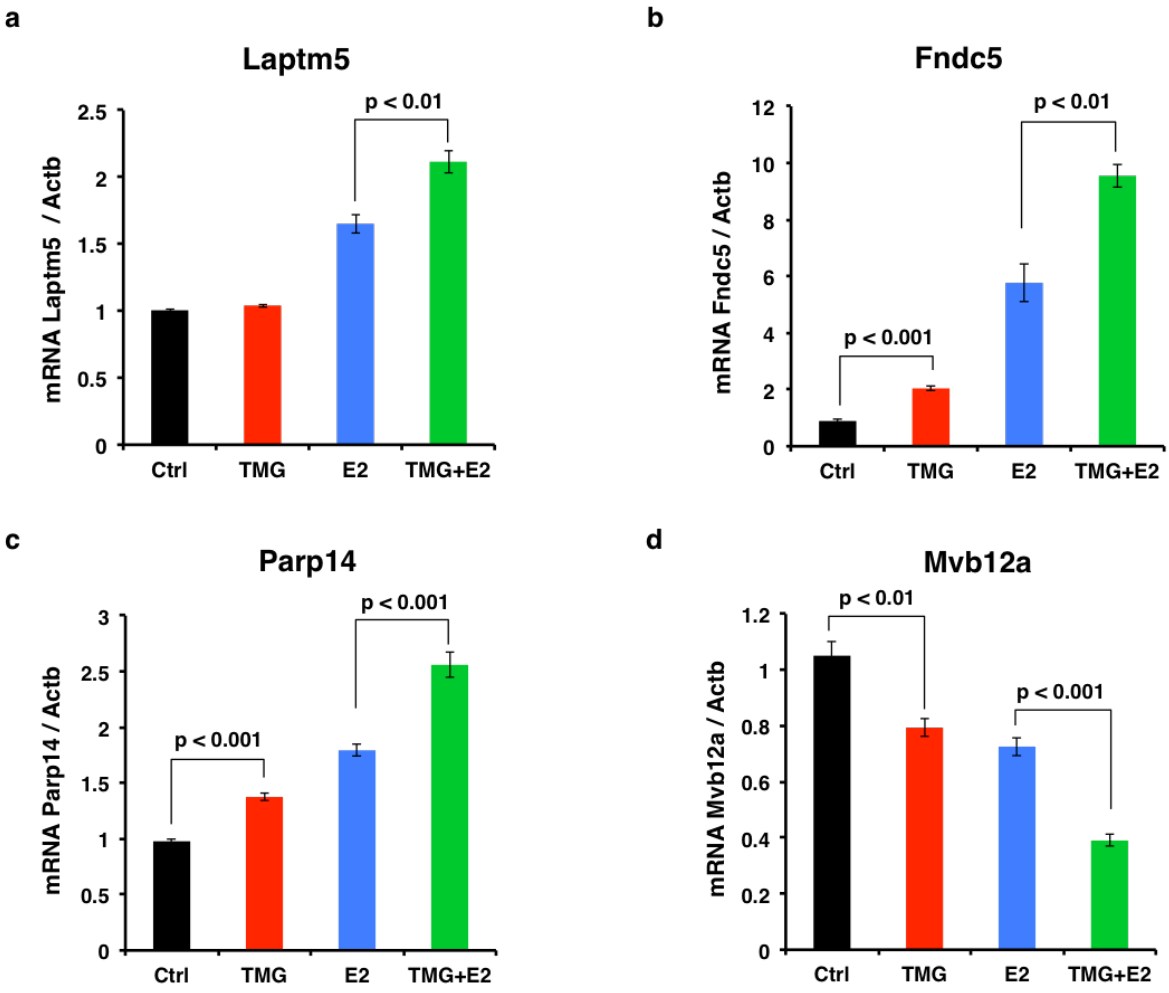

Figure 19:Validation of O-GlcNAc regulated GATA-1 target gene by qPCR. Transcription level of GATA-1 target genes (**a**) *Laptm5*, (**b**) *Fndc5*, (**c**) *Parp14*, and (**d**) *Mvb12a* was analyzed by qPCR after cells were treated with TMG or/and E2. The experiments were repeated 4 times. β-Actin (*Actb*) was used as internal control.

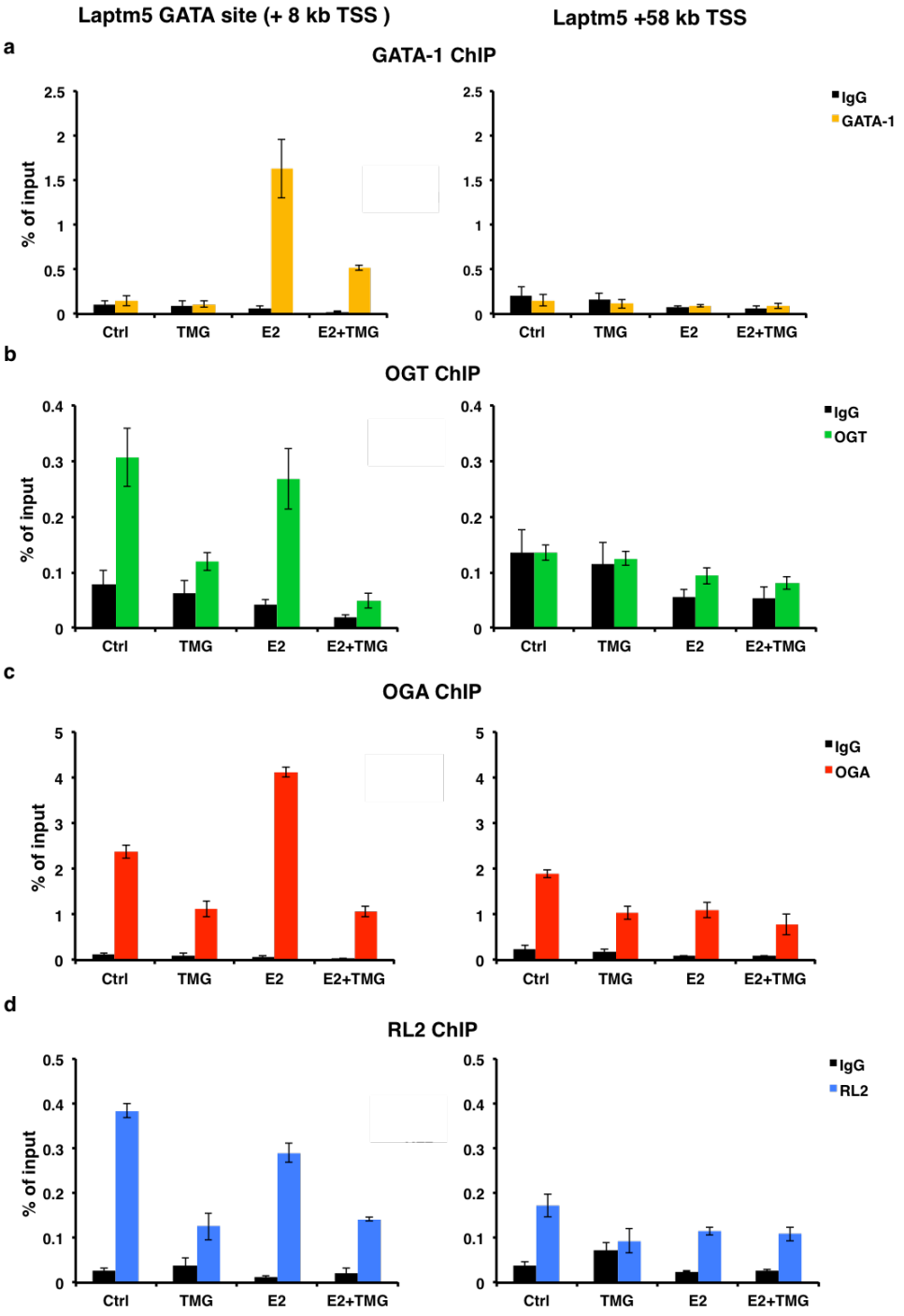

Figure 20: Inhibition of OGA decreases the occupancy of GATA-1, OGT/OGA, and O-GlcNAc level at *Laptm5* GATA Binding Site. (**a**) GATA-1, (**b**) OGT, (**c**) OGA, and (**d**) O-GlcNAc (RL2) ChIP assays were performed using G1E-ER4 cells treated with TMG or/and E2. ChIP DNA was analyzed by qPCR using a set of primer targeting the *Laptm5* GATA binding site (+ 8 kb TSS) and +58 kb TSS respectively. Normal rabbit IgG served as a negative control. All experiments were performed with at least 3 biological replicates.

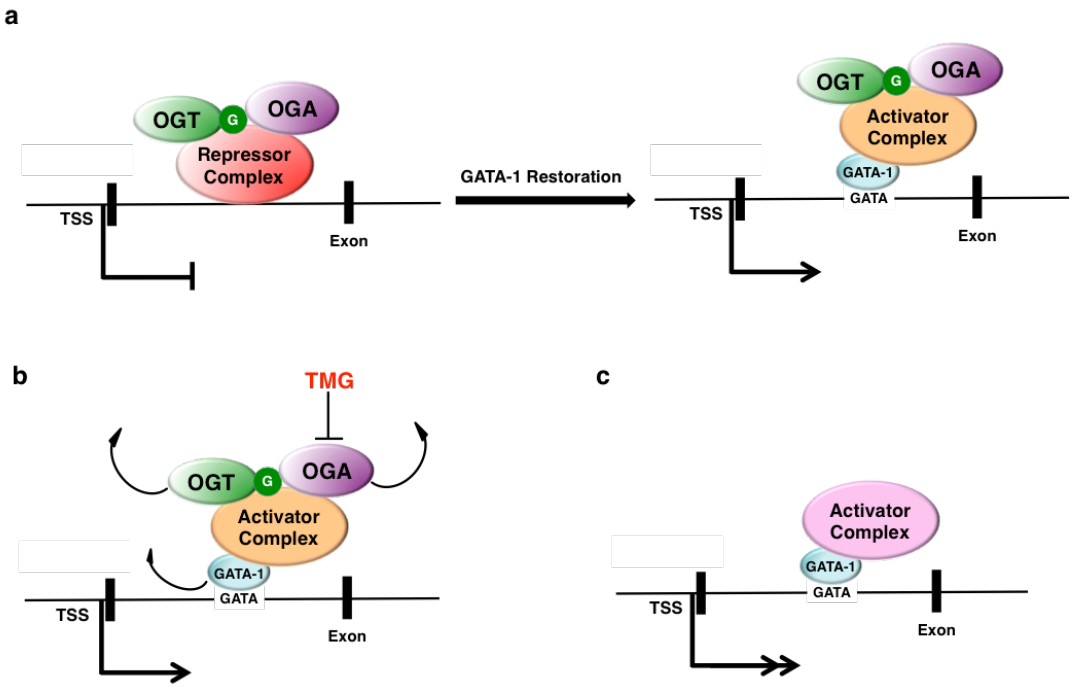

Figure 21: Proposed mechanism of OGT/OGA mediated repression and activation of GATA-1 target gene *Laptm5* during erythropoiesis. In normal G1E-ER4 cells, *Laptm5* transcription is repressed by an O-GlcNAcylated repressor complex (red oval), which localized at the first intron and interacts with OGT/OGA near the GATA binding site. Upon GATA-1 restoration by E2, GATA-1 is recruited to GATA site located in the first intron of *Laptm5*, followed by the recruitment of an O-GlcNAcylated transcription activator complex (orange oval), then *Laptm5* transcription is activated (**a**). When OGA is inhibited by TMG during erythropoiesis, the O-GlcNAc cycling on the activator complex is disrupted, leading to the decreased occupancy of OGT, OGA, and GATA-1 at the GATA binding site (**b**). The remaining activator complex (pink oval) could be more stable and active due to the changes of protein-protein interactions within the complex or chromatin structure resulting in more active transcription of *Laptm5* (**c**).

References

1. Annunziato, A. T. (2008) DNA Packaging: Nucleosomes and Chromatin. *Nature Education* **1(1):26**
2. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-260
3. Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., Lu, Z., Ye, Z., Zhu, Q., Wysocka, J., Ye, Y., Khochbin, S., Ren, B., and Zhao, Y. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**, 1016-1028
4. Bird, A. (2007) Perceptions of epigenetics. *Nature* **447**, 396-398
5. Falkenberg, K. J., and Johnstone, R. W. (2014) Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov* **13**, 673-691
6. Zhao, B. S., Roundtree, I. A., and He, C. (2017) Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* **18**, 31-42
7. Woychik, N. A. (1998) Fractions to functions: RNA polymerase II thirty years later. *Cold Spring Harb Symp Quant Biol* **63**, 311-317
8. Conaway, J. W., Shilatifard, A., Dvir, A., and Conaway, R. C. (2000) Control of elongation by RNA polymerase II. *Trends Biochem Sci* **25**, 375-380
9. Shilatifard, A., Conaway, R. C., and Conaway, J. W. (2003) The RNA polymerase II elongation complex. *Annual review of biochemistry* **72**, 693-715
10. Torres, C. R., and Hart, G. W. (1984) Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *The Journal of biological chemistry* **259**, 3308-3317
11. Hart, G. W., Slawson, C., Ramirez-Correa, G., and Lagerlof, O. (2011) Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annual review of biochemistry* **80**, 825-858
12. Dias, W. B., and Hart, G. W. (2007) O-GlcNAc modification in diabetes and Alzheimer's disease. *Molecular bioSystems* **3**, 766-772
13. Slawson, C., and Hart, G. W. (2011) O-GlcNAc signalling: implications for cancer cell biology. *Nature reviews. Cancer* **11**, 678-684
14. Bond, M. R., and Hanover, J. A. (2013) O-GlcNAc cycling: a link between metabolism and chronic disease. *Annual review of nutrition* **33**, 205-229
15. Swinburne, I. A., and Silver, P. A. (2008) Intron delays and transcriptional timing during development. *Dev Cell* **14**, 324-330
16. Myers, S. A., Panning, B., and Burlingame, A. L. (2011) Polycomb repressive complex 2 is necessary for the normal site-specific O-GlcNAc distribution in mouse embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 9490-9495
17. Kelly, W. G., and Hart, G. W. (1989) Glycosylation of chromosomal proteins: localization of O-linked N-acetylglucosamine in Drosophila chromatin. *Cell* **57**, 243-251
18. Sakabe, K., Wang, Z., and Hart, G. W. (2010) Beta-N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 19915-19920

19. Gambetta, M. C., and Muller, J. (2015) A critical perspective of the diverse roles of O-GlcNAc transferase in chromatin. *Chromosoma* **124**, 429-442
20. Fong, J. J., Nguyen, B. L., Bridger, R., Medrano, E. E., Wells, L., Pan, S., and Sifers, R. N. (2012) beta-N-Acetylglucosamine (O-GlcNAc) is a novel regulator of mitosis-specific phosphorylations on histone H3. *The Journal of biological chemistry* **287**, 12195-12203
21. Fujiki, R., Hashiba, W., Sekine, H., Yokoyama, A., Chikanishi, T., Ito, S., Imai, Y., Kim, J., He, H. H., Igarashi, K., Kanno, J., Ohtake, F., Kitagawa, H., Roeder, R. G., Brown, M., and Kato, S. (2011) GlcNAcylation of histone H2B facilitates its monoubiquitination. *Nature* **480**, 557-560
22. Kelly, W. G., Dahmus, M. E., and Hart, G. W. (1993) RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. *The Journal of biological chemistry* **268**, 10416-10424
23. Comer, F. I., and Hart, G. W. (2001) Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II. *Biochemistry* **40**, 7845-7852
24. Ranuncolo, S. M., Ghosh, S., Hanover, J. A., Hart, G. W., and Lewis, B. A. (2012) Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription in vitro and in vivo. *The Journal of biological chemistry* **287**, 23549-23561
25. Lewis, B. A., and Hanover, J. A. (2014) O-GlcNAc and the epigenetic regulation of gene expression. *The Journal of biological chemistry* **289**, 34440-34448
26. Hardiville, S., and Hart, G. W. (2016) Nutrient regulation of gene expression by O-GlcNAcylation of chromatin. *Curr Opin Chem Biol* **33**, 88-94
27. Yang, X., Zhang, F., and Kudlow, J. E. (2002) Recruitment of O-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein O-GlcNAcylation to transcriptional repression. *Cell* **110**, 69-80
28. Sinclair, D. A., Syrzycka, M., Macauley, M. S., Rastgardani, T., Komljenovic, I., Vocadlo, D. J., Brock, H. W., and Honda, B. M. (2009) Drosophila O-GlcNAc transferase (OGT) is encoded by the Polycomb group (PcG) gene, super sex combs (sxc). *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13427-13432
29. Chen, Q., Chen, Y., Bian, C., Fujiki, R., and Yu, X. (2013) TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature* **493**, 561-564
30. Bauer, D. E., and Orkin, S. H. (2011) Update on fetal hemoglobin gene regulation in hemoglobinopathies. *Curr Opin Pediatr* **23**, 1-8
31. Sankaran, V. G. (2011) Targeted therapeutic strategies for fetal hemoglobin induction. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program* **2011**, 459-465
32. Ginder, G. D. (2015) Epigenetic regulation of fetal globin gene expression in adult erythroid cells. *Transl Res* **165**, 115-125
33. Amaya, M., Desai, M., Gnanapragasam, M. N., Wang, S. Z., Zu Zhu, S., Williams, D. C., Jr., and Ginder, G. D. (2013) Mi2beta-mediated silencing of the fetal gamma-globin gene in adult erythroid cells. *Blood* **121**, 3493-3501
34. Harju-Baker, S., Costa, F. C., Fedosyuk, H., Neades, R., and Peterson, K. R. (2008) Silencing of Agamma-globin gene expression during adult definitive erythropoiesis mediated by GATA-1-FOG-1-Mi2 complex binding at the -566 GATA site. *Molecular and cellular biology* **28**, 3101-3113

35. Costa, F. C., Fedosyuk, H., Chazelle, A. M., Neades, R. Y., and Peterson, K. R. (2012) Mi2beta is required for gamma-globin gene silencing: temporal assembly of a GATA-1-FOG-1-Mi2 repressor complex in beta-YAC transgenic mice. *PLoS genetics* **8**, e1003155
36. Letting, D. L., Chen, Y. Y., Rakowski, C., Reedy, S., and Blobel, G. A. (2004) Context-dependent regulation of GATA-1 by friend of GATA-1. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 476-481
37. Miccio, A., and Blobel, G. A. (2010) Role of the GATA-1/FOG-1/NuRD pathway in the expression of human beta-like globin genes. *Molecular and cellular biology* **30**, 3460-3470
38. Miccio, A., Wang, Y., Hong, W., Gregory, G. D., Wang, H., Yu, X., Choi, J. K., Shelat, S., Tong, W., Poncz, M., and Blobel, G. A. (2010) NuRD mediates activating and repressive functions of GATA-1 and FOG-1 during blood development. *The EMBO journal* **29**, 442-456
39. Rodriguez, P., Bonte, E., Krijgsveld, J., Kolodziej, K. E., Guyot, B., Heck, A. J., Vyas, P., de Boer, E., Grosveld, F., and Strouboulis, J. (2005) GATA-1 forms distinct activating and repressive complexes in erythroid cells. *The EMBO journal* **24**, 2354-2366
40. Fujiwara, T., O'Geen, H., Keles, S., Blahnik, K., Linnemann, A. K., Kang, Y. A., Choi, K., Farnham, P. J., and Bresnick, E. H. (2009) Discovering hematopoietic mechanisms through genome-wide analysis of GATA factor chromatin occupancy. *Mol Cell* **36**, 667-681
41. Tripic, T., Deng, W., Cheng, Y., Zhang, Y., Vakoc, C. R., Gregory, G. D., Hardison, R. C., and Blobel, G. A. (2009) SCL and associated proteins distinguish active from repressive GATA transcription factor complexes. *Blood* **113**, 2191-2201
42. Yu, M., Riva, L., Xie, H., Schindler, Y., Moran, T. B., Cheng, Y., Yu, D., Hardison, R., Weiss, M. J., Orkin, S. H., Bernstein, B. E., Fraenkel, E., and Cantor, A. B. (2009) Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Mol Cell* **36**, 682-695
43. Bresnick, E. H., Lee, H. Y., Fujiwara, T., Johnson, K. D., and Keles, S. (2010) GATA switches as developmental drivers. *The Journal of biological chemistry* **285**, 31087-31093
44. Hasegawa, A., and Shimizu, R. (2016) GATA1 Activity Governed by Configurations of cis-Acting Elements. *Front Oncol* **6**, 269
45. Cheng, Y., Wu, W., Kumar, S. A., Yu, D., Deng, W., Tripic, T., King, D. C., Chen, K. B., Zhang, Y., Drautz, D., Giardine, B., Schuster, S. C., Miller, W., Chiaromonte, F., Zhang, Y., Blobel, G. A., Weiss, M. J., and Hardison, R. C. (2009) Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. *Genome Res* **19**, 2172-2184
46. Wu, W., Cheng, Y., Keller, C. A., Ernst, J., Kumar, S. A., Mishra, T., Morrissey, C., Dorman, C. M., Chen, K. B., Drautz, D., Giardine, B., Shibata, Y., Song, L., Pimkin, M., Crawford, G. E., Furey, T. S., Kellis, M., Miller, W., Taylor, J., Schuster, S. C., Zhang, Y., Chiaromonte, F., Blobel, G. A., Weiss, M. J., and Hardison, R. C. (2011) Dynamics of the epigenetic landscape during erythroid differentiation after GATA1 restoration. *Genome Res* **21**, 1659-1671
47. Haltiwanger, R. S., Holt, G. D., and Hart, G. W. (1990) Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-

acetylglucosamine:peptide beta-N-acetylglucosaminyltransferase. *The Journal of biological chemistry* **265**, 2563-2568
48. Dong, D. L., and Hart, G. W. (1994) Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol. *The Journal of biological chemistry* **269**, 19321-19330
49. Kearse, K. P., and Hart, G. W. (1991) Lymphocyte activation induces rapid changes in nuclear and cytoplasmic glycoproteins. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 1701-1705
50. Zachara, N. E., O'Donnell, N., Cheung, W. D., Mercer, J. J., Marth, J. D., and Hart, G. W. (2004) Dynamic O-GlcNAc modification of nucleocytoplasmic proteins in response to stress. A survival response of mammalian cells. *The Journal of biological chemistry* **279**, 30133-30142
51. Cheung, W. D., and Hart, G. W. (2008) AMP-activated protein kinase and p38 MAPK activate O-GlcNAcylation of neuronal proteins during glucose deprivation. *The Journal of biological chemistry* **283**, 13009-13020
52. Love, D. C., Kochan, J., Cathey, R. L., Shin, S. H., and Hanover, J. A. (2003) Mitochondrial and nucleocytoplasmic targeting of O-linked GlcNAc transferase. *J Cell Sci* **116**, 647-654
53. Tan, E. P., Caro, S., Potnis, A., Lanza, C., and Slawson, C. (2013) O-linked N-acetylglucosamine cycling regulates mitotic spindle organization. *The Journal of biological chemistry* **288**, 27085-27099
54. Tan, E. P., Villar, M. T., E, L., Lu, J., Selfridge, J. E., Artigues, A., Swerdlow, R. H., and Slawson, C. (2014) Altering O-linked beta-N-acetylglucosamine cycling disrupts mitochondrial function. *The Journal of biological chemistry* **289**, 14719-14730
55. Slawson, C., Zachara, N. E., Vosseller, K., Cheung, W. D., Lane, M. D., and Hart, G. W. (2005) Perturbations in O-linked beta-N-acetylglucosamine protein modification cause severe defects in mitotic progression and cytokinesis. *The Journal of biological chemistry* **280**, 32944-32956
56. Kazemi, Z., Chang, H., Haserodt, S., McKen, C., and Zachara, N. E. (2010) O-linked beta-N-acetylglucosamine (O-GlcNAc) regulates stress-induced heat shock protein expression in a GSK-3beta-dependent manner. *The Journal of biological chemistry* **285**, 39096-39107
57. Fardini, Y., Dehennaut, V., Lefebvre, T., and Issad, T. (2013) O-GlcNAcylation: A New Cancer Hallmark? *Frontiers in endocrinology* **4**, 99
58. Chesterton, C. J., Coupar, B. E., Butterworth, P. H., and Green, M. H. (1975) Studies on the control of ribosomal RNA synthesis in HeLa cells. *Eur J Biochem* **57**, 79-83
59. Kim, H. R., Kang, H. S., and Kim, H. D. (1999) Geldanamycin induces heat shock protein expression through activation of HSF1 in K562 erythroleukemic cells. *IUBMB Life* **48**, 429-433
60. Sawicki, S. G., and Godman, G. C. (1971) On the differential cytotoxicity of actinomycin D. *J Cell Biol* **50**, 746-761
61. DiTacchio, L., Le, H. D., Vollmers, C., Hatori, M., Witcher, M., Secombe, J., and Panda, S. (2011) Histone lysine demethylase JARID1a activates CLOCK-BMAL1 and influences the circadian clock. *Science* **333**, 1881-1885

62. Slawson, C., Lakshmanan, T., Knapp, S., and Hart, G. W. (2008) A mitotic GlcNAcylation/phosphorylation signaling complex alters the posttranslational state of the cytoskeletal protein vimentin. *Mol Biol Cell* **19**, 4130-4140
63. Yuzwa, S. A., Macauley, M. S., Heinonen, J. E., Shan, X., Dennis, R. J., He, Y., Whitworth, G. E., Stubbs, K. A., McEachern, E. J., Davies, G. J., and Vocadlo, D. J. (2008) A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo. *Nat Chem Biol* **4**, 483-490
64. Obrig, T. G., Culp, W. J., McKeehan, W. L., and Hardesty, B. (1971) The mechanism by which cycloheximide and related glutarimide antibiotics inhibit peptide synthesis on reticulocyte ribosomes. *The Journal of biological chemistry* **246**, 174-181
65. Ruan, H. B., Dietrich, M. O., Liu, Z. W., Zimmer, M. R., Li, M. D., Singh, J. P., Zhang, K., Yin, R., Wu, J., Horvath, T. L., and Yang, X. (2014) O-GlcNAc transferase enables AgRP neurons to suppress browning of white fat. *Cell* **159**, 306-317
66. Yang, W. H., Kim, J. E., Nam, H. W., Ju, J. W., Kim, H. S., Kim, Y. S., and Cho, J. W. (2006) Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. *Nat Cell Biol* **8**, 1074-1083
67. Shi, F. T., Kim, H., Lu, W., He, Q., Liu, D., Goodell, M. A., Wan, M., and Songyang, Z. (2013) Ten-eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells. *The Journal of biological chemistry* **288**, 20776-20784
68. Yehezkel, G., Cohen, L., Kliger, A., Manor, E., and Khalaila, I. (2012) O-linked beta-N-acetylglucosaminylation (O-GlcNAcylation) in primary and metastatic colorectal cancer clones and effect of N-acetyl-beta-D-glucosaminidase silencing on cell phenotype and transcriptome. *The Journal of biological chemistry* **287**, 28755-28769
69. Gloster, T. M., Zandberg, W. F., Heinonen, J. E., Shen, D. L., Deng, L., and Vocadlo, D. J. (2011) Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells. *Nat Chem Biol* **7**, 174-181
70. Park, K., Saudek, C. D., and Hart, G. W. (2010) Increased expression of beta-N-acetylglucosaminidase in erythrocytes from individuals with pre-diabetes and diabetes. *Diabetes* **59**, 1845-1850
71. Wells, L., Whelan, S. A., and Hart, G. W. (2003) O-GlcNAc: a regulatory post-translational modification. *Biochem Biophys Res Commun* **302**, 435-441
72. Whisenhunt, T. R., Yang, X., Bowe, D. B., Paterson, A. J., Van Tine, B. A., and Kudlow, J. E. (2006) Disrupting the enzyme complex regulating O-GlcNAcylation blocks signaling and development. *Glycobiology* **16**, 551-563
73. Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E., Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Kolpakov, F. A., Podkolodny, N. L., and Kolchanov, N. A. (1998) Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucleic Acids Res* **26**, 362-367
74. Le Mee, S., Fromigue, O., and Marie, P. J. (2005) Sp1/Sp3 and the myeloid zinc finger gene MZF1 regulate the human N-cadherin promoter in osteoblasts. *Exp Cell Res* **302**, 129-142
75. Ma, Z., Vocadlo, D. J., and Vosseller, K. (2013) Hyper-O-GlcNAcylation is anti-apoptotic and maintains constitutive NF-kappaB activity in pancreatic cancer cells. *The Journal of biological chemistry* **288**, 15121-15130

76. de Queiroz, R. M., Carvalho, E., and Dias, W. B. (2014) O-GlcNAcylation: The Sweet Side of the Cancer. *Front Oncol* **4**, 132
77. Orkin, S. H. (1990) Globin gene regulation and switching: circa 1990. *Cell* **63**, 665-672
78. Fraser, P., and Grosveld, F. (1998) Locus control regions, chromatin activation and transcription. *Current opinion in cell biology* **10**, 361-365
79. Li, Q., Peterson, K. R., Fang, X., and Stamatoyannopoulos, G. (2002) Locus control regions. *Blood* **100**, 3077-3086
80. Verduzco, L. A., and Nathan, D. G. (2009) Sickle cell disease and stroke. *Blood* **114**, 5117-5125
81. Gladwin, M. T., and Sachdev, V. (2012) Cardiovascular abnormalities in sickle cell disease. *Journal of the American College of Cardiology* **59**, 1123-1133
82. Hart, G. W. (2014) Three Decades of Research on O-GlcNAcylation - A Major Nutrient Sensor That Regulates Signaling, Transcription and Cellular Metabolism. *Frontiers in endocrinology* **5**, 183
83. Lanza, C., Tan, E. P., Zhang, Z., Machacek, M., Brinker, A. E., Azuma, M., and Slawson, C. (2016) Reduced O-GlcNAcase Expression Promotes Mitotic Errors and Spindle Defects. *Cell Cycle*, 0
84. Shafi, R., Iyer, S. P., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W., and Marth, J. D. (2000) The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 5735-5739
85. Yang, Y. R., Song, M., Lee, H., Jeon, Y., Choi, E. J., Jang, H. J., Moon, H. Y., Byun, H. Y., Kim, E. K., Kim, D. H., Lee, M. N., Koh, A., Ghim, J., Choi, J. H., Lee-Kwon, W., Kim, K. T., Ryu, S. H., and Suh, P. G. (2012) O-GlcNAcase is essential for embryonic development and maintenance of genomic stability. *Aging cell* **11**, 439-448
86. Zhu, Y., Shan, X., Yuzwa, S. A., and Vocadlo, D. J. (2014) The emerging link between O-GlcNAc and Alzheimer disease. *The Journal of biological chemistry* **289**, 34472-34481
87. Akan, I., Love, D. C., Harwood, K., Bond, M. R., and Hanover, J. A. (2016) Drosophila O-GlcNAcase deletion globally perturbs chromatin O-GlcNAcylation. *The Journal of biological chemistry*
88. Li, M. D., Ruan, H. B., Hughes, M. E., Lee, J. S., Singh, J. P., Jones, S. P., Nitabach, M. N., and Yang, X. (2013) O-GlcNAc signaling entrains the circadian clock by inhibiting BMAL1/CLOCK ubiquitination. *Cell metabolism* **17**, 303-310
89. Ramakrishnan, P., Clark, P. M., Mason, D. E., Peters, E. C., Hsieh-Wilson, L. C., and Baltimore, D. (2013) Activation of the transcriptional function of the NF-kappaB protein c-Rel by O-GlcNAc glycosylation. *Science signaling* **6**, ra75
90. Dehennaut, V., Leprince, D., and Lefebvre, T. (2014) O-GlcNAcylation, an Epigenetic Mark. Focus on the Histone Code, TET Family Proteins, and Polycomb Group Proteins. *Frontiers in endocrinology* **5**, 155
91. Peterson, K. R., Costa, F. C., Fedosyuk, H., Neades, R. Y., Chazelle, A. M., Zelenchuk, L., Fonteles, A. H., Dalal, P., Roy, A., Chaguturu, R., Li, B., and Pace, B. S. (2014) A cell-based high-throughput screen for novel chemical inducers of fetal hemoglobin for treatment of hemoglobinopathies. *PloS one* **9**, e107006

92. Zhang, Z., Tan, E. P., VandenHull, N. J., Peterson, K. R., and Slawson, C. (2014) O-GlcNAcase Expression is Sensitive to Changes in O-GlcNAc Homeostasis. *Frontiers in endocrinology* **5**, 206
93. Cioe, L., McNab, A., Hubbell, H. R., Meo, P., Curtis, P., and Rovera, G. (1981) Differential expression of the globin genes in human leukemia K562(S) cells induced to differentiate by hemin or butyric acid. *Cancer research* **41**, 237-243
94. Rao, F. V., Dorfmueller, H. C., Villa, F., Allwood, M., Eggleston, I. M., and van Aalten, D. M. (2006) Structural insights into the mechanism and inhibition of eukaryotic O-GlcNAc hydrolysis. *The EMBO journal* **25**, 1569-1578
95. Lahlil, R., Lecuyer, E., Herblot, S., and Hoang, T. (2004) SCL assembles a multifactorial complex that determines glycophorin A expression. *Molecular and cellular biology* **24**, 1439-1452
96. Merika, M., and Orkin, S. H. (1993) DNA-binding specificity of GATA family transcription factors. *Molecular and cellular biology* **13**, 3999-4010
97. Chou, T. Y., Hart, G. W., and Dang, C. V. (1995) c-Myc is glycosylated at threonine 58, a known phosphorylation site and a mutational hot spot in lymphomas. *The Journal of biological chemistry* **270**, 18961-18965
98. Li, X., Molina, H., Huang, H., Zhang, Y. Y., Liu, M., Qian, S. W., Slawson, C., Dias, W. B., Pandey, A., Hart, G. W., Lane, M. D., and Tang, Q. Q. (2009) O-linked N-acetylglucosamine modification on CCAAT enhancer-binding protein beta: role during adipocyte differentiation. *J Biol Chem* **284**, 19248-19254
99. Dephoure, N., Zhou, C., Villen, J., Beausoleil, S. A., Bakalarski, C. E., Elledge, S. J., and Gygi, S. P. (2008) A quantitative atlas of mitotic phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 10762-10767
100. Mertins, P., Qiao, J. W., Patel, J., Udeshi, N. D., Clauser, K. R., Mani, D. R., Burgess, M. W., Gillette, M. A., Jaffe, J. D., and Carr, S. A. (2013) Integrated proteomic analysis of post-translational modifications by serial enrichment. *Nat Methods* **10**, 634-637
101. Rigbolt, K. T., Prokhorova, T. A., Akimov, V., Henningsen, J., Johansen, P. T., Kratchmarova, I., Kassem, M., Mann, M., Olsen, J. V., and Blagoev, B. (2011) System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. *Science signaling* **4**, rs3
102. Sharma, K., D'Souza, R. C., Tyanova, S., Schaab, C., Wisniewski, J. R., Cox, J., and Mann, M. (2014) Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Rep* **8**, 1583-1594
103. Yi, T., Zhai, B., Yu, Y., Kiyotsugu, Y., Raschle, T., Etzkorn, M., Seo, H. C., Nagiec, M., Luna, R. E., Reinherz, E. L., Blenis, J., Gygi, S. P., and Wagner, G. (2014) Quantitative phosphoproteomic analysis reveals system-wide signaling pathways downstream of SDF-1/CXCR4 in breast cancer stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E2182-2190
104. Trinidad, J. C., Barkan, D. T., Gulledge, B. F., Thalhammer, A., Sali, A., Schoepfer, R., and Burlingame, A. L. (2012) Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. *Molecular & cellular proteomics : MCP* **11**, 215-229

105. Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V., Orkin, S. H., and Costantini, F. (1991) Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349**, 257-260
106. Orkin, S. H. (1992) GATA-binding transcription factors in hematopoietic cells. *Blood* **80**, 575-581
107. Fujiwara, Y., Browne, C. P., Cunniff, K., Goff, S. C., and Orkin, S. H. (1996) Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 12355-12358
108. Yamamoto, M., Ko, L. J., Leonard, M. W., Beug, H., Orkin, S. H., and Engel, J. D. (1990) Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. *Genes Dev* **4**, 1650-1662
109. Welch, J. J., Watts, J. A., Vakoc, C. R., Yao, Y., Wang, H., Hardison, R. C., Blobel, G. A., Chodosh, L. A., and Weiss, M. J. (2004) Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood* **104**, 3136-3147
110. Kingsley, P. D., Greenfest-Allen, E., Frame, J. M., Bushnell, T. P., Malik, J., McGrath, K. E., Stoeckert, C. J., and Palis, J. (2013) Ontogeny of erythroid gene expression. *Blood* **121**, e5-e13
111. Soler, E., Andrieu-Soler, C., de Boer, E., Bryne, J. C., Thongjuea, S., Stadhouders, R., Palstra, R. J., Stevens, M., Kockx, C., van Ijcken, W., Hou, J., Steinhoff, C., Rijkers, E., Lenhard, B., and Grosveld, F. (2010) The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. *Genes Dev* **24**, 277-289
112. Zhang, Y., Wang, Z., Zhang, J., and Lim, S. H. (2009) Core promoter sequence of SEMG I spans between the two putative GATA-1 binding domains and is responsive to IL-4 and IL-6 in myeloma cells. *Leuk Res* **33**, 166-169
113. Dey, A., Seshasayee, D., Noubade, R., French, D. M., Liu, J., Chaurushiya, M. S., Kirkpatrick, D. S., Pham, V. C., Lill, J. R., Bakalarski, C. E., Wu, J., Phu, L., Katavolos, P., LaFave, L. M., Abdel-Wahab, O., Modrusan, Z., Seshagiri, S., Dong, K., Lin, Z., Balazs, M., Suriben, R., Newton, K., Hymowitz, S., Garcia-Manero, G., Martin, F., Levine, R. L., and Dixit, V. M. (2012) Loss of the tumor suppressor BAP1 causes myeloid transformation. *Science* **337**, 1541-1546
114. Deplus, R., Delatte, B., Schwinn, M. K., Defrance, M., Mendez, J., Murphy, N., Dawson, M. A., Volkmar, M., Putmans, P., Calonne, E., Shih, A. H., Levine, R. L., Bernard, O., Mercher, T., Solary, E., Urh, M., Daniels, D. L., and Fuks, F. (2013) TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *The EMBO journal* **32**, 645-655
115. Chu, C. S., Lo, P. W., Yeh, Y. H., Hsu, P. H., Peng, S. H., Teng, Y. C., Kang, M. L., Wong, C. H., and Juan, L. J. (2014) O-GlcNAcylation regulates EZH2 protein stability and function. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 1355-1360
116. Maury, J. J., El Farran, C. A., Ng, D., Loh, Y. H., Bi, X., Bardor, M., and Choo, A. B. (2015) RING1B O-GlcNAcylation regulates gene targeting of polycomb repressive complex 1 in human embryonic stem cells. *Stem Cell Res* **15**, 182-189

117. Resto, M., Kim, B. H., Fernandez, A. G., Abraham, B. J., Zhao, K., and Lewis, B. A. (2016) O-GlcNAcase Is an RNA Polymerase II Elongation Factor Coupled to Pausing Factors SPT5 and TIF1beta. *The Journal of biological chemistry* **291**, 22703-22713
118. Zhang, Z., Costa, F. C., Tan, E. P., Bushue, N., DiTacchio, L., Costello, C. E., McComb, M. E., Whelan, S. A., Peterson, K. R., and Slawson, C. (2016) O-Linked N-Acetylglucosamine (O-GlcNAc) Transferase and O-GlcNAcase Interact with Mi2beta Protein at the Agamma-Globin Promoter. *The Journal of biological chemistry* **291**, 15628-15640
119. Weiss, M. J., Yu, C., and Orkin, S. H. (1997) Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. *Molecular and cellular biology* **17**, 1642-1651
120. Adra, C. N., Zhu, S., Ko, J. L., Guillemot, J. C., Cuervo, A. M., Kobayashi, H., Horiuchi, T., Lelias, J. M., Rowley, J. D., and Lim, B. (1996) LAPTM5: a novel lysosomal-associated multispanning membrane protein preferentially expressed in hematopoietic cells. *Genomics* **35**, 328-337
121. Glowacka, W. K., Alberts, P., Ouchida, R., Wang, J. Y., and Rotin, D. (2012) LAPTM5 protein is a positive regulator of proinflammatory signaling pathways in macrophages. *The Journal of biological chemistry* **287**, 27691-27702
122. Grass, J. A., Jing, H., Kim, S. I., Martowicz, M. L., Pal, S., Blobel, G. A., and Bresnick, E. H. (2006) Distinct functions of dispersed GATA factor complexes at an endogenous gene locus. *Molecular and cellular biology* **26**, 7056-7067
123. Nelson, J. D., Denisenko, O., and Bomsztyk, K. (2006) Protocol for the fast chromatin immunoprecipitation (ChIP) method. *Nat Protoc* **1**, 179-185
124. Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., Houstis, N., Daly, M. J., Patterson, N., Mesirov, J. P., Golub, T. R., Tamayo, P., Spiegelman, B., Lander, E. S., Hirschhorn, J. N., Altshuler, D., and Groop, L. C. (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34**, 267-273
125. Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15545-15550
126. Lamonica, J. M., Deng, W., Kadauke, S., Campbell, A. E., Gamsjaeger, R., Wang, H., Cheng, Y., Billin, A. N., Hardison, R. C., Mackay, J. P., and Blobel, G. A. (2011) Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. *Proceedings of the National Academy of Sciences of the United States of America* **108**, E159-168
127. Manavathi, B., Lo, D., Bugide, S., Dey, O., Imren, S., Weiss, M. J., and Humphries, R. K. (2012) Functional regulation of pre-B-cell leukemia homeobox interacting protein 1 (PBXIP1/HPIP) in erythroid differentiation. *The Journal of biological chemistry* **287**, 5600-5614

128. Villen, J., Beausoleil, S. A., Gerber, S. A., and Gygi, S. P. (2007) Large-scale phosphorylation analysis of mouse liver. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 1488-1493
129. Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T., Rad, R., Beausoleil, S. A., Villen, J., Haas, W., Sowa, M. E., and Gygi, S. P. (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143**, 1174-1189
130. Cantor, A. B., and Orkin, S. H. (2002) Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene* **21**, 3368-3376
131. Rouillard, A. D., Gundersen, G. W., Fernandez, N. F., Wang, Z., Monteiro, C. D., McDermott, M. G., and Ma'ayan, A. (2016) The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database (Oxford)* **2016**
132. Consortium, E. P. (2004) The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* **306**, 636-640
133. Olivier-Van Stichelen, S., and Hanover, J. A. (2015) You are what you eat: O-linked N-acetylglucosamine in disease, development and epigenetics. *Curr Opin Clin Nutr Metab Care* **18**, 339-345
134. Howerton, C. L., Morgan, C. P., Fischer, D. B., and Bale, T. L. (2013) O-GlcNAc transferase (OGT) as a placental biomarker of maternal stress and reprogramming of CNS gene transcription in development. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 5169-5174
135. Speakman, C. M., Domke, T. C., Wongpaiboonwattana, W., Sanders, K., Mudaliar, M., van Aalten, D. M., Barton, G. J., and Stavridis, M. P. (2014) Elevated O-GlcNAc levels activate epigenetically repressed genes and delay mouse ESC differentiation without affecting naive to primed cell transition. *Stem Cells* **32**, 2605-2615
136. Sohn, K. C., Lee, E. J., Shin, J. M., Lim, E. H., No, Y., Lee, J. Y., Yoon, T. Y., Lee, Y. H., Im, M., Lee, Y., Seo, Y. J., Lee, J. H., and Kim, C. D. (2014) Regulation of keratinocyte differentiation by O-GlcNAcylation. *J Dermatol Sci* **75**, 10-15
137. Kim, H. S., Park, S. Y., Choi, Y. R., Kang, J. G., Joo, H. J., Moon, W. K., and Cho, J. W. (2009) Excessive O-GlcNAcylation of proteins suppresses spontaneous cardiogenesis in ES cells. *FEBS Lett* **583**, 2474-2478
138. Ogawa, M., Mizofuchi, H., Kobayashi, Y., Tsuzuki, G., Yamamoto, M., Wada, S., and Kamemura, K. (2012) Terminal differentiation program of skeletal myogenesis is negatively regulated by O-GlcNAc glycosylation. *Biochim Biophys Acta* **1820**, 24-32
139. Andres-Bergos, J., Tardio, L., Larranaga-Vera, A., Gomez, R., Herrero-Beaumont, G., and Largo, R. (2012) The increase in O-linked N-acetylglucosamine protein modification stimulates chondrogenic differentiation both in vitro and in vivo. *The Journal of biological chemistry* **287**, 33615-33628
140. Koyama, T., and Kamemura, K. (2015) Global increase in O-linked N-acetylglucosamine modification promotes osteoblast differentiation. *Exp Cell Res* **338**, 194-202
141. Yang, X., and Qian, K. (2017) Protein O-GlcNAcylation: emerging mechanisms and functions. *Nat Rev Mol Cell Biol*
142. Lewis, B. A., Burlingame, A. L., and Myers, S. A. (2016) Human RNA Polymerase II Promoter Recruitment in Vitro Is Regulated by O-Linked N-

Acetylglucosaminyltransferase (OGT). *The Journal of biological chemistry* **291**, 14056-14061