Site-specific interplay between O-GlcNAcylation and phosphorylation in cellular regulation

Ping Hu, Shino Shimoji, Gerald W. Hart

Department of Biological Chemistry, Johns Hopkins University, School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205-2185, USA

A R T I C L E   I N F O

Article history:
Received 12 April 2010
Accepted 15 April 2010
Available online 22 April 2010

Edited by Stefan Hohmann

Keywords:
O-GlcNAcylation
Phosphorylation
OGT
OGA
Signalling
Mass spectrometry
O-GlcNAc

A B S T R A C T

Ser(Thr)-O-linked β-N-acetylglucosamine (O-GlcNac) is a ubiquitous modification of nucleocytoplasmic proteins. Extensive crosstalk exists between O-GlcNAcylation and phosphorylation, which regulates signaling in response to nutrients/stress. The development of novel O-GlcNAc detection and enrichment methods has improved our understanding of O-GlcNAc functions. Mass spectrometry has revealed O-GlcNAc’s many interactions with phosphorylation-mediated signaling. However, mechanisms regulating O-GlcNAcylation and phosphorylation are quite different. Phosphorylation is catalyzed by hundreds of distinct kinases. In contrast, in mammals, uridine diphospho-β-N-acetylglucosamine:polypeptide β-N-acetylglucosaminyl transferase (OGT) and β-D-N-acetylglucosaminidase (OGA) are encoded by single highly conserved genes. Both OGT’s and OGA’s specificities are determined by their transient associations with many other proteins to create a multitude of specific holoenzymes. The extensive crosstalk between O-GlcNAcylation and phosphorylation represents a new paradigm for cellular signaling.

© 2010 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.

1. Introduction

O-GlcNAcylation is a common posttranslational modification (PTM) of nuclear and cytosolic proteins. O-GlcNAcylation was serendipitously discovered in experiments designed to detect terminal GlcNAc residues on the surfaces of living lymphocytes, but surprisingly nearly all of the O-GlcNAc was found to be on nucleocytoplasmic proteins, where dogma said protein glycosylation was absent [1,2]. O-GlcNAcylation occurs in some bacteria, many protozoa, viruses, filamentous fungi and in all metazoans, including plants, but thus far has not been definitively documented in yeast. O-GlcNAcylation is distinguished from “classical” glycosylation in the following respects: (1) It is nearly exclusively on cytoplasmic and nuclear proteins, compared to the extracellular and luminal localization for most other protein-bound glycosyl moieties. (2) O-GlcNAc is generally not modified or elongated to more complex structures. (3) Like phosphorylation, O-GlcNAcylation occurs at stoichiometric amounts and rapidly cycles on and off polypeptides at different rates and at specific sites in response to different physiological stimuli, such as hormones, growth factors, and mitogens [3–6] (more in review [7]).

OGT and OGA are the paired enzymes responsible for the addition and removal of O-GlcNAc, respectively. Although O-GlcNAcylation has many properties similar to phosphorylation, the two PTMs are regulated very differently. Phosphorylation is determined by hundreds of specific kinases, each encoded by a distinct gene. In
contrast, both OGT and OGA are encoded by a single highly conserved gene in animals (two alleles of OGT in plants [8]), which contains their catalytic centers and numerous protein–protein interaction domains [9–11]. Like RNA polymerase II (RNAP II), phosphatases [12] and other enzymes, OGT's and OGA's substrate specificity is mostly determined by its transient interactions with many binding partners to form various substrate specific holoenzymes [13,14]. The OGT catalytic subunit also has pronounced specificity to peptide sequence around the attachment site, but no absolute consensus sequence exists. In addition, OGT's enzymatic activity is highly responsive to the concentration of uridine diphospho-N-acetylglucosamine (UDP-GlcNAc), the donor substrate for O-GlcNAcylation [15].

OGT is found in many organisms, from Caenorhabditis elegans, Drosophila, plants to human, and shares very high homology among species [7]. The Ogt gene resides on the chromosome near the centromere. Knock-out studies have shown that Ogt is required for the viability of embryonic stem (ES) cells and other cells, even in cell culture [16]. Tissue targeted Ogt mutations in mice and several cell lines, such as T-lymphocytes, neuronal cells, and fibroblasts, using the Cre-loxP recombination system, demonstrated that loss of O-GlcNAcylation results in concomitant loss of cell function and eventually cell death [16]. In contrast to plants and mammals, in C. elegans, Ogt knockout, while producing profound metabolic changes, such as insulin resistance and abnormal metabolism, does not cause the death of the worm [17]. This and a recent genetic study in C. elegans support a role for O-GlcNAcylation in regulating cellular signaling, insulin resistance, longevity, stress as well as immunity [18]. For humans, there are at least three transcript isoforms of OGT: 110 kilodalton (kDa) (p110), 78 kDa (p78) and 103 kDa (p103), with variation from 2.5 to 13.5 tetra- and pentacopeptide repeats (TPRs). The catalytic p110 subunit can form either a homodimer or heterotrimer with p78 in the cytoplasm and nucleus depending on cell type, while p103 is localized to the mitochondria [7,9,19]. OGT's catalytic subunit is composed of two separate domains. Its N-terminal contains TPRs, which mediate protein–protein interactions, thus functioning in OGT substrate recognition and multimerization. The C-terminus of p110 has the catalytic domain, possessing glycosyltransferase activity [20,21]. Structural analysis of the N-terminal TPR domain shows a right-handed superhelix of about 100 Å long and 35 Å wide, which likely provides for substrate binding on its concave surface [22]. The crystal structure of Xanthomonas campestris OGT, a bacterial orthologue of human OGT, shows that OGT has a conserved UDP-GlcNAc binding pocket. This structure also shows that the last few TPR repeats are, closely associated with the C-terminal active site, forming an elongated, continuous groove from the donor substrate binding site to the end of the TPR superhelix [23,24]. Molecular modeling of the bacterial structure and the mammalian TPR repeat structure strongly supports the structure proposed for the mammalian OGT [13–15].

OGA was initially identified as hexosaminidase C, with β-linked GlcNAc substrate specificity, neutral pH optima, and mainly a cytosolic localization, unlike lysosomal hexosaminidases [25–27]. Strikingly, Oga knockouts in C. elegans showed similar metabolic defects as the Ogt knockouts, such as elevated stores of glycogen, and decreased lipid storage. This finding suggests that O-GlcNAc cycling rates and not its absolute stoichiometry may be important in its regulation of cellular metabolism. Human OGA (hOGA) is a 92 kDa bifunctional protein, its N-terminus contains O-GlcNAc hydrolase activity, and its C-terminus bears a putative histone acetyltransferase (HAT) domain. Caspase-3, a cysteine-aspartic protease that serves as the "executioner protease" in apoptosis, cleaves these two domains during apoptosis without affecting the O-GlcNAc hydrolase catalytic activity [28,29]. hOGA catalyzes O-GlcNAc hydrolisis using a substrate-assisted mechanism, which requires the involvement of the 2-acetamido group of GlcNAc as the nucleophile [30]. Similar catalytic mechanisms were also confirmed on bacterial OGA orthologues from Clostridium perfringens NagJ and Bacteroides thetaiotaomicron BtGH84 [31,32].

2. O-GlcNAcylation is a nutrient and stress sensor

As stated above, UDP-GlcNAc, the immediate donor substrate for O-GlcNAcylation, is a terminal product of the hexosamine biosynthetic pathway (HBP) (Fig. 1). This sugar nucleotide is second only to adenosine triphosphate (ATP) in abundance for high-energy small molecules in cells. When HBP flux changes, UDP-GlcNAc levels also change rapidly, altering the extent of O-GlcNAcylation of many proteins. HBP flux and UDP-GlcNAc availability are affected directly by many different nutrients, such as glucose, fatty acids, and amino acids. Exogenously added glucosamine is efficiently ‘saliaged’ by cells, and small amounts can dramatically increase UDP-GlcNAc pools in cells. When glucose enters cells, besides being used for glycogen synthesis or glycolysis, 2–5% is converted to UDP-GlcNAc [33]. Free fatty acids can increase HBP flux by inhibiting glycolysis, resulting in elevated fructose-6-phosphate levels. Acetyl-CoA, produced by fatty acid metabolism, serves as the donor for the acetylation of glucosamine in the formation of UDP-GlcNAc. 1-glutamine: d-fructose-6-phosphate aminotransferase (GFAT) is the key rate-limiting enzyme of the HBP. GFAT converts fructose-6-phosphate to glucosamine-6-phosphate, using glutamine as the nitrogen donor. GFAT is bypassed by glucosamine, which enters directly into the HBP by being phosphorylated to the GlcNAc substrate.

![Fig. 1. Hexosamine biosynthetic pathway (HBP) and O-GlcNAcylation. In addition to glycogen and glycogen synthesis, 2–5% of cellular glucose is processed through HBP. GFAT mediated conversion of fructose-6-phosphate to glucosamine-6-phosphate is the rate-limiting step of HBP. The presence of elevated FFA has been linked to inhibition of glycolysis and elevated glucose flux through HBP. UDP-GlcNAc is the end product of the HBP. OGT catalyzes the transfer of O-GlcNAc from UDP-GlcNAc onto the hydroxyl moiety of serine or threonine residues, while OGA catalyzes the removal of the sugar. Addition of glutamine or glucosamine elevates global O-GlcNAcylation. OGT = uridine diphospho-N-acetylglucosamine:polypeptide β-N-acetylglucosaminyl transferase, OGA = β-D-N-acetylglucosaminidase, UDP-GlcNAc = uridine diphospho-N-acetylglucosamine, GFAT = L-glutamine: D-fructose-6-phosphate aminotransferase, and FFA = free fatty acid.](image-url)
form glucosamine-6-phosphate [34]. Ob/ob mice, which lack of leptin, a hormone that regulates nutrient uptake, have elevated GFAT activity and increased UDP-GlcNac levels in muscle. UDP-GlcNac levels are increased in tissues by hyperglycemia, hyperlipidemia, as well as by glucosamine [35,36]. The substrate specificity of OGT, as well as its enzymatic activity are regulated by different levels of UDP-GlcNac [15]. In hyperglycemia, elevated HBP flux cause increased O-GlcNAcylation on specific insulin signaling molecules, such as Akt, insulin receptor substrate 1 (IRS1), and others, to negatively regulate the insulin pathway. Excessive O-GlcNAcylation of insulin signaling molecules results in less insulin-stimulated glucose transport [34]. Hyperglycemia and hyperlipidemia also affect transcription by increasing the O-GlcNAcylation of transcription factors, such as Sp1. For example, the saturated fatty acid, palmitate, which causes increased HBP flux, increases Sp1’s DNA binding activity, as well as its transcription activity by elevating its O-GlcNAcylation state [37].

O-GlcNAcylation is also a sensor of cellular stress. Multiple forms of stress increase glucose uptake and thus HBP flux and concomitantly increase UDP-GlcNac concentrations, leading to rapidly increased global O-GlcNAcylation. Many different stresses, such as oxidative, osmotic, ultraviolet light, ethanol, and heat shock induce rapid and global increased O-GlcNAcylation. Elevated O-GlcNAcylation induced by thermal stress occurs rapidly and returns back to basal levels after removing the stress. Thermal stress increases OGT’s enzymatic activity but does not increase its expression. In contrast, OGT expression is increased by ethanol treatment or by osmotic stress. Ogt knockout through Cre-LoxP in mouse embryonic fibroblasts cells or knockdown of Ogt by RNA interference (RNAi) in mouse neuroblastoma cells (Neuro-2a) both decreased thermotolerance. Artificially increasing cellular O-GlcNAc levels by pharmacologically inhibiting OGA activity, using O-(2-acetamido-2-deoxy-o-glucopyranosylidene)amino-N-phenylcarbamate (PUGNAC), resulted in increased thermotolerance. This protective effect of O-GlcNAcylation on thermal stress is partially explained by elevating heat shock protein 70 (Hsp70) and Hsp40 expression [38], chaperone complex of these two proteins is well known as a heat protector in mammalian cells [39]. Hsp70, which protects cells from stress, is also an endogenous O-GlcNAc binding lectin [40].

Paradoxically, despite lower UDP-GlcNac and ATP levels, global O-GlcNAcylation is strikingly increased when neuronal cells are deprived of glucose. When Neuro-2a cells are deprived of glucose, increased OGT mRNA and protein expression are observed. Increased targeting of OGT to a population of substrate proteins, rather than an increase in its catalytic activity was found. This increase of OGT towards glucose-deprivation dependent population of proteins in Neuro-2a cells was dependent upon AMP-activated protein kinase (AMPK) and mitogen activated protein kinase (MAPK)/p38 [13]. However, in human hepatocellular carcinoma cells (HepG2), glucose-deprivation dependent elevation of O-GlcNacylation results mainly from increased OGT mRNA and protein expression, as well as decreased OGA expression, but does not appear to be dependent upon specific kinases [41,42]. Glucose deprivation increases O-GlcNAcylation in human non-small cell lung carcinoma A549 cells, by yet another mechanism. In A549 cancer cells, glucose-deprivation results in increased HBP flux, as well as increased UDP-GlcNac levels. The source of glucose driving the increased HBP flux is from increased glycogen phosphorylase activity, the rate-limiting step in mobilizing glycogen. In A549 cancer cells glucose-deprivation also results in increased GFAT activity. In addition, glucose-deprivation in A549 cells results in higher OGT enzymatic activity and lower enzymatic activity for OGA without changing OGT and OGA expression levels [43]. The different mechanisms elevating O-GlcNAcylation in response to glucose-deprivation may result from the distinct difference in regulating energy metabolism among these cell types. Thus, there are complex and intimate relationships between nutrient status, stress and O-GlcNAcylation. O-GlcNAcylation not only responds to nutrient status and stress, but also regulates cell’s response to variations in nutrients and stress.

3. Detection and site mapping of O-GlcNac

O-GlcNAcylation can be detected by using pan-specific anti-O-GlcNac antibodies (CTD 110.6, RL2, HGAC 85) or lectins (succeylnated Wheat-Germ Agglutinin, sWGA). CTD 110.6 is an IgM monoclonal antibody produced using a chemically synthesised O-GlcNAcylated peptide based upon a single repeat of the RNAP II carboxyl terminal domain (CTD) motif as the antigen [34]. RL2 is an IgG monoclonal antibody produced using nuclear pore protein complex fractions as the antigen. HGAC 85 is an IgG monoclonal antibody made against streptococcal group A polysaccharide, which recognizes O-GlcNac [44]. All of these antibodies favor the detection of highly abundant O-GlcNAcylated proteins. While mostly recognizing O-GlcNac, these antibodies differ in their dependency upon peptide structure and bind different subsets of O-GlcNAcylated proteins [45,46]. Recently, several O-GlcNac peptide specific monoclonal antibodies have been described, which may prove valuable in further defining the biological roles of O-GlcNAcylation [47]. The lectin, sWGA, is able to recognize any terminal GlcNAc, including those on N-glycans, O-glycans or O-GlcNac, and has a preference for clustered O-GlcNAcylation. Regardless of which probe is used, additional controls, such as inhibition of binding by excess GlcNAc, prior treatment with hexosaminidase to remove terminal GlcNAc in N-glycans, or ‘capping’ the epitope with galactosyltransf erase should be used [34] to validate specificity.

As with any PTM, mapping the attachment site(s) is a prerequisite toward understanding the modification’s biological functions. Site mapping O-GlcNAcylation is very difficult due to its special characteristics: (1) O-GlcNAcylation usually has low stoichiometry at each site on proteins. For example, stoichiometry of O-GlcNAcylated α-crystallin, a structural protein in the eye lens, is less than 2% [48,49]. (2) In mass spectrometry (MS), the presence of even a small proportion of unmodified peptide ions severely suppresses ionization of O-GlcNAcylated peptides. (3) The β-linkage between O-GlcNac moiety and Ser/Thr (Thr) is labile. Upon fragmentation by standard collision induced dissociation (CID) or collision activated dissociation (CAD), the glycosidic bond breaks first, resulting in poor fragmentation of the glycopeptide and loss of any site information. With the development of sample enrichment methods and new mass spectrometry fragmentation methods, such as electron capture dissociation (ECD) and electron transfer dissociation (ETD), it is now possible to map and quantify both O-GlcNac and phosphosites simultaneously in complicated samples using relatively limited amounts of protein [50].

3.1. Early approaches to map O-GlcNac sites

In early studies, mapping O-GlcNac sites was a tedious and time-consuming process, often taking nearly one-year. In these early studies, O-GlcNac site mapping was achieved by first tagging the O-GlcNac protein or peptide by enzymatically attaching a [1H5]-galactose, using bovine milk β-1,4-galactosyltransferase (GalT1) [1]. Multiple rounds of reverse phase high performance liquid chromatography (RP-HPLC) of tryptic peptides were then used to obtain pure radiolabeled glycopeptides. Automated Edman degradation was used to sequence the peptide, and followed by manual Edman degradation to determine which residue contained the
radioactivity [51]. Several O-GlcNAc sites were mapped in this way, such as Thr48 and Thr431 on neurofilament M (NF-M), Thr21 and Ser27 on NF-L [52,53]. Sites on estrogen receptor (ER), as well as synapsin I were also mapped via this method [54–56]. However, this method had many problems: (1) It contains too many steps and requires expensive radioactive sugar nucleotide. (2) Many rounds of HPLC and Edman degradation are needed to definitively identify the major O-GlcNAc sites on a single protein. (3) The overall procedure is time consuming and difficult. (4) Large amounts of starting sample are required. For example, in order to map O-GlcNAc sites, at least 10 pmol of pure, radiolabeled glycopeptide is required. Due to sample loss, low stoichiometry of O-GlcNAcylation as well as limits in detecting [3H]-galactose, typically nanomole to even micromole amounts of starting protein are often required.

3.2. Mapping O-GlcNAc sites on recombinant or in vitro expressed proteins

It often is difficult to obtain enough of a low abundance endogenous regulatory protein, such as a transcription factor, to map O-GlcNAc sites, even with the best methods currently available. Three approaches have proven useful, at least as a starting point: (1) For low abundance proteins where DNA is available, the protein can be transcribed and translated in a reticulocyte lysate system, which efficiently O-GlcNAcylates nascent proteins [57]. Lectin affinity chromatography on SWGA is then used to detect O-GlcNAcylation [56,58,59]. (2) Overexpression of the protein in Escherichia coli, which does not O-GlcNAcylate proteins, allows for production of a polypeptide substrate, that can be O-GlcNAcylated in vitro using recombinent OGT later. This approach allows for site mapping, which can readily be confirmed in vivo by site-directed mutagenesis, combined with anti-O-GlcNAc Western blotting. (3) Fairly large quantities of in vivo O-GlcNAcylated proteins can be obtained by overexpressing affinity-tagged proteins (eg, His, Flag, or GST). This can be done in either insect cells via baculovirus infection [54,56,60–62] or transfection in HEK293 mammalian cells [63–65]. Thus far, sites mapped to be O-GlcNAcylated in these model systems are also found existing in other cell types. However, putative O-GlcNAc sites need to be directly confirmed in the cell type studied, usually via site-directed mutagenesis and/or by MS.

3.3. O-GlcNAc enrichment methods

Using large-scale protein preparations, while mitigating the low stoichiometry of O-GlcNAc, increases the risk of ion-suppression by non-glycosylated peptides. Therefore, it is almost always necessary to purify O-GlcNAcylated proteins or peptides away from the unmodified peptides.

3.3.1. Immunoaffinity/lectin based enrichment methods

The simplest O-GlcNAc enrichment approach is to purify O-GlcNAcylated proteins using immunoaffinity or lectin chromatography. O-GlcNAcylated proteins from cell lysates are efficiently enriched by both anti-O-GlcNAc antibody affinity chromatography or by SWGA affinity chromatography [66–68]. However, both methods favor enrichment of high abundance proteins (e.g. nuclear pore proteins) or those with multiple clustered O-GlcNAc residues. Low abundance proteins with single or widely separated O-GlcNAc sites are difficult to purify using these tools, because their binding to the affinity solid support is often competed by more abundant or higher affinity O-GlcNAcylated proteins.

3.3.2. Metabolic in vivo tagging of O-GlcNAc for enrichment (tagging-via-substrate (TAS))

TAS based enrichment method takes advantage of the fact that the cellular machinery will take up and use N-azidoacetylgalucos-amine (GlcNAz) as an alternative substrate to GlcNAc. In order to facilitate uptake by living cells, GlcNAz is peracetylated. Upon uptake, the peracetylated GlcNAz is rapidly deacetylated by intracellular esterases. GlcNAz enters the HBP just like glucosamine, but at lower efficiency, and is converted into UDP-N-azidoacetylglucosamine (UDP-GlcNAz). UDP-GlcNAz is utilized by OGT, and O-GlcNAz is attached to Ser(Thr) sites. The GlcNAz on proteins enables conjugation via Staudinger ligation to enrichment tags, such as a biotinylated phosphine [69]. After ligation, the labeled sample can be selectively purified using avidin/streptavidin beads. By using TAS, 199 putative O-GlcNAcylated proteins were identified in Hela cells (human cervical cancer cell line) through nano-HPLC-MS/MS [70]. The high affinity between biotin and avidin/streptavidin allows for a high degree of enrichment by permitting stringent washes to eliminate nonspecifically bound materials, therefore minimizing contamination from unmodified peptides, decreasing ion suppression effect. The major disadvantage of this approach is the poor utilization efficiency of GlcNAz by cells. All of the enzymes in the cellular machinery use their natural substrates much more efficiently. In addition, high concentrations of endogenous UDP-GlcNAc that are present in most cells effectively competes for utilization of exogenous UDP-GlcNAz. Furthermore, the lack of sensitivity is further exacerbated by the often low stoichiometry of O-GlcNAcylation. Unfortunately, lowering endogenous UDP-GlcNAc levels to increase efficiency of GlcNAz incorporation will more or less affect the biological process of interest. Nonetheless, the TAS approach is useful for studying O-GlcNAcylation on abundant proteins and perhaps on those where the saccharide is most rapidly cycling.

3.3.3. β-Elimination followed by Michael Addition-based enrichment

The O-β-GlcNAc glycosidic linkage is much more alkali labile than O-α-GalNAc or O-phosphate. Since O-GlcNAc is very sensitive to β-elimination by alkali, and is very labile in the gas phase upon CID fragmentation, early site mapping studies used mild β-elimination of O-GlcNAc peptides and the resulting loss of water at the Ser(Thr) attachment site (18 m/z loss) to map sites [71]. Later, this method was improved by using Michael Addition chemistry to attach dithiothreitol (DTT) to the β-eliminated O-GlcNAc sites (m/z increase of 136). This method is called, β-elimination followed by Michael Addition with DTT (BEMAD) [66]. BEMAD provides several advantages: (1) DTT allows for affinity enrichment of modified peptides using thiol-affinity chromatography. (2) The attached DTT moiety is stable under standard MS/MS conditions, such as CID, and modified peptides fragment normally. (3) Deuterium containing DTT is readily available allowing for comparative semi-quantitative mass spectrometry [66,72,73]. The biggest disadvantage of the BEMAD method for site mapping is that it is indirect. In theory, alkali-induced elimination of any moiety at a Ser(Thr) residue could result in a false positive assignment as an O-GlcNAc site. Therefore, the use of BEMAD for site mapping requires many careful controls [72,74,75], and confirmation by independent methods.

3.3.4. Chemoenzymatic enrichment of O-GlcNAc peptides

Most recently, chemoenzymatic approaches have been used to tag and enrich O-GlcNAcylated peptides. A commercially available kit (“Click it™ kit, Invitrogen) has simplified the use of this approach. The method uses a genetically engineered (Y289L mutation) GaIT that has a larger donor-substrate binding pocket. The mutant GaIT can efficiently transfer N-azidoacetylgalactosamine (GalNAz) to the O-GlcNAc moiety [76]. Biotin is then quantitatively attached to GalNAz via click chemistry using a copper catalyst [77]. The labeled proteins/peptides can then be enriched through avidin/streptavidin affinity chromatography. Even though this tagging and enrichment method greatly enriches the O-GlcNAcylated
proteins/peptides, the enriched sample is still not suitable for site mapping by direct CID/CAD MS/MS for the following reasons: (1) The large mass of the tag attached to the O-GlcNAc results in even more severe ion suppression compared with non-enriched O-GlcNAcylated peptides. (2) The biotin moiety fragments excessively during MS/MS, complicating interpretation of the spectra. (3) The saccharide moiety bearing the GalNAz-biotin is still very labile during MS/MS fragmentation, preventing direct site mapping. The combined use of BEMAD and the chemoenzymatic enrichment does allow for direct site mapping and the high specificity of the enrichment gives a much higher confidence in the validity of sites identified. The combined approaches have been used to map O-GlcNAc sites on vimentin and calcium/calmodulin-dependent kinase IV (CaMKIV) [65,67]. A complex mixture of human erythrocyte proteins was also analysed using this chemoenzymatic enrichment followed by BEMAD, resulting in the identification of 35 O-GlcNAc sites [73]. A similar enrichment method that alternatively uses an UDP-ketogalactose and hydrazide coupling to biotin also works well, and has been used in cell lysates and brain tissue [78,79].

3.3.5. Photocleavable biotin (PC-biotin) based enrichment of O-GlcNAc peptides
A major problem with enriching O-GlcNAc using biotin-avidin based affinity methods is the difficulty of getting quantitative recovery of the bound glycopeptides. Thus, building upon Olegnik’s report of a PC-biotin reagent [80], a PC-biotin-alkyne reagent suitable for use in O-GlcNAc site mapping was developed [50]. The same chemoenzymatic attachment method (described above) is used, but the tagged glycopeptides are qualitatively released from the avidin/streptavidin columns by a cleavage reaction catalyzed by ultraviolet light (365 nm). A major advantage of this method, in addition to its efficiency, is that the released peptides now carry a basic aminomethyltriazolyl acetylglactosamine group, which results in a +3 or higher charge state for the tryptic O-GlcNAc peptides. The charged moiety remaining after ultraviolet light cleavage not only facilitates ETD MS, but also produces diagnostic fragment ions (m/z of 300.2 and 503.1) allowing detection of O-GlcNAc peptides in complex mixtures by standard CID/CAD MS [50].

3.4. Advances in MS that facilitate analyses of O-GlcNAc

The enrichment methods, described above, have partially solved the analysis problems associated with the low stoichiometry and the ion suppression by unmodified peptides. In addition, increased sensitivity of modern mass spectrometers further aid in O-GlcNAc site mapping.

3.4.1. CID/CAD

As noted above, O-GlcNAc attachment to Ser(Thr) is labile. In CID/CAD, the energy required to break the O-GlcNAc linkage is much lower than that required to break the peptide bond. CID/CAD almost always results in a neutral loss of the saccharide, which can indicate the existence of O-GlcNAcylation on the peptide without any site information [81–83]. In addition, due to the loss of kinetic energy associated with the neutral loss of the saccharide, CID/CAD is generally very poor at producing enough peptide fragmentation of the O-GlcNAc peptide for sequencing. However, in ion-trap mass spectrometers it is possible to perform an MS<sup>3</sup> CID/CAD and obtain good quality peptide sequence, but only after the sugar is lost. In contrast, DTT attached by the BEMAD procedure, is not lost during CID/CAD and the DTT-peptides fragment as well as unmodified peptides. A combination of BEMAD and CID/CAD, has been used to map sites on synapsin I, nuclear pore complex proteins, actin, myosin, and IRS-1 [84,85].

3.4.2. ECD/ETD

ECD uses low energy electrons to react with peptide cations in the magnetic field of a Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer [86]. Unlike CID/CAD mass spectrometry, the O-GlcNAc linkage is not broken by ECD of peptide ions [68]. However, the high cost and high level of expertise required to operate and maintain FT-ICR mass spectrometers limits their widespread use. Recently, an ionization method similar to ECD, ETD has been developed which works well in lower cost ion trap mass spectrometers [87,88]. During ETD an electron is transferred in the gas phase from a radical anion to a protonated peptide to induce cleavage of the C=N bond. Like ECD, ETD preserves the O-GlcNAc and other peptide modifications, while efficiently fragmenting the peptide. The development of ETD is a breakthrough in the detection and analysis of O-GlcNAcylation. ETD mass spectrometry has recently been used to map O-GlcNAc on paxillin, a focal adhesion adaptor protein, IRS-1 and Forkhead box O1 (FoxO1), and peroxisome proliferator activated receptor γ coactivator-1α (PGC-1α) [14,64,89,90]. By combining ETD with sWGA chromatography enrichment, more than 50 previously unknown O-GlcNAc sites on murine postsynaptic density proteins were mapped [91]. Chemoenzymatic enrichment using a PC-biotin tag, combined with ETD MS/MS, enabled the mapping of over 140 O-GlcNAc sites on proteins involved in the regulation of cell division [92].

3.5. Quantitative analysis of O-GlcNAc site occupancy

3.5.1. Stable isotope labeling with amino acids in cell culture (SILAC)
In order to quantitatively compare O-GlcNAc levels on different proteins or at different sites, SILAC [93,94], when combined with the MS methods discussed above, has proven to be a method of choice. SILAC was used to test quantitative O-GlcNAcylation changes in COS7 cells after only a single kinase, glycogen synthase kinase-3 (GSK-3), was inhibited [92]. Major disadvantages of SILAC include the high cost of isotope containing amino acids, and that the metabolic incorporation of the labeled amino acids requires living samples, either in culture or at the whole organism level.

3.5.2. In vitro methods for isotopically labeling O-GlcNAc peptides
Recently, two similar in vitro methods have been used for quantitative analysis of O-GlcNAcylation by isotopically labeling peptides with different tags [73,95]. One approach is “tandem mass tags” (TMT) and the other is “isobaric tags for relative and absolute quantification” (iTRAQ) [96,97]. Peptides are labeled with different tags, which upon tandem mass spectrometry are cleaved into different m/z reporter ions. Each tag introduces a unique mass shift for the labeled peptides. The advantage of these tags is that they allow for multiplexing, and peptides obtained from tissues or other sources can be efficiently labeled in vitro. A similar peptide labeling method makes use of reductive amination to incorporate stable isotopes into N-terminal amines and e-amino groups of lysine residues [98]. Using this approach, a quantitative isotopic and chemoenzymatic tagging (QUIC-tag) method was developed for O-GlcNAc analysis. In this method, O-GlcNAcylated peptides are first enriched by the modified chemoenzymatic enrichment method (discussed above), and then the biotin labeled O-GlcNAcylated peptides are dimethylly isotopically labeled. By using the QUIC-tag approach, many quantitative O-GlcNAc changes on proteins after inhibition of O-GlcNAcase, as well as after different stimulations of the brain were identified [99].

3.5.3. β-Elimination Michael Addition with density tagged DTT
Using chemoenzymatic enrichment (as above) combined with BEMAD, but using deuterated DTT (DTT-[D<sub>6]</sub>) for one sample and light DTT for the other samples, allows for direct comparison of O-GlcNAc site occupancy, as long as the necessary controls are
performed. For example, this approach was used effectively to compare O-GlcNAc site occupancy ratio between sites on proteins from erythrocytes between normal and diabetic patients [73]. The advantage of this approach is that it can be performed on inexpensive ion trap mass spectrometers using CID/CAD tandem mass spectrometry. The disadvantage (as discussed above) is that detection of O-GlcNAc is indirect.

With the development of these new methods, the pace of identification and site mapping of O-GlcNAcylated proteins is rapidly increasing. In fact, the total O-GlcNAc sites mapped within the past two years exceeds that known during the past 20 years [1]. Fig. 2 summarizes the current approaches available for identification and site-mapping O-GlcNAc on proteins.

4. Extensive crosstalk exists between phosphorylation and O-GlcNAcylation

O-GlcNAcylation occurs on Ser(Thr) side chains of proteins at sites that may also be phosphorylated. Thus far, virtually every O-GlcNAcylated protein can also be phosphorylated [100]. After several O-GlcNAc sites were mapped and discovered to also be phosphorylation sites, a “Yin–Yang” model was proposed for the relationship between O-GlcNAcylation and phosphorylation [101]. This model suggested that O-GlcNAcylation is reciprocal to phosphorylation with mutually exclusive occupancy at the same site on a polypeptide. While largely correct for many proteins, the “Yin–Yang” model is an oversimplification. Now hundreds of O-GlcNAc sites have been mapped, and some cellular stimuli were shown to increase both modifications, it is clear that the interplay between O-GlcNAcylating and phosphorylation is both complex and very extensive.

4.1. Global interactions between O-GlcNAcylation and phosphorylation

4.1.1. Global interactions between O-GlcNAcylation and phosphorylation may be antagonistic or synergistic

Reciprocal relationships between O-GlcNAcylating and phosphorylation have indeed been found in several global analyses. Global O-GlcNAc level changes were detected after using various kinase activators and inhibitors. For example, activation of protein kinase A (PKA) or protein kinase C decreased global O-GlcNAcylation and inhibitors of these kinases increased global O-GlcNAcylation [102]. Okadaic acid (OA), a broad-spectrum phosphatase inhibitor, has been used to explore the global interplay between O-GlcNAcylating and phosphorylation. Indeed, in some cells, short-term OA treatment decreases O-GlcNAcylating [103]. However, since OA also induces a stress response that increases O-GlcNAcylation of many proteins by other mechanisms, these types of global treatments are not very informative.

In hippocampal neuronal synapses, pharmacological elevation of O-GlcNAc in vivo, concomitantly increases phosphorylation of synapsin I/II at serine6 (PKA site), serine62/67 (Erk1/2 site), and serine608 (CaMK II site) [104]. These and other studies indicate that the nature and direction of the crosstalk between these two post-translational modifications depends upon not only the protein but also the specific modification site.

4.1.2. Proteomic studies of global interactions between O-GlcNAcylation and phosphorylation

Chemoenzymatic enrichment methods and advances in mass spectrometry have allowed global comparisons of O-GlcNAcylating and phosphorylation at the individual site level. In studies to determine the effects of inhibiting GSK-3 on O-GlcNAcylation, of the 45 identified O-GlcNAcylated proteins, at least 10 proteins (mostly cytoskeletal proteins) had increased O-GlcNAcylation, while 19 other proteins (mostly RNA binding proteins) had decreased O-GlcNAcylation [67]. A recent study examined the affects of globally increasing O-GlcNAcylation, by inhibiting OGA, on the quantitative site occupancy of phosphate at over 700 phosphorylation sites in non-stimulated NIH3T3 cells [105]. Forty eight percentage of the phospho-sites were not actively cycling in these non-stimulated cells, as judged by insensitivity to OA. However, of the remaining actively cycling phospho-sites nearly all had their level of phosphorylation affected by altered O-GlcNAcylating. Elevated O-GlcNAcylating significantly reduced phosphorylation at about 280 sites and significantly increased phosphorylation at 149 sites. These studies revealed the surprisingly extensive level of crosstalk between these two most abundant PTMs. This extensive crosstalk not only results from competition at the same or proximal attachment sites, but also it is now clear that phosphorylation regulates OGT activity. In addition, O-GlcNAc is being found to regulate an increasing number of kinases [65,106–108].

Overexpression of OGT causes defective cytokinesis increasing polyplody in cells, a feature common to many cancer cells [109]. In order to understand the molecular basis for OGT’s induction of...
defective cell division, an O-GlcNAc, phospho-SILAC proteomic analysis, using chemoenzymatic enrichment with PC-biotin and a combination of CID and ETD MS/MS was carried out on spindles/ midbody preparations isolated from HeLa cells either overexpressing GFP (control) or OGT (experimental) [92]. A comparable number of phosphorylation and O-GlcNAcylation sites were mapped (350 phosphorylation sites on 190 proteins vs. 141 O-GlcNAcylated peptides on 64 proteins). Seventeen percentage of the phospho-
sites showed more than a 50% decrease upon a twofold over-
expression of OGT and the resulting increase in global O-GlcNAc-
lyation. Some (7%) phosphorylation sites had a more than 200% increase in site occupancy. These changes resulted in pronounced
inhibition of the cyclin dependent kinase 1 signaling pathway and alterations in polo kinase as well as aurora kinase signaling, which partially explained the cytokinesis defects induced by OGT overex-
pression. These data clearly exemplify the importance of the inter-
play between O-GlcNAcylation and phosphorylation in cellular
signaling pathways.

4.2. Physiological significance of the crosstalk between
O-GlcNAcylation and phosphorylation

4.2.1. Crosstalk between O-GlcNAcylation and phosphorylation at the
protein level

Hyperphosphorylation of the microtubule bundling protein, tau, is a characteristic of Alzheimer’s disease (AD). It has also been
known for some time that tau is heavily O-GlcNAcylated in normal
brain [110]. These findings led to the proposal that defective glu-
cose metabolism in the brain results in lower O-GlcNAcylation of
brain proteins, including tau, leading to abnormally increased
hyperphosphorylation associated with AD [111]. Hyperphospho-
ylated tau forms visible tangles within neurons, and this is re-
ferred to as paired helical filamentous tau (PHF-tau). To date,
more than 30 phosphorylation sites have been mapped on PHF-
tau, and many of them have been used to produce site-specific
phosphorylation antibodies [112]. While there are more than 12
putative O-GlcNAc sites on tau, only a few sites have been mapped
and no site specific antibodies have been made [110]. Increased O-
GlcNAcylation negatively regulates tau phosphorylation at multi-
ple sites, and hyperphosphorylated tau has decreased O-GlcNAc-
lyation. Several studies have shown that decreased flux through the
HBP or decreased expression of OGT result in tau hyperphosphory-
lation in animal models [113,114]. The potent OGA inhibitor, Thia-
met G (TMG), increases global O-GlcNAcylation effectively in
whole animals and crosses the blood–brain barrier [115]. TMG
treatment increased O-GlcNAcylation and decreased phosphoryla-
tion at pathologically relevant sites on tau, Thr231 and Ser196 in
PC12 cells and Ser422 in rat cortex and hippocampus [115].

Upon activation of the insulin signal cascade, phosphorylation of
AKT at Thr308 is increased. Concomitantly, phosphorylations on
IRS-1 at Ser207 and Ser322/323 are decreased. When OGT is overex-
pressed in mouse liver, increased O-GlcNAcylation of both AKT
expression. These data clearly exemplify the importance of the inter-
play between O-GlcNAcylation and phosphorylation in cellular
signaling pathways.

4.2.2. Reciprocal crosstalk of O-GlcNAcylation and phosphorylation
at same site

Despite many exceptions, there are also numerous proteins, such as c-Myc, (ER-β) and Akt [106], in which the O-GlcNAc moiety and the phosphate residue compete for the same hydroxyl group, as
proposed in the “Yin–Yang” model. One of the best-studied examples is the c-Myc oncogene protein. The protooncogene, c-myc, regulates transcription of genes involved in cell prolifera-
tion, apoptosis and metabolism [119]. The major O-GlcNAcylation
site on c-Myc was mapped to Thr58, which is also a GSK-3β phos-
phorylation site within the transaction domain [61]. Frequent
mutation of this site is found in human lymphomas, demonstrating
the importance of this site in tumor progression [119]. Antibodies
specifically detecting O-GlcNAcylated Thr458, phosphorylated Thr58,
as well as Thr58-unmodified c-Myc were used to study the respective
roles of each modification. Serum starvation stops cell growth and
greatly increased Thr58 O-GlcNAc. In contrast, when
were stimulated to grow by serum addition, O-GlcNAcylation
at Thr58 very rapidly disappears and is replaced by increased Thr58
phosphorylation by GSK-3β. The proximal site, Ser62, is a priming
phosphorylation site and a substrate for MAPK. Phosphorylation at
Ser47 is necessary for Thr58 phosphorylation by GSK-3β. When
Ser47 is mutated to alanine, not only does Thr58 no longer become
phosphorylated, but also the extent of O-GlcNAcyllation at Thr58
increases dramatically. Lithium chloride, a GSK-3β inhibitor, also
increased O-GlcNAcylation at Thr58 similar to the Ser62-Ala mutation
[120].

ER-β is reciprocally O-GlcNAcylated/phosphorylated at Ser16
[56,121]. Ser16 is within a region of ER-
which is related to rapid protein degradation. Site-directed mutagenesis comparisons between wild type ER-β to mutants, S16A, S16E suggested that the O-GlcNAcylated ER-β was much more stable, but less transcriptionally active, while the phos-
phorylated ER-β was transcriptionally more active, but rapidly de-
graded [122]. Consistent with the biological studies, NMR analysis
and molecular dynamics simulation of synthesised peptides, sug-
gested that phosphorylation in Ser16 discourages β-turn formation
at the N-terminus of the protein, resulting in a more extended
structure, while Ser16 O-GlcNAcylated peptides promoted β-turn
formation, adopting a more stable structure, which to some degree
might explain why O-GlcNAcylation on Ser16 can stabilize the
protein [123].

Endothelial nitric oxide synthase (eNOS) function is also regu-
lated by reciprocal O-GlcNAcylation/phosphorylation crosstalk. Ser177
phosphorylation of eNOS by AKT activates the enzyme and is important in positively regulating penile erection, and phos-
phorylation on Thr495 of eNOS acts as a negative regulator. Ser615
and Ser633 phosphorylation can also activate eNOS. In diabetic rats,
excessive O-GlcNAcylated Ser177 prevents AKT activation of
eNOS without affecting phosphorylation at Thr495, Ser615 and Ser633
[124], and in humans with diabetes O-GlcNAcylation of eNOS
contributes to erectile dysfunction [124].

Another example of direct reciprocity is β-catenin kinase β (IKKβ).
IKKβ is O-GlcNAcylated at Ser273 in its α-c-terminus, which is also an
inactivating phosphorylation site. The reciprocal interaction in this site
is responsible for regulating IKKβ catalytic activity [108].

4.2.3. Crosstalk of O-GlcNAcylation and phosphorylation at proximal sites

Perhaps more common than direct site competition is the inter-
play that occurs between O-GlcNAcylation and phosphorylation at proximal sites. It should be noted that the Stoke’s radius of an
O-GlcNAc is several times larger than that of a phosphate moiety

4.2.4. O-GlcNAcylation and phosphorylation at the nucleotide-binding site

O-GlcNAc decreases the affinity of eNOS for ATP, while phospho-
ychosphorylation increases the affinity of eNOS for ATP. This change in affinity
leads to decreased activity of eNOS

4.2.5. Crosstalk of O-GlcNAcylation and phosphorylation at the pocket

In contrast, some studies have shown that O-GlcNAcylated eNOS
is more active than phosphorylated eNOS. This is consistent with
the idea that O-GlcNAcylation may play a role in the regulation of eNOS
activity.

4.2.6. O-GlcNAcylation and phosphorylation at the C-terminal domain

O-GlcNAcylation and phosphorylation at the C-terminal domain
of eNOS is important for the regulation of eNOS activity.

4.2.7. O-GlcNAcylation and phosphorylation at the aminoterminal domain

O-GlcNAcylation and phosphorylation at the aminoterminal domain
of eNOS is important for the regulation of eNOS activity.

4.2.8. O-GlcNAcylation and phosphorylation at the carboxy-terminal domain

O-GlcNAcylation and phosphorylation at the carboxy-terminal domain
of eNOS is important for the regulation of eNOS activity.

4.2.9. O-GlcNAcylation and phosphorylation at the catalytic domain

O-GlcNAcylation and phosphorylation at the catalytic domain
of eNOS is important for the regulation of eNOS activity.

4.2.10. O-GlcNAcylation and phosphorylation at the regulatory domain

O-GlcNAcylation and phosphorylation at the regulatory domain
of eNOS is important for the regulation of eNOS activity.
The important tumor suppressor, p53, bears proximal O-GlcNAc (Ser 149) and phosphate (Thr155) residues. O-GlcNAcylation of Ser149 reduces phosphorylation at Thr155, which subsequently blocks the interaction between Mdm2, an E3 ubiquitin ligase, and p53, a prerequisite for proteasomal degradation of p53. Thus, increased O-GlcNAcylation of p53 at Ser149 results in decreased p53 ubiquitination and stabilizes the protein p53\[125\]. Mutation of the Ser149 O-GlcNAc site on p53 to alanine increases phosphorylation at Thr155.

Three-dimensional structural analysis shows that these two sites are located on a flexible loop with close proximity of only 11.58 Å, suggesting that there might not be enough space for dual occupancy of both phosphate and O-GlcNAc groups at these two sites.

A similar dynamic interplay between O-GlcNAcylation and phosphorylation occurs in the C-terminal domain (CTD) of RNAP II. The CTD of RNAP II is extensively phosphorylated during the transition from the preinitiation complex to the assembly of the elongation complex during transcription. On a subset of the RNAP II, the CTD contains no phosphate, but is extensively O-GlcNAcylated at Thr1618 \[126\]. And Thr1616 and Thr1619 are both phosphorylation sites \[127\]. The reciprocal crosstalk of the two modifications on CTD was shown in vitro. OGT was completely unable to O-GlcNAcylate a CTD substrate that contained even a single phosphate moiety. Likewise, the CTD kinase was unable to phosphorylate a CTD containing only a single O-GlcNAc, despite many other sites being available \[128\].

Site mapping studies have recently revealed a similar proximal interplay between O-GlcNAcylation and phosphorylation on CaMKIV, a key kinase regulating gene expression in neurons and other cells (Fig. 3). CaMKIV is predominantly expressed in brain, T-lymphocytes and in the testis. It is a multifunctional calcium/calmodulin-dependent Ser(Thr) kinase, that regulates downstream protein activities according to changes in intracellular Ca\(^{2+}\) concentrations. CaMKIV is O-GlcNAcylated and phosphorylated on several sites. Five O-GlcNAcylation sites were mapped on CaMKIV and O-GlcNAc on Ser189, Thr200/Ser189 were evaluated for their interactions with phosphorylation, mainly at the key regulatory phosphorylation site at Thr200. Phosphorylation at Thr200, which is the target of CaMK kinase increases CaMKIV kinase enzymatic activity. Upon neuronal depolarization using ionomycin to increase intracellular Ca\(^{2+}\), phosphorylation at Thr200 rapidly increases peaking at 2 min, and decreasing after 5 min. In contrast, O-GlcNAcylation at Ser189 reciprocally decreased at 2 min, and returned to basal level at around 10 min. Mutating Ser189 to alanine dramatically increases both the phosphorylation of CaMKIV at Thr200 and its constitutive enzymatic activity. Mutation of Thr200 to alanine inactivates
the kinase. Analysis of the three-dimensional structure of CaMKIV revealed that all three of these O-GlcNAc sites are located within the catalytic cleft, near the ATP binding residues Lys75 and Asp185 (Fig. 3) [65].

4.2.4. Crosstalk between O-GlcNAcylation and phosphorylation at distant sites

In addition to the reciprocal and synergistic crosstalk at same or proximal sites of the proteins discussed above, crosstalk between O-GlcNAcylation and phosphorylation also exists among distantly located sites. IRS1 is an example of this case. Ser(Thr) phosphorylation sites on IRS1 are mostly located in the N-terminus. Phosphorylation on Ser307, Ser612, Ser632/635 negatively regulate insulin signaling, while Ser102, Ser628 (human numbering) have positive regulatory effects on insulin signaling [129]. However, the O-GlcNAcylation sites exist mainly in the C-terminus, such as Ser1009, Ser1036, Ser1041. While crosstalk between specific O-GlcNAcylation and specific phosphorylation among these sites is still unknown, however, as mentioned above, global O-GlcNAcylation on IRS1 has a negative regulation on insulin signal [116,130], therefore on IRS1 the two PTMs are thought to be both reciprocally and synergistically regulated.

In Fig. 4, we summarize different types of crosstalk between O-GlcNAcylation and phosphorylation based upon known modification sites. These can reciprocally or synergistically interact with each other on the same sites or proximal sites, as well as at distant sites within proteins.

4.3. Crosstalk between the cycling enzymes regulating O-GlcNAcylation and phosphorylation

As discussed above, O-GlcNAcylation and phosphorylation certainly have extensive crosstalk due to competition of the modifications themselves at the same or proximal sites on proteins. However, they also influence each other by regulating the activities or localizations of the other's cycling enzymes. For example, OGT is directly activated by tyrosine phosphorylation and is itself O-GlcNAcylated [15], and CaMKIV is directly inhibited by O-GlcNAcylation [65]. Src, a well known tyrosine kinase important to cancer, phosphorylates OGT in vitro [130]. Upon insulin stimulation, OGT is recruited to the plasma membrane by binding phosphatidylinositol 3,4,5-triphosphate (PIP3). The insulin receptor phosphorylates OGT on tyrosine, increasing its activity [116,130]. OGT also forms a stable and active complex with Ser(Thr) protein phosphatase PP1β and PP1γ in rat brain [131]. Additionally, PP1 and the mitotic kinase Aurora B form a complex with OGT and OGA during M phase of the cell cycle. The Aurora B kinase inhibitor, ZM447439 disrupts Aurora B's localization to the midbody during cytokinesis, as well as its co-localization with OGT. The transient complex that forms at the midbody during cytokinesis regulates both the phosphorylation and O-GlcNAcylation of many proteins that regulate cell division, including vimentin [6].

5. How does a single catalytic polypeptide (OGT) specifically modify more than a 1000 different proteins?

As mentioned above, there are 385 known Ser(Thr) kinases in humans with each kinase functioning on its own specific substrates and with its own regulatory signals [132]. However, in mammals there is only one known gene for OGT and only two known in plants. Thus, the site specificity of phosphorylation and O-GlcNAcylation are determined by very different mechanisms. How does this single protein specifically modify so many different sequences within proteins? There is precedence for this problem in biology. For example, there is only a single gene encoding RNAP II's catalytic subunit. How then does RNAP II specifically transcribe in a regulated way from so many different promoters? It is well-documented that RNAP II's activity is determined by its ability to form many transient protein complexes as holoenzymes with precise specificity. Current data, suggests that this mechanism is also how OGT's specificity is determined. That is, in a living cell many different holoenzymes, that are comprised of OGT and other polypeptides, form under specific conditions and the resulting OGT complexes serve to specifically target the catalytic subunit to its target proteins (Fig. 5). In addition, OGT also has specific sequence preference, but an absolute consensus sequence has not been identified.

Early work showed that adaptor or targeting proteins for OGT bind to its TPR repeats to target the enzyme to transcription complexes and other targets [133]. Many transcription factors, such as Sp1, c-myc, are O-GlcNAcylated [61,134]. O-GlcNAcylation of Sp1 inhibits its binding ability with other Sp1 proteins as well as TATA-binding protein-associated factor II 110 (TAFII110) by preventing the hydrophobic interactions between theses proteins [134,135]. Homomultimerization of Sp1 is required to induce

![Fig. 4. Interplay between O-GlcNAcylation and phosphorylation. (A) Single site reciprocal interplay – one serine or threonine (ser/thr) site is modified either by O-GlcNAc or phosphate. (B) Proximal sites reciprocal interplay – two or more immediate or proximal neighboring Ser/Thr sites are reciprocally modified by O-GlcNAc or O-phosphate. (C) Spatially proximal sites with reciprocal interplay – two or more Ser/Thr sites that are distant in primary sequence are brought together and with its own regulatory signals [132]. However, in mammals there is only one known gene for OGT and only two known in plants. Thus, the site specificity of phosphorylation and O-GlcNAcylation are determined by very different mechanisms. How does this single protein specifically modify so many different sequences within proteins? There is precedence for this problem in biology. For example, there is only a single gene encoding RNAP II's catalytic subunit. How then does RNAP II specifically transcribe in a regulated way from so many different promoters? It is well-documented that RNAP II's activity is determined by its ability to form many transient protein complexes as holoenzymes with precise specificity. Current data, suggests that this mechanism is also how OGT's specificity is determined. That is, in a living cell many different holoenzymes, that are comprised of OGT and other polypeptides, form under specific conditions and the resulting OGT complexes serve to specifically target the catalytic subunit to its target proteins (Fig. 5). In addition, OGT also has specific sequence preference, but an absolute consensus sequence has not been identified. Early work showed that adaptor or targeting proteins for OGT bind to its TPR repeats to target the enzyme to transcription complexes and other targets [133]. Many transcription factors, such as Sp1, c-myc, are O-GlcNAcylated [61,134]. O-GlcNAcylation of Sp1 inhibits its binding ability with other Sp1 proteins as well as TATA-binding protein-associated factor II 110 (TAFII110) by preventing the hydrophobic interactions between theses proteins [134,135]. Homomultimerization of Sp1 is required to induce
transcription activation, and interaction with TAFII110 is needed for binding with general transcription factor II D (TFIID), and conducting DNA polymerase II dependent transcription [136]. The role of OGT in transcription repression was later clarified. By tethering OGT to the Gal4 DNA binding domain, it was shown recruitment of OGT to gene promoters enhanced the inhibitory effect in both basal and Sp1-activated transcription. Later analysis showed that OGT was recruited to promoters by another transcription repressor, mSin3A, where HDACs interacted with the complex, resulting in transcription repression. Truncation experiments indicated that the first six TPRs of OGT, which mediates protein-protein interactions, are the main motif interacting with the putative paired amphipathic helix domains of mSin3A [9,133].

Through yeast two-hybrid screens, many putative adaptor/targeting proteins for OGT have been identified. Firstly, two highly homologous proteins were isolated, (gamma-aminobutyric acid) GABA<sub>a</sub> receptor-associated protein (GRIF-1) and OIP106 (now it is known as Trak1). The interaction between OGT and either of these two proteins were confirmed both in vivo and in vitro. Consistent with previous results, the first six TPR domains of OGT are required for interaction with these two adaptor proteins [137]. Confocal and electron microscopy data showed that OIP106 co-localized with RNAP II, which was further confirmed by detection of a complex among OIP106, RNAP II and OGT in HeLa nuclear extracts. This raised the possibility that OIP106 may serve to target OGT to RNAP II. Similarly, GRIF-1 is also thought to target OGT to GABA<sub>a</sub> [138].

A yeast two-hybrid screen using human fetal brain cDNA library, identified 27 putative OGT binding proteins. These proteins belong to a wide range of functional classes, such as cytoskeletal proteins, transcription factors and proteins involved in development, and metabolism. Myosin phosphatase targeting subunit 1 (MYPT1), a known targeting regulatory subunit of PIP<sub>2</sub>; and coactivator-associated arginine methyltransferase 1 (CARM1) were confirmed to interact with OGT directly and affect OGT substrate specificity in vitro. When recombinant MYPT1 or CARM1 were mixed with OGT, in vitro O-GlcNAcylation of several proteins within brain extract was increased. In contrast, in vivo knock down of MYPT1 by RNAi decreased O-GlcNAcylation of many proteins even though the activity of OGT and OGA were not changed. These results clearly document the importance of protein targeting to OGT’s reaction specificity.

Several other adaptor proteins have also been reported in recent years. Ataxin-10 interacts with OGT p110 in brain tissue, and this interaction is competed by the mitochondrial OGT isoform, OGT p78. In vitro OGT activity assay, using PC12 cell extracts, demonstrates that Ataxin-10 increases OGT activity by twofold on certain substrates [139]. These data suggest that Ataxin-10 may function as a potential adaptor protein for OGT in vivo.

This holoenzyme complex hypothesis was also supported by studies of the O-GlcNAcylation of several specific proteins [13,14]. For example, when Neuro-2a cells were deprived of glucose, OGT interacts with the MAPK/p38. Through a large-scale co-immunoprecipitation, NF-H was identified as a substrate for both p38 and OGT. Pharmacologically inhibiting p38 activity abolished O-GlcNAcylation of NF-H induced by glucose-deprivation, even though the interaction of OGT with NF-H was not changed [13].

Another well-documented OGT-adaptor protein-target protein complex formation comes from studies of the O-GlcNAcylation of the transcription factor FoxO1 and its coactivator, PGC-1α. The interaction among these three proteins was confirmed in cultured rat hepatoma Fao cells. In vitro OGT assays showed that PGC-1α increased OGT’s activity towards FoxO1 through increasing the interaction between OGT and FoxO1. High concentrations of glucose increases O-GlcNAcylation of FoxO1 and also its transcriptional activity in FAO cells [90]. This glucose induced increase in O-GlcNAcylation of FoxO1 and transcription activity were also PGC-1α dependent.

The examples above illustrated not only the existence of OGT adaptor/targeting proteins, but also showed that these transient interactions are highly responsive to external signals, such as glucose concentrations. Understanding the mechanisms regulating OGT and OGA targeting to their substrates through adaptor proteins will continue to be a key question for future investigations.

6. Conclusion

O-GlcNAcylation plays an essential role in regulating many cellular processes, such as transcription, cell signaling, metabolism and protein stability (for reviews, Refs. [7,140,141]). Thus it is not surprising that O-GlcNAcylation is involved in various diseases, like cancer, diabetes, as well as neurodegenerative disease and cardiovascular disease (for reviews, Refs. [141–143]).

The development of various rapid, highly efficient technologies for O-GlcNac site mapping, has enabled more rapid advances toward understanding this abundant PTM. Novel detection and site mapping methods have uncovered the extraordinarily extensive crosstalk or interplay between O-GlcNAcylation and phosphorylation. Development of more direct research tools, such as site-specific O-GlcNac antibodies, and increases in our knowledge about the interplay between O-GlcNAcylation and phosphorylation, will lead to a major paradigm shift in our understanding of signal transduction.

Acknowledgements

We wish to acknowledge all the members of Hart Laboratory, especially Quira Zeidan, Dr. Ronald J. Copeland, Dr. Chad Slawson and Dr. Kaoru Sakabe for critical reading of the manuscript. Original research is supported by NIH R01 CA42486 and R01 DK61671 to G.W.H.
References


[33] Liu, R., Dorfman, M. L., Heynh have


