INVESTIGATING THE ROLE OF O-GLCNAC IN THE REGULATION OF HUMAN OCT4

by

SANDRA NICOLE BRIMBLE

(Under the Direction of Lance Wells)

ABSTRACT

O-linked β-*N*-acetylglucosamine (O-GlcNAc) is a single sugar modification found on many different classes of nuclear and cytoplasmic proteins. Addition of this modification, by the enzyme O-linked *N*-acetylglucosamine transferase (OGT), is dynamic and inducible. There is mounting evidence that O-GlcNAc plays a role in regulation of development but the mechanism is not clearly understood. One major class of proteins modified by O-GlcNAc is transcription factors. O-GlcNAc regulates transcription factor properties through a variety of different mechanisms including localization, stability and transcriptional activation. Maintenance of embryonic stem (ES) cell pluripotency requires tight regulation of several key transcription factors, many of which are modified by O-GlcNAc. Pou5f1 (Oct4) is one of the transcription factors required for pluripotency of ES cells and more recently, the generation of induced pluripotent stem (iPS) cells. The action of Oct4 is modulated by the addition of several post-translational modifications, including O-GlcNAc. Previous studies in mouse found a single site of O-GlcNAc addition responsible for transcriptional

regulation. This study was designed to determine if this mechanism is conserved in human. We mapped 10 novel sites of O-GlcNAc attachment on human Oct4, and confirmed a role for OGT in transcriptional activation of Oct4 at a site distinct from that found in mouse that allows distinction between different Oct4 promoters. Additionally, we uncovered a potential new role for OGT that does not include its catalytic function. These results confirm that human Oct4 activity is being regulated by OGT by a mechanism that is independent of O-GlcNAc and distinct from mouse Oct4.

INDEX WORDS: O-GlcNAc, Oct4, embryonic stem cells, transcription factor

INVESTIGATING THE ROLE OF O-GLCNAC IN THE REGULATION OF HUMAN OCT4

by

SANDRA NICOLE BRIMBLE

B.Sc., University of Adelaide, Australia, 1999

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

INVESTIGATING THE ROLE OF O-GLCNAC IN THE REGULATION OF HUMAN OCT4

by

SANDRA NICOLE BRIMBLE

Major Professor:	Lance Wells
Committee:	Stephen Dalton Allan Robins Michael Tiemeyer

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
August 2014

DEDICATION

I would like to dedicate this dissertation to my family, both blood related and extended. You have given me strength and I could not have done this without your unwavering support. To my father, Graham Brimble, who always helped me with my homework even when you didn't even know anything about what I was studying. To my life partner and husband, Stephen Constable, who always believed in me without question, made me laugh and supported me through the tough times. Finally, to my daughter Lilian Constable, you are the light in my life, and my motivation to continue when times got tough. I hope that I will be an inspiration for you in the future to do what ever your heart desires.

ACKNOWLEDGEMENTS

There were so many people that helped me with this undertaking that it will be hard to list them all. First I would like to thank my adviser Lance Wells for the opportunity and for believing in me when no one else did. Thank you to all my committee members for all the helpful discussions and advice when I needed it, especially Allan Robins for always pushing me to be a better person and scientist and never being shy about pointing out when I was wrong. All the members of the Wells' lab for all your helpful discussions both scientific and personal, in particular Rob Bridger who would always get me the help and reagents I needed with a smile, sometimes at short notice. To the members of the CCRC, thank you for your openness to help out with scientific discussion, reagents, or equipment. Michael Kulik for the differentiation work, and Jae-Min Lim for the amazing work done on the Mass Spec. Finally I want to thank Krithika Vaidyanathan for all the late night lab sessions, the hours of scientific discussion and your amazing friendship over the 6 years we were together

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Early Embryonic Development

Development of an entire animal from a single cell requires a finely tuned series of events. Most of what we know about mammalian development comes from studying mouse embryos, with advances in human development from embryos obtained through assisted reproductive technology [1]. After fertilization, the cell undergoes the process of cleavage, a series of mitotic divisions in which the cytoplasm is divided into smaller portions. After three cleavage events have occurred, the resulting blastomeres undergo a process known as compaction. During compaction, cells express the cell surface marker e-cadherin which allow close cell-cell contacts between the cells [2]. This is approximately the same stage when the embryonic genome is actively transcribed [1]. These cells then divide to produce a 16 cell morula consisting of a small group of internal cells surrounded by larger group of external cells. These outer cells differentiate to become the trophectoderm cells that form part of the placenta and the chorion. The remaining cells, the inner cell mass (ICM), contribute to the embryo proper and the extra embryonic tissues including the yolk sac and extra embryonic mesoderm. The cells of the ICM then differentiate into two populations: the epiblast and the primitive endoderm cells. After implantation, through the process of gastrulation, the epiblast cells form the three germ layers which differentiate to form the mature cells in the animal: ectoderm

(nerves, skin), endoderm (internal organs) and mesoderm (muscle, blood) [2] (Figure 1.1).

Embryonic Stem Cells

In 1981, the ability to culture pluripotent cells from the inner cell mass (ICM) of murine embryos [3, 4] brought the field to a new era of discovery. Almost two decades later, the first stem cell lines from human embryos were derived [5]. When provided with the correct combination of growth conditions and growth factors, embryonic stem cells have two important properties: they can expand indefinitely (self-renewal) and still retain the potential to differentiate into any cell type found in the three germ layers (pluripotency) [6-9]. The major promise of human stem cells is the ability to generate pure cell populations for cell therapy purposes for a variety of diseases (Figure 1.1). To date there are several examples of human ES cells being appropriately turned into multiple cells required to cure certain diseases such as insulin producing cells for the therapy of diabetes [10], dopaminergic neurons to cure Parkinson's Disease [11, 12], oligodendrocytes for spinal crush injury [13] and muscle for cardiovascular disease [14]. For these therapies to be successful, knowledge of the mechanisms required for maintenance of pluripotency and differentiation is essential.

Genes Required For Maintenance Of Pluripotency

Maintenance of pluripotent cells requires an intricate gene system that allows self-renewal but prevents differentiation. Oct4 [15-17], Sox 2 (SRY-related HMG box containing protein) [18, 19], and Nanog [20-22] have been

recognized as key transcription factors that make up the core of this regulation. Oct4 is the first factor discovered to be important for the pluripotent state [16] and is the focus of this study, so will be discussed in more detail below. Sox2 was investigated due to its relationship with Oct4. The expression pattern of Oct4 and Sox2 overlap in mouse embryos [19], and Oct4 and Sox2 cooperate to express the *fgf4* gene, which is also developmentally regulated [23, 24]. The role for Sox2 was subsequently solidified during several studies that determined the binding sites of Oct4, Sox2 and nanog [25, 26]. Sox2 bound over half of the Oct4 target genes and of these, almost all were also bound by nanog [15, 25, 26]. Two different groups first discovered Nanog during a screen to determine transcription factors that could promote pluripotency [20, 22]. Like Oct4, Nanog is expressed exclusively in the ICM and epiblast cells, and is absent in the trophoblast and primitive endoderm cells [20]. One striking outcome of these studies is the revelation that all three transcription factors bind to their own promoters in an auto regulatory loop to maintain pluripotency [25].

Induced Pluripotent Stem cells

In 2012, the Nobel Prize in Physiology or Medicine was awarded jointly to John B. Gurdon and Shinya Yamanaka for the discovery that mature cells can be reprogrammed to become pluripotent. Gurdon discovered that he could replace the nucleus from a frog egg with the nucleus from a frog intestinal cell and it still contained all the information required to develop normally into a normal tadpole [27]. About 40 years later, Yamanaka discovered that he could take mature mouse cells and create embryonic stem cell-like cells, named induced pluripotent

stem (iPS) cells, using forced expression of several factors involved in stem cell pluripotency [28]. These factors include previously discussed Oct4, Sox 2 and Nanog, along with Kruppel-like factor 4 (Klf4) [29], c-myc [30] and Lin28 [31]. Although different combinations of factors can be used, they all require Oct4 suggesting it is essential in this process [32]. iPS cells are generated from fully or partially differentiated cells, and have the same characteristics and differentiation potential as ES cells. These factors "reprogram" somatic cells back to their pluripotent state by a mechanism that is not yet well understood [33]. Soon after, the technology was available for human cells opening the door for the use of these cells in cell therapies [34, 35]. Use of these cells may allow patient specific therapies and circumvent rejection and ethical issues that occur using embryonic cells.

O-GlcNAc

Discovered in the 1980's by Hart and coworkers [36], O-linked β-N-acetylglucosamine (O-GlcNAc) is a single sugar modification of serine and threonine residues of cytoplasmic and nuclear proteins. Unlike most other glycosylation events occurring in the secretory pathway, it is not elongated and is dynamic and inducible. The enzyme for addition, O-GlcNAc transferase (OGT) [37] and the enzyme that removes O-GlcNAc, O-GlcNAcase (OGA) [38, 39], are the products of only one gene in most species. OGT is conserved throughout evolution and is essential for embryonic and somatic cell survival in mammalian cells [40, 41], *Drosophila melanogaster* [42] and *Arabidopsis* [43]. Conversely,

the nematode *C. elegans* does not required OGT or OGA [44, 45] suggesting that the presence of these enzymes is not essential for this organism's survival.

The nucleotide sugar donor for OGT, UDP-GlcNAc, is the endpoint of the hexosamine biosynthetic pathway (HBP) that links intracellular levels of UDP-GlcNAc to the influx of glucose into the cell. This allows O-GlcNAc to function as a nutrient and energy sensor, giving cells a way to monitor and respond to their own energy levels [46]. O-GlcNAc levels can be manipulated in several ways: modulation of HBP by low or high glucose or addition of glucosamine [46], inhibition of OGA by the inhibitors PUGNAc [47], GlcNAcstatin [48], NButGT [49], or Thiamet G [50] or genetic manipulation by overexpressing or knocking down OGT and OGA.

O-GlcNAc regulates many different cellular processes such as cell cycle control [51, 52], stress response [53, 54], cell signaling pathways [55-58] and chromatin remodeling [59-62]. The major class of proteins regulated by O-GlcNAc is transcription factors [63-67]. Regulation of transcription factors occurs by a variety of different mechanisms; stability of the protein [68], nuclear localization [69, 70], DNA binding [71], transcriptional activation [72] and protein-protein interactions [73], reviewed in [63] (Chapter 2).

Since phosphorylation and O-GlcNAc both attach to serine and threonine residues, the two modifications form a complex interplay [74-76]. Some proteins are modified with either O-GlcNAc or phosphate but not both. This can occur on either the same site of attachment, or adjacent sites or attachment. Other proteins are modified by both phosphate and O-GlcNAc modifications on

separate residues at the same time. This suggests that one potential role of O-GlcNAc is not simply to act as a block to phosphate regulation, but to allow the system to fine tune regulation in response to nutrient abundance.

O-GlcNAc and Development

The study of O-GlcNAc in vertebrate development has not been easy using traditional knockout strategies. Generation of OGT knockout mice was unsuccessful, and proved to be lethal at the embryonic stem cell level [40]. Targeted knockouts of OGT using Cre-Lox technology failed to provide any insight either as all targeted cell types underwent cell death [41]. To try and circumvent this issue, a knockdown strategy was undertaken in *Danio rerio* (zebrafish) [77]. Using morpholinos directed toward OGT, O-GlcNAc levels were only moderately decreased, by 30% - 40%, but embryos still showed severe developmental defects. All lineages had reduced cell numbers and a more marked reduction in the amount of Sox17 positive cells indicates that definitive endoderm was most affected [77]. Conversely, the phenotype of the O-GlcNAcase (OGA) knockout mouse is not as severe. OGA homozygous null mouse embryos are born smaller than WT due to a developmental delay, but surprisingly do not have any gross defects at birth. However, OGA null mice die within one day of birth due to respiratory failure [78]. The smaller size of these animals is likely due to cell cycle defects due to increased O-GlcNAc levels [78], similar to those seen with OGT overexpression [79].

Since studying and manipulating O-GlcNAc levels *in vivo* has provided a challenge, most of our knowledge of the role O-GlcNAc plays in differentiation comes from studies undertake in cell culture models.

There are several pieces of evidence that O-GlcNAc is regulating neural differentiation. Strong OGT expression is restricted to the head region in zebrafish [77] and O-GlcNAc, OGT and OGA levels increase until birth in rat brains [80]. Although high levels of OGT are seen in neural cells during development *in vivo*, increasing O-GlcNAc modification itself is detrimental to neural differentiation *in vitro*. Growing neural progenitor cells, or differentiating human ES cells, in PUGNAc or Thiamet G, blocks the ability to express the markers of neural differentiation, Pax6, Msx1, Sox1, NGN-2, TBR2, and LHX3 [81]. OGA expression in primary chick forebrain neurons increased axon branching and outgrowth while the OGA inhibitor, 9D, suppresses this outgrowth through a cAMP dependent mechanism [82]. Both OGT and OGA are found associated with promoter of *hcrt,* a gene responsible for the specification of orexin neurons. OGA associates with p300 and CBP in the active state, and OGT associates with Sirt1, mSin3A and Ezh2 in the inactive state [83]. This result was not surprising since OGT is known to combine with mSin3A and HDAC to decrease transcriptional activity [59]. Furthermore, Ezh2 is a known component of the polycomb repressive complex, PCR2 [84] and the *Drosophila melanogaster* OGT homologue was discovered to be the Polycomb group (PcG) gene super sex combs (*sxc*) [61, 62]. Although these results may seem to be contradictory, the nervous system is extremely complex, with multiple cell types

present. It is entirely possible that increased OGT seen in the brain of zebrafish and rats is required to silence genes required for specialized neuronal types, and studies on a cellular level do not allow for the intricacies of the organ development. Further investigation into the role of O-GlcNAc in multiple cell types is required to fully understand this role.

Use of 3T3-L1 cell is a well established and acceptable model of adipocyte differentiation [85]. When cells are allowed to differentiate using established protocols, O-GlcNAc levels increase globally [86, 87] and on specific proteins involved in adipogenesis including: vimentin, EWS, Nup62, pyruvate carboxylase (PC), C/EBPa, C/EBPβ and PPARγ. Reduction of GFAT activity either genetically or with inhibitors, blocks adipocyte differentiation [86, 87] and decreases expression of C/EBPβ and PPARγ [86]. When human ES cells were differentiated in the presence of PUGNAc/Thiamet G, C/EBPa and PPARγ expression was increased [81]. Taken together this evidence suggests that elevated O-GlcNAc levels enhance adipogenesis.

Recent studies reveal a protective role for O-GlcNAc in cardiac complications in diabetic patients [88], so it is not surprising to find evidence of O-GlcNAc modulation of muscle differentiation. Levels of O-GlcNAc decrease during differentiation of cardiac cells from mouse ES cells [89] or during myotube formation from C2C12 cells [90]. This implies that low levels of O-GlcNAc are required for correct muscle cell specification. Indeed, when these cells were treated with PUGNAc, STZ or Thiamet G, differentiation was inhibited [89, 90]. One such mechanism may be the regulation of cell specific factors of

differentiation. Nkx2.5 is expressed in early heart progenitor cells and signals a commitment to the myocardial lineage [91]. Levels of Nkx2.5 were decreased in heart tissue of STZ treated diabetic mice [92], and when cardiac progenitor cells are formed from mouse ES cells in the presence of STZ or PUGNAc [89]. It was further shown that Nkx2.5 is destabilized when coexpressed in HEK293T cells with OGT [92]. Treatment of C2C12 cells with Thiamet G decreases the ability of a transcription factor required for myogenesis, Mef2D, to bind the myogenin promoter [93]. These studies suggest that increased O-GlcNAc is preventing differentiation to muscle related linages.

Although there is clear evidence that O-GlcNAc regulates differentiation at later time points, not much work has been done studying the regulation in embryonic stem cells, and specification in early differentiation. Oct4, Sox2 and Nanog bind to the OGA promoter in human ES cells [25] and many of the transcription factors important for the regulation of pluripotency and differentiation of ES cells are modified by O-GlcNAc. These include: Oct4 [77, 94], Sox2 [94, 95], c-myc [96, 97], b-catenin [70], p53 [98], and Klf4 [94]. The hypothesis that O-GlcNAc is regulating embryonic stem cells and early differentiation events, and the reasoning for this study, came from a study in zebrafish [77]. When OGT is overexpressed in zebrafish embryos, the epiboly defect and the reduction of endoderm mimics the phenotype seen in embryos deficient for the zebrafish homologue of mammalian Oct4, spiel ohne grenzen (spg)/pou2 [99-101], clearly suggesting a role for Oct4 as a target for O-GlcNAc.

Oct4 - Structure and Function

The gene *Pou5f1,* encoding the transcription factor Oct4 (also known as Oct3 or Oct3/4), was discovered in 1989 in embryonal carcinoma (EC) cells [102] and cloned by three groups [103-105]. Oct4 belongs to the family of transcription factors known as POU (Pit-Oct-Unc) [105, 106] which to date has 15 members in six different classes, sorted by homology of the DNA binding domain [107]. Oct1 is ubiquitously expressed in all cell types, whereas many of the other members are restricted to the brain, liver and skin [107]. Oct4 is the most well known and is expressed exclusively in the developing embryo at the 4 cell stage, the inner cell mass (ICM), embryonic stem cells (which are derived from the ICM), and is restricted to germ cells during differentiation [102]. Knockout experiments in mice show it is important for early embryogenesis, generation of ES cells and maintenance of pluripotency [16, 108]. Along with its role in generating iPS cells, Oct4 expression has recently been discovered to also play a role in enhancing multiple different cancers including ovarian cancer [109], cervical cancer [110], pancreatic carcinoma [111], bladder cancer [112] and breast cancer [113].

The structure of Oct4 is shown in Figure 1.2. It contains two DNA binding motifs, a 75 amino acid N-terminal POU specific region (POU_S) and a 60 amino acid C-terminal POU homeodomain (POU_H), joined by a flexible linker region [104]. Both domains bind DNA using helix-loop-helix motifs to the octamer consensus sequence AGTCAAAT [114, 115], discovered originally in immunoglobulin promoters [114]. This family of proteins can bind DNA as monomers, dimers (PORE or MORE) [116, 117] or heterodimers. The best

characterized being members of the Sox (SRY-related HMG-box) family of proteins [118]. The POU domain of Oct4 has also been shown to participate in protein/protein interactions with E1A [119], Ets-2 [120] and Sox2 [23, 24].

Both the N-terminal domain and the C-terminal domain contain transactivation activity. The N-terminal domain contains constitutive activity and the C-terminal domain contains POU specific, cell line specific activation [121]. Complementation experiments in mouse ES cells suggest that these domains have mainly redundant roles in gene activation, but show some differences in activation of certain reporters (monomer) and genes (LEFTY1) [108] suggesting that modulation of these domains can alter the promoter specificity of Oct4.

Chromatin IP experiments demonstrate that Oct4 can bind to several hundred genes in the genome [15, 25, 26]. Using microarray technology, many of these genes show differential gene expression when Oct4 is knocked down in human or mouse ES cells. These include decreased expression of pluripotency associated genes *Oct4, Sox2, Nanog, DPPA4, LEFTY1/2* and *TDGF1* as well as increased expression of genes associated with differentiation: *WT1, Cdx2, EOMES, DKK1,* and *GSC* [26, 122]. Direct evidence that Oct4 functions as a transcriptional activator for genes required for the maintenance of pluripotency include: *Sox2* [123], *Nanog* [124], *Fgf4* [23], *Utf1* [125], as well as *Oct4* itself [123], and as a repressor for genes associated with early differentiation such as *cdx2* [126] and *opn* [116].

Known Regulation of Oct4 via Post Translational Modifications

Knockdown or overexpression of Oct4 promotes differentiation of both human and mouse ES cells [127]. Increased expression of Oct4 in mouse ES cells can induce differentiation to primitive endoderm and mesoderm whereas a decrease induces trophectoderm [127]. Interestingly knockdown of Oct4 in human cells leads to differentiation to endoderm and mesoderm [128, 129] suggesting that Oct4 action is not equivalent in these two cell types. This points to the need for tight regulation. Like many transcription factors, Oct4 is regulated by the addition of several post-translational modifications: ubiquitination [130-132], SUMOylation [133, 134], phosphorylation [130, 135-137], and O-GlcNAc [77, 94] (Figure 1.2).

Ubiquitin's mostly known for tagging proteins for degradation via the 26S proteasome, although there is also evidence that it is involved in DNA repair, kinase activation, chromatin structure, transcriptional regulation and transport of membrane proteins [138]. It consists of a 76 residue peptide that is covalently added to target proteins by a three-step process consisting of three enzymes: a Ub-activating enzyme (E1), a Ub-conjugating enzyme (E2) and a Ub-protein ligase (E3). The final step in this process creates an isopeptide bond between the terminal glycine of the ubiquitin peptide and the ε-amino group of the lysine residue on the target protein [139]. In most cases ubiquitin molecules are added as a chain (polyubiquitin), however, single additions are found, usually when a non-proteosomal function is involved [138]. Although the actual site of ubiquitination on Oct4 has yet to be determined, Oct4 physically interacts with

the E3 ligase WWP2 at both its N-terminal and C-terminal regions [132]. Coexpression of Oct4 and WWP2 decreases the protein stability of Oct4 [131, 132] in a manner that is dependent on the proteasome [130, 131].

SUMOylation (small ubiquitin-related modifier) as the name suggests, is the addition of a 10kDa peptide, similar to ubiquitin, to mainly nuclear proteins. It is also added in a three-step process, using similar but not identical E1, E2 and E3 enzymes. However, unlike ubiquitin, it is usually only added as a monomeric structure. It's function is hard to predict and there is evidence of it regulating localization, activity or stability of the target protein (reviewed in [140]. On mouse Oct4, it is present on the N-terminal transactivation domain at K118. SUMOylated Oct4 has increased protein stability, DNA binding and transcriptional activation [133]. Further studies show that SUMOylated Oct4 increases Nanog expression in F9 EC cells [141], but its role in human ES cells has not yet been determined.

Phosphorylation is probably the most commonly studied post-translational modification. Addition or removal of a phosphate group on serine, threonine or tyrosine residues by a protein kinase or phosphatase, can regulate transcription factor function by altering DNA binding, cellular localization, protein/protein interactions, protein stability and chromatin structure (reviewed in [142, 143]). Early studies undertaken on mouse Oct4 revealed that when expressed in monkey Cos-1 cells, total protein phosphorylation does not change the DNA binding capacity to the monomer promoter [144], however, phosphopeptide analysis in multiple other cell lines revealed that activation of this

promoter correlates to the phosphorylation status of the C-terminal transactivation domain [121]. Currently, 14 sites of phosphorylation have been mapped on human Oct4 [135, 137]; however, the function of only three of these sites has been determined [130, 145, 146]. Oct4 is phosphorylated on the boundary of the POU homeodomain at residue hS236 (mS229) by an as yet undetermined kinase [130, 135, 147]. Presence of this modification specifically blocks dimer (PORE) promoter reporter activation completely, presumably by blocking dimer formation [130, 148]. Similarly, phosphorylation in the POUs domain of human Oct4 at S193 is responsible for blocking binding to MORE sequences [136]. Extracellular Signal-regulated Kinase (ERK) phosphorylates human Oct4 at S111, which alters localization, increases the amount of ubiquitin present and enhances degradation [135, 145]. ERK has also been shown to phosphorylate human Oct4 at T118, and S355 but no analysis has been done on the consequence of these modifications [135]. Finally, Protein kinase B/Akt phosphorylation of human Oct4 at T235 increases stability and interaction with Sox2, leading to an increase in expression of Oct4 and Nanog [146].

O-GlcNAc modification regulates many transcription factors through a variety of different mechanisms (reviewed in [63] / Chapter 1). One O-GlcNAc residue was mapped to T228 on the POU_H domain of Oct4 purified from mouse ES cells [94] which positively regulates transcriptional activity. Oct4 isolated from human ES cells is modified by O-GlcNAc [77], but the function or location has not yet been determined.

Purpose Of Study

Many of the transcription factors required for pluripotency, including Oct4, are modified by O-GlcNAc, but the role that O-GlcNAc plays has not yet been investigated. This study was designed to specifically determine the site of attachment of O-GlcNAc on human Oct4 and determine the role that is plays in the regulation of Oct4 function, with particular attention to transcriptional activation.

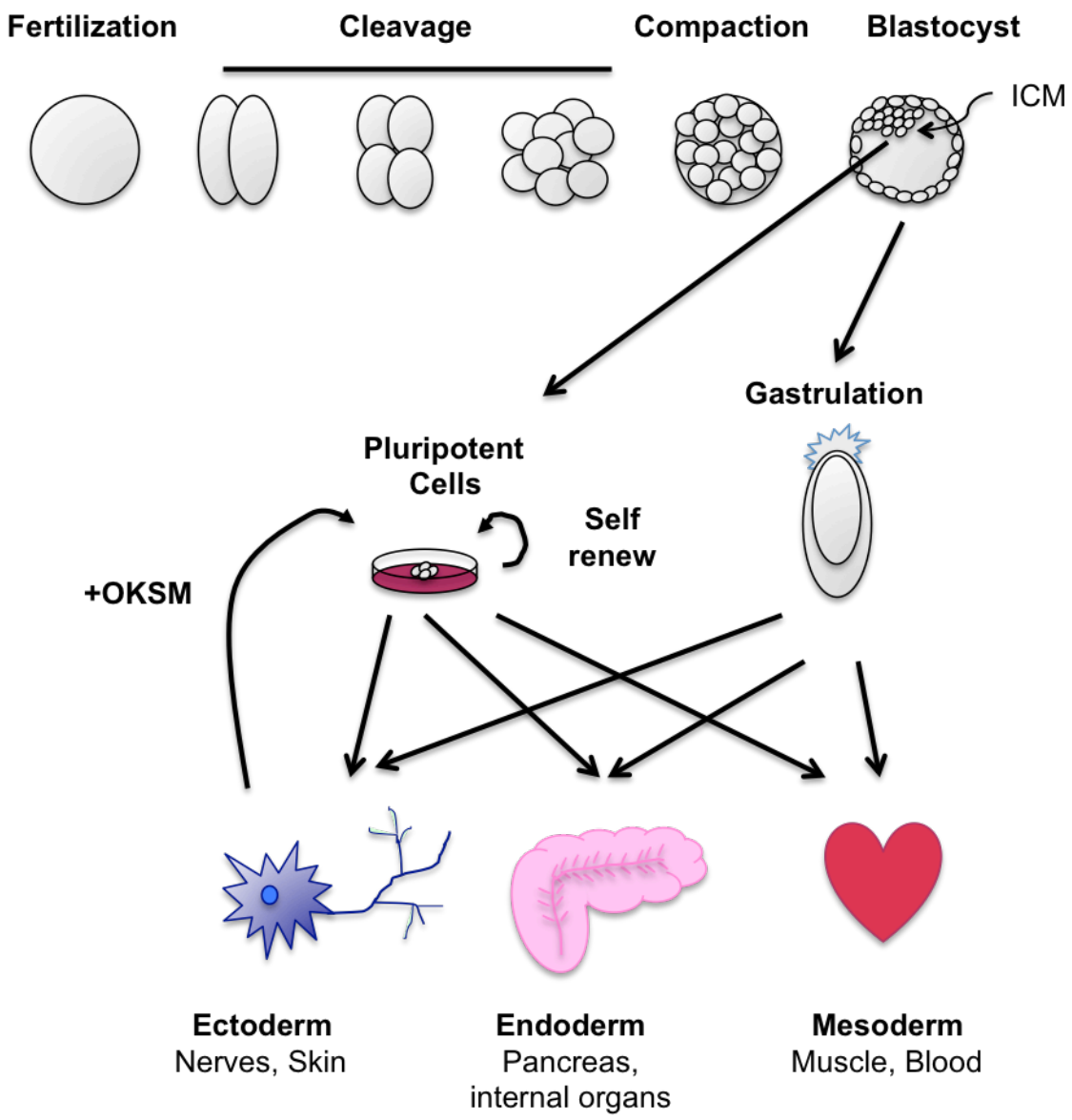


Figure 1.1 – Early Development and Embryonic Stem Cells/Pluripotent Stem cells
After fertilization, the embryo divides in a series of cleavage events to form the blastocyst. Inner cell mass (ICM) cells can be removed and used for derivation of embryonic stem cells. These cells can be used to derive cells from the three major germ layers in the body. Exogenous expression of reprogramming factors Oct4, Klf4, Sox2 and c-Myc (OKSM) can revert differentiated cells back to pluripotent state.

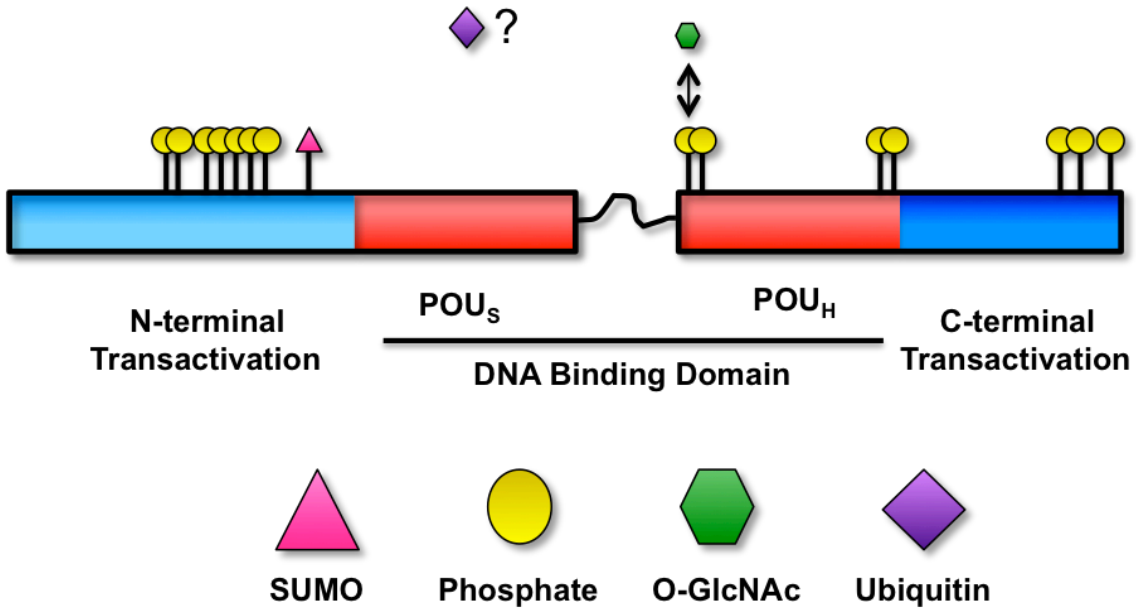


Figure 1.2 - Post Translational Modifications Of Oct4.
Graphical representation of the overall structure of Oct4, including the location of the known post-translational modifications.

CHAPTER 2

THE ROLE OF THE O-GLCNAC MODIFICATION IN REGULATING EUKARYOTIC GENE EXPRESSION

Brimble SN, Wollaston-Hayden EE, Teo CF, Morris AC, and Wells L, 2010 *Curr Signal Transduct Ther*.,. 5(1): p. 12-24.
Reprinted here with permission of the publisher

Abstract

O-linked β-N-acetylglucosamine (O-GlcNAc) modification of proteins has been shown to be involved in many different cellular processes, such as cell cycle control, nutrient sensing, signal transduction, stress response and transcriptional regulation. Cells have developed complex regulatory systems in order to regulate gene expression appropriately in response to environmental and intracellular cues. Control of eukaryotic gene transcription often involves post-translational modification of a multitude of proteins including transcription factors, basal transcription machinery, and chromatin remodeling complexes to modulate their functions in a variety of manners. In this review we describe the emerging functional roles for and techniques to detect and modulate the O-GlcNAc modification and illustrate that the O-GlcNAc modification is intricately involved in at least seven different general mechanisms for the control of gene transcription.

Introduction

Cells have developed a highly regulated system to respond to environmental and intracellular signals to specifically and coordinately express gene products [149, 150]. Surprisingly, the number of protein-coding genes in a genome does not reflect organism complexity, thus it has been hypothesized that increased complexity in gene regulation leads to increased organism complexity [151]. Indeed, the regulation of eukaryotic gene transcription involves a multitude of proteins including transcription factors, basal transcription machinery, and chromatin remodeling complexes [152]. An additional layer of complexity results from a wide variety of post-translational modifications on regulatory proteins [153]. Herein, we describe the emerging role of the O-GlcNAc post-translational modification of nuclear/cytosolic proteins in the regulation of transcription.

In the 1980's, Hart and coworkers reported a nucleocytoplasmic, post-translational sugar modification on serine and/or threonine residues of polypeptides, O-GlcNAc [154-156]. All metazoans currently studied contain the O-GlcNAc modification on proteins involved in many different cell processes, such as cell cycle control [51, 52, 79], nutrient sensing [46], signal transduction [55-58], stress response [53, 54], and transcriptional regulation (the focus of this review) [64-67, 75]. Furthermore, O-GlcNAc transferase (OGT) [157-159], the enzyme required for O-GlcNAc addition, is required for mouse embryonic stem cell viability, emphasizing the importance of this modification [40]. O-GlcNAc is more akin to phosphorylation than complex glycosylation in that it is not elongated, its cycling enzymes, OGT and O-GlcNAcase (OGA) [38, 39], are

nucleocytoplasmic, it is dynamic and inducible, and it can regulate intracellular protein activity, localization, stability, and molecular interactions. O-GlcNAc is often found on the same residues as known phosphorylation sites, suggesting reciprocity between the modifications in some cases, Fig. (**1**) [56, 160, 161]. However, unlike phosphorylation, which is modulated by hundreds of kinases and phosphatases, the cycling of the O-GlcNAc modification is accomplished by the gene products of single genes for OGT and OGA in most metazoans.

O-GlcNAc modification of transcription regulatory proteins could fine tune their regulation in response to nutrient levels in the cell because the synthesis of its sugar donor, UDP-GlcNAc, via the hexosamine biosynthetic pathway (HBP), responds to amino acid, fatty acid, nucleotide and glucose metabolism [46, 65]. There are several ways to modulate O-GlcNAc levels on proteins (for review see [162]) Figure (2.1). OGT is responsive to physiological levels of UDP-GlcNAc, so increased HBP flux by hyperglycemia or by the addition of glucosamine results in globally elevated levels of O-GlcNAc modification [37]. Decreased O-GlcNAc levels can be achieved by blocking glutamine-fructose-6-phosphate transaminase (GFAT), the rate limiting enzyme of the HBP, using the pharmacological inhibitors azaserine or 6-diazo-5-oxo-L-norleucine (DON) or by decreasing glucose levels. However, the alteration of HBP flux may lead to off-target effects as azaserine and DON are general amidotransferase inhibitors. A more specific way to alter global O-GlcNAc levels is by the use of pharmacological OGA inhibitors such as the widely used O-(2-acetamido-2-deoxy-D-glucopyrano-sylidene)amino-N-phenylcarbamate (PUGNAc) [47], or the

more specific inhibitors 1,2-dideoxy-2'-propyl-α- D -glucopyranoso-[2,1- D] -Δ2'-thiazoline (NButGT) [49] and GlcNAcstatin [48, 163]. Several OGT inhibitors have also been recently characterized [164] although their specificity and *in vivo* utility has not been adequately explored. Alternatively, O-GlcNAc steady state levels can be modulated genetically by over expression or knockdown of OGT and/or OGA.

O-GlcNAc Detection and Site Mapping

Over the last 20 years more than 400 proteins have been shown to be modified by O-GlcNAc using a variety of detection methods [65, 165-168]. Interestingly, most RNA Polymerase II transcription factors are glycosylated; many of which respond to nutrient abundance [64, 169]. There are several methods to identify O-GlcNAc modification of proteins [170] and the relevant methods will be briefly discussed here. The first step in identifying O-GlcNAc modified proteins generally involves modification-specific enrichment. Detection or enrichment of O-GlcNAc modified proteins can be achieved using O-GlcNAc specific antibodies, such as RL2 [171, 172] and CTD110.6 [173], and by lectin-blotting or chromatography using succinylated Wheat Germ Agglutinin, a terminal GlcNAc-binding lectin. The presence of O-GlcNAc on proteins can also be determined by labeling with radiolabeled galactose using purified b-1,4-galactosyltransferase (GalT), a galactosyltransferase that transfers galactose onto terminal GlcNAc moieties [154]. Click-iT™ chemistry available from Invitrogen offers two different approaches for *in vitro* and *in vivo* labeling of O-GlcNAc residues. *In vitro* labeling takes advantage of a mutant form of GalT that

transfers ketone-modified galactose onto the GlcNAc residues of proteins [174]. The ketone group introduces a chemically reactive group that can be tagged with biotin and then enriched with streptavidin [174]. Using an *in vivo* approach, introduction of N-azidoacetylglucosamine (GlcNAz) into the cells allows this azidosugar to be converted via the salvage pathway to UDP-GlcNAz and transferred onto proteins by OGT [175]. The azido group of GlcNAz acts as a bio-orthogonal handle for enrichment by the addition of functional groups using the Staudinger ligation [176]. However, there are limitations to using the *in vivo* approach, since it requires the UDP-GlcNAz to compete with the existing UDP-GlcNAc in the cell. These O-GlcNAc enrichment techniques can be combined with mass spectrometry to identify the actual residues of modification [54-56]. Proteomic efforts in this area have identified hundreds of modified polypeptides with proteins involved in transcriptional regulation being a major class [45, 48, 54-56]; however, only about 75 proteins have had their sites of modification mapped. The modification is extremely labile, small, uncharged, and usually substoichiometric [162, 177] making detection difficult using standard mass spectrometry techniques.

Recently, several methods have been developed to make O-GlcNAc site-mapping by mass spectrometry (MS) feasible in biologically relevant tissues. O-GlcNAc enrichment techniques can be combined with mass spectrometry to identify the actual residues of modification [167, 178]. Collision-induced dissociation (CID) mass spectrometry tends to cleave PTMs, so a non-labile tag added to the site of O-GlcNAc modification facilitates identification. Site-mapping

studies using b-elimination followed by Michael addition with dithiothreitol attach a non-labile tag to the site of O-GlcNAc modification so it can be identified by CID MS [168]. In addition, enrichment of O-GlcNAc containing peptides by chemoenzymatic labeling assists in detection [165, 174, 175, 177]. An advantage of these methods is that more O-GlcNAc peptides, which are generally substoichiometric in a total peptide pool, can be detected leading to a more prolific site mapping experiment. The development of electron transfer dissociation fragmentation and related dissociation techniques that often retain CID-labile PTMs have allowed for the identification of O-GlcNAc modified fragments directly [179, 180]. In (Fig. 2.2), we show an example of an electron dissociation technique (electron capture dissociation) for definitively mapping a site of O-GlcNAc on UL32, a synthetic glycosylated peptide, to one particular residue on a peptide containing three potential sites of attachment. Unlike CID, the fragmented peptides containing the modified amino acid retain the mass of the sugar. Electron dissociation techniques are an emerging technology for O-GlcNAc site-mapping that show great promise [177, 180-183].

O-GlcNAc Regulation of Eukaryotic Gene Expression

Specialized transcription factor regulation occurs through the actions of multiple post-translational modifications (PTMs) (reviewed in [153, 184]) such as phosphorylation [142, 143], SUMOylation [185], acetylation [186], and the focus of this review, O-GlcNAc modification. Transcriptional control can occur via at least seven different general mechanisms, (Fig. 2.3), and examples of O-GlcNAc modification participation in each of these regulatory steps are explored below.

Chromatin Remodeling

Chromatin not only provides compact packaging for DNA, it also regulates transcription. For transcription to occur, nucleosomes, the histone proteins/DNA subunits of chromatin, must be positioned to allow transcriptional machinery to access both the promoter and upstream regulatory elements and to allow transcriptional elongation [187]. Access to DNA is regulated by chromatin remodeling enzymes, which recognize PTM's on histones [187, 188]. Acetylation, the most well studied histone PTM, is added by histone acetyltransferases and removed by histone deacetylases (HDACs) [187, 188]. Transcriptional regulation is associated with altered histone acetylation and movement, restructuring, and ejection of nucleosomes [187]. Methylation of certain histone lysines by histone methyltransferases also plays a role in both gene silencing and activation [189]. The actual chromatin remodeling enzymes are thought to be regulated by PTM's such as phosphorylation and acetylation [187]. Several studies have found glycosylation affects the regulation of chromatin remodeling [59, 60].

The first evidence for O-GlcNAc's role in transcriptional regulation was the observation that *Drosophila melanogaster* polytene chromosomes contain more O-GlcNAc modified proteins at the transcriptionally repressed condensed regions than at the active puff regions of the chromatin [190]. Further studies implicated OGT in transcriptional repression through the identification of an interaction between mSin3a and OGT [59]. mSin3a is a corepression scaffolding protein that forms a multi-protein complex with HDAC and can be recruited by

transcription factors to modify histones and repress transcription [191]. Several transcription factors involved in cell survival and apoptosis, such as p53, an O-GlcNAc modified protein [98, 192], recruit mSin3a [193]. The paired amphipathic helix domain 4 of mSin3A was shown to bind to the tetracopeptide repeat (TPR) domain of OGT, suggesting a mechanism where mSin3a recruits OGT for gene silencing [59]. Although both the TPR and catalytic domain of OGT promote transcriptional repression, catalytically active OGT is required for full transcriptional repression [59]. The other proteins in the repression complex, mSin3A and HDAC1, were also found to be O-GlcNAc modified [59] and, although the functional significance is still to be elucidated, may explain why the catalytic activity of OGT is necessary. In agreement with the data seen in *Drosophila melanogaster* polytene chromosomes, a chromatin immunoprecipitation assay showed an increase in both O-GlcNAc modified proteins and mSin3a presence on the promoters of silenced genes [59]. In another study, OGT was found to interact with both mSin3A and Sp3 and was associated with the prevention of transcriptional repression of angiopoietin-2 during hyperglycemic conditions [194]. However, it is unclear whether the association of mSin3A with Sp3 or the direct O-GlcNAc modification of Sp3 was responsible for the transcriptional activation of angiopoietin-2 [194].

A landmark study recently identified a key role for O-GlcNAc in modulating the activity of MLL5, a histone lysine methyltransferase [60]. MLL5 was found to co-activate RARα (retinoic acid receptor α) induction of promyelocyte-like differentiation into granulocyte-like HL60 cells. OGT forms a complex with MLL5.

Elevation of O-GlcNAc levels in undifferentiated HL60 cells increase retinoic acid (RA) stimulated differentiation. Upon RA stimulation, RARα activates the expression of C/EBPε, a major differentiation facilitating transcription factor. Expression of a T440A, the major site of O-GlcNAc modification, mutant of MLL5 failed to activate C/EBPε expression and enhancement of the RA effect on differentiation. Further experiments established that OGT is necessary for MLL5 methylation of H3K4, which allows the transcriptional activation of pro-differentiation genes [60]. Thus, this manuscript clearly illustrates a causal relationship between O-GlcNAc modification of a protein and its enzymatic activity, which is directly involved in chromatin remodeling.

Transcriptional Initiation and Elongation

O-GlcNAc modification has also been implicated in regulating transcriptional initiation via RNA Polymerase II (RNAP II). Transcriptional initiation is achieved in part by several general transcription factors that recruit hypophosphorylated RNAP II to the core promoter and form a pre-initiation complex [195]. RNAP II has a carboxyl terminal domain (CTD) that consists of several tandem consensus sequence repeats that are modified by phosphate and O-GlcNAc [173, 195]. The phosphorylation of the CTD is involved in promoter clearance, passage through promoter proximal pause sites, stabilization of the elongation complex, and recruitment of mRNA processing machinery [150]. The CTD exists in two states with regards to its phosphorylation status; IIO is the phosphorylated form and is found

predominantly in the elongation complex, while IIA is the unphosphorylated form generally found in the initiation complex [195].

When purified fractions of RNAP II were labeled with GalT, it was shown that only the IIA form, the unphosphorylated form, of CTD was modified with O-GlcNAc [196]. In an additional study, OGT failed to label a CTD consensus sequence that had been phosphorylated in vitro by CTD kinase, and CTD kinase would not label a CTD consensus sequence that had been synthetically glycosylated on the Thr 4 of each repeat, suggesting mutual exclusivity between the modifications [173]. This yin-yang relationship between phosphorylation and O-GlcNAc on the CTD suggests that the O-GlcNAc modification may prevent elongation from occurring by blocking phosphorylation or may help to recycle RNAP II after elongation has occurred to allow the complex to reattach to the promoter [173]. Further *in vivo* investigation is needed to clarify the function of glycosylation on the CTD of RNAP II; however, the suggestion that glycosylation regulates transcription initiation is not unprecedented.

Degradation

The proper maintenance of transcription factor levels in cells is often accomplished by degradation via the ubiquitin-proteasome system [197]. Degradation is achieved by two steps: first, ubiquitin is added by an E3 ubiquitin ligase to lysine residues on proteins targeted for destruction, and second, the polyubiquitinylated proteins are degraded by the 26S proteasome [198]. The 26S proteasome is comprised of two major subcomplexes: two 19S regulatory particle caps and the 20S catalytic core [198]. The 20S core catalyzes the

proteolysis of protein substrates. The 19S particle caps contain six ATPases that work to recognize and unfold substrates for entry into the 20S core [139, 198]. Glycosylation and phosphorylation have been suggested to regulate both the activity of the proteasome and the targeting of proteins to the proteasome [67, 199].

The most well-studied O-GlcNAc modified transcription factor is Sp1, a ubiquitous transcription factor for TATA-less genes. Sp1 target genes are involved in many different processes including metabolism, cell proliferation and oncogenesis [200]. In 1988, Jackson and Tjian determined that Sp1 is O-GlcNAc modified [201]. Since then, glycosylation has been described to affect Sp1 function by modulating its stability, protein-protein interactions, DNA binding, and localization [67, 201]. An initial study found that glucose starvation plus adenylate cyclase activation in normal rat kidney cells resulted in decreased Sp1 protein levels and Sp1 hypoglycosylation [68]. The authors suggested that hypoglycosylation of Sp1 promotes degradation through a proteasome-like mechanism [68]. However, it was subsequently shown that the degree of Sp1 glycosylation was independent of its degradation, and instead it was discovered that OGT inhibits and OGA activates the ATPase activity of the 19S regulatory particle caps of the proteasome [199]. OGT catalytic activity is necessary for this inhibition of the proteasome [199]. O-GlcNAc modification of Rpt2, one of the six ATPases present in the 19S cap, blocks the ATPase activity that provides the energy for hydrophobic proteins to unfold and be translocated inside the catalytic core of the proteasome for degradation [199]. Subsequently, in the 26S

proteasome of *Drosophila melanogaster*, five out of nineteen regulatory subunits of the 19S cap and nine out of fourteen subunits of the 20S catalytic core were shown to be O-GlcNAc modified by immunoblotting with monoclonal antibodies and wheat germ agglutinin [202]. O-GlcNAc modification of the proteasome may function to regulate protein degradation in response to nutrient availability, which could potentially regulate transcription by altering transcription factor steady-state levels of transcription factors, such as in the case of Sp1.

Besides its global effect on proteasome function, O-GlcNAc modification is also associated with altered stability of individual transcription factors such as c-Myc, estrogen receptor b (ER-b), and p53. These transcription factors have been shown to be regulated by the ubiquitin proteasome pathway via phosphorylation [98, 203, 204]. A reciprocal relationship between phosphorylation and O-GlcNAc modification is observed for both c-Myc and ER-β [96, 203-205].

c-Myc, a proto-oncogene, was one of the earliest proteins to be site-mapped for O-GlcNAc modification. c-Myc is O-glycosylated on Thr 58 in the N-terminal transcriptional activation domain region [96, 97]. Thr 58 is in the major region of mutation seen in Burkitt's lymphomas [206], and phosphorylation at this site leads to c-Myc polyubiquitinylation and degradation [207]. T58A mutants have increased stability, suggesting that glycosylation via blocking of phosphorylation on this residue may result in increased stability, although the specific mechanism is not known [204]. c-Myc is targeted by several signaling pathways and regulates a plethora of target genes involved in cell proliferation,

differentiation, and apoptosis [208]. Thus, PTM's on c-Myc including phosphorylation and glycosylation appear to influence the specificity and stability of c-Myc [207].

ER-β, an ER-α homologue, is important in many processes such as growth and development, response to stress, and control of energy balance [209, 210]. Phosphorylation of ER-α by GSK-3 (glycogen synthase kinase-3) promotes its stability and full transcriptional activation, and this regulation of ER-α has emerged as an important theme in estrogen signaling [211, 212]. Although this theme is not as well-studied for ER-β, phosphorylation of the ER-β AF-1 domain has been shown to affect its proteasome-dependent degradation [213]. Glycosylation may also play a role in regulating ER-β stability. Ser 16 of ER-b is reciprocally glycosylated and phosphorylated [205]. S16A and S16E mutants were generated to mimic no modification and constitutive phosphorylation, respectively. The S16A mutant had a longer half-life (15-16 hours) and the S16E mutant had a shorter half-life (4-5 hours) than the wild type ER-β (7-8 hours), which suggests that glycosylation may promote ER-β stability by blocking phosphorylation and subsequent targeting for degradation [203].

p53 is a tumor suppressor gene required for cell cycle arrest and apoptosis. Normally, cellular p53 levels, which are highly regulated, are kept very low via degradation by the ubiquitin-dependent proteosome system [214]. Factors such as DNA damage or the activation of oncogenes induce increased p53 stability and activation [214]. p53 is found to be mutated and dysfunctional in many human cancers [214]. An early study determined p53 is O-GlcNAc

modified and the presence of the modification was suggested to increase p53's ability to bind DNA [192]. A later study determined a role for O-GlcNAc modification in p53 stability [98]. p53 is O-GlcNAc modified on Ser 149, which is located on the DNA binding domain. Mutation of Ser 149 to alanine increases Thr 155 phosphorylation. Since elevated Thr 155 phosphorylation is associated with increased degradation of p53, Ser 149 glycosylation has been hypothesized to play an important role in p53 stabilization [98].

Localization

Several papers have been published showing a functional relationship between O-GlcNAc modification and nuclear or cytoplasmic localization [69, 70, 215, 216]. Transcription factors must localize to the nucleus to activate transcription, so sequestering latent transcription factors to the cytoplasm provides an additional mechanism of transcriptional regulation. In response to signals, latent cytoplasmic transcription factors are activated by several mechanisms, many of which depend on phosphorylation or other PTM's, such as glycosylation [149].

The transducer of regulated cyclic adenosine 3'-5' monophosphate response element (CREB) protein (CRTC2) associates with CREB to regulate gluconeogenic genes, including glucose-6-phosphatase (G6Pase), in response to insulin and glucagon [217]. Gluconeogenic genes fail to be inactivated during chronic hyperglycemic conditions, leading to gluconeogenesis during energy prevalent conditions. CRTC2 associates with CREB to bind the cAMP response element on the G6Pase promoter. When insulin is present, SIK2 (salt-induced

kinase 2) is activated by Akt and phosphorylates Ser 171 of CRTC2, which allows it to be sequestered in the cytoplasm by 14-3-3 proteins and targeted for degradation [218]. Glucagon signaling prevents SIK2 from phosphorylating CRTC2 [219]. The dephosphorylated form of CRTC2 is no longer sequestered in the cytosol by 14-3-3 proteins and is free to translocate to the nucleus and activate transcription of target genes. CRTC2 is reciprocally modified by O-GlcNAc and phosphate on Ser 171 and Ser 70, suggesting alternative roles for the modifications. Hyperglycemia or elevating O-GlcNAc levels via genetic or pharmacological methods decreases CRTC2 phosphorylation and increases its O-GlcNAc modification, nuclear localization, and G6Pase promoter activation [69]. Mutation of these sites to aspartate, which simulates phosphorylation, prevents hyperglycemic stimulation of G6Pase promoter activation. Overexpression of OGA in the liver of diabetic *db/db* mice restores their gluconeogenic profiles to nearly normal levels, suggesting that elevated O-GlcNAc levels contribute to the nuclear localization of CRTC2 and the subsequent deregulation of gluconeogenesis during hyperglycemic conditions [69].

O-GlcNAc modification appears to be required for the nuclear localization of NeuroD1 (neurogenic differentiation 1). NeuroD1 is required for the terminal differentiation of neurons and for the development and insulin production of pancreatic b-cells [220]. Hyperglycemia results in increased phosphorylation of NeuroD1 on Ser 274, nuclear translocation, and increased NeuroD1 binding to the insulin promoter. Mutation to S274A results in the cytoplasmic accumulation

of NeuroD1 even in hyperglycemic conditions [215]. Elevation of global O-GlcNAc levels using PUGNAc increased NeuroD1 nuclear localization, binding to the insulin promoter, and insulin expression even in normoglycemic conditions, suggesting that phosphorylation and O-GlcNAc modification are acting cooperatively. This result may be due to a similar increase in NeuroD1 glycosylation in both hyperglycemic and PUGNAc-treated conditions. OGT was found to associate with NeuroD1 in hyperglycemic conditions and OGA was found to associate in normoglycemic conditions [216]. Identifying the NeuroD1 glycosylation sites would help to distinguish whether the effect on localization and subsequent insulin transcriptional activation results from the specific glycosylation of NeuroD1, the interplay between glycosylation and phosphorylation, or from the alteration of global O-GlcNAc levels [216].

b-catenin glycosylation has been shown to regulate its cellular localization [70]. b-catenin plays two major roles in the cell: first, it associates with E-cadherin to form cellular adhesions, and secondly, it is the major downstream signaling molecule for the canonical arm of the Wnt signaling pathway. Wnt signaling pathways are involved in cell growth, movement, and cell survival and are associated with several types of cancer [221]. GSK-3 phosphorylation of b-catenin on its N-terminus targets it for ubiquitination and degradation. Wnt-activated signaling regulates b-catenin by inactivating GSK-3, allowing for the accumulation of b-catenin and its translocation to the nucleus. Here it can activate transcription of target genes by activating TCF (T-cell factor) and recruiting chromatin remodeling proteins [221]. b-catenin has been shown to be

O-GlcNAc modified [222]. PUGNAc treatment of several cancer cell lines resulted in the redistribution of glycosylated b-catenin from the nucleus to the cytoplasm without affecting total protein levels [70]. The increase in cytoplasmic localization was associated with decreased expression of two downstream targets genes, cyclin D and vascular endothelial growth factor A, and decreased promoter activation [70]. More work is needed to determine how the glycosylation of b-catenin influences its interaction with many regulatory binding partners, such as GSK-3 and TCF, and in turn the role of O-GlcNAc in regulating its degradation and transcriptional activation [70].

DNA Binding and Transcriptional Activation

All classical transcription factors share two features: a DNA binding domain for binding to a specific sequence of DNA and a transactivation domain for response to regulatory factors. Sequence-specific transcription factors recruit coactivators to initiate transcription. These coactivators include chromatin remodeling enzymes that are needed to allow the basal transcription machinery to access the DNA and form the pre-initiation complex with the help of additional regulatory proteins [152]. PTM's, such as glycosylation, can affect the ability of transcription factors to bind DNA and activate transcription [153].

The transcription factors PDX-1 (pancreatic/duodenal homeobox-1) protein, NeuroD1, and V-maf musculoaponeurotic fibrosarcoma oncogene homologue A co-regulate insulin transcription. The exact mechanisms of regulation are not clear, which is probably due to the number and complexity of post-translational modifications and cofactor interactions. PDX-1 is necessary for

pancreatic development, and it activates several b-cell specific genes, such as insulin [220]. In response to changing glucose concentrations, PDX-1 recruits chromatin remodeling enzymes and other cofactors and regulates transcriptional elongation. PDX-1 phosphorylation is associated with its translocation to the nucleoplasm and its transactivation potential [220]. PDX-1 is also O-GlcNAc modified on at least two sites [71]. Hyperglycemia or PUGNAc treatment of MIN6 mouse insulinoma cells increases global O-GlcNAc protein levels, enhances PDX-1 binding to the insulin promoter, and is associated with an increase in insulin secretion [71]. The addition of azaserine, which inhibits GFAT and results in lower UDP-GlcNAc levels, decreases global O-GlcNAc levels and glucose-stimulated insulin secretion [71]. Treatment with siRNA against OGT also results in decreased glucose-stimulated insulin secretion, suggesting that the O-GlcNAc modification modulates insulin secretion, perhaps by activating PDX-1 binding to the insulin promoter [71, 223]. O-GlcNAc seems to be extensively involved in β-cell transcription factor regulation and may play an important role in controlling gene expression in response to glucose levels.

Like CRTC2, the forkhead transcription factor family, of which FoxO1 is a member, plays a major role in regulating energy homeostasis [224]. In the liver, FoxO1 and its coactivator, peroxisome proliferator activated receptor γ co-activator 1α (PGC1α), participate in the regulation of gluconeogenesis by activating the expression of G6Pase and phosphoenolpyruvate carboxykinase [225, 226]. Insulin signaling induces Akt to phosphorylate FoxO1 on residues Thr 24, Ser 256, and Ser 319, which results in FoxO1 cytoplasmic localization

[227]. FoxO1 is subject to many PTM's, including glycosylation [228]. Increasing global O-GlcNAc levels by hyperglycemia, PUGNAc, or overexpression of OGT in HEK293 or rat hepatoma cells increases FoxO1 activation of a G6Pase promoter reporter construct [72, 229]. A triple alanine mutant of the Akt phosphorylation sites on FoxO1 is still able to be glycosylated, suggesting that the FoxO1 O-GlcNAc sites are not directly reciprocal with the Akt phosphorylation sites [72]. Consistent with this result, O-GlcNAc modification does not seem to be required for FoxO1 translocation to the nucleus [72, 229]. FoxO1 has been shown to be O-GlcNAc modified on the following residues: Ser 550, Thr 648, Ser 654, and either Thr 317 or Ser 318 [72]. These sites were mutated to alanine and tested for activation of the G6Pase promoter. Only the T317A mutant had a small decrease in promoter activation under hyperglycemic conditions [72]. A follow-up study found that PGC1α interacts with OGT and enhances both OGT interaction and modification of FoxO1 [230]. Coexpression of PGC1α and FoxO1 in HEK293 cells cooperatively increases promoter activation in response to hyperglycemia [230].

Protein/Protein Interactions

Modification of proteins by O-GlcNAc has been shown to modulate protein-protein interactions that regulate nuclear localization [69, 231], stability [98], chromatin remodeling [59, 60], and transcriptional activation [70, 73, 232].

O-GlcNAc modification of Sp1 and b-catenin has been shown to decrease transcriptional activity possibly through inhibition of binding to co-activators [70, 232]. In addition, O-GlcNAc modification of a small peptide segment of Sp1 has

been shown *in vitro* to prevent binding to the general transcription factor TAF110 (TATA-binding-protein-associated factor) [233].

Glycosylation of STAT5a (signal transducer and activator of transcription 5a) was found to be important for its interaction with CREB-binding protein (CBP) [73]. STAT proteins are activated by tyrosine phosphorylation in response to various cytokines and growth factors [234]. They initiate downstream transcriptional activation by dimerizing, translocating to the nucleus and activating transcription partly through the binding to co-activator molecules, such as CBP, that have histone acetyltransferase activity [235]. Mass spectrometry analysis and mutational studies of STAT5a showed that Thr 92 and potentially Thr 97 are O-GlcNAc modified [73]. The mutant T92A prevented STAT5a interaction with CBP and transactivation without affecting DNA binding [73].

NFκB (Nuclear factor κB) signaling has been implicated in a wide range of cellular processes, such as cell immune response, survival, differentiation, and proliferation. In the canonical NFκB signaling pathway, NFκB is normally bound to IκB and sequestered in the cytoplasm [236]. Phosphorylation of IκB by IκB kinase (IKK) leads to IκB degradation via the ubiquitin-proteasome pathway and this allows NFκB to translocate to the nucleus where it can activate transcription [236]. PTM of NFκB subunits can alter transcriptional activation by affecting interactions with transcriptional coactivators and corepressors. NFκB is activated by many pathways, so differential PTMs may specify the particular targets of NFκB. IKK is also regulated by PTMs [236].

Manipulation of the HBP in mesangial cells showed that hyperglycemia increases glycosylation of the p65 subunit of NFκB and promoter activation of a target gene, VCAM-1 (vascular cell adhesion molecule 1) [237]. Hyperglycemia or OGT overexpression decreased the association of the p65 subunit of NFκB with IκB and increased NFκB nuclear localization. Overexpression of OGA in rat vascular smooth muscle cells resulted in lower global O-GlcNAc levels and the reversal of NFκB activation by hyperglycemia. OGT overexpression resulted in the same effects as NFκB activation by hyperglycemia. Mutation of an NFκB O-GlcNAc modification site, Thr 352, to an alanine was found to abrogate promoter activation, DNA binding affinity, association with IκB, nuclear localization, and the expression of VCAM-1 induced by PUGNAc or OGT overexpression [231]. The primary effect of NFκB O-GlcNAc modification may be to prevent p65/IκB interaction, which would lead to nuclear localization and downstream target activation however, more investigation is needed to target the exact mechanism.

A recent paper tied p53 repression of NFκB activation to the O-GlcNAc modification of IKKβ [238]. p53 inactivation leads to an increase in glycolysis through enhanced NFκB activation and results in a positive feedback loop where glycolysis further activates NFκB signaling [239]. The authors proposed that O-GlcNAc modification of IKKβ could be acting as a glucose-sensor to potentiate the feedback loop. In a hepatic cancer cell line, hyperglycemia enhanced IKKβ O-GlcNAc modification and TNFa (tumor necrosis factor α) -stimulated NFκB promoter activation and prolonged NFκB DNA binding and IKKβ activity. Since phosphorylation of IKKβ at Ser 733 is known to inhibit its activation [240], O-

GlcNAc modification of Ser 733 is suggested to prevent phosphorylation-stimulated inactivation leading to an increased in activation of NFκB in transformed cells [238]. These studies establish a clear role for O-GlcNAc in the activation of NFκB.

OGT/OGA Targeting to Substrates – A Special Case of Protein/Protein Interactions

O-GlcNAc modification regulates the function of many target proteins, so aberrant modification by OGT needs to be avoided for proper cellular function. However, the mechanism by which OGT selects its targets is not currently known. No consensus sequence for O-GlcNAc attachment has been found, so it has been proposed that interaction with OGT's TPR domain may determine which proteins it modifies [159, 241, 242]. OGT may also use adaptor proteins that help to modulate its specificity and increase the complexity of its regulation [161, 243]. Cheung *et al.* used a yeast two-hybrid screen to identify proteins that interact with OGT from a human fetal brain cDNA library [243]. Two of the twenty-seven putative OGT-interacting proteins identified, MYPT1 (myosin phosphatase target subunit 1) and CARM1 (coactivating arginine methyltransferase), were shown to interact with OGT and be O-GlcNAc modified by independent methods [243]. Knockdown of MYPT1 using siRNA in Neuro-2a cells reduced the O-GlcNAc levels of several proteins, suggesting that MYPT1 might target OGT to substrates *in vivo* [243]. CARM1 is a histone methyltransferase and functions as part of the p160 coactivator complex, which contributes to chromatin remodeling and transcriptional activation [244]. CARM1

may help to target OGT to substrates that are involved in transcriptional activation [243]. Trak1 (also known as OIP106) was identified by another yeast two-hybrid screen of OGT interacting proteins [242]. Trak1 associates with RNAP II, so it has been proposed that Trak1 targets OGT to the transcriptional machinery [242, 245]. Finally, as mentioned above, PGC-1a may act as an adaptor protein for OGT recruitment to FoxO1 [230].

Although little is known about targeting of OGT to its substrates, even less is known about the regulation of OGA [246]. In some cases, OGT and OGA are found in the same complex [247]. As described above, NeuroD1 can associate with either OGT or OGA depending on glucose concentration [216]. The identification of more OGA-interacting proteins might provide insight into the mechanism of deglycosylation. Using a similar strategy as the OGT experiments, we used a yeast two-hybrid assay obtained from Proquest to identify human OGA binding partners using a cDNA library from human skeletal muscle. Proteins not in frame, proteins identified only once, and proteins known to commonly give false positives were removed from the results. A total of ten proteins were identified by this screen as shown in (Table 2.1). Several of these proteins, including Fragile X mental retardation-related protein 1 (FXR1), Interferon-related developmental regulator 1 (IFRD1), and TANK-Binding Kinase 1 (TBK1)-binding protein 1 (TBKBP1), are relevant to eukaryotic gene expression.

The leading cause of inherited mental retardation is Fragile X syndrome, which is caused by the reduction in an RNA binding protein, Fragile X Mental

Retardation protein (FMRP) [248]. FMRP binds polyribosomes and suppresses translation [249]. FMRP has two homologs, FXR1 and FXR2, which share about 60% sequence homology to FMRP and have been shown to repress TNF translation [250]. Several other RNA-binding proteins, including Ewing-sarcoma RNA-binding protein, eukaryotic initiation factor 4A1, elongation factor 1, and the small and large ribosomal subunits, have been shown to be O-GlcNAc modified, suggesting a possible functional role for O-GlcNAc in post-transcriptional regulation as well [53, 65, 168].

IFRD1 has been shown to play a role in development by induction of differentiation by repression of a specific set of genes through interactions with the co-repressor complex mSin3B/HDAC1 [251, 252]. IFRD1 is implicated in the prevention of Sp1 binding to a common DNA element in IFRD1 regulated genes. It has also been implicated in recruiting HDAC to b-catenin in order to repress its transcriptional activity on downstream targets, such as osteopontin [253, 254]. Since IFRD1 interacts with already known O-GlcNAc targets, it will be interesting to see if the interaction with OGA is required to modify these targets for their function or for interaction with IFRD1.

TBKBP1 was found to interact with TBK1 and inducible IκB kinase (IKKi), which are members of the IKK family that regulate interferon regulatory factor (IRF) [255]. IRF and NFκB coordinate to regulate innate antiviral immunity [256]. TBK1 and IKKi phosphorylate and activate IRF in response to TLR3 (Toll-like receptor 3) activation. Like NFκB, upon activation, IRF dimerizes and translocates to the nucleus to initiate transcriptional activation. TBKBP1, which is

also named Similar to NAP1 TBK1 adaptor (SINTBAD), along with two other cofactors, TANK and NAP1, are needed for full activation of IRF3 in response to the Sendai virus [255]. These cofactors might serve as a link between downstream signaling from TLR3 and activation of TBK1 and IKKi [255]. Since OGA interacts with TBKBP1 and the O-GlcNAc modification is intricately involved in NFκB signaling that is similar to the IRF pathway, it is plausible that the IRF pathway is also regulated by O-GlcNAc modification. Future work will need to establish the relevance of this hypothesis.

Summary

The O-GlcNAc modification of nuclear and cytoplasmic proteins plays a variety of roles in transcription factor regulation including recruiting chromatin remodeling factors, affecting protein stability, changing nuclear localization, and altering DNA binding and transcriptional activation. O-GlcNAc modification can either exert its effects directly on the modified transcription factor or indirectly by altering protein-protein interactions with other modified co-factors. It is becoming increasingly clear that transcription factors do not function in a solely “on” or “off” state but are subject to a number of modifications, such as O-GlcNAc, that fine-tune their regulation [74]. This is advantageous to the cell because transcription factors must interpret a wide range of signals, including nutrient/metabolic signals, and specifically respond to regulate a subset of target genes.

A key feature of the O-GlcNAc modification is that the levels of its sugar donor, UDP-GlcNAc, are directly responsive to the changes in cellular glucose flux. A nutrient sensing ability is valuable for the cell because it prevents it from

being a slave to its extracellular environment [46]. Because altering glucose flux readily modulates global protein O-GlcNAc levels and not just the O-GlcNAc modification on specific proteins, many O-GlcNAc studies to date are correlative. Specific mechanistic and functional studies that show O-GlcNAc modification is indispensable for protein function are beginning to appear in the literature, primarily in relationship to transcriptional control (illustrated above). Advances in the O-GlcNAc site-mapping technology along with the initial experiments for understanding targeting mechanisms for OGT and OGA substrate recognition and the highlighted recent “smoking gun” experiments should facilitate increased interest in understanding functional mechanisms for O-GlcNAc on a wider range of proteins in an increasing number of systems.

Acknowledgments

Thanks to the members of the Wells group for the critical reading of this manuscript. This work was supported in part by a grant from NIH/NIDDK (1RO1DK075069 to LW). CFT is an American Heart Association predoctoral fellow (Southeast affiliate, 0715377B). LW is a Georgia Cancer Coalition Distinguished Scholar.

Table 2.1 **Putative OGA-interacting proteins identified by yeast two-hybrid screen**

gene ID	Symbol	Full Name	Description
12654856	IFRD1	Interferon-related developmental regulator 1	Interacts with corepressor complex
7662301	TBKBP1	ProSAPiP2 protein; TBK1-binding protein 1	NFκB Signaling
33504653	FXR1	Fragile X mental retardation-related protein 1	RNA binding protein, RNA transport
6288762	REV1L	Rev1-like protein	Scaffold for translesion synthesis (TLS) of damaged DNA
18426896	GNAS	GNAS complex locus	G-protein signaling
1730283	COPS8	COP9 signalosome subunit 8	Vesicular transport
34190677	KCNS3	Shab-related delayed-rectifier K+ channel alpha subunit 3	Voltage-gated potassium channel
13528788	MYOZ2	myozenin 2; calcineurin-binding protein calsarcin-1	Interacts with calcineurin: a phosphatase (S/T) calcium/calmodulin dependent
50345685	ATP5B	ATP synthase, H+ transporting, mitochondrial F1 complex	Membrane spanning component
33440538	CAPN7	Calpain 7	Calcium-dependent, cysteine protease

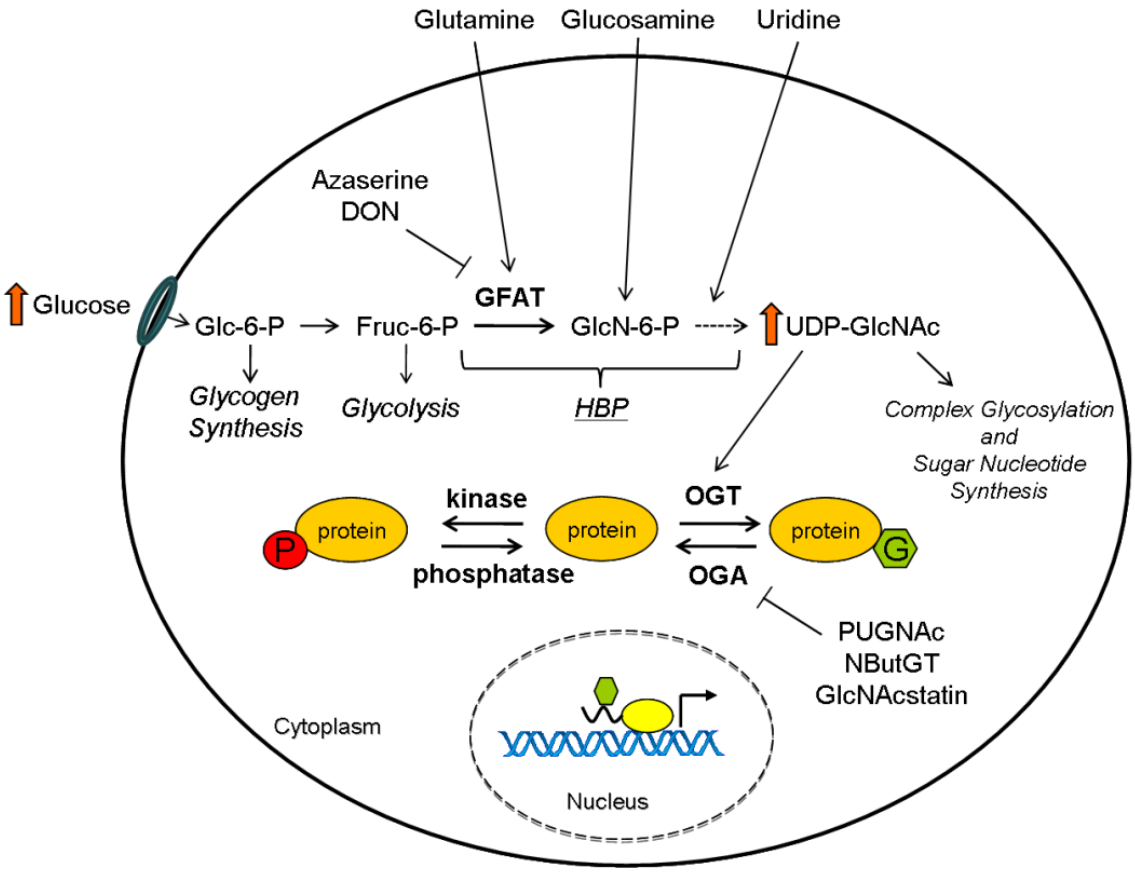


Figure 2.1 - Modulation of cellular O-GlcNAc levels using HBP flux and specific enzyme inhibitors.
The end product of the HBP, UDP-GlcNAc, is sensitive to changes in nutrient levels. Glucosamine enters the HBP downstream of the rate-limiting enzyme GFAT to elevate UDP-GlcNAc levels. The use of the amidotransferase inhibitors azaserine or DON decreases UDP-GlcNAc levels. Proteins can be reciprocally modified by glycosylation and phosphorylation. However, unlike phosphorylation, which is regulated by hundreds of kinases and phosphatases, O-GlcNAc modification is cycled by the result of gene products from only two genes, *ogt* and *oga*. OGT transfers the GlcNAc onto serine and threonine residues of nuclear and cytosolic proteins and is responsive to changes in UDP-GlcNAc concentrations. Global O-GlcNAc levels can also be raised by the use of OGA inhibitors PUGNAc, NButGT and GlcNAcstatin. Enzymes are depicted in bold and biological pathways are in italics.

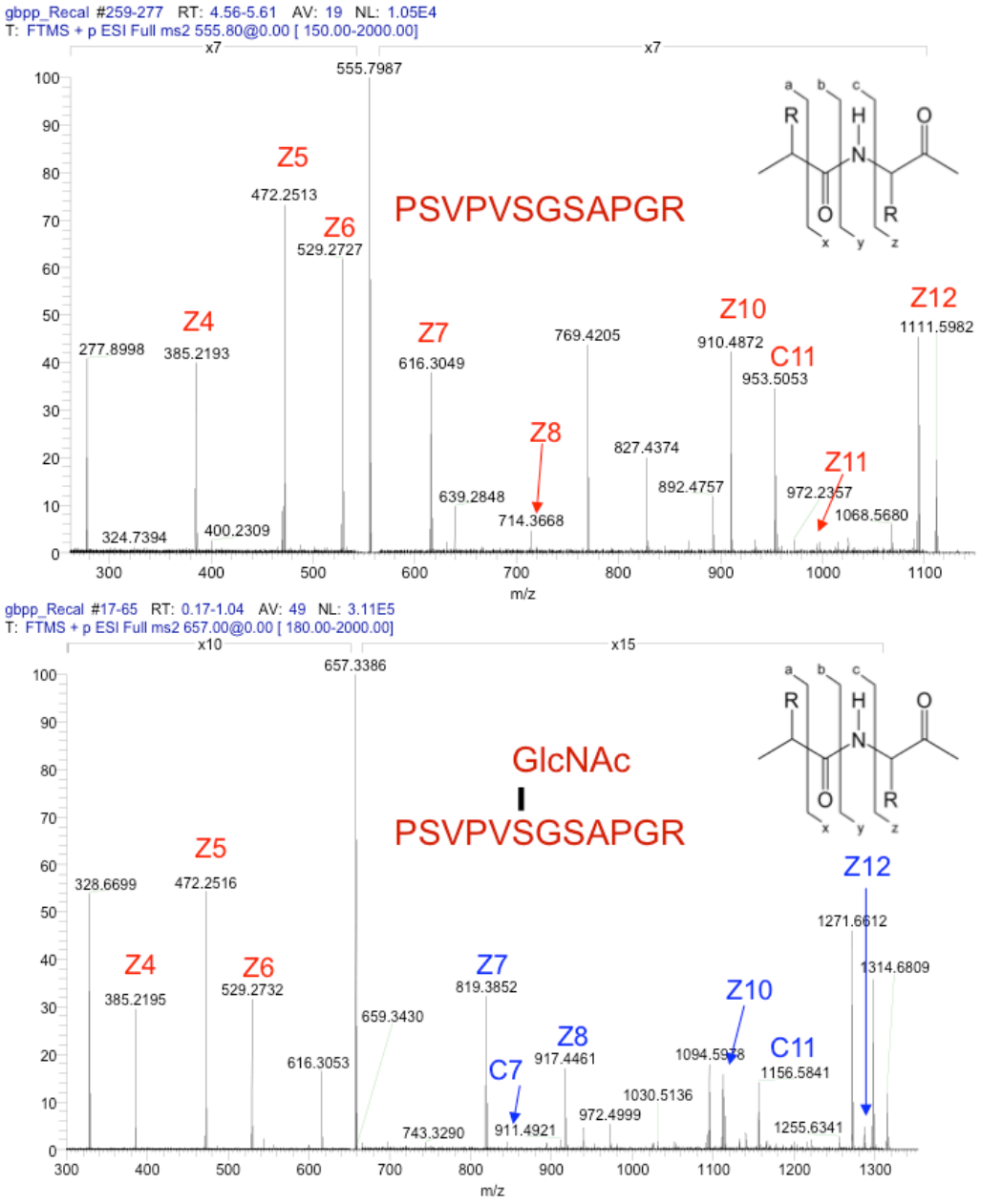


Figure 2.2 Site-mapping of O-GlcNAc sites is facilitated by electron dissociation techniques.

UL32, a synthetic O-GlcNAc modified protein, is efficiently fragmented and the site of modification (from three possible sites) is easily assigned via electron capture dissociation. When comparing the spectra from unglycosylated (top) and glycosylated peptide (bottom), singly charged fragments retaining the O-GlcNAc modified serine (shown in BLUE) show an increase in mass to charge of 203 daltons, the weight of a single GlcNAc residue

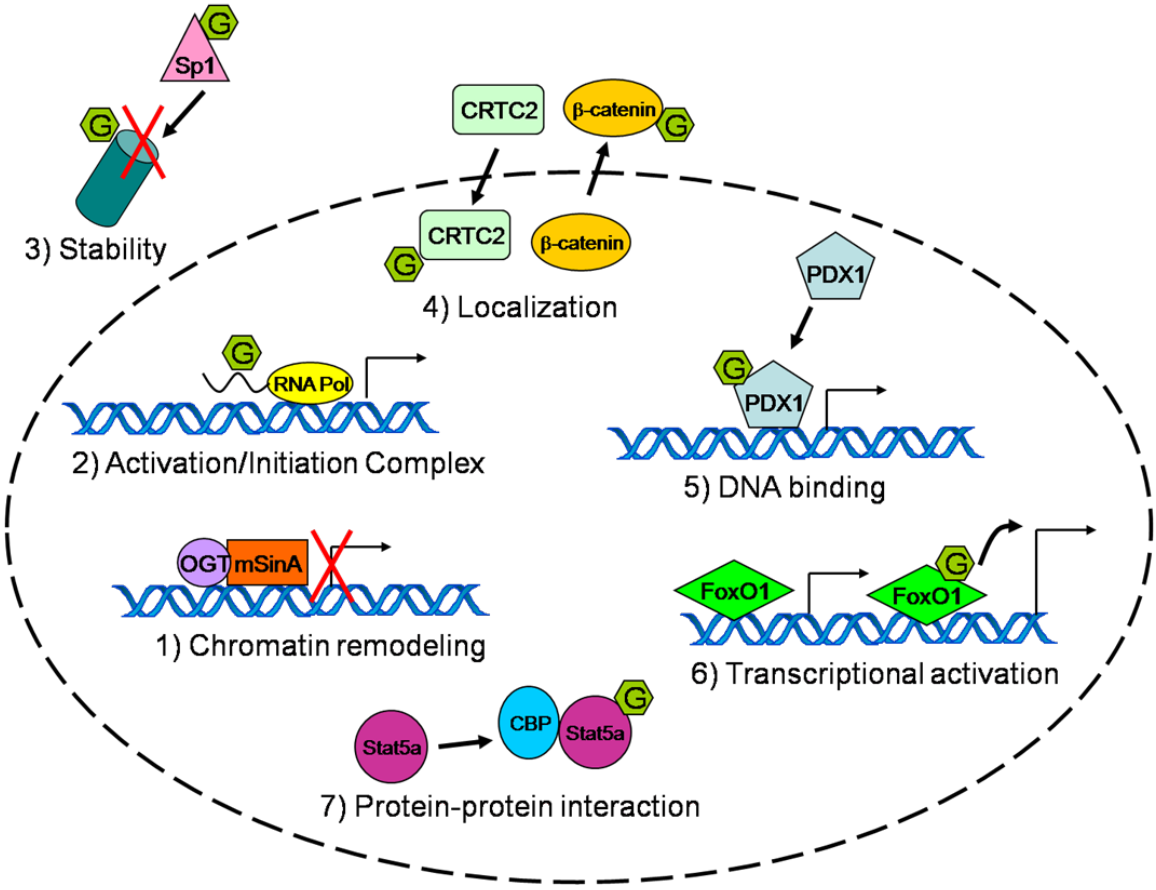


Figure 2.3 - Transcriptional regulation by O-GlcNAc can occur via seven different mechanisms.
The O-GlcNAc modification has been demonstrated to regulate transcription by modulating proteins involved in chromatin remodeling and transcriptional initiation, as well as protein-protein associations, localization, stability, DNA binding, and transactivation capacity of individual transcription factors.

CHAPTER 3

INTERACTION OF HUMAN OCT4 WITH O-LINKED β-*N*-ACETYLGLUCOSAMINE TRANSFERASE REGULATES ITS TRANSCRIPTIONAL ACTIVITY

Brimble SN, Lim JM, Kulik M, Teo CF, Vaidyanathan K, Dalton S and Wells L, To be submitted to *Glycobiology*

Abstract

O-linked β-*N*-acetylglucosamine (O-GlcNAc) is a single sugar modification found on many different classes of nuclear and cytoplasmic proteins. Addition of this modification, by the enzyme O-linked *N*-acetylglucosamine transferase (OGT), is dynamic and inducible. There is mounting evidence that O-GlcNAc plays a role in regulation of development but the mechanism is not clearly understood. One major class of proteins modified by O-GlcNAc is transcription factors. O-GlcNAc regulates transcription factor properties through a variety of different mechanisms including localization, stability and transcriptional activation. Maintenance of embryonic stem (ES) cell pluripotency requires tight regulation of several key transcription factors, many of which are modified by O-GlcNAc. Pou5f1 (Oct4) is one of the transcription factors required for pluripotency of ES cells and more recently, the generation of induced pluripotent stem (iPS) cells. The action of Oct4 is modulated by the addition of several post-translational modifications, including O-GlcNAc. Previous studies in mouse found a single site of O-GlcNAc addition responsible for transcriptional regulation. This study was designed to determine if this mechanism is conserved in human. We mapped 10 novel sites of O-GlcNAc attachment on human Oct4, and confirmed a role for OGT in transcriptional activation of Oct4 at a site distinct from that found in mouse that allows distinction between different Oct4 promoters. Additionally, we uncovered a potential new role for OGT that does not include its catalytic function. These results confirm that human Oct4 activity

is being regulated by OGT by a mechanism that appears independent of O-GlcNAc and distinct from mouse Oct4.

Introduction

Discovered in the 1980's by Hart and coworkers [36], O-linked-*N*-acetylglucosamine (O-GlcNAc) is found in all higher eukaryotes and is a dynamic, single sugar modification found on many different classes of nuclear and cytoplasmic proteins [65, 67, 257]. O-GlcNAc and phosphorylation have a dynamic interplay since they both occur on serine and threonine residues [160, 258]. Unlike phosphorylation, there is only one enzyme required for the addition of O-GlcNAc, O-GlcNAc transferase (OGT) [37], and one for the removal, O-GlcNAcase (OGA) [38, 39]. OGT is essential for embryonic and somatic cell survival in mammalian cells [40, 41], *Drosophila melanogaster* [42] and *Arabidopsis* [43] but interesting not in *C. elegans* [45]. Loss of OGA in mice shows neonatal lethality and developmental delay [78] while OGA mutants are viable in *C. elegans* and *D. melanogaster* [44, 259].

O-GlcNAc regulates many different cellular processes such as: cell cycle control [51, 52], stress response [53, 54], cell signaling pathways [55-58] and chromatin remodeling [59-62]. The major class of proteins regulated by O-GlcNAc is transcription factors and related gene-expression modulators [63-67, 260]. Regulation of transcription factors via O-GlcNAc modification occurs by a variety of different mechanisms (reviewed in Brimble, S.N., Wollaston-Hayden, E.E. et al. 2010) including examples of altering protein stability [68], nuclear localization [69, 70], DNA binding [71], transcriptional activation [72] and protein-protein interactions [73].

During development, Oct4 is expressed only in the oocyte, the inner cell mass and during differentiation is restricted to the germ cells [102]. Oct4 is required for early embryogenesis and maintenance of pluripotency [16], and has been further shown to be one of the key regulatory transcription factors required for pluripotency in mammalian embryonic stem (ES) cells [15, 128, 129]. More recent work has provided a role for Oct4 in the production of induced pluripotent stem (iPS) cells [28, 34, 35]. It belongs to the POU family of transcription factors (Pit-Oct-Unc) [106] that are categorized by their ability to bind to the DNA sequence AGTCAAAT [114]. Oct proteins can bind DNA as a monomer, or in different dimer configurations [148]. Oct4 also frequently works in concert with Sox2 proteins to bind to Oct-Sox DNA element [261]. Oct4 can activate or repress multiple genes which play a role in pluripotency or early differentiation including: *Sox2* [123], *Nanog* [124], *Fgf4* [23], *Utf1* [125], *cdx2* [126], *opn* [116] as well as *Oct4* itself [123]. Small changes in expression level of Oct4 can induce differentiation [127-129] leading to the need for tight regulation. The function of Oct4 protein is regulated by the addition of several post-translational modifications. SUMOylation of Oct4 on the N-terminal transactivation domain, increases protein stability, DNA binding and activation [133, 134]. The site of ubiquitination of Oct4 has not yet been determined, but modification with ubiquitin can negatively impact transcriptional activity and increase protein turnover in human ES cells [130-132]. Oct4 has many phosphorylation sites, most of which have not yet been fully characterized. The sites that have been analyzed thus far

show evidence of transcriptional regulation, cellular distribution and degradation [130, 135-137]. Finally, Oct4 is modified with O-GlcNAc [77, 94].

Several papers have been published providing evidence that O-GlcNAc may regulate Oct4. The first came from our study involving the developmental effects of O-GlcNAc in zebrafish [77]. Overexpression of OGT in zebrafish mimicked the phenotype seen in embryos deficient for the Oct4 homologue spiel ohne grenzen (spg)/pou2 [77, 99-101]. This study also showed that Oct4 protein, isolated from human ES cells, is reactive to an O-GlcNAc specific antibody [77]. More recently, Jang and colleagues mapped one site of O-GlcNAc attachment was mapped to residue T228 on Oct4 purified from mouse ES cells. They showed that transcriptional activation of mouse Oct4 correlates with the level of O-GlcNAc present [94]. Finally, two independent groups found Oct4 associated with OGT during proteomic screens designed to determine the Oct4 interactome [262, 263]. Oct4 is conserved in both mouse and human embryonic stem cells, though its targets and function vary depending on the species [264] suggesting a need to study human Oct4 regulation. Although human Oct4 is modified by O-GlcNAc [77], the actual site of attachment or the functional implications of this modification have not yet been determined. This study was designed to map the site of O-GlcNAc attachment on human Oct4, and to determine the impact O-GlcNAc has on human Oct4 transcriptional activation. In addition to mapping 10 unique O-GlcNAc sites and confirming its role in transcriptional regulation, we discover that OGT catalytic function is not required for activation of certain promoters, opening up a new area of investigation in this field.

Results

Reactivity of O-GlcNAc specific antibody on Oct4 protein can be increased by OGT overexpression

To determine the effect that O-GlcNAc has on human Oct4, we developed several test systems. We first looked at H9 human embryonic stem cells. O-GlcNAc modification of Oct4 isolated from H9 cells was confirmed by western blot analysis of immunoprecipitated Oct4 proteins using the O-GlcNAc specific antibody CTD110.6 (described previously [265]) (Fig 3.1A). Overexpressing OGT in H9 cells increases the amount of reactivity with O-GlcNAc antibody on many different cellular proteins (Fig 3.1B). However, to determine the effect that O-GlcNAc has on Oct4, we needed to ensure that our manipulations would change the amount of O-GlcNAc present on Oct4 itself. Oct4 immunoprecipitated from H9 cells overexpressing OGT show an increase in 110.6 reactivity, validating our system (Fig 3.1C). Since transfection efficiency in human ES cells is low, approximately 30% (data not shown), we moved our analysis into HEK293T cells. These cells are more easily manipulated and have no endogenous Oct4 background [16, 102]. Similarly to the H9 cells, OGT overexpression caused a large increase in 110.6 reactive proteins in HEK293T cell lysates (Fig 3.1E). Immunoprecipitation of exogenously expressed human Oct4 reveals that Oct4 is modified with O-GlcNAc by endogenous OGT (Fig 3.1D). When OGT is coexpressed, the reactivity to the O-GlcNAc antibody increases (Fig 3.1F). This allows us to use HEK293T cells since it is a more

easily manipulated system with null background for Oct4 to characterize the properties of Oct4 when modified by O-GlcNAc.

OGT overexpression increases Oct4 transcriptional activation in HEK293T cells

The transcriptional activation of mouse Oct4 increases when OGT is coexpressed [94]. This property was tested with human Oct4. We used three different Oct4 activating luciferase reporters: a promoter that binds Oct4 as a monomer (6W), a homodimer (PORE) and a heterodimer in which Oct4 co-operates with a Sox protein family member to activate (Oct/Sox). Luciferase activity was increased over control (EGFP) with the monomer promoter when Oct4 was expressed alone. This activity could be enhanced 2.5-fold with co-expression of OGT (Fig 3.2A). Expression of Oct4 alone did not appear to activate the homodimer promoter, however coexpression with OGT increased activity by 5-fold (Fig 3.2B). Using the heterodimer (Oct/Sox) promoter, expression of Oct4 alone showed no change in activity and did not appear to activate the promoter above control levels (EGFP) (Fig 3.2C). Expression of Sox2 alone showed an increase over control, presumably due to its interaction with Oct1 which is present in all cell types [266] and which has previously been shown to interact with Sox2 [267]. Coexpression with OGT does not increase this activity. When Oct4 is co-expressed with Sox2 there is no increase over the result seen with Sox2 alone, but co-expression of OGT with both Sox2 and Oct4 increases the activity 2-fold. Taken that the condition with Sox2 alone does not increase when OGT is coexpressed, the activation seen can be attributed to

Oct4. These results together suggest that human Oct4 transcription is activated by OGT overexpression at a variety of different promoter types.

hOct4 is Modified Beyond Known mThr228 Site

Using the sequence comparison tool ClustalW2, human Oct4 is 100% conserved in the region of the previously mapped O-GlcNAc site on mouse Oct4 [94] (Fig 3.3A). Since human Oct4 transcription is also activated by OGT, we wanted to test if the known O-GlcNAc modification site on mouse Oct4 was also required for the transcriptional activation seen with human Oct4. Using site-directed mutagenesis, we mutated this site to an alanine to prevent O-GlcNAc addition. Since OGT has been shown to be promiscuous in its addition of O-GlcNAc [203] we also mutated S236 and T235/S236 (TSAA) in combination. Immunoprecipitation and Western blot analysis of the mutant Oct4 in HEK293T cells revealed that all the mutants are still 110.6 reactive (Fig 3.3B). This confirms there are more sites of O-GlcNAc on this protein. Surprisingly, densitometry quantification of the western blots reveals that S236A and the TSAA double mutant show higher levels of modification than WT and T235A Oct4 (Fig 3.3C). To test the transcriptional activation of these mutants we used both the monomer and Oct/Sox promoters. As was seen previously with the mouse constructs [94], the T235A and TSAA mutants showed a decrease in transcriptional activation (Fig 3.3D). The transcriptional activation for all of the mutants using the heterodimer Oct/Sox promoter was the same as WT (data not shown). One major issue in the field is using mutation to alanine to study O-GlcNAc effects. This mutation prevents both O-GlcNAc and phosphate addition

making the results ambiguous. Many people use aspartic or glutamic acid substitution to mimic phosphorylation, however no such substitution is available for O-GlcNAc. To circumvent this issue, we used coexpression of OGT to distinguish between the two. If O-GlcNAc modification at the site of mutation is responsible for the decrease in transcriptional activity, then coexpression with OGT should not be able to induce transcription activation. When we used this principle in our system, coexpression of OGT increased transcriptional activity with all of the constructs tested (Fig 3.3E). An increase was also seen when the TSAA mutant was co-expressed with OGT and Sox2 using the Oct/Sox promoter (Fig 3.3F). Although T235A showed a decrease in transcriptional activation using the monomer promoter, the increase in activity seen when the mutant hOct4 is coexpressed with OGT suggests that O-GlcNAc is acting at a site other than T235 to activate transcription.

Oct4 is Modified at Multiple Sites

The ability of OGT to induce transcriptional activation with our TSAA mutant human Oct4 prompted us to determine the other sites of O-GlcNAc attachment. Site mapping O-GlcNAc sites on proteins such as transcription factors is extremely difficult due to the low stoichiometry and low abundance of many proteins in the cell. Initial mass spectrometry analysis undertaken on Oct4 protein immunopurified from H9 cells showed evidence of several peptides modified by O-GlcNAc, as evidenced by neutral loss. However, definitive sites of attachment could not be determined from this data (data not shown). Instead, our study used immunoprecipitated proteins from HEK293T cells expressing

Oct4 protein coexpressed with OGT to increase the abundance of O-GlcNAc modification on Oct4. Previous studies have shown that coexpression of OGT can increase the stoichiometry and will add O-GlcNAc to bona fide O-GlcNAc sites [268]. Peptides containing 10 novel O-GlcNAc sites and 3 unique phosphorylation sites, one of which is novel, were found using a mixture of CID and ETD techniques (Fig 3.4G & H, Table 3.1). Representative spectra from two peptides are shown: Full MS showing parent mass (Fig 3.4A & D), CID spectra showing neutral loss of the HexNAc ion(s) (Fig 3.4B & E) and ETD spectra showing the peptide sequence with O-GlcNAc assignment (Fig 3.4C & F). ETD analysis of the peptide TLVQARKRKRTSIE mapped an O-GlcNAc residue to S236. This site is adjacent to the homologous site in the mouse [94] and is a known site of phosphorylation [135, 137]. The C-terminal peptide, GEAFPPVSVTTLGSPMHSN, was modified with multiple O-GlcNAc and phosphorylation residues (See Table 3.1). Of note, three O-GlcNAc residues were mapped simultaneously to T351, T352 and S359.

Transcriptional ability of O-GlcNAc mutants is unchanged.

To determine the impact of O-GlcNAc on transcriptional activation of the sites mapped, we undertook site directed mutagenesis to make the mutants summarized in Table 3.2. Again due to the promiscuous nature of OGT, we mutated serine or threonine residues that are adjacent or close to the mapped site. For one peptide we could not assign the exact site of O-GlcNAc attachment, however, this peptide only has three possible sites of attachment, S288, S289 or S290, so all three were mutated for analysis (Table 3.2). Due to

the large number of modifications present on the C-terminal peptide, none of the sites were included in this analysis. All of the mutant proteins expressed and immunoprecipitated in HEK293T cells show 110.6 reactivity, confirming multiple sites of O-GlcNAc modification (Supp Figure 3.1). Transcriptional analysis undertaken as previously described in Figure 3.2, also revealed that all of the mutants show increased activity when coexpressed with OGT (data not shown). This revealed to us that the sites mutated in this study are not exclusively involved in transcriptional activation due to OGT overexpression.

OGA inhibitor GlcNAcstatin does not alter embryonic stem cell self renewal or early differentiation

Oct4 transcription is tightly regulated in ES cells, and small changes in Oct4 expression levels have a big impact on self-renewal and differentiation [127-129]. Since OGT enhances Oct4 transcriptional activity (Fig 3.2), we hypothesized that this would have an effect on embryonic stem cell pluripotency or differentiation. We overexpressed OGT in H9 cells for 48 hours and looked at RNA and protein levels of pluripotency markers. Although we saw an 8-fold increase of OGT expression in these cells, pluripotency cell markers Oct4, Sox2 and Nanog were unchanged (Supp Fig 3.2). As mentioned previously, the transfection efficiency of H9 cells is only about 30%, so the lack of change in the two populations is not surprising, as the remaining cells may mask the results. To circumvent this problem, we treated cells with the specific O-GlcNAcase inhibitor GlcNAcstatin (GNS) [48, 163]. After 24 hours of GNS treatment, global O-GlcNAc levels on proteins increase (Fig 3.5A), as well as levels of O-GlcNAc

on Oct4 (Fig 3.5B). To determine the effect of GNS treatment on pluripotency and early differentiation markers, H9 ES cells were treated for 3 days with or without GNS. These cells showed elevated O-GlcNAc levels as determined by 110.6 western (Fig 3.5C). However, pluripotency markers Oct4, Sox2 and Nanog did not change from untreated cells for both transcript levels (Fig 3.5D) and steady state protein levels (Supp Fig 3.3). This result was confirmed by immunofluorescence of H9 cells using early differentiation marker brachyury (T) and pluripotency marker Nanog (Fig 3.5E). Transcript levels of early differentiation markers were also unchanged (Supp Fig 3.4). This suggests that treatment of ES cells with GNS does not affect pluripotency genes, nor does it promote early differentiation. Next we wanted to see if programmed differentiation could be affected by GNS treatment. H9 cells were differentiated to definitive endoderm (DE), or cardiac progenitor cells (Isl1) using established protocols [269](Dalton, unpublished) with or without GNS treatment. As expected, day 4 differentiated cardiac progenitor cells were positive for the cardiac marker Isl1, but not the DE marker FoxA2 (Fig 3.6A). Treatment with GNS had no effect on these markers (Fig 3.6B). Similarly, GNS treatment during DE differentiation had no effect (Fig 3.6 C&D). There was no change in the differentiation timing as day 2 differentiated cells showed similar results (Supp Fig 3.5). Taken together, these results strongly suggest that elevation of O-GlcNAc levels using the inhibitor GNS does not alter pluripotency or early differentiation ability of human ES cells.

OGT activity is not required for transcriptional activation.

Since GNS showed no effect on pluripotency or differentiation of ES cells we wanted to determine if the transcriptional activation of Oct4 was altered by GNS. To our surprise, treatment of H9 cells with GNS does not increase Oct4 reporter constructs (Figure 3.9A). To confirm this and test cell type specific difference, we repeated these experiments in HEK293T cells with GNS and saw no transcriptional induction (Supp Fig 3.6). This suggested to us that OGT increases transcriptional activation by inducing a normally unmodified O-GlcNAc site to become modified, it uses an indirect method of activation (modifies another protein not normally glyocsylated) or by a mechanism that is OGT dependent but O-GlcNAc independent. To distinguish between these possibilities we repeated the experiment using a catalytically inactivated mutant, OGT H567A, described previously [270] (Supp Fig 3.7). In H9 cells, the monomer (6W) promoter shows a moderate decrease in activity when OGT is expressed. When OGT H567A is expressed, there is no change from the control (Fig 3.7B). This suggests that catalytic activity of OGT is important for this promoter. The heterodimer (Oct/Sox) promoter showed an increase in activity regardless of which OGT was co-expressed suggesting that the catalytic activity of this enzyme is not important for promoting activation of this promoter. Together these results demonstrate that OGT and O-GlcNAc modification can alter the transcriptional ability of Oct4 via different mechanisms for different promoters.

Discussion

Site mapping of human Oct4 revealed 10 unique O-GlcNAc sites that have the potential to fine-tune Oct4 function as they are involved in, or are in proximity to, other characterized modifications. The reciprocal nature of O-GlcNAc with phosphorylation has been well documented [160, 258] and although a direct link between SUMO and O-GlcNAc has not yet been discovered, there is a known link between O-GlcNAc and ubiquitin [271, 272], which has a similar mechanism of attachment. With the exception of S335 and S349, all other residues mapped in this study are also modified by phosphorylation, shown either in this study or a previously published study [135]. T235, S355 and S289/S290 are modified by AKT, ERK and PKA/PIM1 respectively [135, 146], although the actual consequence of phosphorylation at these sites has not yet been investigated. Considering the importance these kinases play in signaling pathways responsible for maintaining pluripotency/differentiation [273], blocking phosphorylation by modification with O-GlcNAc at these sites is likely to play an important regulatory role. One peptide containing O-GlcNAc modification at T116 also contains several other known modifications that modulate transcriptional activation. T116 is in proximity to the homologous O-GlcNAc site we previously mapped on Oct1 [274] which, in combination with another un-conserved site, was shown to be important for transcriptional repression. K123 is homologous to the mouse Oct4 SUMOylation site found to enhance DNA binding and transactivation activity [133, 275], specifically an Oct4 dependent increase of Nanog expression was seen in F9 embryonal carcinoma cells [141].

In our study we mapped the O-GlcNAc on the serine adjacent to the threonine mapped in mouse [94]. Since both T235 and S236 are also phosphorylated, changes in cell signaling pathways between the different cell types may activate different kinases impacting the yin-yang relationship of phosphorylation and O-GlcNAc. Transcriptional activity of mouse Oct4 is increased when O-GlcNAc modified at T228 [94]. Unlike the mouse protein, the homologous site on human Oct4 is not solely responsible for its transcriptional activity when OGT is expressed. Furthermore, we did not see any correlation between O-GlcNAc reactivity and transcriptional activation (Fig 3.3C and D). There was also a decrease in transcriptional activation by OGT in H9 cells (Fig 3.7) rather than the increase seen in HEK293T cells using the monomer promoter (Fig 3.2). The difference shown here may be due to a difference in signaling pathways between the two cell types. It has been shown previously that Oct4 is differentially phosphorylated in different cell types, which correlates to its activity [121]. The difference of modification in different cell types should be investigated further in future studies and will make selection of a system for studying function important.

Our data set showed an abundance of O-GlcNAc modifications on the C-terminal peptide (Fig 3.4). 5 of the 10 sites mapped lie within the terminal 12 residues, with one peptide being modified by as many as three O-GlcNAc residues. Although unusual it is not unprecedented. Recent papers have shown peptides containing three O-GlcNAc sites in close proximity on Host cell factor 1 (HCF1) [276] and histone 2B [277], and HCF [95] and C/EBPb [278] both have

O-GlcNAc sites on adjacent residues. Due to the number of sites mapped in this region, it was not investigated in this study but future studies should concentrate on this region. Previous studies have shown that the C-terminal transactivation domain is required for full activation [108], and based on our inability to modulate OGT-induced increases in activation may suggest that O-GlcNAc modifying this region will have a significant impact on its transcriptional activity.

OGA inhibitors are widespread in the field and are usually used interchangeably with OGT overexpression to modulate O-GlcNAc levels. Originally streptozotocin and PUGNAc were used, but more recently their use is limited due to the off target effects seen [49, 279]. Currently, newly designed inhibitors such as GlcNAcstatin and Thiamet G [50] have taken over as the inhibitors of choice as they are much more specific [48]. Curiously, although the use of GlcNAcstatin yielded increased O-GlcNAc levels, it did not yield the same results as OGT overexpression in our transcriptional activation assays (Fig 3.7). Furthermore, there was no effect seen on self-renewal or early cell differentiation to mesoderm or endoderm when we treated cells with GlcNAcstatin (Fig 3.5 and 3.6). One study published during the preparation of this paper confirmed that self-renewal of human ES cells is not altered in the presence of PUGNAc, and that spontaneously differentiated cells are only affected after 7-14 days of treatment with the most prominent lineage affected being the neural linage which we did not investigate in this study [81]. Another recent study saw an impairment in neural differentiation of mouse ES cells when in the presence of GNS [280]. These studies suggest our results were due to the short time frame of the

experiments, or that the neural linage is the most susceptible to perturbed O-GlcNAc levels. There is mounting evidence pointing toward O-GlcNAc being important in neural development. During zebrafish development, strong OGT expression is restricted to the head region [77] and high OGT levels are seen in mouse brains compared to other adult tissue [40]. OGA increases axon branching in chick neurons when overexpressed [82], and was found associated with factors that specify orexin neurons derived from mouse ES cells [83]. O-GlcNAc has also been found to be important in neurological diseases such as Alzheimer's [281], and mutations in OGT have been seen in X-Linked intellectual disability patients (Vaidyanathan, in press). Future studies should focus on the role in neural development.

Although GNS increased O-GlcNAc modification on Oct4 (Fig 3.5), unlike OGT overexpression, GNS failed to alter transcriptional activity of Oct4 (Fig 3.7). This suggests there is another function of OGT separate from its ability to modify Oct4 with O-GlcNAc and cautions the use of these two methods interchangeably. Unlike the studies undertaken with OGA inhibitors, knockdown of OGT in mouse ES cells saw increases in the following genes important for differentiation: Gata6, Sox17, Brachyury, MixL1, Cdx2, EOMES, Sox1, Mash1 [282]. Although both methods lead to an increase of O-GlcNAc, they do so by very different mechanisms. The inhibition of OGA increases O-GlcNAc levels by breaking the cycle and preventing removal of O-GlcNAc. Overexpression of OGT increases O-GlcNAc levels by changing the ratio of OGT to OGA in the cell. The use of an inactive OGT mutant in our transcription experiments points to another function

of OGT that does not rely on its catalytic activity. mSin3A has been shown to enhance transcription of the Oct/Sox promoter [283], and OGT has been shown to interact with mSin3A/HDAC complex to regulate transcription [59]. It is entirely possible that OGT acts as a bridging protein between mSin3A, Oct4 and Sox2 to bring them to this promoter. Indeed, Oct4 has been shown to require a bridging factor for full activation [119], and two independent groups found OGT bound to Oct4 when looking for interaction partners suggesting these proteins form a complex [262, 263]. Further investigation into the non-enzymatic functions of OGT and defining its interactome will be required in future studies.

Conclusion

In this study we discovered that human Oct4 is highly modified by O-GlcNAc and assigned 10 unique sites of attachment, including 2 adjacent sites present on the C-terminal transactivation domain. Coexpression of OGT increases the transactivation ability of Oct4 in both HEK293T cells and H9 cells, however the mechanism may not be linked to O-GlcNAc modification of Oct4. Additionally, we uncovered a potential new role for OGT that does not include its catalytic function.

Materials and methods

Cell Culture and Transfections

H9 ES cells were maintained on Matrigel™ (BD biosciences) in StemPro® hESC media (Life Technologies) using Accutase™ passaging (ICT). H9 cells were differentiated to definitive endoderm using protocol previously published

[269] and Isl1 cells (Stephen Dalton, manuscript in process). HEK293T cells were maintained in 10% FBS/DMEM. Transfections were carried out using X-tremeGENE HP DNA Transfection Reagent (Roche) or JetPRIME (Polyplus) using manufacture protocols. Cells were treated with GlcNAcstatin (GNS) (Gift from Daan van Aalten, University of Dundee, UK) by adding 100nM directly to the appropriate media every 24 hours.

Immunoprecipitation, Immunofluorescence, Western Blotting and Quantification

Immunoprecipitations were carried out using 1mg of protein in Tris buffers containing 1%NP40, 0.1%SDS. Immunofluorescence and western blotting was carried out using standard conditions. Antibodies used in this study: Oct4 (Santa Cruz), 110.6 and HA (gift from Gerald Hart), Sox2 and Nanog (R&D Systems-WB, ReproCELL-IF), Brachyury and Isl1 (R&D Systems), FoxA2 (Millipore). ImageJ software (NIH) was used for the quantification of film exposures. p values were determined using standard Student's t-test undertaken on at least 3 biological replicates.

Luciferase Assays

6W and PORE luciferase constructs were kindly donated by Dr Jonathan Saxe [130]. Oct/Sox promoter was obtained from Addgene (plasmid 15686) [284]. Luciferase expression was detected using Promega Dual Glo® Luciferase Assay System according to manufacture instructions. Student's t-test was carried out in excel on a minimum of biological triplicate samples.

Sample preparation for analysis of mass spectrometry

Human Oct4 was co-expressed in HEK293T cells with human OGT. Immunoprecipitation of ten 10cm plates was carried out as described above, eluted with 0.1M Glycine pH 2.5 and neutralized to pH 8.0 with Tris. The eluted samples were reduced with 10 mM dithiothreitol (DTT) for 1 hour at 56 °C, carboxyamidomethylated with 55 mM iodoacetamide (ICH_2CONH_2, Sigma) in the dark for 45 minutes, and then digested with 3 µg of sequence grade Glu-C (Promega) in 100 mM phosphate buffer at pH 7.0 overnight at 37 °C. After digestion, the peptides were acidified with 1% trifluoroacetic acid (TFA). Desalting was subsequently performed with C18 spin columns (Vydac Silica C18, The Nest Group, Inc.) and the resulting peptides were dried down in a Speed Vac and stored at -20 °C until analysis.

O-GlcNAc site mapping of Oct4 in HEK by LC-MS/MS

The peptides resuspended with 19.5 µL of mobile phase A (0.1% formic acid, FA, in water) and 0.5 µL of mobile phase B (80% acetonitrile, ACN, and 0.1% formic acid in water) and filtered with 0.2 µm filters (Nanosep, PALL). The samples were loaded off-line onto a nanospray tapered capillary column/emitter (360 × 75 × 15 µm, PicoFrit, New Objective, 15 cm column) that was self-packed with C18 reverse phase (RP) resin (Waters) in a nitrogen pressure bomb for 10 minutes at 1000 psi (~5 µL load). The peptides were separated using the Dionex UltiMate 3000 nano-LC system (ThermoFisher) with a 180 min linear gradient of increasing mobile phase B at a flow rate of 120 nL/min. The LC-MS/MS analysis was performed using the Orbitrap Fusion Tribrid MS (ThermoFisher) equipped with a Nanospray Flex Ion Source at 2.2 kV spray voltage and 280 °C ion

transfer tube temperature. The full FTMS (Fourier transform mass spectrometry) spectrum, typically recorded at 120,000 of resolution in positive ion and profile mode, was acquired at 300-2000 m/z followed by the MS/MS spectra of ITMS (ion trap mass spectrometry) on the 15 most intense ions from the targeted mass lists or data dependent MS/MS spectra on the most intense ion with dynamic exclusion at 30s duration time. The targeted ions were isolated by the quadruple at 1.5 m/z isolation window for CID (Collision-induced Dissociation) and 3.0 m/z for ETD (Electron-transfer Dissociation) and fragmented by decision-tree algorithm by alternating between CID at 38% normalized collision energy and ETD at 80 ms of reaction time for above triply charged and 150 ms of reaction time with 40% of supplemental activation for doubly charged ions.

Detection of O-linked glycosylation

The raw files were searched against the Oct4 database including contaminant database (along with reversed proteins as decoys) using Proteomic Discoverer (Thermo Scientific) with a peptide tolerance of 30 ppm; a MS/MS tolerance of 0.8 Da; the carbamidomethylated cysteine; oxidation of methionine and phosphorylation and O-linked glycosylation (HexNAc) of serine and threonine as variable modifications. The peptide sequences were identified by Proteomic Discoverer from the CID and ETD spectra and verified manually. The glycosylations and phosphorylations on the peptides were verified by the presence of corresponding neutral loss fragment ions of sugar and phosphate such as the HexNAc at 203.08 Da and phosphate at 79.97 and 97.98 Da calculating charge states in CID spectra. In total, 34 O-GlcNAc or

phosphorylation sites were observed via multiple LC-MS/MS runs in each experiment. The best scored glyco- and phospho-peptides based on XCorr value that were manually validated for neutral loss peaks are listed in Supplementary Table 3.1. Representative MS and MS/MS spectra are shown in Figure 3.4.

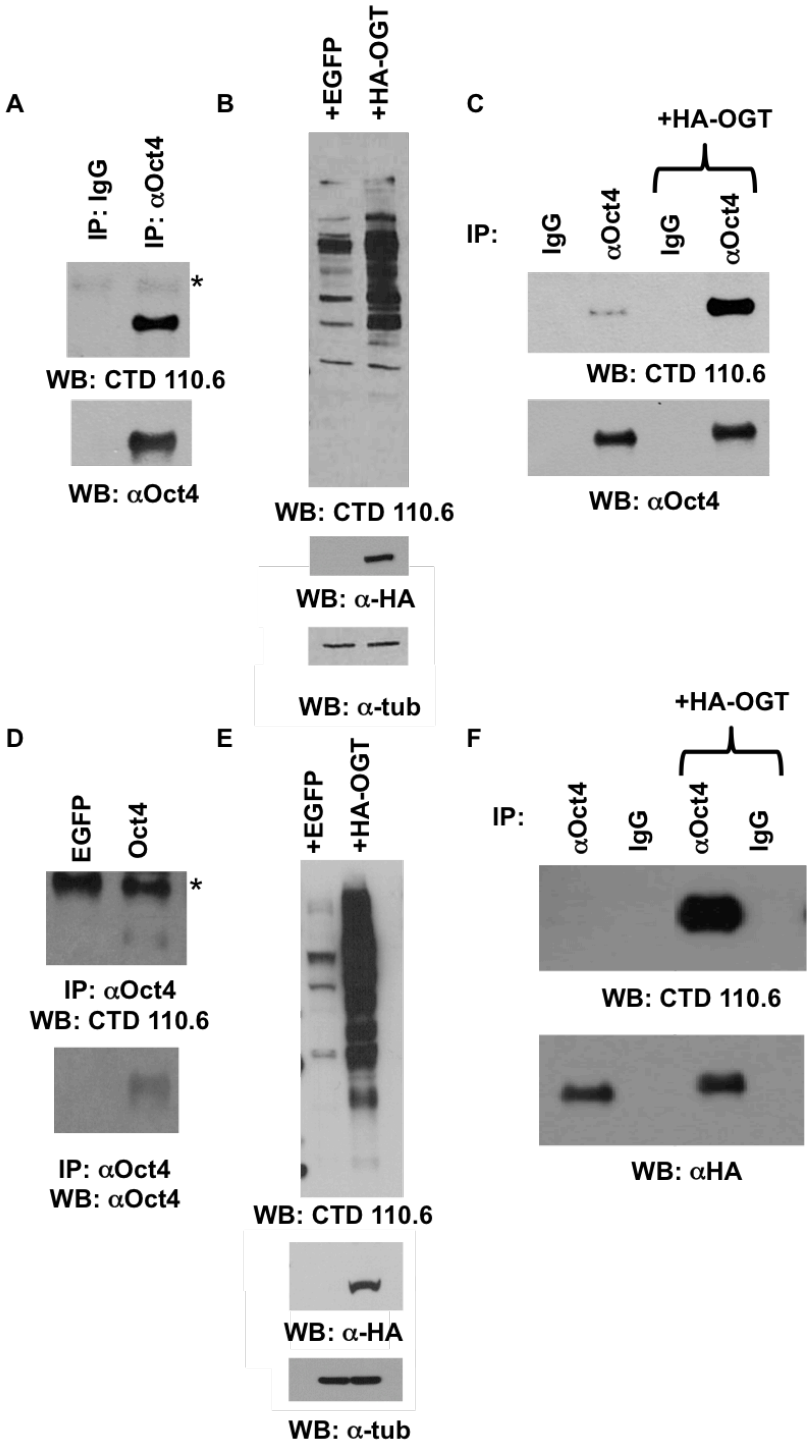
RNA Purification, cDNA synthesis and qPCR

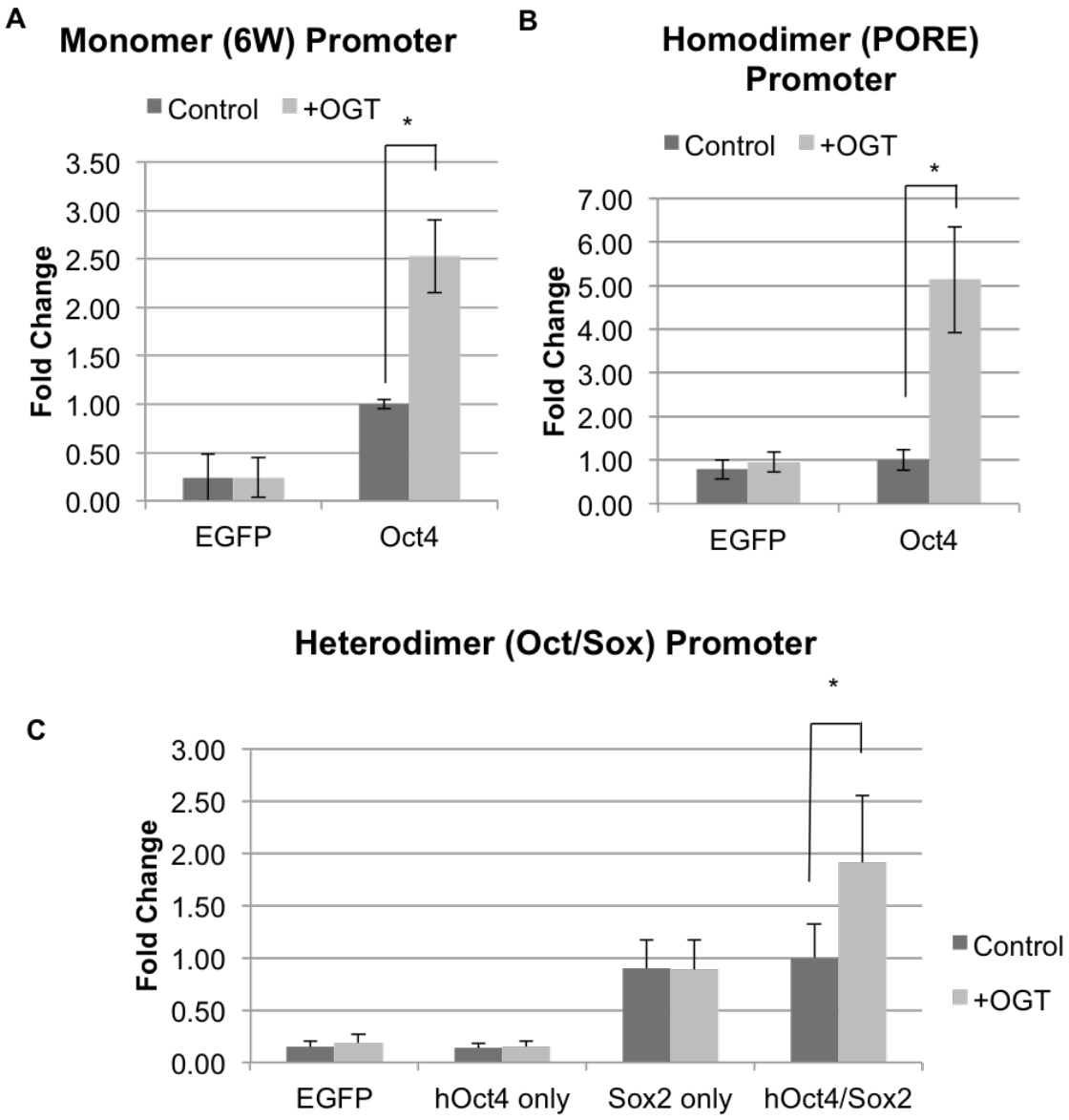
RNA was isolated from cells using QIAGEN RNeasy Mini Kit. cDNA was synthesized using Bio-Rad iScript™ cDNA Synthesis Kit. qPCR was carried out on a BioRad iCycler using BioRad SYBR green. Ct values were compared using the DDCT method.

Acknowledgments

A portion of this research was supported by NIH/NIGMS (National Center for Biomedical Glycomics, P41GM103490, LW Senior Investigator). This research was also supported in part by Basic Science Research Program and the Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (grant number: NRF 2011-0013961 and NRF 2010-0029634, JML).

Figure 3.1 - The amount of O-GlcNAc found on Oct4 increases with OGT overexpression.
Western blot analysis to determine the presence of O-GlcNAc modification. O-GlcNAc specific antibody CTD110.6 on immunoprecipitated endogenous Oct4 protein from H9 hES cells (A) or expressed Oct4 protein in HEK293T cells (D). Crude lysates of H9 (B) or HEK293T (E) cells expressing either EGFP (control) or OGT. Immunoprecipitation of Oct4 from H9 cells (C) or HEK293T cells (F) expressing EGFP (control) or OGT using Oct4 antibody or IgG control to show specificity. * heavy chain


A
IP: IgG
IP: αOct4
*
WB: CTD 110.6
WB: αOct4
B
+EGFP
+HA-OGT
WB: CTD 110.6
WB: α-HA
WB: α-tub
C
+HA-OGT
IP:
IgG
αOct4
IgG
αOct4
WB: CTD 110.6
WB: αOct4
D
EGFP
Oct4
*
IP: αOct4
WB: CTD 110.6
IP: αOct4
WB: αOct4
E
+EGFP
+HA-OGT
WB: CTD 110.6
WB: α-HA
WB: α-tub
F
+HA-OGT
IP:
αOct4
IgG
αOct4
IgG
WB: CTD 110.6
WB: αHA



<u>Figure 3.2</u> - Over expression of OGT alters transcriptional activity of Oct4. The transcriptional activity of Oct4 determined by luciferase reporter constructs. WT human Oct4 was expressed in HEK293T cells with either EGFP (control) or human OGT construct. Promoters bind Oct4 in either monomer (A) or two different dimer (B) & (C) configurations. Experiments using the Oct/Sox promoter included expression of Sox2 either alone or in combination with Oct4. *$p<0.05$

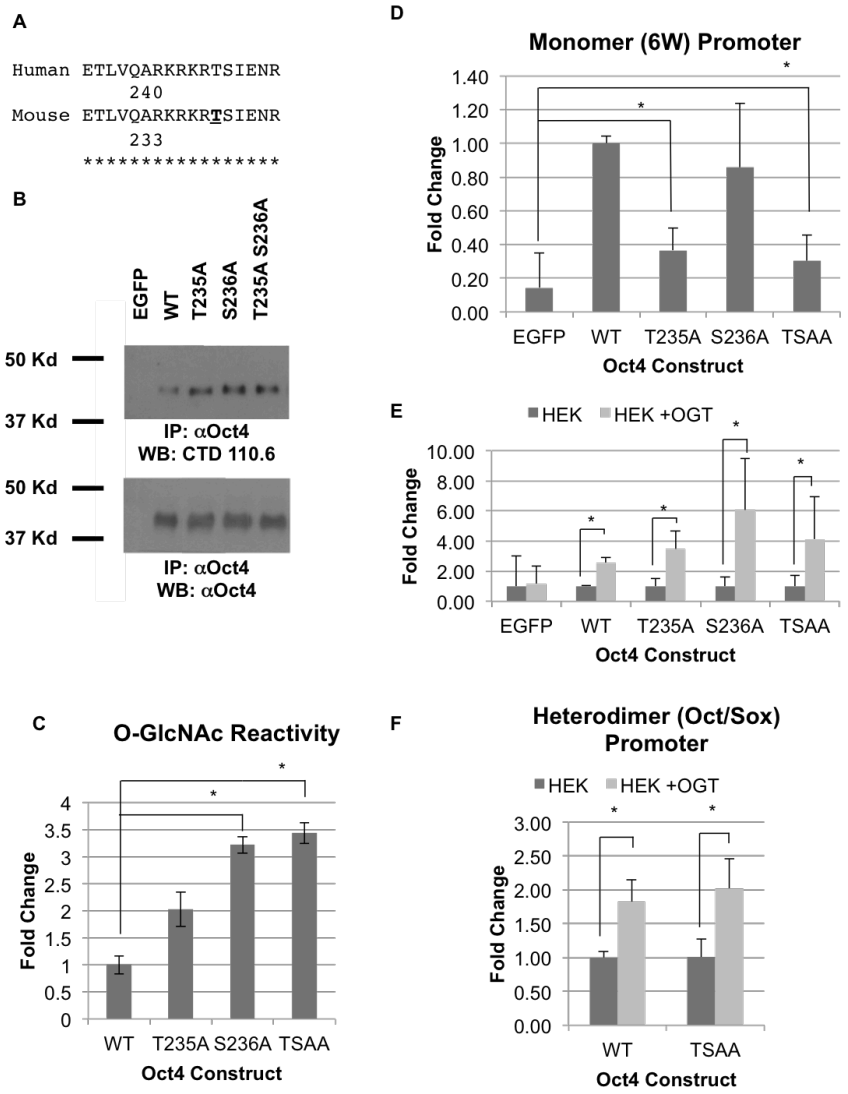


Figure 3.3 - hOct4 is Modified Beyond Known mT228 Site.
Sequence analysis of human and mouse sequence around mapped mouse O-GlcNAc site (A). Western blot analysis to determine the amount of O-GlcNAc on immunoprecipitated mutated Oct4 constructs (B), quantified in (C). Luciferase expression of mutant constructs in HEK293T cells using monomer promoter showing fold change over WT (D). Luciferase expression of mutant constructs using monomer promoter in the presence of EGFP (control) or OGT showing fold change over control for each mutant (E). Luciferase expression of WT and double mutant construct using heterodimer promoter in the presence of EGFP (control) or OGT showing fold change over control for each mutant (F). *p<0.05

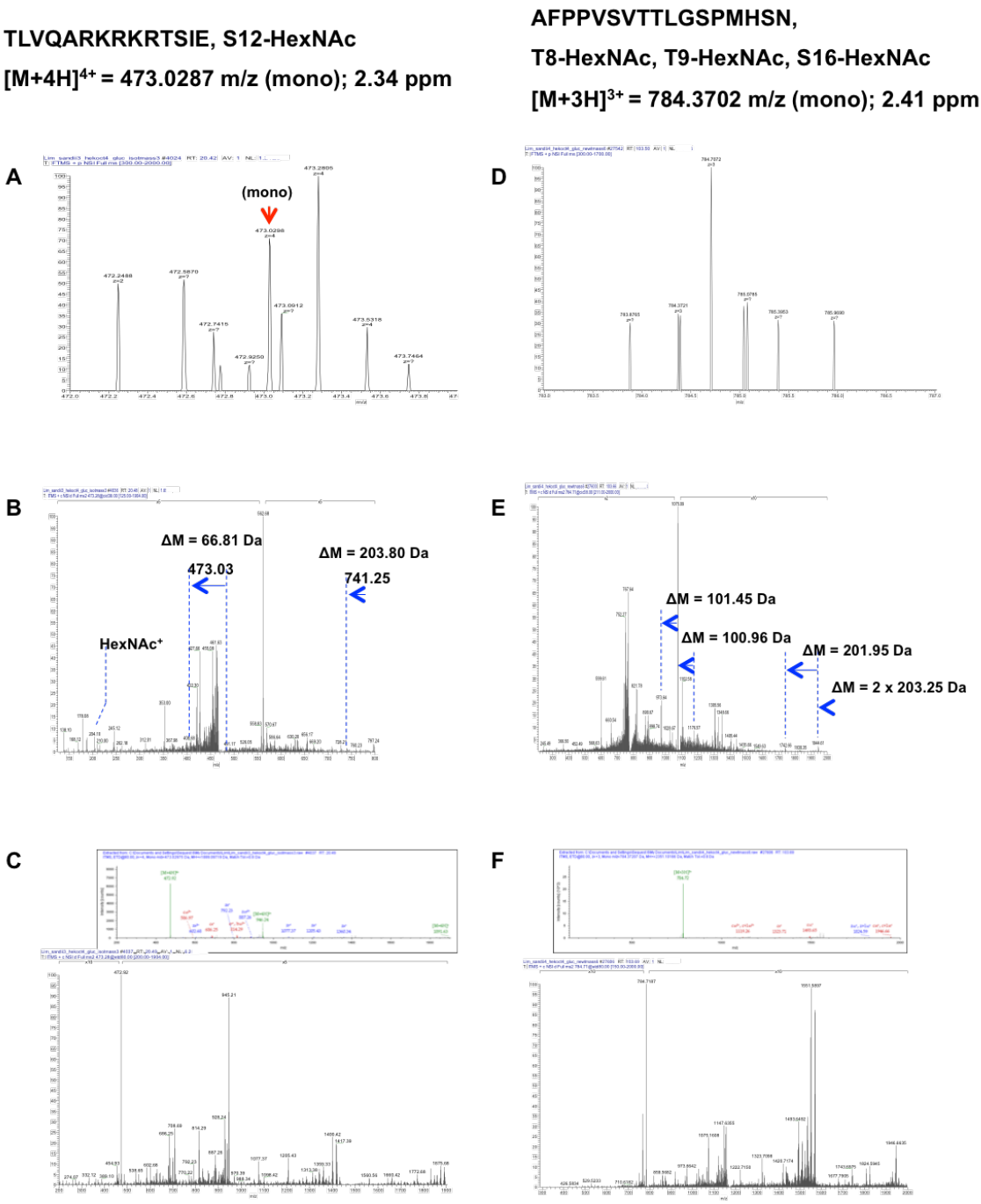


Figure 3.4 - hOct4 is Modified With Multiple O-GlcNAc Residues.
MS spectra for two representative peptides: TLVQARKRKRTSIE (A, B, C) and AFPPVSVTTLGSPMHSN (D, E, F). Full MS showing parent mass (A) & (D), CID spectra showing neutral loss of one or more HexNAc ions (B) & (E) and ETD spectra to confirm the sequence and site assignment (C) & (F).

G

1 MAGHLASDFAFSPPPGGGGDGPGGPEPGWVDPRTWLSFQGPPGGPGIGPGVGPGSEVWGI
61 PPCPPPYEFCGGMAYCGPQVGVGLVPQGGLETSQPEGEAGVGVESNSDGASPEPCTgVTPG
121 AVKLEKEKLEQNPEESQDIKALQKELEQFAKLLKQKRITLGYTQADVGLTLGVLFGKVFS
181 QTTICRFEALQLSFKNMCKLRPLLQKWVEEADNNENLQEICKAETgpLVQARKRKRTpSgIENR
241 VRGNLENLFLQCPKPTLQQISHIAQQLGLEKDVVRVWFCNRRQKGKRSSSDYAQREDFEA
301 AGSPFSGGPVSFPLAPGPHFGTPGYGSPHFTALYSgSVPFPEGEAFPPVSgVTgpTgLGSgPMHSgN

H

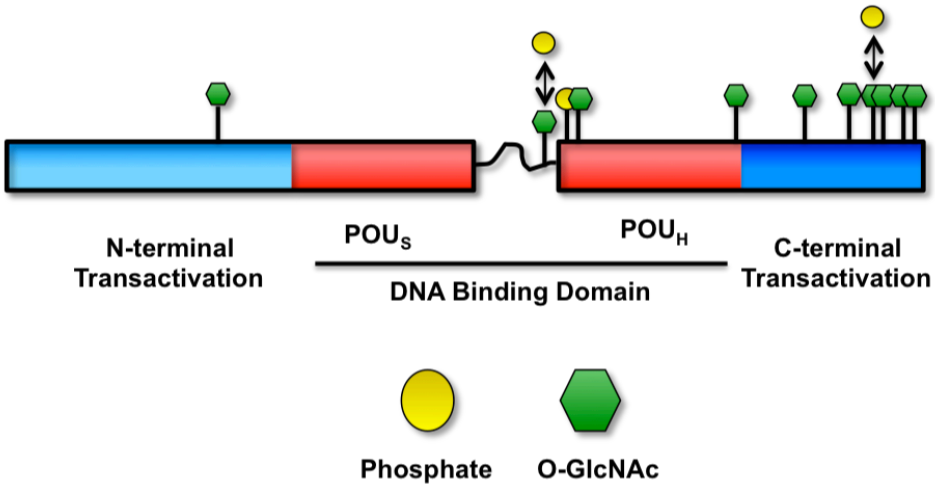


Figure 3.4 (cont.) - hOct4 is Modified With Multiple O-GlcNAc Residues. Full human Oct4 sequence showing mapped O-GlcNAc sites (denoted in red with a lowercase g) and phosphorylation sites (denoted in blue with a lower case p). Sites that were mapped with both O-GlcNAc and phosphate are shown in green (G). Graphical representation of G (H).

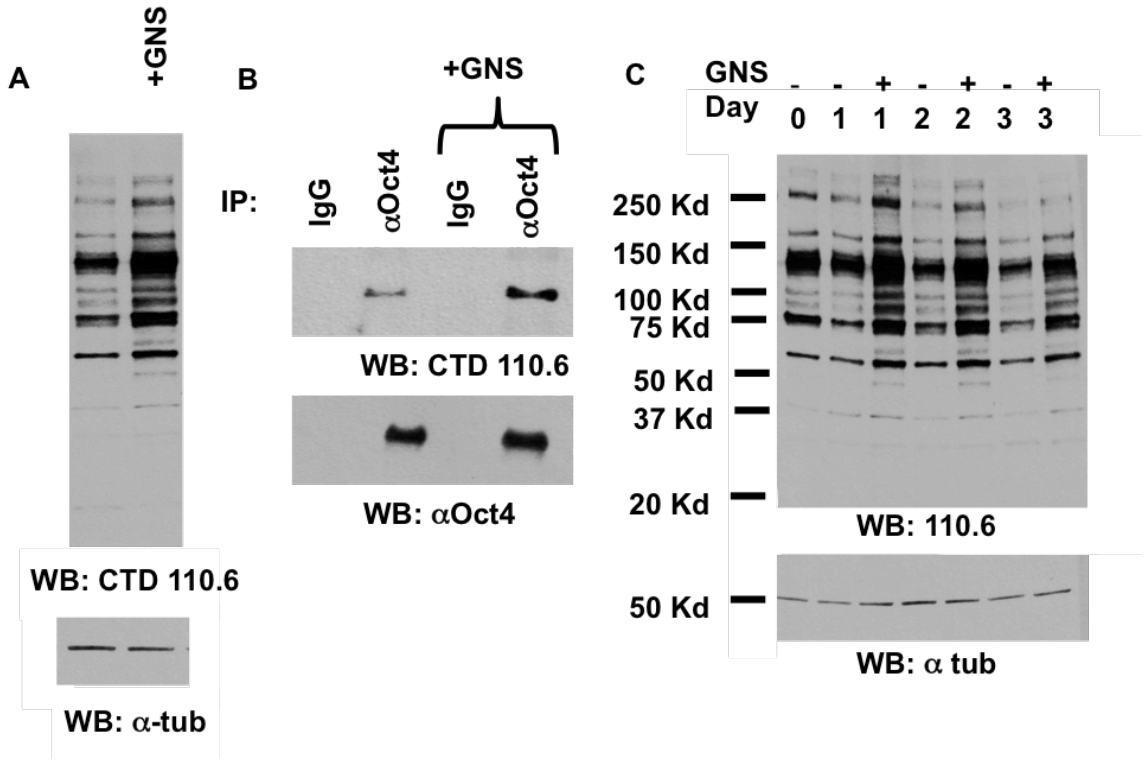


Figure 3.5 - GlcNAcstatin treatment also increases O-GlcNAc levels in H9 hES Cells but shows no effect on pluripotency or early differentiation. Western blot analysis to determine the presence of O-GlcNAc modification. O-GlcNAc specific antibody CTD110.6 on crude lysates of H9 cells treated with 100nM GlcNAcstatin (GNS) (A), or Immunoprecipitated endogenous Oct4 protein from H9 hES cells (B). CTD 110.6 reactivity of crude lysates from H9 cells over three days of GNS treatment (C).

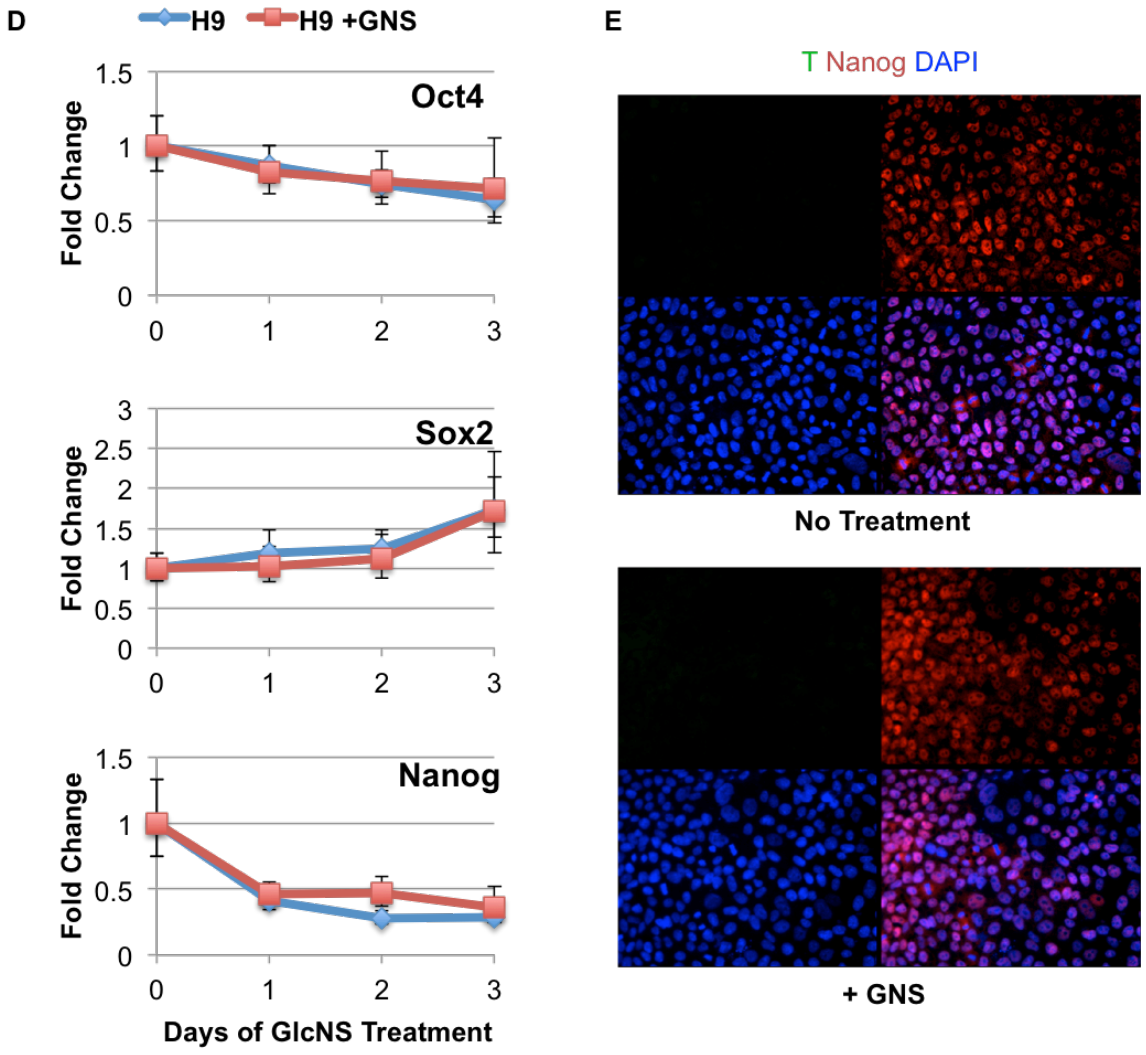


Figure 3.5 (cont.) - GlcNAcstatin treatment also increases O-GlcNAc levels in H9 hES Cells but shows no effect on pluripotency or early differentiation. RNA transcript levels of key pluripotency markers of 100nM GNS treated or untreated H9 cells using qPCR (D). Immunofluorescence of untreated H9 cells (upper panel) or 100nM GNS treated cells (lower panel). Brachyury (T) in green and nanog in red. Nuclei are stained with DAPI in blue. Merged image is shown in bottom right corner of each panel (E).

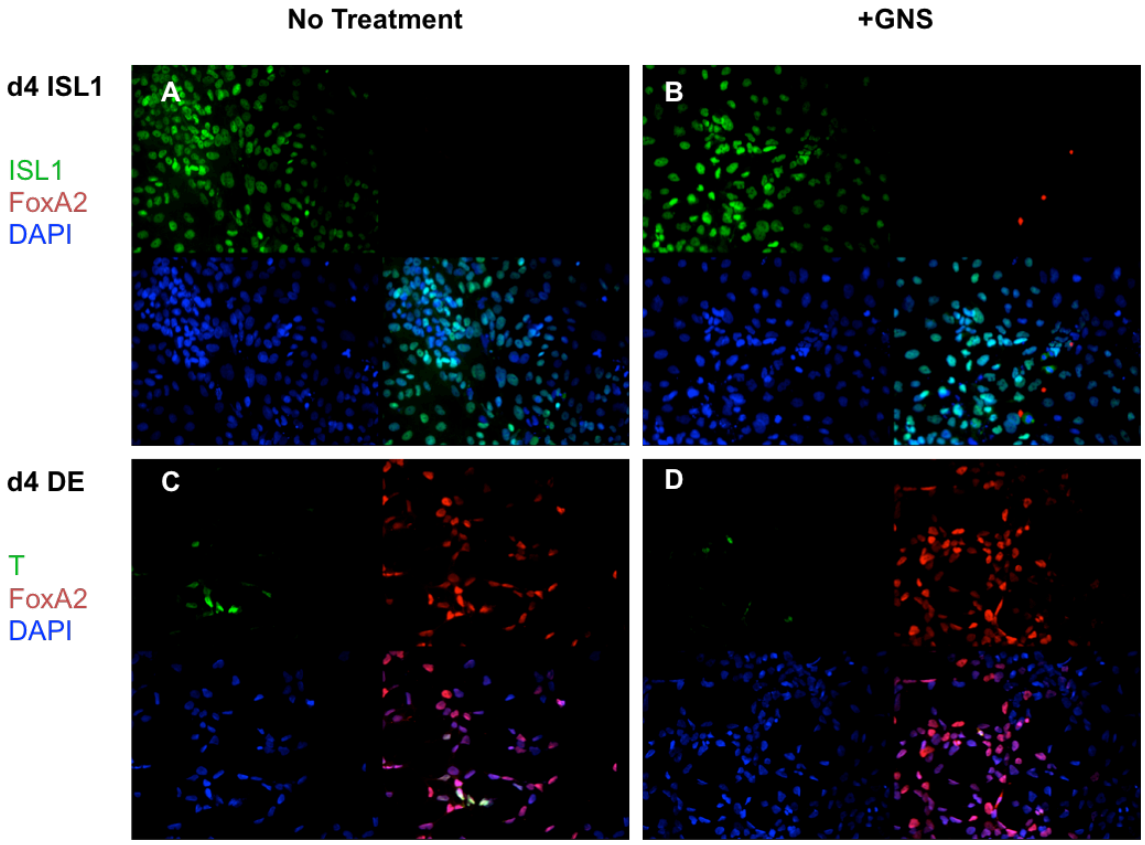


Figure 3.6 - Early differentiation markers of ES cells are unchanged in the presence of GNS.
Immunofluorescence of untreated (A) and GNS treated (B) day 4 differentiated Isl1 cells and untreated (C) and GNS treated (D) day 4 differentiated definitive endoderm (DE) cells. Isl1 and brachyury (T) are in green, FoxA2 in red and DAPI in blue.

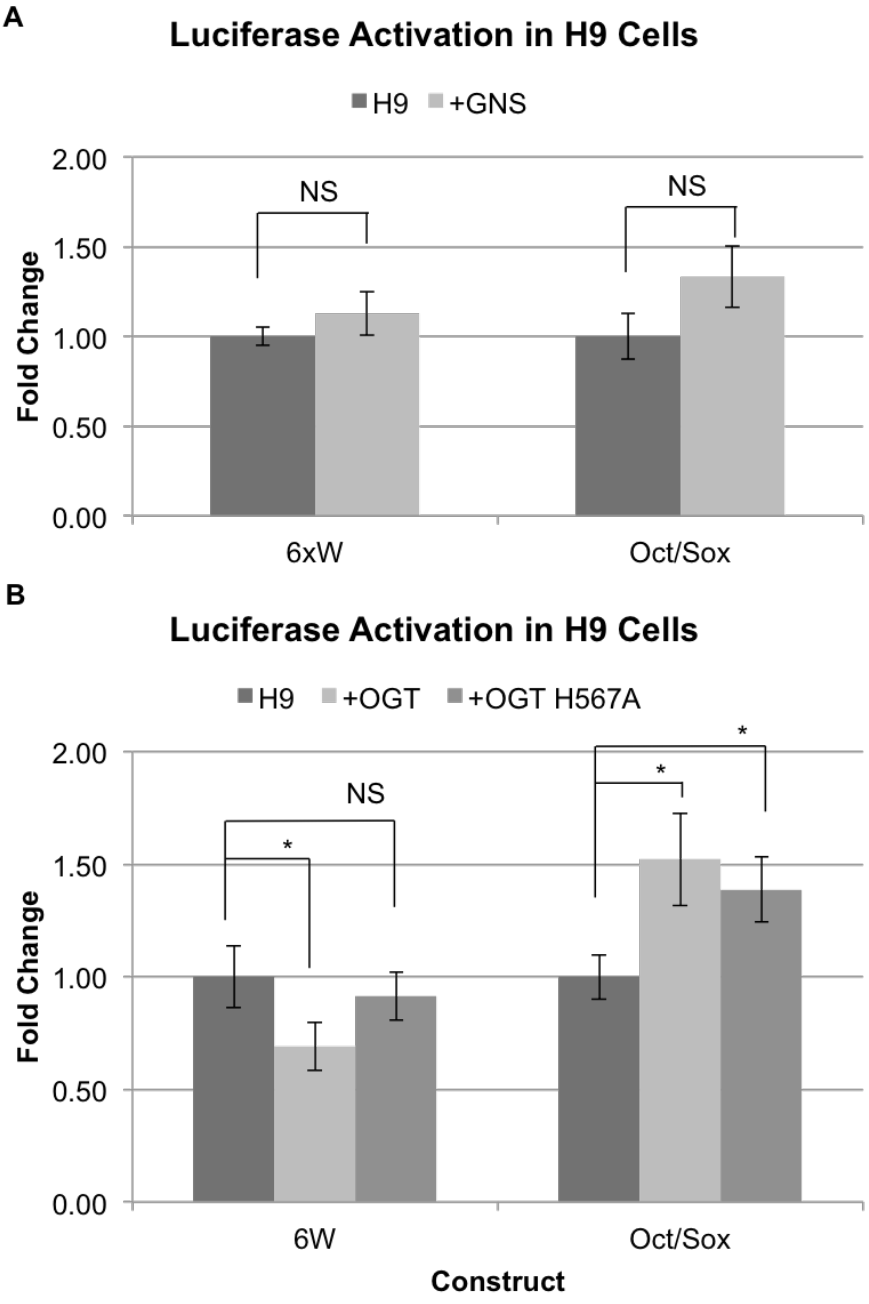


Figure 3.7 - Association with OGT alters transcriptional activation of Oct4 specific luciferase constructs in H9 hES cells.
Luciferase activity in H9 cells treated with or without 100nM GNS (A), or expressing EGFP (control), OGT or catalytically dead OGT (OGT H567A) (B). *p<0.05

Table 3.1: O-GlcNAc and Phosphorylation Sites Mapped on Oct4

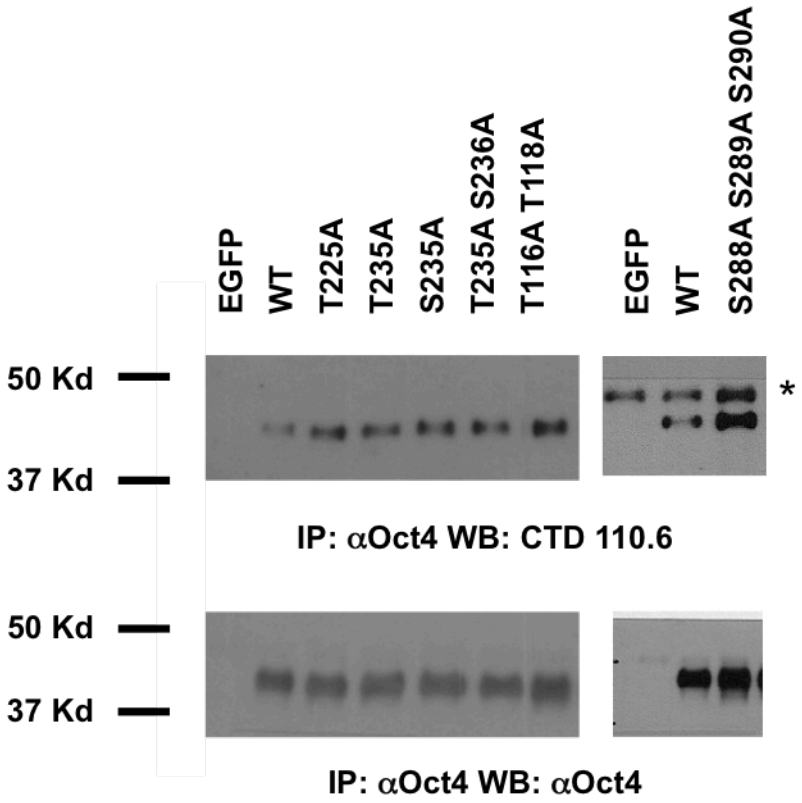
Sequence	Modifications
GASPEPC**T**VTPGAVKLE	T116-HexNAc
TLVQARKRKRT**S**IE	T225-HexNAc, S236-HexNAc T225-Phospho, T235-Phospho,
KDVVRVWFCNRRQKGKR**SSS**D	S288 or S289 or S290
AAGSPFSGGPVSFPLAPGPHFGTPGY GSPHFTALY**S**SVPFPEGE	S335-HexNAc
GEAFPPV**S**V**TT**LG**S**PMH**S**N	S349-HexNAc, S349-HexNAc/S355-HexNAc, S349-HexNAc/T351-Phospho, T351-HexNAc/S359-HexNAc, T351-HexNAc/T352-HexNAc/S359-HexNAc

Commas denote separate modifications on the same peptide, forward slash denotes modifications found on the same peptide

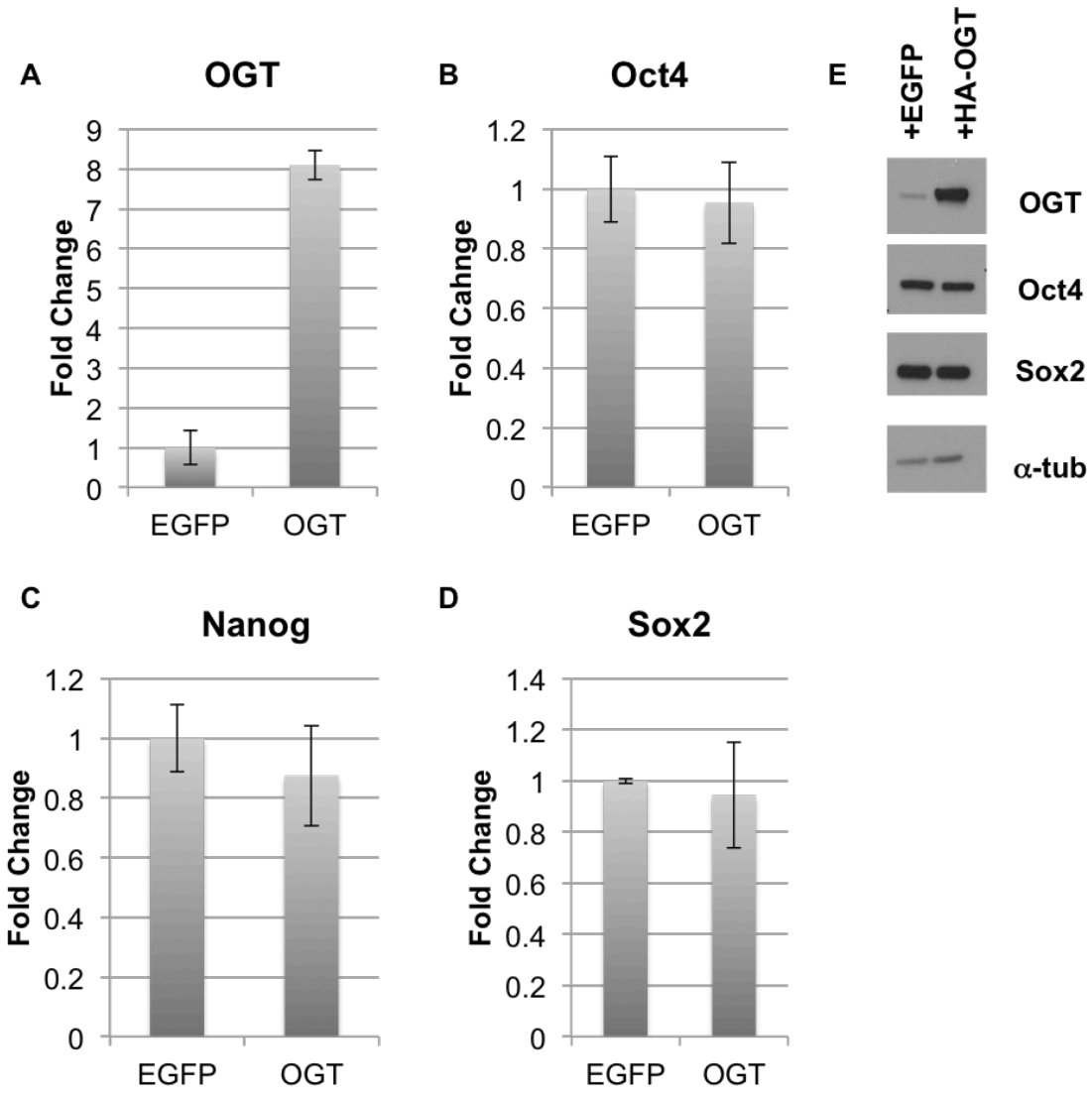
Table 3.2: Oct4 Mutant Properties

hOct4 Construct	110.6 Reactive	6W Promoter	Oct/Sox Promoter
WT	Yes	Increase with OGT	Increase with OGT
T116A T118A (TT)	Yes	N/D	Increase with OGT
T225A	Yes	N/D	Increase with OGT
T235A S236A (TSAA)	Yes	Increase with OGT	Increase with OGT
S288A S289A S290A (SSS)	Yes	Increase with OGT	Increase with OGT

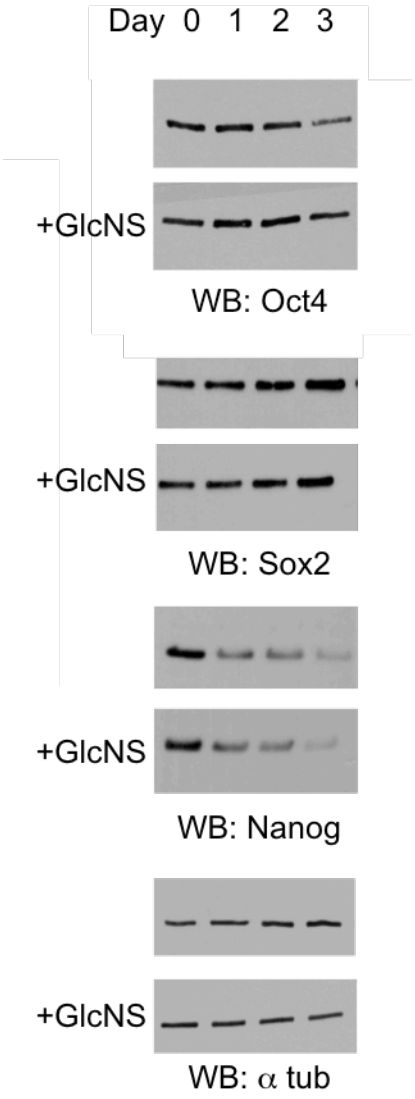
N/D – Not determined.



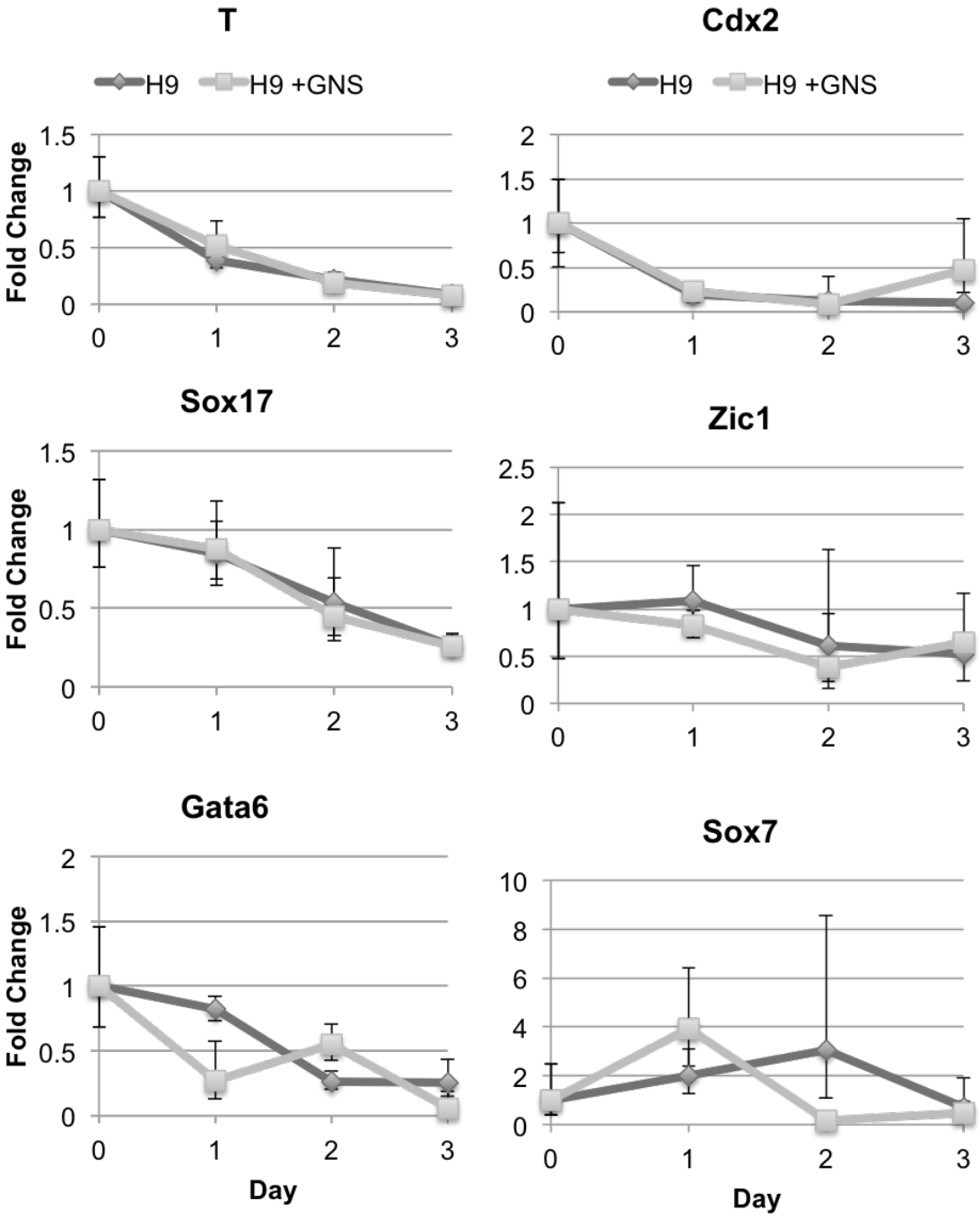
Supplemental Figure 3.1 - hOct4 mutants are still 110.6 reactive. Western blot analysis to determine the amount of O-GlcNAc on immunoprecipitated mutated Oct4 constructs. *Heavy Chain



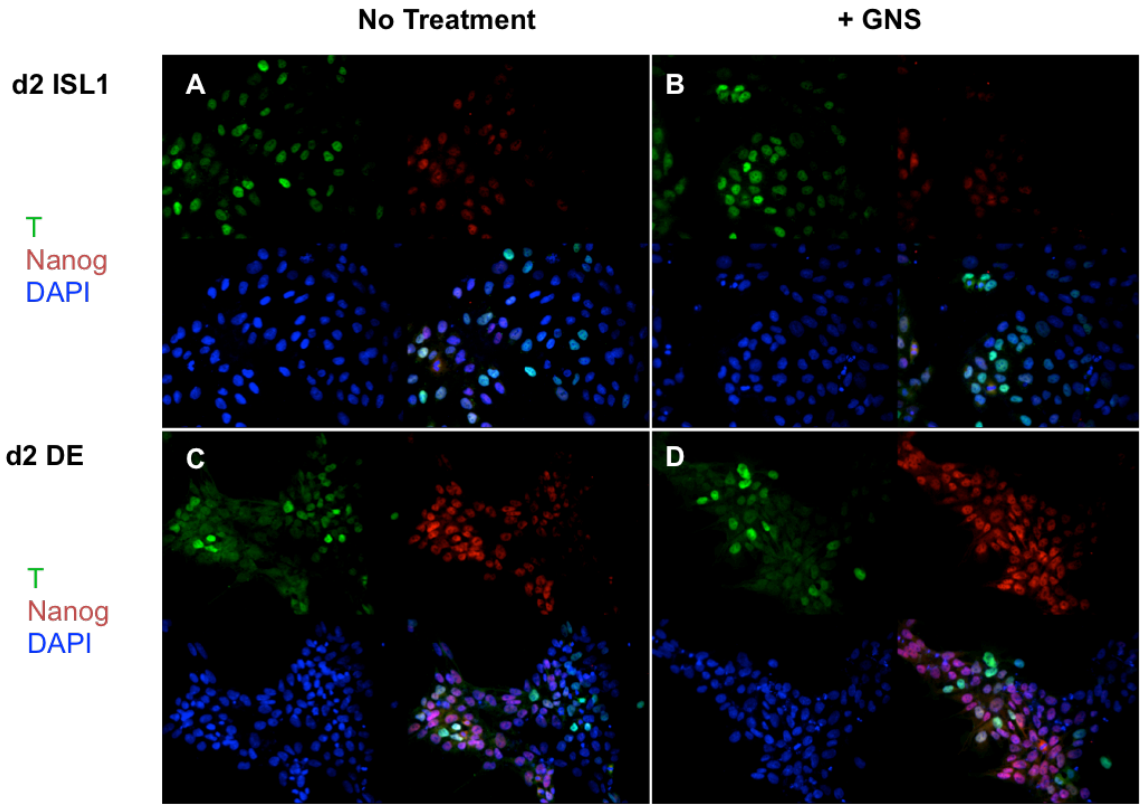
<u>Supplemental Figure 3.2</u> - OGT overexpression in H9 hES cells does not alter pluripotency gene expression.
RNA transcript levels of OGT (A) and key pluripotency markers (B, C & D) and protein levels (E) in H9 cells expressing EGFP (control) or OGT.



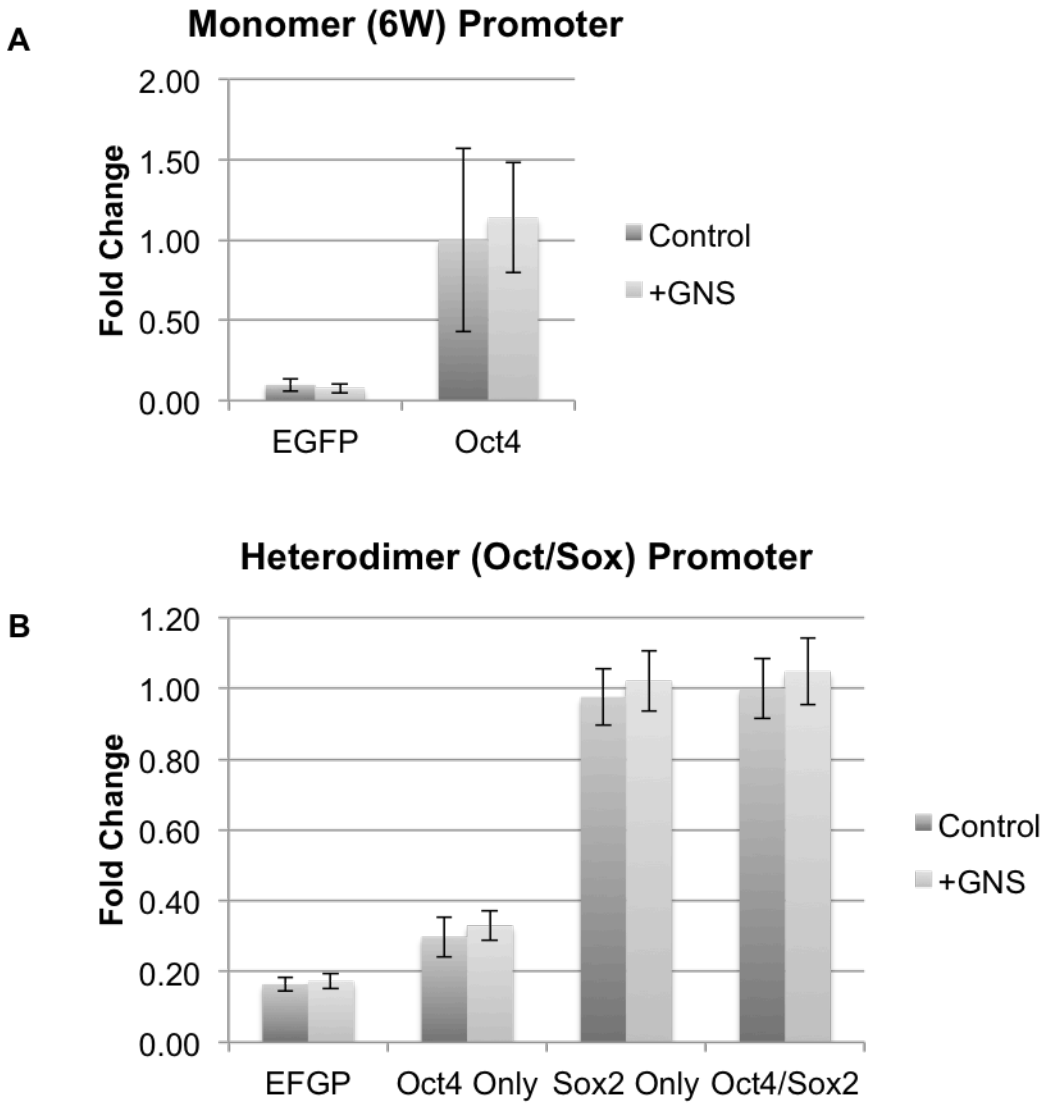
Supplemental Figure 3.3 - GlcNAcstatin treatment of H9 cells does not alter pluripotency gene expression.
Western blot analysis of protein levels of key pluripotency genes in 100nM GNS treated H9 cells.



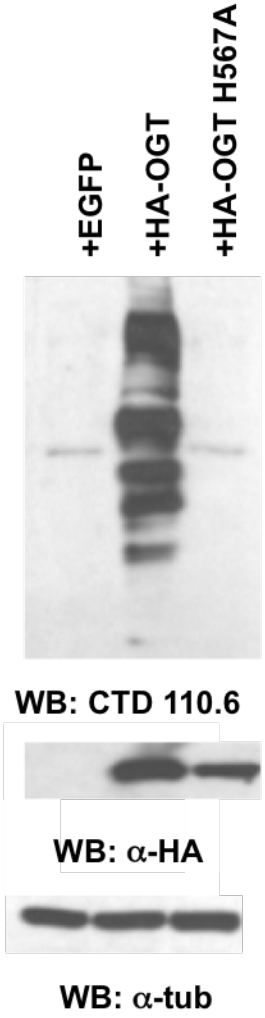
<u>Supplemental Figure 3.4</u> - GlcNAcstatin treatment of H9 ES cells does not alter early differentiation gene expression.
RNA transcript levels of early differentiation markers from untreated or 100nM GNS treated H9 cells grown in ES cell media using qPCR.



Supplemental Figure 3.5 - GlcNAcstatin treatment of H9 ES cells does not alter early differentiation events.
Immunofluorescence of untreated (A) and GNS treated (B) day 2 differentiated Isl1 cells and untreated (C) and GNS treated (D) day 2 differentiated definitive endoderm (DE) cells. Brachyury (T) is in green, Nanog in red and DAPI in blue. Merged image is shown in bottom right corner of each panel.



<u>Supplemental Figure 3.6 - GlcNAcstatin treatment does not alter luciferase promoter activity in HEK293T cells.</u>
The transcriptional activity of Oct4 determined by luciferase reporter constructs. WT human Oct4 was expressed in HEK293T cells with either EGFP (control) or human OGT construct. Promoters bind Oct4 in either monomer (A) or dimer (B) configurations. Experiments using the Oct/Sox promoter included expression of Sox2 either alone or in combination with Oct4



Supplemental Figure 3.7 - OGT H567A does not increase O-GlcNAc levels. Western blot analysis of HEK293T cells expressing EGFP (control), WT OGT or OGT H567A.

Supplementary Table 3.1: Manually validated glycol- and phosphor-peptides based on XCorr value

No.	Sequence	Modifications	Start	End	Charge	MH^+ (mono)	m/z (mono, theoretical)	m/z (mono, experimental)	ΔM (ppm)	XCorr	Activation Type
1	GASPEPCTVTPGAVKLE	1 x HexNAc, C7-Carbamidomethyl	109	125	2	1915.9317	958.4695	958.4640	-5.81	2.50	CID
2	GASPEPCTVTPGAVKLE	T8-HexNAc	109	125	3	1858.9103	620.3083	620.3030	-8.50	1.17	ETD
3	TLVQARKRKRTSIE	S12-HexNAc	225	238	4	1889.0927	473.0287	473.0298	2.34	3.05	ETD
4	TLVQARKRKRTSIE	T11-Phospho	225	238	4	1765.9796	442.2504	442.2517	2.99	3.52	ETD
5	TLVQARKRKRTSIE	T1-HexNAc	225	238	4	1889.0927	473.0287	473.0295	1.70	4.06	ETD
6	TLVQARKRKRTSIE	T1-Phospho	225	238	3	1765.9796	589.3314	589.3388	5.49	2.81	ETD
7	KDVVRVWFCNRRQKGKRSSSD	1 x HexNAc	271	291	3	2754.4216	918.8121	918.8164	4.62	1.37	CID
8	AAGSPFSGGPVSFPLAPGPHFGTPGY GSPHFTALYSSVPFPE	1 x HexNAc	299	241	3	4392.0924	1464.7023	1464.7025	0.09	4.07	CID
9	AAGSPFSGGPVSFPLAPGPHFGTPGY GSPHFTALYSSVPFPEGE	1 x HexNAc	299	243	3	4578.1565	1526.7237	1526.7318	5.29	3.85	CID
10	AAGSPFSGGPVSFPLAPGPHFGTPGY GSPHFTALYSSVPFPE	S36-HexNAc	299	241	4	4392.0924	1098.7786	1098.7830	3.96	3.92	ETD
11	AAGSPFSGGPVSFPLAPGPHFGTPGY GSPHFTALYSSVPFPEGE	S36-HexNAc	299	243	4	4578.1565	1145.2946	1145.2976	2.61	3.05	ETD
12	GEAFPPVSVTTLGSPMHSN	1 x HexNAc	342	360	2	2131.0012	1066.0043	1066.0054	1.01	2.74	CID
13	GEAFPPVSVTTLGSPMHSN	1 x HexNAc, 1 x Phospho	342	360	2	2210.9675	1105.9874	1105.9899	2.18	3.18	CID
14	GEAFPPVSVTTLGSPMHSN	1 x HexNAc, M16-Oxidation	342	360	2	2146.9961	1074.0017	1074.0017	-0.04	0.97	CID
15	GEAFPPVSVTTLGSPMHSN	3 x HexNAc	342	360	2	2537.1600	1269.0837	1269.0824	-1.00	1.36	CID
16	GEAFPPVSVTTLGSPMHSN	M16-Oxidation, 1 x Phospho	342	360	2	2023.8831	1012.4452	1012.4612	15.77	2.77	CID
17	GEAFPPVSVTTLGSPMHSN	S8-HexNAc (potential)	342	360	2	2131.0012	1066.0043	1066.0006	-3.45	3.51	ETD
18	GEAFPPVSVTTLGSPMHSN	S8-HexNAc, M16-Oxidation	342	360	3	2146.9961	716.3369	716.3511	19.77	2.86	ETD
19	AFPPVSVTTLGSPMHSN	1 x HexNAc	344	360	2	1944.9372	972.9723	972.9769	4.81	2.68	CID
20	AFPPVSVTTLGSPMHSN	1 x HexNAc, 1 x Phospho	344	360	2	2024.9035	1012.9554	1012.9542	-1.18	2.48	CID
21	AFPPVSVTTLGSPMHSN	1 x HexNAc, M14-Oxidation	344	360	2	1960.9321	980.9697	980.9681	-1.60	2.65	CID
22	AFPPVSVTTLGSPMHSN	2 x HexNAc, M14-Oxidation	344	360	2	2164.0114	1082.5094	1082.5098	0.34	1.16	CID
23	AFPPVSVTTLGSPMHSN	3 x HexNAc	344	360	2	2351.0959	1176.0516	1176.0558	3.54	1.05	CID
24	AFPPVSVTTLGSPMHSN	3 x HexNAc, M14-Oxidation	344	360	2	2367.0908	1184.0491	1184.0463	-2.38	1.48	CID
25	AFPPVSVTTLGSPMHSN	S6-HexNAc	344	360	3	1944.9372	648.9839	648.9829	-1.67	4.32	ETD

Supplementary Table 3.1 (Cont.): Manually validated glycol- and phosphor-peptides based on XCorr value

26	AFPPVSVTTLGSPMHSN	S6-HexNAc (potential)	344	360	2	1944.9372	972.9723	972.9728	0.60	4.24	ETD
27	AFPPVSVTTLGSPMHSN	S6-HexNAc, M14-Oxidation	344	360	3	1960.9321	654.3156	654.3328	26.31	4.16	ETD
28	AFPPVSVTTLGSPMHSN	S6-HexNAc, S12-HexNAc	344	360	3	2148.0165	716.6770	716.6782	1.63	3.06	ETD
29	AFPPVSVTTLGSPMHSN	S6-HexNAc, S12-Phospho (potential)	344	360	2	2024.9035	1012.9554	1012.9546	-0.82	2.97	ETD
30	AFPPVSVTTLGSPMHSN	S6-HexNAc, T8-Phospho	344	360	3	2024.9035	675.6394	675.6387	-0.95	2.64	ETD
31	AFPPVSVTTLGSPMHSN	T8-HexNAc, S16-HexNAc	344	360	3	2148.0165	716.6770	716.6752	-2.63	3.02	ETD
32	AFPPVSVTTLGSPMHSN	T8-HexNAc, T9-HexNAc (potential)	344	360	2	2148.0165	1074.5119	1074.5199	7.41	2.52	ETD
33	AFPPVSVTTLGSPMHSN	T8-HexNAc, T9-HexNAc, M14-Oxidation, S16-HexNAc	344	360	3	2367.0908	789.7018	789.6969	-6.23	3.06	ETD
34	AFPPVSVTTLGSPMHSN	T8-HexNAc, T9-HexNAc, S16-HexNAc	344	360	3	2351.0959	784.3702	784.3721	2.41	3.47	ETD

CHAPTER 4

DISCUSSION

O-GlcNAc modification plays a large role in regulating the properties of transcription factors [63], and enhances the transcriptional activity of mouse Oct4 [94] as well as being involved in epigenetic control of gene expression [285]. Although human Oct4 is modified by O-GlcNAc [77], no sites of attachment or functions had been assigned. This study was designed to map the sites of O-GlcNAc attachment and determine the role O-GlcNAc plays in regulating Oct4 transcriptional activation. Using a combination of CID and ETD, we mapped 10 novel sites of O-GlcNAc attachment, and one novel site of phosphorylation. The previous study used truncation constructs expressed in HEK293T cells and found the modified region of Oct4 to be exclusively in the POU domain [94]. However, our study mapped residues on both the POU domain and the C-terminal transactivation domain (Figure 3.4). In fact, over half of the modifications mapped to the C-terminal domain. The discrepancy between these results could be due to the way the samples were prepared and assayed, or it is due to the difference in species. More in-depth analysis of mouse Oct4 would reveal if these sites were conserved.

The interplay between O-GlcNAc and phosphorylation has been well documented [162, 286, 287]. 8 of the 10 O-GlcNAc sites mapped in this study are also phosphorylated by kinases important for stem cell pluripotency [135,

146]. Furthermore, treatment of cells with the GSKβ inhibitor lithium, a common ingredient in stem cell media [8], causes changes to O-GlcNAc levels on multiple proteins [258]. Although GSKβ has not yet been shown to modify Oct4, it is possible that the interplay between these two modifications is playing a role in regulation of this protein, and in turn pluripotency. Indeed, increased phosphorylation of Oct4 has been correlated with a decrease in transcriptional activation [121] hence O-GlcNAc modification will prevent this phosphorylation and promote transcriptional activation. This theory may also explain the difference we saw between H9 and HEK cells with the monomer promoter as these cells are not being governed by the same signaling pathways. Future research should include the investigation of the interplay between signaling pathways and nutrient intake [46, 160].

We also showed that transcriptional activity of human Oct4 is increased in the presence of OGT, concurrent with the current research [94]. However, the previously mapped site did not solely regulate OGT transcriptional activation of human Oct4. This suggests there is either another site involved or another mechanism. For one of the promoters used, we also saw an increase of activation in the absence of OGT catalytic activity. One theory is that OGT and Oct4 are affecting transcription in complex with the chromatin remodeling enzyme complex Sin3A/HDAC. The Sin3A/HDAC complex traditionally removes acetyl groups from histones H3 and H4 leading to transcriptional repression; however, there has been recent evidence that this complex targets non-histone complexes [288]. Furthermore, Sin3A/HDAC has been shown to enhance

transcription of the Oct/Sox element of the nanog promoter through interaction with Sox2 [283]. Two independent groups found OGT bound to Oct4 when looking for interaction partners [262, 263], and OGT interacts with Sin3A/HDAC to repress transcription [59]. This evidence suggests that OGT acts as a bridging protein between Sin3A, Oct4 and Sox2 to regulate transcription. This concept is not too far fetched since Oct4 has already been shown to require a bridging factor for full activation [119]. This is an emerging role for OGT. OGT is found as part of the polycomb repression complex [61, 62], and more recently with the Ten-eleven translocation family of proteins, involved in demethylation of chromatin and transcriptional activation [282, 289-292].

Conclusion and Future Directions

In summary, this study enhanced our knowledge of Oct4 regulation. Human Oct4 is transcriptionally regulated by OGT by a different mechanism to the mouse counterpart. We mapped 10 new sites of O-GlcNAc addition, giving us multiple targets for further studies into the role O-GlcNAc addition plays into protein functionality. We also uncovered evidence of an emerging role for OGT in transcriptional regulation they may involve formation of complexes to directly regulate transcription through chromatin remodeling. Any information gained about the regulation of Oct4 has major implications on embryonic stem cells culture, the induction of pluripotent stem cells and ultimately cell therapies.

REFERENCES

1. Niakan, K.K., et al., *Human pre-implantation embryo development.* Development, 2012. **139**(5): p. 829-41.

2. Gilbert, S.F., *Developmental Biology*. 8 ed. 2006, Sunderland, MA 01375 USA: Sinauer Associates. 817.

3. Evans, M.J. and M.H. Kaufman, *Establishment in culture of pluripotential cells from mouse embryos.* Nature, 1981. **292**(5819): p. 154-6.

4. Martin, G.R., *Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.* Proc Natl Acad Sci U S A, 1981. **78**(12): p. 7634-8.

5. Thomson, J.A., et al., *Embryonic stem cell lines derived from human blastocysts.* Science, 1998. **282**(5391): p. 1145-7.

6. Wang, L., et al., *Self-renewal of human embryonic stem cells requires insulin-like growth factor-1 receptor and ERBB2 receptor signaling.* Blood, 2007. **110**(12): p. 4111-9.

7. Ludwig, T.E., et al., *Feeder-independent culture of human embryonic stem cells.* Nat Methods, 2006. **3**(8): p. 637-46.

8. Ludwig, T.E., et al., *Derivation of human embryonic stem cells in defined conditions.* Nat Biotechnol, 2006. **24**(2): p. 185-7.

9. Brimble, S.N., et al., *Karyotypic stability, genotyping, differentiation, feeder-free maintenance, and gene expression sampling in three human embryonic stem cell lines derived prior to August 9, 2001.* Stem Cells Dev, 2004. **13**(6): p. 585-97.

10. Kelly, O.G., et al., *Cell-surface markers for the isolation of pancreatic cell types derived from human embryonic stem cells.* Nat Biotechnol, 2011. **29**(8): p. 750-6.

11. Ambasudhan, R., et al., *Potential for cell therapy in Parkinson's disease using genetically programmed human embryonic stem cell-derived neural progenitor cells.* J Comp Neurol, 2014. **522**(12): p. 2845-56.

12. Chiba, S., et al., *Noggin enhances dopamine neuron production from human embryonic stem cells and improves behavioral outcome after transplantation into Parkinsonian rats.* Stem Cells, 2008. **26**(11): p. 2810-20.

13. Faulkner, J. and H.S. Keirstead, *Human embryonic stem cell-derived oligodendrocyte progenitors for the treatment of spinal cord injury.* Transpl Immunol, 2005. **15**(2): p. 131-42.

14. Matsa, E., K. Sallam, and J.C. Wu, *Cardiac stem cell biology: glimpse of the past, present, and future.* Circ Res, 2014. **114**(1): p. 21-7.

15. Loh, Y.H., et al., *The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells.* Nat Genet, 2006. **38**(4): p. 431-40.

16. Nichols, J., et al., *Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4.* Cell, 1998. **95**(3): p. 379-91.

17. Pan, G.J., et al., *Stem cell pluripotency and transcription factor Oct4.* Cell Res, 2002. **12**(5-6): p. 321-9.

18. Fong, H., K.A. Hohenstein, and P.J. Donovan, *Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells.* Stem Cells, 2008. **26**(8): p. 1931-8.

19. Avilion, A.A., et al., *Multipotent cell lineages in early mouse development depend on SOX2 function.* Genes Dev, 2003. **17**(1): p. 126-40.

20. Chambers, I., et al., *Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells.* Cell, 2003. **113**(5): p. 643-55.

21. Chambers, I., et al., *Nanog safeguards pluripotency and mediates germline development.* Nature, 2007. **450**(7173): p. 1230-4.

22. Mitsui, K., et al., *The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells.* Cell, 2003. **113**(5): p. 631-42.

23. Yuan, H., et al., *Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3.* Genes Dev, 1995. **9**(21): p. 2635-45.

24. Ambrosetti, D.C., et al., *Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer.* J Biol Chem, 2000. **275**(30): p. 23387-97.

25. Boyer, L.A., et al., *Core transcriptional regulatory circuitry in human embryonic stem cells.* Cell, 2005. **122**(6): p. 947-56.

26. Sharov, A.A., et al., *Identification of Pou5f1, Sox2, and Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data.* BMC Genomics, 2008. **9**: p. 269.

27. Gurdon, J.B., *The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles.* J Embryol Exp Morphol, 1962. **10**: p. 622-40.

28. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.

29. Jiang, J., et al., *A core Klf circuitry regulates self-renewal of embryonic stem cells.* Nat Cell Biol, 2008. **10**(3): p. 353-60.

30. Cartwright, P., et al., *LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism.* Development, 2005. **132**(5): p. 885-96.

31. Richards, M., et al., *The transcriptome profile of human embryonic stem cells as defined by SAGE.* Stem Cells, 2004. **22**(1): p. 51-64.

32. Muller, L.U., G.Q. Daley, and D.A. Williams, *Upping the ante: recent advances in direct reprogramming.* Mol Ther, 2009. **17**(6): p. 947-53.

33. David, L. and J.M. Polo, *Phases of reprogramming.* Stem Cell Res, 2014. **12**(3): p. 754-761.

34. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors.* Cell, 2007. **131**(5): p. 861-72.

35. Yu, J., et al., *Induced pluripotent stem cell lines derived from human somatic cells.* Science, 2007. **318**(5858): p. 1917-20.

36. Torres, C.R. and G.W. Hart, *Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc*, in *J Biol Chem*. 1984. p. 3308-17.

37. Haltiwanger, R.S., M.A. Blomberg, and G.W. Hart, *Glycosylation of nuclear and cytoplasmic proteins. Purification and characterization of a*

uridine diphospho-N-acetylglucosamine:polypeptide beta-N-acetylglucosaminyltransferase. J Biol Chem, 1992. **267**(13): p. 9005-13.

38. Dong, D.L. and G.W. Hart, *Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol.* J Biol Chem, 1994. **269**(30): p. 19321-30.

39. Gao, Y., et al., *Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic beta-N-acetylglucosaminidase from human brain.* J Biol Chem, 2001. **276**(13): p. 9838-45.

40. Shafi, R., et al., *The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny.* Proc Natl Acad Sci U S A, 2000. **97**(11): p. 5735-9.

41. O'Donnell, N., et al., *Ogt-dependent X-chromosome-linked protein glycosylation is a requisite modification in somatic cell function and embryo viability.* Mol Cell Biol, 2004. **24**(4): p. 1680-90.

42. Ingham, P.W., *A gene that regulates the bithorax complex differentially in larval and adult cells of Drosophila.* Cell, 1984. **37**(3): p. 815-23.

43. Hartweck, L.M., C.L. Scott, and N.E. Olszewski, *Two O-linked N-acetylglucosamine transferase genes of Arabidopsis thaliana L. Heynh. have overlapping functions necessary for gamete and seed development.* Genetics, 2002. **161**(3): p. 1279-91.

44. Forsythe, M.E., et al., *Caenorhabditis elegans ortholog of a diabetes susceptibility locus: oga-1 (O-GlcNAcase) knockout impacts O-GlcNAc cycling, metabolism, and dauer.* Proc Natl Acad Sci U S A, 2006. **103**(32): p. 11952-7.

45. Hanover, J.A., et al., *A Caenorhabditis elegans model of insulin resistance: altered macronutrient storage and dauer formation in an OGT-1 knockout.* Proc Natl Acad Sci U S A, 2005. **102**(32): p. 11266-71.

46. Wells, L., K. Vosseller, and G.W. Hart, *A role for N-acetylglucosamine as a nutrient sensor and mediator of insulin resistance.* Cell Mol Life Sci, 2003. **60**(2): p. 222-8.

47. Haltiwanger, R.S., K. Grove, and G.A. Philipsberg, *Modulation of O-linked N-acetylglucosamine levels on nuclear and cytoplasmic proteins in vivo using the peptide O-GlcNAc-beta-N-acetylglucosaminidase inhibitor O-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-N-phenylcarbamate.* J Biol Chem, 1998. **273**(6): p. 3611-7.

48. Dorfmueller, H.C., et al., *GlcNAcstatin: a picomolar, selective O-GlcNAcase inhibitor that modulates intracellular O-glcNAcylation levels.* J Am Chem Soc, 2006. **128**(51): p. 16484-5.

49. Macauley, M.S., et al., *O-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors.* J Biol Chem, 2005. **280**(27): p. 25313-22.

50. Yuzwa, S.A., et al., *A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo.* Nat Chem Biol, 2008. **4**(8): p. 483-90.

51. Dehennaut, V., et al., *O-linked N-acetylglucosaminyltransferase inhibition prevents G2/M transition in Xenopus laevis oocytes.* J Biol Chem, 2007. **282**(17): p. 12527-36.

52. Dehennaut, V., et al., *Microinjection of recombinant O-GlcNAc transferase potentiates Xenopus oocytes M-phase entry.* Biochem Biophys Res Commun, 2008. **369**(2): p. 539-46.

53. Ohn, T., et al., *A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly.* Nat Cell Biol, 2008. **10**(10): p. 1224-31.

54. Zachara, N.E. and G.W. Hart, *O-GlcNAc a sensor of cellular state: the role of nucleocytoplasmic glycosylation in modulating cellular function in response to nutrition and stress.* Biochim Biophys Acta, 2004. **1673**(1-2): p. 13-28.

55. Vosseller, K., et al., *Elevated nucleocytoplasmic glycosylation by O-GlcNAc results in insulin resistance associated with defects in Akt activation in 3T3-L1 adipocytes.* Proc Natl Acad Sci U S A, 2002. **99**(8): p. 5313-8.

56. Wells, L., K. Vosseller, and G.W. Hart, *Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc.* Science, 2001. **291**(5512): p. 2376-8.

57. Gandy, J.C., A.E. Rountree, and G.N. Bijur, *Akt1 is dynamically modified with O-GlcNAc following treatments with PUGNAc and insulin-like growth factor-1.* FEBS Lett, 2006. **580**(13): p. 3051-8.

58. Yang, X., et al., *Phosphoinositide signalling links O-GlcNAc transferase to insulin resistance.* Nature, 2008. **451**(7181): p. 964-9.

59. Yang, X., F. Zhang, and J.E. Kudlow, *Recruitment of O-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein O-GlcNAcylation to transcriptional repression.* Cell, 2002. **110**(1): p. 69-80.

60. Fujiki, R., et al., *GlcNAcylation of a histone methyltransferase in retinoic-acid-induced granulopoiesis.* Nature, 2009.

61. Gambetta, M.C., K. Oktaba, and J. Muller, *Essential role of the glycosyltransferase sxc/Ogt in polycomb repression.* Science, 2009. **325**(5936): p. 93-6.

62. Sinclair, D.A., et al., *Drosophila O-GlcNAc transferase (OGT) is encoded by the Polycomb group (PcG) gene, super sex combs (sxc).* Proc Natl Acad Sci U S A, 2009. **106**(32): p. 13427-32.

63. Brimble, S.N., et al., *The role of the O-GlcNAc modification in regulating eukaryotic gene expression.* Curr Signal Transduct Ther., 2010. **5**(1): p. 12-24.

64. Comer, F.I. and G.W. Hart, *O-GlcNAc and the control of gene expression.* Biochim Biophys Acta, 1999. **1473**(1): p. 161-71.

65. Love, D.C. and J.A. Hanover, *The hexosamine signaling pathway: deciphering the "O-GlcNAc code".* Sci STKE, 2005. **2005**(312): p. re13.

66. Vosseller, K., et al., *Diverse regulation of protein function by O-GlcNAc: a nuclear and cytoplasmic carbohydrate post-translational modification.* Curr Opin Chem Biol, 2002. **6**(6): p. 851-7.

67. Zachara, N.E. and G.W. Hart, *Cell signaling, the essential role of O-GlcNAc!* Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids, 2006. **1761**(5-6): p. 599-617.

68. Han, I. and J.E. Kudlow, *Reduced O glycosylation of Sp1 is associated with increased proteasome susceptibility.* Mol Cell Biol, 1997. **17**(5): p. 2550-8.

69. Dentin, R., et al., *Hepatic glucose sensing via the CREB coactivator CRTC2.* Science, 2008. **319**(5868): p. 1402-5.

70. Sayat, R., et al., *O-GlcNAc-glycosylation of beta-catenin regulates its nuclear localization and transcriptional activity.* Exp Cell Res, 2008. **314**(15): p. 2774-87.

71. Gao, Y., J. Miyazaki, and G.W. Hart, *The transcription factor PDX-1 is post-translationally modified by O-linked N-acetylglucosamine and this modification is correlated with its DNA binding activity and insulin secretion in min6 beta-cells.* Arch Biochem Biophys, 2003. **415**(2): p. 155-63.

72. Housley, M.P., et al., *O-GlcNAc regulates FoxO activation in response to glucose.* J Biol Chem, 2008. **283**(24): p. 16283-92.

73. Gewinner, C., et al., *The coactivator of transcription CREB-binding protein interacts preferentially with the glycosylated form of Stat5.* J Biol Chem, 2004. **279**(5): p. 3563-72.

74. Hart, G.W., M.P. Housley, and C. Slawson, *Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins.* Nature, 2007. **446**(7139): p. 1017-22.

75. Dias, W.B. and G.W. Hart, *O-GlcNAc modification in diabetes and Alzheimer's disease.* Mol Biosyst, 2007. **3**(11): p. 766-72.

76. Butkinaree, C., K. Park, and G.W. Hart, *O-linked beta-N-acetylglucosamine (O-GlcNAc): Extensive crosstalk with phosphorylation*

to regulate signaling and transcription in response to nutrients and stress. Biochim Biophys Acta, 2009.

77. Webster, D.M., et al., *O-GlcNAc modifications regulate cell survival and epiboly during zebrafish development.* BMC Dev Biol, 2009. **9**: p. 28.

78. Yang, Y.R., et al., *O-GlcNAcase is essential for embryonic development and maintenance of genomic stability.* Aging Cell, 2012. **11**(3): p. 439-48.

79. Slawson, C., et al., *Perturbations in O-linked beta-N-acetylglucosamine protein modification cause severe defects in mitotic progression and cytokinesis.* J Biol Chem, 2005. **280**(38): p. 32944-56.

80. Liu, Y., et al., *Developmental regulation of protein O-GlcNAcylation, O-GlcNAc transferase, and O-GlcNAcase in mammalian brain.* PLoS One, 2012. **7**(8): p. e43724.

81. Maury, J.J., et al., *Excess of O-linked N-acetylglucosamine modifies human pluripotent stem cell differentiation.* Stem Cell Res, 2013. **11**(2): p. 926-37.

82. Francisco, H., et al., *O-GLcNAc post-translational modifications regulate the entry of neurons into an axon branching program.* Dev Neurobiol, 2009. **69**(2-3): p. 162-73.

83. Hayakawa, K., et al., *Epigenetic switching by the metabolism-sensing factors in the generation of orexin neurons from mouse embryonic stem cells.* J Biol Chem, 2013.

84. Laugesen, A. and K. Helin, *Chromatin Repressive Complexes in Stem Cells, Development, and Cancer.* Cell Stem Cell, 2014. **14**(6): p. 735-751.

85. Green, H. and O. Kehinde, *An established preadipose cell line and its differentiation in culture. II. Factors affecting the adipose conversion.* Cell, 1975. **5**(1): p. 19-27.

86. Hsieh, T.J., et al., *Suppression of Glutamine:fructose-6-phosphate amidotransferase-1 inhibits adipogenesis in 3T3-L1 adipocytes.* J Cell Physiol, 2012. **227**(1): p. 108-15.

87. Ishihara, K., et al., *Characteristic increase in nucleocytoplasmic protein glycosylation by O-GlcNAc in 3T3-L1 adipocyte differentiation.* Biochem Biophys Res Commun, 2010.

88. Zachara, N.E., *The roles of O-linked beta-N-acetylglucosamine in cardiovascular physiology and disease.* Am J Physiol Heart Circ Physiol, 2012. **302**(10): p. H1905-18.

89. Kim, H.S., et al., *Excessive O-GlcNAcylation of proteins suppresses spontaneous cardiogenesis in ES cells.* FEBS Lett, 2009. **583**(15): p. 2474-8.

90. Ogawa, M., et al., *Terminal differentiation program of skeletal myogenesis is negatively regulated by O-GlcNAc glycosylation.* Biochim Biophys Acta, 2012. **1820**(1): p. 24-32.

91. Lints, T.J., et al., *Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants.* Development, 1993. **119**(3): p. 969.

92. Kim, H.S., et al., *Cardiac transcription factor Nkx2.5 is downregulated under excessive O-GlcNAcylation condition.* PLoS One, 2012. **7**(6): p. e38053.

93. Ogawa, M., Y. Sakakibara, and K. Kamemura, *Requirement of decreased O-GlcNAc glycosylation of Mef2D for its recruitment to the myogenin promoter.* Biochem Biophys Res Commun, 2013. **433**(4): p. 558-62.

94. Jang, H., et al., *O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network.* Cell Stem Cell, 2012. **11**(1): p. 62-74.

95. Myers, S.A., B. Panning, and A.L. Burlingame, *Polycomb repressive complex 2 is necessary for the normal site-specific O-GlcNAc distribution in mouse embryonic stem cells.* Proc Natl Acad Sci U S A, 2011. **108**(23): p. 9490-5.

96. Chou, T.Y., C.V. Dang, and G.W. Hart, *Glycosylation of the c-Myc transactivation domain.* Proc Natl Acad Sci U S A, 1995. **92**(10): p. 4417-21.

97. Chou, T.Y., G.W. Hart, and C.V. Dang, *c-Myc is glycosylated at threonine 58, a known phosphorylation site and a mutational hot spot in lymphomas.* J Biol Chem, 1995. **270**(32): p. 18961-5.

98. Yang, W.H., et al., *Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability.* Nat Cell Biol, 2006. **8**(10): p. 1074-83.

99. Lachnit, M., E. Kur, and W. Driever, *Alterations of the cytoskeleton in all three embryonic lineages contribute to the epiboly defect of Pou5f1/Oct4 deficient MZspg zebrafish embryos.* Dev Biol, 2008. **315**(1): p. 1-17.

100. Lunde, K., H.G. Belting, and W. Driever, *Zebrafish pou5f1/pou2, homolog of mammalian Oct4, functions in the endoderm specification cascade.* Curr Biol, 2004. **14**(1): p. 48-55.

101. Reim, G., et al., *The POU domain protein spg (pou2/Oct4) is essential for endoderm formation in cooperation with the HMG domain protein casanova.* Dev Cell, 2004. **6**(1): p. 91-101.

102. Scholer, H.R., et al., *A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor.* EMBO J, 1989. **8**(9): p. 2543-50.

103. Okamoto, K., et al., *A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells.* Cell, 1990. **60**(3): p. 461-72.

104. Rosner, M.H., et al., *A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo.* Nature, 1990. **345**(6277): p. 686-92.

105. Scholer, H.R., et al., *New type of POU domain in germ line-specific protein Oct-4.* Nature, 1990. **344**(6265): p. 435-9.

106. Scholer, H.R., *Octamania: the POU factors in murine development.* Trends Genet, 1991. **7**(10): p. 323-9.

107. Tantin, D., *Oct transcription factors in development and stem cells: insights and mechanisms.* Development, 2013. **140**(14): p. 2857-66.

108. Niwa, H., et al., *Phenotypic complementation establishes requirements for specific POU domain and generic transactivation function of Oct-3/4 in embryonic stem cells.* Mol Cell Biol, 2002. **22**(5): p. 1526-36.

109. Peng, S., N.J. Maihle, and Y. Huang, *Pluripotency factors Lin28 and Oct4 identify a sub-population of stem cell-like cells in ovarian cancer.* Oncogene, 2010. **29**(14): p. 2153-9.

110. Wang, Y.D., et al., *OCT4 promotes tumorigenesis and inhibits apoptosis of cervical cancer cells by miR-125b/BAK1 pathway.* Cell Death Dis, 2013. **4**: p. e760.

111. Wen, J., et al., *Oct4 and Nanog expression is associated with early stages of pancreatic carcinogenesis.* Pancreas, 2010. **39**(5): p. 622-6.

112. Atlasi, Y., et al., *OCT-4, an embryonic stem cell marker, is highly expressed in bladder cancer.* Int J Cancer, 2007. **120**(7): p. 1598-602.

113. Strizzi, L., et al., *Emerging roles of nodal and Cripto-1: from embryogenesis to breast cancer progression.* Breast Dis, 2008. **29**: p. 91-103.

114. Kemler, I. and W. Schaffner, *Octamer transcription factors and the cell type-specificity of immunoglobulin gene expression.* Faseb J, 1990. **4**(5): p. 1444-9.

115. Klemm, J.D., et al., *Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules.* Cell, 1994. **77**(1): p. 21-32.

116. Botquin, V., et al., *New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2.* Genes Dev, 1998. **12**(13): p. 2073-90.

117. Tomilin, A., et al., *Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration.* Cell, 2000. **103**(6): p. 853-64.

118. Aksoy, I., et al., *Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm.* EMBO J, 2013. **32**(7): p. 938-53.

119. Scholer, H.R., T. Ciesiolka, and P. Gruss, *A nexus between Oct-4 and E1A: implications for gene regulation in embryonic stem cells.* Cell, 1991. **66**(2): p. 291-304.

120. Ezashi, T., D. Ghosh, and R.M. Roberts, *Repression of Ets-2-induced transactivation of the tau interferon promoter by Oct-4.* Mol Cell Biol, 2001. **21**(23): p. 7883-91.

121. Brehm, A., K. Ohbo, and H. Scholer, *The carboxy-terminal transactivation domain of Oct-4 acquires cell specificity through the POU domain.* Mol Cell Biol, 1997. **17**(1): p. 154-62.

122. Babaie, Y., et al., *Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells.* Stem Cells, 2007. **25**(2): p. 500-10.

123. Chew, J.L., et al., *Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells.* Mol Cell Biol, 2005. **25**(14): p. 6031-46.

124. Rodda, D.J., et al., *Transcriptional regulation of nanog by OCT4 and SOX2.* J Biol Chem, 2005. **280**(26): p. 24731-7.

125. Nishimoto, M., et al., *The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2.* Mol Cell Biol, 1999. **19**(8): p. 5453-65.

126. Strumpf, D., et al., *Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst.* Development, 2005. **132**(9): p. 2093-102.

127. Niwa, H., J. Miyazaki, and A.G. Smith, *Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.* Nat Genet, 2000. **24**(4): p. 372-6.

128. Rodriguez, R.T., et al., *Manipulation of OCT4 levels in human embryonic stem cells results in induction of differential cell types.* Exp Biol Med (Maywood), 2007. **232**(10): p. 1368-80.

129. Hay, D.C., et al., *Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells.* Stem Cells, 2004. **22**(2): p. 225-35.

130. Saxe, J.P., et al., *Post-translational regulation of Oct4 transcriptional activity.* PLoS ONE, 2009. **4**(2): p. e4467.

131. Xu, H., et al., *WWP2 promotes degradation of transcription factor OCT4 in human embryonic stem cells.* Cell Res, 2009. **19**(5): p. 561-73.

132. Xu, H.M., et al., *Wwp2, an E3 ubiquitin ligase that targets transcription factor Oct-4 for ubiquitination.* J Biol Chem, 2004. **279**(22): p. 23495-503.

133. Wei, F., H.R. Scholer, and M.L. Atchison, *Sumoylation of Oct4 enhances its stability, DNA binding, and transactivation.* J Biol Chem, 2007. **282**(29): p. 21551-60.

134. Tsuruzoe, S., et al., *Inhibition of DNA binding of Sox2 by the SUMO conjugation.* Biochem Biophys Res Commun, 2006. **351**(4): p. 920-6.

135. Brumbaugh, J., et al., *Phosphorylation regulates human OCT4.* Proc Natl Acad Sci U S A, 2012. **109**(19): p. 7162-8.

136. Kang, J., et al., *A general mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and oxidative stress.* Genes Dev, 2009. **23**(2): p. 208-22.

137. Swaney, D.L., et al., *Human embryonic stem cell phosphoproteome revealed by electron transfer dissociation tandem mass spectrometry.* Proc Natl Acad Sci U S A, 2009. **106**(4): p. 995-1000.

138. Johnson, E.S., *Ubiquitin branches out.* Nat Cell Biol, 2002. **4**(12): p. E295-8.

139. Glickman, M.H. and A. Ciechanover, *The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction.* Physiol Rev, 2002. **82**(2): p. 373-428.

140. Geiss-Friedlander, R. and F. Melchior, *Concepts in sumoylation: a decade on.* Nat Rev Mol Cell Biol, 2007. **8**(12): p. 947-56.

141. Wu, Y., et al., *SUMOylation represses Nanog expression via modulating transcription factors Oct4 and Sox2.* PLoS One, 2012. **7**(6): p. e39606.

142. Whitmarsh, A.J. and R.J. Davis, *Regulation of transcription factor function by phosphorylation.* Cell Mol Life Sci, 2000. **57**(8-9): p. 1172-83.

143. Holmberg, C.I., et al., *Multisite phosphorylation provides sophisticated regulation of transcription factors.* Trends Biochem Sci, 2002. **27**(12): p. 619-27.

144. Rosfjord, E., et al., *Phosphorylation and DNA binding of the octamer binding transcription factor Oct-3.* Biochem Biophys Res Commun, 1995. **212**(3): p. 847-53.

145. Spelat, R., F. Ferro, and F. Curcio, *Serine 111 phosphorylation regulates OCT4A protein subcellular distribution and degradation.* J Biol Chem, 2012. **287**(45): p. 38279-88.

146. Lin, Y., et al., *Reciprocal regulation of Akt and Oct4 promotes the self-renewal and survival of embryonal carcinoma cells.* Mol Cell, 2012. **48**(4): p. 627-40.

147. Swaney, D.L., et al., *Human embryonic stem cell phosphoproteome revealed by electron transfer dissociation tandem mass spectrometry.* Proc Natl Acad Sci U S A, 2009.

148. Remenyi, A., et al., *Differential dimer activities of the transcription factor Oct-1 by DNA-induced interface swapping.* Mol Cell, 2001. **8**(3): p. 569-80.

149. Brivanlou, A.H. and J.E. Darnell, Jr., *Signal transduction and the control of gene expression.* Science, 2002. **295**(5556): p. 813-8.

150. Orphanides, G. and D. Reinberg, *A unified theory of gene expression.* Cell, 2002. **108**(4): p. 439-51.

151. Levine, M. and R. Tjian, *Transcription regulation and animal diversity.* Nature, 2003. **424**(6945): p. 147-51.

152. Heintzman, N.D. and B. Ren, *The gateway to transcription: identifying, characterizing and understanding promoters in the eukaryotic genome.* Cell Mol Life Sci, 2007. **64**(4): p. 386-400.

153. Khidekel, N. and L.C. Hsieh-Wilson, *A 'molecular switchboard'--covalent modifications to proteins and their impact on transcription.* Org Biomol Chem, 2004. **2**(1): p. 1-7.

154. Torres, C.R. and G.W. Hart, *Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc.* J Biol Chem, 1984. **259**(5): p. 3308-17.

155. Holt, G.D. and G.W. Hart, *The subcellular distribution of terminal N-acetylglucosamine moieties. Localization of a novel protein-saccharide linkage, O-linked GlcNAc.* J Biol Chem, 1986. **261**(17): p. 8049-57.

156. Holt, G.D., et al., *Erythrocytes contain cytoplasmic glycoproteins. O-linked GlcNAc on Band 4.1.* J Biol Chem, 1987. **262**(31): p. 14847-50.

157. Haltiwanger, R.S., G.D. Holt, and G.W. Hart, *Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-acetylglucosamine:peptide beta-N-acetylglucosaminyltransferase.* J Biol Chem, 1990. **265**(5): p. 2563-8.

158. Kreppel, L.K., M.A. Blomberg, and G.W. Hart, *Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats.* J Biol Chem, 1997. **272**(14): p. 9308-15.

159. Lubas, W.A. and J.A. Hanover, *Functional expression of O-linked GlcNAc transferase. Domain structure and substrate specificity.* J Biol Chem, 2000. **275**(15): p. 10983-8.

160. Comer, F.I. and G.W. Hart, *O-Glycosylation of nuclear and cytosolic proteins. Dynamic interplay between O-GlcNAc and O-phosphate.* J Biol Chem, 2000. **275**(38): p. 29179-82.

161. Wells, L., et al., *O-GlcNAc transferase is in a functional complex with protein phosphatase 1 catalytic subunits.* J Biol Chem, 2004. **279**(37): p. 38466-70.

162. Copeland, R.J., J.W. Bullen, and G.W. Hart, *Cross-talk between GlcNAcylation and phosphorylation: roles in insulin resistance and glucose toxicity.* Am J Physiol Endocrinol Metab, 2008. **295**(1): p. E17-28.

163. Dorfmueller, H.C., et al., *GlcNAcstatins are nanomolar inhibitors of human O-GlcNAcase inducing cellular hyper-O-GlcNAcylation.* Biochem J, 2009.

164. Gross, B.J., J.G. Swoboda, and S. Walker, *A strategy to discover inhibitors of O-linked glycosylation.* J Am Chem Soc, 2008. **130**(2): p. 440-1.

165. Vosseller, K., et al., *O-linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry.* Mol Cell Proteomics, 2006. **5**(5): p. 923-34.

166. Khidekel, N., et al., *Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain.* Proc Natl Acad Sci U S A, 2004. **101**(36): p. 13132-7.

167. Khidekel, N., et al., *Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics.* Nat Chem Biol, 2007. **3**(6): p. 339-48.

168. Wells, L., et al., *Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications.* Mol Cell Proteomics, 2002. **1**(10): p. 791-804.

169. Issad, T. and M. Kuo, *O-GlcNAc modification of transcription factors, glucose sensing and glucotoxicity.* Trends Endocrinol Metab, 2008. **19**(10): p. 380-9.

170. Zachara, N.E., *Detecting the "O-GlcNAc-ome"; detection, purification, and analysis of O-GlcNAc modified proteins.* Methods Mol Biol, 2009. **534**: p. 251-79.

171. Hanover, J.A., et al., *O-linked N-acetylglucosamine is attached to proteins of the nuclear pore. Evidence for cytoplasmic and nucleoplasmic glycoproteins.* J Biol Chem, 1987. **262**(20): p. 9887-94.

172. Snow, C.M., A. Senior, and L. Gerace, *Monoclonal antibodies identify a group of nuclear pore complex glycoproteins.* J Cell Biol, 1987. **104**(5): p. 1143-56.

173. Comer, F.I. and G.W. Hart, *Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II.* Biochemistry, 2001. **40**(26): p. 7845-52.

174. Khidekel, N., et al., *A chemoenzymatic approach toward the rapid and sensitive detection of O-GlcNAc posttranslational modifications.* J Am Chem Soc, 2003. **125**(52): p. 16162-3.

175. Vocadlo, D.J., et al., *A chemical approach for identifying O-GlcNAc-modified proteins in cells.* Proc Natl Acad Sci U S A, 2003. **100**(16): p. 9116-21.

176. Saxon, E. and C.R. Bertozzi, *Cell surface engineering by a modified Staudinger reaction.* Science, 2000. **287**(5460): p. 2007-10.

177. Whelan, S.A. and G.W. Hart, *Identification of O-GlcNAc sites on proteins.* Methods Enzymol, 2006. **415**: p. 113-33.

178. Nandi, A., et al., *Global identification of O-GlcNAc-modified proteins.* Anal Chem, 2006. **78**(2): p. 452-8.

179. Zubarev, R.A., N.L. Kelleher, and F.W. McLafferty, *Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process.* Journal of the American Chemical Society, 1998. **120**(13): p. 3265-3266.

180. Syka, J.E., et al., *Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry.* Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9528-33.

181. Mikesh, L.M., et al., *The utility of ETD mass spectrometry in proteomic analysis.* Biochim Biophys Acta, 2006. **1764**(12): p. 1811-22.

182. Wiesner, J., T. Premsler, and A. Sickmann, *Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications.* Proteomics, 2008. **8**(21): p. 4466-83.

183. Viner, R.I., et al., *Quantification of post-translationally modified peptides of bovine alpha-crystallin using tandem mass tags and electron transfer dissociation.* J Proteomics, 2009.

184. O'Malley, B.W., J. Qin, and R.B. Lanz, *Cracking the coregulator codes.* Curr Opin Cell Biol, 2008. **20**(3): p. 310-5.

185. Lyst, M.J. and I. Stancheva, *A role for SUMO modification in transcriptional repression and activation.* Biochem Soc Trans, 2007. **35**(Pt 6): p. 1389-92.

186. Spange, S., et al., *Acetylation of non-histone proteins modulates cellular signalling at multiple levels.* Int J Biochem Cell Biol, 2009. **41**(1): p. 185-98.

187. Clapier, C.R. and B.R. Cairns, *The Biology of Chromatin Remodeling Complexes.* Annu Rev Biochem, 2009.

188. Lee, T.I. and R.A. Young, *Transcription of eukaryotic protein-coding genes.* Annu Rev Genet, 2000. **34**: p. 77-137.

189. Zhang, Y. and D. Reinberg, *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.* Genes Dev, 2001. **15**(18): p. 2343-60.

190. Kelly, W.G. and G.W. Hart, *Glycosylation of chromosomal proteins: localization of O-linked N-acetylglucosamine in Drosophila chromatin.* Cell, 1989. **57**(2): p. 243-51.

191. Dannenberg, J.H., et al., *mSin3A corepressor regulates diverse transcriptional networks governing normal and neoplastic growth and survival.* Genes Dev, 2005. **19**(13): p. 1581-95.

192. Shaw, P., et al., *Regulation of specific DNA binding by p53: evidence for a role for O-glycosylation and charged residues at the carboxy-terminus.* Oncogene, 1996. **12**(4): p. 921-30.

193. Murphy, M., et al., *Transcriptional repression by wild-type p53 utilizes histone deacetylases, mediated by interaction with mSin3a.* Genes Dev, 1999. **13**(19): p. 2490-501.

194. Yao, D., et al., *High glucose increases angiopoietin-2 transcription in microvascular endothelial cells through methylglyoxal modification of mSin3A.* J Biol Chem, 2007. **282**(42): p. 31038-45.

195. Corden, J.L., *Tails of RNA polymerase II.* Trends Biochem Sci, 1990. **15**(10): p. 383-7.

196. Kelly, W.G., M.E. Dahmus, and G.W. Hart, *RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc.* J Biol Chem, 1993. **268**(14): p. 10416-24.

197. Desterro, J.M., M.S. Rodriguez, and R.T. Hay, *Regulation of transcription factors by protein degradation.* Cell Mol Life Sci, 2000. **57**(8-9): p. 1207-19.

198. Ravid, T. and M. Hochstrasser, *Diversity of degradation signals in the ubiquitin-proteasome system.* Nat Rev Mol Cell Biol, 2008. **9**(9): p. 679-90.

199. Zhang, F., et al., *O-GlcNAc modification is an endogenous inhibitor of the proteasome.* Cell, 2003. **115**(6): p. 715-25.

200. Wierstra, I., *Sp1: Emerging roles--Beyond constitutive activation of TATA-less housekeeping genes.* Biochemical and Biophysical Research Communications, 2008. **372**(1): p. 1-13.

201. Jackson, S.P. and R. Tjian, *O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation.* Cell, 1988. **55**(1): p. 125-33.

202. Sumegi, M., et al., *26S proteasome subunits are O-linked N-acetylglucosamine-modified in Drosophila melanogaster.* Biochem Biophys Res Commun, 2003. **312**(4): p. 1284-9.

203. Cheng, X. and G.W. Hart, *Alternative O-glycosylation/O-phosphorylation of serine-16 in murine estrogen receptor beta: post-translational regulation of turnover and transactivation activity.* J Biol Chem, 2001. **276**(13): p. 10570-5.

204. Sears, R., et al., *Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability.* Genes Dev, 2000. **14**(19): p. 2501-14.

205. Cheng, X., et al., *Alternative O-glycosylation/O-phosphorylation of the murine estrogen receptor beta.* Biochemistry, 2000. **39**(38): p. 11609-20.

206. Yustein, J.T. and C.V. Dang, *Biology and treatment of Burkitt's lymphoma.* Curr Opin Hematol, 2007. **14**(4): p. 375-81.

207. Vervoorts, J., J. Luscher-Firzlaff, and B. Luscher, *The ins and outs of MYC regulation by posttranslational mechanisms.* J Biol Chem, 2006. **281**(46): p. 34725-9.

208. Adhikary, S. and M. Eilers, *Transcriptional regulation and transformation by Myc proteins.* Nat Rev Mol Cell Biol, 2005. **6**(8): p. 635-645.

209. Kuiper, G.G. and J.A. Gustafsson, *The novel estrogen receptor-beta subtype: potential role in the cell- and promoter-specific actions of estrogens and anti-estrogens.* FEBS Lett, 1997. **410**(1): p. 87-90.

210. Zhao, C., K. Dahlman-Wright, and J.A. Gustafsson, *Estrogen receptor beta: an overview and update.* Nucl Recept Signal, 2008. **6**: p. e003.

211. Grisouard, J., et al., *Glycogen synthase kinase-3 protects estrogen receptor alpha from proteasomal degradation and is required for full transcriptional activity of the receptor.* Mol Endocrinol, 2007. **21**(10): p. 2427-39.

212. Reid, G., et al., *Cyclic, proteasome-mediated turnover of unliganded and liganded ERalpha on responsive promoters is an integral feature of estrogen signaling.* Mol Cell, 2003. **11**(3): p. 695-707.

213. Picard, N., et al., *Phosphorylation of activation function-1 regulates proteasome-dependent nuclear mobility and E6-associated protein ubiquitin ligase recruitment to the estrogen receptor beta.* Mol Endocrinol, 2008. **22**(2): p. 317-30.

214. Whibley, C., P.D. Pharoah, and M. Hollstein, *p53 polymorphisms: cancer implications.* Nat Rev Cancer, 2009. **9**(2): p. 95-107.

215. Petersen, H.V., et al., *Glucose induced MAPK signalling influences NeuroD1-mediated activation and nuclear localization.* FEBS Lett, 2002. **528**(1-3): p. 241-5.

216. Andrali, S.S., Q. Qian, and S. Ozcan, *Glucose mediates the translocation of NeuroD1 by O-linked glycosylation.* J Biol Chem, 2007. **282**(21): p. 15589-96.

217. Koo, S.H., et al., *The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism.* Nature, 2005. **437**(7062): p. 1109-11.

218. Dentin, R., et al., *Insulin modulates gluconeogenesis by inhibition of the coactivator TORC2.* Nature, 2007. **449**(7160): p. 366-9.

219. Screaton, R.A., et al., *The CREB coactivator TORC2 functions as a calcium- and cAMP-sensitive coincidence detector.* Cell, 2004. **119**(1): p. 61-74.

220. Andrali, S.S., et al., *Glucose regulation of insulin gene expression in pancreatic beta-cells.* Biochem J, 2008. **415**(1): p. 1-10.

221. Barker, N., *The canonical Wnt/beta-catenin signalling pathway.* Methods Mol Biol, 2008. **468**: p. 5-15.

222. Zhu, W., B. Leber, and D.W. Andrews, *Cytoplasmic O-glycosylation prevents cell surface transport of E-cadherin during apoptosis.* EMBO J, 2001. **20**(21): p. 5999-6007.

223. Akimoto, Y., et al., *Elevation of the post-translational modification of proteins by O-linked N-acetylglucosamine leads to deterioration of the glucose-stimulated insulin secretion in the pancreas of diabetic Goto-Kakizaki rats.* Glycobiology, 2007. **17**(2): p. 127-40.

224. Zhang, W., et al., *FoxO1 regulates multiple metabolic pathways in the liver: effects on gluconeogenic, glycolytic, and lipogenic gene expression.* J Biol Chem, 2006. **281**(15): p. 10105-17.

225. Finck, B.N. and D.P. Kelly, *PGC-1 coactivators: inducible regulators of energy metabolism in health and disease.* J Clin Invest, 2006. **116**(3): p. 615-22.

226. Puigserver, P., et al., *Insulin-regulated hepatic gluconeogenesis through FOXO1-PGC-1alpha interaction.* Nature, 2003. **423**(6939): p. 550-5.

227. Barthel, A., D. Schmoll, and T.G. Unterman, *FoxO proteins in insulin action and metabolism.* Trends Endocrinol Metab, 2005. **16**(4): p. 183-9.

228. Gross, D.N., A.P. van den Heuvel, and M.J. Birnbaum, *The role of FoxO in the regulation of metabolism.* Oncogene, 2008. **27**(16): p. 2320-36.

229. Kuo, M., et al., *O-glycosylation of FoxO1 increases its transcriptional activity towards the glucose 6-phosphatase gene.* FEBS Lett, 2008. **582**(5): p. 829-34.

230. Housley, M.P., et al., *A PGC-1{alpha}-O-GlcNAc Transferase Complex Regulates FoxO Transcription Factor Activity in Response to Glucose.* J Biol Chem, 2009. **284**(8): p. 5148-57.

231. Yang, W.H., et al., *NFkappaB activation is associated with its O-GlcNAcylation state under hyperglycemic conditions.* Proc Natl Acad Sci U S A, 2008. **105**(45): p. 17345-50.

232. Yang, X., et al., *O-linkage of N-acetylglucosamine to Sp1 activation domain inhibits its transcriptional capability.* Proc Natl Acad Sci U S A, 2001. **98**(12): p. 6611-6.

233. Roos, M.D., et al., *O glycosylation of an Sp1-derived peptide blocks known Sp1 protein interactions.* Mol Cell Biol, 1997. **17**(11): p. 6472-80.

234. Levy, D.E. and J.E. Darnell, Jr., *Stats: transcriptional control and biological impact.* Nat Rev Mol Cell Biol, 2002. **3**(9): p. 651-62.

235. Korzus, E., et al., *Transcription factor-specific requirements for coactivators and their acetyltransferase functions.* Science, 1998. **279**(5351): p. 703-7.

236. Hayden, M.S. and S. Ghosh, *Shared principles in NF-kappaB signaling.* Cell, 2008. **132**(3): p. 344-62.

237. James, L.R., et al., *Flux through the hexosamine pathway is a determinant of nuclear factor kappaB- dependent promoter activation.* Diabetes, 2002. **51**(4): p. 1146-56.

238. Kawauchi, K., et al., *Loss of p53 enhances catalytic activity of IKKbeta through O-linked beta-N-acetyl glucosamine modification.* Proc Natl Acad Sci U S A, 2009. **106**(9): p. 3431-6.

239. Kawauchi, K., et al., *p53 regulates glucose metabolism through an IKK-NF-kappaB pathway and inhibits cell transformation.* Nat Cell Biol, 2008. **10**(5): p. 611-8.

240. Schomer-Miller, B., et al., *Regulation of IkappaB kinase (IKK) complex by IKKgamma-dependent phosphorylation of the T-loop and C terminus of IKKbeta.* J Biol Chem, 2006. **281**(22): p. 15268-76.

241. Kreppel, L.K. and G.W. Hart, *Regulation of a cytosolic and nuclear O-GlcNAc transferase. Role of the tetratricopeptide repeats.* J Biol Chem, 1999. **274**(45): p. 32015-22.

242. Iyer, S.P., Y. Akimoto, and G.W. Hart, *Identification and cloning of a novel family of coiled-coil domain proteins that interact with O-GlcNAc transferase.* J Biol Chem, 2003. **278**(7): p. 5399-409.

243. Cheung, W.D., et al., *O-linked beta-N-acetylglucosaminyltransferase substrate specificity is regulated by myosin phosphatase targeting and other interacting proteins.* J Biol Chem, 2008. **283**(49): p. 33935-41.

244. Stallcup, M.R., et al., *The roles of protein-protein interactions and protein methylation in transcriptional activation by nuclear receptors and their coactivators.* J Steroid Biochem Mol Biol, 2003. **85**(2-5): p. 139-45.

245. Iyer, S.P. and G.W. Hart, *Roles of the tetratricopeptide repeat domain in O-GlcNAc transferase targeting and protein substrate specificity.* J Biol Chem, 2003. **278**(27): p. 24608-16.

246. Hurtado-Guerrero, R., H.C. Dorfmueller, and D.M. van Aalten, *Molecular mechanisms of O-GlcNAcylation.* Curr Opin Struct Biol, 2008. **18**(5): p. 551-7.

247. Whisenhunt, T.R., et al., *Disrupting the enzyme complex regulating O-GlcNAcylation blocks signaling and development.* Glycobiology, 2006. **16**(6): p. 551-63.

248. Verkerk, A.J., et al., *Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome.* Cell, 1991. **65**(5): p. 905-14.

249. Garber, K., et al., *Transcription, translation and fragile X syndrome.* Curr Opin Genet Dev, 2006. **16**(3): p. 270-5.

250. Garnon, J., et al., *Fragile X-related protein FXR1P regulates proinflammatory cytokine tumor necrosis factor expression at the post-transcriptional level.* J Biol Chem, 2005. **280**(7): p. 5750-63.

251. Vietor, I. and L.A. Huber, *Role of TIS7 family of transcriptional regulators in differentiation and regeneration.* Differentiation, 2007. **75**(9): p. 891-7.

252. Vietor, I., et al., *TIS7 interacts with the mammalian SIN3 histone deacetylase complex in epithelial cells.* EMBO J, 2002. **21**(17): p. 4621-31.

253. Vietor, I., et al., *TIS7 regulation of the beta-catenin/Tcf-4 target gene osteopontin (OPN) is histone deacetylase-dependent.* J Biol Chem, 2005. **280**(48): p. 39795-801.

254. Wick, N., et al., *Inhibitory effect of TIS7 on Sp1-C/EBPalpha transcription factor module activity.* J Mol Biol, 2004. **336**(3): p. 589-95.

255. Ryzhakov, G. and F. Randow, *SINTBAD, a novel component of innate antiviral immunity, shares a TBK1-binding domain with NAP1 and TANK.* EMBO J, 2007. **26**(13): p. 3180-90.

256. Hiscott, J., et al., *Manipulation of the nuclear factor-kappaB pathway and the innate immune response by viruses.* Oncogene, 2006. **25**(51): p. 6844-67.

257. Teo, C.F., et al., *Generation of O-GlcNAc Specific Monoclonal Antibodies Using a Novel Synthetic Immunogen.* **In Press**.

258. Wang, Z., A. Pandey, and G.W. Hart, *Dynamic interplay between O-linked N-acetylglucosaminylation and glycogen synthase kinase-3-dependent phosphorylation.* Mol Cell Proteomics, 2007. **6**(8): p. 1365-79.

259. Sekine, O., et al., *Blocking O-linked GlcNAc cycling in Drosophila insulin-producing cells perturbs glucose-insulin homeostasis.* J Biol Chem, 2010. **285**(49): p. 38684-91.

260. Teo, C.F., et al., *Glycopeptide-specific monoclonal antibodies suggest new roles for O-GlcNAc.* Nat Chem Biol, 2010.

261. Remenyi, A., et al., *Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers.* Genes Dev, 2003. **17**(16): p. 2048-59.

262. Pardo, M., et al., *An expanded Oct4 interaction network: implications for stem cell biology, development, and disease.* Cell Stem Cell, 2010. **6**(4): p. 382-95.

263. van den Berg, D.L., et al., *An Oct4-centered protein interaction network in embryonic stem cells.* Cell Stem Cell, 2010. **6**(4): p. 369-81.

264. Schnerch, A., C. Cerdan, and M. Bhatia, *Distinguishing between mouse and human pluripotent stem cell regulation: the best laid plans of mice and men.* Stem Cells, 2010. **28**(3): p. 419-30.

265. Comer, F.I., et al., *Characterization of a mouse monoclonal antibody specific for O-linked N-acetylglucosamine.* Anal Biochem, 2001. **293**(2): p. 169-77.

266. Ryan, A.K. and M.G. Rosenfeld, *POU domain family values: flexibility, partnerships, and developmental codes.* Genes Dev, 1997. **11**(10): p. 1207-25.

267. Di Rocco, G., et al., *The recruitment of SOX/OCT complexes and the differential activity of HOXA1 and HOXB1 modulate the Hoxb1 auto-regulatory enhancer function.* J Biol Chem, 2001. **276**(23): p. 20506-15.

268. Yuzwa, S.A., et al., *Mapping O-GlcNAc modification sites on tau and generation of a site-specific O-GlcNAc tau antibody.* Amino Acids, 2011. **40**(3): p. 857-68.

269. Singh, A.M., et al., *Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation.* Cell Stem Cell, 2012. **10**(3): p. 312-26.

270. Martinez-Fleites, C., et al., *Structure of an O-GlcNAc transferase homolog provides insight into intracellular glycosylation.* Nat Struct Mol Biol, 2008. **15**(7): p. 764-5.

271. Fujiki, R., et al., *GlcNAcylation of histone H2B facilitates its monoubiquitination.* Nature, 2011. **480**(7378): p. 557-60.

272. Guinez, C., et al., *Protein ubiquitination is modulated by O-GlcNAc glycosylation.* Faseb J, 2008. **22**(8): p. 2901-11.

273. Dalton, S., *Signaling networks in human pluripotent stem cells.* Curr Opin Cell Biol, 2013. **25**(2): p. 241-6.

274. Kang, J., et al., *Regulation of Oct1/Pou2f1 transcription activity by O-GlcNAcylation.* FASEB J, 2013. **27**(7): p. 2807-17.

275. Campbell, P.A. and M.A. Rudnicki, *Oct4 interaction with Hmgb2 regulates Akt signaling and pluripotency.* Stem Cells, 2013. **31**(6): p. 1107-20.

276. Capotosti, F., et al., *O-GlcNAc transferase catalyzes site-specific proteolysis of HCF-1.* Cell, 2011. **144**(3): p. 376-88.

277. Hahne, H., A. Moghaddas Gholami, and B. Kuster, *Discovery of O-GlcNAc-modified proteins in published large-scale proteome data.* Mol Cell Proteomics, 2012. **11**(10): p. 843-50.

278. Li, X., et al., *O-linked N-acetylglucosamine modification on CCAAT enhancer-binding protein beta: role during adipocyte differentiation.* J Biol Chem, 2009. **284**(29): p. 19248-54.

279. Szkudelski, T., *The mechanism of alloxan and streptozotocin action in B cells of the rat pancreas.* Physiol Res, 2001. **50**(6): p. 537-46.

280. Speakman, C.M., et al., *Elevated O-GlcNAc levels activate epigenetically repressed genes and delay mouse ES cell differentiation without affecting naive to primed cell transition.* Stem Cells, 2014.

281. Yuzwa, S.A. and D.J. Vocadlo, *O-GlcNAc and neurodegeneration: biochemical mechanisms and potential roles in Alzheimer's disease and beyond.* Chem Soc Rev, 2014.

282. Shi, F.T., et al., *Ten-eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells.* J Biol Chem, 2013. **288**(29): p. 20776-84.

283. Baltus, G.A., et al., *A positive regulatory role for the mSin3A-HDAC complex in pluripotency through Nanog and Sox2.* J Biol Chem, 2009. **284**(11): p. 6998-7006.

284. Tokuzawa, Y., et al., *Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development.* Mol Cell Biol, 2003. **23**(8): p. 2699-708.

285. Hanover, J.A., M.W. Krause, and D.C. Love, *Bittersweet memories: linking metabolism to epigenetics through O-GlcNAcylation.* Nat Rev Mol Cell Biol, 2012. **13**(5): p. 312-21.

286. Hart, G.W., et al., *Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease.* Annu Rev Biochem, 2011. **80**: p. 825-58.

287. Wang, Z., M. Gucek, and G.W. Hart, *Cross-talk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated O-GlcNAc.* Proc Natl Acad Sci U S A, 2008. **105**(37): p. 13793-8.

288. Kadamb, R., et al., *Sin3: insight into its transcription regulatory functions.* Eur J Cell Biol, 2013. **92**(8-9): p. 237-46.

289. Chen, Q., et al., *TET2 promotes histone O-GlcNAcylation during gene transcription.* Nature, 2013. **493**(7433): p. 561-4.

290. Deplus, R., et al., *TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS.* EMBO J, 2013. **32**(5): p. 645-55.

291. Ito, R., et al., *TET3-OGT interaction increases the stability and the presence of OGT in chromatin.* Genes Cells, 2014. **19**(1): p. 52-65.

292. Vella, P., et al., *Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells.* Mol Cell, 2013. **49**(4): p. 645-56.