Published in final edited form as:

Curr Signal Transduct Ther. 2010 January ; 5(1): 74–91.

Mechanism, Structure, and Inhibition of O-GlcNAc Processing Enzymes

Tracey M. Gloster^{1,2} and David J. Vocadlo^{1,*}

¹Department of Chemistry, Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada

²York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York, YO10 5YW, UK

Abstract

The post-translational modification of nucleocytoplasmic proteins with O-linked 2-acetamido-2-deoxy-D-glucopyranose (O-GlcNAc) is a topic of considerable interest and attracts a great deal of research effort. O-GlcNAcylation is a dynamic process which can occur multiple times over the lifetime of a protein, sometimes in a reciprocal relationship with phosphorylation. Several hundred proteins, which are involved in a diverse range of cellular processes, have been identified as being modified with the monosaccharide. The control of the O-GlcNAc modification state on different protein targets appears to be important in the aetiology of a number of diseases, including type II diabetes, neurodegenerative diseases and cancer. Two enzymes are responsible for the addition and removal of the O-GlcNAc modification: uridine diphospho-N-acetylglucosamine:polypeptide β -N-acetylglucosaminyltransferase (OGT) and O-GlcNAcase (OGA), respectively. Over the past decade the volume of information known about these two enzymes has increased significantly. In particular, mechanistic studies of OGA, in conjunction with structural studies of bacterial homologues of OGA have stimulated the design of inhibitors and offered a rationale for the binding of certain potent and selective inhibitors. Mechanistic information about OGT lags a little way behind OGA, but the recent deduction of the structure of an OGT bacterial homologue should now drive these studies forward.

Keywords

Post-translational modification; O-GlcNAc; O-GlcNAcase; O-GlcNAc transferase; enzyme mechanism; inhibition; three-dimensional structure

INTRODUCTION

The post-translational modification of proteins with a single 2-acetamido-2-deoxy-D-glucopyranose (GlcNAc) residue was first detected 25 years ago [1]. Despite considerable research efforts in the field during this time, the precise functional roles this simple monosaccharide plays *in vivo* remain a topic of considerable continuing interest. The GlcNAc moiety is attached through a β -glycosidic O-linkage (thus O-GlcNAc) to the serine or threonine residues of proteins. Unlike all other forms of eukaryotic protein glycosylation, which occurs within the secretory pathway, proteins modified with O-GlcNAc are found in

the nucleus or cytoplasm. The modification is ubiquitous through all higher eukaryotes. While disruption of the gene encoding the enzyme responsible for transfer of *O*-GlcNAc showed it was essential for embryonic stem cell viability in mice [2] and more recently for development in *Drosophila melanogaster* [3, 4], loss of the genes encoding the enzymes for either addition or removal of the monosaccharide in *Caenorhabditis elegans* results in growth defects [5, 6].

The extent of the protein targets that get *O*-GlcNAc modified is most likely not yet fully established, but to date several hundred have been identified [7-9]. The protein targets have widespread functions in cells, and include classes of proteins such as transcription factors (such as Sp1, p53 and c-myc [10, 11]), cytoskeletal proteins (such as talin, α -tubulin and the neurofilament proteins [12-15]), nuclear pore proteins (such as p62 and p180 [16-18]), chaperones (such as HSP70 [12]), as well as enzymes such as RNA polymerase II [19] and the proteasome [20]. The control of the *O*-GlcNAc modified state of different protein targets appears important in the aetiology of a number of diseases, including type II diabetes, neurodegenerative diseases, and cancer. Control of global *O*-GlcNAc levels is believed to play a role in mediating insulin resistance [21-23], a key feature of type II diabetes, although recently this widely held view has been questioned [24-26] suggesting that more research in this area may clarify the processes involved. The balance between *O*-GlcNAcylation and phosphorylation on tau is implicated in the development of Alzheimer's disease; tau normally stabilizes microtubules but it is thought that its hyperphosphorylation leads to its aggregation into paired helical filaments which, in turn, aggregate to form the neurofibrillary tangles that constitute one half of the pathology of this disease [27]. Control of *O*-GlcNAc levels may also be aberrant in cancer, since, for example, the tumour suppressor protein p53 and protooncogene product c-myc are modified [11, 28].

The *O*-GlcNAc modification is dynamic, with cycling of the addition and removal of the monosaccharide occurring several times over the lifetime of a target protein [29]. In addition, *O*-GlcNAcylation has some interplay with phosphorylation (Fig. 1) [30]; the modifications have been observed to occur independently on the same serine or threonine residue of a protein (such as c-myc [31]) or competitively at adjacent or nearby residues (such as p53 [28]), which may allow control of cellular signalling. Although to date only a limited number of *O*-GlcNAc sites have been mapped, there appears to be no consensus sequence, unlike *N*-linked glycosylation. It has been noted, however, that Pro-Val-Ser is present in a significant portion of the sites mapped, which is a sequence similar to that recognized by the proline-directed kinases [8, 32, 33].

The cycling of the *O*-GlcNAc modification is modulated by two enzymes; uridine diphospho-*N*-acetylglucosamine: polypeptide β -*N*-acetylglucosaminyltransferase (OGT) [34, 35] is responsible for addition of the monosaccharide and a β -*N*-acetylglucosaminidase, known as *O*-GlcNAcase (OGA) [36, 37], catalyses removal of the moiety (Fig. 1). The presence of just one enzyme for the addition and one enzyme for the removal of *O*-GlcNAc in an organism differs greatly from the multitude of kinases and phosphatases that are responsible for phosphorylation [38, 39]. The question of how these two enzymes can recognise and modify such a large number of completely unrelated proteins, with no recognisable consensus sequence, has started to be addressed over the last decade or so. The deduction of the structure of homologues of each enzyme and a number of sophisticated recognition and mechanistic studies have started to elucidate some of these details, but there is still some way to go before the regulation and mode of action of OGT and OGA are fully understood.

OGA: REMOVAL OF THE O-GlcNAc MODIFICATION

OGA, or as it was first described hexosaminidase C [36, 40-43], was isolated and characterized at least a decade before the O-GlcNAc modification on proteins itself was elucidated [1]. It was later purified from bovine brain, sequenced and subsequently the gene was cloned, protein over-expressed recombinantly and characterized [37, 44]; the cloned sequence was identical to the *MGEA5* (meningioma expressed antigen 5) gene which had been identified in human meningiomas and preliminarily characterized as a hyaluronidase [45]. The properties of the recombinant enzyme cloned by the Hart group were in agreement with previous studies of the purified OGA [46]; the enzyme is a monomer [44], and has an optimal activity at neutral pH, as expected for an enzyme located in the nucleocytoplasm. In addition, OGA showed an apparent absolute selectivity for GlcNAc-containing substrates (including O-GlcNAc modified synthetic peptides), and did not hydrolyse GalNAc-configured substrates [37]. OGA activity was found to be present in both the nuclear and cytoplasmic fractions of cells, although when the gene encoding OGA was transfected into cells for over-expression, the recombinant tagged enzyme was observed to localize primarily to the cytoplasm. OGA is found in all tissue types, with highest expression levels in the brain, placenta, and pancreas [37].

The kinetic and physical properties of OGA contrast with the human lysosomal hexosaminidases which perform optimally at acidic pH, consistent with their localization to the lysosome, and are able to hydrolyse substrates containing either GlcNAc or GalNAc [37]. There are two genes encoding the lysosomal hexosaminidases in humans, *HEXA* encodes the α subunit and *HEXB* encodes the β subunit. The protein products are dimeric isozymes *in vivo*: hexosaminidase A (HexA) is a heterodimer of an α and β subunit, hexosaminidase B (HexB) is a homodimer of two β subunits, and hexosaminidase S, which is unstable, is a homodimer of two α subunits [47]. Mutations in these hexosaminidases are responsible for the G_{M2} gangliosidosis observed in Tay-Sachs and Sandhoff diseases [47]. A fourth hexosaminidase, named hexosaminidase D (HexD), has also been purified from brain tissues [42, 48, 49], but is less well characterised and the function *in vivo* currently unknown. The gene encoding HexD (*HEXD*) has recently been cloned from mouse and human and the protein expressed recombinantly [50]. These studies demonstrate that this enzyme, like OGA, operates under neutral conditions and is localized to the nucleus and cytoplasm. HexD differs from OGA in that it prefers substrates containing GalNAc, although it appears to also process GlcNAc-configured substrates. HexD, however, shows no activity toward O-GlcNAc modified synthetic peptides and so is likely to fulfil a cellular role distinct from that of OGA [50].

The development of a coherent and reliable classification system for carbohydrate processing enzymes, by primary sequence similarity, into different 'families' has been a great boon to those studying these enzymes. The families of enzymes are listed in the Carbohydrate Active enZyme (CAZy) database [51] (available at <http://afmb.cnrs-mrs.fr/CAZY/>). A feature of virtually all CAZy families is that, because the primary sequence dictates structure, and structure ultimately determines function, the catalytic mechanism is conserved within a family [52]. It is interesting to note that HexA, HexB, and HexD all belong to glycoside hydrolase family 20 (GH20), whereas HexC (OGA) is a member of family 84 (GH84). Family GH20 has been well characterized, both structurally and mechanistically for a number of bacterial enzymes [53, 54] and the human hexosaminidases HexA and HexB [55, 56]. The GH20 enzymes, which catalyse hydrolysis with retention of anomeric configuration, were shown to use a substrate participation mechanism using the acetamido group for nucleophilic attack at the anomeric carbon to create an oxazoline intermediate (Fig. 2). The intermediate is hydrolysed by a water molecule, activated by an enzymic general base, which attacks the anomeric carbon. This mechanism differs from the

'classical' glycoside hydrolase mechanism first postulated by Koshland, where the nucleophile is an enzymic residue [57], most commonly a carboxylate group [58].

Full length OGA is a large 103 kDa (916 amino acid) protein [37], which has been reported to stably interact with several proteins [37]. A 75 kDa splice variant also exists that arises from failure to splice out intron 10; this intron contains a stop codon so the protein encoded by the resulting transcript lacks the *C*-terminus of the enzyme, but does contain an additional 15 residues encoded by the intron [59]. This shorter isoform localizes predominantly to the nucleus [59], but when cloned recombinantly initially appeared to be inactive [44]. Interestingly, although there was originally debate, primarily based on bioinformatics studies and sequence alignments, about whether the active site (for OGA) was likely to reside at the *N*-terminal region [60, 61] or the *C*-terminus [62], OGA activity was shown conclusively to occur in a recombinant protein comprising residues 1-350, whereas a protein consisting of residues 351-916 was inactive. The rate of catalysis for this *N*-terminal recombinant enzyme, however, was around 1000-fold lower than for the full length protein, which is consistent with a functionally important role for the *C*-terminus [63]. A different study of the mouse OGA also confirmed the activity is encoded by a polypeptide comprising the *N*-terminal region, between residues 63-283 [64]. More recently, the short nuclear splice variant has been shown to have catalytic activity, albeit at a much lower rate than the full length enzyme [65, 66].

Sequence alignments indicated the *C*-terminal region has amino acid sequence similarity to known acetyltransferases [59]; OGA, when over-expressed in mammalian cells, was later reported to possess acetyltransferase activity towards histone synthetic substrates and free core histones [67]. More recently, however, efforts to independently verify this activity have been unsuccessful [68]. OGA also contains a caspase-3 cleavage site [44], an executioner caspase in apoptosis [69], which might explain the early observation of a 65 kDa fragment in cells [44]. This proteolytic processing by caspase-3 was shown to occur both *in vitro* and *in vivo* under conditions of Fas-induced apoptosis [68]. Surprisingly, after caspase cleavage, the OGA activity is retained, at least *in vitro* [44]. It is interesting to note, however, that if the *N*-terminal region (before the caspase cleavage site) is over-expressed recombinantly in cells alone there appears to be no OGA activity, but if co-over-expressed with the *C*-terminal region there is abundant activity present; this observation suggests an interaction between the two domains may be required for OGA to be optimally active [68]. The precise roles played by this *C*-terminal domain, however, remain to be uncovered.

MECHANISTIC STUDIES ON OGA

Consistent with the *N*-terminal domain of OGA harbouring the catalytic machinery required for *O*-GlcNAc processing, this domain bears similarity to around 70 open reading frames, predominantly found within bacteria, that place this enzyme into family GH84 of the CAZy system [51]. The sequence identity between the *N*-terminal domain of OGA found in eukaryotes, and in particular mammals, is extremely high (for example 97% between human and mouse; 53% between human and *Drosophila melanogaster*), but lower between eukaryotes and prokaryotes (the highest sequence identity is between human and *Streptomyces pyogenes* at 36%). Notably, GH84 enzymes bear no obvious sequence similarity to family GH20 enzymes, which includes the lysosomal β -hexosaminidases.

A significant advance in the analysis of OGA and other GH84 members came from a study which deduced the mechanism of the enzyme and paved the way for the design of potent and selective inhibitors [70]. OGA hydrolyses β -*N*-acetylglucosamine substrates with retention of anomeric configuration [70], a proposal verified in a study of the *Streptococcus pyogenes* GH84 homologue [71]. Measurement of the dependence of catalytic rates on pH

produced a bell-shaped curve, reflecting the ionization of important catalytic residues, with a pH optimum of 6.5 and kinetic pK_a values for the two enzymic residues of approximately 5.0 and 7.5 [63, 64, 72]. To determine if OGA used a 'classical' hydrolytic mechanism, as found for some enzymes processing substrates with a 2-acetamido group, such as the GH22 lysozymes [73] and the GH3 β -*N*-acetylhexosaminidases [74, 75], or the less common substrate participation mechanism which involves the acetamido group acting as a nucleophile (Fig. 2), Macauley *et al.* conducted Taft-like linear free energy relationships (for a brief overview on this approach see ref. [58]). This involved using a series of substrates bearing varying numbers of fluorine atoms substituted at the methyl group of the *N*-acetyl substituent in order to determine the effect on the rate of catalysis [70]. The electronegative fluorine atoms lower the basicity of the carbonyl group and, if OGA used a substrate participation mechanism as described for the GH20 enzymes, a decrease in rate of the enzyme catalyzed reaction would be expected. Such a trend was observed for both the human β -hexosaminidase and OGA; the slopes of these trends, however, were not equivalent, suggesting differences between these human enzymes either in the steric constraints in the active site or the position of the transition state along the reaction coordinate [70].

OGA was demonstrated to be able to hydrolyse both *O*- and *S*-linked synthetic substrates, using the same substrate participation mechanism, with similar efficiencies [72] (unlike most other glycosidases where this comparison has been made). Interestingly, Taft-like linear free energy analysis on the *S*-linked substrates showed a steeper slope than with the *O*-linked substrates suggesting a greater involvement of the 2-acetamido group. This is in agreement with secondary deuterium kinetic isotope effect (KIE) measurements (for a review on kinetic isotope effects see ref. [76]) on the second order rate constant (k_{cat}/K_M) where the k_H/k_D ratios for substrates containing either an *O*- or an *S*-glycoside are both large, suggesting dissociative transition states for both substrates. The KIE is slightly larger for the *S*-glycoside, suggesting the transition state may be later, with accordingly greater involvement of the 2-acetamido group [72]. In addition, Brønsted linear free energy analyses (for a brief overview on this approach see ref. [58]) were conducted to provide insight into the development of relative charge at either the glycosidic oxygen or sulphur atom at these transition states. This study was carried out by measuring the values of the second order rate constant for OGA processing series of both *O*- and *S*-linked substrates bearing phenol or thiophenol leaving groups, respectively, with a range of pK_a values. Interestingly, the slopes of the plotted lines (equal to β_{LG}) differed between the two series. The slope was shallower (corresponding to a smaller β_{LG} value) for the *O*-linked series suggesting there was little charge accumulation at the transition state, which is consistent with the kinetic isotope effect data that also suggest an advanced transition state where proton donation and glycosidic bond cleavage are significantly advanced. In contrast, the slope for the *S*-linked substrates was steep (with a large β_{LG} close to -1) which is consistent with extensive relative negative charge accumulation occurring on the leaving group thiol at the transition state, and hence either no, or extremely little, general acid catalysis. The lack of general acid catalysis for OGA catalyzed cleavage of *S*-linked substrates was also suggested in the plot of the second order rate constant against pH, which showed the presence of only one ionizable group as opposed to two for *O*-linked substrates [72].

The deduction of the catalytic mechanism used by OGA was aided by data from other families that use the same, but less common, substrate-assisted mechanism for hydrolysis. In addition to GH84, substrate-assisted hydrolysis has been shown to occur in the GH18 family of chitinases [77-80], GH20 family of chitobias and β -hexosaminidases [53, 54, 56, 79, 81], GH56 family of hyaluronidases [82], very recently in the GH85 family of endo- β -*N*-acetylglucosaminidases [83-85] (Fig. 3d) and is also likely in GH25 enzymes [86]. Obviously a requirement for the enzymes using this mechanism is the presence of an

acetamido group in the substrate which acts as the nucleophile to attack the anomeric position.

Two key enzymic groups (virtually always the carboxyl group of either an aspartate or glutamate residue) have been shown to be important for catalysis during the substrate-assisted mechanism. The mechanistic studies discussed above provide good support for the proposed mechanism with only some subtle variation between GH families. During the first step of the mechanism, one residue which is in its carboxylate form in the resting enzyme orients and polarizes the 2-acetamido group to increase its nucleophilicity, and hence facilitates attack at the anomeric centre to form an oxazoline intermediate. The other residue, a carboxylic acid in the resting enzyme, acts to aid fission of the glycosidic bond and leaving group departure by general acid catalysis (Fig. 2). During the second step, this second carboxyl residue now acts as a general base to facilitate the attack of a water molecule at the anomeric centre of the oxazoline intermediate, while the polarizing residue aids expulsion of the 2-acetamido group from the anomeric position. The net result of these processes is cleavage of the glycosidic bond to liberate the leaving group (the hydroxyl group of the protein in the case of OGA) and the hemiacetal of *O*-GlcNAc with retained stereochemistry. There is debate as to the absolute role the polarizing residue plays in catalysis and this remains unclear at present; some postulate for some families of glycoside hydrolases using substrate-assisted catalysis that the carboxylate residue electrostatically stabilizes the oxazolinium intermediate [78, 87], whereas others suggest it is more likely to function as a general acid/base to aid formation of the oxazoline intermediate [63, 70].

Structural and kinetic methods have been valuable for defining the important residues required for catalysis in these families using substrate participation. In the family GH18 [80, 88] and GH56 [82, 89] enzymes, there is a highly conserved DXE (where X can be any amino acid) motif; in the cases examined, the glutamate served as the acid/base residue and the aspartate acted as the polarizing residue of the 2-acetamido group. The GH85 enzymes appear to have a similar motif to the GH18 and GH56 enzymes, but instead of possessing an aspartate residue, have a highly conserved asparagine residue in the equivalent position (and thus have the motif NXE) [60, 84, 85, 90]. The GH85 crystal structure suggests, however, that a hydrogen bond is formed with the nitrogen atom of the amide moiety of the asparagine residue, and the authors of one report propose that the residue may exist in its imidic acid form, which by acting as a general base is able to aid formation of the oxazoline intermediate [85]. Further studies, however, will be needed to support or discard this proposal. The family GH20 and GH84 enzymes differ from GH18, GH56, and GH85 in that the key catalytic residues are adjacent to each other, in a DE motif; the glutamate acts as the acid/base residue and the aspartate as the polarizing residue of the 2-acetamido group [91-93]. It is interesting to note that families possessing a motif with an amino acid between the two catalytic residues hydrolyse endoglycosidic bonds, whereas GH20 and GH84, which have adjacent catalytic residues, hydrolyse exoglycosidic bonds; this difference is also reflected in the size of the active site pocket (Fig. 3d).

Given the conservation of the substrate-assisted mechanism with these other families, the key catalytic residues in OGA were likely to be found in an analogous motif. One bioinformatics study showed that a conserved DD motif in GH84 enzymes could be aligned with the catalytic DE motif found in GH20 enzymes, and postulated that these two aspartate residues would constitute the residues involved in catalysis [60]. There is, however, also a highly conserved DXE motif as found in the GH18 and GH56 enzymes which had also been suggested may include the catalytic residues [64]. The residues in the DD motif, which equate to Asp174 and Asp175 in the human OGA, were mutated to alanine and displayed significant decreases in activity [63]. Perturbations were also observed in the pH dependence of the mutant proteins. The pH dependence for the wild type enzyme is bell-shaped, a shape

governed by the two catalytic residues where in the resting enzyme the polarizing residue is likely to be deprotonated and the acid/base residue protonated. Mechanistic studies indicate the acidic limb of the curve is most reasonably assigned to ionization of the polarizing residue, and the basic limb assigned to ionization of the acid/base residue [63]. The pH dependence of the D174A mutant showed only the basic limb, which is consistent with titration of the acid/base residue, and thus implicates Asp174 as the polarizing residue of the acetamido group in catalysis. The pH dependence of D175A was more ambiguous, but rates of hydrolysis were lower at higher pH, which is consistent with Asp175 acting as the acid/base residue. This proposal is in agreement with Brønsted analyses of the D175A mutant, which gave a β_{LG} of -1 , indicating the absence of acid catalysis [63]. In addition, the activity of the D175A mutant toward substrates with a good leaving group (i.e. where general acid catalysis is not necessary) was increased upon addition of azide, an exogenous nucleophile, and also yielded a β -azide product. Previous studies have shown that in the absence of an acid/base residue, a small molecule such as azide can attack the intermediate with greater effect than a water molecule since the water molecule cannot be activated by the general base [94-96]. The rate of hydrolysis of *O*- and *S*-linked glycosides by the D175A mutant was also similar [63], which is consistent with acid catalysis providing no benefit for cleavage of *S*-linked substrates with leaving groups that have a high pK_a value, but being of greater importance for corresponding *O*-linked substrates. The role for the polarizing residue, Asp174, acting as a general acid/base rather than as a stabilizer of the oxazolinium ion is consistent with the large change in pK_a of the protonated nitrogen in the acetamido group upon changing from an amide to an oxazoline; this dramatic change in pK_a , which crosses the pK_a value of the polarizing residue, suggests complete proton transfer from the amide and thus formation of an oxazoline intermediate [63]. The role of the polarizing residue acting as a general acid/base, but not as a stabilizer of an oxazolinium ion, would also agree with the GH85 enzymes possessing an asparagine residue in this position; in an imidic acid form this asparagine residue would be able to accept a hydrogen bond from the nitrogen atom in the acetamido group, but as it is neutral in its resting state it is not in a position to stabilize the charge present on an oxazolium ion [85].

STRUCTURAL INSIGHTS USING BACTERIAL HOMOLOGUES OF OGA

Ideally, a three-dimensional structure of the human OGA would serve to support or challenge a number of conclusions from the mechanistic studies, offer insight into the potency of selective inhibitors of OGA, inform the design of future inhibitors against the human enzyme, and facilitate studies directed at understanding the involvement of OGA in protein-protein interactions. Despite significant and continuing efforts, OGA has not proved amenable to analysis by X-ray diffraction to date. A number of prokaryotic homologues in family GH84, however, show significant homology to the *N*-terminal region of OGA, which has the catalytic activity. The structures of two GH84 bacterial homologues of OGA, one from *Bacteroides thetaiotaomicron* (*BtGH84*) [97] and the other from *Clostridium perfringens* (*CpGH84*) [98, 99] were published in 2006. The functional role of these bacterial enzymes *in vivo* remains unknown since prokaryotes do not have the *O*-GlcNAc post-translational modification, but both enzymes were shown to be capable of cleaving *O*-GlcNAc from proteins in cell lysates [97, 98].

BtGH84 and *CpGH84* display an *N*-terminal α/β fold domain, followed by a canonical $(\beta/\alpha)_8$ fold domain, and ending in three (in *BtGH84*) or one (in *CpGH84*) further domain(s) at the *C*-terminus, which are specific to only the prokaryotic enzymes (Fig. 3a) [97, 98]. The structures reveal a remarkable structural homology to the GH20 enzymes [53, 54, 56], despite a lack of sequence identity. The $(\beta/\alpha)_8$ fold domain houses the catalytic machinery in a shallow depression, and the key DD motif which comprises the residues required for catalysis, is situated at the end of $\beta 4$ strand; strikingly the positioning of these key residues

is identical in the GH20 enzymes. The presence of the two carboxylate residues in the active site were judged to be appropriately positioned (using complexes with inhibitors, see below for more details) to act as the key catalytic residues, consistent with previous proposals made for the human enzyme [60, 63, 70, 72]; the first of these two aspartates acts as the polarizing residue of the 2-acetamido group and the latter as the acid/base residue. Site-directed mutagenesis of the bacterial enzymes also supports assignment of these catalytic residues as discussed above [97, 98]. The active site contains two conserved tryptophan residues, which are predicted to carefully position the acetamido group, and line a sizable pocket which is capped by a cysteine (in *BtGH84* and OGA, but not *CpGH84*) and a methionine residue [97, 98] (Fig. 3b and 3c). It is postulated that a conserved tyrosine residue is in a suitable position to stabilize the transition state during catalysis, through hydrogen bonding with the glycosidic oxygen (where partial negative charge would develop at the transition state) [98]. The structures also provided evidence, in addition to biochemical experiments [71], that these GH84 enzymes are not likely to act as hyaluronidases (as had been initially proposed [59]), as there is not sufficient room in the active site pocket to bind a substrate any larger than a monosaccharide (Fig. 3d). It is important to note that despite the relatively modest sequence similarity at the amino acid level in comparison to OGA, the important active site residues of these bacterial structures differ by only one (for *BtGH84*) or two (for *CpGH84*) residues; these structures therefore likely provide a highly accurate picture of the OGA catalytic machinery.

INHIBITION OF OGA

Given the number of biological roles proposed for *O*-GlcNAc and the relatively early stage in our understanding of the functional roles of this common post-translational modification, inhibition of OGA has been a topic of some interest since these inhibitors can be valuable research tools. As is the case with virtually all inhibitors, high potency is a necessity that facilitates avoiding off-target effects. More of a challenge with OGA, however, has been obtaining inhibitors selective for this enzyme over the functionally related human β -hexosaminidase, which as discussed above uses the same catalytic mechanism. The design of such selective inhibitors has, however, progressed over the past few years.

The vast majority of early studies in the field, carried out before design of selective inhibitors, involved increasing cellular *O*-GlcNAc levels using streptozotocin (*N*-methyl-*N*-nitrosoureido *D*-glucosamine; STZ) [100] and *O*-(2-acetamido-2-deoxy-*D*-glucopyransylidene)-amino-*N*-phenylcarbamate (PUGNAc) [101] (Fig. 4). STZ was shown to be a modest inhibitor of OGA *in vitro* [70, 102, 103] and has long served as a way to induce experimental type I diabetes [104] through its ability to induce apoptosis of pancreatic β -cells [105, 106]. These observations, coupled with the abundance of *O*-GlcNAc in β -cells, prompted the reasonable hypothesis that STZ induces apoptosis of pancreatic β -cells due to increased *O*-GlcNAc modification of proteins following inhibition of OGA [102, 103, 107, 108]. The mode of inhibition of OGA has been elusive, with suggestions that STZ is a suicide substrate *via* *S*-nitrosylation of a cysteine residue in the active site of the enzyme [108], or that OGA can assist in the conversion of STZ into a more potent tight binding transition state analogue [109]. A greater body of data, however, supports the ability of STZ to cause apoptosis through other means, such as free radical production (for example nitric oxide from break down of the nitroso moiety) leading to DNA damage [110, 111] and methylation of DNA [112, 113]. The apoptosis-inducing effect of STZ being ascribed to inhibition of OGA was drawn into question when comparisons by two independent groups were made with PUGNAc; both compounds caused increases in cellular *O*-GlcNAc levels, but only STZ caused cell death [114, 115]. Furthermore, STZ is a low millimolar inhibitor of OGA, and inhibition has been shown to be reversible and hence it is not a suicide substrate as once proposed [70, 116]. The structures of *CpGH84* and

BtGH84 were later solved in complex with STZ, and showed the absence of any covalent attachment with an active site cysteine residue, and also a lack of conversion into a different compound as once proposed [116, 117]. Further, a recent compelling study showed the *galacto*-epimer of STZ, which did not inhibit OGA or increase cellular *O*-GlcNAc levels, induced apoptosis of pancreatic β -cells suggesting the ability of STZ to induce apoptosis is independent of OGA and attributable to the reactive nitrosourea group [116]. The use of STZ in studies of the functional role of *O*-GlcNAc therefore presents obvious and serious problems and should be discouraged.

PUGNAc is a potent inhibitor of OGA, with a K_i value of around 50 nM [46, 70, 118], but is also a potent inhibitor of the GH20 human β -hexosaminidases. PUGNAc is able to increase cellular *O*-GlcNAc levels by inhibiting OGA [118], and as such has been used to implicate aberrant *O*-GlcNAc cycling as a cause of insulin resistance [23, 119]. Interpretation of the results with this inhibitor are, however, somewhat hampered by its poor selectivity for OGA [70]. The structure of both *BtGH84* and *CpGH84* have been solved in complex with PUGNAc [24, 98]; in each case the *N*-phenylcarbamate moiety forms hydrogen bond interactions with the active site residues, including one between the exocyclic oxime nitrogen atom and the catalytic acid/base residue, and the phenyl moiety, which protrudes from the active site pocket, takes advantage of hydrophobic interactions with aromatic side chains. It is interesting to note that 2-acetamido-2-deoxy- β -gluconhydroximo-1,5-lactone (LOGNAc), which is similar to PUGNAc but lacks the phenyl carbamate group, is around 30 times less potent [46] demonstrating that this moiety makes an important contribution to the binding. The use of PUGNAc in studies has very recently been questioned, however, as it has been found to have a number of off-target effects; not only can it inhibit β -hexosaminidases, which leads to an increase in ganglioside levels that could impact insulin resistance [120], but it has also been shown to inhibit enzymes from other enzyme families [121] and impact cell growth [122]. The results of a study by Macauley *et al.* show that a highly selective inhibitor of OGA (see below) did not induce insulin resistance in 3T3-L1 adipocytes as PUGNAc did [24]. In addition, PUGNAc appears unable to cross the blood-brain barrier [123].

Despite the prolonged use of STZ and PUGNAc as inhibitors of OGA both *in vitro* and in cells to induce a diabetic or insulin resistant state, it is apparent that both compounds have off-target effects and interpreting the results obtained using either compound should be done with caution. The deduction of the substrate-assisted mechanism led to the use of 1,2-dideoxy-2'-methyl- α - β -glucopyranoso-[2,1-*d*]- $\Delta 2'$ -thiazoline (NAG-thiazoline) [124] as an inhibitor of OGA [70]. This is a mimic of the oxazoline intermediate (or a closely related transition state), with a sulphur atom in place of the oxygen in the oxazoline ring, and had been shown to be a good inhibitor of GH20 enzymes which use the same mechanism [54, 124]. NAG-thiazoline inhibits OGA with a K_i value of around 70 nM, which is similar to the value measured against the human β -hexosaminidase and so, as seen with PUGNAc, this inhibitor shows little selectivity [70]. The structure of *BtGH84* in complex with NAG-thiazoline, where the pyranose ring adopts a slightly distorted chair (4C_1) conformation, shows a hydrogen bond interaction between Asp242 (predicted to be the polarizing residue) and the nitrogen of the thiazoline ring [97]. The issue of finding a selective inhibitor for OGA over the β -hexosaminidase was addressed in the design of a series of inhibitors where an aliphatic chain of increasing length was added to the thiazoline ring of NAG-thiazoline [70]. At the time this showed considerable foresight as the bacterial homologue structures had not yet been elucidated. The idea was based on the knowledge that the Taft-like analysis suggested the bulk of the *N*-acyl group may have been important for the substrate recognition in OGA, that GH20 enzymes possessed a small pocket which fitted the acetamido group snugly, and that OGA could process a chromogenic substrate bearing an azidoacetyl group [125]. Indeed, as the aliphatic chain on the thiazoline ring was increased,

the inhibitors were shown to dramatically increase in selectivity for OGA over the β -hexosaminidase, with only a small penalty in the potency of the compounds, and also were able to increase cellular *O*-GlcNAc levels with no effect on cell growth or morphology [70]. The structural solution of two bacterial homologues of OGA later corroborated these biochemical findings, and revealed that indeed they possessed a deeper pocket into which the 2-*N*-acyl group protrudes [97, 98] (Fig. 3c). The observation of the thiazoline ring bearing an *N*-butyl chain (1,2-dideoxy-2'-propyl- α -D-glucopyranoso-[2, 1-*d*]- Δ 2'-thiazoline; NButGT) in the active site of BtGH84 showed it bound in an identical position to NAG-thiazoline, with the acyl chain nestled in the active site pocket [126]; this compound has a K_i value against OGA of 230 nM and displays a 1500-fold selectivity over β -hexosaminidase. Efforts were made to exploit PUGNAc in a similar way with elongation of the aliphatic chain from the *N*-acyl moiety; there were some modest gains in selectivity for OGA over the lysosomal β -hexosaminidase, but were accompanied by a sacrifice in potency, and these compounds were significantly less potent than the thiazoline series [127, 128]. Others have also functionalized thiazoline inhibitors; a number of these also show apparent high potency coupled with good selectivity for OGA [129]. Arguably the most successful attempt to generate more potent and selective inhibitors for OGA came with the advent of 1,2-dideoxy-2'-ethylamino- α -D-glucopyranoso-[2,1-*d*]- Δ 2'-thiazoline (Thiamet-G) [130]. It was speculated that a compound in which the thiazoline ring had a higher pK_a would be protonated at physiological pH (NAG-thiazoline itself has a pK_a of approximately 5) and make more favourable electrostatic interactions in the active site with the carboxylate of the polarizing residue; thus a compound similar to NButGT with the substituent next to the thiazoline ring linked *via* a nitrogen atom was designed. It was proposed this would increase the basicity of the endocyclic nitrogen atom, as well as retain the steric considerations required for selective binding. A highly potent inhibitor resulted, with a K_i value of 21 nM and a dramatic 37,000-fold selectivity over the lysosomal β -hexosaminidase [130]. The structure of Thiamet-G in complex with BtGH84 shows that Asp242 (the key residue important for polarizing the acetamido group during catalysis) engages in favourable electrostatic interactions with both the endocyclic and exocyclic nitrogen atoms in the inhibitor (Fig. 3b); a consequence of these interactions is the subtle rearrangement of some of the active site residues which contracts the active site and draws the enzyme and inhibitor closer together [130]. Thiamet-G is capable of increasing *O*-GlcNAc levels in cells and *in vivo*, including in the brain as it is able to cross the blood-brain barrier, and in addition shows greater chemical stability than the other thiazoline analogues. As such it is able to increase the level of *O*-GlcNAcylation on the protein tau, and thus cause a decrease in phosphorylation; an effect that may enable a better understanding of the hyperphosphorylation of this protein in various neurodegenerative diseases [130].

A different approach to the development of highly potent OGA inhibitors has been based on the naturally occurring β -hexosaminidase inhibitor nagstatin [131]. Following its discovery, nagstatin was synthesized [132] and the scaffold manipulated to make potent β -glucosidase inhibitors [133, 134]. These tetrahydroimidazopyridine compounds were substituted with a number of different functional groups in order to promote favourable interactions in the active site of enzymes of interest [135-138], and these have proven to be among the most potent β -glycosidase inhibitors studied [139]. The potency of the tetrahydroimidazopyridine compounds are attributed to the protonated nitrogen atom in the imidazole forming a strong hydrogen bond interaction with the general acid/base residue, and the pseudo-glycoside ring adopting a half chair/envelope conformation which mimics the proposed oxocarbenium transition state [140]. The *gluco*-epimer of nagstatin inhibited OGA with a K_i value of 420 nM, but was more potent against β -hexosaminidase [141]. Subsequent development of a PUGNAc-imidazole hybrid produced a low micromolar inhibitor of human OGA. The PUGNAc-imidazole hybrid was observed by crystallography in complex with BtGH84 and when overlapped with the CpGH84 complex with PUGNAc showed the phenyl ring was

forced into a position where it was more solvent exposed and made fewer interactions with the enzyme, and that the amide carbonyl was in the opposite orientation which prevented it from forming a hydrogen bond with the general acid/base residue [141]. Along similar lines, the van Aalten group designed GlcNAc-configured nagstatin derivatives (GlcNAcstatins) such as GlcNAcstatin C, which bears the tetrahydroimidazopyridine scaffold with a phenethyl functional group attached, and an isobutanamido moiety at C2. This was shown to be low picomolar inhibitor against *CpGH84* [142], reported to be a 4 nM inhibitor of human OGA, and found to display 160-fold selectivity over the lysosomal β -hexosaminidase [143]. In complex with *CpGH84*, GlcNAcstatin C was observed in an envelope conformation and the imidazole made a hydrogen bond interaction with the general acid/base residue as predicted. The phenethyl group extended out of the active site and formed hydrophobic interactions with a tryptophan residue [142]. GlcNAcstatin D, which bears an *N*-propionyl group at C2, has a K_i value of 0.74 nM but only 4-fold selectivity over β -hexosaminidase; it is shown to bind in the same manner as GlcNAcstatin C [143]. Both compounds were shown to be able to increase cellular *O*-GlcNAc levels [142, 143].

Several other unrelated inhibitors have been reported in the literature. These include alloxan, an analogue of uracil which is also capable of inhibiting OGT (see below) [144], but has some toxic effects and should therefore be used with caution [145], α -linked GlcNAc thiosulfonates, which appear to be better inhibitors of the short isoform of OGA than the full length enzyme [146], 7-membered ring azepanes [147], and an analogue of 6-*epi*-valienamine bearing an acetamido group [148]. The observation of two of the azepane compounds and the 6-*epi*-valienamine analogue in the active site of *BtGH84* showed they all bound in an unusual distorted conformation which differs from their predicted conformation in solution [147, 148].

The question of which glycosidase inhibitors are true mimics of the transition state has started to be addressed in recent years (for some recent examples see refs. [139, 149, 150]). Pauling first proposed in the 1940s that the highest affinity inhibitors for an enzyme would be those containing features of the transition state [151, 152]. More recently, Wolfenden demonstrated that glycoside-degrading enzymes can enhance the rate of catalysis 10^{17} fold over the uncatalysed reaction, making them a remarkably proficient class of enzyme [153]. If binding of a glycosidase to a transition state analogue could capture all of the binding energy proposed to be available by Wolfenden, this would translate to an estimated dissociation constant for the transition state of approximately 10^{-22} M, an obviously extraordinary affinity [153, 154]. Therefore if inhibitors were designed which truly mimicked the transition state, in theory a significant portion of this binding potential could be harnessed for potent inhibition. In the case of glycosidases, the widely proposed oxocarbenium ion-like transition state possesses double bond character along the C1-O5 bond causing the C1, C2, O5 and C5 atoms of the pyranose ring to lie in a plane. Accompanying this flattening of the pyranose ring would be significant positive charge delocalization along the C1-O5 bond [155, 156]. A study by Whitworth *et al.* addressed this issue of transition state mimicry for the OGA inhibitors PUGNAc and NAG-thiazoline using linear free energy relationships [126]. Such linear free energy relationships predict that, for a genuine transition state analogue, alterations made in the chemical structure of the substrate should perturb the transition state for the enzyme catalyzed reaction to an equivalent extent as parallel and analogous changes to a transition state analogue affect its binding to the enzyme. This can be assessed by introducing systematic changes at one position in an inhibitor and measuring the resultant K_i value, and also making the equivalent change in the substrate and determining the k_{cat}/K_M value, which reflects on the transition state structure in the first step of catalysis. Assuming these modifications do not change the mode of substrate binding or reactivity, or the rate-determining step during hydrolysis, the plot of the inverse logarithm of K_i against the inverse logarithm of k_{cat}/K_M should give a correlation

with a slope equal to one, if the inhibitor is a transition state mimic [157]. The correlation for OGA was assessed using series of inhibitors and substrates which varied in the volume of the alkyl group in the *N*-acyl moiety at the C2 position. The resulting plots of inverse logarithm K_i against inverse logarithm k_{cat}/K_M for the NAG-thiazoline series revealed a strong correlation, with a slope of one within error, indicating that it is a true transition state mimic. A different case, however, was observed for the PUGNAc series, which showed a weaker correlation suggesting this inhibitor may be a worse transition state analogue [126]. These observations were surprising given the sp^2 anomeric centre in PUGNAc would be predicted to be more 'transition state-like' than the sp^3 centre in NAG-thiazoline, which may be considered more closely resembling the oxazoline intermediate, at least from a geometric perspective. It has been postulated this apparent anomaly may be attributable to the longer C-S bond length in the thiazoline ring of the inhibitor as compared to that of the C-O bond in the oxazoline ring of the intermediate. This increased length results in some distortion of the pyranose ring and might mimic the partial, but significant, bond order between the nucleophilic carbonyl oxygen atom and anomeric carbon of a very late transition state in which the acetamido group is heavily involved [126].

Theoretical studies have also been devoted to understanding the binding of PUGNAc and NAG-thiazoline to OGA. In one such study, molecular dynamics methods were used to investigate the conformational similarities between inhibitors and a calculated transition state for the uncatalyzed formation of the oxazoline intermediate from *N*-acetylglucosaminides [158]. These simulations suggested that NAG-thiazoline was a poor conformational mimic of the transition state, but that GlcNAcstatin C was a better analogue of the modelled solution phase reaction. PUGNAc showed a high degree of conformational flexibility, and appeared to be only planar at C1 rather than C1 and O5 as would be likely at the transition state [158]. Electrostatic potentials have also been calculated for PUGNAc and NAG-thiazoline, and predictions made about the likely interactions with active site residues that drive binding of inhibitors suggest these may influence the design of inhibitors in the future [159].

OGT: ADDITION OF THE O-GlcNAc MODIFICATION

OGT was first purified from rat liver, where it was suggested to form a heterotrimer consisting of two 110 kDa subunits and one 78 kDa subunit [160]. Subsequently, the genes encoding a 110 kDa polypeptide (1046 amino acids) in rat, human, and *C. elegans* were cloned and the proteins over-expressed recombinantly [35, 161]. OGT is found in all tissue types, although there are at least four transcripts or isoforms in rat and human, some of which appear to be tissue-specific [35, 161, 162]; the 110 kDa polypeptide itself, however, is observed in virtually every tissue examined [161], with particularly high levels in pancreatic tissue [35]. A splice variant encoding a smaller OGT isoform has also been identified which is localized to the mitochondria [163, 164]. OGT was shown to be present, both endogenously and when transfected into cells, in the nucleus and cytoplasm; this observation is consistent with the presence of O-GlcNAc on nucleoplasmic proteins and the localization of OGA, as well as with the fact that OGT possesses a nuclear localization sequence [35, 161]. It was shown, by recombinantly expressing the 110 kDa subunit of OGT (which exists as a homotrimer [165]) in *Escherichia coli*, that this protein product alone was capable of modifying targets *in vitro* [165, 166]. OGT itself can be modified with O-GlcNAc, which may be one way in which cycling is regulated *in vivo*, and can also be tyrosine phosphorylated [161]. In addition it has been proposed that OGT possesses a phosphoinositide-binding region [22]. The gene encoding OGT has been shown in mammals to be vital for embryonic stem cell viability and embryonic development [2].

Sequence homology has shown OGT comprises two domains; an *N*-terminal domain consisting of 13.5 tetratricopeptide repeat (TPR) motifs (residues 1-464) [35, 161, 167, 168] and a *C*-terminal domain where the glycosyltransferase activity resides [165, 166]. The TPR domains likely modulate binding to a number of different protein substrates or accessory proteins [165], and have been demonstrated to interact with a variety of proteins including, for example, a histone deacetylation complex by binding to the corepressor mSin3A [169]. The TPR motif is found as a domain in many proteins, and appears to be ubiquitous through many organisms and cellular locations [167]. These repeats play an important role in modulating protein-protein interactions in a range of cellular processes including protein transport, transcription and cell cycle control [167, 168, 170]. The TPR motif itself consists of a 34 amino acid repeat, which contains eight residues throughout the sequence that are highly conserved, and the rest tend to be highly consistent in residue 'type' [167, 168]. The TPR motifs are often found in serial arrays having a different number of repeats (usually between 3 and 16), which structurally form a superhelix that is responsible for mediating the interactions with other proteins. The helix has a concave face which presents the precise binding residues defined in the overall structure to the target in a specific fashion [168]. It is likely to be the number of repeats, and hence the size of the TPR groove(s) generated, that determines the number of different protein targets with which the TPR domains can interact [167]. Consistent with the presence of so many TPR domains on OGT, numerous binding partners of OGT have been proposed [169, 171-175]. Interestingly, OGT does not appear to recognize a consensus sequence directing modification of target proteins, although a preference for Pro-Val-Ser-Thr and Thr-Thr-Ala has been proposed based on proteomics studies in which several *O*-GlcNAc sites were identified on a number of different proteins [33]. Systematic amino acid variation of residues surrounding an *O*-GlcNAc site on an α -crystalline peptide has also suggested OGT has some modest preferences for particular residues at certain positions [176]. The TPR domains have also been found to play a role in the activity of OGT. The deletion of a number of the TPR domains from the *N*-terminal end of OGT demonstrated that some, but not all, were required for enzymatic transfer of *O*-GlcNAc onto peptides; those missing the first three or six TPR domains had full activity but this region aided trimerization of OGT, whereas those missing the first nine or eleven TPR domains were catalytically inactive [165]. Lubas and Hanover report a difference in the activity with TPR deletion mutants between peptide and protein substrates, although they appear to have used the mitochondrial OGT which has only nine TPR domains. They found that while the *N*-terminal deletion of the three most *N*-terminal TPR domains resulted in an enzyme having unchanged OGT activity on peptide substrates, OGT activity of the same construct measured using nucleoporin p62 (np62) as a substrate was impaired, suggesting these TPR domains are involved in substrate recognition [166]. In a different study, the interaction of OGT (110 kDa polypeptide) with the interaction domain (OID) of OIP106 (OGT-interacting protein, 106 kDa [177]) was investigated. If OGT was missing the most *N*-terminal 5.5 TPR domains it was unable to interact with OID (and was catalytically inactive), whereas if it was missing only the first 2.5 TPR domains it was still able to bind, yet remained catalytically impaired [171]. Recombinantly expressed and purified TPR domains of OGT alone were capable of inhibiting OID glycosylation by competing for the same binding site on OGT; the same is not true, however, for a peptide substrate, which reaffirms the role of the TPR domains as a docking or binding domain [171]. In addition, OID and np62 also compete with each other in *in vitro* glycosylation assays using OGT [171]. More recently, an appealing hypothesis has emerged that OGT may also be positively regulated through its interaction with other protein partners such as PGC1- α [175], mSin3A [169], myosin phosphatase targeting subunit 1 [173], co-activator-associated arginine methyltransferase 1 [173], OIP106 [177] and p38 MAP kinase [178].

MECHANISTIC STUDIES OF OGT

OGT is classified into family GT41 of the CAZy system [51], which it currently shares with around 270 prokaryotic sequences and 70 other eukaryotic sequences. OGT catalyzes glycosyl transfer with inversion of stereochemistry to generate the β -*O*-GlcNAc linkage to serine and threonine residues of substrate proteins. The amino acid sequence identity of human OGT with OGT proteins from other higher eukaryotes is very high (>95%), but much lower when compared with prokaryotic sequences encoding related proteins (around 35-40% at most). Mechanistic studies of OGT lag behind those of OGA and related homologues, which reflects the nature of the carbohydrate enzyme field in general [179, 180]. Primarily, the relative lack of knowledge about glycosyltransferases stems from the fact that many of these enzymes are larger in size than hydrolases and/or membrane bound, which makes them harder to express recombinantly, characterize, and crystallize for structural studies. In addition, the availability of small molecule tools or probes, such as synthetic substrates and inhibitors, is more limited for the glycosyltransferases.

The Leloir donor substrate of OGT is uridine diphospho-*N*-acetylglucosamine (UDP-GlcNAc) [34], which is the end product of the hexosamine biosynthetic pathway. It is estimated that 2-3% of the total incoming glucose to the cell is diverted from catabolism into this pathway to form UDP-GlcNAc [181]. OGT shows an optimum catalytic activity at pH 6, but below pH 6 and above pH 7.5 enzyme activity decreases significantly [160]. Unlike many other glycosyltransferases, OGT does not require the presence of a divalent metal ion for efficient catalysis [34, 160]. Early kinetic studies of OGT revealed unusual Michaelis-Menten kinetics for this enzyme. When using UDP-GlcNAc and a peptide substrate containing an endogenous *O*-GlcNAc-modifiable site, no saturation was observed and three K_M values for UDP-GlcNAc were reported (6, 35, and 217 μ M). In addition, it was observed that OGT had a different affinity for peptide substrates at different UDP-GlcNAc concentrations, suggesting that UDP-GlcNAc levels can differentially modulate the affinity OGT has for different peptide acceptor sites [165]. Others, however, have observed that when a protein such as np62 is used as the acceptor, only one K_M for UDP-GlcNAc is observed (K_M for UDP-GlcNAc of 0.5 μ M; K_M for np62 of 1.2 μ M); these kinetics appear to have reported on the mitochondrial OGT however, and so a direct comparison may not be entirely appropriate [166]. The use of a whole protein substrate rather than a peptide may, however, be more physiologically relevant since the natural substrates of OGT *in vivo* are intact proteins. Dissection of the kinetic parameters on OGT using a peptide substrate suggested the enzyme used a random bi-bi mechanism [165]. Further studies are needed to provide greater insight into the mechanisms by which OGT recognizes and acts to modify its protein substrates.

STRUCTURAL INSIGHTS USING A BACTERIAL HOMOLOGUE OF OGT

Full length human OGT has, to date, eluded crystallization and structure solution. Efforts to realize the structure of OGT are a topic of intense current interest and a structure would significantly aid attempts at dissecting the mechanistic details and perhaps aid inhibitor design. Jínek and coworkers, however, were able to solve the structure of the TPR domains from OGT [182], which has offered some insight into the molecular basis for substrate binding and recognition (Fig. 5c). The crystallized protein lacked the catalytic domain and consisted of 11.5 out of the 13.5 TPR domains (residues 16-400) of OGT. This protein construct was shown to contain the region responsible for recognition of np62, as demonstrated by competition assays with the full length OGT. The TPR domains form a homodimer (which differs from the trimeric state proposed for full length OGT [165]) of right-handed superhelices. Each TPR domain consists of two antiparallel helices, which possess the hydrophobic residues in conserved positions as seen in other proteins containing

TPR motifs [182]. The convex face of the superhelices is responsible for dimerization and this process is mediated primarily by hydrophobic interactions; mutation of a tryptophan and isoleucine which appear to be important for this dimerization causes the TPR domains to appear monomeric by size exclusion chromatography (whereas the wild type TPR domains run at a mass consistent with being dimeric) [182]. The majority of the conserved surface exposed residues line the inner concave surface of the superhelix; in particular the central part of the groove is mostly lined with asparagine residues, which form a continuous 'ladder' through the superhelix [182]. These conserved asparagine residues on the inner surface of the superhelix bear a marked similarity to the ARM-repeat proteins importin- α and β - catenin; in these proteins the asparagine residues are important for recognition of their target peptides and form bidentate hydrogen bonds with the protein main chain of the substrate [183-185] (Fig. 5d). It may be possible that different pockets of the TPR domains of OGT recognise and bind to different protein substrates; if this is the case the asparagine residues may be important in forming the interaction, but it is likely to be the neighbouring residues that confer specificity [182].

Despite the lack of a full length OGT structure, there are prokaryotic homologues in GT41 which show significant homology to OGT and, in particular, the C-terminal domain. The structure of a GT41 bacterial homologue of OGT from the plant pathogen *Xanthomonas campestris* (*XcGT41*), which has 36% sequence identity to the human enzyme, was published last year by two independent groups [186, 187]. The substrates of the bacterial enzymes are as yet unknown, but it was shown that, like OGT, *XcGT41* was able to catalyze the transfer of UDP-GlcNAc to water *in vitro* [186]. No activity of *XcGT41* was detected towards human or bacterial cell lysates or synthetic peptides, but the enzyme may be able to modify a protein in *Arabidopsis thaliana* cell lysates with *O*-GlcNAc [187]. The structure reveals that *XcGT41* comprises an N-terminal domain consisting of 5.5 TPR domains, and a C-terminal catalytic domain which displays the GT-B topology of glycosyltransferases (Fig. 5a). The GT-B topology comprises two (α/β)-fold domains, between which the UDP-GlcNAc binds; these observations are consistent with previous predictions for this enzyme [188, 189]. The bacterial enzymes, however, lack a 120 amino acid insertion between the two domains at the N- and C-terminus of the GT-B fold, which appears in all mammalian sequences [186, 187]. The structure of *XcGT41* in complex with UDP revealed the likely residues which are important for binding of the nucleotide-phosphate donor, and these were highly conserved in the eukaryotic sequences. Systematic mutation of these residues in human OGT support the structural model and revealed residues important for binding UDP-GlcNAc as well as those that are necessary for catalytic activity [186, 187] (Fig. 5b). In addition, a histidine residue (His558 in human OGT, which is equivalent to His218 in *XcGT41*) was identified as the likely general base for catalysis [186], based on both the fact that when deleted there is no activity, and by structural homology to other glycosyltransferases having the GT-B fold for which this role has been assigned [190]; Clarke *et al.* suggest the base may be a tyrosine residue, however, and a histidine residue may be important for stabilization of the leaving group [187]. Glycosyltransferases that catalyze transfer with inversion of stereochemistry are believed to use a dissociative S_N2 -like mechanism; an enzymic residue acts as a general base to deprotonate the nucleophile of the acceptor, which facilitates displacement of the activated nucleotide-phosphate leaving group. In a number of glycosyltransferases for which the three-dimensional structure has been elucidated, the general base is predicted to be a histidine, which interacts with an adjacent aspartate or glutamate residue [190]. Further studies will be required to clarify the identities of the catalytic residues and their roles in OGT. Notably, the structure of *XcGT41* also reveals an unusual and intimate relationship between the TPR domains and the C-terminal glycosyltransferase domain. The last TPR motif is atypical, which allows a large contact area with the enzymatic domain and orients the catalytic domain to enable the superhelical groove of the TPR domains to be continuous with the active site cleft [186,

187], suggesting that substrates binding to the TPR domains may be positioned within the groove to be exposed to the catalytic machinery.

INHIBITION OF OGT

While the design of potent and selective inhibitors for OGA has flourished in recent years, unfortunately the same cannot be said for OGT. However, the advent of the structure of a bacterial homologue of OGT may aid efforts. It has long been established that UDP, UDP-GlcNAc, and UTP are potent (around 200 nM) inhibitors of OGT [160] (Fig. 6). In addition, alloxan, a uracil analogue, was also demonstrated to be an OGT inhibitor, with complete inhibition at 1 mM and half of the maximum inhibition at 0.1 mM; the mechanism for inhibition is unclear, however, and alloxan may disrupt cysteine residues rather than bind in place of UDP-GlcNAc [191]. Sophisticated high throughput studies have been performed with the smaller splice variant of OGT, which can be over-expressed recombinantly in higher yield than the full length protein, in order to find potent inhibitors. Compounds were screened using a fluorescent UDP-GlcNAc analogue displacement assay, which led to the exciting discovery of three structurally unrelated compounds that inhibit OGT [192]. It is, as yet, unknown if these inhibitors act *in vivo*, but if so, they could prove to be a useful tool to decrease *O*-GlcNAc levels and study the physiological role of *O*-GlcNAc. These molecules are poorly water soluble but generation of derivatives having improved physical properties should be feasible. Others have explored a C1 phosphonate analogue of UDP-GlcNAc as an inhibitor of OGT, but this proved to be a poor inhibitor with an IC₅₀ of more than 5 mM, suggesting the presence of the glycosidic oxygen in UDP-GlcNAc is important [193]. Despite the poor affinity, this compound was successfully soaked into crystals of XcGT41; the solution of the structure in complex with the inhibitor compared to the apo-structure showed that a loop was positioned differently in each suggesting this loop may act as an 'active site lid' [187].

WHAT DOES THE FUTURE HOLD FOR THE *O*-GLcNAc PROCESSING ENZYMES?

This review has demonstrated that work involving the *O*-GlcNAc processing enzymes is highly active at present and reflects the research efforts of a number of groups. There is, however, still a number of unanswered questions. Despite the apparent simplicity of having only one enzyme for the addition and one enzyme for the removal of *O*-GlcNAc, the system of regulation and substrate specificity may be highly complex. These two enzymes are responsible for the *O*-GlcNAc cycling on hundreds of nucleocytoplasmic proteins, in the absence of a clear consensus sequence. Some regulatory mechanism is presumably in place to prevent futile *O*-GlcNAc cycling *in vivo*, but this is not yet understood, nor are the precise roles of the addition of this small sugar molecule to proteins. In addition it would be interesting to address why prokaryotes possess enzymes homologous to OGA and OGT, yet do not appear to have the *O*-GlcNAc modification of proteins; presumably they have gained (and retained) the genes encoding these proteins from their eukaryotic hosts for some, as yet unexplained, reasons.

The OGA field has flourished in recent years, with the deduction of the catalytic mechanism, the generation of highly potent and selective inhibitors, and the solution of the structures of two bacterial homologues of this enzyme. The question remains, however, of the role played by the *C*-terminal region of OGA which is absent in the bacterial sequences. Although it has been suggested this region has the ability to acetylate histones, this function has not been related to its OGA activity, and one cannot help but think it might also play another role; particularly since confirmation of this activity has proven elusive. In addition, the structures have not really aided in the knowledge of how the protein substrate is

recognised or binds to OGA, which may be useful for understanding how this enzyme is capable of modifying such a wide array of substrates. Ultimately the structure of the human OGA may address both of these questions.

Studies of OGT do lag behind those of OGA, which unfortunately reflects the general situation of glycosyltransferases *vs.* glycoside hydrolases. The solution of the structure of a bacterial homologue very recently, however, should now drive the field forward in the quest for a more detailed understanding of the catalytic mechanism and could aid design of potent inhibitors. Potent and selective inhibitors, such as those that have now been developed for OGA, will be an invaluable tool to use in cells to alter *O*-GlcNAc levels, which will ultimately aid elucidation of the way the enzyme works and is regulated.

Acknowledgments

T. M. G. is a Sir Henry Wellcome postdoctoral fellow and a Michael Smith for Health Research (MSFHR) trainee award holder. D. J. V. is a scholar of the MSFHR and holds a Canada Research Chair in Chemical Glycobiology.

ABBREVIATIONS

<i>O</i>-GlcNAc	<i>O</i> -linked 2-acetamido-2-deoxy- β -D-glucopyranose
OGA	<i>O</i> -GlcNAcase
OGT	uridine diphospho- <i>N</i> -acetylglucosamine:polypeptide β - <i>N</i> -acetylglucosaminyltransferase
CAZy	Carbohydrate Active enZyme
GH	glycoside hydrolase
GT	glycosyltransferase
BtGH84	GH84 enzyme from <i>Bacteroides thetaiotaomicron</i>
CpGH84	GH84 enzyme from <i>Clostridium perfringens</i>
STZ	streptozotocin or <i>N</i> -methyl- <i>N</i> -nitrosourea β -D-glucosamine
PUGNAc	<i>O</i> -(2-acetamido-2-deoxy- β -D-glucopyransylidene)-amino- <i>N</i> -phenylcarbamate
NAG-thiazoline	1,2-dideoxy-2'-methyl- α -D-glucopyranoso-[2,1- <i>d</i>]- Δ 2'-thiazoline
NButGT	1,2-dideoxy-2'-propyl- α -D-glucopyranoso-[2,1- <i>d</i>]- Δ 2'-thiazoline
Thiamet-G	1,2-dideoxy-2'-ethylamino- α -D-glucopyranoso-[2,1- <i>d</i>]- Δ 2'-thiazoline
TPR	tetratricopeptide repeat
np62	nucleoporin p62
OIP106	OGT-interacting protein, 106 kDa
OID	interaction domain of OIP106
UDP-GlcNAc	uridine diphospho- <i>N</i> -acetylglucosamine
XcGT41	GT41 enzyme from <i>Xanthomonas campestris</i>

REFERENCES

1. Torres CR, Hart GW. Topography and polypeptide distribution of terminal *N*-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for *O*-linked GlcNAc. J Biol Chem. 1984; 259:3308–17. [PubMed: 6421821]

2. Shafi RI, Iyer SPR, Ellies LG, 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 USA*. 2000; 97:5735–9. [PubMed: 10801981]
3. Gambetta MC, Oktaba K, Muller J. Essential role of the glycosyltransferase *sxc/Ogt* in polycomb repression. *Science*. 2009; 325:93–6. [PubMed: 19478141]
4. Sinclair DAR, Syrzycka M, Macauley MS, et al. *Drosophila* *O*GlcNAc transferase (OGT) is encoded by the Polycomb group (PcG) gene, super sex combs (*sxc*). *Proc Natl Acad Sci USA*. 2009; 106:13427–32. [PubMed: 19666537]
5. Hanover JA, Forsythe ME, Hennessey PT, et al. A *Caenorhabditis elegans* model of insulin resistance: Altered macronutrient storage and dauer formation in an OGT-1 knockout. *Proc Natl Acad Sci USA*. 2005; 102:11266–271. [PubMed: 16051707]
6. Forsythe ME, Love DC, Lazarus BD, 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 USA*. 2006; 103:11952–7. [PubMed: 16882729]
7. Wells L, Vosseller K, Hart GW. Glycosylation of nucleocytoplasmic proteins: Signal transduction and *O*-GlcNAc. *Science*. 2001; 291:2376–8. [PubMed: 11269319]
8. Hart GW, Housley MP, Slawson C. Cycling of *O*-linked beta-*N*-acetylglucosamine on nucleocytoplasmic proteins. *Nature*. 2007; 446:1017–22. [PubMed: 17460662]
9. Love DC, Hanover JA. The hexosamine signaling pathway: deciphering the “*O*-GlcNAc code”. *Sci STKE*. 2005; 2005:re13. [PubMed: 16317114]
10. Jackson SP, Tjian R. *O*-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. *Cell*. 1988; 55:125–33. [PubMed: 3139301]
11. Chou TY, Dang CV, Hart GW. Glycosylation of the c-Myc transactivation domain. *Proc Natl Acad Sci USA*. 1995; 92:4417–21. [PubMed: 7753821]
12. Walgren JL, Vincent TS, Schey KL, Buse MG. High glucose and insulin promote *O*-GlcNAc modification of proteins, including alpha-tubulin. *Am J Physiol Endocrinol Metab*. 2003; 284:E424–34. [PubMed: 12397027]
13. Hagmann J, Grob M, Burger MM. The cytoskeletal protein talin is *O*-glycosylated. *J Biol Chem*. 1992; 267:14424–8. [PubMed: 1629228]
14. Dong DL, Xu ZS, Hart GW, Cleveland DW. Cytoplasmic *O*-GlcNAc modification of the head domain and the KSP repeat motif of the neurofilament protein neurofilament-H. *J Biol Chem*. 1996; 271:20845–52. [PubMed: 8702840]
15. Dong DL, Xu ZS, Chevrier MR, et al. Glycosylation of mammalian neurofilaments. Localization of multiple *O*-linked *N*-acetylglucosamine moieties on neurofilament polypeptides L and M. *J Biol Chem*. 1993; 268:16679–87. [PubMed: 8344946]
16. Hanover JA, Cohen CK, Willingham MC, Park MK. *O*-linked *N*-acetylglucosamine is attached to proteins of the nuclear pore. Evidence for cytoplasmic and nucleoplasmic glycoproteins. *J Biol Chem*. 1987; 262:9887–94. [PubMed: 3110163]
17. Davis LI, Blobel G. Nuclear pore complex contains a family of glycoproteins that includes p62: glycosylation through a previously unidentified cellular pathway. *Proc Natl Acad Sci USA*. 1987; 84:7552–6. [PubMed: 3313397]
18. Holt GD, Snow CM, Senior A, et al. Nuclear pore complex glycoproteins contain cytoplasmically disposed *O*-linked *N*-acetylglucosamine. *J Cell Biol*. 1987; 104:1157–64. [PubMed: 3571327]
19. Kelly WG, Dahmus ME, Hart GW. RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by *O*-GlcNAc. *J Biol Chem*. 1993; 268:10416–24. [PubMed: 8486697]
20. Zhang F, Su K, Yang X, et al. *O*-GlcNAc modification is an endogenous inhibitor of the proteasome. *Cell*. 2003; 115:715–25. [PubMed: 14675536]
21. McClain DA, Lubas WA, Cooksey RC, et al. Altered glycan-dependent signaling induces insulin resistance and hyperleptinemia. *Proc Natl Acad Sci USA*. 2002; 99:10695–9. [PubMed: 12136128]
22. Yang X, Ongusaha PP, Miles PD, et al. Phosphoinositide signalling links *O*-GlcNAc transferase to insulin resistance. *Nature*. 2008; 451:964–9. [PubMed: 18288188]

23. Vosseller K, Wells L, Lane MD, Hart GW. Elevated nucleocytoplasmic glycosylation by *O*-GlcNAc results in insulin resistance associated with defects in Akt activation in 3T3-L1 adipocytes. *Proc Natl Acad Sci USA*. 2002; 99:5313–8. [PubMed: 11959983]
24. Macauley MS, Bubb AK, Martinez-Fleites C, Davies GJ, Vocadlo DJ. Elevation of global *O*-GlcNAc levels in 3T3-L1 adipocytes by selective inhibition of *O*-GlcNAcase does not induce insulin resistance. *J Biol Chem*. 2008; 283:34687–95. [PubMed: 18842583]
25. Buse MG. Hexosamines, insulin resistance, and the complications of diabetes: current status. *Am J Physiol Endocrinol Metab*. 2006; 290:E1–E8. [PubMed: 16339923]
26. Robinson KA, Ball LE, Buse MG. Reduction of *O*-GlcNAc protein modification does not prevent insulin resistance in 3T3-L1 adipocytes. *Am J Physiol Endocrinol Metab*. 2007; 292:E884–90. [PubMed: 17122093]
27. Liu F, Iqbal K, Grundke-Iqbal I, Hart GW, Gong CX. *O*-GlcNAcylation regulates phosphorylation of tau: a mechanism involved in Alzheimer's disease. *Proc Natl Acad Sci USA*. 2004; 101:10804–9. [PubMed: 15249677]
28. Yang WH, Kim JE, Nam HW, et al. Modification of p53 with *O*-linked *N*-acetylglucosamine regulates p53 activity and stability. *Nat Cell Biol*. 2006; 8:1074–83. [PubMed: 16964247]
29. Slawson C, Housley MP, Hart GW. *O*-GlcNAc cycling: how a single sugar post-translational modification is changing the way we think about signaling networks. *J Cell Biochem*. 2006; 97:71–83. [PubMed: 16237703]
30. Wang Z, Gucek M, Hart GW. Cross-talk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated *O*-GlcNAc. *Proc Natl Acad Sci USA*. 2008; 105:13793–8. [PubMed: 18779572]
31. Kamemura K, Hayes BK, Comer FI, Hart GW. Dynamic interplay between *O*-glycosylation and *O*-phosphorylation of nucleocytoplasmic proteins: alternative glycosylation/phosphorylation of THR-58, a known mutational hot spot of c-Myc in lymphomas, is regulated by mitogens. *J Biol Chem*. 2002; 277:19229–35. [PubMed: 11904304]
32. Rechsteiner M, Rogers SW. PEST sequences and regulation by proteolysis. *Trends Biochem Sci*. 1996; 21:267–71. [PubMed: 8755249]
33. Vosseller K, Trinidad JC, Chalkley RJ, et al. *O*-linked *N*-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Mol Cell Proteomics*. 2006; 5:923–34. [PubMed: 16452088]
34. Haltiwanger RS, Holt GD, Hart GW. 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; 265:2563–8. [PubMed: 2137449]
35. Lubas WA, Frank DW, Krause M, Hanover JA. *O*-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. *J Biol Chem*. 1997; 272:9316–24. [PubMed: 9083068]
36. Braidman I, Carroll M, Dance N, et al. Characterisation of human *N*-acetyl-beta-hexosaminidase C. *FEBS Lett*. 1974; 41:181–4. [PubMed: 4859245]
37. Gao Y, Wells L, Comer FI, Parker GJ, Hart GW. 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; 276:9838–45. [PubMed: 11148210]
38. Walsh CT, Garneau-Tsodikova S, Gatto GJ Jr. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chem Int Ed Engl*. 2005; 44:7342–72. [PubMed: 16267872]
39. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science*. 2002; 298:1912–34. [PubMed: 12471243]
40. Poenaru L, Dreyfus JC. Electrophoretic study of hexosaminidases. Hexosaminidase C. *Clin Chim Acta*. 1973; 43:439–42. [PubMed: 4690914]
41. Swallow DM, Evans L, Saha N, Harris H. Characterization and tissue distribution of *N*-acetyl hexosaminidase C: suggestive evidence for a separate hexosaminidase locus. *Ann Hum Genet*. 1976; 40:55–66. [PubMed: 9025]

42. Izumi T, Suzuki K. Neutral hydrolases of rat brain. Preliminary characterization and developmental changes of neutral beta-*N*-acetylhexosaminidases. *Biochim Biophys Acta*. 1980; 615:402–13. [PubMed: 6448075]
43. Overdijk B, van der Kroef WM, van Steijn GJ, Lisman JJ. Isolation and further characterization of bovine brain hexosaminidase C. *Biochim Biophys Acta*. 1981; 659:255–66. [PubMed: 7260095]
44. Wells L, Gao Y, Mahoney JA, et al. Dynamic *O*-glycosylation of nuclear and cytosolic proteins: further characterization of the nucleocytoplasmic beta-*N*-acetylglucosaminidase, *O*-GlcNAcase. *J Biol Chem*. 2002; 277:1755–61. [PubMed: 11788610]
45. Heckel D, Comtesse N, Brass N, et al. Novel immunogenic antigen homologous to hyaluronidase in meningioma. *Hum Mol Genet*. 1998; 7:1859–72. [PubMed: 9811929]
46. Dong DL, Hart GW. Purification and characterization of an *O*-GlcNAc selective *N*-acetyl-beta-d-glucosaminidase from rat spleen cytosol. *J Biol Chem*. 1994; 269:19321–30. [PubMed: 8034696]
47. Mahuran DJ. Biochemical consequences of mutations causing the GM2 gangliosidosis. *Biochim Biophys Acta*. 1999; 1455:105–38. [PubMed: 10571007]
48. Izumi T, Suzuki K. Neutral beta-*N*-acetylhexosaminidases of rat brain. Purification and enzymatic and immunological characterization. *J Biol Chem*. 1983; 258:6991–9. [PubMed: 6222051]
49. Overdijk B, van der Kroef WM, Veltkamp WA, Hooghwinkel GJ. The separation of bovine brain beta-*N*-acetyl-d-hexosaminidases. Abnormal gel-filtration behaviour of beta-*N*-acetyl-d-glucosaminidase C. *Biochem J*. 1975; 151:257–61. [PubMed: 1240755]
50. Gutternigg M, Rendic D, Voglauer R, Iskratsch T, Wilson IB. Mammalian cells contain a second nucleocytoplasmic hexosaminidase. *Biochem J*. 2009; 419:83–90. [PubMed: 19040401]
51. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active enZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res*. 2008; 37:D233–8. [PubMed: 18838391]
52. Henrissat B, Davies G. Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol*. 1997; 7:637–44. [PubMed: 9345621]
53. Tews I, Perrakis A, Oppenheim A, Dauter Z, Wilson KS, Vorgias CE. Bacterial chitinase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. *Nat Struct Biol*. 1996; 3:638–48. [PubMed: 8673609]
54. Mark BL, Vocadlo DJ, Knapp S, Triggs-Raine BL, Withers SG, James MN. Crystallographic evidence for substrate-assisted catalysis in a bacterial beta-hexosaminidase. *J Biol Chem*. 2001; 276:10330–7. [PubMed: 11124970]
55. Lemieux MJ, Mark BL, Cherney MM, Withers SG, Mahuran DJ, James MN. Crystallographic structure of human beta-hexosaminidase A: interpretation of Tay-Sachs mutations and loss of GM2 ganglioside hydrolysis. *J Mol Biol*. 2006; 359:913–29. [PubMed: 16698036]
56. Mark BL, Mahuran DJ, Cherney MM, Zhao D, Knapp S, James MN. Crystal structure of human beta-hexosaminidase B: understanding the molecular basis of Sandhoff and Tay-Sachs disease. *J Mol Biol*. 2003; 327:1093–109. [PubMed: 12662933]
57. Koshland DE Jr. Stereochemistry and the mechanism of enzymatic reactions. *Biol Rev*. 1953; 28:416–36.
58. Vocadlo DJ, Davies GJ. Mechanistic insights into glycosidase chemistry. *Curr Opin Chem Biol*. 2008; 12:539–55. [PubMed: 18558099]
59. Comtesse N, Maldener E, Meese E. Identification of a nuclear variant of MGEA5, a cytoplasmic hyaluronidase and a beta-*N*-acetylglucosaminidase. *Biochem Biophys Res Commun*. 2001; 283:634–40. [PubMed: 11341771]
60. Rigden DJ, Jedrzejewski MJ, de Mello LV. Identification and analysis of catalytic TIM barrel domains in seven further glycoside hydrolase families. *FEBS Lett*. 2003; 544:103–11. [PubMed: 12782298]
61. Hanover JA. Glycan-dependent signaling: *O*-linked *N*-acetylglucosamine. *FASEB J*. 2001; 15:1865–76. [PubMed: 11532966]
62. Schultz J, Pils B. Prediction of structure and functional residues for *O*-GlcNAcase, a divergent homologue of acetyltransferases. *FEBS Lett*. 2002; 529:179–82. [PubMed: 12372596]
63. Cetinbas N, Macauley MS, Stubbs KA, Drapala R, Vocadlo DJ. Identification of Asp174 and Asp175 as the key catalytic residues of human *O*-GlcNAcase by functional analysis of site-directed mutants. *Biochemistry*. 2006; 45:3835–44. [PubMed: 16533067]

64. Toleman C, Paterson AJ, Kudlow JE. Location and characterization of the *O*-GlcNAcase active site. *Biochim Biophys Acta*. 2006; 1760:829–39. [PubMed: 16517082]
65. Kim EJ, Kang DO, Love DC, Hanover JA. Enzymatic characterization of *O*-GlcNAcase isoforms using a fluorogenic GlcNAc substrate. *Carbohydr Res*. 2006; 341:971–82. [PubMed: 16584714]
66. Macauley MS, Vocadlo DJ. Enzymatic characterization and inhibition of the nuclear variant of human *O*-GlcNAcase. *Carbohydr Res*. 2009; 344:1079–84. [PubMed: 19423084]
67. Toleman C, Paterson AJ, Whisenhunt TR, Kudlow JE. Characterization of the histone acetyltransferase (HAT) domain of a bifunctional protein with activable *O*-GlcNAcase and HAT activities. *J Biol Chem*. 2004; 279:53665–73. [PubMed: 15485860]
68. Butkinaree C, Cheung WD, Park S, Park K, Barber M, Hart GW. Characterization of beta-*N*-acetylglucosaminidase cleavage by caspase-3 during apoptosis. *J Biol Chem*. 2008; 283:23557–66. [PubMed: 18586680]
69. Wolf BB, Green DR. Suicidal tendencies: apoptotic cell death by caspase family proteinases. *J Biol Chem*. 1999; 274:20049–52. [PubMed: 10400609]
70. Macauley MS, Whitworth GE, Debowski AW, Chin D, Vocadlo DJ. *O*-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors. *J Biol Chem*. 2005; 280:25313–22. [PubMed: 15795231]
71. Sheldon WL, Macauley MS, Taylor EJ, et al. Functional analysis of a group A streptococcal glycoside hydrolase Spy1600 from family 84 reveals it is a beta-*N*-acetylglucosaminidase and not a hyaluronidase. *Biochem J*. 2006; 399:241–7. [PubMed: 16822234]
72. Macauley MS, Stubbs KA, Vocadlo DJ. *O*-GlcNAcase catalyzes cleavage of thioglycosides without general acid catalysis. *J Am Chem Soc*. 2005; 127:17202–3. [PubMed: 16332065]
73. Vocadlo DJ, Davies GJ, Laine R, Withers SG. Catalysis by hen egg-white lysozyme proceeds *via* a covalent intermediate. *Nature*. 2001; 412:835–8. [PubMed: 11518970]
74. Vocadlo DJ, Mayer C, He S, Withers SG. Mechanism of action and identification of Asp242 as the catalytic nucleophile of *Vibrio furnisii* *N*-acetyl-beta-d-glucosaminidase using 2-acetamido-2-deoxy-5-fluoro-alpha-L-idopyranosyl fluoride. *Biochemistry*. 2000; 39:117–26. [PubMed: 10625486]
75. Mayer C, Vocadlo DJ, Mah M, et al. Characterization of a beta-*N*-acetylhexosaminidase and a beta-*N*-acetylglucosaminidase/beta-glucosidase from *Cellulomonas fimi*. *FEBS J*. 2006; 273:2929–41. [PubMed: 16762038]
76. Berti P, McCann JAB. Toward a detailed understanding of base excision repair enzymes: Transition state and mechanistic analyses of *N*-glycoside hydrolysis and *N*-glycoside transfer. *Chem Rev*. 2006; 106:506–55. [PubMed: 16464017]
77. Terwisscha van Scheltinga AC, Armand S, Kalk KH, Isogai A, Henrissat B, Dijkstra BW. Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and X-ray structure of a complex with allosamidin: evidence for substrate assisted catalysis. *Biochemistry*. 1995; 34:15619–23. [PubMed: 7495789]
78. van Aalten DM, Komander D, Synstad B, Gåseidnes S, Peter MG, Eijssink VG. Structural insights into the catalytic mechanism of a family 18 exo-chitinase. *Proc Natl Acad Sci USA*. 2001; 98:8979–84. [PubMed: 11481469]
79. Tews I, Terwisscha van Scheltinga AC, Perrakis A, Wilson KS, Dijkstra B. Substrate-assisted catalysis unifies two families of chitinolytic enzymes. *J Am Chem Soc*. 1997; 119:7954–9.
80. Terwisscha van Scheltinga AC, Hennig M, Dijkstra BW. The 1.8 Å resolution structure of hevamine, a plant chitinase/lysozyme, and analysis of the conserved sequence and structure motifs of glycosyl hydrolase family 18. *J Mol Biol*. 1996; 262:243–57. [PubMed: 8831791]
81. Drouillard S, Armand S, Davies GJ, Vorgias CE, Henrissat B. *Serratia marcescens* chitobiase is a retaining glycosidase utilizing substrate acetamido group participation. *Biochem J*. 1997; 328:945–9. [PubMed: 9396742]
82. Markovi -Housley Z, Miglierini G, Soldatova L, Rizkallah PJ, Müller U, Schirmer T. Crystal structure of hyaluronidase, a major allergen of bee venom. *Structure*. 2000; 8:1025–35. [PubMed: 11080624]

83. Fujita M, Shoda S, Haneda K, Inazu T, Takegawa K, Yamamoto K. A novel disaccharide substrate having 1,2-oxazoline moiety for detection of transglycosylating activity of endoglycosidases. *Biochim Biophys Acta*. 2001; 1528:9–14. [PubMed: 11514092]
84. Umekawa M, Huang W, Li B, et al. Mutants of *Mucor hiemalis* endo-beta-N-acetylglucosaminidase show enhanced transglycosylation and glycosynthase-like activities. *J Biol Chem*. 2008; 283:4469–79. [PubMed: 18096701]
85. Abbott DW, Macauley MS, Vocadlo DJ, Boraston AB. *Streptococcus pneumoniae* endohexosaminidase D: Structural and mechanistic insight into substrate-assisted catalysis in family 85 glycoside hydrolases. *J Biol Chem*. 2009; 284:11676–689. [PubMed: 19181667]
86. Martinez-Fleites C, Korczynska JE, Davies GJ, Cope MJ, Turkenburg JP, Taylor EJ. The crystal structure of a family GH25 lysozyme from *Bacillus anthracis* implies a neighboring-group catalytic mechanism with retention of anomeric configuration. *Carbohydr Res*. 2009; 344:1753–7. [PubMed: 19595298]
87. Greig IR, Zahariev F, Withers SG. Elucidating the nature of the *Streptomyces plicatus* beta-hexosaminidase-bound intermediate using *ab initio* molecular dynamics simulations. *J Am Chem Soc*. 2008; 130:17620–8. [PubMed: 19053474]
88. Lu Y, Zen K-C, Muthukrishnan S, Kramer KJ. Site-directed mutagenesis and functional analysis of active site acidic amino acid residues D142, D144 and E146 in *Manduca sexta* (tobacco hornworm) chitinase. *Insect Biochem Mol Biol*. 2002; 32:1369–82. [PubMed: 12530205]
89. Arming S, Strobl B, Wechselberger C, Kreil G. *In vitro* mutagenesis of PH-20 hyaluronidase from human sperm. *Eur J Biochem*. 1997; 247:810–4. [PubMed: 9288901]
90. Ling Z, Suits MD, Bingham RJ, et al. The X-ray crystal structure of an *Arthrobacter protophormiae* endo-beta-N-acetylglucosaminidase reveals a (beta/alpha)(8) catalytic domain, two ancillary domains and active site residues key for transglycosylation activity. *J Mol Biol*. 2009; 389:1–9. [PubMed: 19327363]
91. Hou Y, Vocadlo DJ, Leung A, Withers SG, Mahuran D. Characterization of the Glu and Asp residues in the active site of human beta-hexosaminidase B. *Biochemistry*. 2001; 40:2201–9. [PubMed: 11329289]
92. Williams SJ, Mark BL, Vocadlo DJ, James MN, Withers SG. Aspartate 313 in the *Streptomyces plicatus* hexosaminidase plays a critical role in substrate-assisted catalysis by orienting the 2-acetamido group and stabilizing the transition state. *J Biol Chem*. 2002; 277:40055–65. [PubMed: 12171933]
93. Prag G, Papanikolaou Y, Tavlas G, Vorgias CE, Petratos K, Oppenheim AB. Structures of chitinase mutants complexed with the substrate Di-N-acetyl-d-glucosamine: the catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540. *J Mol Biol*. 2000; 300:611–7. [PubMed: 10884356]
94. MacLeod AM, Lindhorst T, Withers SG, Warren RA. The acid/base catalyst in the exoglucanase/xylanase from *Cellulomonas fimi* is glutamic acid 127: evidence from detailed kinetic studies of mutants. *Biochemistry*. 1994; 33:6371–6. [PubMed: 7910761]
95. Viladot JL, de Ramon E, Durany O, Planas A. Probing the mechanism of *Bacillus* 1,3-1,4-beta-d-glucan 4-glucanohydrolases by chemical rescue of inactive mutants at catalytically essential residues. *Biochemistry*. 1998; 37:11332–42. [PubMed: 9698381]
96. Zechel DL, Withers SG. Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Acc Chem Res*. 2000; 33:11–8. [PubMed: 10639071]
97. Dennis RJ, Taylor EJ, Macauley MS, et al. Structure and mechanism of a bacterial beta-glucosaminidase having O-GlcNAcase activity. *Nat Struct Mol Biol*. 2006; 13:365–71. [PubMed: 16565725]
98. Rao FV, Dorfmueller HC, Villa F, Allwood M, Eggleston IM, van Aalten DM. Structural insights into the mechanism and inhibition of eukaryotic O-GlcNAc hydrolysis. *EMBO J*. 2006; 25:1569–78. [PubMed: 16541109]
99. Ficko-Blean E, Boraston AB. Cloning, recombinant production, crystallization and preliminary X-ray diffraction studies of a family 84 glycoside hydrolase from *Clostridium perfringens*. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2005; 61:834–6.

100. Herr R, Jahnke HK, Argoudelis AD. Structure of streptozotocin. *J Am Chem Soc.* 1967; 89:4808–9. [PubMed: 6074800]
101. Horsch M, Hoesch L, Vasella A, Rast DM. *N*-acetylglucosaminono-1,5-lactone oxime and the corresponding (phenylcarbamoyl)oxime. Novel and potent inhibitors of beta-*N*-acetylglucosaminidase. *Eur J Biochem.* 1991; 197:815–8. [PubMed: 2029909]
102. Roos MD, Xie W, Su K, et al. Streptozotocin, an analog of *N*-acetylglucosamine, blocks the removal of *O*-GlcNAc from intracellular proteins. *Proc Assoc Am Physicians.* 1998; 110:422–32. [PubMed: 9756093]
103. Hanover JA, Lai Z, Lee G, Lubas WA, Sato SM. Elevated *O*-linked *N*-acetylglucosamine metabolism in pancreatic beta-cells. *Arch Biochem Biophys.* 1999; 362:38–45. [PubMed: 9917327]
104. Mansford KR, Opie L. Comparison of metabolic abnormalities in diabetes mellitus induced by streptozotocin or by alloxan. *Lancet.* 1968; 1:670–1. [PubMed: 4170654]
105. O'Brien BA, Harmon BV, Cameron DP, Allan DJ. Beta-cell apoptosis is responsible for the development of IDDM in the multiple low-dose streptozotocin model. *J Pathol.* 1996; 178:176–81. [PubMed: 8683386]
106. Saini KS, Thompson C, Winterford CM, Walker NI, Cameron DP. Streptozotocin at low doses induces apoptosis and at high doses causes necrosis in a murine pancreatic beta cell line, INS-1. *Biochem Mol Biol Int.* 1996; 39:1229–36. [PubMed: 8876977]
107. Liu K, Paterson AJ, Chin E, Kudlow JE. Glucose stimulates protein modification by *O*-linked GlcNAc in pancreatic beta cells: linkage of *O*-linked GlcNAc to beta cell death. *Proc Natl Acad Sci USA.* 2000; 97:2820–5. [PubMed: 10717000]
108. Konrad RJ, Mikolaenko I, Tolar JF, Liu K, Kudlow JE. The potential mechanism of the diabetogenic action of streptozotocin: inhibition of pancreatic beta-cell *O*-GlcNAc-selective *N*-acetyl-beta-d-glucosaminidase. *Biochem J.* 2001; 356:31–41. [PubMed: 11336633]
109. Toleman C, Paterson AJ, Shin R, Kudlow JE. Streptozotocin inhibits *O*-GlcNAcase *via* the production of a transition state analog. *Biochem Biophys Res Commun.* 2006; 340:526–34. [PubMed: 16376298]
110. Turk J, Corbett JA, Ramanadham S, Bohrer A, McDaniel ML. Biochemical evidence for nitric oxide formation from streptozotocin in isolated pancreatic islets. *Biochem Biophys Res Commun.* 1993; 197:1458–64. [PubMed: 7904159]
111. Kaneto H, Fujii J, Seo HG, et al. Apoptotic cell death triggered by nitric oxide in pancreatic beta-cells. *Diabetes.* 1995; 44:733–8. [PubMed: 7540572]
112. Bennett RA, Pegg AE. Alkylation of DMA in rat tissues following administration of streptozotocin. *Cancer Res.* 1981; 41:2786–90. [PubMed: 6454479]
113. Bolzan AD, Bianchi MS. Genotoxicity of streptozotocin. *Mutat Res.* 2002; 512:121–34. [PubMed: 12464347]
114. Gao Y, Parker GJ, Hart GW. Streptozotocin-induced beta-cell death is independent of its inhibition of *O*-GlcNAcase in pancreatic Min6 cells. *Arch Biochem Biophys.* 2000; 383:296–302. [PubMed: 11185566]
115. Okuyama R, Yachi M. Cytosolic *O*-GlcNAc accumulation is not involved in beta-cell death in HIT-T15 or Min6. *Biochem Biophys Res Commun.* 2001; 287:366–71. [PubMed: 11554736]
116. Pathak S, Dorfmueller HC, Borodkin VS, van Aalten DM. Chemical dissection of the link between streptozotocin, *O*-GlcNAc, and pancreatic cell death. *Chem Biol.* 2008; 15:799–807. [PubMed: 18721751]
117. He Y, Martinez-Fleites C, Bubb A, Gloster TM, Davies GJ. Structural insight into the mechanism of streptozotocin inhibition of *O*-GlcNAcase. *Carbohydr Res.* 2009; 344:627–31. [PubMed: 19217614]
118. Haltiwanger RS, Grove K, Philipsberg GA. 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; 273:3611–7. [PubMed: 9452489]

119. Arias EB, Kim J, Cartee GD. Prolonged incubation in PUGNAc results in increased protein *O*-linked glycosylation and insulin resistance in rat skeletal muscle. *Diabetes*. 2004; 53:921–30. [PubMed: 15047606]
120. Nojiri H, Stroud M, Hakomori S. A specific type of ganglioside as a modulator of insulin-dependent cell growth and insulin receptor tyrosine kinase activity. Possible association of ganglioside-induced inhibition of insulin receptor function and monocytic differentiation induction in HL-60 cells. *J Biol Chem*. 1991; 266:4531–7. [PubMed: 1999434]
121. Ficko-Blean E, Stubbs KA, Nemirovsky O, Vocadlo DJ, Boraston AB. Structural and mechanistic insight into the basis of mucopolysaccharidosis IIIB. *Proc Natl Acad Sci USA*. 2008; 105:6560–5. [PubMed: 18443291]
122. Slawson C, Zachara NE, Vosseller K, Cheung WD, Lane MD, Hart GW. Perturbations in *O*-linked beta-*N*-acetylglucosamine protein modification cause severe defects in mitotic progression and cytokinesis. *J Biol Chem*. 2005; 280:32944–56. [PubMed: 16027160]
123. Zou L, Yang S, Hu S, Chaudry IH, Marchase RB, Chatham JC. The protective effects of PUGNAc on cardiac function after trauma-hemorrhage are mediated *via* increased protein *O*-GlcNAc levels. *Shock*. 2007; 27:402–8. [PubMed: 17414423]
124. Knapp S, Vocadlo DJ, Gao Z, et al. NAG-thiazoline, an *N*-acetyl- β -hexosaminidase inhibitor that implicates acetamido participation. *J Am Chem Soc*. 1996; 118:6804–5.
125. Vocadlo DJ, Hang HC, Kim EJ, Hanover JA, Bertozzi CR. A chemical approach for identifying *O*-GlcNAc-modified proteins in cells. *Proc Natl Acad Sci USA*. 2003; 100:9116–21. [PubMed: 12874386]
126. Whitworth GE, Macauley MS, Stubbs KA, et al. Analysis of PUGNAc and NAG-thiazoline as transition state analogues for human *O*-GlcNAcase: mechanistic and structural insights into inhibitor selectivity and transition state poise. *J Am Chem Soc*. 2007; 129:635–44. [PubMed: 17227027]
127. Kim EJ, Ferreira M, Thomas CJ, Hanover JA. An *O*-GlcNAcase-specific inhibitor and substrate engineered by the extension of the *N*-acetyl moiety. *J Am Chem Soc*. 2006; 128:4234–5. [PubMed: 16568991]
128. Stubbs KA, Zhang N, Vocadlo DJ. A divergent synthesis of 2-acyl derivatives of PUGNAc yields selective inhibitors of *O*-GlcNAcase. *Org Biomol Chem*. 2006; 4:839–45. [PubMed: 16493467]
129. Knapp S, Abdo M, Ajayi K, et al. Tautomeric modification of GlcNAc-thiazoline. *Org Lett*. 2007; 9:2321–4. [PubMed: 17508759]
130. Yuzwa SA, Macauley MS, Heinonen JE, et al. A potent mechanism-inspired *O*-GlcNAcase inhibitor that blocks phosphorylation of tau *in vivo*. *Nat Chem Biol*. 2008; 4:483–90. [PubMed: 18587388]
131. Aoyagi T, Suda H, Uotani K, et al. Nagstatin, a new inhibitor of *N*-acetyl-beta-d-glucosaminidase, produced by *Streptomyces amakusaensis* MG846-ff3. Taxonomy, production, isolation, physico-chemical properties and biological activities. *J Antibiot*. 1992; 45:1404–8. [PubMed: 1429224]
132. Tatsuta K, Miura S. Total synthesis of nagstatin, an *N*-acetyl-beta-d-glucosaminidase inhibitor. *Tetrahedron Lett*. 1995; 36:6721–24.
133. Tatsuta K, Miura S, Ohta S, Gunji H. Total syntheses of debranched nagstatin and its analogs having glycosidase inhibiting activities. *Tetrahedron Lett*. 1995; 36:1085–8.
134. Granier T, Panday N, Vasella A. Structure-activity relations for imidazo-pyridine-type inhibitors of β -d-glucosidases. *Helv Chim Acta*. 1997; 80:979–87.
135. Panday N, Vasella A. Synthesis of glucose- and mannose-derived *N*-acetylamino imidazopyridines and their evaluation as inhibitors of glycosidases. *Synthesis*. 1999; SI:1459–68.
136. Terinek M, Vasella A. Synthesis of tetrahydropyridoimidazole-2-acetates: effect of carboxy and methoxycarbonyl groups at C(2) on the inhibition of some β - and α -glycosidases. *Helv Chim Acta*. 2004; 87:3035–49.
137. Panday N, Canac Y, Vasella A. Very strong inhibition of glucosidases by C(2)-substituted tetrahydroimidazopyridines. *Helv Chim Acta*. 2000; 83:58–79.

138. Shanmugasundaram B, Vasella A. Synthesis of new *C*(2)-substituted *gluco*-configured tetrahydroimidazopyridines and their evaluation as glucosidase inhibitors. *Helv Chim Acta*. 2005; 88:2593–602.
139. Gloster TM, Meloncelli P, Stick RV, Zechel D, Vasella A, Davies GJ. Glycosidase inhibition: an assessment of the binding of eighteen putative transition-state mimics. *J Am Chem Soc*. 2007; 129:2345–54. [PubMed: 17279749]
140. Heightman TD, Vasella AT. Recent insights into inhibition, structure, and mechanism of configuration-retaining glycosidases. *Angew Chem Int Ed Engl*. 1999; 38:750–70.
141. Shanmugasundaram B, Debowski AW, Dennis RJ, Davies GJ, Vocadlo DJ, Vasella A. Inhibition of O-GlcNAcase by a gluco-configured nagstatin and a PUGNAc-imidazole hybrid inhibitor. *Chem Commun*. 2006; (42):4372–4.
142. Dorfmueller HC, Borodkin VS, Schimpl M, Shepherd SM, Shpiro NA, van Aalten DM. GlcNAcstatin: a picomolar, selective O-GlcNAcase inhibitor that modulates intracellular O-GlcNAcylation levels. *J Am Chem Soc*. 2006; 128:16484–5. [PubMed: 17177381]
143. Dorfmueller HC, Borodkin VS, Schimpl M, van Aalten DM. GlcNAcstatins are nanomolar inhibitors of human O-GlcNAcase inducing cellular hyper-O-GlcNAcylation. *Biochem J*. 2009; 420:221–7. [PubMed: 19275764]
144. Lee TN, Alborn WE, Knierman MD, Konrad RJ. Alloxan is an inhibitor of O-GlcNAc-selective *N*-acetyl-beta-d-glucosaminidase. *Biochem Biophys Res Commun*. 2006; 350:1038–43. [PubMed: 17045574]
145. Lenzen S. The mechanisms of alloxan- and streptozotocin-induced diabetes. *Diabetologia*. 2008; 51:216–26. [PubMed: 18087688]
146. Kim EJ, Amorelli B, Abdo M, et al. Distinctive inhibition of O-GlcNAcase isoforms by an alpha-GlcNAc thiolsulfonate. *J Am Chem Soc*. 2007; 129:14854–5. [PubMed: 17994748]
147. Marcelo F, He Y, Yuzwa SA, et al. Molecular basis for inhibition of GH84 glycoside hydrolases by substituted azepanes: Conformational flexibility enables probing of substrate distortion. *J Am Chem Soc*. 2009; 131:5390–2. [PubMed: 19331390]
148. Scaffidi A, Stubbs KA, Dennis RJ, et al. A 1-acetamido derivative of 6-epi-valienamine: an inhibitor of a diverse group of beta-*N*-acetylglucosaminidases. *Org Biomol Chem*. 2007; 5:3013–9. [PubMed: 17728868]
149. Tailford LE, Offen WA, Smith NL, et al. Structural and biochemical evidence for a boat-like transition state in beta-mannosidases. *Nat Chem Biol*. 2008; 4:306–12. [PubMed: 18408714]
150. Wicki J, Williams SJ, Withers SG. Transition-state mimicry by glycosidase inhibitors: a critical kinetic analysis. *J Am Chem Soc*. 2007; 129:4530–1. [PubMed: 17385869]
151. Pauling L. Molecular architecture and biological reactions. *Chem Eng News*. 1946; 24:1375–7.
152. Pauling L. Nature of forces between large molecules of biological interest. *Nature*. 1948; 161:707–9. [PubMed: 18860270]
153. Wolfenden R, Lu X, Young G. Spontaneous hydrolysis of glycosides. *J Am Chem Soc*. 1998; 120:6814–5.
154. Wolfenden R, Snider MJ. The depth of chemical time and the power of enzymes as catalysts. *Acc Chem Res*. 2001; 34:938–45. [PubMed: 11747411]
155. Davies, GJ.; Sinnott, ML.; Withers, SG. Glycosyl transfer. In: Sinnott, ML., editor. Comprehensive biological catalysis. A mechanistic reference. Academic Press; London: 1998. p. 119-209.
156. Sinnott ML. Catalytic mechanisms of enzymic glycosyl transfer. *Chem Rev*. 1990; 90:1171–202.
157. Mader MM, Bartlett PA. Binding energy and catalysis: the implications for transition-state analogs and catalytic antibodies. *Chem Rev*. 1997; 97:1281–301. [PubMed: 11851452]
158. Greig IR, Williams IH. Glycosidase inhibitors as conformational transition state analogues. *Chem Commun*. 2007; (36):3747–9.
159. Lameira J, Alves CN, Moliner V, Martí S, Kanaan N, Tuñón I. A quantum mechanics/molecular mechanics study of the protein-ligand interaction of two potent inhibitors of human O-GlcNAcase: PUGNAc and NAG-thiazoline. *J Phys Chem B*. 2008; 112:14260–6. [PubMed: 18939790]

160. Haltiwanger RS, Blomberg MA, Hart GW. Glycosylation of nuclear and cytoplasmic proteins. Purification and characterization of a uridine diphospho-*N*-acetylglucosamine:polypeptide beta-*N*-acetylglucosaminyltransferase. *J Biol Chem.* 1992; 267:9005–13. [PubMed: 1533623]
161. Kreppel LK, Blomberg MA, Hart GW. Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique *O*-GlcNAc transferase with multiple tetratricopeptide repeats. *J Biol Chem.* 1997; 272:9308–15. [PubMed: 9083067]
162. Nolte D, Muller U. Human *O*-GlcNAc transferase (OGT): genomic structure, analysis of splice variants, fine mapping in Xq13.1. *Mamm Genome.* 2002; 13:62–4. [PubMed: 11773972]
163. Hanover JA, Yu S, Lubas WB, et al. Mitochondrial and nucleocytoplasmic isoforms of *O*-linked GlcNAc transferase encoded by a single mammalian gene. *Arch Biochem Biophys.* 2003; 409:287–97. [PubMed: 12504895]
164. Love DC, Kochan J, Cathey RL, Shin SH, Hanover JA. Mitochondrial and nucleocytoplasmic targeting of *O*-linked GlcNAc transferase. *J Cell Sci.* 2003; 116:647–54. [PubMed: 12538765]
165. Kreppel LK, Hart GW. Regulation of a cytosolic and nuclear *O*-GlcNAc transferase. Role of the tetratricopeptide repeats. *J Biol Chem.* 1999; 274:32015–22. [PubMed: 10542233]
166. Lubas WA, Hanover JA. Functional expression of *O*-linked GlcNAc transferase. Domain structure and substrate specificity. *J Biol Chem.* 2000; 275:10983–8. [PubMed: 10753899]
167. Blatch GL, Lassle M. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays.* 1999; 21:932–9. [PubMed: 10517866]
168. D'Andrea LD, Regan L. TPR proteins: the versatile helix. *Trends Biochem Sci.* 2003; 28:655–62. [PubMed: 14659697]
169. Yang X, Zhang F, Kudlow JE. Recruitment of *O*-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein *O*-GlcNAcylation to transcriptional repression. *Cell.* 2002; 110:69–80. [PubMed: 12150998]
170. Lamb JR, Tugendreich S, Hieter P. Tetratricopeptide repeat interactions: to TPR or not to TPR? *Trends Biochem Sci.* 1995; 20:257–9. [PubMed: 7667876]
171. Iyer SP, Hart GW. Roles of the tetratricopeptide repeat domain in *O*-GlcNAc transferase targeting and protein substrate specificity. *J Biol Chem.* 2003; 278:24608–16. [PubMed: 12724313]
172. Wells L, Kreppel LK, Comer FI, Wadzinski BE, Hart GW. *O*-GlcNAc transferase is in a functional complex with protein phosphatase 1 catalytic subunits. *J Biol Chem.* 2004; 279:38466–70. [PubMed: 15247246]
173. Cheung WD, Sakabe K, Housley MP, Dias WB, Hart GW. *O*-linked beta-*N*-acetylglucosaminyltransferase substrate specificity is regulated by myosin phosphatase targeting and other interacting proteins. *J Biol Chem.* 2008; 283:33935–41. [PubMed: 18840611]
174. Andrali SS, Marz P, Ozcan S. Ataxin-10 interacts with *O*-GlcNAc transferase OGT in pancreatic beta cells. *Biochem Biophys Res Commun.* 2005; 337:149–53. [PubMed: 16182253]
175. Housley MP, Udeshi ND, Rodgers JT, et al. A PGC-1alpha-*O*-GlcNAc transferase complex regulates FoxO transcription factor activity in response to glucose. *J Biol Chem.* 2009; 284:5148–57. [PubMed: 19103600]
176. Leavy TM, Bertozzi CR. A high-throughput assay for *O*-GlcNAc transferase detects primary sequence preferences in peptide substrates. *Bioorg Med Chem Lett.* 2007; 17:3851–4. [PubMed: 17531489]
177. Iyer SP, Akimoto Y, Hart GW. Identification and cloning of a novel family of coiled-coil domain proteins that interact with *O*-GlcNAc transferase. *J Biol Chem.* 2003; 278:5399–409. [PubMed: 12435728]
178. Cheung WD, Hart GW. AMP-activated protein kinase and p38 MAPK activate *O*-GlcNAcylation of neuronal proteins during glucose deprivation. *J Biol Chem.* 2008; 283:13009–20. [PubMed: 18353774]
179. Coutinho PM, Deleury E, Davies GJ, Henrissat B. An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol.* 2003; 328:307–17. [PubMed: 12691742]
180. Davies GJ. Sweet secrets of synthesis. *Nat Struct Biol.* 2001; 8:98–100. [PubMed: 11175889]
181. Marshall S, Bacote V, Traxinger RR. Discovery of a metabolic pathway mediating glucose-induced desensitization of the glucose transport system. Role of hexosamine biosynthesis in the induction of insulin resistance. *J Biol Chem.* 1991; 266:4706–12. [PubMed: 2002019]

182. Jínek M, Rehwinkel J, Lazarus BD, Izaurralde E, Hanover JA, Conti E. The superhelical TPR-repeat domain of *O*-linked GlcNAc transferase exhibits structural similarities to importin alpha. *Nat Struct Mol Biol.* 2004; 11:1001–7. [PubMed: 15361863]
183. Conti E, Uy M, Leighton L, Blobel G, Kuriyan J. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell.* 1998; 94:193–204. [PubMed: 9695948]
184. Huber AH, Weis WI. The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. *Cell.* 2001; 105:391–402. [PubMed: 11348595]
185. Conti E, Kuriyan J. Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure.* 2000; 8:329–38. [PubMed: 10745017]
186. Martinez-Fleites C, Macauley MS, He Y, Shen DL, Vocadlo DJ, Davies GJ. Structure of an *O*-GlcNAc transferase homolog provides insight into intracellular glycosylation. *Nat Struct Mol Biol.* 2008; 15:764–5. [PubMed: 18536723]
187. Clarke AJ, Hurtado-Guerrero R, Pathak S, et al. Structural insights into mechanism and specificity of *O*-GlcNAc transferase. *EMBO J.* 2008; 27:2780–8. [PubMed: 18818698]
188. Lazarus BD, Roos MD, Hanover JA. Mutational analysis of the catalytic domain of *O*-linked N-acetylglucosaminyl transferase. *J Biol Chem.* 2005; 280:35537–44. [PubMed: 16105839]
189. Wrabl JO, Grishin NV. Homology between *O*-linked GlcNAc transferases and proteins of the glycogen phosphorylase superfamily. *J Mol Biol.* 2001; 314:365–74. [PubMed: 11846551]
190. Lairson LL, Henrissat B, Davies GJ, Withers SG. Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem.* 2008; 77:521–55. [PubMed: 18518825]
191. Konrad RJ, Zhang F, Hale JE, Knierman MD, Becker GW, Kudlow JE. Alloxan is an inhibitor of the enzyme *O*-linked N-acetylglucosamine transferase. *Biochem Biophys Res Commun.* 2002; 293:207–12. [PubMed: 12054585]
192. Gross BJ, Kraybill BC, Walker S. Discovery of *O*-GlcNAc transferase inhibitors. *J Am Chem Soc.* 2005; 127:14588–9. [PubMed: 16231908]
193. Hajduch J, Nam G, Kim EJ, Fröhlich R, Hanover JA, Kirk KL. A convenient synthesis of the C-1-phosphonate analogue of UDPGlcNAc and its evaluation as an inhibitor of *O*-linked GlcNAc transferase (OGT). *Carbohydr Res.* 2008; 343:189–95. [PubMed: 18039537]
194. Kraulis PJ. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J Appl Cryst.* 1991; 24:946–50.

Fig. (1).

Dynamic interplay between the *O*-GlcNAc modification and phosphorylation on serine and threonine residues which occurs over the lifetime of a protein. Phosphorylation is modulated by a multitude of kinases and phosphatases, whereas the addition and removal of the *O*-GlcNAc modification is catalysed by two enzymes, uridine diphospho-*N*-acetylglucosamine:polypeptide β -*N*-acetylglucosaminyltransferase (OGT) and β -*N*-acetylglucosaminidase (OGA), respectively. *O*-GlcNAc, however, is often found at sites that are not phosphorylated and competition in some cases may be indirect by modification of a nearby residue.

Fig. (2).

Substrate-assisted mechanism used by OGA to catalyze hydrolysis with retention of anomeric configuration. During the first step, the acetamido group attacks the anomeric carbon to create an oxazoline intermediate. A water molecule, which is activated by an enzymic general base, attacks the anomeric carbon to hydrolyze the intermediate during the second step. The numbering shown corresponds to the catalytic residues proposed for human OGA (Asp174 is the polarizing residue of the acetamido group and Asp175 is the general acid/base residue).

Fig. (3).

Three-dimensional structure of a bacterial homologue of OGA. **(a)** Cartoon representation of *Bt*GH84 in complex with Thiamet-G (PDB 2VVN [130]). The $(\beta/\alpha)_8$ fold domain possesses the active site residues in a shallow depression. **(b)** Ball-and-stick representation of the active site of *Bt*GH84 in complex with Thiamet-G which shows the active site residues, including the catalytic residues Asp242 (polarizing residue) and Asp243 (acid/base residue). **(c)** Overlap of the active site of *Bt*GH84 in complex with NAG-thiazoline (PDB 2CHN [97]) and GH20 human HexB in complex with NAG-thiazoline (PDB 1NP0 [56]). This shows the basis for selectivity for GH84 enzymes which possess a larger and deeper pocket than GH20 enzymes. **(d)** Surface representation of the active site for enzyme families which use substrate-assisted catalysis demonstrating the variation in active site pocket at the acetamido position: GH18 (PDB 1LLO [77]), GH20 (PDB 1NP0 [56]), GH56 (PDB 1FCV [82]), GH84 (PDB 2CHN [97]) and GH85 (PDB 2W92 [85]). Figs. **(a)**–**(c)** were drawn using MOLSCRIPT [194] and **(d)** using PYMOL (<https://www.pymol.org>).

Fig. (4). Structure of OGA inhibitors discussed in the text: STZ, PUGNAc, NAG-thiazoline, NButGT, Thiamet-G, *gluco*-epimer of nagstatin, PUGNAc-imidazole hybrid and GlcNAcstatin C.

Fig. (5).

Three-dimensional structure of a bacterial homologue of OGT. **(a)** Cartoon representation of *XcGT41* in complex with a C1 phosphonate analogue of UDP-GlcNAc (PDB 2JLB [187]). The active site lies at the interface of the two (α/β)-fold domains which make up the GT-B topology. **(b)** Ball-and-stick representation of the active site of *XcGT41* in complex with a C1 phosphonate analogue of UDP-GlcNAc, and where significant differences arise *XcGT41* in complex with UDP (PDB 2VSN [186]). His218 has been proposed to be the catalytic base by Martinez-Fleites *et al.* [186], but lies in a different position in the two complexes. **(c)** Cartoon and surface representation of the dimeric TPR domains from OGT (PDB 1W3B [182]); the TPR domains are a homodimer of right-handed superhelices. **(d)** Cartoon and surface representation of importin- α in complex with a nuclear localization sequence (PDB 1EE4 [185]), which has been suggested to have a similar mode of binding to the TPR domains. The asparagine residues on the inner surface of the superhelix have been shown to be important for recognition of their target sequences. Figs. **(a)** and **(b)** were drawn using MOLSCRIPT [194] and **(c)** and **(d)** using PYMOL (<https://www.pymol.org>).

Fig. (6).

Structure of OGT inhibitors discussed in the text: UDP-GlcNAc, alloxan, three compounds identified from high throughput screening, and a C1 phosphonate analogue of UDP-GlcNAc.