DOCTOR OF PHILOSOPHY

Molecular characterisation of O-GlcNAc transferase

Zheng, Xiaowei

Award date:
2013

Awarding institution:
University of Dundee

DOCTOR OF PHILOSOPHY

Molecular characterisation of O-GlcNAc Transferase

Xiaowei Zheng

2013

University of Dundee

Molecular Characterisation of O-GlcNAc Transferase

By

Xiaowei Zheng

A Thesis Submitted for the Degree of

Doctor of Philosophy

University of Dundee

February 2013

Acknowledgements

I would like to thank my PhD supervisor Prof. Daan van Aalten for his mentoring and supervision for the past 4 years, who provided me the opportunity to undertake my post-graduate study in his laboratory. I especially enjoyed the time when several of us were locked away on a small island of Arran for paper writing or an after dinner discussion at his house. Those times were proven to be most productive and very much has been taken home.

Would also like to thank two ladies, Dr. Shalini Pathak and Dr. Marianne Schimpl, who have been sharing the same office for the past three years. Although topics on food of differently kinds were the most popular ones, there were always invaluable discussions on the project. Special thank should be given to Dr. Marianne Schimpl for the help with the mechanist study of hOGT, which makes the story complete. I would also like to thank her for the help through the last few months during the writing.

During my PhD, it has been much collaboration involving different projects. I would like to express my gratitude to Prof. Arno Müller for his support and suggestions for the fly works. I would also like to thank Dr. Daniel Mariyappa for his input on the *Drosophila* project.

I also thank all the people who supported me during my PhD. Dr. Andrew Ferenbach, the 'cloning master', has been a great help with the constructs, where all the proteins used in the project started. Dr. Osama Albrarbarawi and Dr. David Campbell helped a lot for the Mass Spec studies. DSTT and the media kitchen have made the life easier by providing the reagent. I also thank all the DVA lab members, the current and the past.

No matter how sunny Dundee could be, there were always some cloudy days; just like me during my study, there were always 'up and down'. Very often these personal feelings have gone to the persons very close to you. In my case, there were two women, my mum and my wife. I am really glad that they were always there when you needed them and it will never be the same without them by my side.

Declaration

I declare that the following thesis is based on the results of investigations conducted by myself, and that this thesis is of my own composition. Work other than my own is clearly indicated in the text by reference to the relevant researchers or to their publications. This dissertation has not in whole, or in part, been previously submitted for a higher degree.

Xiaowei Zheng

We certify that Xiaowei Zheng has spent the equivalent of at least nine terms in research work at the College of Life Sciences, University of Dundee, and that he has fulfilled the conditions of the Ordinance General No 14 of the University of Dundee and is qualified to submit the accompanying thesis in application for the degree of Doctor of Philosophy.

Prof. Daan M.F. van Aalten

List of Publications

The work described in this thesis has been published in the following articles:

X. Zheng*, D. Mariyappa*, S. Marianne Schimpl, A. Ferenbach, A. Mueller, and D.M.F. van Aalten, 'O-GlcNAc transferase catalytic activity is required for larval development in *Drosophila*' *In preparation.* (*Authors contribute equally)

S. Pathak, O. Albarbarawi, M. Schimpl, V. S. Borodkin, D. E. Blair, **X. Zheng**, A. T. Ferenbach, K. Sanders, V. Gapsys, Bert de Groot, P. Pedrioli and D. M. F. van Aalten 'A degenerate peptide sequon and regional disorder define human O-GlcNAc transferase specificity'. *In preparation*

M. Schimpl*, **X. Zheng***, V.S. Borodkin*, D.E. Blair, A.T. Ferenbach, A.W. Schuettelkopf, I. Navratilova, T. Aristotelous, O. Albarbarawi, D.A. Robinson, M.A. MacNaughtan and D.M.F. van Aalten, "O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis", *Nature Chem.Biol*. (2012), 8, 969-974. (*Authors contribute equally)

L.M. Gay, **X. Zheng** and D.M.F. van Aalten, "O-GlcNAc transfer: size matters", *Nature Chem.Biol*. (2011), 7, 134-135.

H.C. Dorfmueller, V.S. Borodkin, M. Schimpl, **X. Zheng**, R. Kime, K. D. Read and D.M.F. van Aalten, "Cell-penetrant, nanomolar O-GlcNAcase inhibitors selective against lysosomal hexosaminidases", *Chem.Biol.* (2010), 17, 1250-1255.

Summary

Post-translational modification of a large number of intracellular proteins with a single O-linked GlcNAc sugar is involved in various cellular processes. The dynamic nature of this O-GlcNAc modification has demonstrated its relevance in cell signalling and transcription regulations. At the centre of this complex regulatory network are two enzymes with antagonistic action: O-GlcNAc transferase (OGT) that catalyses the transfer of GlcNAc onto Ser/Thr side chains of proteins and O-GlcNAcase that removes the O-GlcNAc.

The work in this thesis was focused on the O-GlcNAc transferase, using a combination of structural biology and cell biology approaches. OGT exhibits two defined domains, the N-terminal domain consisting of tetratricopeptide repeats (TPRs) and the C-terminal glycosyltransferase domain. The first chapter focuses on the functions of the TPR domain of OGT in substrate recognition. It also provides examples of how the TPR domain may affect the substrates through binding to TPRs. During the study using systematic TPR truncations, multiple expression constructs were generated and used for the production of stable and pure hOGT proteins, which led to a detailed study of the catalytic mechanism of hOGT with the aid of structural biology, as described in Chapter II. A novel glycosyltransferase catalytic mechanism is suggested for hOGT. Structural information gained from this thesis and other previous reports was employed to design further experiments using *Drosophila* model, aimed at investigating the involvement of OGT and O-GlcNAcylation during development, as described in Chapter III.

Abbreviations

Å	**Angstrom**
Amp	**Ampicillin**
BSA	**Bovine serum albumin**
°C	**Degree Celsius**
C-terminal	**Carboxy-terminal**
CCP4	**Collaborative computational project number 4 in protein crystallography**
C elegans	***Caenorhabditis elegans***
CAZy	**Carbohydrate Active enZymes**
CID	**Collision-induced dissociation**
CK2α	**Casein kinase 2 alpha**
cpm	**Counts per minute**
***Cp*OGA**	***Clostridium perfringens* O-GlcNAcase**
CREB	**cAMP response element-binding protein**
Da	**Dalton**
ddH2O	**Doubly distilled water, milli-Q water**
dH2O	**Distilled water**
Dm	***Drosophila melanogaster***
DSTT	**Division of Signal Transduction Therapy**
DTT	**Dithiothreitol**
ECL	**Enhanced chemiluminescence**
E. coli	***Escherichia coli***
EDTA	**Ethylenediamine tetraacetic acid**
EGTA	**Ethyleneglycol bis-(2-aminoethylether)-*N,N,N',N'* tetraacetic acid**
FBS	**Fetal bovine serum**
FOXO1	**Forkhead in rhabdomyosarcoma**

g	**Gram (or gravity)**
GalNAc	***N*-acetylgalactosamine**
GlcNAc	***N*-acetylglucosamine**
GlcNAz	***N*-azidoacetylglucosamine**
GSH	**Glutathione**
GSK3β	**Glycogen synthase kinase-3 beta**
GST	**Glutathione-*S*-transferase**
h	**Hour**
HEPES	***N*-(2-Hydroxyethyl)piperazine-*N'*-(2-ethanesulfonic acid)**
HEK293	**Human Embryonic Kidney 293 (cells)**
HRP	**Horseradish peroxidase**
IC50	**Half maximal inhibitory concentration**
IPTG	**Isopropyl β-D-1-thiogalactopyranoside**
K_i	**Inhibition constant**
K_m	**Michaelis constant**
LB	**Lysogeny broth**
LC-MS	**Liquid chromatography–mass spectrometry**
LDS	**Lithium dodecyl sulphate**
OGA	**O-GlcNAcase**
OGT	**O-GlcNAc transferase**
SD	**Standard deviation**
SPR	**Surface plasmon resonance**
PDB	**Protein Data Base**
PEG	**Polyethylene glycol**
PMSF	**Phenylmethanesulfonylfluoride**
p*K*a	**Acid dissociation constant**
PREs	**Polycomb Response Elements**
PVDF	**Polyvinylidene fluoride**
RMSD	**Root mean square deviation**
SH2, 3, 4	**Src homologous 2, 3 or 4**
SPR	**Surface Plasmon Resonance**
SDS-PAGE	**Sodium dodecyl sulfate polyacrylamide gel electrophoresis**

sxc	***Super Sex Combs***
TAB1	**Mitogen-activated protein kinase kinase kinase 7-interacting protein 1**
TBS	**Tris buffered saline**
TPR	**Tetratricopeptide repeat**
UDP-GlcNAc	**Uridine diphopho-N-acetylglucosamine**
UAS	**Upstream Activation Sequence**
Xc	***Xanthomonas campestris***
Yes1	**Tyrosine kinase Yes**

Amino Acid Code

Amino acid	Three letter code	One letter symbol
Alanine	Ala	A
Argineine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histadine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methione	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Theonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Glycosylation

Oligo- and polysaccharides, or glycans, are the most abundant biomolecules on earth. In the linear or branched arrangement consisting of either mono- or heteropolymers of monosaccharide, glycans provide large diversity in live systems. Glycans can be secreted or attached to proteins or lipids, forming glycoconjugates. In human, more than 50% of proteins are glycoproteins through co- or post-translational modification (Apweiler et al, 1999). The attachment of glycans on a glycoproteins fall mainly into two classes depending on the glycol-peptide linkage: attachment of sugar(s) to the nitrogen atom of asparagine (N-glycosylation), and to an oxygen atom (O-glycosylation) predominantly to serine or threonines, although atypically on tyrosine, hydroxylproline and hydroxylysine have also been observed (Spiro, 2002).

1.2. Historical overview of protein O-GlcNAcylation

The post-translational modification of intracellular proteins with the monosaccharide β-D-N-acetyl-glucosamine (GlcNAc) on serine or threonine was discovered in the 1980s, while studying the composition of the terminal N-acetylglucosamine residues on the surface glycans of live

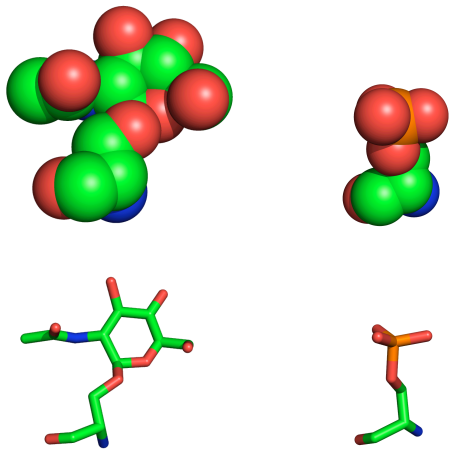


Figure 1.1: Chemical structure of O-GlcNAc modified serine with both phosphorylation and O-GlcNAcylation. Both molecules are shown in balls and sticks representation.

lymphocytes (Torres & Hart, 1984) (Fig. 1.1). The authors probed for terminal GlcNAc moieties with radiolabeled UDP-GalNAc and bovine milk Galactosyltransferase (GalT1). It was found that O-GlcNAcylation is particularly enriched in the nuclei, especially on the nuclear envelope, and the soluble fraction in the rat liver tissue (Holt & Hart, 1986). Its localisation and non-extending monosaccharide modification therefore differ from the classic cell surface glycan. Early studies have shown that the turnover of O-GlcNAcylation is faster than the protein that carries it (Chou et al, 1992; Roquemore et al, 1992); this dynamic feature of O-GlcNAc modification suggested its function in protein regulation. Over the years, with technical advances in O-GlcNAc detection, more than a thousand proteins have been found to be O-GlcNAc modified; and despite early speculation of targeting the carriers to the nuclei (Schindler et al, 1987), which are now believed to be involved in various cellular functions, as discussed later (Fig. 1.2).

1.2.1. Phylogeny of O-GlcNAc Modification

Post-translational modification on intracellular proteins by O-GlcNAc is highly conserved across species from bacteria to human. In the bacterium *Listeria monocytogenes*, flagellar protein FlaA was found to be modified with O-GlcNAc (Schirm et al, 2004; Shen et al, 2006). Although there is no clear evidence for the function of this modification, other O-linked glycosylation on flagellins have been shown to affect the assembly and motility of the flagella (Schirm et al, 2003; Thibault et al, 2001). In eukaryotes, O-GlcNAcylation has been found in organisms from the most ancient and primitive protist *Giardia* (Banerjee et al,

2009), to plants and animals from the most primitive multicellular placozoon *Trichoplax adhaerens* (unpublished) to higher animals like *C. elegans*, *Drosophila*, *Xenopus*, zebrafish, mouse and in human (Olszewski et al, 2010). O-GlcNAcylation has not been reported in yeast to date.

1.2.2. O-GlcNAc is Dynamic and Inducible

Two enzymes carry out O-GlcNAc modification, O-GlcNAc transferase (OGT) adding GlcNAc moiety and O-GlcNAcase (OGA) removing it. The high-energy donor substrate of OGT is UDP-GlcNAc, an end product of the hexosamine biosynthetic pathway (Fig. 1.3). 2-5% of the total glucose consumption is converted into UDP-GlcNAc (Marshall et al, 1991b; Marshall et al, 2004). Many components in the hexosamine biosynthetic pathway, such as acetyl-CoA,

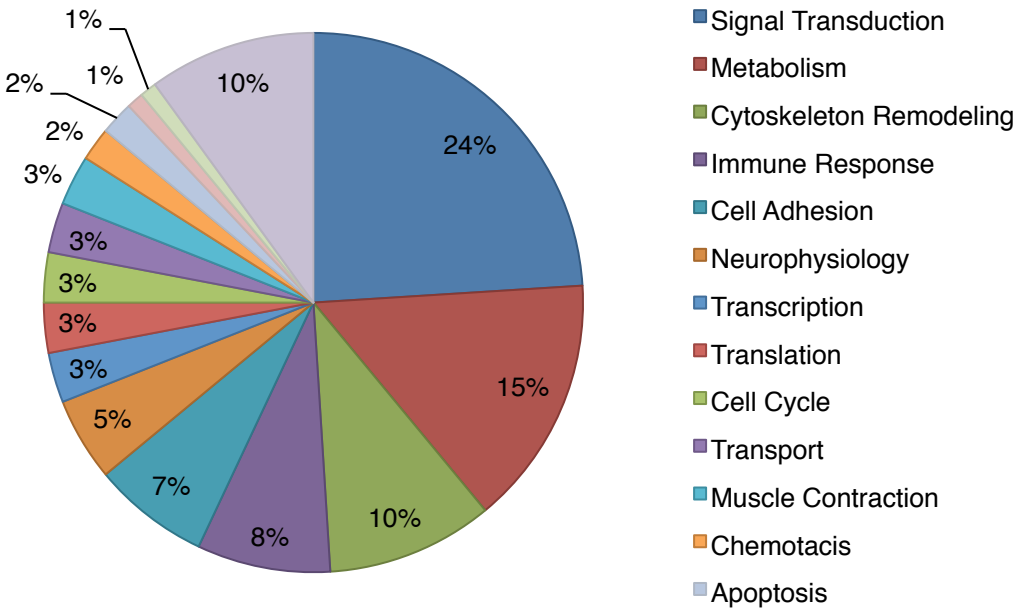


Figure 1.2: A pie chart representation of the intracellular proteins that have been found to be O-GlcNAc modified and their involvement in cellular processes.
Percentage was calculated from the total number of the proteins carrying O-GlcNAcylation. Fig is adopted from Ruan *et al*., 2012, *Cell Metabolism*.

glutamine and UTP, are also involved in other metabolic pathways and play key roles. The synthesis of UDP-GlcNAc is therefore affected by the cross talk between different biosynthetic pathways and is subject to tight regulation. OGT activity is dependent on and regulated by the availability of the donor substrate UDP-GlcNAc. Fluctuation of the intracellular UDP-GlcNAc pool alters the global O-GlcNAc level through OGT. The cellular O-GlcNAc level can also be acutely elevated by the addition of glucosamine, bypassing the rate limiting enzyme in the hexosamine biosynthetic pathway, glutamine:fructose-6-phosphate aminotransferase (Marshall et al, 1991a) (Fig. I.3).

It also appears that cellular O-GlcNAc levels are responsive to insulin, growth hormones and cellular stresses, including UV exposure, ethanol, heavy metals, heat shock, osmotic and oxidative stresses (Zachara et al, 2004). Pathologically, ischemia (Champattanachai et al, 2007; Zou et al, 2007), hypoxia (Ngoh et al., 2009) and arterial injury (Xing et al, 2008) also induces O-GlcNAc flux.

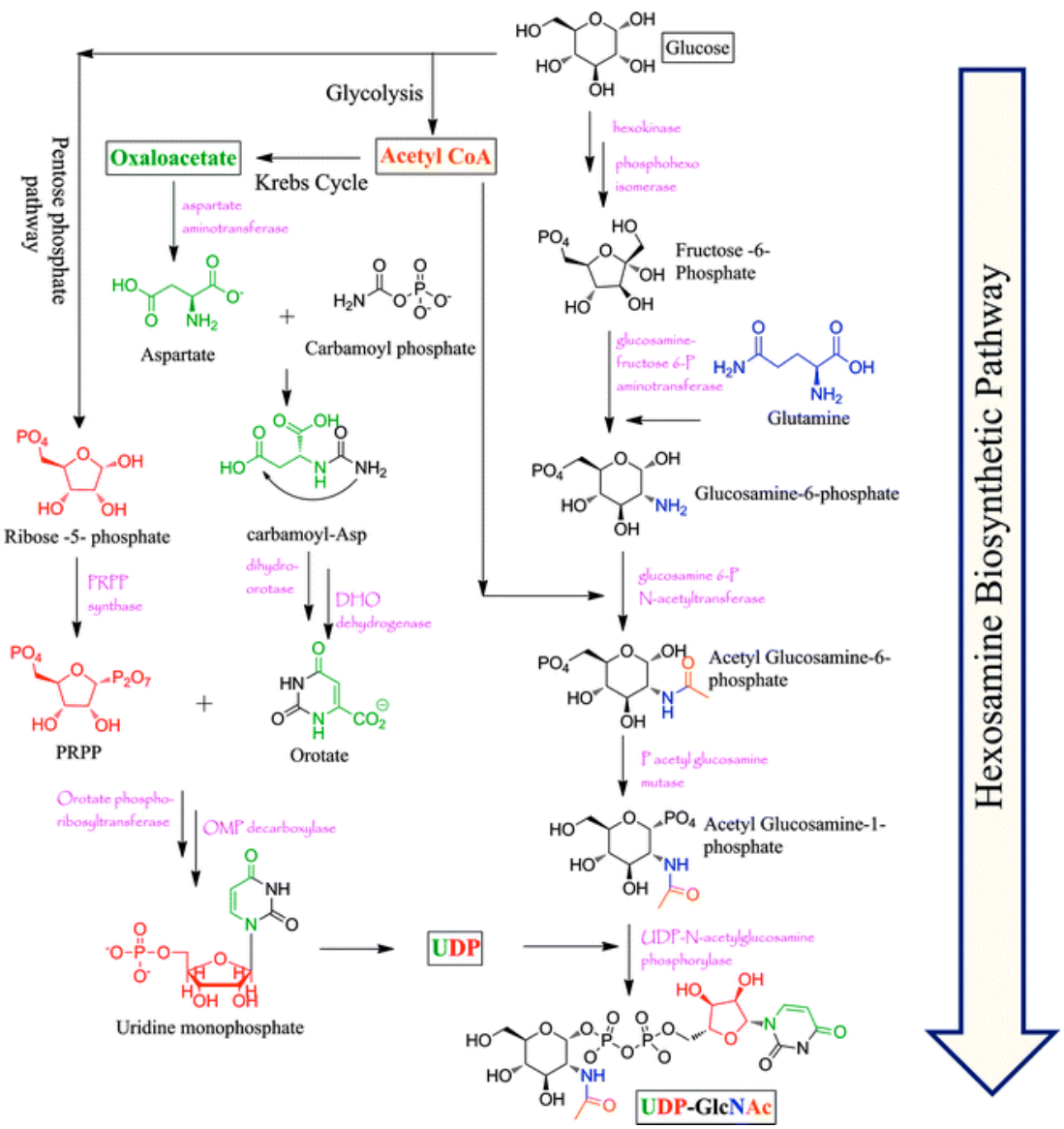


Figure 1.3: Schematic overview of hexosamine biosynthetic pathway (HBP). Glucose is used in the HBP to produce UDP-GlcNAc as the final product. Glucose, as the upstream source of the hexosamine biosynthetic pathway, not only provides the scaffold but also contributes to the synthesis of many components for UDP-GlcNAc synthesis. The origin of each component is colour coded. Arrows indicate the direction of the pathway and enzymatic processes. Key enzymes are labelled next to the reaction (magenta). The figure is adapted from Banerjee *et al.*, 2013, *Chem. Soc. Rev.*

1.3. O-GlcNAc Cycling Enzymes

In cells, O-GlcNAcylation and its turnover is orchestrated by the antagonising actions of two enzymes, O-GlcNAc transferase (OGT) and O-GlcNAc hydrolase (OGA).

1.3.1. O-GlcNAc Transferase (OGT)

O-GlcNAc transferase, the enzyme responsible for the post-translational modification with O-GlcNAc, was affinity-purified using a synthetic substrate peptide from rabbit raticulocytes 8 years after O-GlcNAcylation was discovered (Haltiwanger et al, 1990). OGT was later purified from rat liver using an UDP-hexanolamine-sepharose column (Haltiwanger et al, 1992). A 340 kDa protein complex was identified, which consisted of two 110 kDa and one 78 kDa subunits. These are two of three isoforms known today as nucleocytoplasmic OGT (ncOGT) and short OGT (sOGT). The authors observed at the time that the mitochondrial fraction contained 15% of the OGT activity (Hart et al, 2011; Love

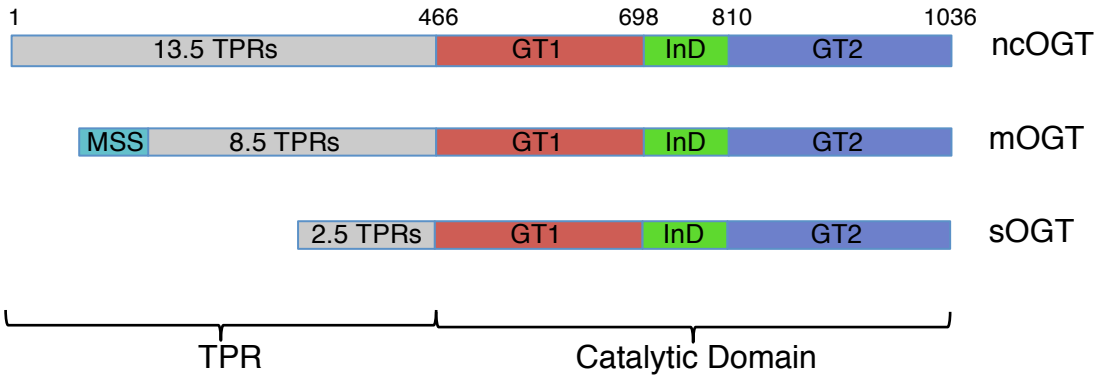


Figure 1.4: Schematic representation of the domain structure of human O-GlcNAc transferase. Three isoforms OGT have been identified: nucleoplasmic OGT (ncOGT), mitochondrial OGT (mOGT) and small OGT (sOGT). All three OGT isoforms contain different length of tetratricopeptide repeats (TPRs) at the N-terminus (gray), except mOGT that starts with a mitochondrial specific sequence (MSS, slate). The catalytic domains of the three OGT isoforms are identical, with a glycosyltransferase domain of two lobe (GT1, red and GT2, blue) and a intervening domain (InD, green).

et al, 2003). Subsequently, an OGT isoform specific to mitochondria was reported and denoted as the mitochondrial specific OGT (mOGT) (Love et al, 2003). The identification of the *ogt* gene demonstrated that all three isoforms are encoded from a single gene as the products of splicing (Shafi et al, 2000). They share an identical C-terminal glycosyltransferase domain and differ in the number of repeats in the Tetratricopeptide repeat (TPR) domain at the N-terminus (Fig. 1.4). An additional mitochondrial targeting sequence in mOGT directs the protein to the mitochondria (Lubas & Hanover, 2000). *ogt* genes from many organisms have been cloned, including *Arabidopsis thaliana (Hartweck et al, 2002)*, *C. elegans* (Hanover et al, 2005), *Drosophila melanogaster* (Gambetta et al, 2009; Sinclair et al, 2009), *Xenopus laevis* (Kenwrick et al, 2004), zebrafish (Webster et al, 2009), mouse (Shafi et al, 2000), and human (Lubas & Hanover, 2000) . Sequence conservation is high throughout the catalytic domain as well as the N-terminal TPR repeats of metazoan OGTs, with the intervening domain being the least well conserved region of the protein (Fig 1.5). This intervening domain is absent in OGTs from bacteria such as *Xanthomonas campestris* and *Listeria monocytogenes*, the domain arrangement is, however, conserved. Uniquely, the catalytic domain of this *Listeria monocytogenes* OGT belongs to a different family of glycosyltransferases (GT2). (See later in *General Classification of Glycosyltransferases*)

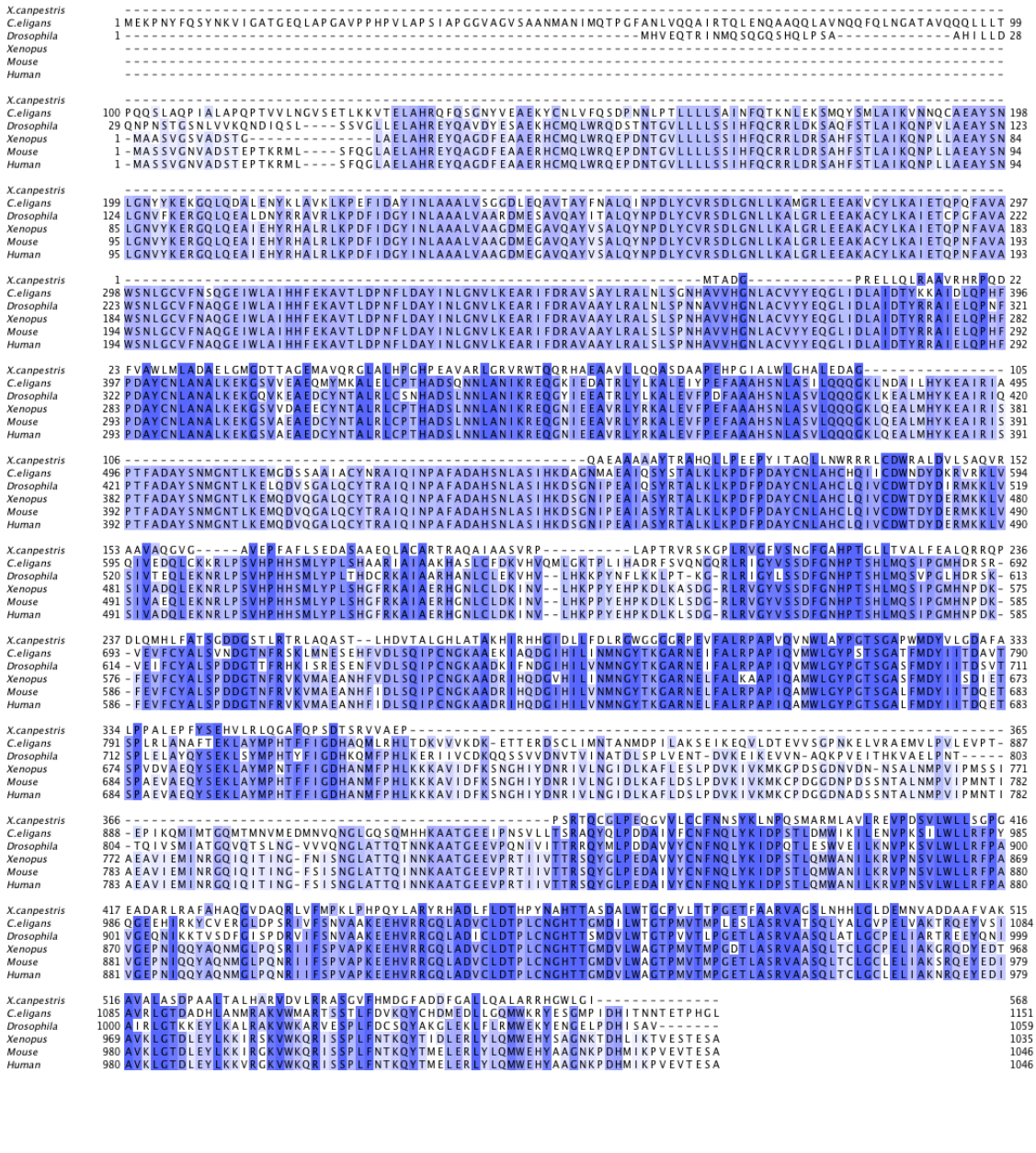


Figure 1.5: Sequence conservation among metazoan O-GlcNAc transferases: Multiple sequence alignment of sequences of O-GlcNAc transferases from *Xanthomonas campestris* (UniProt: Q8PC69), *Caenorhabditis elegans* (O18158), *Drosophila melanogaster* (Q7KJA9), *Xenopus laevis* (Q641I4), *Mus musculus* (Q8CGY8) and *Homo sapiens* (O15294). Blue colour shades indicate the conservation across species.

1.3.1.1. O-GlcNAc Transferase is Involved in Many Cellular Processes

Metazoan Embryogenesis

Nearly all animals analysed appear to have a single *ogt* gene, except for zebrafish, which has two (Sohn & Do, 2005). Studies have shown that *ogt* is essential for the development in several animal models. In mouse, it has shown that the intact OGT gene is essential for the viability of ES cell and no mice with OGT deletion allele was produced with homologous recombination (Shafi et al, 2000); nevertheless, the authors were able to generate mice with allele containing functional *lox-P* flunking *Ogt* gene, which would allow conditional knock out with Cre-*lox* system. By applying this conditional mutagenesis strategy, O'Donnel *et al*. was able to introduce OGT deletions in specific tissues with tissue specific promoter driving the expression of Cre recombinase, which mediates recombination event at the *lox-P* site and excises the target gene, in this case, *Ogt* gene (O'Donnell et al, 2004). T-cell specific knockout of OGT has been found to induce T-cell apoptosis by up to 75% within lymph nodes and spleen prior to emigration into peripheral. In neuronal specific ogt knockout mice, only half carrying deletion allele survived postnatal, but they fail to nurse and die within 10 days. It is also found increased Tau protein levels in the brain and spinal cord; Tau is also found hyperphosphorylated, which is often found associated with early death and neuronal apoptosis, and neurofibrillary tangle formation by hyperphosphorylated Tau is pathological hallmark of neurodegenerative disease Tauopathies (Brich et al, 2003; Kobayashi et al, 2003). Fibroblast *ogt* knockout using retrovirus appeared senesce and die by day 12; biochemical studies has shown that the cells were defective in serum

responses and decrease protein stability on transcription factors Sp1, c-Jun, c-Fos and c-Myc in the absence of OGT and O-GlcNAc. In agreement, protein levels of transcription factor Tet1 is also reduced with depletion of OGT or replacing with an inactive mutant of OGT (Shi et al, 2013)., suggesting O-GlcNAcylation might be important in transcription as discussed later in section *1.7.1 Roles in transcription*. Interestingly, many of the phonotypes were observed in cells undergoes cell division or during development, suggesting OGT and O-GlcNAcylation primarily function in cell division as discussed in section *1.7.2 Rose in cell cycle and cell division*; aberrant levels of OGT or subsequence O-GlcNAc levels would disrupt cell cycle program and lead to apoptosis. Recently, a report showed tissue specific OGT ablation in adult mouse heart did not lead to heart dysfunction (Watson et al, 2010), further support this hypothesis.

The functions of OGT in neuronal development were also observed in zebrafish. In zebrafish, two *ogt* genes encode 6 variants. Var1, 2 express in the very early stages while var 3, 4 appear later, after the dome stage. Var5, 6 express throughout the development (Sohn & Do, 2005). OGT is expressed dominantly in the brain at 24 hours post-fertilization (hpf) and becomes more restricted at 36 hpf and 72 hpf (Webster et al, 2009). Depletion of OGT levels by injecting morpholinos against each of the transcripts in the 1-4 cell stage led to increased cell death and developmental defects, producing embryos with shortened body axis, smaller brain and abnormalities in or absence of eyes. Necrotic tissue in the brain was also observed in these embryos. OGT depletions also caused modest change in global O-GlcNAc levels (Webster et al, 2009).

The outcome is not surprising given that O-GlcNAc modification has been found to be most abundantly in the brain (Dias & Hart, 2007). Recently, a study monitoring OGT expression and global O-GlcNAc levels in the rat brain over a two-year period has been reported (Liu et al, 2012). High levels of O-GlcNAcylation was observed in the early development, followed by a marked decline in O-GlcNAc levels. The total OGT protein levels remain stable with alteration of the isoforms. OGT protein expression in early development has been studied before that. In *Xenopus tropicalis*, *ogt* RNA is found at low level throughout neurula embryos with high level at the anterior neural tube (Gawantka et al, 1998; Kenwrick et al, 2004). The *ogt* gene is later primarily expressed in the head at the tadpole stage. *ogt* depletion by morpholino injection produced embryos with gastrulation defects presenting as bifurcated axis (spina bifida), while the rest were truncated at the posterior end with little or no tail buds. These dramatic morphological defects observed in mouse, zebrafish and frog, produced by OGT depletion either with gene removal or morpholino depletion, suggest that the development, especially the neuronal development, in vertebrate, is sensitive to the change in OGT protein level, and possibly to fluctuations in O-GlcNAc levels.

ogt in *Drosophila* is also known as *super sex combs* (*sxc*), which was first described in 1984 as a Polycomb group (PcG) gene that plays a repressive function on homeotic genes during the development. This suggests a function of OGT in transcription. Mutations of *sxc* led to homeotic transformation (Ingham, 1984). Only recently, the product encoded by *sxc* was proven to be OGT and

mutations of *sxc* leading to OGT ablation also resulted in early death of the flies in pharate adults (Gambetta et al, 2009; Sinclair et al, 2009). The importance of *sxc* in the early embryogenesis is not clearly demonstrated in flies as the maternal transcripts and gene products in the oocyte can compensate for the lack of sxc on repression. Transgenic flies overexpressing sxc and hOGT were able to rescue this lethality (Sinclair et al, 2009), making *Drosophila* an excellent model for studying the functions of OGT and O-GlcNAcylation.

Energy Sensing and Glucose metabolism

Drosophila was also employed as a model system in the study of OGT and its involvement in energy metabolism. Genetically introducing raise in OGT protein levels in flies produces an increase in the body size and weight in larvae and adult (Sekine et al, 2010), which was also observed in the flies with reduced OGA protein levels, suggesting this effect may be caused by the elevation of global O-GlcNAc levels. Conversely, reduction in OGT protein produces smaller larvae and adults. *Drosophila* insulin-like peptide (dilp) expression was also reduced in this OGT KD model, indicating the involvement of OGT and O-GlcNAcylation in glucose metabolism and insulin signalling.

C. elegans with *ogt* null allele is viable and fertile, and phenotypically normal upon simple inspection (Hanover et al, 2005). The genomic rearrangement introduced an in-frame stop codon, producing a possible fragment of OGT with 465 residues that retains the entire TPR repeat but without the catalytic domain. This *ogt-1* mutant strain exhibited an abnormal energy homeostasis with

elevation in trehalose levels and glycogen stores, and a decrease in triglyceride levels; the OGT knockout also suppresses dauer larvae formation induced by a temperature-sensitive allele of the insulin-like receptor gene *daf-2*, also suggesting OGT has a function in energy sensing and the regulation of insulin signalling.

The involvement of OGT in energy homeostasis is also observed in the plant *Arabidopsis thaliana*. There are two ogt genes, *spindly* (*spy*) and *secret agent* (*sec*) (Hartweck et al, 2002; Jacobsen & Olszewski, 1993; Silverstone et al, 2007). Both are involved in plant growth hormone gibberellin signalling. OGT was shown to be essential: Single gene knock-outs produced mild effects, however, removal of both genes led to severe effects on germination, and embryos died before development (Hartweck et al, 2002).

1.3.2. O-GlcNAcase (OGA)

O-GlcNAcase was first identified as a cytosolic and neutral β-N-acetylglucosaminidase and was called hexosaminidase C to distinguish from the lysosomal acidic hexosaminidase A and B (Braidman et al, 1974; Overdijk et al, 1981; Swallow et al, 1976). OGA was first purified from rat spleen cytosol and later rat brain (Dong & Hart, 1994; Gao et al, 2001). Human OGA was cloned in 2001 and was found to be identical to a previously identified gene, meningioma-express antigen 5 (*MGEA5*), which was originally thought to be a hyaluronidase, since expression of these matrix-degrading enzymes is often associated with meningioma (Heckel et al, 1998). OGA is also highly conserved in eukaryotes,

and allows studying many model organisms, such as *C. elegans,* zebrafish, *Drosophila* and plant *Arabidopsis* (Olszewski et al, 2010). Northern blot has shown OGA transcript is expressed in all human tissue examined, with high levels in the brain, placenta and pancreas (Gao et al, 2001).

There are two isoforms of OGA in eukaryotes (Fig. 1.6), with the shorter splice variant predominantly found in the nucleus of glioblastoma cells. The long splice variant contains an N-terminal hydrolase domain and a C-terminal histone acetyl transferase-like domain connected by a domain predicted to adopt mainly α-helix and a less conserved disordered region. This connecting region serves many biological functions; it is the site of caspase cleavage during apoptosis, phosphorylation (Wells et al, 2002; Whisenhunt et al, 2006). Most interestingly, the interaction of OGA and OGT occurs in the connecting region of OGA and allows the formation of a complex, or GlcNAczyme, that possesses antagonising action. The complex has been reported in repressor complex with Sin3A and at the midbody with phosphatase 1 and mitotic kinase Aurora B (Slawson et al, 2008; Whisenhunt et al, 2006). The function of this OGT-OGA complex and how it is coordinated however is not clear understood. The HAT-like domain in OGA was shown to possess acetyltransferase activity on Lys8 of histone H4 and Lys14 of histone H3 (Toleman et al, 2004), but to date this finding has never been reproduced. Nevertheless, in addition to the catalytic domain, the HAT-like domain has been shown in yeast-two-hybrid experiments to contribute to protein-protein interaction as observed in OGT (Butkinaree et al, 2008). The catalytic domain belongs to the GH84 glycoside hydrolase family and adopts the classic

TIM barrel fold demonstrated by the crystal structures of bacterial OGA homologues (Dennis et al, 2006; Rao et al, 2006). The structural information from these bacterial OGAs has provided invaluable insights and contributed to the development of many pharmacological inhibitors, that have been used to study the function of O-GlcNAcylation *in vivo* (Yuzwa et al, 2012). To date, the structure of hOGA is yet to be solved and the biological functions of O-GlcNAc modifications remain to be elucidated.

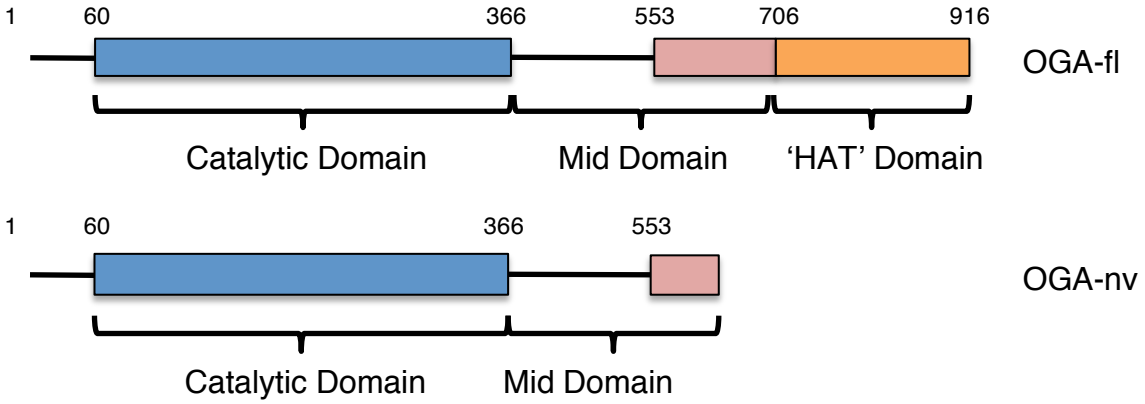


Figure 1.6: Schematic representation of the domain structure of human O-GlcNAcase. Human O-GlcANcase consists of three distinctive domains: N-terminal glycosyl hydrolase domain (blue), C-terminal HAT-like domain (orange) and a middle domain, half of which has no predicted secondary structure. The short nuclear variant OGA (OGA-nv) lacks the HAT-like domain from the full length OGA (OGA-fl).

1.4. Effect of O-GlcNAcylation on Proteins

1.4.1. O-GlcNAc modification induces conformational change

The exact function of O-GlcNAc modification to the secondary structure of the target protein is not clearly understood. The modification leads to the addition of a large (comparing to O-phosphorylation) but uncharged moiety onto the substrate proteins (Fig. 1.1), which may presumably introduce local conformational changes, as occasionally observed for phosphorylation. Studies using biophysical methods, such as nuclear magnetic resonance and circular dichroism, have shown O-GlcNAc modification promotes turn-like structure formation on the C-terminal domain of PolII and the N-terminus of estrogen receptor β (Chen et al, 2006; Simanek et al, 1998). To date, no crystal structure of a protein carrying O-GlcNAc modification has been reported, partially because of the flexible nature of the modification site. The existing crystal structure of p53 (DNA binding domain) may provide valuable information on the effect of O-GlcNAc modification on protein architectures (Cho et al, 1994). Ser149 and Thr155 are found to be O-GlcNAcylated and phosphorylated, respectively (Yang et al, 2006). In crystal structures of the unmodified protein, both residues are located within a short ordered loop with the amino acid side chains buried. In order to be modified, significant rearrangement would be required, which may affect the nature of the protein. Another example is O-GlcNAc modification on histone 2B, where Ser112 is found an O-GlcNAc modification site (Fujiki et al, 2011). Crystal structure of histone 2B containing Ser122 shows that this region, in the absence of any post-translational modification, adopts α-helical conformation (Luger et al, 1997). More interestingly, lys120, a previously reported

mono ubiquitination site, is also packed in the same helix (Fujiki et al, 2011). O-GlcNAclyation on Ser112 induces conformational change and promotes K120 monoubiquitination (Fujiki et al, 2011).

1.4.2. Cross-talk between O-GlcNAcylation and phosphorylation

Modulation of O-GlcNAc level in cells by changing OGT and OGA levels or pharmacological methods result in alteration of the cellular phosphoproteome, and *vice versa* (Slawson & Shaughnessy, 2005; Wang et al, 2008; Wang et al, 2007), demonstrating the interplay between protein O-GlcNAcylation and O-phosphorylation in cells. Since O-GlcNAcylation occurs on serine and threonine residues, competitive site occupancy is possible, and intracellular proteins may be reciprocally modified with protein phosphorylation at the same or adjacent

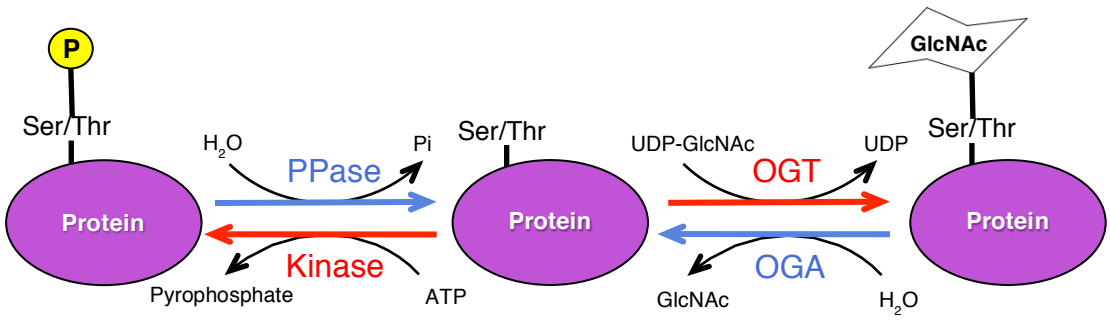


Figure 1.7: Schematic representation of the interplay between O-GlcNAcylation and O-phosphorylation.
Protein substrates are modified at serine or threonine residues. GlcNAcylation and phosphorylation mutually exclude each other at the same modification site. Enzymes performing the transfer reactions, OGT in O-GlcNAcylation and kinase in phosphorylation, are coloured in red; the enzymes performing hydrolysis, OGA and phosphatase (PPase), are in blue.

sites. Almost all the O-GlcNAc modified proteins identified so far are also phosphorylated, with some proteins showing a reciprocal relationship (Fig. 1.7) (Hart et al, 2011).

Many reports have also demonstrated the involvement of O-GlcNAc modification in the regulation of protein stability, enzyme activity, cellular localisation and protein-protein interaction (Hart et al, 2007), however, in most cases the molecular mechanisms of these functions remain unclear. As yet, no O-GlcNAc binding protein or domain has been discovered.

1.5. Roles of O-GlcNAc in Cellular Processes

It is generally considered that O-GlcNAc has a close relationship with energy metabolism as mentioned in sections *1.4 O-GlcNAc is Dynamic and Inducible* and *1.5.2.2. Energy Sensing and Glucose metabolism* through UDP-GlcNAc in HBP. With the interlink with protein phosphorylation, O-GlcNAcylation provides a 'molecular switch' that regulates many cellular processes.

1.5.1. Roles in transcription

Chromatin is enriched with O-GlcNAc modifications as shown by immunostaining of polytene chromosome in *Drosophila*, and later by chromatin immunoprecipitation; almost all of which are located at the transcription active regions (Gambetta et al, 2009; Kelly & Hart, 1989; Sinclair et al, 2009). Early Mass-spec studies have identified large group of proteins involving in protein transcription carry O-GlcNAc modification. One of the key component, the

catalytic subunit of the RNA polymerase II is found heavily O-GlcNAcylated at the C terminal tail, which contains a tandem of C-terminal repeat domains (CTD) with imperfect consensus sequence YSPTSPS (Dahmus, 1996). CTD is originally found to adopt either unphosphorylated or phosphorylated, and it has been understood that it serves to recruit other proteins during elongation (Phatnani & Greenleaf, 2006). Recently, Ranuncolo et al has suggested that O-GlcNAcylated RNA PolII at CTD represents the third status in addition to phosphorylated and unphosphorylated; the O-GlcNAc cycling on CTD is required for promoter activity (Ranuncolo et al, 2012) and O-GlcNAc needs to be removed prior to phosphorylation for elongation to occur as demonstrated previously (Comer & Hart, 1999). Many proteins involve in transcriptional regulation are also found to be heavily O-GlcNAcylated, including Sp1, c-Myc, Sin3A, p53, Tet1, Tet2, HCF1, CREB many of which also been found to directly interact with O-GlcNAc transferase (Chou et al, 1995a; Deplus et al, 2013; Goldberg et al, 2000; Hassig et al, 1997; Rexach et al, 2012; Vella et al, 2013; Williams et al, 2012; Yang et al, 2006); and the roles involve in regulating their protein stability, cellular localisation, interactions with transcription complex. Histones, which regulate chromatin dynamic and gene expression, are also found O-GlcNAc modified (Fujiki et al, 2011).

1.5.2. Roles in cell cycle and cell division

Increasing numbers of evidence have supported the connections of O-GlcNAcylation to cell cycle regulation and cell death. As mentioned previously, completely remove O-GlcNAcylation in mouse by OGT deletion causes

embryonic lethal (O'Donnell et al, 2004; Shafi et al, 2000); null mutations of both *ogt* genes, *Spy* and *Sec*, causes germination defect (Hartweck et al, 2002). In supporting this hypothesis, fibroblast growth arrest has been observed in mice with tissue specific *ogt* knockout (O'Donnell et al, 2004). Studies using *Xenopus* oocytes as model system have shown that microinjection with Galactosyltransferase that caps the O-GlcNAc inhibits aster formation during M-phase (Fang & Miller, 2001). This observation was further supported with the microinjection of morpholinos against OGT into oocytes and found to delay M phase entry; microinjection of the ncOGT isoform that also increase O-GlcNAc levels produced opposite effect by potentiating M-phase entry (Dehennaut et al, 2008). Later the same group has also shown that reduction of O-GlcNAcylation with OGT specific inhibitor slightly delay cell proliferation and they proposed that OGT activity is required for G1 to S transition (Olivier-Van Stichelen et al, 2012). With both genetic and pharcological method to reduce O-GlcNAc levels, it has shown that decreasing O-GlcNAc levels prevent cell division. However, similar results were also observed when increase O-GlcNAc levels. Reduction of O-GlcNAc by microinjection of O-GlcNAcase inhibitor PUGNAc or inhibitor DON that prevent UDP-GlcNAc biosynthesis increases oocytes maturation time (Slawson et al, 2002). Similar experiments were also performed in mammalian cell lines by microinjection of PUGNAc into HeLa or 3T3-L1 cells and growth delay was also observed (Slawson et al, 2005). Despite of the contradictory results produced by manipulating O-GlcNAc levels, it is clear that cell division is under tight regulation and also involves O-GlcNAcylation. Microscopic study has shown that both OGT and OGA are presence at the midbody during mitosis and

complex with phosphatase 1 and mitotic kinase Aurora B. This observation suggests that there is an extensive cross talk between phosphorylation and O-GlcNAcylation at midbody during cytokinesis (Slawson et al, 2008). Further study has identified more than hundreds of proteins that are involved in spindle assembly and cytokinesis (Wang et al, 2010). Most importantly, many of the O-GlcNAcylation sites are the same or adjacent to the phosphorylation site that involves in cytokinesis, supporting the role of O-GlcNAcylation during cell cycle event. In addition, intermediate filament vimentin is also a mitotic OGT substrate protein, whose O-GlcNAcylation was found increased after OGT overexpression and OGA inhibition; increase O-GlcNAcylation on vimentin decreases the phosphorylation that is crucial for vimentin rearrangement during mitosis, further support the involvement of O-GlcNAcylation in cell cycle regulation (Slawson et al, 2008).

1.6. Chronic Diseases Involve O-GlcNAcylation

Providing the potential involvement of O-GlcNAcylation in energy sensing, as well as a broad cellular activities of O-GlcNAcylated substrate proteins, it is not surprising that aberrant levels of O-GlcNAc in cells is often found to be associated with diseases.

1.6.1. Type II diabetes

Prolonged hyperglycaemia leads to insulin resistance, a hallmark of type II diabetes. High glucose levels increase the flux through the hexosamine biosynthetic pathway and increase global O-GlcNAc levels through the increase

of UDP-GlcNAc. Although malfunction of the hexosamine biosynthetic pathway has long been implicated in type II diabetes (Hart et al, 2011), the causality of elevated O-GlcNAc levels in the aetiology of the disease is unclear. The first direct link between global O-GlcNAcylation and type II diabetes came from studies using pharmacological inhibition of OGA activity, which resulted in a reduction of insulin-dependent glucose uptake (Park et al, 2005; Vosseller et al, 2002). However, these studies used the non-specific hexosaminidase inhibitor PUGNAc. A more recently developed OGA specific inhibitor did not induce insulin resistance in 3T3-L1 adipocytes despite the increase of global O-GlcNAc levels (Macauley et al, 2008). Non-pharmacological methods of modulating O-GlcNAc levels also gave contradicting results. Reduction of O-GlcNAcylation by increasing OGA or reducing OGT protein levels in 3T3-L1 adipocytes did not prevent glucose-induced insulin resistance (Robinson et al, 2006). However, transgenic mice with adipocyte and cardiac muscle overexpression of hOGT lead to insulin resistance and hyperleptinemia (McClain et al, 2002). Studies in other animal models, such as in *C. elegans* and *Drosophila*, also support the direct involvement of global O-GlcNAcylation and insulin signalling (McClain et al, 2002). Moreover, upon insulin stimulation OGT was found localized to the plasma membrane, where IRS proteins were O-GlcNAcylated (Yang et al, 2008). O-GlcNAcylation of IRS prevents phosphorylation-dependent activation by insulin stimulation and interaction with PI3K p85 subunit. In the long term, alteration of global O-GlcNAcylation affects the activity of transcription factors, such as Sp1, and therefore their downstream gene expressions, aberrant levels of which have been found involving in diabetes (Love et al, 2010).

1.6.2. Neurodegenerative diseases

The brain exhibits high level of OGT (Kreppel et al, 1997; Lubas et al, 1997). Many studies have reported the link between abnormal global O-GlcNAc levels and neurodegenerative diseases (ND). In a group of ND called tauopathies, the microtubule stabilising protein tau is found hyperphosphorylated and forming neurofibrillary tangles. These intracellular macroscopic aggregations are often found in brain tissue from patients with Alzheimer's disease (AD), while in normal brain tissue, tau is hyper O-GlcNAcylated. This competitive site occupancy is believed to prevent hyper phosphorylation in healthy brain tissue (Arnold et al, 1996; Lefebvre et al, 2010; Liu et al, 2004). Brain specific knock out of OGT protein in mice caused hyperphosphorylation of tau and abnormal locomotor activity (O'Donnell et al, 2004). Another pathology of the AD brain is neuronal plaques, which are caused by the aggregation of the degradation product of amyloid precursor proteins (APP). APP is found to be O-GlcNAcylated in the cytoplasmic domain and the modification may affect the cleavage of APP to form the toxic β-amyloid (Griffith et al, 1995). Many other proteins involved in brain functions are found to be O-GlcNAcylated, such as myriad synaptosomal proteins (Cole & Hart, 1999; Cole & Hart, 2001; Vosseller et al, 2006), AP-18, AP-3 (Yao & Coleman, 1998a; Yao & Coleman, 1998b), ataxin-10 (Marz et al, 2006), neurofilaments H, L and M (Dong et al, 1993; Dong et al, 1996). Although there is no direct report demonstrating the link of O-GlcNAcylation on those proteins to neurodegenerative disease, given the important roles of each protein plays in neuronal functions, O-GlcNAcylation on those proteins may serve functional purpose, especially with the interplay with phosphorylation.

1.6.3. Cancer

Aberrant levels of global O-GlcNAcylation are also implicated in cancer. However, the roles of O-GlcNAcylation in cancer are less clear. Many proteins involved in cancer development are found to be O-GlcNAcylated, including c-Myc (Chou et al, 1995a; Chou et al, 1995b), Rb (Wells et al, 2011), p53 (Yang et al, 2006), estrogen receptor β (Cheng & Hart, 2001). OGT is found localised at the midbody during mitosis, and overexpression leads to mitotic defects such as polyploidy, which is often associated with cancer (Slawson et al, 2005). On one hand, it has been shown that upregulation of OGT and O-GlcNAcylation have been found in certain types of breast cancer; reduced O-GlcNAcylation has been shown to inhibit tumour growth and metastasis *in vivo* (Ball et al, 2006). On the other hand, high O-GlcNAc levels may prevent tumour progression in chronic lymphocytic leukemic lymphocytes (Matthews et al, 2005). In this case, high O-GlcNAc levels may play a protective role in cancer development as opposed to the cancer promoting functions. It is not clear how the O-GlcNAc levels affect cancer cells and its transformation. Possibly, the involvement of O-GlcNAcylation in cancer is depending on the cancer types and their pathology.

1.6.4. Cardiovascular diseases

Elevation of global O-GlcNAc level may also protect against acute cardiovascular diseases. It has been shown that acute hyperglycaemia-induced hyper-O-GlcNAcylation improves the survival of heart tissue from ischemia and hypoxia (Champattanachai et al, 2008; Liu et al, 2007). Chronic hyper-O-GlcNAcylation

however causes cardiomyopathy as observed in diabetes (Hu et al, 2005; Liu et al, 2007).

1.7. General Classification of Glycosyltransferases

Glycosyltransferases (GT) catalyse glycosidic bond formation between activated high-energy glycosyl donors and a wide variety of acceptor substrates ranging from simple or complex carbohydrates to proteins, nucleic acid, lipids, antibiotics and other metabolites. Based on amino acid similarities, only 27 families of glycosyltransferases were documented in 1997 (Campbell et al, 1997; Coutinho et al, 2003). Over the past 15 years, studies have advanced our understanding of glycosyltransferases and allowed more accurate classification according to information such as the protein structure, substrate specificity, metal dependence and stereochemical outcome of the reaction (Haltiwanger et al, 1992). To date, there are 94 families of GTs documented in the CAZy database of carbohydrate-active enzymes (http://www.cazy.org/GlycosylTransferases.html). Most importantly, glycosyl donors are the sugar nucleotides derivatives, which are utilised by large numbers of glycosyltransferases. The sugar nucleotide was first discovered by Luis F. Leloir (Caputto et al, 1950) and the enzymes that use them are often referred to as Leloir enzymes.

1.7.1. Structural Insights into Glycosyl transfer

Structural enzymology of glycosyltransferases is challenging for a variety of reasons, and this class of enzymes is therefore less well understood than others. Many glycosyltransferases are membrane-associated proteins, and a high

content of flexible regions in the proteins also makes crystallisation difficult to achieve. To date, crystal structures of only 137 glycosyltransferases in total have been deposited in the PDB database, most of which in the last 10 years. The structural information has further expanded our understanding of this class of enzyme and the enzymatic action of the glycosyltransferase begins to emerge.

Despite the great diversities of the GTs, structural studies have revealed that all Leloir GTs adopt one of the two general folds, namely GT-A and GT-B folds, both of which contain two Rossmann-like folds, a β/α/β protein motif often found in nucleotide binding proteins and therefore allowing Leloir glycosyltransferase to bind to sugar nucleotide donor substrates (Fig. 1.8).

The first structure of Leloir glycosyltransferase, a DNA-modifying β-

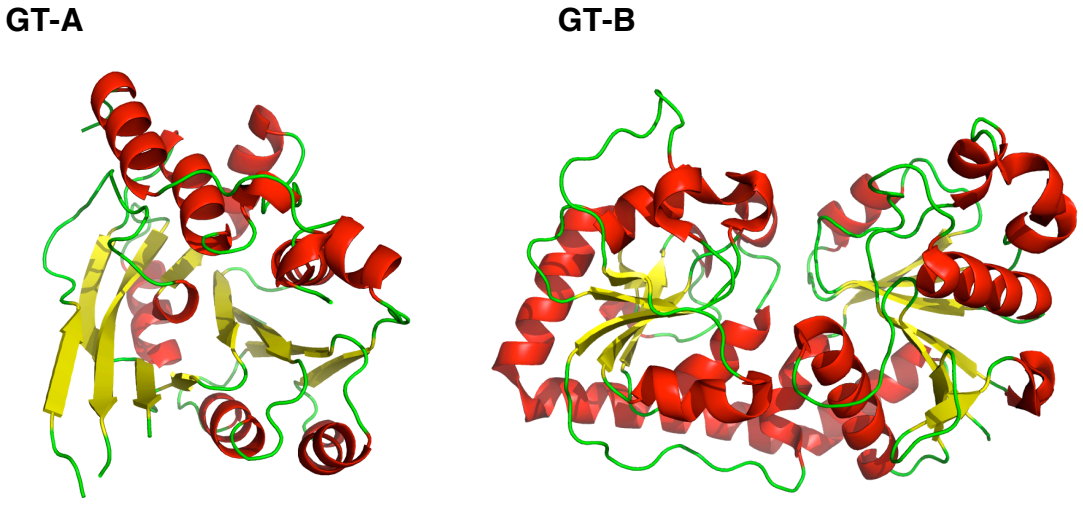


Figure 1.8: Cartoon representation of overall folds observed for glycosyltransferases. Cartoon representation of GT-A fold represented by SpsA from *Bacillus subtilis* (*A;* pdb: 1QGQ) and GT-B fold by bacteriophage T4 β-glucosyltransferase (*B*; pdb: 1JG7). α-helices are in red, β-strands are in yellow and loops are in green.

glucosyltransferase from T4 bacterial phage (BGT), was reported in 1994 (Vrielink et al, 1994). It reveals two less tightly associated Rossmann-like domains and a cleft in between, which was found later to harbour the active site (Lariviere & Morera, 2002). This overall arrangement of the Rossmann-like domains was observed for many other glycosyltransferases and denoted as GT-B fold. The first GT-A fold glycosyltransferase, SpsA from *Bacillus subtilis*, was reported in 1999 (Charnock & Davies, 1999a; Charnock & Davies, 1999b). In contrast to the GT-B fold, the structure revealed two closely associated Rossmann folds forming a single domain with a continuous central β-sheet. Most GT-A glycosyltransferases are metal dependent enzymes and posses an Asp-X-Asp (DXD) motif in which the negative charges of the carboxylates coordinate a divalent cation to interact with the phosphates of the nucleotide and facilitate the leaving group formation. This DXD motif is not often observed in the GT-B glycosyltransferases, which are generally metal-independent. The cation is replaced by positively charged amino acid side chains in the catalytic site. Non-Leloir GTs may adopt the GT-C fold, which contains a hydrophobic domain and embedded in the membrane (Lizak et al, 2011), as well as a fold resembling bacteriophage-lysozyme (Lovering et al, 2007).

1.7.2. Mechanism of Glycosyl Transfer

Stereochemistry of the anomeric carbon in a sugar allows two different outcomes of the glycosidic bond formation; the product either inverts or retains the configuration with respect to the donor substrate. The enzymes that catalyse glycosyl transfer can therefore be classified as either retaining or inverting

glycosyltransferases, depending on the outcome of the reaction. Glycosyltransferases employ either inverting or retaining mechanisms have been reported to adopt both GT-A and GT-B folds (Fig. 1.9).

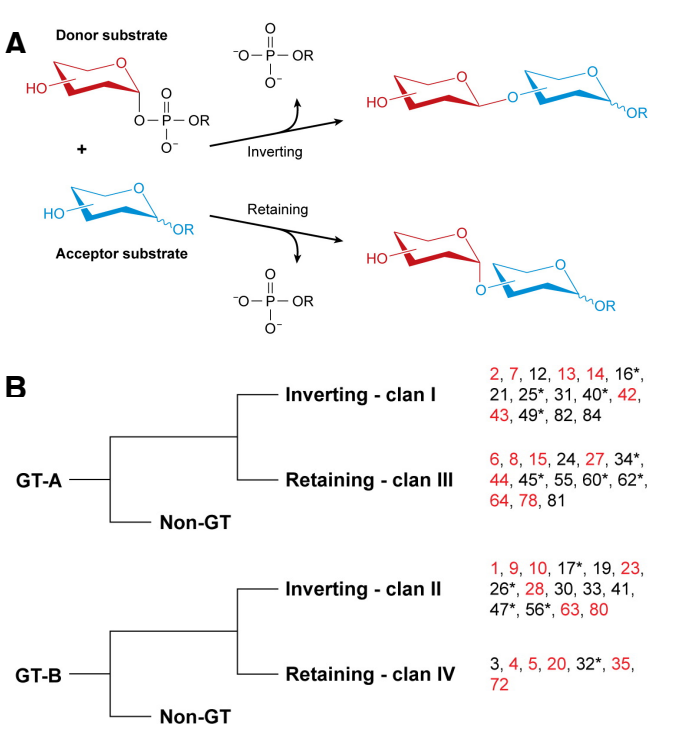


Figure 1.9: Schematic representation of general glycosyltransfer reaction and classification of GT families. The inversion or retention denoted on the anomeric stereochemistry with respect to the donor substrate (*A*). Both GT folds employ inverting or retaining mechanism (*B*). GT families belongs to each clan are indicated on the right. The sorting is based on predictions on the Cazy database and Liu *et al.*, 2003. Families with solved structures are indicated in red. Asterisks mark the protein fold predicted by Liu *et al.*, 2003.

The general principle of the inverting mechanism employed by glycosyltransferases is usually an S_N2-like direct displacement reaction: a catalytic base, usually a side chain from an active site amino acid, deprotonates the incoming hydroxyl group of the acceptor substrate and activates it for nucleophilic attack onto the anomeric carbon (C1) of the sugar moiety. Concomitantly, a positively charged species stabilises the negative charge developing on the leaving group, thus facilitating the substitution at the anomeric carbon; a new covalent bond is in turn formed on the opposite side (Fig. 1.10a). The catalytic base and the positive charges that stabilise the leaving groups are the key components of this process. Identification of this base, as well as the method to facilitate the leaving phosphate group, are essential in understanding an inverting glycosyltransferase.

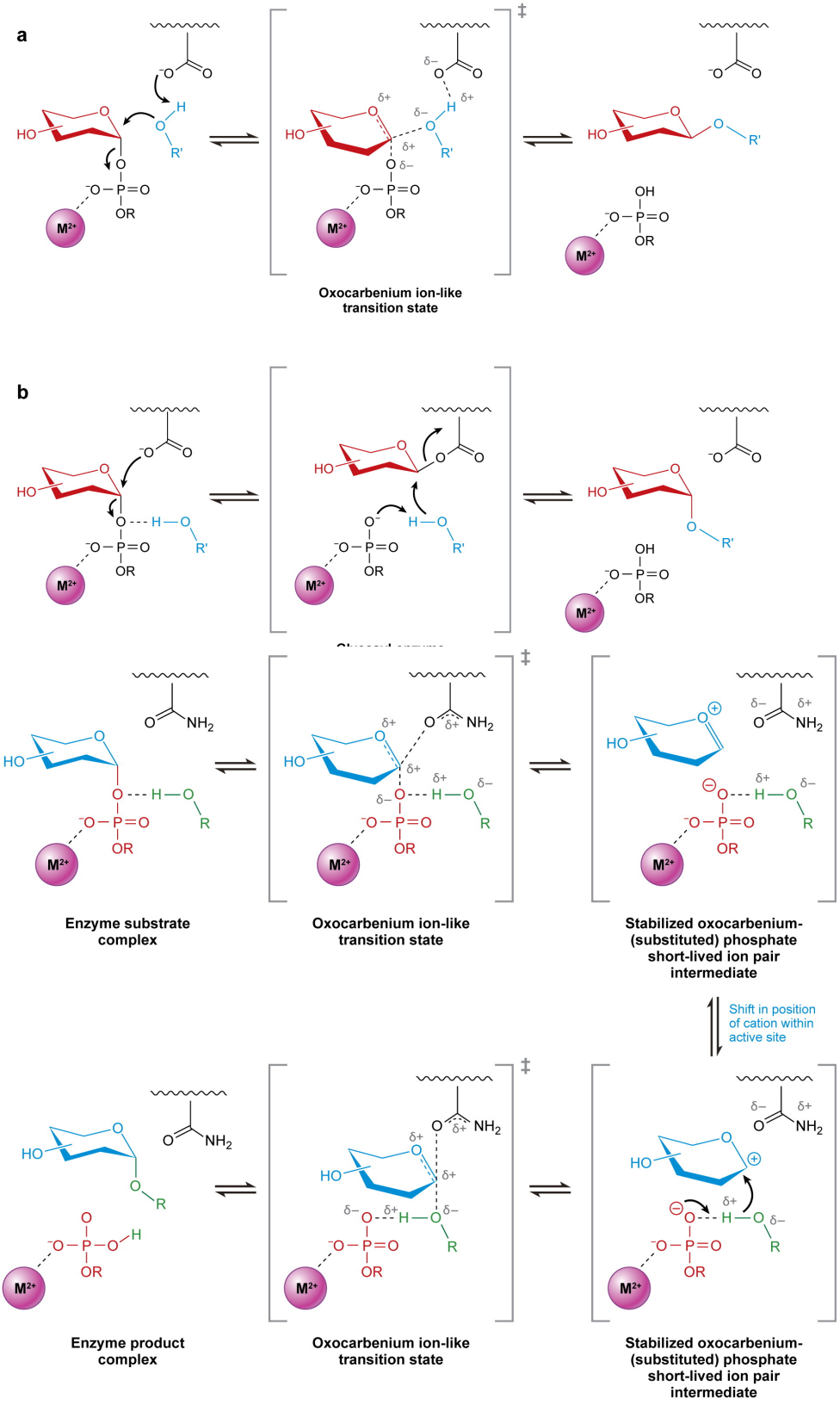


Figure 1.10: Schematic representations of reaction mechanisms of glycosyltransfer. Inverting glycosyltransfer (*a*) utilises a direct-displacement S_N-2 like mechanism resulting in a product with inverting stereochemistry at the anomeric carbon. Two proposed retaining glycosyltransfer mechanisms with (*b*) an SN-2 like double displacement forming temporary covalent sugar-enzyme complex or an alternative (*c*) an S_N-1 like dissociative mechamism by forming ion pair in the transation status. Noticeable that in the S_N-1 like retaining mechanism, the donor and the acceptor are on the same side of the sugar plane compared to the opposite side in the inverting mechanism. Figure is adopted from Lairson *et al.*, 2008, *Annu. Rev. Biochem.*

The aforementioned GT-A fold glycosyltransferase SpsA from *B. subtilis* belongs to an inverting glycosyltransferase family GT2. Despite the fact that both the donor and acceptor substrates are unknown, the crystal structure of SpsA was less informative, because a molecule of glycerol from the cryoprotectant was observed bound to the active site of SpsA, mimicking the acceptor substrate. The study led to the identification of an aspartic residue (Asp191) within hydrogen bonding distance to the glycerol, which was proposed to act as the catalytic base. Interestingly, in several structural studies, either Asp or Glu residues have been found at the equivalent position of the Asp191 in the active site of GT-A fold glycosyltransferases, within hydrogen bonding distance to the acceptors, confirming their involvement in the base catalysis (reviewed in Lairson et al., 2008). Mutagenesis studies at the relevant position of Asp191 led to loss of enzyme activity and demonstrated the essential role of this residue in the glycosyltransfer reaction (Garinot-Schneider et al, 2000).

The enzymatic base catalysis has likewise been reported in inverting GT-B glycosyltransferases. β-glycosyltranferase (BGT) from T4 bacteriophage was the first Leloir type glycosyltransferase using UDP-Glc as donor substrate (Vrielink et al, 1994). A conserved Asp100 was proposed to serve as the catalytic base based on structural and mutagenesis studies (Lariviere & Morera, 2002; Morera et al, 2001; Vrielink et al, 1994). The D100A mutation causes enzyme inactivation and enabled the crystallisation of a protein complex with the intact donor substrate (Lariviere et al, 2002). BGT is a metal-dependent inverting glycosyltransferase. However, the divalent ion serves no function in the

glycosidic bond cleavage but to facilitate product release. The metal ion was only found in the UDP complex, but not in the UDP-Glc complex, replacing the glucose. The negatively charged pyrophosphate group is found neutralised by the positively charged amino acid side chains in the active site and a helix dipole of the enzyme, instead of the metal ion. This arrangement is commonly found in many metal-independent inverting GT-B glycosyltransferases, such as GtfA, GtfB and GtfD (Mulichak et al, 2003; Mulichak et al, 2001; Webb et al, 2004). Proton extraction and activation of the acceptor could also be more indirect than the direct interaction with Asp or Glu side chains described earlier. Many glycosyltransferases of this class, such as UGT71G1 (Shao et al, 2005), VvGTI (Offen et al, 2006), OleD, OleI (Bolam et al, 2007), UGT2B7 (Miley et al, 2007), and bifunctional N/O-glycosyltranferase (Brazier-Hicks et al, 2007), are thought to utilize a catalytic dyad consisting of imidazole side chain of a histidine in conjunction with Asp or Glu.

The mechanism of retaining glycosyl transfer is less clear than the inverting mechanism. In general, two controversial mechanisms have been proposed: a S_N-2 like double-displacement with a covalent enzyme-substrate intermediate was originally proposed by Koshland (Koshland, 1953), and an alternative S_N-1 like oxocabenium ion intermediate has been proposed by Phillips (Phillips, 1966). In spite of extensive research, little evidence for such covalent intermediated exists, which makes the S_N-i-like mechanism more likely. This mechanism dispenses with the need for acid-base catalysis and postulates a late transition state where the enzyme facilitates dissociation of the donor to such an extent that

the reaction constitutes electrophilic migration of the resulting oxocarbenium ion rather than a nucleophilic attack. In retaining GTs the acceptor approaches the sugar from the same side as the leaving group, consequently the product itself has been described as the proton acceptor by Aktories et al (Ziegler et al, 2008).

1.8. Structural Insight into O-GlcNAc Transferase

Based on sequence conservation, OGT was classified into the GT41 family, a group of inverting GlcNAc transferases adopting the GT-B fold. OGT was first purified from animal tissue, and demonstrated that metal ions are not required for its enzymatic activity (Haltiwanger et al, 1992). All members of the GT41 family are believed to transfer sugars directly onto protein substrates, and all members also contain TPR repeats.

1.8.1. Structure of Tetratricopeptide repeats (TPRs)

TPR is a structural motif of 34 amino acids adopting a helix-turn-helix fold. The sequence is characterized by conserved amino acids located at the interface of the two helices, at positions 8, 20 and 27 by small hydrophobic residues and at positions 4, 7, 11 and 24 by large hydrophobic residue; a proline residue is often found in the loop connecting the helices (Allan & Ratajczak, 2011; D'Andrea &

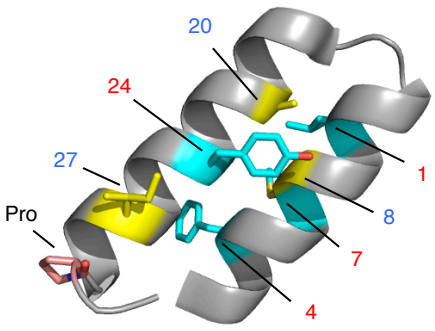


Figure 1.11: Arrangement of TPR motif
Cartoon representation of a single TPR motif from hOGT TPR 12. Large hydrophobic residues are shown by cyan sticks and the positions are numbered in red; while the small hydrophobic residues are shown in yellow sticks and numbered in blue. A proline is positioned at the turn shown by pink sticks.

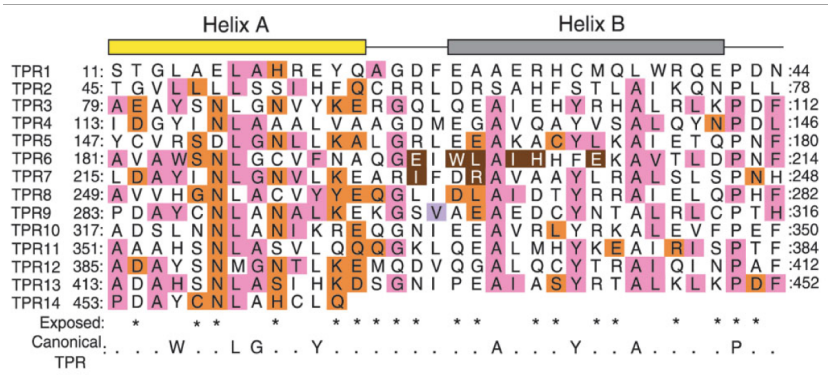


Figure 1.12: Schematic representation of the conservation of concave surface residues in hOGT TPRs.
Note the highly conserved Asp residues at the position 6 in each TPR. Conserved residues with charged side chains are marked in orange, other conserved residues are in pink. Residues forming interactions for dimerization are in brown. Essential amino acid residues for helix-turn-helix are listed in the consensus below each position. Figure is adopted from Jinek *et al*. 2006 *Nat Struct Mol Biol* **11:** 1001-1007.

Regan, 2003; Zeytuni et al, 2012) (Fig. 1.11 & Fig. 1.12). Mutations of these conserved residues disrupt the folding of the TPR (Jinek et al, 2004). Typically, TPR containing proteins consist of 3-16 copies of this motif, arranged in tandem array and forming a right-hand super-helical structure with a concave and a convex surface (Sikorski et al, 1990). Using bioinformatics techniques, more than 5,000 TPR containing proteins have been identified in all organisms, from bacteria to human; many of these proteins are involved in multiprotein complexes (Allan & Ratajczak, 2011; D'Andrea & Regan, 2003; Zeytuni et al, 2012).

The crystal structure of the major portion of the TPR from the long isoform of human ncOGT comprising 11.5 repeats was reported in 2004 (Jinek et al, 2004), providing structural and functional insight into substrate binding. Each of pair of antiparallel helices stacks with adjacent TPR units and the overall structure reveals a ‘corkscrew-like’ right-handed super spiral structure with a groove along the longitudinal axis (Fig. 1.13A). This topology shares great structural similarity

to the HEAT repeat from PP2α and the ARM repeat from importin-α (Conti & Kuriyan, 2000; Groves et al, 1999). Most conserved residues are restricted to the α-helices. More interestingly, highly conserved asparagine residues at the 6th position of each TPR repeat form a ladder along the concave surface (Fig. 1.12). This arrangement was also found in importin-α and shorter TPR-containing proteins Pex5, Hop, PP5 and Tom70 (Kikuchi et al, 2003; Ramasamy et al, 2005; Strahl-Bolsinger et al, 1993; Takahashi et al, 1996), in all of which the Asn ladders are found to interact with the backbone of a peptide that the TPR motifs bind (Fig. 1.13B). The authors also suggested substrate recognition might be at different sets of the concave surface and conserved Asn residues, analogously to importin-α, where different types of nuclear localisation signal are recognised by different sets of the conserved Asn residues, independent of the peptide sequence (Ramakrishnan et al, 2002; Ramakrishnan et al, 2006). The TPR domain of OGT also forms a dimer at TPR 6 and 7 through hydrophobic interaction. Mutations disrupt the interaction and reduce the enzyme activity by 40% (Jinek et al, 2004). Three isoforms of OGT with different length of N-terminal TPR have been shown to recognise different sub-groups of substrates. This isoform-specific substrate recognition has been suggested as an evolutionary solution for the challenge that a single *ogt* gene is available to modify a large number of substrates involved in various functions in the cell (Lazarus et al, 2006).

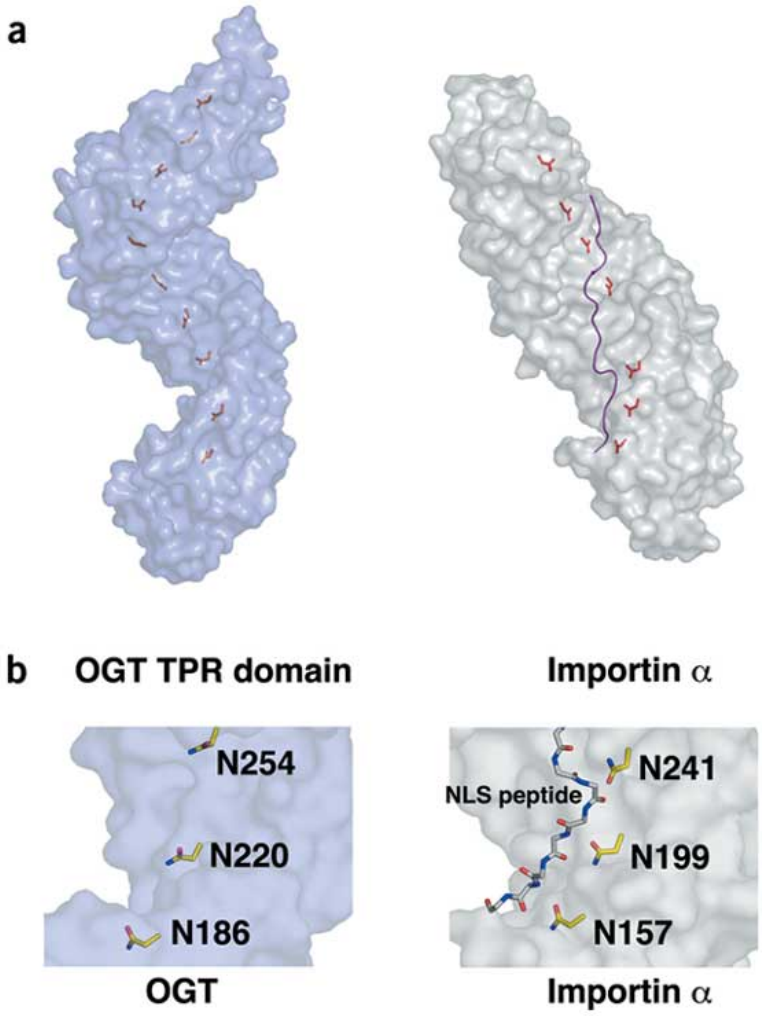


Figure I.13: Surface representation of the overall architecture of hOGT TPR domain in comparison with the ARM repeat from importin-α. The Asp residues at position 6 form a ladder along the concave surface of the superhelix (a, left). This arrangement has been found at the same position in the ARM repeat, in which the Asp residues are responsible for NLS peptide binding (a, right). Zoomed in view of Asp residues and interaction with NLS peptide (b). Figure is adopted from Jinek *et al.* 2006 *Nat Struct Mol Biol* **11:** 1001-1007.

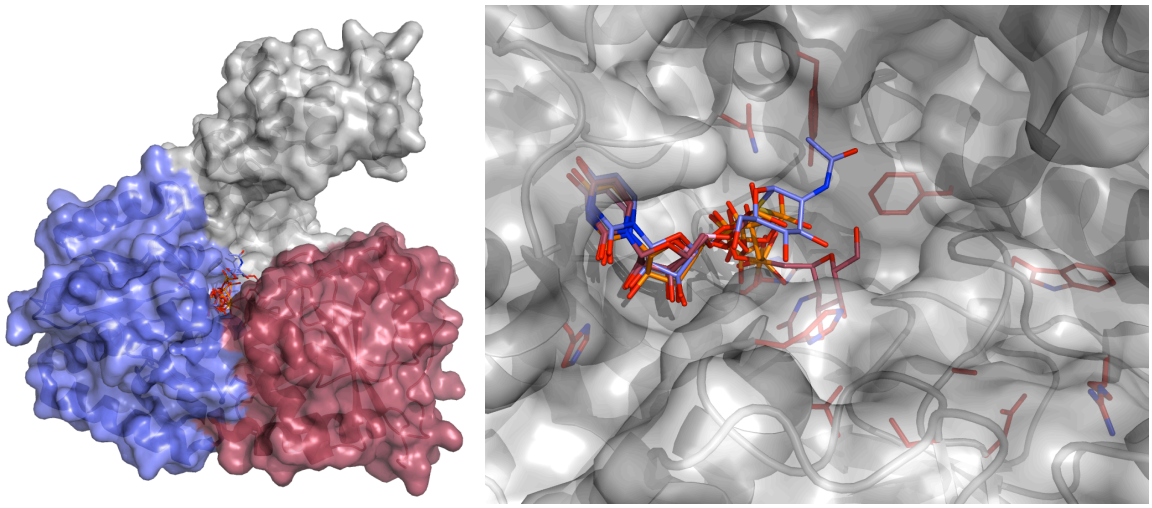


Figure 1.14: Cartoon representation of the XcOGT structure (left) and its catalytic site with UDP and its derivatives (right).
Two lobes of the catalytic domain (blue and red) with closely associated TPR repeat (grey) forming a groove. UDP is deeply buried in the pocket, surrounded by a group of highly conserved residues (red). The uridine from UDP (pdb: 2VSN), UDP-C-GlcNAc (pdb: 2JLB), UDP-S-GlcNAc (pdb: 2XGO) and C-UDP (2XGS) are perfectly superimposed, while the phosphate and sugar moiety are less so.

1.8.2. OGT Glycosyltransferase Domain

The first crystal structures from family GT41 came from an OGT orthologue from the bacterial plant pathogen *Xanthomonas campestris* (Clarke et al, 2008; Martinez-Fleites et al, 2008). Although this bacterial protein lacks the intervening domain found in all metazoan OGTs, *Xc*OGT shows 36% sequence identity to hOGT in the catalytic domain, and has three classic TPRs. The *Xc*OGT structures confirmed that OGT adopts the GT-B fold. Demonstrated by several complexes with UDP and UDP-GlcNAc derivatives, the active site is located in the groove between the two Rossmann-like folds as in all GT-B fold glycosyltransferases (Fig. I.12b). The TPR at the N-terminus abuts on the catalytic domain and the superhelical groove of the TPR motif extends into the

active site, forming a continuous groove in which the substrate peptide may be recognised and bound. Unfortunately, the natural acceptor and donor substrates for *Xc*OGT are not known. In 2011, the crystal structure of a truncated hOGT construct, bearing 4.5 TPRs and the catalytic domain, was solved by the Walker group (Lazarus et al, 2011). Most importantly, two structures of hOGT and UDP with or without acceptor peptide were reported. These complexes provided the first evidence of how OGT recognises the substrate at the peptide level, and demonstrated unambiguously through structural and kinetic studies that OGT uses an ordered bi-bi catalytic mechanism. As observed in the *Xc*OGT structure, the active site groove is a narrow cleft between the TPRs and the catalytic domain in the OGT-UDP structure, however, this is widened slightly in the presence of the acceptor peptide. Although the two complexes were obtained from different crystallisation conditions, this suggests that the TPRs possess

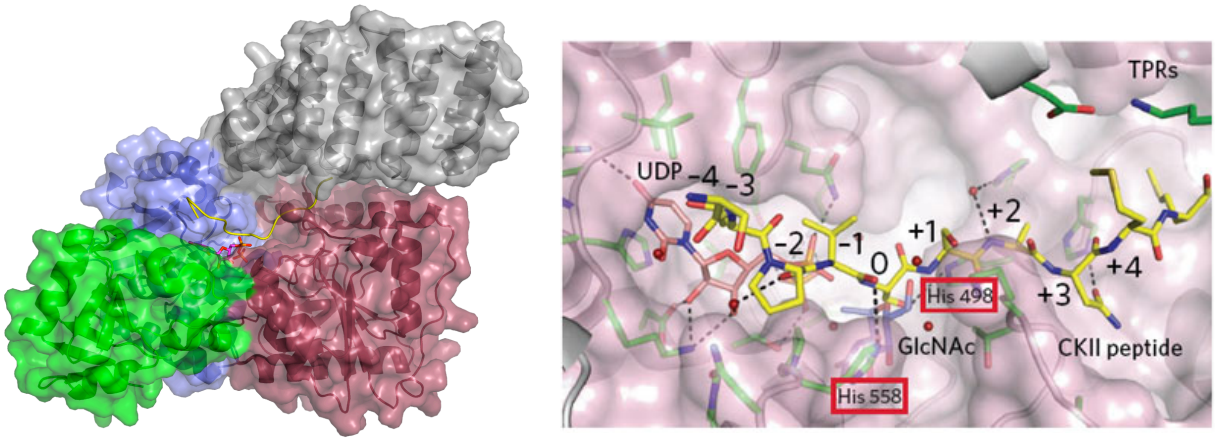


Figure 1.15: Cartoon representation of the hOGT catalytic domain with 4.5 TPR (left) and its catalytic site with UDP and a peptide from hOGT substrate CKII (right).
Structure shows the architecture of catalytic domain (red and blue) with 4.5 TPR (grey) is similar to the XcOGT and the novel intervening domain (green) that is absent in XcOGT. Zoom in view at the catalytic site surface (pink) shows the binding of the CKII peptide (yellow) and its limited interaction with hOGT. CKII peptide covers the UDP-GlcNAc binding pocket, suggesting an ordered *bi-bi* catalytic mechanism. Amino acid residues interact with UDP (salmon) and create a pocket for GlcNAc binding. The position of the GlcNAc moiety (blue) is the result of a modelling experiment. Suggested catalytic bases His498 and His558 have been marked by red boxes.

plasticity and can change conformation to allow substrate binding. No OGT-UDP-GlcNAc complex was obtained due to the hydrolysis, as the authors suggested.

The metazoan-specific intervening domain that was absent in the *Xc*OGT structure was reported to adopt a topologically novel fold containing a seven-strand beta-sheet stabilised by flanking helices. This intervening domain also forms a large positively charged patch of ten lysine residues on the opposite surface from the catalytic groove. These lysine residues were previously suggested to be the PIP_3 binding motif responsible for the hOGT membrane localisation upon insulin stimulation (Yang et al, 2008).

Before the work from this thesis, the mechanistic action of OGT on O-GlcNAc transfer on substrate proteins was unclear. All the speculations were based on the crystal structure of the catalytic domain from bacterial ortholog *Xc*OGT (Clarke et al. 2008; Martinez-Fleites et al., 2008). Also, a recent published crystal structure of hOGT catalytic domain (Lazarus et al., 2011). The new hOGT structure, however, lacks the GlcNAc moiety and clear data to support for the hypothesis of the catalytic mechanism, as well as the involvement of the TPR in the catalysis.

In the following chapters, I have used a combination of structural biology together with the support of biochemical and biophysical to prove the hypothesis and demonstrated the unusual catalytic mechanism employed by hOGT. Later in the

chapters, a *Drosophila* system was employed to demonstrate the important of this catalytic activity for the survival and early development of flies.

2. Chapter I

Functional Study of Tetratricopeptide Repeat Domain in Human O-GlcNAc Transferase

2.1. Project Overview

O-GlcNAc transferase (OGT) post-translationally modifies intracellular proteins on Ser or Thr residues with a single GlcNAc moiety and has been shown to be essential for embryogenesis and early development in metazoa (Hart et al, 2011; Love et al, 2010). The protein is comprised of N-terminal tetratricopeptide repeats (TPRs) and the C-terminal glycosyltransferase domain. Three splice variants, with differing lengths of the N-terminal TPR domains, have been identified encoded by a single gene (Shafi et al, 2000). The longest isoform (ncOGT) contains 13.5 TPRs and the short isoform (sOGT) contains 2.5 TPRs, both these isoforms are uniformly localised in the cytoplasm and nucleus; while in the third isoform the first 5 TPRs are replaced by a mitochondrial specific sequence (MSS) containing 8.5 TPRs, and the protein is localised specifically in the mitochondria (mOGT).

Where structural information is available, it indicates that the ligands binding to TPR domains frequently consist of flexible peptides, which may be found docked in the concave surface formed by the TPR superhelix (Quinaud et al, 2007; Sampathkumar et al, 2008; Wang et al, 2009; Zeytuni et al, 2012). Although TPRs are highly conserved at the above-mentioned positions, the surface-exposed amino acids are less stringently conserved. In contrast to other similar helix domain containing proteins, such as 14-3-3 proteins (Coblitz et al, 2006), no consensus sequence has been identified among the ligands of TPR domains. However, individual TPR domains have been shown to exert some selectivity for the peptide sequence of the binding partner

(Sampathkumar et al, 2008; Zeytuni et al, 2012). However, flexible peptides are not the only type of ligand recognized by TPR domains, other methods of binding have been observed in several structural studies, where TPR motifs were found to bind α-helices at both the concave and convex surfaces, as well as proteins containing multiple TPR domains and dimerization mediated by the TPR (Jinek et al, 2004; Sampathkumar et al, 2008; Zhang et al, 2010).

It has been proposed that interactions with the TPR domain of OGT enables it to act on a large variety of substrates, as well as providing isoform selectivity for specific substrates (Chen et al, 2013; Iyer & Hart, 2003; Lubas & Hanover, 2000; Yang et al, 2002; Zhang et al, 2012b). The available crystal structure of the TPR domain of ncOGT contains 11.5 of the 13.5 repeats, revealing the canonical right-hand superhelical structure (Jinek et al, 2004). In this structure, groups of highly conserved polar or charged residues have been shown to cover the concave surface of the superhelix. Especially, recurrent arginine residues were noticed at position 6 on each TPR motif, forming a ladder that resembles the Asn-ladder observed in the armadillo repeats in importin-α (Conti & Kuriyan, 2000). The conserved concave surface and the Asn-ladder suggest that this surface may represent a binding interface for peptide ligands, which would bind along the axis of the TPR super-helix. Multiple peptides binding to one TPR domain has been demonstrated in peroxin 5 (PEX5), a cargo adaptor protein from *Trypanosoma brucei*. Pex5 contains 7 TPRs, which are able to bind type 1 or type 2 peroxisomal targeting signal peptides; crystal structures of complexes of PEX5 and five

ligand peptides showed the TPRs clamping the peptides in position, all of which adopt similar conformations (Sampathkumar et al, 2008). In addition, adaptor protein Hop contains two individual TPR domains with selective binding affinity for Hop binding partners Hsp70 and Hsp90 (Scheufler et al, 2000). These finding suggest that OGT substrates may bind in different positions on the OGT TPR domain, forming specific interactions with individual repeats.

TPR participation in the recognition of individual substrates by OGT has been the subject of several independent studies (D'Andrea & Regan, 2003; Iyer & Hart, 2003; Love et al, 2010; Quinaud et al, 2007; Zhang et al, 2010). In this chapter, I describe the detailed study of the functions of the TPR domain of OGT by systematic removal of TPRs from the N-terminus, and by determining the effect of these truncations on O-GlcNAcylation of substrate proteins with different structural properties. I propose that substrate interactions with the TPR domain may induce conformational changes in the substrate, thereby giving access to the site of O-GlcNAcylation. This leads to the hypothesis of two OGT substrate-binding models: the TPR-independent recognition of loose peptides or flexible regions as opposed to a model of TPR-induced fit.

2.2. The aims of the work presented here were:

- To identify an expression construct to produce proteins for future crystallization studies
- To study the roles of the TPR domain in OGT substrate recognition by systematic TPR deletion

2.3. Results (Part1):

2.3.1. Design of an OGT expression construct through systematic N-terminal truncations lead to the stable and active ΔTPR hOGT

Three splicing isoforms of human OGT have been reported, and the longest splice variant is found in both cytoplasm and nucleus (ncOGT) (Love et al, 2003). The protein is 1046 amino acids in length with a calculated molecular weight of 117 kDa. ncOGT comprises of two principle domains, the N-terminal thirteen and a half tetratricopeptide repeats (TPRs) and the C-terminal glycosyltransferase (GT) domain. Mitochondrial OGT (mOGT) and short OGT (sOGT) are both shorter than ncOGT with 8.5 and 2.5 TPRs respectively, and with an identical C-terminal GT domain (Fig. 1.4). Previously, recombinant expression of the ncOGT was most successful using the insect cell sf9/sf21 baculovirus system, while bacterial expression has been problematic and resulted in extremely low yield and purity. A model built on the basis of the

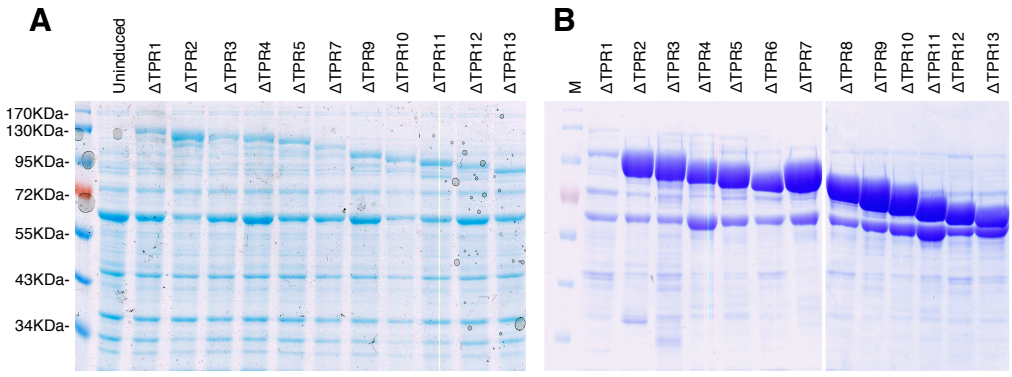


Figure 2.1: Expression and purification of hOGT (nc) TPR truncations in ArcticExpress cells.
A: Total lysate of the *E coli* from 10 μl of the expression culture were boiled in sample buffer. Samples were separated on a SDS-PAGE and the gel was stained by Coomassie blue.
B: hOGT proteins with TPR truncations were purified and the purity was assessed by SDS-PAGE. The gradual change in size was noticeable. ΔTPR1 is expressed but fail to be purified, suggesting the protein is unstable or insoluble.

two structures available at the beginning of this project, namely the TPR domain from ncOGT and the bacterial homologue GT domain has shown that while the last repeat of the TPRs closely abuts on the GT domains, the remaining TPRs form a spiral extending away from the catalytic core. This was taken as indication that the removal of the TPRs from the N-terminus would have little effect on the catalytic activity and that N-terminal truncations could be more amenable to recombinant expression and yet remain suitable for studying the enzymatic function of OGT. Accordingly, I made serial truncations sequentially removing TPR repeats, with construct boundaries derived from the structural information, and expression constructs were generated in pGEX6p1 vector for bacterial expression with a cleavable GST tag.

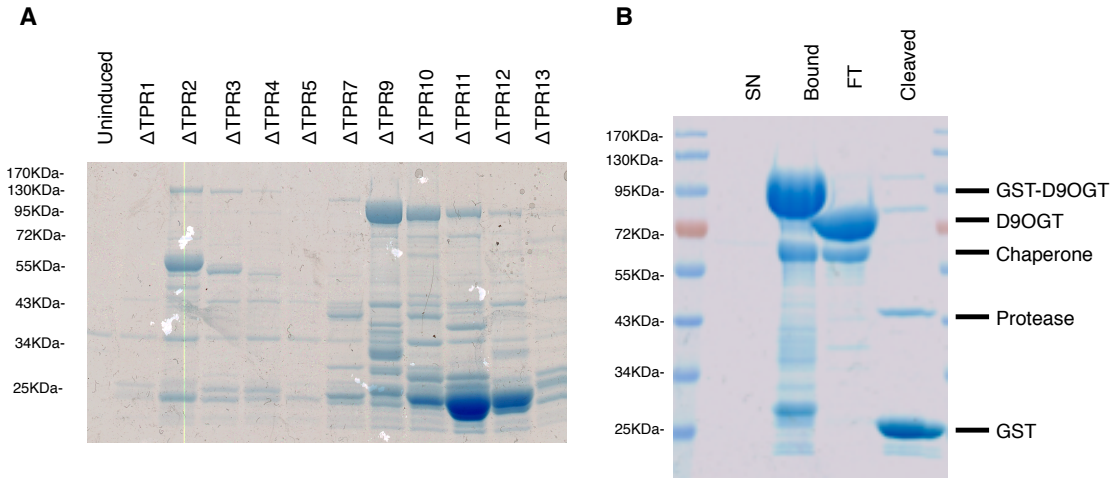


Figure 2.2: the stability of each TPR truncations (A) and primary purification of Δ9 hOGT (B). **A**. Equal amount of hOGT proteins bound to the glutathione beads were taken and incubated at room temperature for 24 hrs without protease inhibitor. The products were analyzed by Coomassie blue protein stain. **B**. Δ9 hOGT was expressed in *ArcticExpress E coli*. GST-tagged protein was purified on glutathione sepharose beads (Bound) and the tag was cleaved off by Prescission protease, releasing D9 hOGT to the flow through (FT) and leaving the protease on the beads.

After comparing the amount of soluble protein obtained from expression in different *E. coli* strains, a commercial ArcticExpress strain, which is engineered to co-express a chaperone protein to help with protein folding at low temperatures, was selected as the expression system (data not shown). Although a chaperone protein was co-purified in the first stage of purification (Fig. 2.1B), this was efficiently removed by anion exchange chromatography (Fig. 2.3A, B). Further purification by size exclusion chromatography was judged unnecessary (Fig. 2.3C, D). All 13 TPR truncations could be produced in this way, but notably, the deletion of 9 TPR repeats (Δ9-hOGT, amino

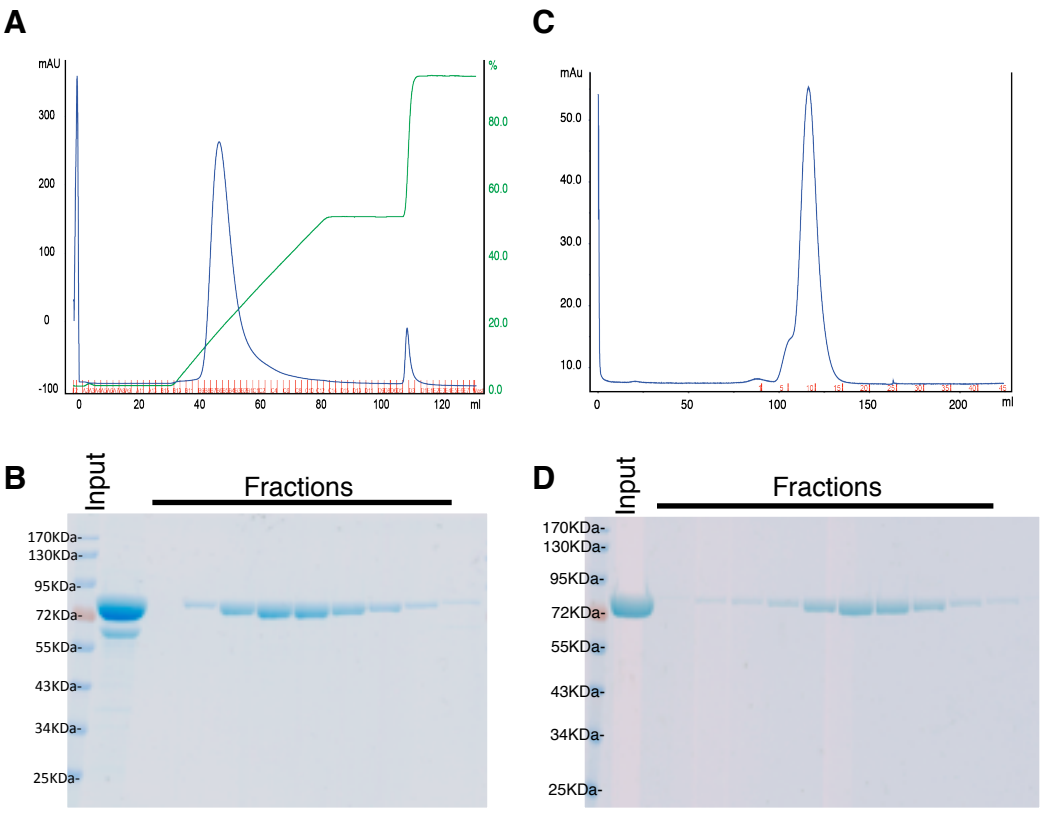


Figure 2.3: Further purification of Δ9 hOGT by anion exchange chromatography and size exclusion chromatography.
A, B: Affinity purified Δ9-hOGT (B, input) was loaded onto HiTrap Q FF inion exchange column. Protein was eluted with a 0-1 M NaCl gradient (Green line). Purity of each fractions were checked by Coomassie blue staining before pooling together. C, D: Δ9-hOGT was further purified by size exclusion chromatography on a *Superdex* 75 column. Purity of the fractions was checked by Coomassie blue staining.

acids (314-end) gave the highest expression levels and was shown to be more stable than the other truncations (Fig. 2.2). From a 12 L culture around 15 mg of pure Δ9-hOGT protein was produced. Δ9-hOGT protein was later used in crystallisations trials described in Chapter II.

2.3.2. hOGT TPR on substrate specificity

As a first step in investigating the roles of individual hOGT TPR motifs in recognition of specific substrates, *in vitro* activity of hOGT with TPR truncations was determined on total HEK293 cell lysate. Mammalian cells grown in culture possess a high level of protein O-GlcNAcylation. To remove this background O-GlcNAcylation and thereby improve the signal margin, the cell lysate was treated with an active bacterial O-GlcNAcase, *Cp*OGA, to remove all O-GlcNAc prior to *in vitro* O-GlcNAcylation with OGT TPR truncation mutants (Fig. 2.4). In order to minimise any interference of

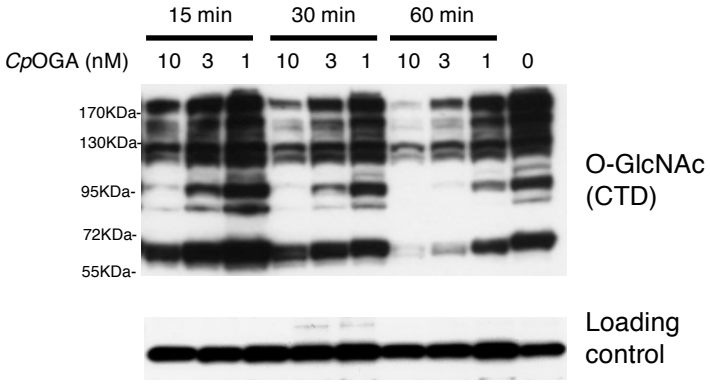


Figure 2.4: removal of O-GlcNAc modification on HEK293 lysate using *Cp*OGA
50 µg of crude lysate from HEK293 cells was treated with indicated amount of CpOGA for different period of time. O-GlcNAcylation was detected by Western blotting using CTD antibody.

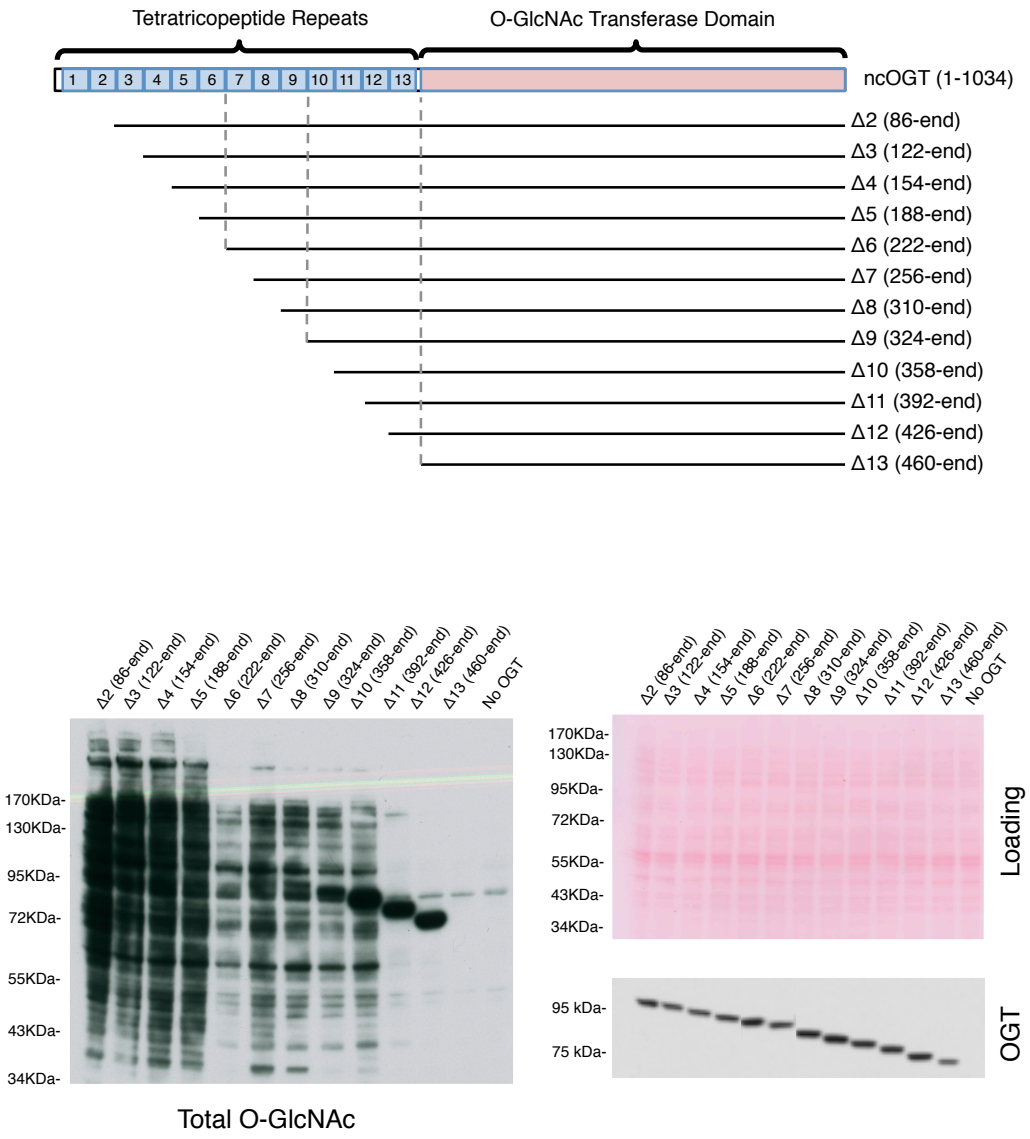


Figure 2.5: *in vitro* activity of hOGT with TPR truncations on 293 cell lysate
A. Schematic representation of the TPR truncations.
B. HEK293 cell lysate has been treated with *Cp*OGA to remove the O-GlcNAcylation. The cell lysate was subjected to recombinant hOGT with C-terminal TPR truncations. Reaction was performed in the presence of UDP-GlcNAc for 30 mins and stopped by boiling in sample buffer before SDS-PAGE and Western blotting. Total O-GlcNAc signal was detected with RL2 antibody. The presence of recombinant OGT was detected by anti OGT antibody. Loading was confirmed by Ponceau Red.

endogenous hOGT and hOGA, the reaction was performed in the presence of 1 µM of the O-GlcNAcase inhibitor GlcNAcstatin C (K_i=4.6 nM, (Dorfmueller et al, 2009)), using high concentrations (up to 0.5 µM) of recombinant hOGT, which allowed for short reaction times (30 mins). Total O-GlcNAc levels did not gradually decrease with progressive removal of TPRs, rather, there appeared to be a step-wise loss of activity (Fig. 2.5): hOGT lacking up to 5 TPRs in hOGT (Δ2-Δ5 hOGT) was able to produce high levels of O-GlcNAc on lysate *in vitro*. A dramatic decrease in total O-GlcNAcylation was observed when TPR 6 was removed. Similarly low levels of O-GlcNAcylation were obtained for all truncations up to Δ11-hOGT. The removal of TPR 12 and TPR 13 completely abolished OGT activity on total cell lysate. Notably, unlike the full-length enzyme, certain truncation mutants were prone to autoglycosylation. Auto-O-GlcNAcylation was observed on hOGT with TPR 10, TPR 11 and TPR 12 deletions (Δ10-, Δ11-, Δ12-hOGT), but not on TPR 13 deletion (Δ13-hOGT). This highlights the preservation of enzymatic activity of

Table 2.1: Summary of protein substrates tested for in vitro O-GlcNAcylation with TPR truncation mutants

Name	Length	Construct	Flexibility	O-GlcNAc site(s)	References
TAB1	504aa	7-402	Yes	S395	Pathak *et al.*, 2011
CK2α	391aa	1-391	Yes	S347	Tarrant *et al.*, 2012
FOXO1	665aa	2-655	Yes	T317, T318 (T648, S654 predicted)	Housley *et al.*, 2008
CREB	341aa	1-283	No	S40, T228, S260	Rexach *et al.*, 2012
GSK3b	420aa	2-420	No	N-terminus	Trinidad *et al.*, 2012
YES1	543aa	1-543	Yes	N-terminus	Trinidad *et al.*, 2012

Δ10-, Δ11-, and Δ12-hOGT, which had showed little activity on cell lysates.

Further in-depth investigations of the effect of hOGT TPR truncations were performed on several *bona fide* O-GlcNAc target proteins. These included the transcription factors FOXO1 and CREB, a kinase GSK3β, the innate immunity signalling protein TAB1, casein kinase 2 subunit α (CK2α), and a Src family kinase (SFK) Yes1 (Housley et al, 2008; Lubas & Hanover, 2000; Pathak et al, 2012; Rexach et al, 2012). Recombinant proteins or relevant domains, CREB (1-283), FOXO1 (2-655), GSK3β (2-420), TAB1 (7-402), CK2 (1-391) and Yes1 (1-543), were expressed in *E. coli* or insect cells and provided by DSTT (the Division of Signal Transduction and Therapy of the University of Dundee). *In vitro* activities of hOGT with TPR truncations were determined on these proteins (Fig. 2.6). The data showed that TPR truncations in OGT has a negligible effect on the O-GlcNAcylation of TAB1 and CK2, as well as full-length FOXO1 (upper band at around 65 kDa); all TPR truncations of hOGT are capable of modifying these proteins, except for the TPR 13 deletion, which did not show any activity on these substrates. The recombinant FOXO1 protein showed significant levels of degradation products, which also appear to be substrates for hOGT, appearing on Western blots as multiple bands below the full-length protein. Compared to the full-length protein, the levels of modification on these lower bands appear to be more sensitive to truncations of the enzyme. OGT activity remains at an evenly high level when 6 TPRs are removed, is reduced when 7 TPRs are removed, and completely abolished upon deletion of more than 7 TPRs. A very similar pattern is seen for CREB,

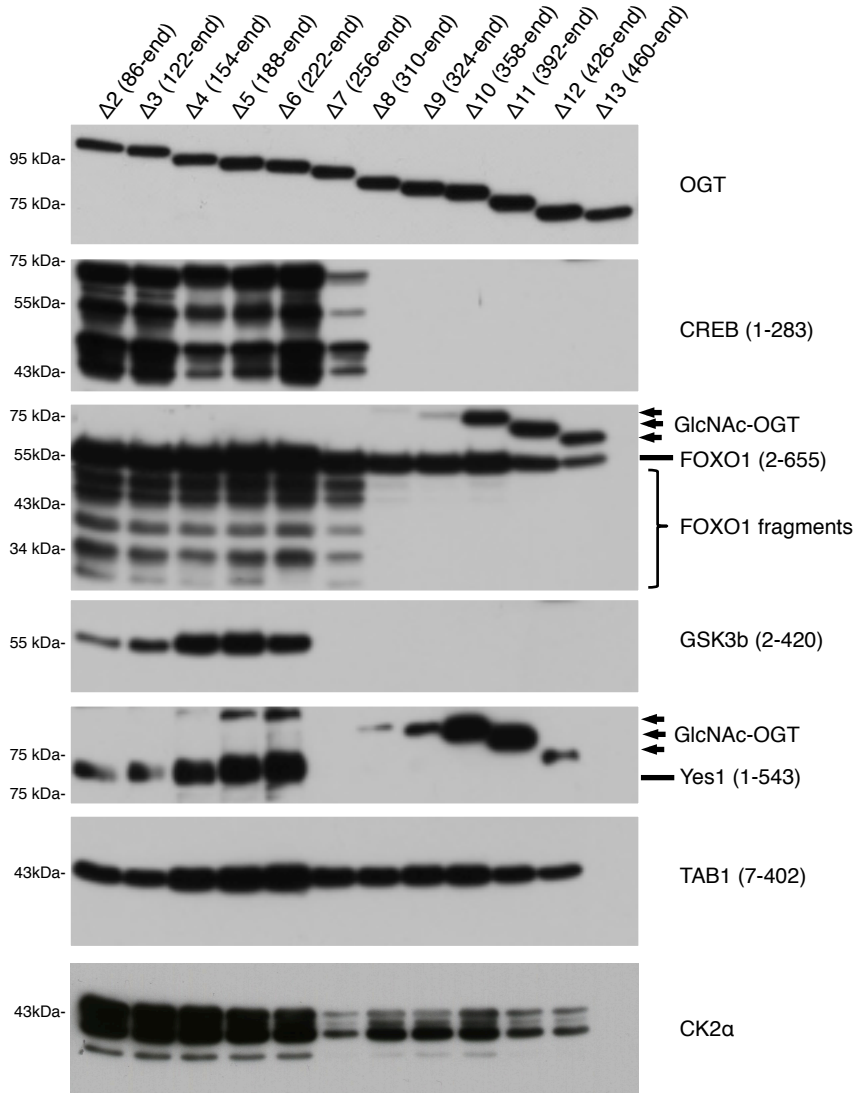


Figure 2.6: Affects of TPR truncations in hOGT on selected OGT substrate proteins

Recombinately expressed OGT substrate proteins were subjected to *in vitro* O-GlcNAcylation with TPR truncation mutants of hOGT. Reaction was stopped by boiling in sample buffer before SDS-PAGE and Western blotting. O-GlcNAcylation on the substrate proteins was detected by anti O-GlcNAc RL2 antibody. It is noticeable that CREB and FOXO1 are degraded and generate protein fragments. Δ10-, Δ11- and Δ12-hOGT O-GlcNAcylated, which was not detectable in other truncation proteins

which is a substrate for full-length OGT and truncations up to TPR6, and to a limited degree up to TPR7. The boundary is even more pronounced for GSK3β and Yes1; removing more than 6 TPRs results in loss of OGT activity on GSK3β and Yes1 (Fig. 2.6). As summarised in Table 2.1, O-GlcNAcylation sites on TAB1 and CK2α are Ser395 and Ser347, both of which are located in a flexible region of the respective protein (Pathak et al, 2012; Tarrant et al, 2012). O-GlcNAcylation on FOXO1 has been reported at Thr317 and Thr318 (Housley et al, 2008) while two additional sites, Thr648 and Ser654, at the C-terminus have also been suggested. Thr317 and Thr318 are predicted in the structured region while the final two sites, Thr648 and Ser654, are close to the predicted disordered C-teminus. CREB is O-GlcNAc modified at Ser40, Thr228 and Ser260, all of which are likely, according to secondary structure predictions, to be located in structured regions of the protein (Rexach et al, 2012). Modification sites on GSK3β and Yes1 have not been reported, however, mass spectrometry studies have suggested that O-GlcNAc modifications are located close to the N-termini of both proteins.

2.4. Discussion (Part1):

2.4.1. Dimerization and substrate recognition

In ncOGT, approximately half of the protein forms the TPR domain, which is shorter in mOGT and sOGT (Shafi et al, 2000). It has been shown that the TPR domain contributes to substrate specificity (Chen et al, 2013; Iyer & Hart, 2003; Lubas & Hanover, 2000; Tseng et al, 2001; Yang et al, 2002; Zhang et al, 2012a). Structural analysis of the TPRs from hOGT showed a superhelical

structure that protrudes away from the catalytic site (Jinek et al, 2004; Lazarus et al, 2011). This architecture creates a large surface area and a large groove running along the inside of the helix, possibly aiding in the interaction with substrate proteins. The extensive surface of the TPR domain suggests that the protein-protein interaction may occur at specific positions on the TPR motifs, which would create a large diversity for OGT substrate recognition. One would expect that systematic removal of individual TPR repeats from the N-terminus might produce a gradual effect in reducing the number proteins that are recognised in a pool of substrates. However, this gradual effect was not observed *in vitro* when the series of truncated hOGT constructs was tested on total cell lysate. Instead, the activity of OGT was decreased in three major steps along with the TPR removals.

In the series of TPR truncations of OGT, deletions of the first TPRs from N-terminus up to TPR 5 retained high activity on cell lysate proteins indistinguishable from the full-length enzyme, but the activity is dramatically decreased after TPR 6 is removed, which suggests its importance in OGT activity on substrate proteins. It has been shown that OGT forms a homodimer in cells and a crystal structure of the TPR domain of hOGT suggests that several residues in TPRs 6 and 7, especially the hydrophobic residues (W198, L199, I201, I230), lie at the dimer interface and are indispensable for dimerization (Jinek et al, 2004); mutations of these residues disrupt the dimerization and lead to OGT activity loss by 50% on nuclear pore protein p62. In our experiment, deletion of TPR 6 could lead to disruption of

the OGT homodimer and decrease OGT activity as in the previous report (Jinek et al, 2004).

The next significant decrease in O-GlcNAc signal was observed upon the deletions of TPRs 12 and 13, which almost abolished the O-GlcNAcylation of lysate proteins. Auto-O-GlcNAcylation on Δ12-hOGT indicates its enzymatic activity, unlike OGT with TPR13 truncation, for which no evidence of enzymatic activity was observed. However, it is not clear whether this auto-O-GlcNAcylation occur in a *cis-* or *trans-*. One way to find it out is to perform an in vitro assay with inactive tagged Δ12-hOGT and detect for O-GlcNAc modification. The position of TPRs 12 and 13 are shown in the crystal structure of the hOGT catalytic domain (Fig. 1.15). TPR 13 in particular is closely associated with the glycosyltransferase domain (Lazarus et al, 2011). From the structure, both TPRs 12 and 13 do not participate directly in forming the substrate-binding groove, but create the start of a channel on the concave surface of the TPR super-helix. The presence of substrate peptides in hOGT crystal structures defines the peptide binding site. Interestingly, the C-terminus of the peptides were pointed towards the channel created by the TPRs, however were not long enough to provide direct evidence for such interactions. A slight overall rotation of the TPR domain with respect to the catalytic domain (6 Å) has been observed when a peptide substrate was bound, creating a more 'open' conformation of the groove (Lazarus et al, 2011). The opening of the TPR groove may allow extended polypeptides to bind and interact with highly conserved surface residues. Deletion of TPRs 12

and 13 could affect the interactions with the peptide and would be particularly likely to affect the activity on larger polypeptides, as observed in our experiment. TPR 13 especially is in close contact with the catalytic domain and might be required for the structural integrity of the enzyme; deletion appears completely inactive the enzyme, as was observed for all substrates tested, including the proteins TAB1, CK2α and FOXO1 (Fig. 2.6).

Giving the fact that three OGT isoforms differ in numbers of TPR present in the N-terminus, it has been shown isoform specific substrate recognition on a selection of substrate proteins (Lazarus et al., 2006). The activity of sOGT was not detected on any of the substrate the authors tested. sOGT contains 2.5 TPRs, equivalents to removal of 10.5 TPRs. It has been demonstrated here that removing of 10 and 11 TPRs from ncOGT abolished substrate recognition on many substrates tested, but still retains activity on substrates, which adopts flexible conformation at the modification site. This observation is in agreement with previous reports by others that sOGT is an active enzyme on peptide (Dehennaut et al., 2008, Ma et al., 2013). This isoform specific substrate recognition was also tested *in vivo* with microinjection into Xenopus oocytes, where ncOGT was able to accelerate germinal vesicle breakdown while sOGT was not (Dehennaut et al. 2008), further suggesting its biological relevance.

2.4.2. Mechanisms of OGT substrate recognitions

Further analysis of O-GlcNAcylation on protein substrates by OGT with TPR truncations suggests two distinct OGT substrate recognition mechanisms. In the experiments, two groups of OGT substrate proteins were selected based on the reported or predicted O-GlcNAcylation sites in correlation to their positions in the protein structure. O-GlcNAcylation on proteins CREB, GSK3β, Yes1 and FOXO1 fragments was dependent on the length of the TPR domain (Fig. 2.6). O-GlcNAc signal on the proteins is greatly reduced or abolished when TPRs 7, or subsequent repeats are removed. One explanation is that TPRs 7 and 8 are involved in substrate interactions. Further molecular analysis has shown that the positions of the O-GlcNAc sites on proteins of this group, except for Yes1, are likely to be located in the structured regions of the proteins. This has also been observed in several other O-GlcNAcylated proteins where sites and structures are known, such as Nup62 (Lubas & Hanover, 2000), p53 (Yang et al, 2006) and histione 2B (Fujiki et al, 2011). In this case, the immediate surroundings of the modification site would have to undergo a conformational change to allow the binding in the OGT active site, as it has been shown in several crystal structure that the substrate peptides adopt extended conformation, and the substrate binding groove is narrow and is unlikely to admit a folded protein substrate (Lazarus et al, 2012; Lazarus et al, 2011; Schimpl et al, 2012). A possible explanation is that the interaction of substrate proteins with the TPR domain induces a conformational change in the substrates, which exposes the modification site. Without this induced-fit mechanism involving the interactions with the TPR domain, O-GlcNAcylation

will not occur. Alternatively, the loss of activity in TPR 7 and 8 truncations may also be due to the disruption of OGT dimerization. In the experiment with total cell lysate, a similar effect was observed upon deletion of TPRs 6 or 7. This apparent shift of the TPR boundary may be the result of different reaction conditions; purified protein substrates are present in much higher concentrations in the reaction compared to individual proteins in cell lysate samples, and the increase in substrate concentration my overcome reduced substrate affinity in the *in vitro* assay.

The second group of OGT substrate proteins, TAB1, FOXO1 and CK2α, contain O-GlcNAcylation sites in regions of the proteins that adopt flexible conformations, such as the C-terminal tail in CK2 and the interdomain region in TAB1. Rearrangement of the O-GlcNAc peptide therefore will not be required in preparation for binding to the OGT active site. OGT substrate interaction and protein O-GlcNAcylation are independent of the TPR domain in this model. This is indeed what we have observed: truncations in the TPR domain have no affect on the OGT O-GlcNAcylation of these protein substrates. Recognition and binding of such substrates will resemble those of a loose peptide, where the sequence immediately surrounding the O-GlcNAc modification site, i.e. the sequon, will predominate in OGT substrate recognition. The fact that no such sequon has been clearly observed for OGT substrate proteins, suggests that sequon recognition is not the complete story in OGT substrate interactions and protein O-GlcNAcylation, which is consistent with our proposal of two distinct substrate recognitions: one

contingent upon, the other independent of participation of the TPR domain of OGT.

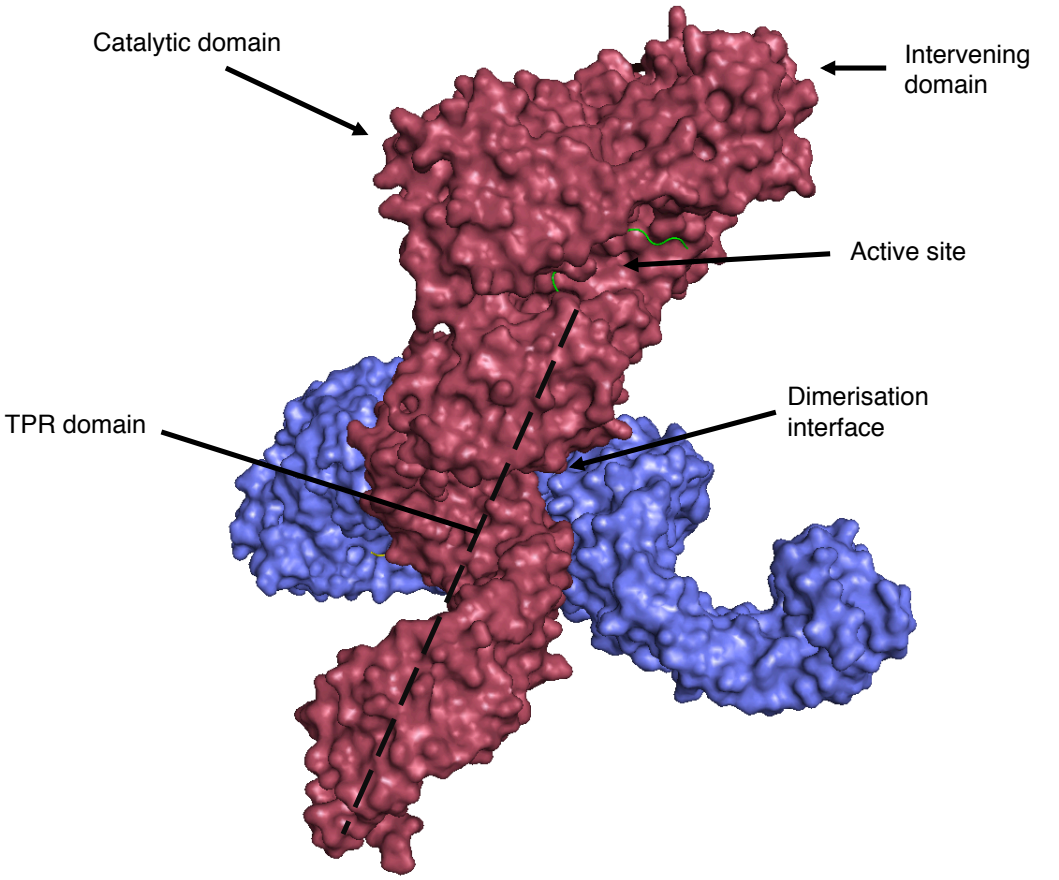


Figure 2.7: Surface representation of ncOGT homodimer model
Two ncOGT homodimerise (blue and red) at the interface in the TPRs 6 and 7, forming a 'clamp'. Green ribbon shows the CK2 peptide in the active site. The model is generated by superimposing the structure of TPR dimer (pdb: 1W3B) and hOGT catalytic domain (pdb: 3PE4).

2.5. Result (Part2):

2.5.1. YES1 is O-GlcNAcylated on Ser26 and Thr37

Previously, it has been reported that Src family non-receptor tyrosine kinase Yes1 is an OGT substrate and exhibits OGT isoform specific O-GlcNAcylation (Lazarus et al, 2006). We have identified the modification by mass spectrometry. Recombinant Yes1 was treated with ncOGT *in vitro* and in-gel digested with trypsin after SDS-PAGE. Peptides were analysed by liquid chromatography-mass spectrometry (LC-MS/MS). A single peptide of ^{16}YRPENTPEPVSTSVSHYGAEPTTVSPCPSSSAK^{49} was identified carrying multiple O-GlcNAc modifications, which is in agreement with previously reported mass spectrometry analysis in mouse synapse samples (Trinidad et al, 2012). The data suggest Ser26 and Thr37 as the likeliest modification sites. The peptide is located in the N-terminal region of Yes1. Both residues and surrounding sequences are not found in other Src family kinases (SFKs), in agreement with the fact that Yes1 is the only member of SFKs that undergoes O-GlcNAc modification. It is known that all SFKs adopt a defined SH3-SH2-kinase-tail core structure, in addition to a flexible N-terminal region comprised of an SH4 domain and a Unique domain. The O-GlcNAc site(s) on Yes1 appear to lie in one of the latter domains in the flexible part of the protein, similar to the modification sites on TAB1, CK2 and FOXO1. However, unlike TAB1 and CK2 (and, to a limited extent, FOXO1), Yes1 exhibits TPR dependent O-GlcNAcylation as was also observed for CREB, in which the modification sites are more likely to be located in structured regions of the protein.

2.5.2. Removal of YES1 kinase domain releases the N-terminal region and O-GlcNAcylation on YES1 becomes TPR-independent

To further investigate roles of OGT TPR motifs and their involvement in Yes1 recognition, I generated C-terminal truncations of Yes1, removing individual domains in order to confirm the site of modification and to investigate whether additional domains played any part in substrate recognition by OGT. The domain structure of full-length Yes1 being SH4-Unique-SH3-SH2-kinase-tail, the constructs SH4-Unique-SH3 (4U3, 1-151), in addition to the full length Yes1 (FL, 1-520), were expressed in *E. coli* with GST-tag, which was removed during the affinity purification on glutathione column to produce pure Yes1 truncation proteins (see Materials and Methods for details). The truncated Yes1 protein was *in vitro* O-GlcNAcylated with the hOGT TPR truncation proteins as described previously and O-GlcNAcylation of Yes1 proteins were determined by Western blotting. As shown previously, it appears that Yes1 O-GlcNAcylation is dependent on TPR 6 of OGT, suggesting that there is an additional enzyme-substrate interaction distant from the modification site. One would expect to see that C-terminal deletion of Yes1 would eventually abolish OGT activity on Yes1, presumably by losing the protein-protein interaction. However, all three Yes1 truncation proteins remain substrate of hOGT with TPR deletions up to TPR 6, as shown previously for the full length Yes1. More surprisingly, the hOGT truncations with more than 6 TPRs removed, that have been shown to be inactive on full length Yes1 protein, become active on Yes1 protein with the C-terminal truncation. That is, short hOGT truncation protein is active on Yes1 protein

that has been C-terminal deleted. It seems that the C-terminal domains prevent Yes1 O-GlcNAcylation by the shorter OGT TPR truncation (TPRs 7-12) but does not prevent O-GlcNAcylation by OGT with longer TPRs. The interaction of the full length OGT, possibly at TPR6, with the full length Yes1 disrupts this inhibitory effect by the kinase domain and allows O-GlcNAc modification at the Unique domain.

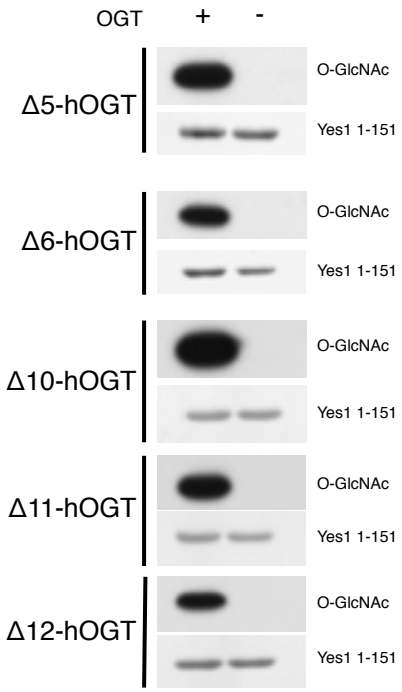


Figure 2.8: Yes1 become an OGT substrate irrespective of TPR deletions after the removal of kinase domain
Activity of selected hOGT with TPR truncations were tested on the C-terminal truncated Yes1 1-151 *in vitro.* *T*he reactions were stoppted by boiling in sample buffer before SDS-PAGE and detected by immunoblotting with anti O-GlcNAc RL2 antibody. Loading of the Yes1 1-151 protein was shown in Coomassie staining.

2.5.3. The TPR domain of OGT activates YES1 kinase

It is know that SFKs are under a tight autoinhibition by the SH3-SH2 clamp; releasing the clamp through displacement with high affinity ligand binding to SH3-SH2 results in activation of the kinase (Boggon & Eck, 2004; Pawson, 2004). SFKs are activated by upstream receptor kinases or high affinity SH3 and SH2 ligand proteins through the displacement of their intramolecular inhibition, allowing auto phosphorylation on the activation loop to achieve full activation. The data above suggest that OGT interacts with Yes1 through its

TPR motifs and may induce the release of the N-terminal region for O-GlcNAcylation. It was interesting to see whether the interaction of OGT with Yes1 plays a similar role as other SH3-SH2 ligands in Yes1 activation. It appears that TPR 6 is crucial in OGT-Yes1 interactions. To avoid the affect that may be produced by OGT catalytic activity, a hOGT construct with only the TPRs (16-400) was designed and purified as described in Jinek *et al* (Jinek et al, 2004). *In vitro* activity of the sf9 insect cell-expressed full length Yes1 was determined on the universal tyrosine kinase substrate peptide (poly4Glu:Tyr) using a radiometric assay. Optimal enzyme concentration and the reaction temperature were determined, and subsequent reactions were performed under conditions satisfying the requirements for steady-state kinetics (Appendix). In initial time course reactions, the presence of TPRs did not show any significant effect on Yes1 kinase activity (Fig. 2.9A). This is perhaps not surprising, given that Yes1 protein was expressed in sf9 cells as a constitutively active kinase. Physiologically, Yes1 is subject to negative regulation by phosphorylation on Y527 in the autoinhibitory tail, its absence leads to the accumulation of autophosphorylation on Tyr416 in the activation loop. To remove the autophosphorylation on Tyr416, Yes1 kinase was treated with GST-tagged lambda phosphatase. After the reaction, lambda phosphatase was removed by glutathione sepharose pull-down and othovanadate was added to inhibit any remaining phosphatase activity. The activity of Yes1 was reduced by 40% after 10 minutes of lambda phosphatase treatment. The treatment further resulted in a temporary lag in initial Yes1 activity, as seen in the time course (Fig. 2.9B). This lag might be caused by

the need to regain autophosphorylation. Most interestingly, however, the kinase activity of phosphatase-treated Yes1 was increased by 1.7 times in the presence of the TPR domain of hOGT.

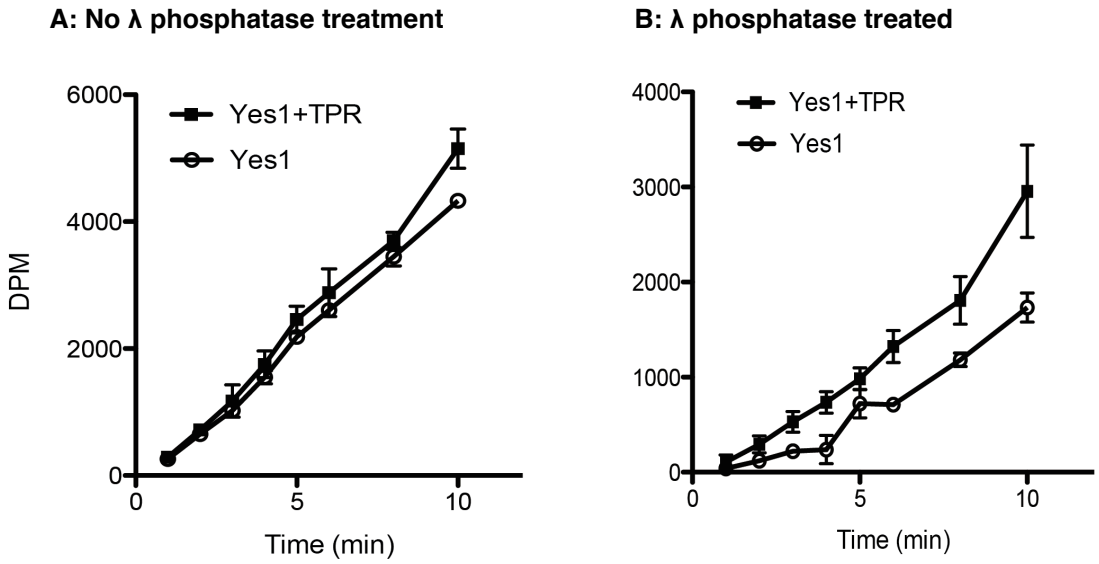


Figure 2.9: Initial activity and autophosphorylation was increased by OGT TPR domain

A: The kinase activity of Yes1 with or without TPR domain was measured using radiometric assay with ^{32}P ATP.

B: Yes1 was treated with GST-tagged λ phosphatase for 30 mins before removed by glutathione beads. Subsequently, the kinase activity of Yes1 was measured as for A in the presence of 0.1 mM sodium orthovanadate. Each data point was measured in triplicate and connected by straight lines. Error bars represent the standard error of the mean.

2.6. Discussion (Part2):

2.6.1. OGT activates Yes1 kinase in three ways

Yes1 belongs to the Src family tyrosine kinases (SFKs). All nine members of this family share high sequence identity in the core region comprised of a 'Src Homologous' 3 (SH3) domain, an SH2 domain followed by a helix linking them to the kinase domain and the C-terminal regulatory tail. SH3 and SH2 domains play central roles in intramolecular regulation of SFKs (Boggon & Eck, 2004; Pawson, 2004). High-resolution crystal structures of Src and Hck reveal a tight auto-inhibitory mechanism (Sicheri et al, 1997; Xu et al, 1999). The interactions of SH3 domain with the PXXP motif in the linker helix and the SH2 domain with the phospho-tyrosine on pY527 in the C-terminal tail region create a 'clamp' that locks the kinase domain in an inactivate conformation. This inactive conformation reduces the flexibility of the kinase domain and restrains the accessibility to the active site, blocking the autophosphorylation of Tyr416 in the activation loop (Sicheri et al, 1997; Xu et al, 1999; Xu et al, 1997). Recruitment of SFKs to ligand-activated receptors or interactions with other high affinity ligands of SH3 or SH2 displace the auto-inhibitory interactions and activate SFKs by releasing the SH3-SH2 clamp, which subsequently leads to autophosphorylation at the activation loop resulting in full activation of the kinase (Boggon & Eck, 2004; Pawson, 2004). This auto-regulatory mechanism is employed by all SFKs, as well as Tec family kinases and Abelson kinase (Abl) (Hantschel et al, 2003). The N-terminal region of SFKs, however, is less well conserved. The SH4 domain, the first 16 amino acids in the N-terminus, is usually modified by acylation, the covalent

attachment of a long fatty acid, which anchors SFKs to the plasma membrane, where they are activated by transmembrane receptor proteins to trigger downstream signalling cascades (reviewed in (Resh, 1999)). The Unique domain, which follows on the SH4, is even less well understood in most SFKs, with the exception of Lck. The N-terminus of Lck has been shown to associate with the cytoplasmic tails of the T-cell receptors CD4 and CD8 through a Zn-finger-like structural interaction (Kim et al, 2003). The Cys residues involved in Zn^{2+} binding in Lck are not conserved. The function of the Unique domain of Yes1 remains unclear.

O-GlcNAc modification has been found on Yes1 protein, but not on other members of SFKs (Lazarus et al, 2006). We have identified the peptide that carries the O-GlcNAc modification, with preliminary site-mapping results suggesting O-GlcNAc sites at Ser26 and Thr37, which lie within the Unique domain of Yes1. The N-terminal region of Yes1 is predicted to adopt a flexible conformation. Despite this, *in vitro* O-GlcNAcylation with TPR truncations of OGT clearly display a TPR-dependent manner of substrate recognition, which appears to be in contradiction with the model mentioned above. However, we were able to demonstrate that the TPRs are no longer required for Yes1 O-GlcNAcylation when the C-terminal domains are removed (Fig. 2.8), suggesting that the kinase domain of Yes1 prevents O-GlcNAcylation when OGT lacks its TPR domain. OGT with an intact TPR domain is able to overcome the inhibition by releasing the N-terminal region of Yes1, possibly through the interaction with the core region consisting of SH3-SH2-kinase-tail.

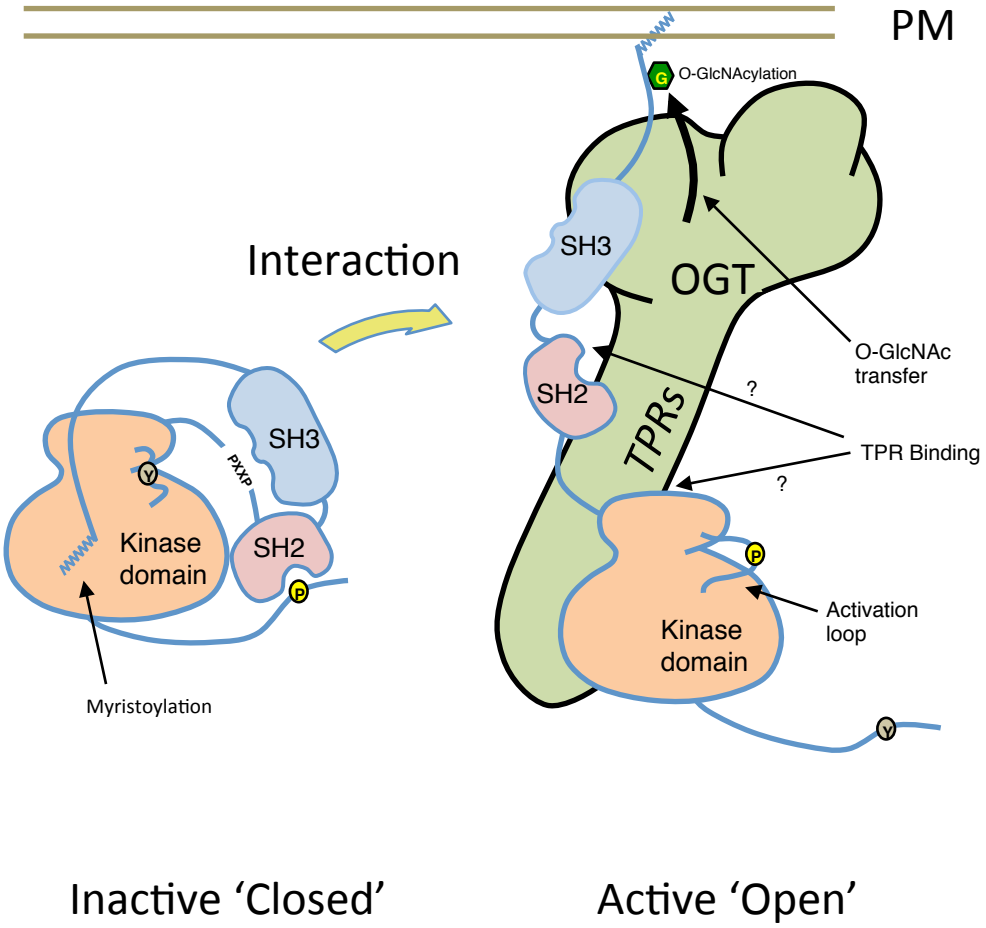


Figure 2.10: Proposed OGT regulation on Yes1 kinase
Schematic representation of proposed Yes1 kinase activation by OGT. Yes1 is in a 'closed' inactive conformation by the classic 'SH3-SH2 clamp' with SH2 domain interacts with phosphotyrosine at the C-terminal 'tail' and SH3 domain interacts with PXXP motif. The N-terminal region with the myristoylation bound to the core region of Yes1. Interaction of Yes1 with OGT release the 'SH3-SH2 clamp', and promote Yes1 to adopt an 'open' active conformation. The interaction also release the N-terminal region, which may promote Yes1 membrane localization. O-GlcNAc modifications have been observed in the N-terminal region of Yes1. It is not known how the TPRs interact with the Yes1 kinase. PM is plasma membrane.

We propose an unusually close association of the N-terminal region in Yes1 with the core region. The N-terminal region of Yes1 has unique functions; SH4-Unique domains are sufficient to inhibit Src-induced changes in cell morphology in experiments with Src/Yes1 chimera proteins (Summy et al, 2003). We have demonstrated that the TPR domain of OGT alone, without the glycosyltransferase activity, is sufficient to promote autophosphorylation of Yes1 in the absence of auto-inhibitory pY527. This activation of Yes1 kinase activity by the TPR motifs might be the direct effect of disrupting the SH3-SH2 clamp, or brought about indirectly through the release of the N-terminal region. Auto-inhibition by the N-terminal region has previously been observed in tyrosine kinase Abl, which shares great sequence homology with SFKs in the first 510 amino acids. Crystal structures of Abl showed the myristoylation buried in a hydrophobic pocket on the kinase domain (Nagar et al, 2003). Removing myristoylation in Abl leads to hyper activation of the kinase; this N-terminal cap (Ncap) has been suggested to be promoting the formation of the SH3-SH2 clamp, thus inactivating kinase activity of Abl (Hantschel et al, 2003).

I hypothesize that OGT may activate Yes1 through three distinct methods (Fig. 2.10): Firstly, interactions of OGT with Yes1 may directly activate Yes1 kinase activity as mentioned above by disrupting the inhibitory conformation caused by the SH2-SH3-clamp. Secondly, interaction of OGT with Yes1 may dislodge the N-terminal region and exposes the acylation to the plasma membrane, therefore promoting Yes1 membrane translocation, where Yes1

would be fully activated. Thirdly, O-GlcNAcylation at the N-terminus may stabilise Yes1 membrane localisation. All SFKs are myristoylated at Gly-2, where a 14-carbon myristoyl group is covalently attached to the alpha-N. This myristoylation promotes SFKs plasma membrane localisation in conjunction with a palmitoylation, a 16-cabon fatty acid is covalently attached to Cys side chain, which also occurs in the N-terminal region. In addition to these acylations, the N-terminal region is rich in positively charged residues, which interact with the plasma membrane through electrostatic interactions. Phosphorylations have been also reported in this region and been proposed to affect the electrostatic effects. O-GlcNAcylation may compete with the phosphorylation sites and maintain the electrostatic interaction.

2.6.2. Two OGT substrate-binding models: loose peptide and TPR-induced fit.

When OGT acts on peptide substrates, the peptide, in the absence of any secondary structure, will be in a flexible conformation, which enables it to adopt an extended conformation when binding to the active site of OGT. Protein substrates where the O-GlcNAc site lies in an extended flexible region, such as the C-terminus of CK2α or the interdomain region of TAB1, may be expected to adopt a similarly extended binding mode. The opportunities for initial contact and binding could be considered very similar for both types of substrates. In such cases, any interaction of substrate protein with the TPR domain would be accessory rather than necessary.

In contrast to this subset of 'easily accessible' substrates, O-GlcNAcylation has been observed in ordered regions of protein substrates, within domains and on secondary structure elements on the molecular surface. To allow OGT to act on these substrates, we suggest that the binding to the TPR domain of OGT may introduce conformational changes, which temporarily expose/give access to the region bearing the modification site. One example is provided here by the study of Yes1 kinase, which is O-GlcNAc modified in its N-terminal domain. It appears that binding of OGT induces the release of the modification site from the core region of Yes1, thereby allowing its O-GlcNAcylation.

The TPR domain of OGT also mediates the homodimerisation of the enzyme, with the dimer interface located primarily on repeats 6 and 7. Our experiments investigating the contributions of individual TPR motifs have highlighted a particular requirement for TPRs 6 and 7 in TPR dependent protein O-GlcNAcylation, which may suggest that dimerization of OGT is required for the TPR-induced fit model of substrate recognition. A composite model of the ncOGT dimer, based on crystal structures of the TPR dimer and of the catalytic domain, reveals the proximity of the catalytic domain to the early TPRs from the partner molecule (Fig. 2.7). This raises the possibility of intermolecular cooperative effects, which have not hitherto been studied for OGT. It would be interesting to obtain the crystal structure of the OGT dimer in complex with a substrate protein, to observe this substrate recognition mechanism in action.

2.7. Material and Methods

Constructs information

Primers used to generate the TPR truncations in hOGT are listed in the Appendix.

Expression of the TPR truncations and Δ9 construct purification

hOGT was expressed in ArcticExpress *E. coli* (Stratagene) by inoculating 1 L of LB medium with 5 ml of overnight culture grown in the presence of 50 µg/l ampicillin and 20 µg/l gentamycin and allowing it to grow at 37 °C until OD600 reached 1.0. The culture was cold shocked at 4 °C for 1 h followed by induction with 125 µM IPTG and incubation at 12 degrees for 48 h. The culture was harvested by centrifugation at 3,000 g for 30 min and the pellet washed with TBS buffer (25 mM Tris-Cl pH 7.5, 150 mM NaCl). The pellet was then resuspended in lysis buffer (25 mM Tris-Cl pH 8.5, 250 mM NaCl, 1 mM DTT, 1 mM benzamidine, 0.2 mM PMSF and 5 µM leupeptin) and lysed using a continuous flow cell disruptor (Constant Systems) at 30 kpsi. The lysate was centrifuged at 40,000 g, the supernatant filtered through 0.2 µm cellulose acetate filters and incubated with glutathione sepharose 4B beads (GE Healthcare) for 2 h at 4 °C on a rotating platform. The beads were then washed with washing buffer (25 mM Tris-Cl pH 8.5, 250 mM NaCl, 1 mM DTT) and incubated with PreScission Protease (200 µg protease per ml of beads) at 4 °C for 16 h. The supernatant containing the cleaved protein was diluted to 25 mM NaCl in Tris-Cl, pH 8.5 before IEX on Q sepharose eluting with a linear gradient of 0.025—0.5 M NaCl. Fractions selected were pooled and buffer exchanged into 25 mM Tris-Cl pH 7.5,

150 mM NaCl and 1 mM THP. The proteins were snap-frozen at 1 g/l concentration in the same buffer containing 10 % glycerol for future use.

Protein stability

20 μg of hOGT proteins of different constructs (Δ1-13 hOGT) that bound to the glutathione beads in 50% slurry in the buffer containing 25 mM Tris-Cl pH 7.5, 150 mM NaCl only was incubated at 22 °C for 24 hours. Proteins were boiled in the 1x sample buffer (Thermo Scientific, containing 10% glycerol, 1% Lithium dodecyl sulphate (LDS), 1% Ficoll-400, 0.2 M triethanolamine-Cl, pH=7.6, 0.005% phenol red, 0.005% Coomassie G250 and 0.5 mM EDTA sodium). Samples was subject to SDS-PAGE and the protein was analysed by Coomassie stain (0.1% Coomassie R250, 10% Acetic acid and 20% methanol).

Cp*OGA stripping*

HEK293 cells were cultured on 10 cm diameter dishes in 10 ml DMEM supplemented with 10% (v/v) fetal bovine serum (FBS), 2 mM L-glutamine, 100 U/ml penicillin, and 0.1 mg/ml streptomycin. Cells were lysed 36 hours post-transfection in 0.1 ml of ice-cold lysis buffer (50 mM Tris-HCl pH=7.5, 1 mM EGTA, 1 mM EDTA, 1% (w/v) Triton-X (Sigma), 1 mM sodium othovanadate, 50 mM sodium fluoride, 5 mM sodium pyrophosphate, 0.27 M sucrose, 1 mM DTT, 1 mM benzamidine and 0.1 mM PMSF). The crude lysates were collected by centrifugation at 20000 g for 15 min at 4 °C, and the supernatants were collected and divided into aliquots, frozen in liquid nitrogen and stored at -20 °C for future use. 200 μg of the total lysate in 100 μl of lysis buffer was treated with 0.5 μg of

GST-*Cp*OGA for 1 hour at 37 °C (*Cp*OGA was previously expressed in *E. coli* and purified on glutathione 4B beads and elute with glutathione. Buffer exchange with concentrator of molecular cut-off of 50 kDa to remove the glutathione and protein is stored in 50 mM Tris-Cl, pH=7.5, 150 mM NaCl and 40% glycerol, frozen at -80 °C for future use). *Cp*OGA was removed by 20 μl of glutathione 4B beads in 50 % slurry for 1 hour at 4 °C; centrifugation collected supernatant, the action was performed twice to ensure the removal of *Cp*OGA and to the final supernatant added final concentraton of 1 μM of GlcNAcstatin C. The O-GlcNAc depleted 293 cell lysate was frozen at -80 °C for future use.

In vitro *OGT reaction on lysate*

30 μg of the striped 293 cell lysate was subjected to *in vitro* O-GlcNAcylation with 0.5 μM of purified TPR truncation hOGT in 30 μl of reaction in buffer containing 50 mM Tris-Cl, pH=7.5, 150 mM NaCl, 1 mM DTT, 1 mM benzamidine, 0.2 mM PMSF and 5 μM leupeptin and 1 μM GlcNAcstatin C. The reaction was stopped after 30 min at 37 °C by boiling in 1x LDS sample buffer. 10 μl of the sample (~10 μg of lysate) was subjected to SDS-PAGE and transferred onto PVDF membrane, which was blocked by 5% BSA at room temperature for 1 h. O-GlcNAcylation on the lysate was detected by anti-O-GlcNAc antibody RL2 overnight at 4 °C. The signal was detected by ECL after amplification with secondary HRP conjuganted anti-mouse antibody. OGT was detected by anti-OGT antibody (DM17) overnight in the same conditions as for O-GlcNAc detection.

Recombinant protein expression and purification

GSK3β (2-420) was expressed in insect cell sf9 using baculovirus system. The N-terminally His (6) and EE (EFMPME) tagged protein was purified with Nickel NTA-agarose and cleaved by TEV protease (ENLYFQG site). Protein was stored in 50 mM Tris-Cl, pH=7.5, 150 mM NaCl and 20% glycerol at -80°C for future use. Yes1 (1-543) was expressed in sf9 cells and purified as for GSK3β. CK2α (1-139), TAB1 (1-401), FOXO1 (2-655) and CREB (1-283) were expressed in *E. coli*. CK2α and TAB1 were purified by glutathione 4B Sepharose; FOXO1 was purified with GSH Sepharose and nickel NTA agarose; CREB was purified with GSH Sepharose. All proteins were stored at -80 °C for future use. Yes1 truncations (1-91, 1-151, 1-260) were expressed in *E. coli.* Cultures with the cell density of 1.0 were induced by 100 μM IPTG. Allowed to grow at 18 °C for 24 hours and lysed with continuous flow cell disruptor. Purification was the same as for hOGT. (Expression and purification of GSK3β, Yes1, FOXO1, CREB was performed by DSTT, CK2α and TAB1 was by Dr. Marianne Schimpl)

In vitro *O-GlcNAcylation on recombinant proteins*

1-2 μg of the recombinant proteins were used in *in vitro* reactions with 1 μM of hOGT TPR truncations proteins in 20 μl reaction containing 1 mM UDP-GlcNAc, 1 mM DTT, 50 mM Tris-Cl, pH=7.5, 150 mM NaCl for 60-90 minutes at 37 °C. The reactions were stopped by boiling in 1x sample buffer. Samples were subject to SDS-PAGE and transferred to PVDF membrane and probed by anti-O-GlcNAc LR2 antibody overnight at 4 °C. The signal was detected by ECL after amplification with secondary HRP conjuganted anti-mouse antibody. OGT was

detected by anti-OGT antibody (DM17) overnight in the same conditions as for O-GlcNAc detection.

Kinase reaction of Yes1

Activity of Yes1 was determined on universal tyrosine kinase peptide substrate poly(4Glu:Tyr) in radiometric assay. The assays were performed by using 0.2 to 5 μg of recombinant Yes1 proteins expressed and purified from sf9 cells as described above. The reaction was in a final volume of 30 μl containing 20 mM Tris-Cl, pH=7.5, 10 mM $MgCl_2$, 1 mM DTT, 0.1 mM [γ-^{32}P] ATP at 0.3 μCi and 1 mg/ml poly(4Glu:Tyr) peptide. The reaction was started by adding the [γ-^{32}P] ATP into the reaction and incubated at 4 or 16 °C for 10 minutes. At each time point, reaction was taken and added into 300 μl of 75 mM ice-cold phosphoric acid to stop the reaction. The samples were then applied onto 96-wells P81 assay plate and the liquid was decanted by vacuum. The P81 was washed 10x with 300 μl of 75 mM ice-cold phosphoric acid and air dried. Radioactivity was measured with plate reader (PerkinElmer, MicroBeta®).

75 μg of Yes1 was treated with 25 μg of GST-tagged lambda phosphatase in the buffer containing 100 mM Tris-Cl, pH=7.5, 150 mM NaCl, 10 mM $MgCl_2$ in the final volume of 300 μl at 30 °C for 30 minute and the phosphatase was removed by adding 50 μl of glutathione 4B Sepherose and incubated for 30 min at 4°C on roller. Repeat the last step and collect the supernatant and added final concentration of 1 mM of Na orthovanadate. Where indicated, Yes1 was

pretreated with a 5-fold molar access of TPR on ice for 1 hour before the kinase reaction.

Mass-spectrometry of O-GlcNAcylation on Yes1

For the site mapping analysis of enzymatically glycosylated Yes1, *in vitro* glycosylation reactions contained 10 mg of recombinantly (sf9) expressed Yes1, 5 mM UDP-GlcNAc and 0.5 μg OGT in 50 μl of 50 mM Tris-Cl, pH=7.5, 150 mM NaCl, 1 mM DTT for 90 min at 37 °C. Protein sample was buffer exchanged into 50 mM Na bicarbonate and alkylated by final concentration of 10 mM iodoaceamide for 30 minutes at dark before trypsinisation with 0.2 ug of trypsin at 30 °C overnight in the filter tube. Digested peptides were collected by centrifugation at 3,000 rpm with tabletop centrifuge. The sample was evaporated to dryness and reconstituted in 0.1 % formic acid, before separation on a PepMap C18 column (75 μ m i.d. x 15 cm x 2 μ m) LC-MS/MS(Bruker Daltonics). The mass spectrometry method used was based on alternating CID measurements. Raw data was processed and used to search against Mascot.

3. Chapter II

Catalytic Mechanism Of Human O-GlcNAc Transferase (OGT)

3.1. Project Overview:

O-GlcNAc transferase (OGT) is a key enzyme in the O-GlcNAcylation process. Many groups have reported the biochemical and biological characterisation of OGTs from many organisms (Hart et al, 2007). At the start of this project, two groups had independently reported the crystal structures of a bacterial OGT ortholog from the plant pathogen *Xanthomonas campestris* (*Xc*OGT). They were the first structures reported from the GT41 family and confirmed that OGT adopts the general GT-B fold as predicted from sequence-based classification (Clarke et al, 2008; Martinez-Fleites et al, 2008). The information provided by the structures allows speculating on the enzymatic actions of OGT utilising UDP-GlcNAc to post-translationally modify intracellular proteins. Therefore, *Xc*OGT became an excellent working model towards the understanding of human OGT (hOGT) at the molecular level. However, several considerations must be taken into account. Firstly, although *Xc*OGT shares high sequence conservation at the catalytic domain with the metazoan OGTs (36% identity to human OGT, Fig 1.5), it lacks the intervening domain within the catalytic domain that is present in all metazoan OGTs. Secondly, no protein substrates have been reported for *Xc*OGT so far, questioning whether it is truly a bacterial form of O-GlcNAc transferase. And finally, the donor substrate UDP-GlcNAc, known as the sugar donor for the metazoan OGTs, shows poor affinity for *Xc*OGT at 15 mM (Dorfmueller et al, 2011). This affinity is similar to that of UDP-Glucuronic acid, which shares the structural similarities with UDP-GlcNAc. *Xc*OGT has been demonstrated to hydrolyse UDP-GlcNAc over a two-day period at relatively high concentrations, this however does not fully support that

*Xc*OGT is a peptido-O-GlcNAc transferase (Martinez-Fleites et al, 2008). The OGT enzyme catalytic mechanism and action of O-GlcNAc transfer therefore remained unknown and structural information from active metazoan OGT was highly desirable. Only recently, the crystal structure of human OGT in a complex with UDP and acceptor peptide has been reported (Lazarus et al, 2011). The complex lacked the position of the GlcNAc moiety; therefore the information did not conclusively explain the catalytic mechanism of hOGT.

Limitations in small selective pharmacological inhibitors of OGT represent a challenge in studying the biological functions of O-GlcNAc modifications. The uracil mimic alloxan inhibits OGT by a) occupying the nucleotide binding pocket and b) binding to the active site possibly through covalent modification at a catalytic site cysteine residue, C917 (Konrad et al, 2002). Because of its chemical nature, it is unlikely that inhibition by alloxan is specific to OGT (Zachara & Hart, 2004). Strategies with high throughput screens (HTS) of fluorescent-based substrate displacement assay have yielded compounds that inhibit human OGT in the low μM range *in vitro* (Gross et al, 2005). Another irreversible inhibitor was reported recently (Jiang et al, 2012). The crystal structure of the inhibitor-hOGT complex shows that this bicarbamate compound binds to the active site of human OGT by covalently targeting conserved active site lysine (K842) and cysteine (C917) residues. An alternative strategy is to rationally design competitive inhibitors by mimicking the natural substrate. The inhibitors bind to the active site but resist glycosyltransfer. Several such UDP-GlcNAc analogues have been reported, however they are relatively poor

inhibitors of hOGT (Dorfmueller et al, 2011). The chemical nature of this class of inhibitors may also result in low cell penetration. A thio-derivative GlcNAc (tetra-acetyl-5SGlcNAc)(Gloster et al, 2011) that replaces the ring oxygen in GlcNAc with sulfur is cell permeable as a per-acetylated prodrug and could hijack the hexosamine biosynthetic pathway, resulting in intracellular synthesis of the hOGT inhibitor UDP-5SGlcNAc (Fig 3.1). However, UDP-5SGlcNAc is also a substrate of hOGT as shown by mass spectrometry (Appendix). This focus on the substrate based inhibitor design has been unsuccessful partially because of insufficient knowledge of the glycosyltransfer mechanism.

This chapter describes the characterization and the enzymatic properties of hOGT, through the structures of ternary product and substrate complexes of hOGT with sugar donor and synthetic peptide analogs. The structures reveal the GlcNAc moiety in the active site, which was absent from previous reports (Lazarus et al, 2011). The unambiguous electron density for a peptide provide information in the peptide binding in addition to the previously reported peptide from casein kinase 2 α (CK2α). An unusual conformation of the UDP-GlcNAc analog was observed, which brings the non-bridging oxygen from the *α*-phosphate in a close proximity to the acceptor serine hydroxyl group. Together with stereoisomers of the *α*-phosphorothioate analog of UDP-GlcNAc and site-

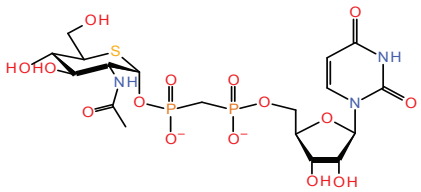


Figure 3.1: Chemical structure of UDP-5SGlcNAc
(uridine diphospho-5-thio-*N*-acetyl-glucosamine). Noted the sulfur atom replacing the ring oxygen of GlcNAc shown in orange.

directed mutagenesis, the data delineate a new catalytic mechanism for inverting glycosyl transfer.

3.2. The aims of the work presented here were:

- To provide structural information of the hOGT catalytic domain
- To elucidate the catalytic mechanism of hOGT glycosyltransfer and identify the catalytic base
- To identify a point mutation that abrogates hOGT activity without losing substrate binding

3.3. Result and Discussion

3.3.1. Crystallisation of hOGT

As described above in Chapter 1, a hOGT construct containing the catalytic domain and 4.5 TPRs (or delta TPR9, Δ9-hOGT) produced the best expression in *E. coli* and produced stable protein. Therefore, this protein was chosen for crystallisation experiments in order to determine the structure of the enzyme by X-ray crystallography. After attempts with several commercial crystallisation screens, the Δ9-hOGT protein in the presence of a 10-fold molar excess of UDP was found to produce microcrystals in multiple conditions, while in about half of the conditions produced mild precipitation. Further optimisation could not improve the crystal formation and no X-ray diffraction was observed indicating that further construct design, such as minimal trimming of the construct might be required. At the same time, Suzanne Walker's laboratory reported a structure of hOGT from a construct spanning amino acids 312-1031, similar to Δ9-hOGT but without the C-terminal 15 amino acids, which were predicted to be flexible and unstructured (Lazarus et al, 2011). To further study the catalytic mechanism of hOGT by crystallography, the 'Walker' construct has been used replacing the existing Δ9-hOGT construct for crystallisation using the expression and purification system described above. Crystallisation of the 'Walker' hOGT in the presence of UDP or UDP-5SGlcNAc could be reproduced. Point mutations in the catalytic site were re-cloned in the 'Walker' construct to reduce confusion, and unless otherwise stated, all the proteins of wild type (wt) and mutants for biochemical and biophysical studies are based on the 'Walker' construct.

The crystallographic study of hOGT reported a trimetric complex containing the reaction product UDP as well as a substrate peptide (pdb: 3PE4) (Lazarus et al, 2011). The position of the sugar moiety was consequently still unknown. Lazarus *et al.* report that they were unable to obtain complex structures with UPD-GlcNAc due to hydrolysis of the substrate. Substrate hydrolysis, i.e. the transfer onto a water molecule, may be catalyzed by glycosyltransferases in the absence of the native acceptor substrate. Attempts at preventing hydrolysis by determining the structure of hOGT with the inhibitor UDP-5SGlcNAc also resulted in lack of electron density for the GlcNAc moiety. Such absence of signal may show that the substrate has undergone hydrolysis, but it may likewise indicate disorder of that part of the molecule. In the *Xc*OGT structure, complexes with UDP-S-GlcNAc and UDP-C-GlcNAc showed two different GlcNAc positions (2XGO and 2JLB). It suggested that the lack of GlcNAc density might not be the result of hydrolysis. Instead, the acceptor peptide may be required to be in place to position the GlcNAc moiety.

(All the crystallisation of the 'Walker' hOGT and the data processing were carried out by Dr Marianne Schimpl)**.**

3.3.2. The product complex of human OGT with UDP and an O-GlcNAc peptide - Active site residues participate in positioning, but not transfer, of GlcNAc

3.3.2.1. Structure of a ternary hOGT product complex defines a conserved peptide-binding mode

Previous attempts at determining the structure of complexes between hOGT and its donor substrate UDP-GlcNAc have resulted in hydrolysis of the substrate (Lazarus et al, 2011). In order to ascertain the precise position of GlcNAc in the active site and gain insight into OGT substrate recognition a synthetic O-GlcNAcylated peptide (gTAB1tide) (Fig. 3.2) derived from the regulatory O-GlcNAc site on the innate immunity signaling protein TAB1 was designed and synthesized by Dr Vladimir Borodkin to trap the enzyme/glycopeptide product complex (Pathak et al, 2012). hOGT in complex with UDP was crystalized, which was then soaked with gTAB1tide before data collection. Electron density maps at 3.15 Å, improved by 4-fold non-crystallographic averaging, revealed unambiguous density for UDP and gTAB1tide. UDP adopts the same conformation as observed in the recently reported hOGT-UDP-peptide complex (max. atom shift = 1.0 Å), tethered by interactions with nine residues that are all conserved in metazoan OGTs (Fig 1.4). The peptide shows ordered density for

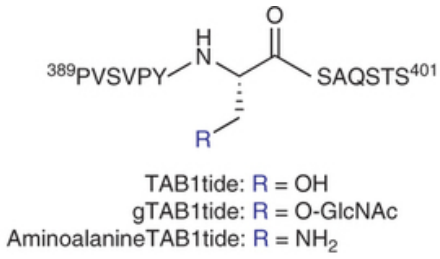


Figure 3.2: Chemical structures of TAB1 peptide and its O-GlcNAc and amine derivatives at the reported O-GlcNAc modified serine.

the –6 to +4 subsites (Fig. 3.3C), and adopts a backbone conformation in the –3 to +3 subsites similar to the previously reported complex of hOGT with a CKII-derived substrate peptide (VPYSSAQ) for gTAB1tide, TPVSSAN for CK2α (Fig 3.3C). This suggests that hOGT may impose structural and/or sequence constraints on the acceptor peptide and that hOGT specificity at the peptide sequence level may be worth exploring.

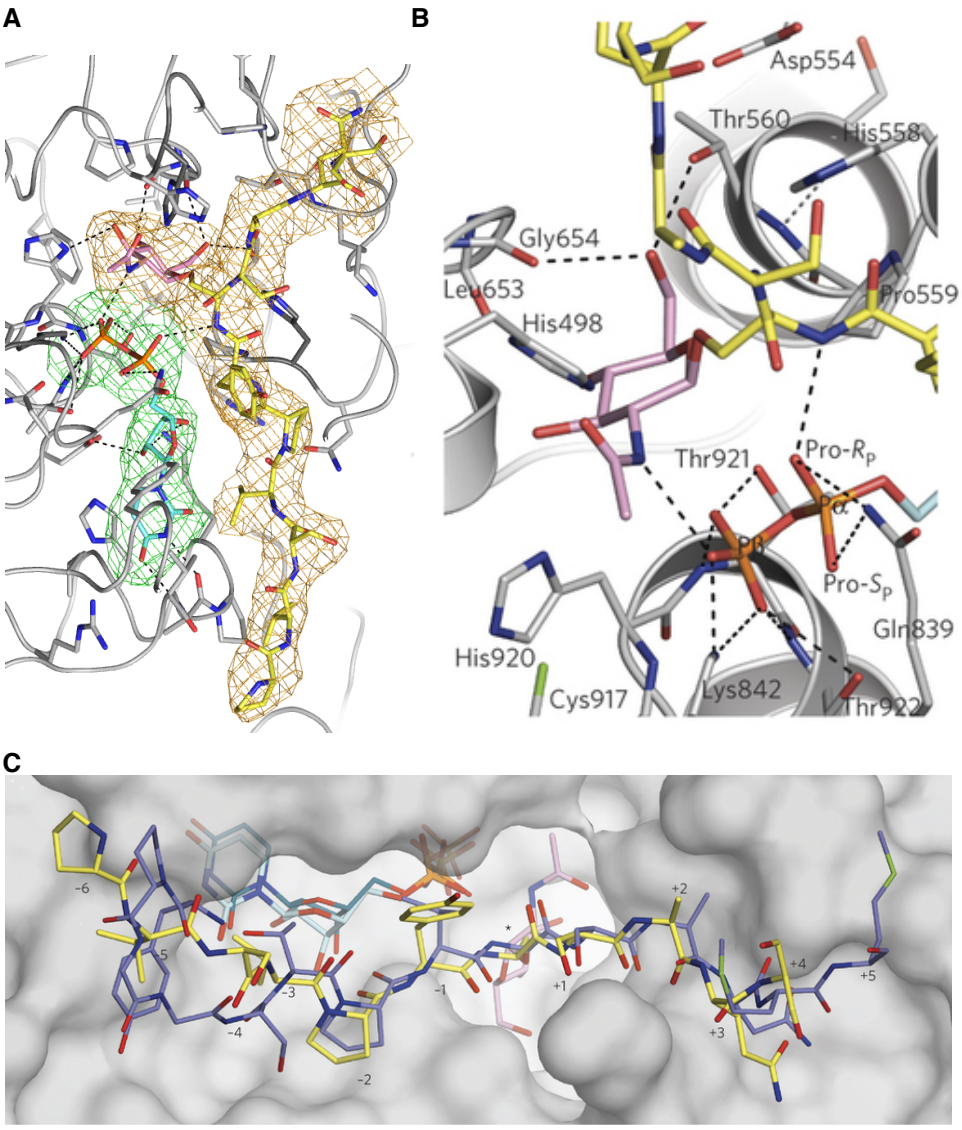


Figure 3.3: Zoom in view of the hOGT catalytic site
A. hOGT product complex with full length glycopeptide (gTAB1tide) ligand(yellow, GlcNAc is in pink) and UDP (cyan) showing unbiased *Fo-Fc* electron density after four-fold averaging (contoured at 2.5 σ).
B. The enzyme is shown in cartoon representation with side chains shown as sticks with grey carbon atoms. Colours for the glycopeptide and UDP are as in A. The α and β phosphates of UDP are labelled with Pα and Pβ, respectively. Hydrogen bonds are depicted by dashed lines.
C. Surface representation of the hOGT active site with reaction products gTAB1tide and UDP shown as sticks (the peptide, sugar and UDP are shown with yellow, pink and turquoise carbon atoms, respectively). The previously reported complex with the unmodified CKII peptide and UDP (PDB code 3PE4) is shown as sticks with blue and dark green carbon atoms, respectively.

3.3.2.2. Active site residues participate in positioning, but not transfer, of GlcNAc

In the gTAB1tide-hOGT complex the β-linked sugar, adopting the $^4C_1$ chair conformation, projects into a conserved pocket, where it is tethered by T560, H920, L653 and G654 (Fig 3.3B). The methyl group of the N-acetyl moiety points between the GT-B catalytic core and the TPR repeats, into a pocket formed by C917, M501 and L502, explaining why hOGT can tolerate a diverse array of UDP-GlcNAc analogs bearing bulky amido substituents, including azido derivatives that are widely used with click-chemistry to identify and enrich for O-GlcNAc proteins (Vocadlo et al, 2003; Wang et al, 2010). The identity of the OGT catalytic base that is thought to activate the Ser/Thr in acceptor proteins has been the subject of a number of studies that propose either of two histidines, H498 or H558, as candidates (Clarke et al, 2008; Lazarus et al, 2011; Martinez-Fleites et al, 2008). Inspection of the OGT glycopeptide product complex described here reveals that both of these histidines are positioned >4.5

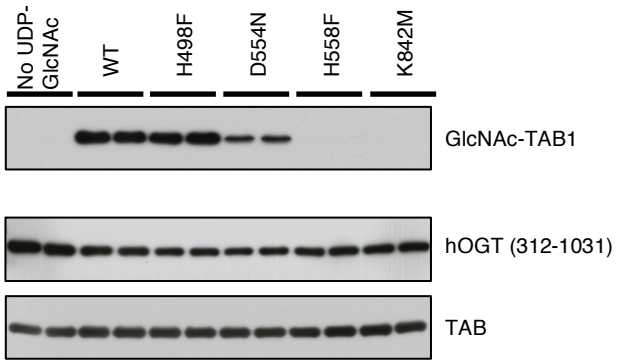


Figure 3.4: Activity of hOGT point mutants in an *in vitro* O-GlcNAcylation assay of TAB1 protein. O-GlcNAc was detected by immunoblotting with a pan-O-GlcNAc antibody (RL2). WT, wild type. Experiments were performed in duplicate, as shown.

Å from the acceptor hydroxyl, and lack interacting residues that would support either of them acting as a catalytic base (Fig. 3.3B). His498 has been most recently proposed as the catalytic base; mutations to alanine or aspartic acid inactivate hOGT. When the His498 was mutated to phenylalanine, the enzyme retains activity (Fig. 3.4). Interestingly, a phenylalanine is found in the equivalent position of the *Xanthomonas campestris* putative OGT, indicating the histidine likely to have structural function other than serving as the catalytic base. H498E mutation also dramatically decreases the binding of UDP-GlcNAc but not UDP, suggesting its involvement in orientating the sugar moiety in the donor (Appendix 3). Most interestingly, the candidate for the catalytic base, H558 as discussed in the previous section (see section 1.2.2), is sandwiched between P559 and D554 with the carboxylate of D554 stacking with the H558 imidazole side chain. The imidazole side chain of H558 accepts a hydrogen bond from a protein backbone amide on (deprotonated) Nδ (Fig. 3.3B), implying that the imidazole Nε, facing the acceptor serine, will be protonated at neutral pH, which is not compatible with a role as a general base. Mutagenesis of H558, however, was not able to demonstrate the regain of enzymatic activity while it was replaced with any other amino acids (Fig. 3.5A). Further attempts in rescuing the histidine to alanine and glycine mutation by adding imidazole aiming to mimic the histidine was also not successful (Fig. 3.5B). His558 to Ala or Phe however did not affect the binding of UDP, UDP-GlcNAc and its analogues (Appendix 3). On the other hand, as discussed previously, sequence and structural comparison with other glycosyltransferases has suggested that D554/H558 may form a catalytic dyad; D554N mutation in hOGT did not abolish

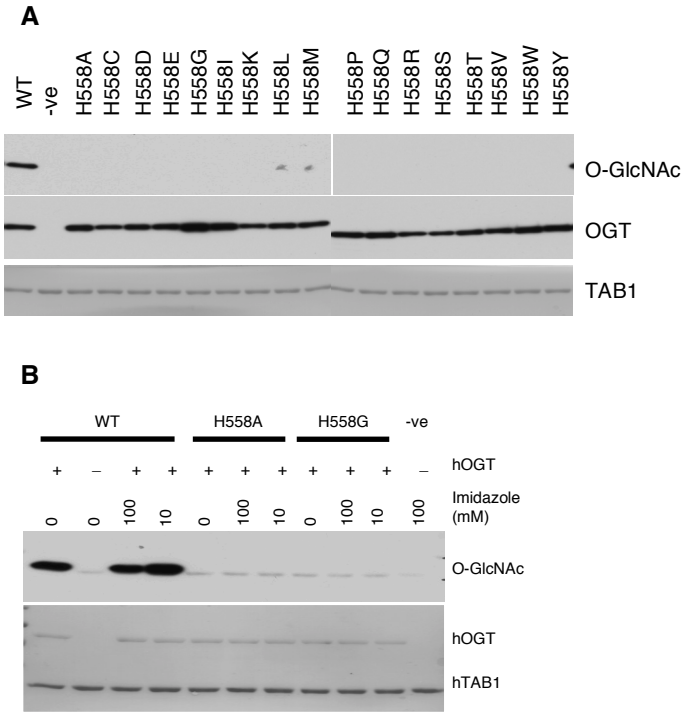


Figure 3.5: Attempts at defining the role of His558 in the catalysis.
A. hOGT His558 has been mutated in an attempt to identify mutations at the site that retain the catalytic activity. Proteins were purified and the in vitro activity was determined by in vitro assay and immunoblotting.
B. Attempts to rescue the activity of hOGT H558A/G mutations by addition of imidazole to compensate for the missing histidine side chain.

its activity. The dyad arrangement was not observed in the structure as Asp554 residue stacks with His558 (Fig 3.3B). These observations revoke the direct involvement of D554 in the catalysis and suggest the effects of the H558 mutation are presumably due to structural reasons. Furthermore, in the latest report by Lazarus et al (Lazarus et al, 2012), authors were also able to trap the GlcNAc moiety in complex with hOGT and proposed another hOGT catalytic mechanism based solely on the structural information: the base catalysis of proton extraction was performed by D554 through the Grotthus chain, where proton is passed through a hydrogen network of water. Our finding with the D554N mutation has therefore disproved this hypothesis. Tyr841 is another residue that has been proposed to be the catalytic base (Martinez-Fleites et al,

2008); hOGT with the Y841F mutation retains catalytic activity. It is likely that Tyr841 forms part of a hydrophobic patch providing structural support for neighboring residues and positioning of H498 and K842, which is important for catalysis as discussed below.

3.3.3. A pseudo-Michaelis complex suggests substrate-assisted catalysis. (UDP-5SGlcNAc-Pep-OGT complex)

3.3.3.1. Trapping the Michaelis complex of hOGT

Surprisingly, it appears that none of the enzyme side chains closest to the acceptor serine can act as a catalytic base. To uncover the identity of this catalytic base, it is more desirable to trap a complex with intact substrates. Two artificial substrate analogues were used to reduce the rate of enzymatic turnover *in crystallo*, namely UDP-5S-GlcNAc, a recently reported hOGT donor substrate analog inhibitor (Gloster et al, 2011), and the aminoalanine derivative of the TAB1 acceptor peptide (aaTAB1tide), where the serine hydroxyl is replaced with a primary amine (Fig. 3.2). Notably, although turnover is reduced, OGT can utilize both UDP-5S-GlcNAc (reported reduction of K_{cat} by factor 14 compared to UDP-GlcNAc, (Gloster et al, 2011)) and aaTAB1tide as

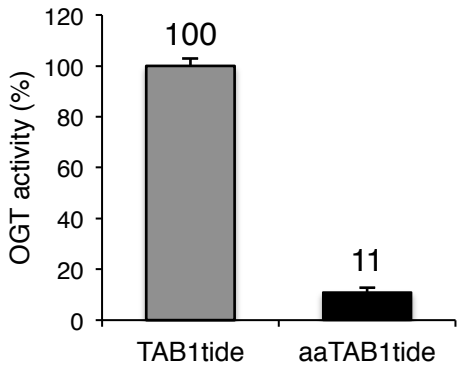


Fig 3.6: Glycosyl transfer to N-terminally biotinylated TAB1tide and aaTAB1tide was determined in a scintillation proximity assay by D.B.

substrates for glycosyl transfer onto the aminoalanine (Fig. 3.6). The observed activity with these pseudo-substrates suggests that they possess catalytically competent binding modes, similar to the natural substrates and would be suitable for the trapping.

3.3.3.2. Structure analysis of hOGT-UDP-5SGlcNAc-peptide pseudo-Michaelis complex

hOGT was crystallized in complex with UDP-5S-GlcNAc and soaked with aaTAB1tide. Synchrotron data were collected to 3.3 Å and electron density maps (improved by 4-fold non-crystallographic averaging) re vealed unambiguous density of a pseudo-Michaelis complex (Fig. 3.8A). The

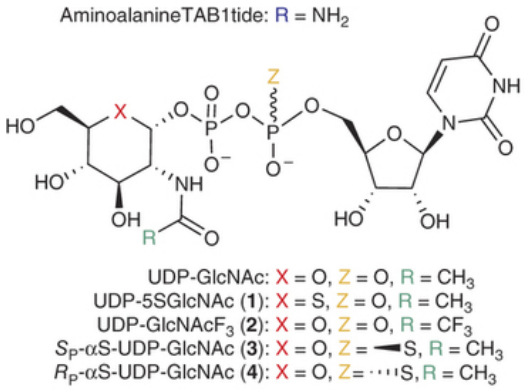


Figure 3.7: Chemical structure scaffold of UDP-GlcNAc and its derivatives.

overall conformation of the enzyme is almost unchanged (RMSD on 698 Cα atoms = 0.3 Å, max. atomic shift of any active site residue after overall superposition = 0.3 Å). The UDP moiety of UDP-5S-GlcNAc adopts a conformation similar to that of UDP in the product complex (RMSD = 0.2 Å, max. atomic shift = 1.0 Å), whereas the sugar is tilted away from the acceptor compared to the product complex (angle of rotation = ~30°), and is tethered by H920/L653, G654/T560 on the O3/O4/O6 hydroxyls respectively (Fig. 3.8B). His920 to Ser mutation disrupt the binding of UDP-GlcNAc. Strikingly, the observed conformation of the donor substrate is remarkably different from

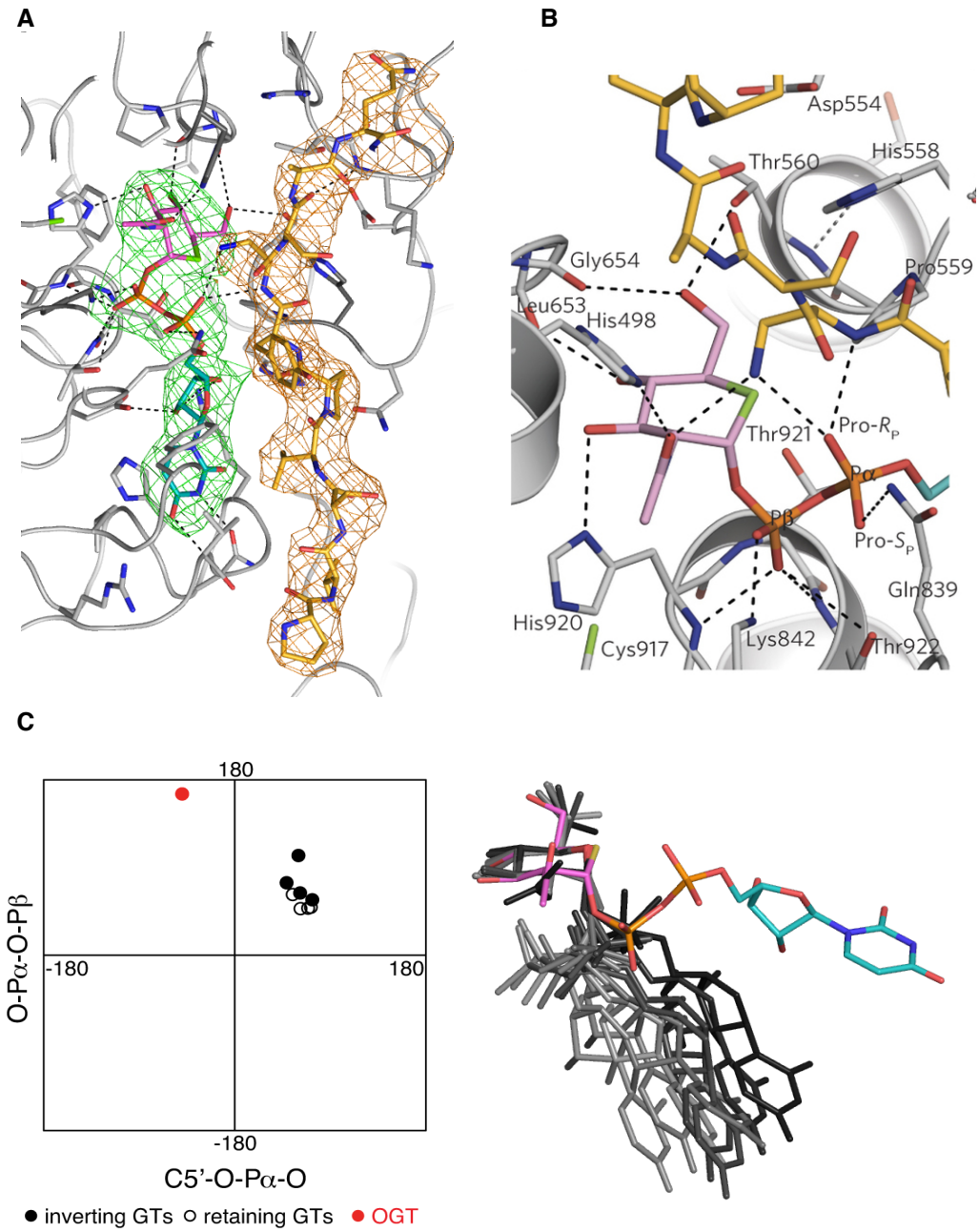


Figure 3.8: The unusual conformation of the sugar nucleotide in the hOGT pseudo-Michaelis complex suggests substrate-assisted catalysis.
A. hOGT pseudo-Michaelis complex with full length peptide ligand (yellow, GlcNAc is in magenta) and UDP (cyan) showing unbiased *Fo-Fc* electron density after four-fold averaging (2.5 σ). **B**. hOGT pseudo-Michaelis complex. Close-up view of the active site of hOGT (gray carbons) in complex with the donor analog and inhibitor UDP-5S-GlcNAc (pink carbons for the sugar, turquoise carbons for the uridine moiety) and the synthetic peptide aaTAB1tide (acceptor serine is replaced with an aminoalanine, yellow carbons). Hydrogen bonds are depicted by dashed lines. The α and β phosphates of UDP-GlcNAc are labeled with Pα and Pβ, **C.** respectively.Comparison of sugar nucleotide conformations from complexes with GT-B family enzymes. The structures of donor substrates in complex with active GT-B family enzymes deposited in the PDB database are shown superimposed on the sugar ring. The donor substrate from the hOGT complex is shown with pink carbons for the sugar and turquoise carbons in the uridine moiety. Sugar nucleotide coordinates were obtained from crystal structures of the following enzymes and are depicted in increasingly lighter shades of gray: (*Continue to Next page*)

Figure 1.7: Continued…
inverting enzymes MurG (PDB code 1NLM), UGT71G1 (2ACW), UGT72B1 (2VCE) and VvGT1 (2C1Z) and retaining enzymes AGT (1Y6F), OtsA (1UQU and 1UQT) and WaaG (2IW1). Selected torsion angles of the donor substrates of active glycosyltransferases belonging to the GT-B family plotted in a two-dimensional graph. Analyzed by M.S.

previously reported structures of GT-B donor complexes (Fig. 3.8C). It appears that binding in the OGT active site induces a 'backbent' UDP-GlcNAc conformation, characterized by unusual torsion angles of the pyrophosphate, which positions the sugar directly opposite the α-phosphate (Fig. 3.8C). This unusual donor conformation brings the α-phosphate pro-R_P oxygen to within 2.8 Å of the aminoalanine amino group mimicking the serine hydroxyl, suggesting that they would form a hydrogen bond. Concomitantly, the carbonyl group of the GlcNAc N-acetyl group approaches the serine analog to within 2.9 Å. In fact, these two substrate moieties approach the serine analog side chain more closely than any atom on the OGT enzyme itself. The acceptor appears to be positioned for nucleophilic attack with in-line displacement on the sugar anomeric carbon (angle nucleophile-anomeric carbon-leaving group = 151°), yet the side chains of H498 and H558 remain > 4.5 Å away from the acceptor serine as in the product complex. Thus, inspection of this pseudo-Michaelis complex leads to the tentative identification of two non-enzymic functional groups, residing on the substrate, as candidates for the elusive catalytic base. The hypothesis was tested by devising derivatives of the substrate, UDP-GlcNAc, analogous to the common practice of testing enzymic bases by site-directed mutagenesis.

	WT	D554N	H558F	K842M
UDP	0.54 ± 0.01	0.60 ± 0.03	0.7 ± 0.1	No binding
UDP-GlcNAc	16.1 ± 0.1	32.1 ± 0.3	42.3 ± 0.4	4.7 ± 0.1
UDP-5S-GlcNac	7.5 ± 0.1	6.2 ± 0.1	14.6 ± 0.1	5.7 ± 0.1
S_P-αS-UDP-GlcNAc	16.2 ± 0.1	10.8 ± 0.1	19.6 ± 0.1	50.0 ± 0.1
R_P-αS-UDP-GlcNAc	11.3 ± 0.1	11.5 ± 0.1	14.0 ± 0.1	2.8 ± 0.1
UDP-GlcNAcF$_3$	12.1 ± 0.1	9.5 ± 0.1	13.8 ± 0.1	29.3 ± 0.1

(Unit = μM)

Table 3.1: Binding affinity (K_d ± s.d.) of UDP, UDP-sugars and the α-phosphorothioate analogs of UDP-GlcNAc as determined by SPR with I.N (sensograms are shown in the appendix 4).

3.3.4. Chemical Probes and Identification of the Catalytic Base

3.3.4.1. N-Acetyl group of GlcNAc contributes to the binding but not catalysis

To investigate the possibility of a mechanism involving substrate-assisted catalysis, a number of UDP-GlcNAc derivatives were prepared by Dr Vladimir Borodkin (Fig. 3.7). The two moieties considered as possible candidates for the catalytic base were the carbonyl oxygen of the *N*-acetyl group, as well as the non-bonding (pro-R_P) oxygen of the α-phosphate, due to their proximity to the acceptor serine analog in the pseudo-Michaelis complex (Fig. 3.8B). While the carbonyl oxygen of the N-acetyl group has been shown to act as the catalytic nucleophile in O-GlcNAc hydrolysis (Macauley et al, 2005), it is not likely to act as a general base catalyst in O-GlcNAc transfer, as the low pKa (approx. – 0.5) of such a moiety makes it an unlikely proton acceptor. Lately, a computational study based on the hOGT-UDP-peptide structure proposed the proton extraction performed by the N-acetyl group (Tvaroska et al, 2012), although the study was based on the early hOGT structure that lacks the sugar moiety. To test the possibility of which as the catalytic base, the previously described *N*-

trifluoroacetyl-UDP-GlcNAc derivative (UDP-GlcNAcF_3, Fig. 3.7) was prepared (Sala et al, 1998), with the aim of affecting the electronegativity of the N-acetyl carbonyl moiety. We found that the UDP-GlcNAcF_3 compound bound the enzyme with a K_d similar to UDP-GlcNAc and proved to be a functional donor substrate for hOGT (Table 3.1), which excludes a significant catalytic role of the N-acetyl group.

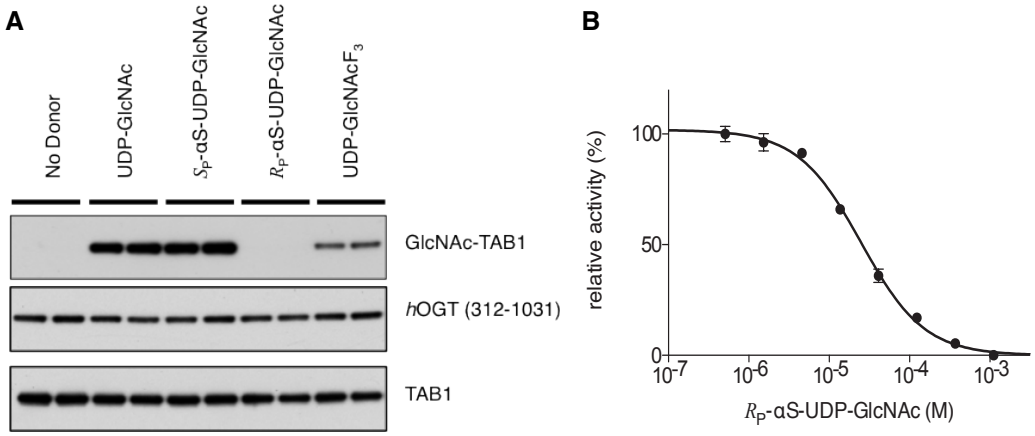


Figure 3.9: In vitro O-GlcNAcylation assay using mechanism-inspired UDP-GlcNAc analogs (A) and Inhibition of hOGT by the donor analog R_P-αS-UDP-GlcNAc (B)

A. O-GlcNAcylation of TAB1 by hOGT (residues 312–1031) was detected by immunoblotting with a site-specific TAB1 O-GlcNAc Ser395–specific antibody. Experiments were performed in duplicate, as shown.

B. Activity of hOGT on biotinylated TAB1tide was determined radiometrically as described in the Experimental Procedures. Measurements were performed by Dr David Blair in triplicate (error bars represent the s.e.m.) and values were fitted to the standard equation for dose-dependent inhibition in GraphPad Prism 5.0. R_P-αS-UDP-GlcNAc (4) inhibited hOGT with an IC50 of 25 μM.

3.3.4.2. Identification of the catalytic base and proposed non-enzymic catalytic mechanism

As expected, both diastereomers were able to form a complex with the enzyme with K_d values similar to that of UDP-GlcNAc (Table 3.1) and underwent non-enzymatic hydrolysis at rates similar to UDP-GlcNAc (Appendix 5). The S_P diastereomer was found to be a functional donor in an O-GlcNAc transfer reaction (Fig. 3.9A). Strikingly, the R_P diastereomer, where the sulfur replaces

the oxygen pointing towards the acceptor serine, was not a substrate (Fig. 3.9A), and indeed inhibited the reaction with an IC50 of 25 μM (Fig. 3.9B), while not affecting binding of the acceptor peptide (Table 3.1). These data suggest that the unusual conformation of UDP-GlcNAc positions the α-phosphate to act as the catalytic base in hOGT catalyzed O-GlcNAc transfer (Fig. 3.11), a mechanism not previously reported for other Leloir-type glycosyltransferases and apparently unique to OGT. Further evidence supports this mechanism. Dr. David Blair has shown that the acidic leg of a pH-activity profile of hOGT (Fig. 3.10) exhibits a pKa of 5.5. In the proposed catalytic mechanism there is also a need to stabilize the developing negative charge on the β-phosphate of the UDP leaving group. Inspection of the substrate/product complexes reveals a perfectly shaped oxyanion hole, constructed by three elements: hydrogen bonds from the backbone amides of H920, T921 and T922, alignment with a α-helical electrostatic dipole and interaction with the evolutionarily conserved K842 (Fig. 1.4). Strikingly, the K842M mutation is

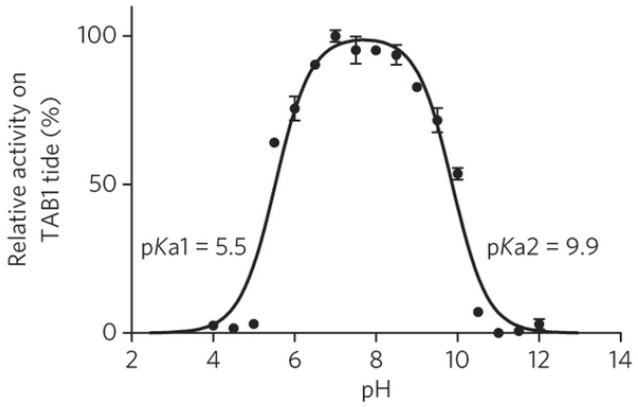


Figure 3.10: pH profile of hOGT.
Activity of wild-type hOGT on N-terminally biotinylated TAB1tide in phosphate buffer with pH 4–12 was measured by scintillation proximity assay by Dr David Blair. Data points show the mean ± s.e.m. of three observations.

catalytically inactive (Fig. 3.4), and no longer binds the UDP product as shown by SPR (Table 3.1). With a theoretical pKa of 10.5, deprotonation of this lysine may represent the basic leg of the pH-activity profile of hOGT (Fig. 3.10). The hOGT pseudo-Michaelis complex reveals that this lysine stabilizes the unusual conformation of the β-phosphate that is required for correct positioning of the α-phosphate opposite the acceptor serine (Fig. 3.8B). Finally, the proposed mechanism explains the surprising observation that UDP -5S-GlcNAc is a potent and specific inhibitor of hOGT, but not of any other human GlcNAc transferases (Gloster et al, 2011). The smaller C5-S-C1 bond angle, the longer C5-S/S-C1 bonds and the larger van der Waals radius of the sulfur (Fig. 3.8B) would all perturb the trajectory of proton transfer from the acceptor serine hydroxyl onto the UDP α-phosphate, a process that so far appears to be unique to hOGT.

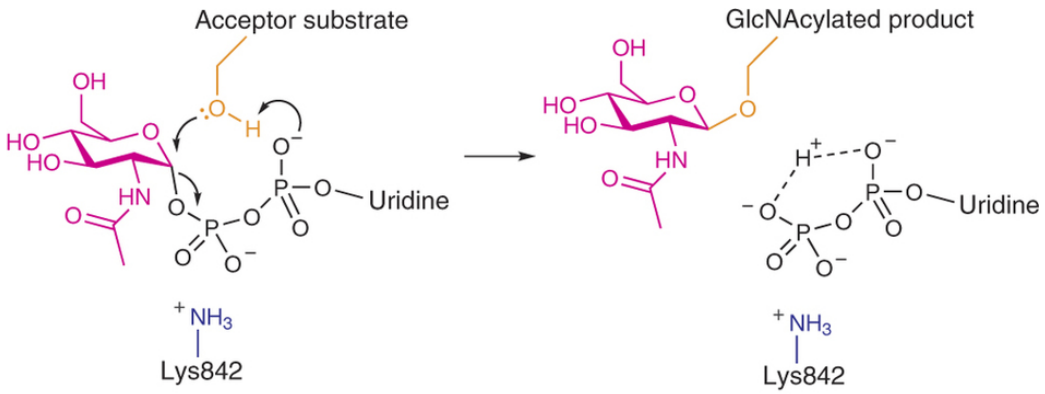


Figure 3.11: Schematic representation of the proposed catalytic mechanism of hOGT, showing substrate-assisted catalysis involving the sugar donor phosphates.

3.4. Conclusion:

OGT is an essential metazoan O-GlcNAc glycosyltransferase that targets specific sites on a large number of protein substrates, utilizing UDP-GlcNAc as the donor substrate. We report crystal structures of ternary substrate and product complexes, providing snapshots of hOGT catalysis. These data show that hOGT binds peptide substrates in a defined orientation and with similar conformations near the site of O-GlcNAc transfer. Together with synthetic probes, these structures define a new mechanism of inverting glycosyl transfer, where the catalytic base is not provided by the enzyme, but by the α-phosphate on the donor substrate itself. This differs from other Leloir-type GTs, where this function is generally performed by a side chain carboxylate or imidazole (Lairson et al, 2008). Participation of substrate phosphates in enzymatic catalysis is an established concept for farnesyl pyrophosphate synthases, and the cumulative evidence of multiple crystal structures of Clostridium *spp* (Hosfield et al, 2004) glucosylating toxins has led to the suggestion of the β-phosphate of UDP-glucose as the base in these inverting GT-A family enzymes (Ziegler et al, 2008). Lately, a similar mechanism has been proposed for the GDP-fucose Protein O-fucosyltransferase (POFUT) (Lira-Navarrete et al, 2011), an inverting GT-B family enzyme, although a recent study reporting the structure of a GT from the same family, POFUT2 proposes an enzyme side chain as the catalytic base (Chen et al, 2012). Intriguingly, the substrates for POFUT1 are specific serine residues on EGF repeats of the Notch receptor, and the clostridial toxins glucosylate Rho GTPases, so like OGT, these enzymes act on protein substrates. The involvement of the UDP-GlcNAc α-phosphate in catalysis may be what enables OGT to act on both Ser and Thr

acceptors, as it provides an evolutionary solution to the simultaneous spatial constraints imposed by a large peptide acceptor, the necessity to recognize the N -acetyl group on the UDP-GlcNAc donor and the need to accommodate the additional γ-methyl group on threonines. The mechanistic insights into hOGT catalysis are invaluable for the design of drug-like inhibitors to facilitate further research into the cell biological role of O-GlcNAc, and the targeting of human diseases such as diabetes and cancer. It is noteworthy that as a part of our mechanistic investigations, we have now identified a hOGT point mutant, K842M, that appears to be unaffected in substrate binding but is catalytically inactive and would be useful for cell biological dissection of the role of OGT as an enzyme versus a scaffolding function through its TPR repeats.

3.5. Material and Methods

Crystalisation of hOGT

Crystallisation experiments were performed by Dr. Marianne Schimpl. Protein used in crystallisation was stored in 25 mM Tris-Cl, pH=7.5, 150 mM NaCl and 1 mM THP (instead of 1 mM DTT). Protein was crystallized in complex with donor substrate/product, and crystals were soaked with acceptor/product prior to freezing. Vapor diffusion crystallization experiments with hanging drops containing 1 µl protein (100 µM in a buffer of 10 mM Tris-Cl pH 8.5, 50 mM NaCl, 0.5 mM THP, 1 mM UDP or UDP-5S-GlcNAc) and 0.6 µl reservoir solution (1.45 M K_2HPO_4, 10 mM EDTA, 1 % xylitol) gave bar-shaped crystals with maximum dimensions of 0.1 × 0.1 × 0.4 mm after 3—4 days at room temperature. These were transferred to a drop of reservoir solution containing 2 mM of the (glyco)peptide (Ac-PVSVPYS(-β-O-GlcNAc)SAQSTS-NH2) for 30 min, then cryoprotected (1.45 M K2HPO4, 10 mM EDTA, 27 % xylitol) and flash-frozen. Data were collected at the European Synchrotron Radiation Facility (ESRF) at 100 K and wavelengths of 0.939 Å and 0.873 Å on beamlines ID14-4 and ID23-2, respectively. Crystals belonged to space group P321 and contained 4 molecules per asymmetric unit. The structure was solved by molecular replacement using the A chain of PDB ID 3PE3 as the search model. Model building was performed in Coot, and various programs of the CCP4 suite were used for structure refinement. Ligand topologies were calculated using PRODRG.

In vitro *Glycosylation of hTAB1*

1 mg/ml of purified hTAB1 (7—402) was incubated with 0.05 mg/ml of purified hOGT in 100 mM potassium phosphate pH 7.5 containing 1 mM UDP-GlcNAc and 1 mM DTT. The reaction was allowed to proceed for 1.5 h at 37 °C and stopped by addition of SDS loading buffer and heating to 95 °C. The samples were subjected to SDS-PAGE and transferred to PVDF membrane followed by immunological detection of OGT, TAB-1 and OGlcNAc with primary and secondary antibodies diluted 1:5,000 in TBS-Tween containing 3 % bovine serum albumin. Site-specific antibody against O-GlcNAc-TAB1 S395 was raised in rabbit against a keyhole limpet hemocyanin (KLH)-conjugated glycopeptide as described by Pathak *et al.*(Pathak et al, 2012). The anti-TAB1 antibody was obtained from the Division of Signal Transduction and Therapy, University of Dundee (DSTT).

hOGT activity measurements by scintillation proximity assay

The experiments were performed with the help from Dr. David Blair. Radiometric detection of OGT activity on peptide substrates was achieved through scintillation proximity technology (PerkinElmer). Assays were conducted in 20 µl format in 384-well polypropylene plates. Reactions contained 200 nM hOGT (312-1031), 2 µM biotinylated substrate peptides and 500 nM UDP-GlcNAc with 0.3 Ci/mmol UDP-[^{3}H]-GlcNAc as a radioactive tracer, 100 mM potassium phosphate, 1 mM DTT and 0.2 mg/ml BSA. The reaction was stopped by addition of 40 µl of 0.75 M phosphoric acid, and transferred to a streptavidin-coated FlashPlate®-384 for detection on a TopCount NXT microplate luminescence counter.

Surface Plasmon Resonance

The experiment was performed with the help form Dr. Iva Navratilova. SPR measurements were collected using a Biacore T100 instrument (GE Healthcare). Streptavidin was immobilized on a CM5 sensor chip using standard amine coupling method. 10 mM HEPES, 150 mM NaCl, pH 7.4 was used as a running buffer for immobilization. hOGT was biotinylated by mixing of hOGT with amine-binding biotin (Pierce) in 1:1 molar ratio. The chip surface was primed with running buffer (25 mM Tris pH 7.5, 150 mM NaCl, 1 mM DTT and 0.05 % Tween 20) and biotinylated hOGT protein was captured on the streptavidin surface. All compounds were injected in duplicates with highest concentrations from 10–500 μM depending on affinity, followed by a 1:3 dilution series. Association was measured for 1 min and dissociation for 2 min. All experiments were run at flow rate 30 μl/min and temperature 25 °C. All data were referenced for blocked streptavidin surface and blank injections of buffer. Srubber 2 (BioLogic Software) was used to process and analyse data. Affinities were calculated using 1:1 equilibrium binding fit.

Chemical synthesis of UDP-GlcNAc analogs and peptides

Experimental procedures have been reported in (Schimpl et al, 2012).

Differential scanning fluorimetry

5 μM of hOGT proteins was mixed with 1:1000 dilution of Sypro Orange in the final volume of 40 μl containing 25 mM Tris-Cl, pH=7.5, 150 mM NaCl. The mixture was contained in clear 96-well PCR plate and the fluorescence was

measured by thermocycler (BioRad) with raising 1 °C/min from 25 to 95 °C. Data analysis was based on (Matulis et al, 2005).

Biolayer interferometry measurements

The experiments were performed with the help from Dr. David Robinson. Peptide binding affinity to hOGT in the presence of donor substrate was measured using the OCTET® RED384 system (ForteBio). hOGT was biotinylated by mixing with amine-binding biotin (Pierce) in 1:1 molar ratio. Biotin-hOGT was immobilised onto Superstreptavidin biosensors followed by blocking with 10 µg/ml biocytin. A second set of Superstreptavidin biosensors without hOGT coupling was blocked with biocytin, and washed and equilibrated in buffer (25 mM Tris-Cl, pH=7.5, 150 mM NaCl, 1 mM DTT) containing UDPGlcNAc or R_P-αS-UDP-GlcNAc at saturating concentration of 500 mM. The biosensors then sampled a 15 dilution series (4 points, 4-fold dilutions) of TAB1tide starting from a top concentration of 2.5 mM. Each experiment was run using two biosensors and at temperature 25 °C. The data were processed with ForteBio Data Analysis v.7.0. Double referencing was used to correct for instrument drift and non-specific interactions between TAB1tide and the biosensor surface. To obtain the dissociation equilibrium constant (K_d) the sensograms were fitted to a global Langmuir 1:1 model using 3 concentrations.

4. Chapter III

Structural And Functional Analysis Of *Drosophila Melanogaster* O-GlcNAc Transferase

4.1. PROJECT OVERVIEW

As aforementioned in the section 1.5.1, OGT and O-GlcNAcylation have been reported to involve in various cellular processes and is essential during embryogenesis using genetic models. Many of studies in these animal models relied on the complete removal of the *ogt* gene, which leaves the question whether the observed lethality is caused directly by the loss of OGT enzymatic activity and consequently the lack of O-GlcNAc modification, or whether it is an effect caused by the absence of OGT protein, which is known to participate in numerous protein-protein complexes and might have an essential scaffolding function. Here, I have used *Drosophila melanogaster* as a model organism to study the importance of O-GlcNAcylation in early embryogenesis and to attempt to separate the enzymatic and non-enzymatic functions of OGT. I report the crystal structure of *Drosophila melanogaster* OGT (*Dm*OGT), and demonstrate that *Drosophila* and human OGTs share structural and enzymatic similarities. Transgenic flies expressing catalytically inactive *Dm*OGT in the *ogt/sxc* null background were unable to rescue the lethality suggesting the enzyme activity, although with almost undetectable amount of O-GlcNAc by antibody, is important during embryogenesis and early development in *Drosophila*.

4.2. The aims of the work presented here were:

- To compare OGTs at structural level
- To identify a peptide substrate enabling *Dm*OGT *in vitro* activity measurements
- To establish Michaelis-Menten kinetics for *Dm*OGT
- To generate and verify point mutations that abolish the catalytic activity and/or the substrate binding of *Dm*OGT
- To use these mutant forms of *Dm*OGT in genetic complementation experiments with the aim of deconvoluting the enzymatic and non-enzymatic roles of OGT *in vivo*

4.3. RESULTS AND DISCUSSION

4.3.1. *Dm*OGT expression and crystallization

OGT is highly conserved between human and *Drosophila* (Fig. 1.5 & 4.1), with 90 % sequence identity in the TPR region and 87% sequence identity in the catalytic domain. The most variable domain appears to be the 'intervening domain', a 135 amino acids insertion between the two Rossmann-like domains that constitute the catalytic domain. Sequence identity in the intervening domain is only 42 %, and the length differs by 8 amino acids. In order to understand to what extent these differences are reflected in the structure and function of *Dm*OGT, I crystalized an N-terminally truncated construct starting at amino acid 353 in TPR 10 (353-end), which is equivalent to the Δ9-hOGT construct. Protein was recombinantly produced in *E. coli* as for hOGT described previously (see Chapter 1). *Dm*OGT expression and purification

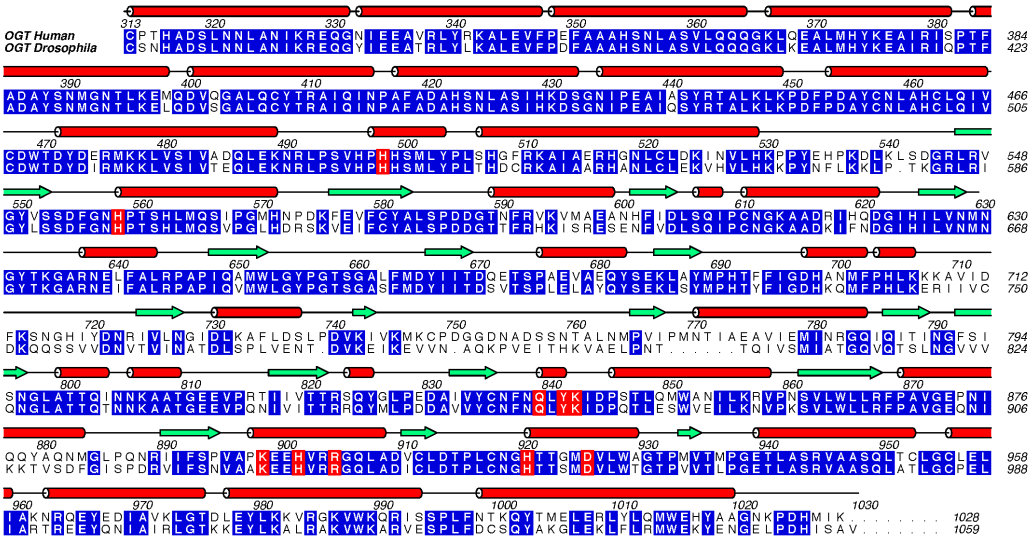


Figure 4.1: Sequence alignment between Drosophila and human OGT at the catalytic domain.
Secondary structures of hOGT (312-1028) and *Dm*OGT (353-end) are shown (sheet arrows: β-strand; barrels: α-helices). Identical residues are shaded in blue and residues interacting with UDP-GlcNAc are shaded in red. Alignment was generated with *Clustal W* and displayed using Aline.

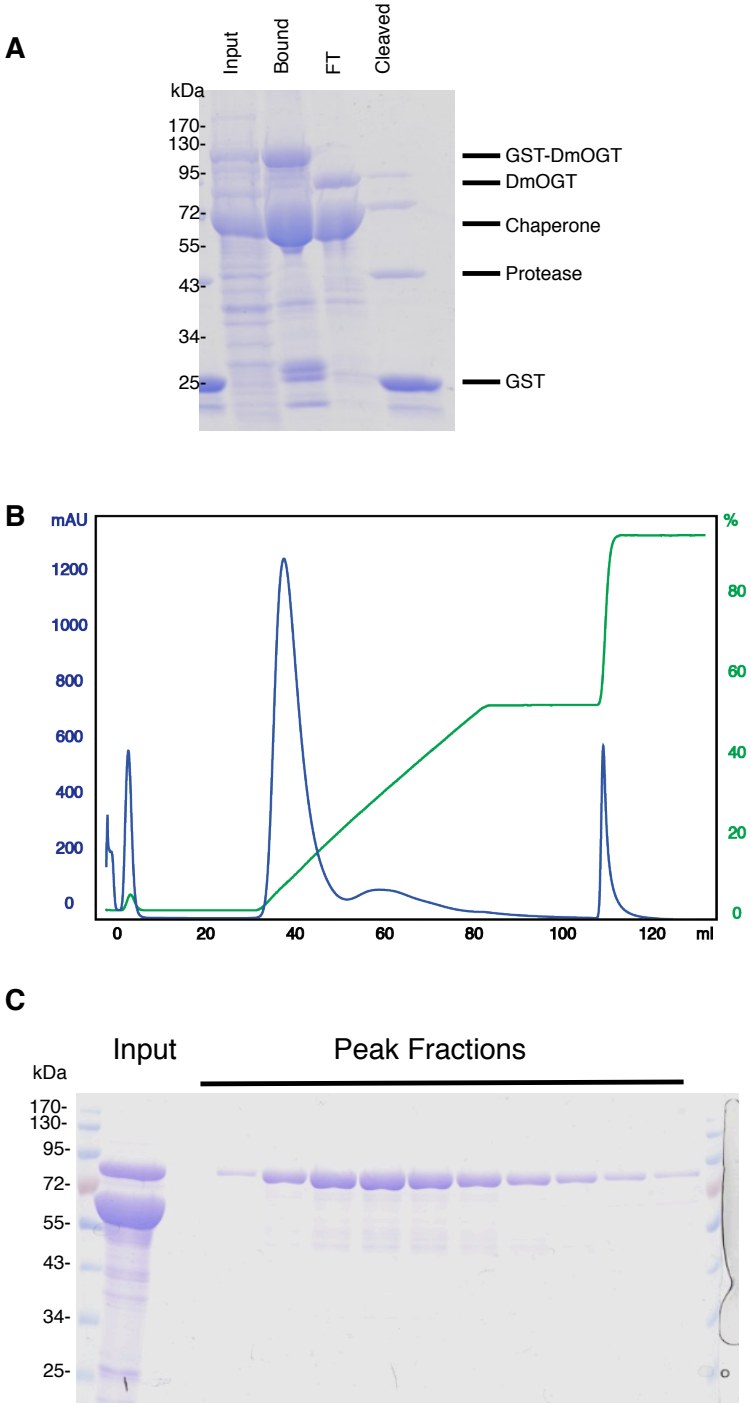


Figure 4.2: Expression and purification of *Dm*OGT.
A. *Dm*OGT was expressed in ArcticExpress *E. coli* as for hOGT. GST-tagged protein was affinity-purified on glutathione sepharose beads (Bound) and the tag was cleaved with Prescission protease, releasing *Dm*OGT to the flow-through (FT) and leaving the protease on the beads (Cleaved). **B.** Affinity purified DmOGT (353-end) (C, input) was loaded onto HiTrap Q FF anion exchange column. Protein was eluted with a 0-1 M NaCl gradient (Green line). Purity of fractions was verified by Coomassie blue staining before pooling (**C**).

proceeded similarly to that of hOGT, with a bacterial chaperone protein co-eluting in the initial affinity chromatography step (Fig. 4.2). This was later removed by anion exchange chromatography, yielding pure *Dm*OGT protein in fair yield (approx. 0.5 mg/l culture volume) (Fig. 4.2). In order to crystallize *Dm*OGT, a crystallization screen was carried out with ~600 different conditions using sitting drop method. Crystals appeared in 4 conditions of the initial screen, all of which contain a combination of high molecular weight polyethyleneglycol (PEG, above 6000 in mw) and low molecular weight alcohols/polyols such as ethylene glycol or glycerol, with pH between 6.5 and 8.5. This is in marked contrast to hOGT, which has been reported in two crystal forms originating from three different buffer conditions containing high concentrations of salts (ammonium or lithium sulfate and potassium phosphate). Further optimization of the crystallization condition led to needle-shaped crystals with maximum dimensions of approximately 0.3 x 0.05 mm,

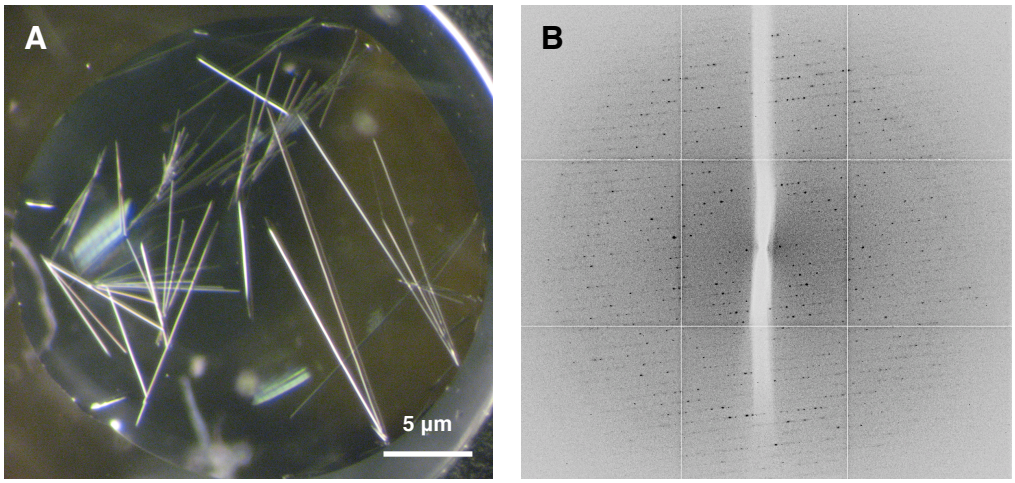


Figure 4.3: Crystallization of *Dm*OGT (353-end) and X-ray diffraction image.
A. Needle-shaped crystals grew from a 5 mg/ml protein solution in 12% PEG8000, 16% glycerol and 0.1 M HEPES 7.5; and appeared after 2 days. **B.** X-ray diffraction images were recorded at Diamond Light Source, UK. Crystals diffract to 3.3 Å.

which belonged to space group H3 and diffracted to >3 Å (Fig. 4.3 & table 4.1). X-ray diffraction data were collected at Diamond Light Source, and the structure of *Dm*OGT (353-end) was solved by molecular replacement using the hOGT structure (pdb: 3PE4) as the search model. Due to the relatively low resolution, the absence of any exploitable non-crystallographic symmetry and regions of disorder, no attempts were made to fully refine this structure. Refinement was terminated after 2 macro cycles.

Table 4.1: Data collection and refinement statistics of DmOGT

Data Collection	
Space group	*H3*
Cell dimensions	
a=b, c (Å)	158.2, 76.2
Resolution (Å)	79.1-3.3 (3.39-3.30)
R_{merge}	0.079 (0.034)
$I/\sigma I$	8.9 (2)
Completeness (%)	99.6 (99.8)
Redundancy	2.9 (2.9)
Refinement	
Resolution (Å)	30-3.3
No. reflections	10118
R_{work}/R_{free}	33.8/41.1
No. atoms	5514

Highest-resolution shell is shown in paretheses

4.3.2. *Drosophila* OGT adopts a canonical GT fold

*Dm*OGT adopts the canonical OGT fold with the bi-lobular arrangement of two Rossmann-like domains of the classic GT-B superfamily glycosyltransferases; as well as the additional TPR-like helices (535-566) N-terminal of the catalytic domain, which lead into to the TPR domain. As a result, the TPRs are in close association with the glycosyltransferase domain and the catalytic site is

aligned with the channel of the TPR superhelix. Superposition of the catalytic domain of *Drosophila* OGT with reported hOGT structure (4AY6) highlights the structural similarities (RMSD=0.43 Å for 698 Calphas). The intervening domain, an insertion between the two Rossman-like folds of the catalytic domain, faces away from the catalytic site and does not appear to be involved in peptide substrate recognition. During structure refinement of *Dm*OGT, we noticed that the electron density in the intervening domain is less well defined than throughout the glycosyltransferase domain, possibly indicating that the intervening domain shares less structural similarities to the search model hOGT (pdb: 3PE4), or conformational flexibility. This is consistent with the lower sequence conservation observed in this region of the protein. The functions of the intervening domain are not clear despite the high-resolution structure of hOGT (Lazarus et al, 2011). As previously suggested in chapter 1, the intervening domain may be involved in substrate protein recognition and the structural differences may contribute to differences in substrate recognition between hOGT and *Dm*OGT.

Sequence conservation among surface-exposed amino acids is considered a sign of functional importance, and can help to identify sites of protein-protein interaction. The surface of the *Dm*OGT structure reveals a patch of conserved residues mostly concentrated around the catalytic site, but also in the TPR domain. The surface-exposed residues involved in UDP-GlcNAc and peptide binding are almost identical to hOGT (Fig. 4.4). This conservation is less profound in the rest of the catalytic domain. In particular, a group of ten lysine

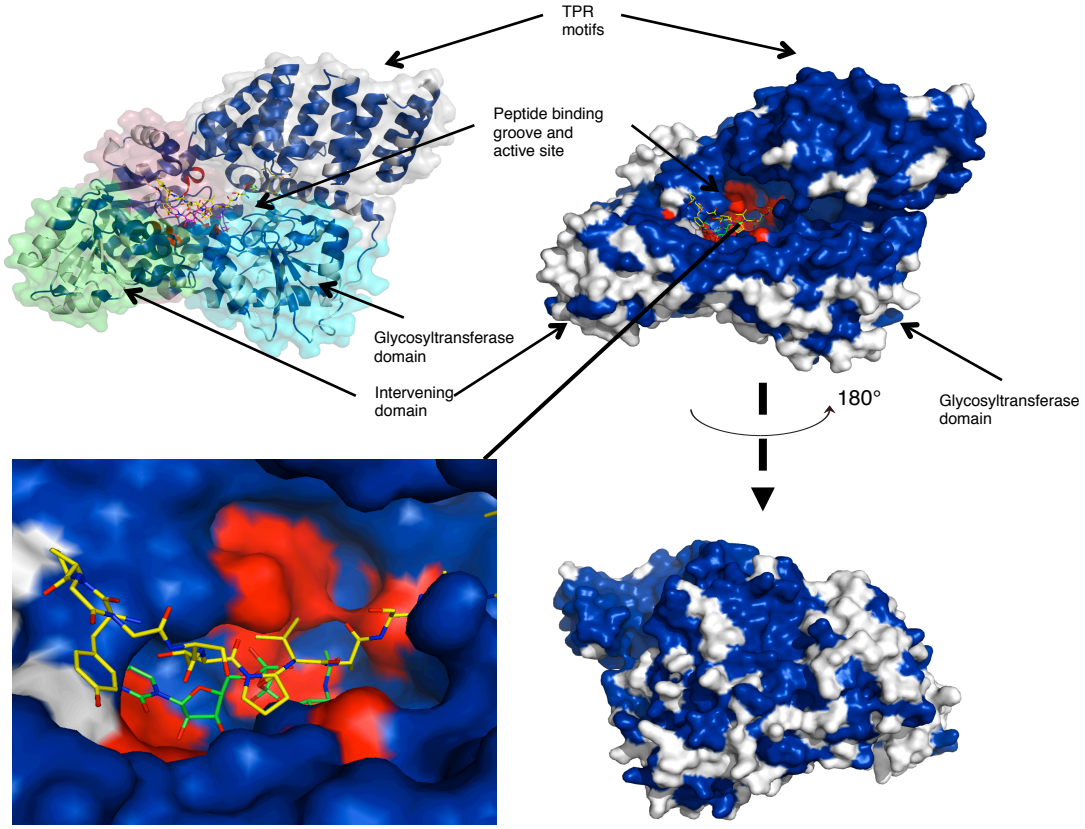


Figure 4.4: overall structure of DmOGT and its conservation between human and Drosophila.
*Dm*OGT adopts the canonical OGT fold with the intervening domain (green surface) and the TPR repeats (gray surface) closely associated with the glycosyltransferase domain (cyan and pink surface) (top left). The surface of *Dm*OGT coloured by sequence conservation with hOGT (right panel). Identical residues are shaded in blue, and non-conserved residues are shown in white. A zoom in view of the catalytic site of *Dm*OGT with UDP-5SGlcNAc (green sticks) and TAB1tide (yellow sticks) from superimposed hOGT structure (pdb: 4AY6) (bottom left). Red surface indicates residues that interact with UDP-GlcNAc.

residues that form a positively charged pocket in hOGT and have been proposed to interact with phosphatidylinositol-(3,4,5)-triphosphate (PIP_3), are not conserved in *Dm*OGT (Fig. 4.1). No such positively charged pocket is observed in the *Dm*OGT structure. Binding of PIP_3 to hOGT has been proposed to mediate translocation of hOGT to the plasma membrane upon insulin stimulation (Yang et al, 2008). However, recent structure-guided mutagenesis studies in hOGT could not confirm the binding of PIP_3 to these lysine residues (Lazarus et al, 2011). The absence of a similar positively charged patch on the surface of *Dm*OGT makes it unlikely that PIP_3 binding could regulate the subcellular localization of OGT in *Drosophila*.

4.3.3. *Dm*OGT and hOGT possess different peptide substrate sequence specificities

Previously, *in vitro* activity of *Dm*OGT has been demonstrated using rat nuclear pore protein p62 as a substrate (Gambetta et al, 2009; Sinclair et al, 2009), and no O-GlcNAc modified *Drosophila* protein has yet been studied in detail. However, no substrates are currently known that would allow determination of steady state kinetics of wild type and mutant forms of the protein. To identify such a substrate, a peptide substrate screen was performed on a library of 720 peptides derived from annotated Ser/Thr phosphorylation sites on human proteins. The library consisted of synthetic 13-amino acid peptides with an N-terminal biotin tag, and each peptide contained at least one Ser or Thr for post-translational modification. This library was screened in 384-well plate format using a scintillation proximity assay

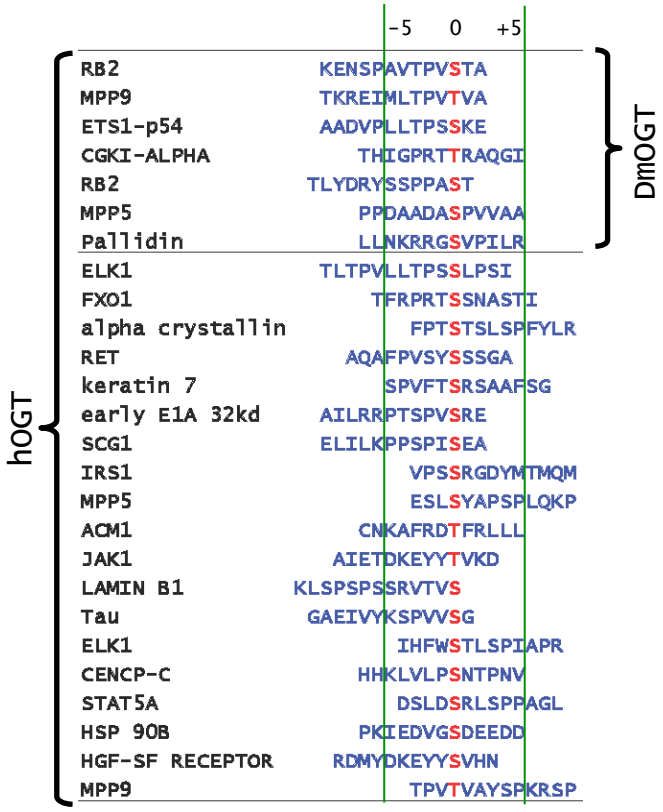


Figure 4.5: Peptide modified by *Drosophila* and human OGT

Peptide sequences identified from an activity assay against a library of synthetic peptides. Listed peptides have been shown to be O-GlcNAc modified by hOGT at specific Ser or Thr identified with ETD mass spectrometry (performed by D.B and O.A). Seven peptides recognized by *Drosophila* and human OGT are shown on top. The peptides are aligned at the modified amino acids, which are coloured in red. Green lines mark the boundaries of peptide from -5 to +5. Notably, the modification sites of substrates for *Dm*OGT are within 5 amino acids from the C-terminus.

developed for high-throughput OGT activity measurements. The highest signal was obtained for a peptide derived from human retinoblastoma-like protein 2 (RBL2) (KENSPCVTPVSTA). The same sequence has also been observed in a similar screen as an optimal hOGT peptide substrate (unpublished data). RBL2 is a retinoblastoma family member, which all share high degree of structural and functional homology, such as cell cycle, cell differentiation and apoptosis (reviewed in (Bellacchio & Paggi, 2013)). RBL2 is conserved in

Drosophila with 40% sequence identity, however the peptide sequence and modification site is not conserved in the *Drosophila* protein. The peptide library screen also gives insights into the substrate peptide selectivity of *Dm*OGT. A similar experiment has previously been performed for hOGT (Dr. Shalini Pathak, unpublished work), thus allowing direct comparison of the peptide substrate specificity of the two enzymes. The average number of Ser or Thr in these 13 amino acid peptides is 2.8 and the site of modification needs to be established independently. For the top 40 hits of the hOGT screen, the O-GlcNAc sites have been unambiguously identified with ETD mass spectrometry. For the purpose of this analysis, I assumed that the modification sites are the same for *Dm*OGT, although this was not experimentally verified. Surprisingly, among the 40 highest-scoring peptides of both screens, only 7 are common for hOGT and *Dm*OGT, so there is indirect evidence for the sites on 7 peptides. In all these *Dm*OGT substrate peptides the modification sites were within 5 amino acids of the C-terminus of the peptide (Fig. 4.5). The limited number of the peptides does not give conclusive results, but leads to speculation of the selectivity of *Dm*OGT on the substrate peptides. Similar to hOGT, there appears to be no strict sequence pattern (or 'sequon') for O-GlcNAc modification of proteins.

Interestingly, *Dm*OGT itself is found to carry an O-GlcNAc modification, which could be removed by a bacterial O-GlcNAcase, *Cp*OGA (Fig 7A). Mass spectrometry analysis located the O-GlcNAc modification (Ser389) predominantly in the region of TPR of *Dm*OGT (Fig. 7B). O-GlcNAcylation on

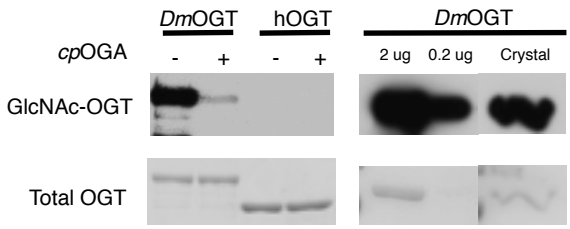


Figure 4.6: O-GlcNAc modification of *Dm*OGT
O-GlcNAc modification was shown by Western blot on bacteria expressed *Drosophila* (353-end) and human (412-end) OGT with or without *cp*OGA treatment, which removes O-GlcNAc modification (left). Right panel shows the O-GlcNAc modified DmOGT in crystals. 5 crystals were taken and washed 5 times with the mother liquor to remove non-crystalised DmOGT and boiled in sample buffer before SDS-PAGE and Western blotting. Loading was shown with Coomassie blue.

the equivalent site has not been reported in hOGT. The position of this O-GlcNAcylation is on TPR 12 and at the interface between TPRs 12 and 13, which is a proposed site of structural flexibility in hOGT (Lazarus et al, 2011). In the currently available crystal structures of OGTs (bacterial, *Drosophila* and human), access to the active site is partly occluded by the TPR domain. This would either presuppose that O-GlcNAcylation could occur only on flexible loops or termini of substrate proteins, or that there is significant movement of the TPRs which would allow access of compact substrate proteins to the active site. Consequently, a 'hinge and latch' mechanism near the catalytic site has been proposed, based on a modest positional difference of the TPRs in the two crystal structures reported by Lazarus et al. (Lazarus et al, 2011). The authors report structures of hOGT in the presence and absence of a peptide substrate, and observe a 6 Å shift of TPR 10, and a slight opening of the catalytic cleft, in the presence of the peptide. This structural shift of the TPRs is due to a hinge-like movement between TPRs 12 and 13. Notably, the two structures originated from different crystal forms with unrelated space groups.

Crystal packing can limit the conformational flexibility of multidomain proteins, and while the observation does indeed support a certain degree of flexibility on the TPRs with respect to the catalytic domain, it is not clear that the movement is caused by substrate binding. Other methods better suited to the observation of dynamic processes will be required to settle the question of the proposed hinge-like movement. Our finding of the O-GlcNAc modification in this position could corroborate such a hinge-like mechanism, because a modification of this site would invariably disrupt the interaction between TPRs 12 and 13 (Fig. 4.7) and lead to a profound rearrangement of the TPR motifs. Although is not unusual that post-translational modifications induce structure rearrangement, it has not been observed in TPRs. O-GlcNAcylation on TPR 12 could stabilize an open conformation of *Dm*OGT, facilitating access to its active site and influencing its catalytic activity. O-GlcNAcylation is also observed in protein from the *Dm*OGT crystals (Fig. 7A). However, no electron density for the O-GlcNAc modification was observed in the current crystal structure, possibly due to the low resolution. There is little evidence, however, for a major rearrangement of the TPRs, although poorer electron density in this region suggests a higher degree of flexibility compared to the catalytic domain. It would be interesting whether this O-GlcNAc modification influent *Dm*OGT selectivity on peptide and protein substrates.

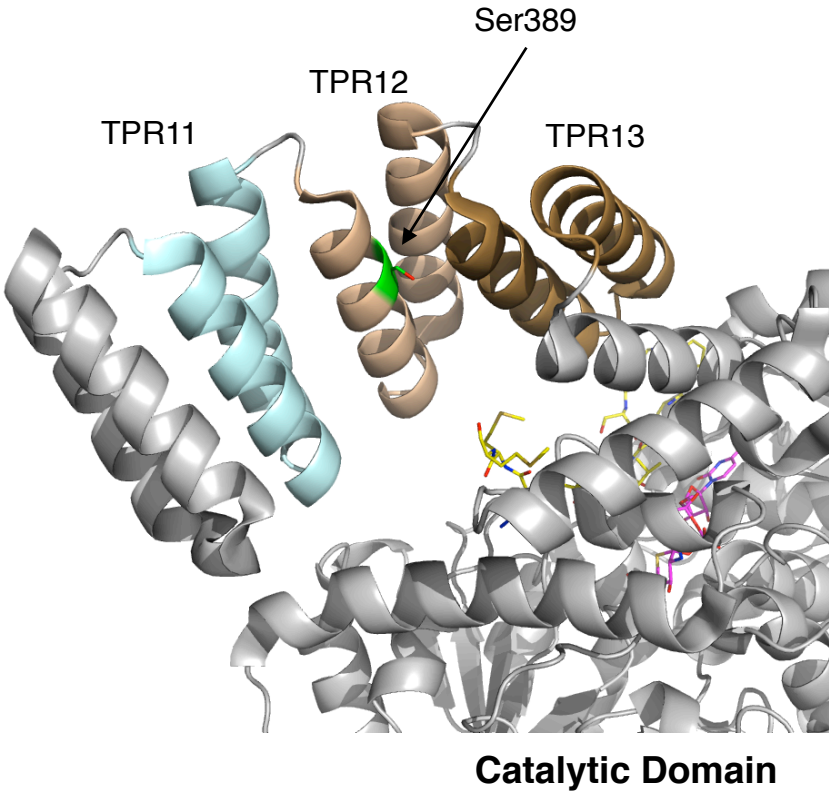


Figure 4.7: position of Ser389 in TPR12
Cartoon representation of the DmOGT catalytic domain (with substrates shown in stick representation) and TPRs 11, 12 and 13 with Ser389, which as been found to be O-GlcNAc modified. Ser389 is positioned at the interface between TPRs 12 and 13.

4.3.4. The hOGT catalytic machinery is conserved in *Dm*OGT

*Dm*OGT crystals diffracted to a maximum resolution of 3.3 Å, which enabled us to solve the structure. Co-crystallization with UDP, UDP-5SGlcNAc and peptide (KENSPCVTPVSTA) did not improve crystallization or the data quality, and no clear electron densities for UDP-5SGlcNAc or the peptide were observed. However, the catalytic site residues that are found to interact with the UDP-GlcNAc and the acceptor peptides in hOGT are conserved throughout (Fig. 4, bottom left), suggesting a similar catalytic mechanism with the same UDP-GlcNAc and acceptor peptide binding modes (Fig. 4). In order to explore the

mechanistic similarities between the two enzymes, Michaelis-Menten kinetic parameters of *Dm*OGT were determined and compared to those of hOGT using the same short expression constructs that has been shown to exert catalytic activity. The RBL 2 peptide that has been identified previously from the peptide screen was synthesised and used as *in vitro* substrate for *Dm*OGT, to determine Michaelis-Menten kinetics with a radiometric assay. Michaelis constant (K_m) of UDP-GlcNAc for *Dm*OGT is 17.8 ± 0.7 µM, which is only three fold higher than the K_m for hOGT at 6.6 ± 0.3 µM (Fig. 5A). Intracellular UDP-GlcNAc concentration is believed to be a key regulator of OGT enzymatic activity and therefore of global O-GlcNAcylation levels. The similar K_m values of UDP-GlcNAc for *Drosophila* and human OGT suggests that both proteins are subject to regulation by intracellular UDP-GlcNAc levels.

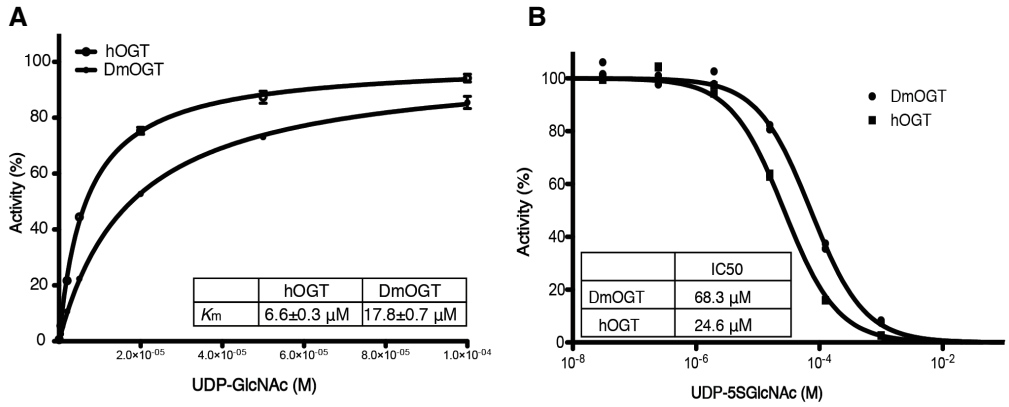


Figure 4.8: Michaelis constant of UDP-GlcNAc and *in vitro* inhibition of hOGT wt and *Dm*OGT wt by UDP-5SGlcNAc

A. Michaelis-Menten constant (K_m) of UDP-GlcNAc on DmOGT wt and hOGT wt were determined in a radiometric *in vitro* assay on RBL2 peptide. Triplicated data points were plotted with a Michaelis-Menten kinetic curve fit and K_m are shown in table. Error bars represent the standard error of the mean **B**. Half maximal inhibition concentration (IC50) of UDP-5SGlcNAc on *Dm*OGT and hOGT were determined using the same assay with UDP-GlcNAc concentration equal to the K_m for each enzyme. Duplicated data points were fitted to a three-parameter equation for dose-dependent inhibition and are shown in the table. Standard curve and enzyme *vs* time relationship were shown in the Appendix.

The proposed similarities in the catalytic mechanism of human and *Drosophila* OGT were further supported by inhibition with UDP-5SGlcNAc. A novel substrate-assisted glycosyltransfer mechanism has been proposed for hOGT as discussed in chapter 2. It has been observed that the donor substrate UDP-GlcNAc adopts an unusual conformation where the non-bridging oxygen of the α-phosphate is brought in close proximity to the acceptor Ser or Thr, which were activated. This unique catalytic mechanism may lead to the specific inhibition of hOGT with a thio-derivative of the donor substrate UDP-5SGlcNAc (Gloster et al, 2011), by perturbing the trajectory of proton transfer during the deprotonation of the nucleophile acceptor Ser or Thr. Just like the substrate-assisted catalytic mechanism, this inhibition by UDP-5SGlcNAc seems to be unique for OGT. UDP-5SGlcNAc was able to inhibit *Dm*OGT activity with inhibition constant (K_i) of 36.2 µM, which is similar to the K_i of hOGT at 13.6 µM (Fig. 3.5B). This observation further supports the assumption that *Drosophila* OGT adopts the same catalytic mechanism as hOGT, in addition to the evidence provided by structural information and Michaelis-Menten kinetics.

4.3.5. Identification of catalytically inactivate *Dm*OGT point mutants

To investigate and differentiate the non-enzymatic functions of OGT from its catalytic activity during embryogenesis and early development, *Dm*OGT active site mutations with impaired catalytic activity were designed, based on the structure of *Dm*OGT and studies on hOGT (Clarke et al, 2008; Martinez-Fleites et al, 2008). The amino acid residues in the active site are highly conserved and especially the ones that have been found to interact with the donor

substrate UDP-GlcNAc and on the surface of the peptide-binding groove are almost identical. It has been found in the structure of hOGT that Asp925 (*Dm*OGT Asp955) interacts with the uridine and an alanine mutant has been shown to abolish hOGT enzymatic activity, primarily by disrupting UDP-GlcNAc binding (Clarke et al, 2008; Deplus et al, 2013). H558F (*Dm*OGT H596F) and K842M (*Dm*OGT K872M) mutants of hOGT have been described previously to be enzymatically inactive but still able to bind to the donor substrate UDP-GlcNAc with similar affinity to the wild type (Table 4.2). H498A mutation in hOGT (DmOGT H537A), on the other hand, has been shown to greatly reduce hOGT enzyme activity *in vitro* and was also selected. All of the aforementioned residues were selected to mutate in *Dm*OGT. *In vitro* O-GlcNAc transferase activity was determined using radiometric assay on the previously identified

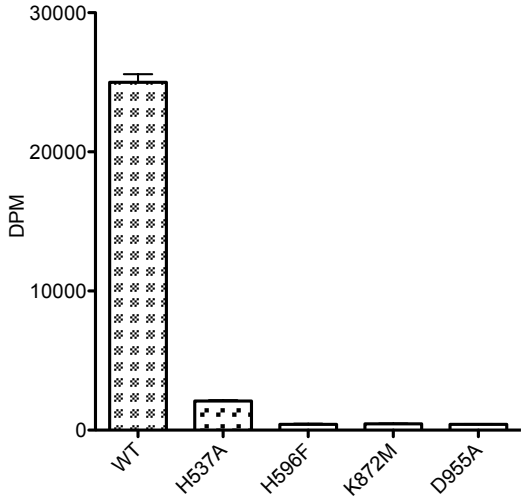


Figure 4.9: Selected mutations in *Dm*OGT abolish the enzymatic activity.
*Dm*OGT (353-end) wt and the mutants were expressed in *E. coli* and purified as described in the *Material and Methods*. The activity of each protein was determined on RBL2 peptide *in vitro* with radiometric assay as described in the *Material and Methods*.

*Dm*OGT substrate RL2 peptide. While the H537A mutation reduces the enzymatic activity of *Dm*OGT to less than 10%, all other three mutations, H596F, K872M and D955A have almost no detectable activity. These results of *Dm*OGT mutations are in agreement with it of the human enzyme tested previously with the same assay set up. Therefore, all of the four mutants were selected and used to generate transgenic flies.

4.3.6. *Dm*OGT catalytic activity is essential for *Drosophila* survival

To further address this question whether inactive *Dm*OGT could rescue the lethality produced by *ogt* null mutation, we have generated transgenic flies with *ogt*-null background and with the *sxc/ogt* gene of either the wt (*sxc/ogt*^{wt}) or abovementioned inactive mutants H537A (*sxc/ogt*^{H537F}), H569F (*sxc/ogt*^{H596F})

Table 4.1: Rescue with Drosophila *sxc/ogt* transgene

Parental genotypes	F1 genotype	Total	Expected	Rescued	% rescue
sxc^{1}/CyO; MKRS/TM6 X *sxc^{6}/Cyo; Tub-GAL4*/ TM6*	*sxc1/sxc6; MKRS/Tub-GAL4*	600	0	0	0
sxc^{1}/Cyo; ogt^{wt}/TM6 X *sxc^{6}/TM6; MKRS/TM6*	*sxc^{1}/sxc^{6}; ogt^{wt}/MKRS*	600	0	0	0
sxc^{1}/CyO; ogt^{wt}/ogt^{wt} X *sxc^{6}/CyO;Tub-GAL4/ TM6*	*sxc1/sxc6; ogt^{wt}/Tub-GAL4* (**OGT wt**)	600	75	16	21%
sxc^{1}/CyO; ogt^{H537A}/ ogt^{H537A} X *sxc^{6}/ CyO;Tub-GAL4/TM6*	*sxc1/sxc6; ogt^{H596F}/ Tub-GAL4* (**OGT H537A**)	600	75	15	20%
sxc^{1}/CyO; ogt^{H596F}/ ogt^{H596F} X *sxc^{6}/ CyO;Tub-GAL4/TM6*	*sxc1/sxc6; ogt^{H596F}/ Tub-GAL4* (**OGT H596F**)	600	75	12	16%
sxc^{1}/CyO; ogt^{K872M}/ ogt^{K872M} X *sxc^{6}/ CyO;Tub-GAL4/TM6*	*sxc1/sxc6; ogt^{K872M}/ Tub-GAL4* (**OGT K872M**)	600	75	1	1%
sxc^{1}/CyO; ogt^{D955A}/ ogt^{D955A} X *sxc^{6}/ CyO;Tub-GAL4/TM6*	*sxc1/sxc6; ogt^{D955A}/ Tub-GAL4* (**OGT D955A**)	600	75	17	23%

* *Tub-GAL4*: Mat-α-Tubulin-VP16 – GAL4

K872M (*sxc/ogt*^{K872M}) and D955A (*sxc/ogt*^{D955A}). The transgenes were integrated at the same position on the third chromosome and later crossed with transgenic flies carrying *sxc*^{1} transgene to generate flies with homozygotic genotype of *sxc*^{1}*/Cyo:ogt*^{x}*/ogt*^{x} (x represents wt or the inactive mutants), which would double the number of expected offspring carrying *ogt* transgene. The rescue experiment would be performed by ubiquitously driving the *sxc/ogt* transgenes with the tubulin-GAL4 (Mat-α-Tubulin-VP16 – GAL4) driver. The maternal tubulin-GAL4 driver was crossed into *sxc*^{6} transgenic flies to generate *sxc*^{6}*/Cyo:Tub-GAL4/TM6*. Female virgins of *sxc*^{1}*/Cyo:ogt*^{x}*/ogt*^{x} were collected to cross with male *sxc*^{6}*/Cyo:Tub-GAL4/TM6.* Flies were kept in 23 °C incubator and 600 embryos were collected and transferred onto fresh fly food and allow the embryos to hatch in room temperature. Hatched flies were transferred and numbers were taken. According to Mendelian genetics, 75 of the embryos should carry the expecting genetic trace, *sxc1/sxc6; ogt/Tub-GAL4.* The transgenic flies with *Dm*OGT wt could support the development to adulthood as rescue observed from the previous report (Sinclair et al, 2009). Rescue to adulthood by expressing wt *Dm*OGT ubiquitously was 21% (16 adults eclosed) as opposed to no recovery in control crosses lacking either the driver or the transgene. Interestingly, 15 embryos of total 20% with *sxc/ogt*^{H537A} transgene, 12 embryos of total 16% with *sxc/ogt*^{H596F} transgene, and 17 embryos of total 23% with *sxc/ogt*^{D955A} transgene were also recovered in *sxc*^{1}*/sxc*^{6} transheterozygotic adults background; only one embryo was collected for the transgenic flies with K872M mutation (Table 3.1). To probe the level of catalytic activity of *Dm*OGT transgenes in the rescued animals, except K872M

mutation, the total O-GlcNAcylation in the flies was determined by immunoblotting (Fig. 4.10). We detected minimal levels of O-GlcNAc in the offspring rescued by the *sxc/ogt*^{H596F} and *sxc/ogt*^{D955A} transgenes. However, flies offspring rescued with the *sxc/ogt*^{H537A} transgene retained total O-GlcNAcylation levels comparable with those rescued with *sxc/ogt*^{wt} transgene or wt flies as expected since the *Dm*OGT with H537A mutation retains reasonable activity *in vitro* as shown previously. In addition to the rescue of transgenic flies with *Dm*OGT H537A and H596F, they were also viable and

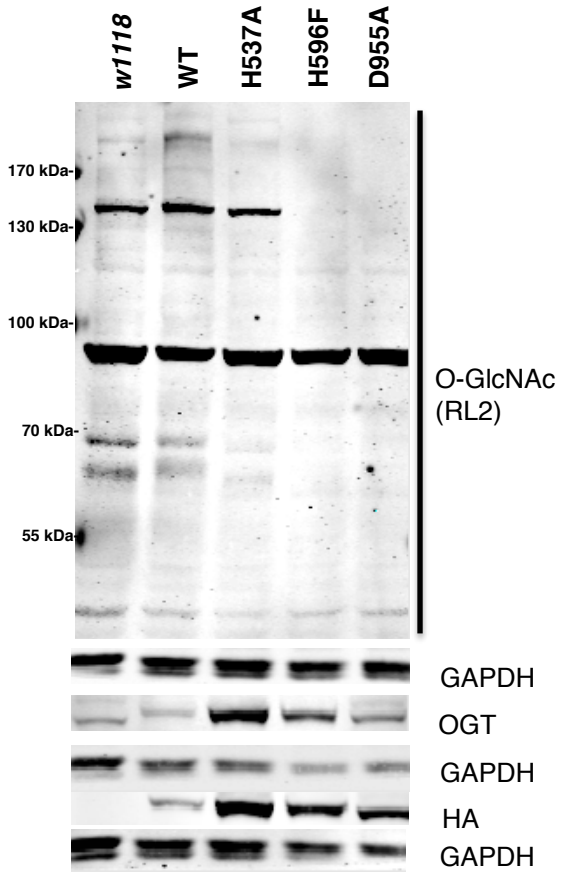


Figure 4.10: O-GlcNAcylation in trans-hetrozygotic flies
Rescued male adult flies collected and lysates were prepared from rescued male flies as detailed in Material and Method. Total O-GlcNAcylation was assessed by immunoblotting with anti-O-GlcNAc antibody (RL2). The total *Dm*OGT was probed by anti OGT antibody (H300) and the recombinant DmOGT was probed by anti-HA antibody. Loading was confirmed by GAPDH.

able to produce offspring. However, *Dm*OGT D955A transgenic flies struggled to produce offspring that survive to adult. There were not enough transgenic flies with K872M mutant to allow crossing. Some noticeable slow development was also observed in all of the above transgenic flies (Data not shown).

We suspected that the current *in vitro* assay set up might not be sensitive enough to completely conclude the activity of *Dm*OGT. To further confirm and confidently demonstrate the effect of DmOGT catalytic activity with the active site mutations, the reactions were performed with exceed amount of enzyme (15 times of the original enzyme concentration) for as long as 10 hours of reaction time (Figure 3.11). This would allow the signal accumulation of O-

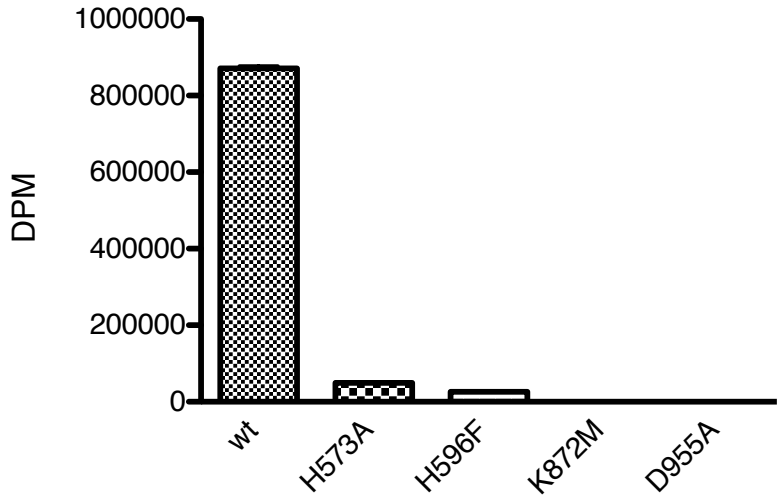


Figure 4.11: *In vitro* activity of *Dm*OGT wt and mutants at high enzyme concentration and longer reaction time.
*Dm*OGT (353-end) wt and the mutants were expressed in *E. coli* and purified as described in the *Material and Methods*. The activity of each protein was determined on RBL2 peptide *in vitro* with radiometric assay in the same conditions as described in the *Material and Methods*, except the reactions were performed with 1.5 µM *Dm*OGT and reaction for 10 hours. Signal for the wt enzyme was calculated from 1 nM enzyme concentration.

GlcNAc transfer and increase the signal to noise ratio. The reaction with DmOGT wt was performed with 1500 times less enzyme to ensure the reaction occurs in the linear phase; later the signal for the wt enzyme was converted back to the same enzyme concentration as for the mutants. Interestingly, we have detected signal for O-GlcNAc transfer with *Dm*OGT H596F, at less than 3% of the signal of wt enzyme, which was not detected previously. However, the signal for both K872M and D955A remain very low and the mutants were more than 11,000 times less active comparing with the wt enzyme. All the signal for the O-GlcNAc transfer with *Dm*OGT mutants are less than 5% of the total UDP-GlcNAc input and therefore the reactions would be considered to be in the linear phase. The results correlate with the rescue experiments that active *Dm*OGT H537A and H596F mutants, although with very low enzyme activity, were able to rescue the lethality phenotype as the wt enzyme does, while the inactive enzyme *Dm*OGT K872M was not, suggesting that the activity of *Dm*OGT is important for survival and development of flies. With the current detection method, *Dm*OGT D955A mutant has also been shown to be 11,000 times less active than the wt enzyme, however, the were able to rescue the lethality by *ogt* null mutation, which differs from the outcome produced by K872M mutant. Interestingly, these transgenic flies struggle to produce offspring that survive till adult stage, indicating the activity of *Dm*OGT and O-GlcNAc indeed play some roles in survival and development.

sxc/ogt is a PcG gene that functions in homeotic gene repressing; loss function mutations of *sxc/ogt* cause severe homeotic transformation and lethality

(Ingham, 1984; Sinclair et al., 2009 and Gambetta et al., 2009). Recent genetic analysis with Chip-ChIP showed O-GlcNAc signal is enriched at the Polycomb Response Elements (PREs) in the middle of the promoter region supporting OGT functions in gene regulations (Sinclair et al., 2009 and Gambetta et al., 2009). It has also been shown that in *sxc/ogt* null fly larvae, a 1.5-2 folds reduction of PcG protein polyhomeotic (Ph) PRE localization and downstream gene regulation (Gambetta et al, 2009). Genetic studies in *C elegans* with the same approach further suggested that *sxc/ogt* also involve in insulin signalling pathway, energy metabolism and stress through gene regulation (Love et al., 2010). OGT has been found directly interacts with many transcription factors, such as Sin3A, Tet1, Tet2, HCF1, by which OGT might be recruited to promoters, many of which have suggested the glycosyltransferase activity of OGT is required for these interactions (Chen et al, 2013; Hassig et al, 1997; Vella et al, 2013; Williams et al, 2012; Wysocka et al, 2003).

Attempts at rescuing the lethality with the TPR domain only in *Arabidopsis* caused by ablation of both OGT orthologs, Spindly and Secret agent, proved unsuccessful. It indicates that OGT without the catalytic domain is not sufficient for the embryogenesis (Tseng et al, 2001). On one hand, the catalytic domain of OGT has been shown to participate in protein-protein interactions. It has been demonstrated that O-GlcNAcylation on neurofilaments is mediated by MAPK, which interacts with OGT though the catalytic domain (Cheung & Hart, 2008). Similarly, the interaction between OGT and Sin3A involves Sp1-dependent transcription repression, and the catalytic domain of OGT (471-end)

lacking the entire TPR domain has been shown to repress transcription (Yang et al, 2002). Similar construct or with more TPRs removed has been shown to inactivate hOGT (unpublished). It therefore appears that the transcription repression occurs through the interaction of Sin3A with the catalytic domain of OGT independent of OGT enzymatic activity. On the other hand, it has also been shown that inhibition of OGT using an ambiguous OGT inhibitor alloxan decreases OGT-HCF1 interaction (Deplus et al., 2013). An elegant approach was taken to investigate the interactions with OGT using a catalytically inactive enzyme H558A (H568A in the original publication); hOGT H558A has been found to destabilize the interaction with Tet1 and also causes protein degradation (Shi et al, 2013).

In our rescue experiments in *Drosophila*, we have generated a structurally intact ncOGT that lacks the catalytic activity. We demonstrated that K872M mutation of *Dm*OGT could dramatically reduced its enzyme activity by 11,000 folds *in vitro* on a peptide substrate and did not rescue the lethality phenotype in *ogt* null transgenic flies. Transgenic flies with another inactive mutant D955A, however, could rescue the lethality but struggle to produce healthy adult offspring. These findings suggest OGT activity and O-GlcNAcylation are required for survival and development of *Drosophila*. We have also identified another mutant H596F, which severely decreased or almost abolished O-GlcNAcylation in the transgenic flies. However, giving the limitations of *in vitro* assays, it does not completely reflect the activity of the enzyme in a long period of time as it has been shown in the longer *in vitro* reaction as well as in

complicated biological systems. Secondly, there might also be a detection limit with the antibody used to detect O-GlcNAc levels in flies. Nevertheless, we were able to demonstrate a dramatic reduction in the overall O-GlcNAc signal in the transgenic flies with active site mutants.

So far, *C. elegans* is the only model organism reported to survive to adulthood with *ogt* null genotype, this has been used in studies of the involvement of O-GlcNAcylation in many biological or pathological conditions, especially the involvement of O-GlcNAcylation in energy metabolism and diabetes (Forsythe et al, 2006; Hanover et al, 2005; Love et al, 2010). However, since *sxc* mutations are lethal, *Drosophila* with abolished or even reduced O-GlcNAc modification would be a good model in addition to *C elegans* for elucidating the details of the physiological functions of OGT. *sxc/ogt* is a Polycomb group gene that has been found to be involved in development of the body axis not only in *Drosophila,* but also in vertebrates (Schwartz & Pirrotta, 2007). It would be interesting to investigate the involvement of O-GlcNAcylation during development, especially in the nervous system where OGT is highly expressed, using these O-GlcNAc null flies.

4.4. Material and Methods

Expression and purification of DmOGT and mutant

*Dm*OGT (353-end) was expressed in ArcticExpress *E. coli* and purified as described for hOGT (314-end) described in Chapter 1.

Crystallisation

Protein used in crystallisation was stored in 25 mM Tris-Cl, pH=7.5, 150 mM NaCl and 1 mM DTT). Protein was crystallized in complex with UDP-5SGlcNAc. Vapor diffusion crystallization experiments with sitting drops containing 0.5 µl protein (at 5 mg/ml in a buffer of 10 mM Tris-Cl pH=7.5, 50 mM NaCl, 1 mM DTT, 1 mM UDP-5S-GlcNAc) and 0.5 µl reservoir solution (12% PEG8000, 16% glycerol and 0.1 M HEPES pH=7.5) gave needle-shaped crystals with approximately 0.3 x 0.05 mm after 2-3 days at 21°C. The crystal was flash-frozen without further cryoprotection as the reservoir solution freezes clear. Data were collected at the Diamond Light Source, UK, at a wavelength of 0.92 Å on beamlines IO4-1, respectively. Crystals belonged to space group H3 and contained 1 molecule per asymmetric unit. The structure was solved by molecular replacement using the A chain of PDB ID 3PE3 as the search model. Initial structure refinement was performed with Refmac5.

Peptide screen

The screen was performed by Dr. David Blair. A library of 720 biotinylated peptide kinase substrates (JPT Peptide Technologies GmbH) consisting of two 384 polypropylene plates each containing 360 peptides (rows A1 to O24) at

concentrations of 0.25 nmols each was solubilised in 62.5 μl of 50 mM Tris (pH 7.5) to give a final concentration of 4 μM using a PlateMate plus liquid handling robot (Thermo Scientific Matrix). A solution of either control (No protein) or *Dm*OGT protein (0.5 μM) and UDP-GlcNAc substrate stock (1 μM), consisting of 910 nM cold UDP-GlcNAc (Sigma) and 90 nM UDP-[^{3}H]GlcNAc radioactive tracer (0.4 Ci/mmol)(American Radiolabeled Chemicals, Inc.) was made up in buffer consisting of 50 mM Tris pH=7.5, 2 mM DTT and 0.1 mg/ml BSA. 10 μl of 4 μM peptide was added to all 384 polypropylene U bottom plates using a PlateMate plus liquid handling robot (Thermo Scientific Matrix). The reaction was then initiated by the addition of 10μl of control (No OGT)/ UDP-GlcNAc to both QC plate (columns 9-16) and assay plates P13 - P24 to provide low controls. All other positions received OGT / UDP-GlcNAc. Plates were incubated at room temperature on a micro titre plate shaker (Heidolph) for 240 minutes.The reaction was stopped by the addition of 40 μl of 100 μM phosphoric acid (pH 4) / 750 μM $MgCl_2$ using a Wellmate liquid handling robot (Thermo Scientific Matrix). The stopped reactions were incubated for a further 10 min at room temperature. 56 μl from the QC and screening plates was transferred to 384 well streptavidin FlashPlates (PerkinElmer) using a PlateMate plus liquid handling robot. Plates were then sealed and incubated over night at room temperature before determining scintillation counts using a Top Count plate reader (PerkinElmer).

Enzyme kinetics (K_m and IC50 and K_i determination) on hOGT and DmOGT

The *in vitro* reactions were performed on RBL2 derived peptide (KENSPCVTPVSTA) identified from the peptide screen using radiometric assay and the signal was detected by scintillation counter. Assays were conducted in the final reaction volume of 20 μl containing 0.05 μM of hOGT (314-end) or 0.1 μM *Dm*OGT (353-end), 0.25 mM RBL2 peptides and UDP-GlcNAc (100 μM, 50 μM, 20 μM, 5 μM, 2 μM, 0.5 μM and 0.05 μM) with 0.3 Ci/mmol UDP-[^{3}H]-GlcNAc as a radioactive tracer, in the buffer containing 100 mM Tris-Cl, pH=7.5, 150 mM NaCl and 1 mM DTT. The reaction was stopped by addition of 200 μl of 0.75 M ice-cold phosphoric acid containing final concentration of 0.5% TFA and 5% ACN. C18 column was used for the binding of peptide. C18 column was activated by 200 μl 100% methanol and equilibrated twice with 200 μl of equilibration buffer containing 0.5% TFA and 5% ACN. Samples were passed through the C18 column by centrifugation at 500 rpm on the table-top centrifuge for 30 seconds and washed 6x with 400 μl of equilibration buffer. Peptide bound to the column was eluted by the addition of 100 μl 100% methanol followed by evaporation with SpeedVac. Radioactivity was detected by the addition of 4 ml scintillation fluid and signal was read by scintillation counter.

Half maximal inhibitory concentration (IC50) for the UDP-5SGlcNAc was determined with the same experimental set up described for the K_m determination. The reactions were conducted at the fixed concentration of UDP-[^{3}H]GlcNAc (0.3 Ci/mmol) with 5 μM for hOGT and 15 μM for *Dm*OGT, both are close to the K_m of UDP-GlcNAc for both proteins (6.6 μM for hOGT

and 17.8 μM for *Dm*OGT). The inhibitory constant (K_i) of UDP-5SGlcNAc for both proteins were calculated by Cheng-Prusoff equation:

$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_m}}$$

Enzyme activity of DmOGT wt and mutants

DmOGT wt and mutants were expressed and purified as described above. Assays were conducted in the final reaction volume of 30 μl containing 1 μM *Dm*OGT (353-end), 0.25 mM RBL2 peptides and UDP-GlcNAc at 100 μM with 0.3 Ci/mmol UDP-[^{3}H]-GlcNAc as a radioactive tracer, in the buffer containing 100 mM Tris-Cl, pH=7.5, 150 mM NaCl and 1 mM DTT, at 37 °C for 90 min on a shaker. The reaction was stopped by addition of 200 μl of 0.75 M ice-cold phosphoric acid containing final concentration of 0.5% TFA and 5% ACN and the peptide were purified with C18 column method described above and signals were obtained by scintillation counter.

O-GlcNAcylation on DmOGT by Mass-spectrometry

The protocol is same to the O-GlcNAc site determination on Yes1 described in Chapter 1.

Drosophila genetics

W^{1118}, *sxc1/CyO*, *sxc6/CyO* fly stocks were used

UAS::sxc/ogt of wt and the mutants transgenic flies were generated by integrating the genes into 86Fb site on the third chromosome and transgenic flies were generated by Rainbow Transgenic Flies, Inc, California.

Fly lysis and western blotting

5 adult flies were anesthetized and frozen on dry ice. The frozen flies were homogenized in 100 μl of lysis buffer containing 50 mM Tris-HCl, pH=8.0, 150 mM NaCl, 1% Triton-X-100, 1 μM GlcNAcstatin C, 5 mM NaF, 2 mM Na othovanadate, 1 mM benzamidine, 0.2 mM PMSF, 5 μM leupeptin and 1 mM DTT) and left on ice for 60 min for protein extraction. The lysates were centrifuged at 16000 g for 10 min and supernatant was collected before Bradford assay to determine the protein concentrations. 10 μg of the crude lysate was subject to SDS-PAGE and transferred onto nitrocellulose membrane before immunoblotting with corresponding antibodies. Signal was detected by Odyssey® Li-Cor infrared imaging system.

Reference:

ALLAN, R. K. & RATAJCZAK, T. 2011. Versatile TPR domains accommodate different modes of target protein recognition and function. *Cell stress & chaperones,* 16**,** 353-67.

ANDRALI, S. S., MARZ, P. & OZCAN, S. 2005. Ataxin-10 interacts with O-GlcNAc transferase OGT in pancreatic beta cells. *Biochemical and biophysical research communications,* 337**,** 149-53.

APWEILER, R., HERMJAKOB, H. & SHARON, N. 1999. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et biophysica acta,* 1473**,** 4-8.

ARNOLD, C. S., JOHNSON, G. V., COLE, R. N., DONG, D. L., LEE, M. & HART, G. W. 1996. The microtubule-associated protein tau is extensively modified with O-linked N-acetylglucosamine. *The Journal of biological chemistry,* 271**,** 28741-4.

BALL, L. E., BERKAW, M. N. & BUSE, M. G. 2006. Identification of the major site of O-linked beta-N-acetylglucosamine modification in the C terminus of insulin receptor substrate-1. *Molecular & cellular proteomics : MCP,* 5**,** 313-23.

BANERJEE, S., ROBBINS, P. W. & SAMUELSON, J. 2009. Molecular characterization of nucleocytosolic O-GlcNAc transferases of Giardia lamblia and Cryptosporidium parvum. *Glycobiology,* 19**,** 331-6.

BELLACCHIO, E. & PAGGI, M. G. 2013. Understanding the targeting of the RB family proteins by viral oncoproteins to defeat their oncogenic machinery. *Journal of cellular physiology,* 228**,** 285-91.

BOGGON, T. J. & ECK, M. J. 2004. Structure and regulation of Src family kinases. *Oncogene,* 23**,** 7918-27.

BOLAM, D. N., ROBERTS, S., PROCTOR, M. R., TURKENBURG, J. P., DODSON, E. J., MARTINEZ-FLEITES, C., YANG, M., DAVIS, B. G., DAVIES, G. J. & GILBERT, H. J. 2007. The crystal structure of two macrolide glycosyltransferases provides a blueprint for host cell antibiotic immunity. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 5336-41.

BRAIDMAN, I., CARROLL, M., DANCE, N. & ROBINSON, D. 1974. Separation and properties of human brain hexosaminidase C. *The Biochemical journal,* 143**,** 295-301.

BRAZIER-HICKS, M., OFFEN, W. A., GERSHATER, M. C., REVETT, T. J., LIM, E. K., BOWLES, D. J., DAVIES, G. J. & EDWARDS, R. 2007. Characterization and engineering of the bifunctional N- and O-glucosyltransferase involved in xenobiotic metabolism in plants.

Proceedings of the National Academy of Sciences of the United States of America, 104**,** 20238-43.

BRICH, J., SHIE, F. S., HOWELL, B. W., LI, R., TUS, K., WAKELAND, E. K., JIN, L. W., MUMBY, M., CHURCHILL, G., HERZ, J. & COOPER, J. A. 2003. Genetic modulation of tau phosphorylation in the mouse. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 23**,** 187-92.

BUTKINAREE, C., CHEUNG, W. D., PARK, S., PARK, K., BARBER, M. & HART, G. W. 2008. Characterization of beta-N-acetylglucosaminidase cleavage by caspase-3 during apoptosis. *The Journal of biological chemistry,* 283**,** 23557-66.

CAMPBELL, J. A., DAVIES, G. J., BULONE, V. & HENRISSAT, B. 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *The Biochemical journal,* 326 (Pt 3)**,** 929-39.

CAPUTTO, R., LELOIR, L. F., CARDINI, C. E. & PALADINI, A. C. 1950. Isolation of the coenzyme of the galactose phosphate-glucose phosphate transformation. *The Journal of biological chemistry,* 184**,** 333-50.

CHAMPATTANACHAI, V., MARCHASE, R. B. & CHATHAM, J. C. 2007. Glucosamine protects neonatal cardiomyocytes from ischemia-reperfusion injury via increased protein-associated O-GlcNAc. *American journal of physiology. Cell physiology,* 292**,** C178-87.

CHAMPATTANACHAI, V., MARCHASE, R. B. & CHATHAM, J. C. 2008. Glucosamine protects neonatal cardiomyocytes from ischemia-reperfusion injury via increased protein O-GlcNAc and increased mitochondrial Bcl-2. *American journal of physiology. Cell physiology,* 294**,** C1509-20.

CHARNOCK, S. J. & DAVIES, G. J. 1999. Structure of the nucleotide-diphospho-sugar transferase, SpsA from Bacillus subtilis, in native and nucleotide-complexed forms. *Biochemistry,* 38**,** 6380-5.

CHARNOCK, S. J. & DAVIES, G. J. 1999. Cloning, crystallization and preliminary X-ray analysis of a nucleotide-diphospho-sugar transferase spsA from Bacillus subtilis. *Acta crystallographica. Section D, Biological crystallography,* 55**,** 677-8.

CHEN, C. I., KEUSCH, J. J., KLEIN, D., HESS, D., HOFSTEENGE, J. & GUT, H. 2012. Structure of human POFUT2: insights into thrombospondin type 1 repeat fold and O-fucosylation. *The EMBO journal,* 31**,** 3183-97.

CHEN, Q., CHEN, Y., BIAN, C., FUJIKI, R. & YU, X. 2013. TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature,* 493**,** 561-4.

CHEN, Y. X., DU, J. T., ZHOU, L. X., LIU, X. H., ZHAO, Y. F., NAKANISHI, H. &

LI, Y. M. 2006. Alternative O-GlcNAcylation/O-phosphorylation of Ser16 induce different conformational disturbances to the N terminus of murine estrogen receptor beta. *Chemistry & biology,* 13**,** 937-44.

CHENG, X. & HART, G. W. 2001. Alternative O-glycosylation/O-phosphorylation of serine-16 in murine estrogen receptor beta: post-translational regulation of turnover and transactivation activity. *The Journal of biological chemistry,* 276**,** 10570-5.

CHEUNG, W. D. & HART, G. W. 2008. AMP-activated protein kinase and p38 MAPK activate O-GlcNAcylation of neuronal proteins during glucose deprivation. *The Journal of biological chemistry,* 283**,** 13009-20.

CHEUNG, W. D., SAKABE, K., HOUSLEY, M. P., DIAS, W. B. & HART, G. W. 2008. O-linked beta-N-acetylglucosaminyltransferase substrate specificity is regulated by myosin phosphatase targeting and other interacting proteins. *The Journal of biological chemistry,* 283**,** 33935-41.

CHO, Y., GORINA, S., JEFFREY, P. D. & PAVLETICH, N. P. 1994. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science,* 265**,** 346-55.

CHOU, C. F., SMITH, A. J. & OMARY, M. B. 1992. Characterization and dynamics of O-linked glycosylation of human cytokeratin 8 and 18. *The Journal of biological chemistry,* 267**,** 3901-6.

CHOU, T. Y., DANG, C. V. & HART, G. W. 1995. Glycosylation of the c-Myc transactivation domain. *Proceedings of the National Academy of Sciences of the United States of America,* 92**,** 4417-21.

CHOU, T. Y., DANG, C. V. & HART, G. W. 1995. Glycosylation of the c-Myc transactivation domain. *Proceedings of the National Academy of Sciences of the United States of America,* 92**,** 4417-21.

CHOU, T. Y., HART, G. W. & DANG, C. V. 1995. c-Myc is glycosylated at threonine 58, a known phosphorylation site and a mutational hot spot in lymphomas. *The Journal of biological chemistry,* 270**,** 18961-5.

CHOU, T. Y., HART, G. W. & DANG, C. V. 1995. c-Myc is glycosylated at threonine 58, a known phosphorylation site and a mutational hot spot in lymphomas. *The Journal of biological chemistry,* 270**,** 18961-5.

CLARKE, A. J., HURTADO-GUERRERO, R., PATHAK, S., SCHUTTELKOPF, A. W., BORODKIN, V., SHEPHERD, S. M., IBRAHIM, A. F. & VAN AALTEN, D. M. 2008. Structural insights into mechanism and specificity of O-GlcNAc transferase. *The EMBO journal,* 27**,** 2780-8.

COBLITZ, B., WU, M., SHIKANO, S. & LI, M. 2006. C-terminal binding: an expanded repertoire and function of 14-3-3 proteins. *FEBS letters,* 580**,** 1531-5.

COLE, R. N. & HART, G. W. 1999. Glycosylation sites flank phosphorylation sites on synapsin I: O-linked N-acetylglucosamine residues are localized within domains mediating synapsin I interactions. *Journal of neurochemistry,* 73**,** 418-28.

COLE, R. N. & HART, G. W. 2001. Cytosolic O-glycosylation is abundant in nerve terminals. *Journal of neurochemistry,* 79**,** 1080-9.

COMER, F. I. & HART, G. W. 1999. O-GlcNAc and the control of gene expression. *Biochimica et biophysica acta,* 1473**,** 161-71.

CONTI, E. & KURIYAN, J. 2000. Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure,* 8**,** 329-38.

CONTI, E., MULLER, C. W. & STEWART, M. 2006. Karyopherin flexibility in nucleocytoplasmic transport. *Current opinion in structural biology,* 16**,** 237-44.

COUTINHO, P. M., DELEURY, E., DAVIES, G. J. & HENRISSAT, B. 2003. An evolving hierarchical family classification for glycosyltransferases. *Journal of molecular biology,* 328**,** 307-17.

D'ANDREA, L. D. & REGAN, L. 2003. TPR proteins: the versatile helix. *Trends in biochemical sciences,* 28**,** 655-62.

DAHMUS, M. E. 1996. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *The Journal of biological chemistry,* 271**,** 19009-12.

DAHMUS, M. E. 1996. Phosphorylation of mammalian RNA polymerase II. *Methods in enzymology,* 273**,** 185-93.

DEHENNAUT, V., HANOULLE, X., BODART, J. F., VILAIN, J. P., MICHALSKI, J. C., LANDRIEU, I., LIPPENS, G. & LEFEBVRE, T. 2008. Microinjection of recombinant O-GlcNAc transferase potentiates Xenopus oocytes M-phase entry. *Biochemical and biophysical research communications,* 369**,** 539-46.

DEHENNAUT, V., LEFEBVRE, T., LEROY, Y., VILAIN, J. P., MICHALSKI, J. C. & BODART, J. F. 2009. Survey of O-GlcNAc level variations in Xenopus laevis from oogenesis to early development. *Glycoconjugate journal,* 26**,** 301-11.

DENNIS, B., AZIZ, K., SHE, L., FARUQUI, A. M., DAVIS, C. E., MANOLIO, T. A., BURKE, G. L. & AZIZ, S. 2006. High rates of obesity and cardiovascular disease risk factors in lower middle class community in Pakistan: the Metroville Health Study. *JPMA. The Journal of the Pakistan Medical Association,* 56**,** 267-72.

DEPLUS, R., DELATTE, B., SCHWINN, M. K., DEFRANCE, M., MENDEZ, J.,

MURPHY, N., DAWSON, M. A., VOLKMAR, M., PUTMANS, P., CALONNE, E., SHIH, A. H., LEVINE, R. L., BERNARD, O., MERCHER, T., SOLARY, E., URH, M., DANIELS, D. L. & FUKS, F. 2013. TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *The EMBO journal*.

DEPLUS, R., DELATTE, B., SCHWINN, M. K., DEFRANCE, M., MENDEZ, J., MURPHY, N., DAWSON, M. A., VOLKMAR, M., PUTMANS, P., CALONNE, E., SHIH, A. H., LEVINE, R. L., BERNARD, O., MERCHER, T., SOLARY, E., URH, M., DANIELS, D. L. & FUKS, F. 2013. TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *The EMBO journal,* 32**,** 645-55.

DIAS, W. B. & HART, G. W. 2007. O-GlcNAc modification in diabetes and Alzheimer's disease. *Molecular bioSystems,* 3**,** 766-72.

DONG, D. L. & HART, G. W. 1994. Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol. *The Journal of biological chemistry,* 269**,** 19321-30.

DONG, D. L., XU, Z. S., CHEVRIER, M. R., COTTER, R. J., CLEVELAND, D. W. & HART, G. W. 1993. Glycosylation of mammalian neurofilaments. Localization of multiple O-linked N-acetylglucosamine moieties on neurofilament polypeptides L and M. *The Journal of biological chemistry,* 268**,** 16679-87.

DONG, D. L., XU, Z. S., HART, G. W. & CLEVELAND, D. W. 1996. Cytoplasmic O-GlcNAc modification of the head domain and the KSP repeat motif of the neurofilament protein neurofilament-H. *The Journal of biological chemistry,* 271**,** 20845-52.

DORFMUELLER, H. C., BORODKIN, V. S., BLAIR, D. E., PATHAK, S., NAVRATILOVA, I. & VAN AALTEN, D. M. 2011. Substrate and product analogues as human O-GlcNAc transferase inhibitors. *Amino acids,* 40**,** 781-92.

DORFMUELLER, H. C., BORODKIN, V. S., SCHIMPL, M. & VAN AALTEN, D. M. 2009. GlcNAcstatins are nanomolar inhibitors of human O-GlcNAcase inducing cellular hyper-O-GlcNAcylation. *The Biochemical journal,* 420**,** 221-7.

DU, X. L., EDELSTEIN, D., ROSSETTI, L., FANTUS, I. G., GOLDBERG, H., ZIYADEH, F., WU, J. & BROWNLEE, M. 2000. Hyperglycemia-induced mitochondrial superoxide overproduction activates the hexosamine pathway and induces plasminogen activator inhibitor-1 expression by increasing Sp1 glycosylation. *Proceedings of the National Academy of Sciences of the United States of America,* 97**,** 12222-6.

FANG, B. & MILLER, M. W. 2001. Use of galactosyltransferase to assess the

biological function of O-linked N-acetyl-d-glucosamine: a potential role for O-GlcNAc during cell division. *Experimental cell research,* 263**,** 243-53.

FORSYTHE, M. E., LOVE, D. C., LAZARUS, B. D., KIM, E. J., PRINZ, W. A., ASHWELL, G., KRAUSE, M. W. & HANOVER, J. A. 2006. Caenorhabditis elegans ortholog of a diabetes susceptibility locus: oga-1 (O-GlcNAcase) knockout impacts O-GlcNAc cycling, metabolism, and dauer. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 11952-7.

FUJIKI, R., HASHIBA, W., SEKINE, H., YOKOYAMA, A., CHIKANISHI, T., ITO, S., IMAI, Y., KIM, J., HE, H. H., IGARASHI, K., KANNO, J., OHTAKE, F., KITAGAWA, H., ROEDER, R. G., BROWN, M. & KATO, S. 2011. GlcNAcylation of histone H2B facilitates its monoubiquitination. *Nature,* 480**,** 557-60.

GAMBETTA, M. C., OKTABA, K. & MULLER, J. 2009. Essential role of the glycosyltransferase sxc/Ogt in polycomb repression. *Science,* 325**,** 93-6.

GAO, Y., WELLS, L., COMER, F. I., PARKER, G. J. & HART, G. W. 2001. Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic beta-N-acetylglucosaminidase from human brain. *The Journal of biological chemistry,* 276**,** 9838-45.

GARINOT-SCHNEIDER, C., LELLOUCH, A. C. & GEREMIA, R. A. 2000. Identification of essential amino acid residues in the Sinorhizobium meliloti glucosyltransferase ExoM. *The Journal of biological chemistry,* 275**,** 31407-13.

GAWANTKA, V., POLLET, N., DELIUS, H., VINGRON, M., PFISTER, R., NITSCH, R., BLUMENSTOCK, C. & NIEHRS, C. 1998. Gene expression screening in Xenopus identifies molecular pathways, predicts gene function and provides a global view of embryonic patterning. *Mechanisms of development,* 77**,** 95-141.

GLOSTER, T. M., ZANDBERG, W. F., HEINONEN, J. E., SHEN, D. L., DENG, L. & VOCADLO, D. J. 2011. Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells. *Nature chemical biology,* 7**,** 174-81.

GOLDBERG, H. J., SCHOLEY, J. & FANTUS, I. G. 2000. Glucosamine activates the plasminogen activator inhibitor 1 gene promoter through Sp1 DNA binding sites in glomerular mesangial cells. *Diabetes,* 49**,** 863-71.

GRIFFITH, L. S., MATHES, M. & SCHMITZ, B. 1995. Beta-amyloid precursor protein is modified with O-linked N-acetylglucosamine. *Journal of neuroscience research,* 41**,** 270-8.

GROSS, B. J., KRAYBILL, B. C. & WALKER, S. 2005. Discovery of O-GlcNAc

transferase inhibitors. *Journal of the American Chemical Society,* 127**,** 14588-9.

GROVES, M. R., HANLON, N., TUROWSKI, P., HEMMINGS, B. A. & BARFORD, D. 1999. The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell,* 96**,** 99-110.

HALTIWANGER, R. S., BLOMBERG, M. A. & HART, G. W. 1992. Glycosylation of nuclear and cytoplasmic proteins. Purification and characterization of a uridine diphospho-N-acetylglucosamine:polypeptide beta-N-acetylglucosaminyltransferase. *The Journal of biological chemistry,* 267**,** 9005-13.

HALTIWANGER, R. S., HOLT, G. D. & HART, G. W. 1990. Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-acetylglucosamine:peptide beta-N-acetylglucosaminyltransferase. *The Journal of biological chemistry,* 265**,** 2563-8.

HANOVER, J. A., FORSYTHE, M. E., HENNESSEY, P. T., BRODIGAN, T. M., LOVE, D. C., ASHWELL, G. & KRAUSE, M. 2005. A Caenorhabditis elegans model of insulin resistance: altered macronutrient storage and dauer formation in an OGT-1 knockout. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 11266-71.

HANTSCHEL, O., NAGAR, B., GUETTLER, S., KRETZSCHMAR, J., DOREY, K., KURIYAN, J. & SUPERTI-FURGA, G. 2003. A myristoyl/phosphotyrosine switch regulates c-Abl. *Cell,* 112**,** 845-57.

HART, G. W., HOUSLEY, M. P. & SLAWSON, C. 2007. Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins. *Nature,* 446**,** 1017-22.

HART, G. W., SLAWSON, C., RAMIREZ-CORREA, G. & LAGERLOF, O. 2011. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annual review of biochemistry,* 80**,** 825-58.

HARTWECK, L. M., SCOTT, C. L. & OLSZEWSKI, N. E. 2002. Two O-linked N-acetylglucosamine transferase genes of Arabidopsis thaliana L. Heynh. have overlapping functions necessary for gamete and seed development. *Genetics,* 161**,** 1279-91.

HASSIG, C. A., FLEISCHER, T. C., BILLIN, A. N., SCHREIBER, S. L. & AYER, D. E. 1997. Histone deacetylase activity is required for full transcriptional repression by mSin3A. *Cell,* 89**,** 341-7.

HECKEL, D., COMTESSE, N., BRASS, N., BLIN, N., ZANG, K. D. & MEESE, E.

1998. Novel immunogenic antigen homologous to hyaluronidase in meningioma. *Human molecular genetics,* 7**,** 1859-72.

HIROMURA, M., CHOI, C. H., SABOURIN, N. A., JONES, H., BACHVAROV, D. & USHEVA, A. 2003. YY1 is regulated by O-linked N-acetylglucosaminylation (O-glcNAcylation). *The Journal of biological chemistry,* 278**,** 14046-52.

HOANG, A. T., LUTTERBACH, B., LEWIS, B. C., YANO, T., CHOU, T. Y., BARRETT, J. F., RAFFELD, M., HANN, S. R. & DANG, C. V. 1995. A link between increased transforming activity of lymphoma-derived MYC mutant alleles, their defective regulation by p107, and altered phosphorylation of the c-Myc transactivation domain. *Molecular and cellular biology,* 15**,** 4031-42.

HOLT, G. D. & HART, G. W. 1986. The subcellular distribution of terminal N-acetylglucosamine moieties. Localization of a novel protein-saccharide linkage, O-linked GlcNAc. *The Journal of biological chemistry,* 261**,** 8049-57.

HOSFIELD, D. J., ZHANG, Y., DOUGAN, D. R., BROUN, A., TARI, L. W., SWANSON, R. V. & FINN, J. 2004. Structural basis for bisphosphonate-mediated inhibition of isoprenoid biosynthesis. *The Journal of biological chemistry,* 279**,** 8526-9.

HOUSLEY, M. P., RODGERS, J. T., UDESHI, N. D., KELLY, T. J., SHABANOWITZ, J., HUNT, D. F., PUIGSERVER, P. & HART, G. W. 2008. O-GlcNAc regulates FoxO activation in response to glucose. *The Journal of biological chemistry,* 283**,** 16283-92.

HU, Y., BELKE, D., SUAREZ, J., SWANSON, E., CLARK, R., HOSHIJIMA, M. & DILLMANN, W. H. 2005. Adenovirus-mediated overexpression of O-GlcNAcase improves contractile function in the diabetic heart. *Circulation research,* 96**,** 1006-13.

HUANG, F., LI, W., ZHANG, B., CUI, X., HAN, Z., FANG, Z., CAI, S., YIN, L. & WANG, L. 2001. Effects of free radicals and amyloid beta protein on the currents of expressed rat receptors in Xenopus oocytes. *Chinese medical journal,* 114**,** 244-7.

HUANG, F. N., ZHANG, B. L., LI, W. B., CUI, X., HAN, Z. T. & FANG, Z. Y. 2001. [Effects of free radicals and amyloid beta protein on the currents of expressed rat receptors in Xenopus oocytes]. *Zhongguo ying yong sheng li xue za zhi = Zhongguo yingyong shenglixue zazhi = Chinese journal of applied physiology,* 17**,** 109-12.

INGHAM, P. W. 1984. A gene that regulates the bithorax complex differentially in larval and adult cells of Drosophila. *Cell,* 37**,** 815-23.

IYER, S. P. & HART, G. W. 2003. Roles of the tetratricopeptide repeat domain in O-GlcNAc transferase targeting and protein substrate specificity. *The Journal of biological chemistry,* 278**,** 24608-16.

JACOBSEN, S. E. & OLSZEWSKI, N. E. 1993. Mutations at the SPINDLY locus of Arabidopsis alter gibberellin signal transduction. *The Plant cell,* 5**,** 887-96.

JIANG, J., LAZARUS, M. B., PASQUINA, L., SLIZ, P. & WALKER, S. 2012. A neutral diphosphate mimic crosslinks the active site of human O-GlcNAc transferase. *Nature chemical biology,* 8**,** 72-7.

JINEK, M., REHWINKEL, J., LAZARUS, B. D., IZAURRALDE, E., HANOVER, J. A. & CONTI, E. 2004. The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha. *Nature structural & molecular biology,* 11**,** 1001-7.

KELLY, S. E. & CARTWRIGHT, I. L. 1989. Perturbation of chromatin architecture on ecdysterone induction of Drosophila melanogaster small heat shock protein genes. *Molecular and cellular biology,* 9**,** 332-5.

KELLY, S. E. & CARTWRIGHT, I. L. 1989. Perturbation of chromatin architecture on ecdysterone induction of Drosophila melanogaster small heat shock protein genes. *Molecular and cellular biology,* 9**,** 332-5.

KELLY, W. G. & HART, G. W. 1989. Glycosylation of chromosomal proteins: localization of O-linked N-acetylglucosamine in Drosophila chromatin. *Cell,* 57**,** 243-51.

KELLY, W. G. & HART, G. W. 1989. Glycosylation of chromosomal proteins: localization of O-linked N-acetylglucosamine in Drosophila chromatin. *Cell,* 57**,** 243-51.

KENWRICK, S., AMAYA, E. & PAPALOPULU, N. 2004. Pilot morpholino screen in Xenopus tropicalis identifies a novel gene involved in head development. *Developmental dynamics : an official publication of the American Association of Anatomists,* 229**,** 289-99.

KIKUCHI, N., KWON, Y. D., GOTOH, M. & NARIMATSU, H. 2003. Comparison of glycosyltransferase families using the profile hidden Markov model. *Biochemical and biophysical research communications,* 310**,** 574-9.

KIM, S. W., FANG, X., JI, H., PAULSON, A. F., DANIEL, J. M., CIESIOLKA, M., VAN ROY, F. & MCCREA, P. D. 2002. Isolation and characterization of XKaiso, a transcriptional repressor that associates with the catenin Xp120(ctn) in Xenopus laevis. *The Journal of biological chemistry,* 277**,** 8202-8.

KIM, U., SIEGEL, R., REN, X., GUNTHER, C. S., GAASTERLAND, T. & ROEDER, R. G. 2003. Identification of transcription coactivator OCA-B-

dependent genes involved in antigen-dependent B cell differentiation by cDNA array analyses. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 8868-73.

KOBAYASHI, K., NAKANO, H., HAYASHI, M., SHIMAZAKI, M., FUKUTANI, Y., SASAKI, K., SUGIMORI, K. & KOSHINO, Y. 2003. Association of phosphorylation site of tau protein with neuronal apoptosis in Alzheimer's disease. *Journal of the neurological sciences,* 208**,** 17-24.

KONRAD, R. J., ZHANG, F., HALE, J. E., KNIERMAN, M. D., BECKER, G. W. & KUDLOW, J. E. 2002. Alloxan is an inhibitor of the enzyme O-linked N-acetylglucosamine transferase. *Biochemical and biophysical research communications,* 293**,** 207-12.

KOSHLAND, M. E. 1953. The origin of fecal antibody and its relationship to immunization with adjuvant. *Journal of immunology,* 70**,** 359-65.

KREPPEL, L. K., BLOMBERG, M. A. & HART, G. W. 1997. Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. *The Journal of biological chemistry,* 272**,** 9308-15.

LAIRSON, L. L., HENRISSAT, B., DAVIES, G. J. & WITHERS, S. G. 2008. Glycosyltransferases: structures, functions, and mechanisms. *Annual review of biochemistry,* 77**,** 521-55.

LARIVIERE, L., KURZECK, J., ASCHKE-SONNENBORN, U., RUGER, W. & MORERA, S. 2002. Crystallization and preliminary crystallographic study of a ternary complex between the T4 phage beta-glucosyltransferase, uridine diphosphoglucose and a DNA fragment containing an abasic site. *Acta crystallographica. Section D, Biological crystallography,* 58**,** 1484-6.

LARIVIERE, L. & MORERA, S. 2002. A base-flipping mechanism for the T4 phage beta-glucosyltransferase and identification of a transition-state analog. *Journal of molecular biology,* 324**,** 483-90.

LAZARUS, B. D., LOVE, D. C. & HANOVER, J. A. 2006. Recombinant O-GlcNAc transferase isoforms: identification of O-GlcNAcase, yes tyrosine kinase, and tau as isoform-specific substrates. *Glycobiology,* 16**,** 415-21.

LAZARUS, M. B., JIANG, J., GLOSTER, T. M., ZANDBERG, W. F., WHITWORTH, G. E., VOCADLO, D. J. & WALKER, S. 2012. Structural snapshots of the reaction coordinate for O-GlcNAc transferase. *Nature chemical biology,* 8**,** 966-8.

LAZARUS, M. B., NAM, Y., JIANG, J., SLIZ, P. & WALKER, S. 2011. Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature,* 469**,** 564-7.

LEFEBVRE, T., DEHENNAUT, V., GUINEZ, C., OLIVIER, S., DROUGAT, L., MIR, A. M., MORTUAIRE, M., VERCOUTTER-EDOUART, A. S. & MICHALSKI, J. C. 2010. Dysregulation of the nutrient/stress sensor O-GlcNAcylation is involved in the etiology of cardiovascular disorders, type-2 diabetes and Alzheimer's disease. *Biochimica et biophysica acta,* 1800**,** 67-79.

LIRA-NAVARRETE, E., VALERO-GONZALEZ, J., VILLANUEVA, R., MARTINEZ-JULVEZ, M., TEJERO, T., MERINO, P., PANJIKAR, S. & HURTADO-GUERRERO, R. 2011. Structural insights into the mechanism of protein O-fucosylation. *PloS one,* 6**,** e25365.

LIU, J., MARCHASE, R. B. & CHATHAM, J. C. 2007. Glutamine-induced protection of isolated rat heart from ischemia/reperfusion injury is mediated via the hexosamine biosynthesis pathway and increased protein O-GlcNAc levels. *Journal of molecular and cellular cardiology,* 42**,** 177-85.

LIU, K., PATERSON, A. J., ZHANG, F., MCANDREW, J., FUKUCHI, K., WYSS, J. M., PENG, L., HU, Y. & KUDLOW, J. E. 2004. Accumulation of protein O-GlcNAc modification inhibits proteasomes in the brain and coincides with neuronal apoptosis in brain areas with high O-GlcNAc metabolism. *Journal of neurochemistry,* 89**,** 1044-55.

LIU, Y., LI, X., YU, Y., SHI, J., LIANG, Z., RUN, X., LI, Y., DAI, C. L., GRUNDKE-IQBAL, I., IQBAL, K., LIU, F. & GONG, C. X. 2012. Developmental regulation of protein O-GlcNAcylation, O-GlcNAc transferase, and O-GlcNAcase in mammalian brain. *PloS one,* 7**,** e43724.

LIZAK, C., GERBER, S., NUMAO, S., AEBI, M. & LOCHER, K. P. 2011. X-ray structure of a bacterial oligosaccharyltransferase. *Nature,* 474**,** 350-5.

LOVE, D. C., GHOSH, S., MONDOUX, M. A., FUKUSHIGE, T., WANG, P., WILSON, M. A., ISER, W. B., WOLKOW, C. A., KRAUSE, M. W. & HANOVER, J. A. 2010. Dynamic O-GlcNAc cycling at promoters of Caenorhabditis elegans genes regulating longevity, stress, and immunity. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 7413-8.

LOVE, D. C., KOCHAN, J., CATHEY, R. L., SHIN, S. H. & HANOVER, J. A. 2003. Mitochondrial and nucleocytoplasmic targeting of O-linked GlcNAc transferase. *Journal of cell science,* 116**,** 647-54.

LOVE, D. C., KRAUSE, M. W. & HANOVER, J. A. 2010. O-GlcNAc cycling: emerging roles in development and epigenetics. *Seminars in cell & developmental biology,* 21**,** 646-54.

LOVERING, A. L., DE CASTRO, L. H., LIM, D. & STRYNADKA, N. C. 2007. Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis. *Science,* 315**,** 1402-5.

LUBAS, W. A., FRANK, D. W., KRAUSE, M. & HANOVER, J. A. 1997. O-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. *The Journal of biological chemistry,* 272**,** 9316-24.

LUBAS, W. A. & HANOVER, J. A. 2000. Functional expression of O-linked GlcNAc transferase. Domain structure and substrate specificity. *The Journal of biological chemistry,* 275**,** 10983-8.

MACAULEY, M. S., BUBB, A. K., MARTINEZ-FLEITES, C., DAVIES, G. J. & VOCADLO, D. J. 2008. Elevation of global O-GlcNAc levels in 3T3-L1 adipocytes by selective inhibition of O-GlcNAcase does not induce insulin resistance. *The Journal of biological chemistry,* 283**,** 34687-95.

MACAULEY, M. S., WHITWORTH, G. E., DEBOWSKI, A. W., CHIN, D. & VOCADLO, D. J. 2005. O-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors. *The Journal of biological chemistry,* 280**,** 25313-22.

MARSHALL, S., BACOTE, V. & TRAXINGER, R. R. 1991. Discovery of a metabolic pathway mediating glucose-induced desensitization of the glucose transport system. Role of hexosamine biosynthesis in the induction of insulin resistance. *The Journal of biological chemistry,* 266**,** 4706-12.

MARSHALL, S., BACOTE, V. & TRAXINGER, R. R. 1991. Complete inhibition of glucose-induced desensitization of the glucose transport system by inhibitors of mRNA synthesis. Evidence for rapid turnover of glutamine:fructose-6-phosphate amidotransferase. *The Journal of biological chemistry,* 266**,** 10155-61.

MARSHALL, S., NADEAU, O. & YAMASAKI, K. 2004. Dynamic actions of glucose and glucosamine on hexosamine biosynthesis in isolated adipocytes: differential effects on glucosamine 6-phosphate, UDP-N-acetylglucosamine, and ATP levels. *The Journal of biological chemistry,* 279**,** 35313-9.

MARTINEZ-FLEITES, C., MACAULEY, M. S., HE, Y., SHEN, D. L., VOCADLO, D. J. & DAVIES, G. J. 2008. Structure of an O-GlcNAc transferase homolog provides insight into intracellular glycosylation. *Nature structural & molecular biology,* 15**,** 764-5.

MARZ, P., PROBST, A., LANG, S., SCHWAGER, M., ROSE-JOHN, S., OTTEN, U. & OZBEK, S. 2004. Ataxin-10, the spinocerebellar ataxia type 10 neurodegenerative disorder protein, is essential for survival of cerebellar neurons. *The Journal of biological chemistry,* 279**,** 35542-50.

MARZ, P., STETEFELD, J., BENDFELDT, K., NITSCH, C., REINSTEIN, J., SHOEMAN, R. L., DIMITRIADES-SCHMUTZ, B., SCHWAGER, M.,

LEISER, D., OZCAN, S., OTTEN, U. & OZBEK, S. 2006. Ataxin-10 interacts with O-linked beta-N-acetylglucosamine transferase in the brain. *The Journal of biological chemistry,* 281**,** 20263-70.

MATTHEWS, J. A., ACEVEDO-DUNCAN, M. & POTTER, R. L. 2005. Selective decrease of membrane-associated PKC-alpha and PKC-epsilon in response to elevated intracellular O-GlcNAc levels in transformed human glial cells. *Biochimica et biophysica acta,* 1743**,** 305-15.

MATULIS, D., KRANZ, J. K., SALEMME, F. R. & TODD, M. J. 2005. Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry,* 44**,** 5258-66.

MCCLAIN, D. A., LUBAS, W. A., COOKSEY, R. C., HAZEL, M., PARKER, G. J., LOVE, D. C. & HANOVER, J. A. 2002. Altered glycan-dependent signaling induces insulin resistance and hyperleptinemia. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 10695-9.

MILEY, M. J., ZIELINSKA, A. K., KEENAN, J. E., BRATTON, S. M., RADOMINSKA-PANDYA, A. & REDINBO, M. R. 2007. Crystal structure of the cofactor-binding domain of the human phase II drug-metabolism enzyme UDP-glucuronosyltransferase 2B7. *Journal of molecular biology,* 369**,** 498-511.

MORERA, S., LARIVIERE, L., KURZECK, J., ASCHKE-SONNENBORN, U., FREEMONT, P. S., JANIN, J. & RUGER, W. 2001. High resolution crystal structures of T4 phage beta-glucosyltransferase: induced fit and effect of substrate and metal binding. *Journal of molecular biology,* 311**,** 569-77.

MULICHAK, A. M., LOSEY, H. C., LU, W., WAWRZAK, Z., WALSH, C. T. & GARAVITO, R. M. 2003. Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 9238-43.

MULICHAK, A. M., LOSEY, H. C., WALSH, C. T. & GARAVITO, R. M. 2001. Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. *Structure,* 9**,** 547-57.

NAGAR, B., HANTSCHEL, O., YOUNG, M. A., SCHEFFZEK, K., VEACH, D., BORNMANN, W., CLARKSON, B., SUPERTI-FURGA, G. & KURIYAN, J. 2003. Structural basis for the autoinhibition of c-Abl tyrosine kinase. *Cell,* 112**,** 859-71.

O'DONNELL, N., ZACHARA, N. E., HART, G. W. & MARTH, J. D. 2004. Ogt-dependent X-chromosome-linked protein glycosylation is a requisite modification in somatic cell function and embryo viability. *Molecular and*

cellular biology, 24**,** 1680-90.

OFFEN, W., MARTINEZ-FLEITES, C., YANG, M., KIAT-LIM, E., DAVIS, B. G., TARLING, C. A., FORD, C. M., BOWLES, D. J. & DAVIES, G. J. 2006. Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *The EMBO journal,* 25**,** 1396-405.

OLIVIER-VAN STICHELEN, S., DROUGAT, L., DEHENNAUT, V., EL YAZIDI-BELKOURA, I., GUINEZ, C., MIR, A. M., MICHALSKI, J. C., VERCOUTTER-EDOUART, A. S. & LEFEBVRE, T. 2012. Serum-stimulated cell cycle entry promotes ncOGT synthesis required for cyclin D expression. *Oncogenesis,* 1**,** e36.

OLSZEWSKI, N. E., WEST, C. M., SASSI, S. O. & HARTWECK, L. M. 2010. O-GlcNAc protein modification in plants: Evolution and function. *Biochimica et biophysica acta,* 1800**,** 49-56.

OVERDIJK, B., VAN DER KROEF, W. M., LISMAN, J. J., PIERCE, R. J., MONTREUIL, J. & SPIK, G. 1981. Demonstration and partial characterization of endo-N-acetyl-beta-D-glucosaminidase in human tissues. *FEBS letters,* 128**,** 364-6.

PARK, S. Y., RYU, J. & LEE, W. 2005. O-GlcNAc modification on IRS-1 and Akt2 by PUGNAc inhibits their phosphorylation and induces insulin resistance in rat primary adipocytes. *Experimental & molecular medicine,* 37**,** 220-9.

PATHAK, S., BORODKIN, V. S., ALBARBARAWI, O., CAMPBELL, D. G., IBRAHIM, A. & VAN AALTEN, D. M. 2012. O-GlcNAcylation of TAB1 modulates TAK1-mediated cytokine release. *The EMBO journal,* 31**,** 1394-404.

PAWSON, T. 2004. Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell,* 116**,** 191-203.

PHATNANI, H. P. & GREENLEAF, A. L. 2006. Phosphorylation and functions of the RNA polymerase II CTD. *Genes & development,* 20**,** 2922-36.

PHILLIPS, D. C. 1966. The three-dimensional structure of an enzyme molecule. *Scientific American,* 215**,** 78-90.

QUINAUD, M., PLE, S., JOB, V., CONTRERAS-MARTEL, C., SIMORRE, J. P., ATTREE, I. & DESSEN, A. 2007. Structure of the heterotrimeric complex that regulates type III secretion needle formation. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 7803-8.

RAMAKRISHNAN, B., BOEGGEMAN, E. & QASBA, P. K. 2002. Beta-1,4-galactosyltransferase and lactose synthase: molecular mechanical devices. *Biochemical and biophysical research communications,* 291**,** 1113-8.

RAMAKRISHNAN, B., RAMASAMY, V. & QASBA, P. K. 2006. Structural snapshots of beta-1,4-galactosyltransferase-I along the kinetic pathway. *Journal of molecular biology,* 357**,** 1619-33.

RAMASAMY, V., RAMAKRISHNAN, B., BOEGGEMAN, E., RATNER, D. M., SEEBERGER, P. H. & QASBA, P. K. 2005. Oligosaccharide preferences of beta1,4-galactosyltransferase-I: crystal structures of Met340His mutant of human beta1,4-galactosyltransferase-I with a pentasaccharide and trisaccharides of the N-glycan moiety. *Journal of molecular biology,* 353**,** 53-67.

RANUNCOLO, S. M., GHOSH, S., HANOVER, J. A., HART, G. W. & LEWIS, B. A. 2012. Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription in vitro and in vivo. *The Journal of biological chemistry,* 287**,** 23549-61.

RANUNCOLO, S. M., PITTALUGA, S., EVBUOMWAN, M. O., JAFFE, E. S. & LEWIS, B. A. 2012. Hodgkin lymphoma requires stabilized NIK and constitutive RelB expression for survival. *Blood,* 120**,** 3756-63.

RAO, F. V., DORFMUELLER, H. C., VILLA, F., ALLWOOD, M., EGGLESTON, I. M. & VAN AALTEN, D. M. 2006. Structural insights into the mechanism and inhibition of eukaryotic O-GlcNAc hydrolysis. *The EMBO journal,* 25**,** 1569-78.

RESH, M. D. 1999. Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochimica et biophysica acta,* 1451**,** 1-16.

REXACH, J. E., CLARK, P. M., MASON, D. E., NEVE, R. L., PETERS, E. C. & HSIEH-WILSON, L. C. 2012. Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. *Nature chemical biology,* 8**,** 253-61.

REXACH, J. E., CLARK, P. M., MASON, D. E., NEVE, R. L., PETERS, E. C. & HSIEH-WILSON, L. C. 2012. Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. *Nature chemical biology,* 8**,** 253-61.

ROBINSON, K. A., YACOUB WASEF, S. Z. & BUSE, M. G. 2006. At therapeutic concentrations, olanzapine does not affect basal or insulin-stimulated glucose transport in 3T3-L1 adipocytes. *Progress in neuro-psychopharmacology & biological psychiatry,* 30**,** 93-8.

ROQUEMORE, E. P., DELL, A., MORRIS, H. R., PANICO, M., REASON, A. J., SAVOY, L. A., WISTOW, G. J., ZIGLER, J. S., JR., EARLES, B. J. & HART, G. W. 1992. Vertebrate lens alpha-crystallins are modified by O-linked N-acetylglucosamine. *The Journal of biological chemistry,* 267**,** 555-63.

RUAN, H. B., HAN, X., LI, M. D., SINGH, J. P., QIAN, K., AZARHOUSH, S., ZHAO, L., BENNETT, A. M., SAMUEL, V. T., WU, J., YATES, J. R., 3RD & YANG, X. 2012. O-GlcNAc transferase/host cell factor C1 complex regulates gluconeogenesis by modulating PGC-1alpha stability. *Cell metabolism,* 16**,** 226-37.

SALA, R. F., MACKINNON, S. L., PALCIC, M. M. & TANNER, M. E. 1998. UDP-N-trifluoroacetylglucosamine as an alternative substrate in N-acetylglucosaminyltransferase reactions. *Carbohydrate research,* 306**,** 127-36.

SAMPATHKUMAR, P., ROACH, C., MICHELS, P. A. & HOL, W. G. 2008. Structural insights into the recognition of peroxisomal targeting signal 1 by Trypanosoma brucei peroxin 5. *Journal of molecular biology,* 381**,** 867-80.

SCHEUFLER, C., BRINKER, A., BOURENKOV, G., PEGORARO, S., MORODER, L., BARTUNIK, H., HARTL, F. U. & MOAREFI, I. 2000. Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell,* 101**,** 199-210.

SCHIMPL, M., ZHENG, X., BORODKIN, V. S., BLAIR, D. E., FERENBACH, A. T., SCHUTTELKOPF, A. W., NAVRATILOVA, I., ARISTOTELOUS, T., ALBARBARAWI, O., ROBINSON, D. A., MACNAUGHTAN, M. A. & VAN AALTEN, D. M. 2012. O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis. *Nature chemical biology,* 8**,** 969-74.

SCHINDLER, M., HOGAN, M., MILLER, R. & DEGAETANO, D. 1987. A nuclear specific glycoprotein representative of a unique pattern of glycosylation. *The Journal of biological chemistry,* 262**,** 1254-60.

SCHIRM, M., KALMOKOFF, M., AUBRY, A., THIBAULT, P., SANDOZ, M. & LOGAN, S. M. 2004. Flagellin from Listeria monocytogenes is glycosylated with beta-O-linked N-acetylglucosamine. *Journal of bacteriology,* 186**,** 6721-7.

SCHIRM, M., SOO, E. C., AUBRY, A. J., AUSTIN, J., THIBAULT, P. & LOGAN, S. M. 2003. Structural, genetic and functional characterization of the flagellin glycosylation process in Helicobacter pylori. *Molecular microbiology,* 48**,** 1579-92.

SCHWARTZ, Y. B. & PIRROTTA, V. 2007. Polycomb silencing mechanisms and the management of genomic programmes. *Nature reviews. Genetics,* 8**,** 9-22.

SEKINE, O., LOVE, D. C., RUBENSTEIN, D. S. & HANOVER, J. A. 2010. Blocking O-linked GlcNAc cycling in Drosophila insulin-producing cells perturbs glucose-insulin homeostasis. *The Journal of biological chemistry,*

285, 38684-91.

SHAFI, R., IYER, S. P., ELLIES, L. G., O'DONNELL, N., MAREK, K. W., CHUI, D., HART, G. W. & MARTH, J. D. 2000. The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proceedings of the National Academy of Sciences of the United States of America,* 97, 5735-9.

SHAO, H., HE, X., ACHNINE, L., BLOUNT, J. W., DIXON, R. A. & WANG, X. 2005. Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from Medicago truncatula. *The Plant cell,* 17, 3141-54.

SHEN, A., KAMP, H. D., GRUNDLING, A. & HIGGINS, D. E. 2006. A bifunctional O-GlcNAc transferase governs flagellar motility through anti-repression. *Genes & development,* 20, 3283-95.

SHI, F., KIM, H., LU, W., HE, Q., LIU, D., GOODELL, M. A., WAN, M. & SONGYANG, Z. 2013. Ten-eleven translocation 1 (Tet1) Is Regulated by O-GlcNAc Transferase (Ogt) for Target Gene Repression in Mouse Embryonic Stem Cells. *The Journal of biological chemistry*.

SICHERI, F., MOAREFI, I. & KURIYAN, J. 1997. Crystal structure of the Src family tyrosine kinase Hck. *Nature,* 385, 602-9.

SIKORSKI, R. S., BOGUSKI, M. S., GOEBL, M. & HIETER, P. 1990. A repeating amino acid motif in CDC23 defines a family of proteins and a new relationship among genes required for mitosis and RNA synthesis. *Cell,* 60, 307-17.

SILVERSTONE, A. L., TSENG, T. S., SWAIN, S. M., DILL, A., JEONG, S. Y., OLSZEWSKI, N. E. & SUN, T. P. 2007. Functional analysis of SPINDLY in gibberellin signaling in Arabidopsis. *Plant physiology,* 143, 987-1000.

SIMANEK, E. E., MCGARVEY, G. J., JABLONOWSKI, J. A. & WONG, C. H. 1998. Selectinminus signCarbohydrate Interactions: From Natural Ligands to Designed Mimics. *Chemical reviews,* 98, 833-862.

SINCLAIR, D. A., SYRZYCKA, M., MACAULEY, M. S., RASTGARDANI, T., KOMLJENOVIC, I., VOCADLO, D. J., BROCK, H. W. & HONDA, B. M. 2009. Drosophila O-GlcNAc transferase (OGT) is encoded by the Polycomb group (PcG) gene, super sex combs (sxc). *Proceedings of the National Academy of Sciences of the United States of America,* 106, 13427-32.

SLAWSON, C., LAKSHMANAN, T., KNAPP, S. & HART, G. W. 2008. A mitotic GlcNAcylation/phosphorylation signaling complex alters the posttranslational state of the cytoskeletal protein vimentin. *Molecular biology of the cell,* 19, 4130-40.

SLAWSON, C., SHAFII, S., AMBURGEY, J. & POTTER, R. 2002. Characterization of the O-GlcNAc protein modification in Xenopus laevis oocyte during oogenesis and progesterone-stimulated maturation. *Biochimica et biophysica acta,* 1573**,** 121-9.

SLAWSON, C., ZACHARA, N. E., VOSSELLER, K., CHEUNG, W. D., LANE, M. D. & HART, G. W. 2005. Perturbations in O-linked beta-N-acetylglucosamine protein modification cause severe defects in mitotic progression and cytokinesis. *The Journal of biological chemistry,* 280**,** 32944-56.

SLAWSON, D. C. & SHAUGHNESSY, A. F. 2005. Teaching evidence-based medicine: should we be teaching information management instead? *Academic medicine : journal of the Association of American Medical Colleges,* 80**,** 685-9.

SOHN, K. C. & DO, S. I. 2005. Transcriptional regulation and O-GlcNAcylation activity of zebrafish OGT during embryogenesis. *Biochemical and biophysical research communications,* 337**,** 256-63.

STRAHL-BOLSINGER, S., IMMERVOLL, T., DEUTZMANN, R. & TANNER, W. 1993. PMT1, the gene for a key enzyme of protein O-glycosylation in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America,* 90**,** 8164-8.

SUMMY, J. M., QIAN, Y., JIANG, B. H., GUAPPONE-KOAY, A., GATESMAN, A., SHI, X. & FLYNN, D. C. 2003. The SH4-Unique-SH3-SH2 domains dictate specificity in signaling that differentiate c-Yes from c-Src. *Journal of cell science,* 116**,** 2585-98.

SWALLOW, D. M., EVANS, L., SAHA, N. & HARRIS, H. 1976. Characterization and tissue distribution of N-acetyl hexosaminidase C: suggestive evidence for a separate hexosaminidase locus. *Annals of human genetics,* 40**,** 55-66.

TAKAHASHI, M., INOUE, N., OHISHI, K., MAEDA, Y., NAKAMURA, N., ENDO, Y., FUJITA, T., TAKEDA, J. & KINOSHITA, T. 1996. PIG-B, a membrane protein of the endoplasmic reticulum with a large lumenal domain, is involved in transferring the third mannose of the GPI anchor. *The EMBO journal,* 15**,** 4254-61.

TARRANT, M. K., RHO, H. S., XIE, Z., JIANG, Y. L., GROSS, C., CULHANE, J. C., YAN, G., QIAN, J., ICHIKAWA, Y., MATSUOKA, T., ZACHARA, N., ETZKORN, F. A., HART, G. W., JEONG, J. S., BLACKSHAW, S., ZHU, H. & COLE, P. A. 2012. Regulation of CK2 by phosphorylation and O-GlcNAcylation revealed by semisynthesis. *Nature chemical biology,* 8**,** 262-9.

THIBAULT, P., LOGAN, S. M., KELLY, J. F., BRISSON, J. R., EWING, C. P.,

TRUST, T. J. & GUERRY, P. 2001. Identification of the carbohydrate moieties and glycosylation motifs in Campylobacter jejuni flagellin. *The Journal of biological chemistry,* 276**,** 34862-70.

TOLEMAN, C., PATERSON, A. J., WHISENHUNT, T. R. & KUDLOW, J. E. 2004. Characterization of the histone acetyltransferase (HAT) domain of a bifunctional protein with activable O-GlcNAcase and HAT activities. *The Journal of biological chemistry,* 279**,** 53665-73.

TORRES, C. R. & HART, G. W. 1984. Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *The Journal of biological chemistry,* 259**,** 3308-17.

TRINIDAD, J. C., BARKAN, D. T., GULLEDGE, B. F., THALHAMMER, A., SALI, A., SCHOEPFER, R. & BURLINGAME, A. L. 2012. Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. *Molecular & cellular proteomics : MCP,* 11**,** 215-29.

TSENG, T. S., SWAIN, S. M. & OLSZEWSKI, N. E. 2001. Ectopic expression of the tetratricopeptide repeat domain of SPINDLY causes defects in gibberellin response. *Plant physiology,* 126**,** 1250-8.

TVAROSKA, I., KOZMON, S., WIMMEROVA, M. & KOCA, J. 2012. Substrate-assisted catalytic mechanism of O-GlcNAc transferase discovered by quantum mechanics/molecular mechanics investigation. *Journal of the American Chemical Society,* 134**,** 15563-71.

VELLA, P., SCELFO, A., JAMMULA, S., CHIACCHIERA, F., WILLIAMS, K., CUOMO, A., ROBERTO, A., CHRISTENSEN, J., BONALDI, T., HELIN, K. & PASINI, D. 2013. Tet Proteins Connect the O-Linked N-acetylglucosamine Transferase Ogt to Chromatin in Embryonic Stem Cells. *Molecular cell,* 49**,** 645-656.

VOCADLO, D. J., HANG, H. C., KIM, E. J., HANOVER, J. A. & BERTOZZI, C. R. 2003. A chemical approach for identifying O-GlcNAc-modified proteins in cells. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 9116-21.

VOSSELLER, K., TRINIDAD, J. C., CHALKLEY, R. J., SPECHT, C. G., THALHAMMER, A., LYNN, A. J., SNEDECOR, J. O., GUAN, S., MEDZIHRADSZKY, K. F., MALTBY, D. A., SCHOEPFER, R. & BURLINGAME, A. L. 2006. O-linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Molecular & cellular proteomics : MCP,* 5**,** 923-34.

VOSSELLER, K., WELLS, L., LANE, M. D. & HART, G. W. 2002. Elevated nucleocytoplasmic glycosylation by O-GlcNAc results in insulin resistance

associated with defects in Akt activation in 3T3-L1 adipocytes. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 5313-8.

VRIELINK, A., RUGER, W., DRIESSEN, H. P. & FREEMONT, P. S. 1994. Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *The EMBO journal,* 13**,** 3413-22.

WANG, J., DYE, B. T., RAJASHANKAR, K. R., KURINOV, I. & SCHULMAN, B. A. 2009. Insights into anaphase promoting complex TPR subdomain assembly from a CDC26-APC6 structure. *Nature structural & molecular biology,* 16**,** 987-9.

WANG, Z., GUCEK, M. & HART, G. W. 2008. Cross-talk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated O-GlcNAc. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 13793-8.

WANG, Z., PANDEY, A. & HART, G. W. 2007. Dynamic interplay between O-linked N-acetylglucosaminylation and glycogen synthase kinase-3-dependent phosphorylation. *Molecular & cellular proteomics : MCP,* 6**,** 1365-79.

WANG, Z., UDESHI, N. D., SLAWSON, C., COMPTON, P. D., SAKABE, K., CHEUNG, W. D., SHABANOWITZ, J., HUNT, D. F. & HART, G. W. 2010. Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. *Science signaling,* 3**,** ra2.

WATSON, L. J., FACUNDO, H. T., NGOH, G. A., AMEEN, M., BRAINARD, R. E., LEMMA, K. M., LONG, B. W., PRABHU, S. D., XUAN, Y. T. & JONES, S. P. 2010. O-linked beta-N-acetylglucosamine transferase is indispensable in the failing heart. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 17797-802.

WEBB, N. A., MULICHAK, A. M., LAM, J. S., ROCCHETTA, H. L. & GARAVITO, R. M. 2004. Crystal structure of a tetrameric GDP-D-mannose 4,6-dehydratase from a bacterial GDP-D-rhamnose biosynthetic pathway. *Protein science : a publication of the Protein Society,* 13**,** 529-39.

WEBSTER, D. M., TEO, C. F., SUN, Y., WLOGA, D., GAY, S., KLONOWSKI, K. D., WELLS, L. & DOUGAN, S. T. 2009. O-GlcNAc modifications regulate cell survival and epiboly during zebrafish development. *BMC developmental biology,* 9**,** 28.

WELLS, L., GAO, Y., MAHONEY, J. A., VOSSELLER, K., CHEN, C., ROSEN, A. & HART, G. W. 2002. Dynamic O-glycosylation of nuclear and cytosolic proteins: further characterization of the nucleocytoplasmic beta-N-acetylglucosaminidase, O-GlcNAcase. *The Journal of biological*

chemistry, 277**,** 1755-61.

WELLS, L., KREPPEL, L. K., COMER, F. I., WADZINSKI, B. E. & HART, G. W. 2004. O-GlcNAc transferase is in a functional complex with protein phosphatase 1 catalytic subunits. *The Journal of biological chemistry,* 279**,** 38466-70.

WELLS, L., SLAWSON, C. & HART, G. W. 2011. The E2F-1 associated retinoblastoma-susceptibility gene product is modified by O-GlcNAc. *Amino acids,* 40**,** 877-83.

WHISENHUNT, T. R., YANG, X., BOWE, D. B., PATERSON, A. J., VAN TINE, B. A. & KUDLOW, J. E. 2006. Disrupting the enzyme complex regulating O-GlcNAcylation blocks signaling and development. *Glycobiology,* 16**,** 551-63.

WILLIAMS, K., CHRISTENSEN, J. & HELIN, K. 2012. DNA methylation: TET proteins-guardians of CpG islands? *EMBO reports,* 13**,** 28-35.

WYSOCKA, J., MYERS, M. P., LAHERTY, C. D., EISENMAN, R. N. & HERR, W. 2003. Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes & development,* 17**,** 896-911.

XING, D., FENG, W., NOT, L. G., MILLER, A. P., ZHANG, Y., CHEN, Y. F., MAJID-HASSAN, E., CHATHAM, J. C. & OPARIL, S. 2008. Increased protein O-GlcNAc modification inhibits inflammatory and neointimal responses to acute endoluminal arterial injury. *American journal of physiology. Heart and circulatory physiology,* 295**,** H335-42.

XU, W., DOSHI, A., LEI, M., ECK, M. J. & HARRISON, S. C. 1999. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Molecular cell,* 3**,** 629-38.

XU, W., HARRISON, S. C. & ECK, M. J. 1997. Three-dimensional structure of the tyrosine kinase c-Src. *Nature,* 385**,** 595-602.

YANG, W. H., KIM, J. E., NAM, H. W., JU, J. W., KIM, H. S., KIM, Y. S. & CHO, J. W. 2006. Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. *Nature cell biology,* 8**,** 1074-83.

YANG, W. H., KIM, J. E., NAM, H. W., JU, J. W., KIM, H. S., KIM, Y. S. & CHO, J. W. 2006. Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. *Nature cell biology,* 8**,** 1074-83.

YANG, X., ONGUSAHA, P. P., MILES, P. D., HAVSTAD, J. C., ZHANG, F., SO, W. V., KUDLOW, J. E., MICHELL, R. H., OLEFSKY, J. M., FIELD, S. J. & EVANS, R. M. 2008. Phosphoinositide signalling links O-GlcNAc transferase to insulin resistance. *Nature,* 451**,** 964-9.

YANG, X., ZHANG, F. & KUDLOW, J. E. 2002. Recruitment of O-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein O-GlcNAcylation to transcriptional repression. *Cell,* 110**,** 69-80.

YAO, P. J. & COLEMAN, P. D. 1998. Reduced O-glycosylated clathrin assembly protein AP180: implication for synaptic vesicle recycling dysfunction in Alzheimer's disease. *Neuroscience letters,* 252**,** 33-6.

YAO, P. J. & COLEMAN, P. D. 1998. Reduction of O-linked N-acetylglucosamine-modified assembly protein-3 in Alzheimer's disease. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 18**,** 2399-411.

YUZWA, S. A., SHAN, X., MACAULEY, M. S., CLARK, T., SKOROBOGATKO, Y., VOSSELLER, K. & VOCADLO, D. J. 2012. Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation. *Nature chemical biology,* 8**,** 393-9.

ZACHARA, N. E. & HART, G. W. 2004. O-GlcNAc modification: a nutritional sensor that modulates proteasome function. *Trends in cell biology,* 14**,** 218-21.

ZACHARA, N. E., O'DONNELL, N., CHEUNG, W. D., MERCER, J. J., MARTH, J. D. & HART, G. W. 2004. Dynamic O-GlcNAc modification of nucleocytoplasmic proteins in response to stress. A survival response of mammalian cells. *The Journal of biological chemistry,* 279**,** 30133-42.

ZEYTUNI, N., BARAN, D., DAVIDOV, G. & ZARIVACH, R. 2012. Inter-phylum structural conservation of the magnetosome-associated TPR-containing protein, MamA. *Journal of structural biology,* 180**,** 479-87.

ZHANG, F., SNEAD, C. M. & CATRAVAS, J. D. 2012. Hsp90 regulates O-linked beta-N-acetylglucosamine transferase: a novel mechanism of modulation of protein O-linked beta-N-acetylglucosamine modification in endothelial cells. *American journal of physiology. Cell physiology,* 302**,** C1786-96.

ZHANG, L., YI, Y., GUO, Q., SUN, Y., MA, S., XIAO, S., GENG, J., ZHENG, Z. & SONG, S. 2012. Hsp90 interacts with AMPK and mediates acetyl-CoA carboxylase phosphorylation. *Cellular signalling,* 24**,** 859-65.

ZHANG, Z., KULKARNI, K., HANRAHAN, S. J., THOMPSON, A. J. & BARFORD, D. 2010. The APC/C subunit Cdc16/Cut9 is a contiguous tetratricopeptide repeat superhelix with a homo-dimer interface similar to Cdc27. *The EMBO journal,* 29**,** 3733-44.

ZIEGLER, M. O., JANK, T., AKTORIES, K. & SCHULZ, G. E. 2008. Conformational changes and reaction of clostridial glycosylating toxins. *Journal of molecular biology,* 377**,** 1346-56.

ZOU, L., YANG, S., HU, S., CHAUDRY, I. H., MARCHASE, R. B. & CHATHAM,

J. C. 2007. The protective effects of PUGNAc on cardiac function after trauma-hemorrhage are mediated via increased protein O-GlcNAc levels. *Shock,* 27**,** 402-8.

Appendix

Primers used to generate TPR truncations on ncOGT

N-term	**CTGGGATCC** ATGGCGTCTTCCGTGGGCAACG
d TPR1	**CTGGGATCC** CCAGACAATACTGGTGTGC
d TPR2	**CTGGGATCC** CCCCTTCTGGCAGAAGCTTATTCG
d TPR3	**CTGGGATCC** CCTGATTTCATCGATGGTTATATTAACCTGC
d TPR4	**CTGGGATCC** CCTGATTTGTACTGTGTTCGC
d TPR5	**CTGGGATCC** CCGAACTTTGCAGTAGCTTGGAG
d TPR6	**CTGGGATCC** CCAAACTTTCTGGATGCTTATATC
d TPR7	**CTGGGATCC** CCAAATCACGCAGTGGTGCACG
d TPR8	**CTGGGATCC** CCACATTTCCCTGATGCTTACTGC
d TPR9	**CTGGGATCC** CCCACCCATGCAGACTCTCTG
d TPR10	**CTGGGATCC** CCAGAGTTTGCTGCTGCCCATTC
d TPR11	**CTGGGATCC** CCTACCTTTGCTGATGCCTAC
d TPR12	**CTGGGATCC** CCTGCATTTGCAGATGCACATAGC
d TPR13	**CTGGGATCC** CCTGATTTTCCTGATGCTTATTGTAAC
Reverse	**CCGCTCGAG** TTATGCTGACTCAGTGACTTCAACAGGC
Reverse	**TAGAGTCGCGGCCGC** TTATGCTGACTCAGTGACTTCAACAGGC

Mutagenesis primers for hOGT

hOGT (walker) construct in pGEX

H506F: RLPSVH*PHHS
f
gaagaataggttgccttctgtgTTtcctcatcatagtatgctatatc
r
gatatagcatactatgatgaggaAAcacagaaggcaacctattcttc

TMT2 H506M: RLPSVH*PHHS
f
gagaagaataggttgccttctgtgATGcctcatcatagtatgctatatcc
r
ggatatagcatactatgatgaggCATcacagaaggcaacctattcttctc

TMT4 H508N: SVHPH*HSML
f
ggttgccttctgtgcatcctAatcatagtatgctatatcctc
r
gaggatatagcatactatgatTaggatgcacagaaggcaacc

TMT5 H509F: VHPHH*SMLY
f
gttgccttctgtgcatcctcatTTtagtatgctatatcctctttc
r
gaaagaggatatagcatactaAAatgaggatgcacagaaggcaac

TMT6 H527W: IAERH*GNLCL
f
ggaaggctattgctgagaggTGGggcaacctgtgcttagataag
r
cttatctaagcacaggttgccCCAcctctcagcaatagccttcc

TMT7 D564N: VSSD*FGNHP
f
gtgtaggatatgtgagttccAactttgggaatcatcctacttc
r
gaagtaggatgattcccaaagtTggaactcacatatcctacac

TMT8 H568N: SDFGNH*PTSHL
f
gagttccgactttgggaatAatcctacttctcaccttatgc
r
gcataaggtgagaagtaggatTattcccaaagtcggaactc

TMT9 T570A: GNHPT*SHLMQ
f
gactttgggaatcatcctGcttctcaccttatgcagtc
r
gactgcataaggtgagaagCaggatgattcccaaagtTMT10 P666L: MWLGYP*GTSG
f
caatgtggctgggataccTtgggacgagtggtgcgc
r
gcgcaccactcgtcccaAggtatcccagccacattg

TMT11 Q849L: CNFNQ*LYKID
f
Gatgccatcgtatactgtaactttaat TTgttgtataaaattgacccttctactttg
r
caaagtagaagggtcaattttatacaacAAattaaagttacagtatacgatggcatC
TMT12 Y851F: NFNQLY*KIDPS
f
gtatactgtaactttaatcagttgtTtaaaattgacccttctactttgcag
r
ctgcaaagtagaagggtcaattttaAacaactgattaaagttacagtatac

TMT13 K852M: FNQLYK*IDPST
f
gtatactgtaactttaatcagttgtataTGattgacccttctactttgcagatgtgg
r
ccacatctgcaaagtagaagggtcaatCAtatacaactgattaaagttacagtatac

TMT14 V905M: IFSPV*APKEE
f
gaaccgtatcatttttcacctAtGgctcctaaagaggaacacgtcagg
r
cctgacgtgttcctctttaggagcCaTaggtgaaaaaatgatacggttc

TMT15 K908M: SPVAPK*EEHVR
f
catttttcacctgttgctcctaTGgaggaacacgtcaggagaggccag
r
ctggcctctcctgacgtgttcctcCAtaggagcaacaggtgaaaaaatg

TMT16 R914M: EEHVRR*GQLADV
f
ctaaagaggaacacgtcaggaTGggccagctggctgatgtctg
r
cagacatcagccagctggccCAtcctgacgtgttcctctttag

TMT17 C927L: CLDTPLC*NGHTT
f
ctgcttggacactccactcCTtaatgggcacaccacaggg
r
ccctgtggtgtgcccattaAGgagtggagtgtccaagcag

TMT18 T931A: CNGHT*TGMDVL
f
cactctgtaatgggcacGccacagggatggatgtc
r
gacatccatccctgtggCgtgcccattacagagtg

TMT19 T932A: CNGHTT*GMDVL
f
ctgtaatgggcacaccGcagggatggatgtcctc
r
gaggacatccatccctgCggtgtgcccattacag

TMT20 D935N: TTGMD*VLWAG
f
cacaccacagggatgAatgtcctctgggcaggg
r

ccctgcccagaggacatTcatccctgtggtgtg

TMT22 H508D: SVHPH*HSML
f
ggttgccttctgtgcatcct**Gat**catagtatgctatatcctc
r
gaggatatagcatactatgatCaggatgcacagaaggcaacc

TMT23 H508E: SVHPH*HSML
f
ggttgccttctgtgcatcct**GaA**catagtatgctatatcctctttc
r
gaaagaggatatagcatactatgTtCaggatgcacagaaggcaacc

TMT24 H508Q: SVHPH*HSML
f
ggttgccttctgtgcatcct**caA**catagtatgctatatcctctttc
r
gaaagaggatatagcatactatgTtgaggatgcacagaaggcaacc

Mutagenesis Primers for *Drosophila* OGT

H596 → A and F

H596A

Aaggaaggcttcgaataggttatctta

DmOGT_H596A_fwd
gctcggacttcgggaat**GCt**cctacctcgcatttaatgc

DmOGT_H596A_rev
gcattaaatgcgaggtaggaGCattcccgaagtccgagc

aatctgtgccaggtcttcat

H596F

Aaggaaggcttcgaataggttatctta

DmOGT_H596F_fwd
gctcggacttcgggaat**TTt**cctacctcgcatttaatgc

DmOGT_H596F_rev
gcattaaatgcgaggtaggaAAattcccgaagtccgagc

aatctgtgccaggtcttcat

K872 K872M

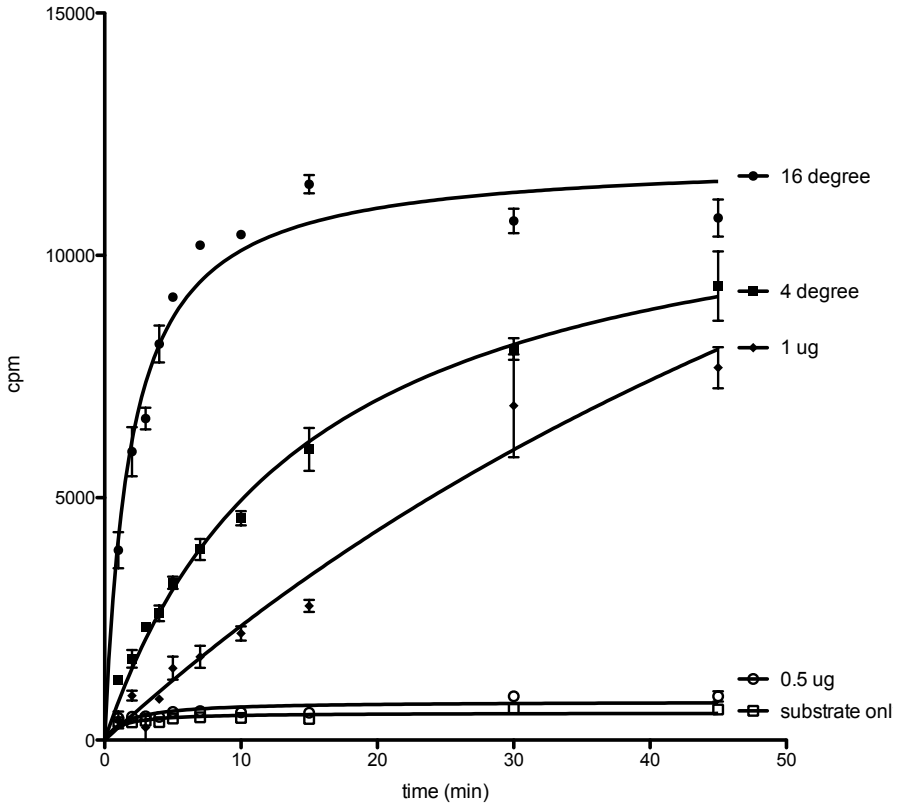
Agatgacgctgttgtgtatt

DmOGT_K872M_fwd
gcaattttaatcagctataca**TG**atcgatccacaaaccctcgagtc

DmOGT_K872M_rev
gactcgagggtttgtggatcgatCAtgtatagctgattaaaattgc

atgggtagaaattttaaaaaatg

in vitro kinase reaction with recombinant Yes1 protein. Protein concentration vs temperature vs time was to ensure optimised reaction condition and steady state kinetics

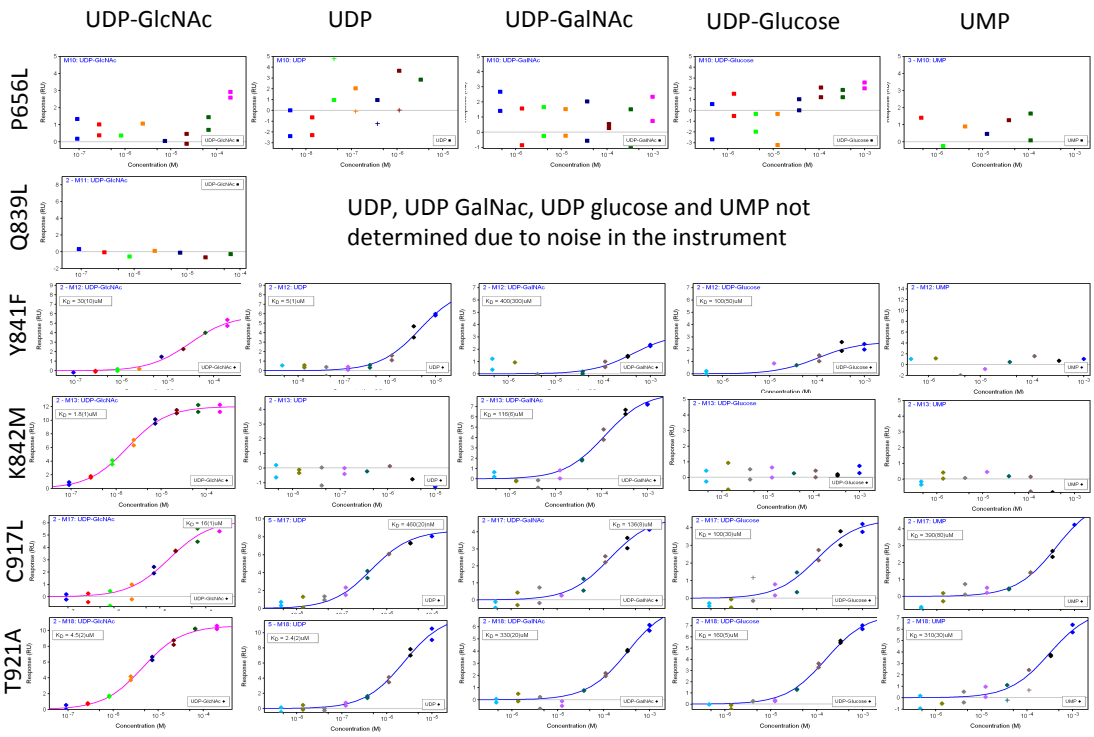


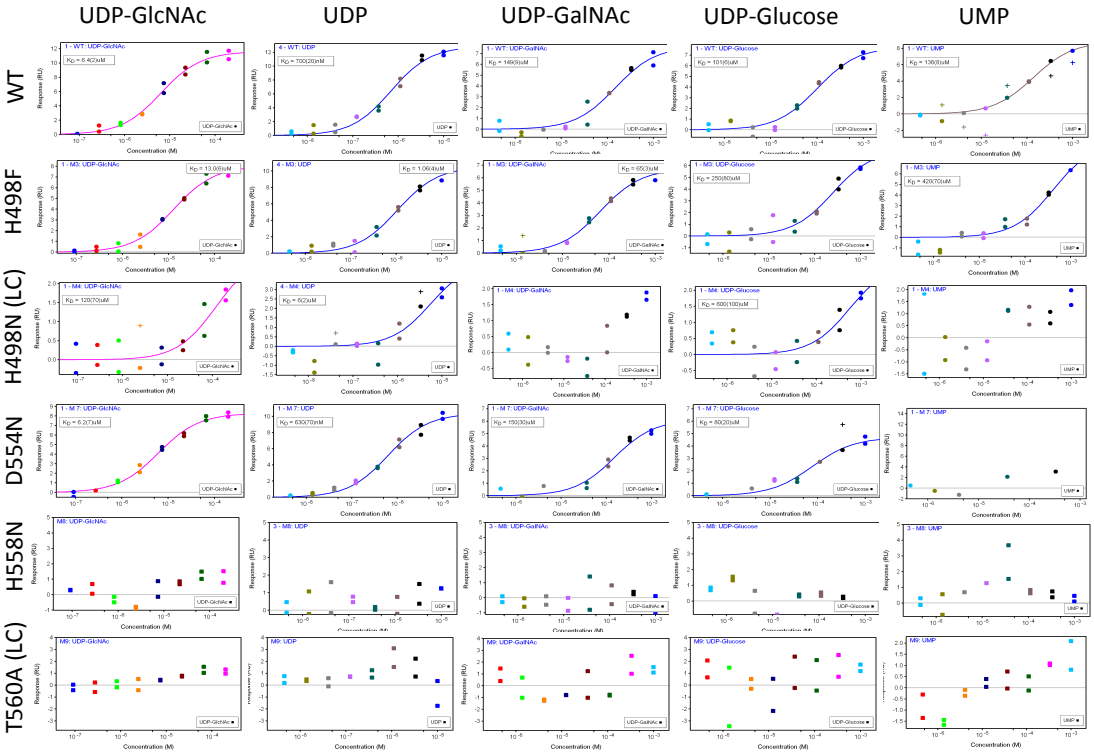
Summary of the activity, protein folding and binding to UDP, UDP-GlcNAc and their analogs.

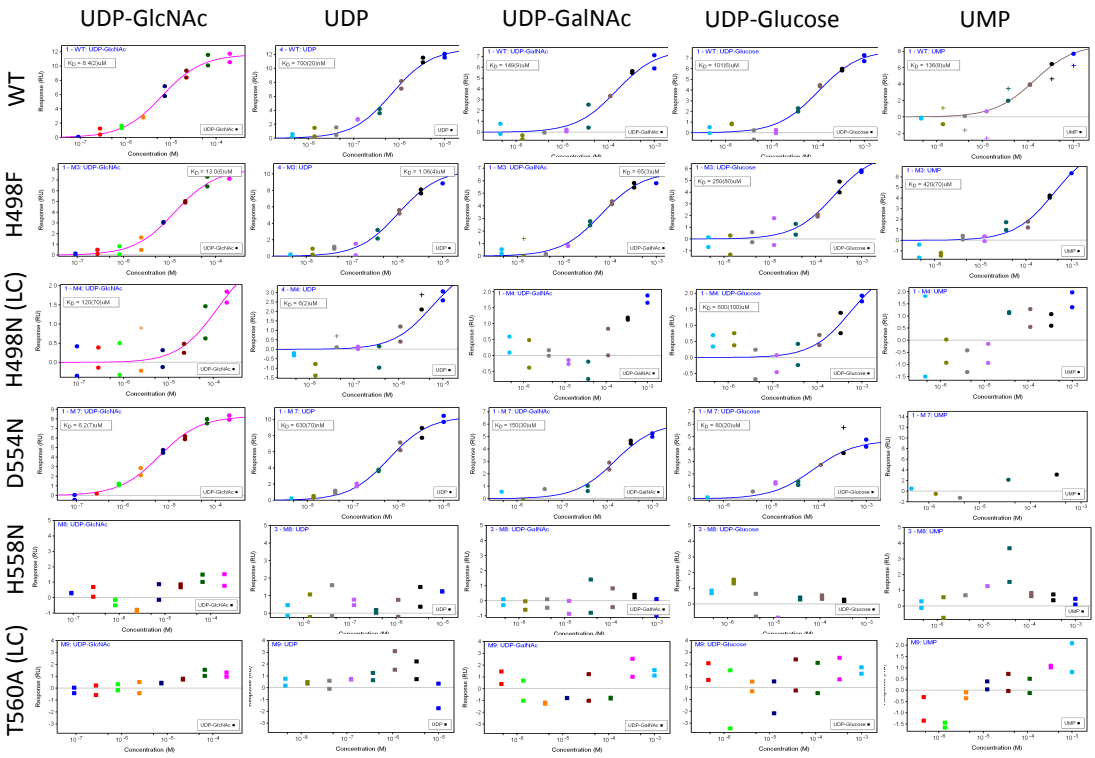
D9 hOGT	Activity	Folding	UDP-GlcNAc	UDP	UDP-GalNAc	UDP-Glucose	UMP
WT	✓	✓	**6.4**	**0.7**	**149**	**101**	**136**
H498F	✓	✓	**13**	**1.06**	**65**	**250**	**420**
H498N	✗	✓	6.2	0.63	150	80	NB
H498E	✗	✓	200	5	NB	900	290
D554N	✓	✓	NB	NB	NB	NB	NB
H558F	✗	✓	8.7	0.81	380	160	45
H558N	✗	✗	NB	NB	NB	NB	NB
T560A	✗	✗	NB	NB	NB	NB	NB
P656L	✗	✗	NB	NB	NB	NB	NB
Q839L	✗	✗	NB	ND	ND	ND	ND
Y841F	✓	✓	**30**	**5**	**400**	**100**	**NB**
K842M	✗	✓	**1.8**	**NB**	**116**	**NB**	**NB**
C917L	✓	✓	16	460	136	100	390
H920S	✗	✓	100	50	140	200	470
T921A	✗	✓	4.5	2.4	330	160	310
H922A	✗	✓	12	NB	NB	NB	NB

Affinities in uM, ND= not determined due to instrument noise, - = no binding, LC=low capture (could be low resolution for binding)

Curve fitting for the SPR data of binding of small molecules to hOGT and the mutants





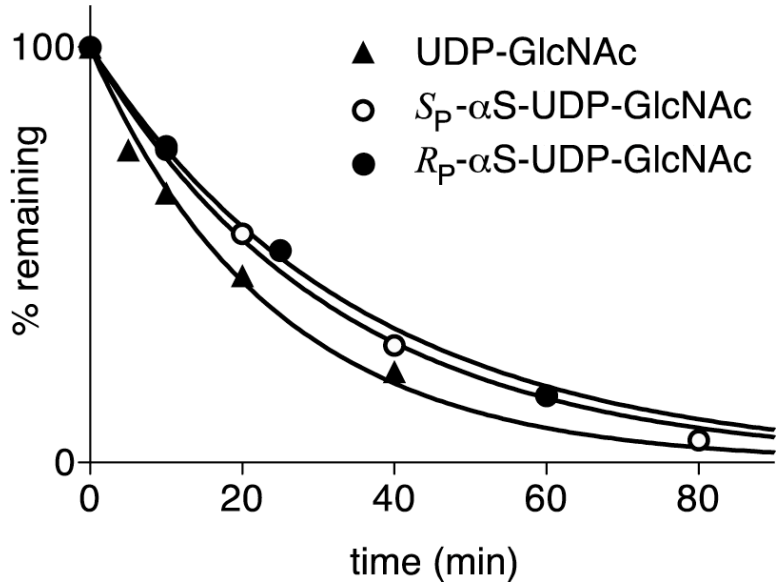

UDP-GlcNAc
UDP
UDP-GalNAc
UDP-Glucose
UMP
WT
H498F
H498N (LC)
D554N
H558N
T560A (LC)

Data collection and refinement statistics of hOGT

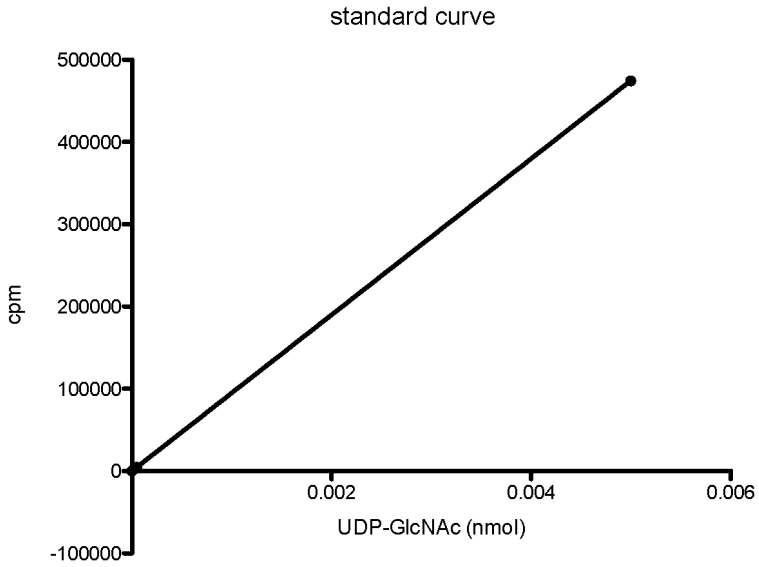
	Product complex	Michaelis complex
	hOGT+UDP+gTAB1tide	hOGT+UDP-5SGlcNAc+aaTAB1tide
Data collection		
Space group	*P*321	*P*321
Cell dimensions		
a = *b*, *c* (Å)	274.2, 142.2	275.4, 142.6
Resolution (Å)	30—3.15 (3.32—3.15) *	40—3.30 (3.39—3.30) *
R_{merge}	0.113 (0.481)	0.131 (0.496)
$I / \sigma I$	6.7 (2.2)	7.7 (1.7)
Completeness (%)	99.7 (99.9)	98.2 (96.9)
Redundancy	3.6 (3.3)	2.2 (2.2)
Refinement		
Resolution (Å)	30—3.15	40—3.30
No. reflections	105108	91074
R_{work} / R_{free}	17.4 / 20.4	22.7 / 27.2
No. atoms		
Protein	22052	22052
Nucleotide(sugar)	96	152
(Glyco)peptide	364	364
B-factors		
Protein	60.32	46.68
Nucleotide(sugar)	62.6	46.5
(Glyco)peptide	89.9	64.31
R.m.s. deviations		
Bond lengths (Å)	0.017	0.010
Bond angles (°)	2.04	1.31

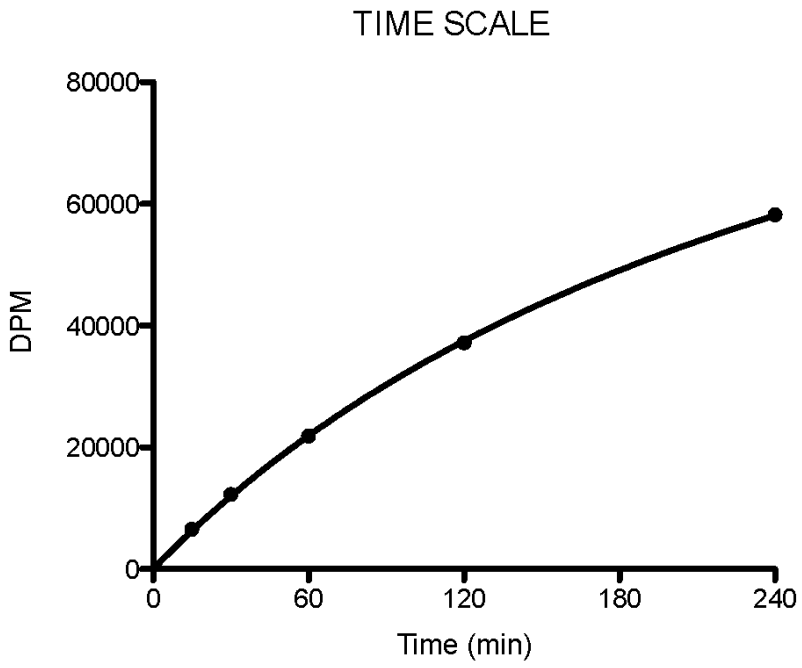
*Highest-resolution shell is shown in parentheses.

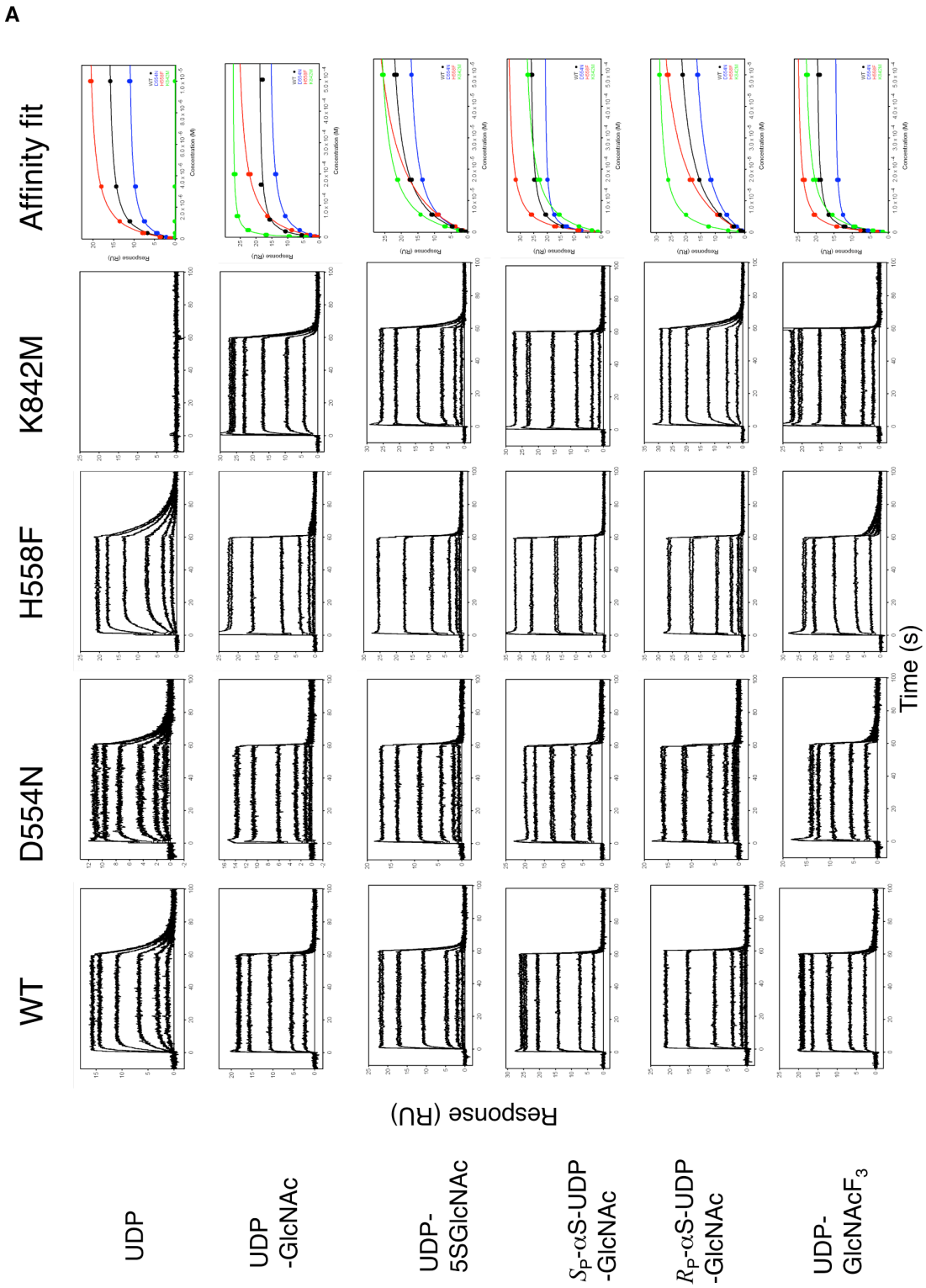
Solvolysis of phosphorothioate analogues of UDP-GlcNAc, reaction conditions could be found in Schimpl et al., 2012.



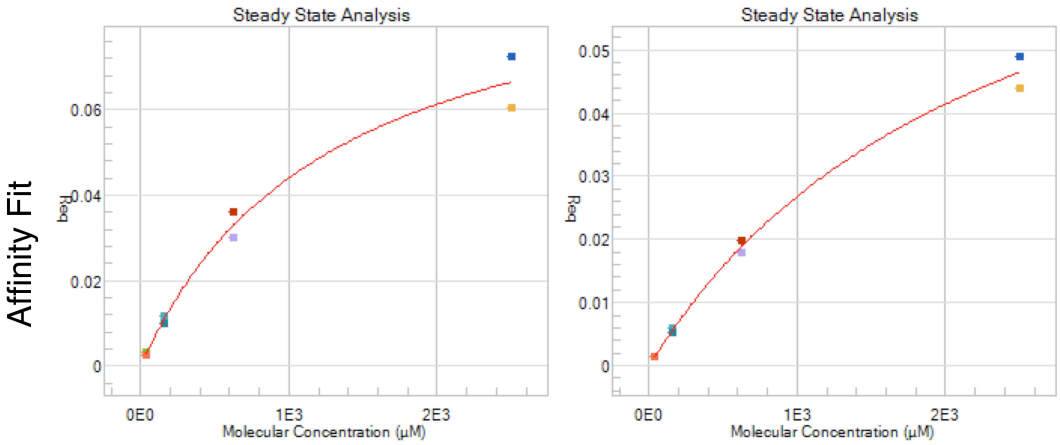
Top: Standard of UDP-[^{3}H]GlcNAc using scintillation counter.
Bottom: Time course of *Dm*OGT *in vitro* reaction.






A
WT
D554N
H558F
K842M
Affinity fit
UDP
UDP -GlcNAc
UDP-5SGlcNAc
S_P-αS-UDP -GlcNAc
R_P-αS-UDP -GlcNAc
UDP-GlcNAcF$_3$
Response (RU)
Time (s)

Curve fitting for



UDP-GlcNAc

R_P-αS-UDP-GlcNAc

Differential scanning fluormetry demonstrate the folding of hOGT and its mutants