Novel OGT Inhibitors Reveal O-GlcNAc Regulates Splicing

Novel OGT Inhibitors Reveal O-GlcNAc Regulates Splicing

A dissertation presented

by

Zhi Wei Tan

to

The Committee on Higher Degrees in Chemical Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Chemical Biology

Harvard University

Cambridge Massachusetts

Nov 2019

Dissertation Advisors: Professor Suzanne Walker Zhi Wei Tan

Novel OGT Inhibitors Reveal O-GlcNAc Regulates Splicing

Abstract

Reversible glycosylation of nuclear and cytoplasmic proteins is an important regulatory mechanism across metazoans. One enzyme, O-linked N-acetylglucosamine transferase (OGT), is responsible for all nucleocytoplasmic protein glycosylation and there is a well-known need for potent, cell-permeable inhibitors to interrogate OGT function. Here we report the structure-based evolution of OGT inhibitors culminating in compounds with low nanomolar inhibitory potency and on-target cellular activity. The structures we report provide insight into how to inhibit glycosyltransferases, a family of enzymes that has been notoriously refractory to inhibitor development.

Intron detention in precursor RNAs serves to regulate expression of a substantial fraction of genes in eukaryotic genomes. How detained intron (DI) splicing is controlled is poorly understood. Here we show that O-GlcNAc, which is thought to integrate signaling pathways as nutrient conditions fluctuate, controls detained intron splicing. Using specific inhibitors of OGT and the enzyme that removes O-GlcNAc (OGA), we first show that O-GlcNAc regulates splicing of the highly conserved detained introns in *OGT* and *OGA* to control mRNA abundance in order to buffer O-GlcNAc changes. We show that *OGT* and *OGA* represent two distinct paradigms for how DI splicing can control gene expression. We also show that when DI splicing of the O-

GlcNAc-cycling genes fails to restore O-GlcNAc homeostasis, there is a global change in detained intron levels. Strikingly, almost all detained introns are spliced more efficiently when O-GlcNAc levels are low, yet other alternative splicing pathways change minimally. Our results demonstrate that O-GlcNAc controls detained intron splicing to tune system-wide gene expression, providing a means to couple nutrient conditions to the cell's transcriptional regime.

Contents

Introduction

1.1.1 Glycosyltransferases catalyze a diverse range of reactions

Glycosyltransferases (GTs) are a large family of enzymes that catalyze the transfer of a mono- or oligosaccharide from a glycosyl donor to an acceptor [1] (Figure 1.1.1a). Glycosyl donors are usually sugar-nucleotides or lipid phosphosugars [2] while acceptors can range from sugars, proteins, lipids, DNA, to diverse secondary metabolites. Among the sugar donors, uridine diphosphate glucose (UDP-Glc), uridine diphosphate N-acetylglucosamine (UDP-GlcNAc), uridine diphosphate N-acetylgalactosamine (UDP-GalNAc), cytidine monophosphate N-acetylneuramic acid (CMP-NeuAc) and guanosine diphosphate mannose (GDP-Mannose) are more commonly used (Figure 1.1.1b). For example, the DNA beta-glucosyltransferase (BGT) uses UDP-Glc to transfer glucose to hydroxymethylcytosine residues in DNA [3], the β-1,4-GlcNAc transferase (MurG) transfers GlcNAc from UDP-GlcNAc onto the C4 hydroxyl of undecaprenyl-pyrophosphoryl-MurNAc-pentapeptide (Lipid I) [4], the polypeptide-α-GalNAc transferase (pp-GAlNAc-T1) transfers GalNAc from UDP-GalNAc to Ser/Thr residues [5] on proteins, the sialyltransferase (CstII) transfer sialic acid from CMP-NeuAc to the C3 hydroxyl of the galactose in lactose from nascent oligosaccharide [6-7] and mannosylglycerate synthase (MGS) utilizes GDP-mannose to catalyze the condensation of GDP-Mannose with D-glycerate to form alpha-mannosyl-D-glycerate [8-9] (Table 1). In addition to the wide diversity of sugar donors and acceptors that GTs accept, transfer of the sugar can also occur with inversion or retention of stereochemistry at the anomeric carbon, depending on the active site of the enzyme [10]. Hence, in terms of functions, glycosyltransferases are considered as one of the largest and most diverse family of enzymes present in nature.

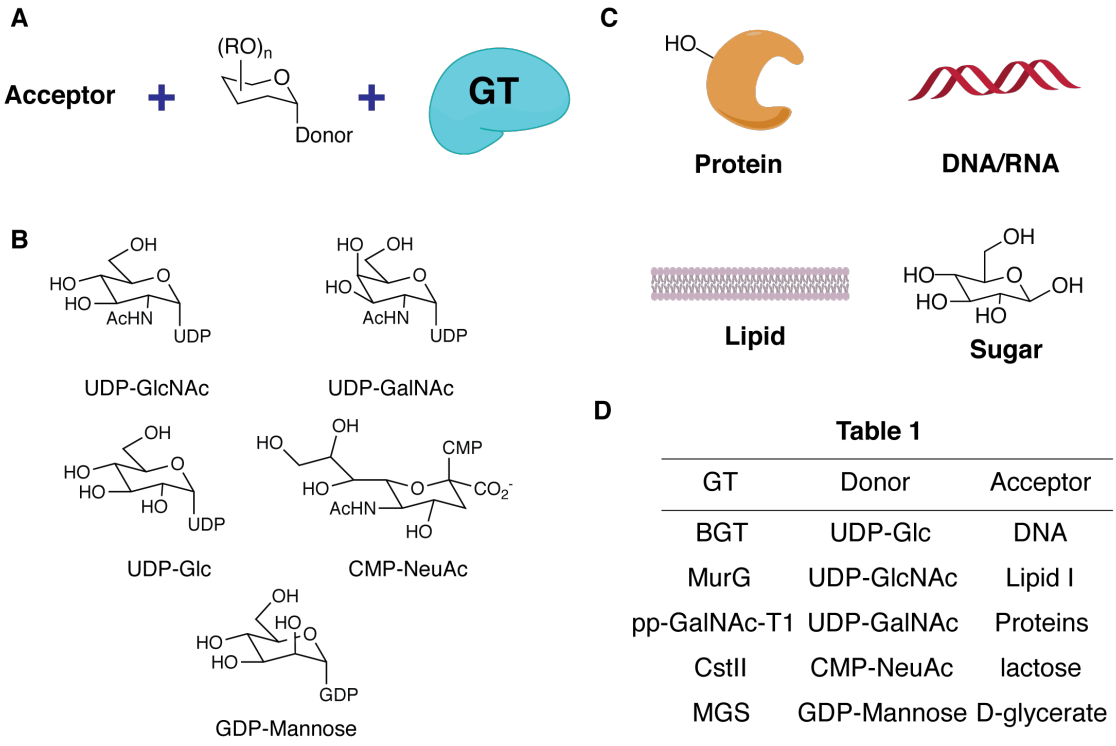


Fig 1.1.1. GTs are functionally diverse. A: Schematic of a GT reaction. B: Common sugar donors of GTs. C: Representative of GTs' acceptors, ranging from protein to sugar acceptors. D: Table 1 showing an example of GT using a different donor.

1.1.2 GTs adopt mainly 3 types of structural folds

In addition to their impressive functional diversity, most GTs fall into one of the three structural classes, or superfamilies: GT-A, GT-B and GT-C [1]. In short, GT-A enzymes require a divalent metal ion for catalysis and consist of two abutting Rossmann-like fold domains [11]. GT-B enzymes are generally metal ion-independent and comprise two Rossmann-like fold domains facing each other [12]. Unlike GT-A and GT-B enzymes which are soluble proteins, GT-C enzymes contain a large hydrophobic integral membrane domain with 8-12 transmembrane helixes [13-14].

The three structural classes differ from each other mainly through the arrangement or absence of Rossman-like domain and catalytic site. GT-A enzymes consist of two tightly associated β/α/β domains that resemble a Rossmann fold and possess an Asp-X-Asp (DXD) signature, which interacts primarily with the phosphate groups of nucleotide through the coordination of the divalent cation. An example of GT-A enzyme is the glycosyltransferase SpsA from *Bacillus subtili* (Figure 1.1.2a) [15]. Unlike GT-A enzymes, GT-B enzymes consist of two separate Rossmann domains connected together and a catalytic site that spans the two domains. Compared to GT-A, where the leaving group is promoted through the interaction of the coordinated divalent cation, GT-B enzymes use positively charged side chains and helix dipoles to stabilize the leaving group and thus do not require metal cation for catalysis. An example of GT-B enzyme is MurG (Figure 1.1.2b) [16]. GT-C enzymes, on the other hand, are very different from either GT-A or GT-B enzymes. GT-C enzymes do not contain any Rossmann-like domains. Instead, GT-C enzymes possess a large hydrophobic integral membrane domain with the active site generally located within a long-loop region. An example of GT-C enzyme is the bacteria oligosaccharyltransferase (PgIB) [17] (Figure 1.1.2c).

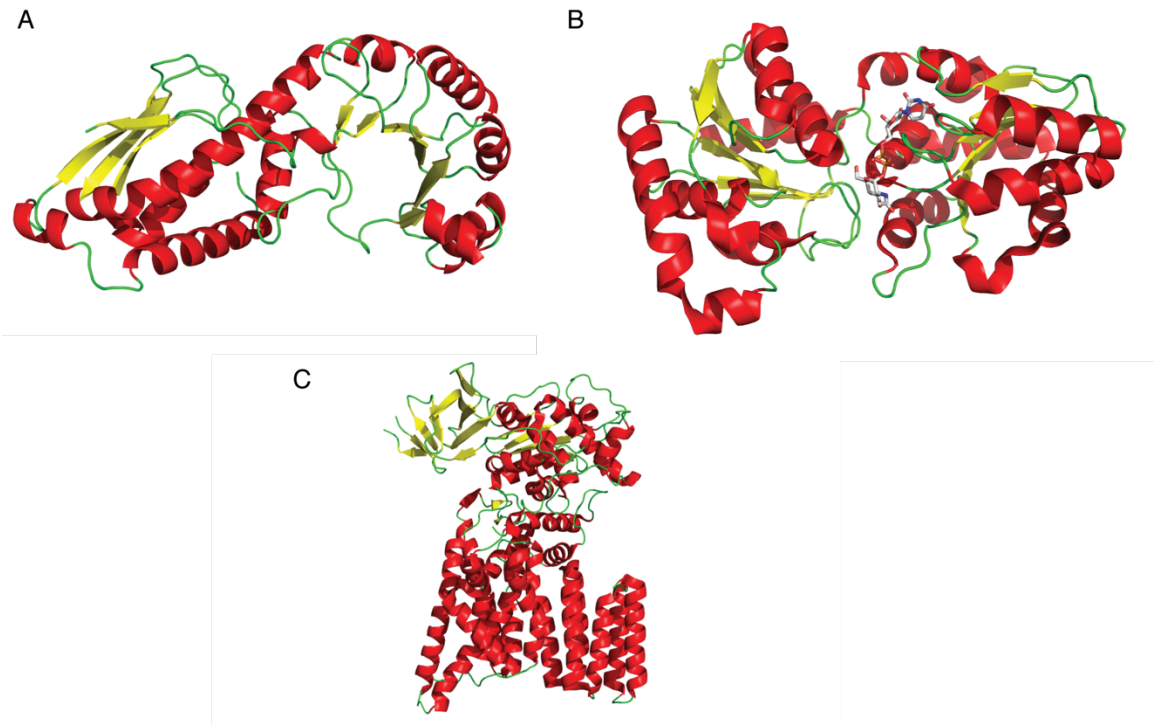


Fig 1.1.2. Ribbon cartoon diagram for 3 different glycosyltransferase. A: GT-A fold, Apo form of SpsA (PDB code 1QPG). B: GT-B fold, MurG bounded with UDP-GlcNAc (PDB code 1NLM). C: GT-C fold, PgIB in complex with magnesium and peptide substrate (not highlighted). (PDB code 3RCE)

1.1.3 GTs are potential drug targets

Because of the diversity of biology that GTs are involved in, there have been increasing interests in developing inhibitors for GTs. GTs play important roles in many fundamental biological processes including cellular adhesion, cell signaling, cellular transport, bacterial and viral infection and transcription [18-21]. Consequently, GTs have emerged as potential drug targets for a number of indications including infectious diseases, cancers, and inflammation of various tissues [22]. For example, β-galactoside α-2,6-sialyltransferase 1 (ST6Gal-I) sialyltransferase was found to be upregulated in numerous cancers and shown to be a functional driver of cancer stem cell phenotype. This suggests that blocking ST6Gal-I function could be a therapeutic option for cancers. Indeed, knockdown of ST6Gal-I was shown to reduce tumor growth in numerous cancers [23-24]. Inhibitors could also be used as biological tools to investigate the functions of GTs.

For example, tunicamycin, which broadly blocks eukaryotic N-Glycosylation by inhibiting UDP-N-acetylglucosamine-dolichol phosphate N-acetylglucosamine-1-phosphate transferase (GPT), have been extensively used to study the role and function of N-glycans and endoplasmic reticulum (ER) stress in mammalian cells [25-26] (Figure 1.1.3).

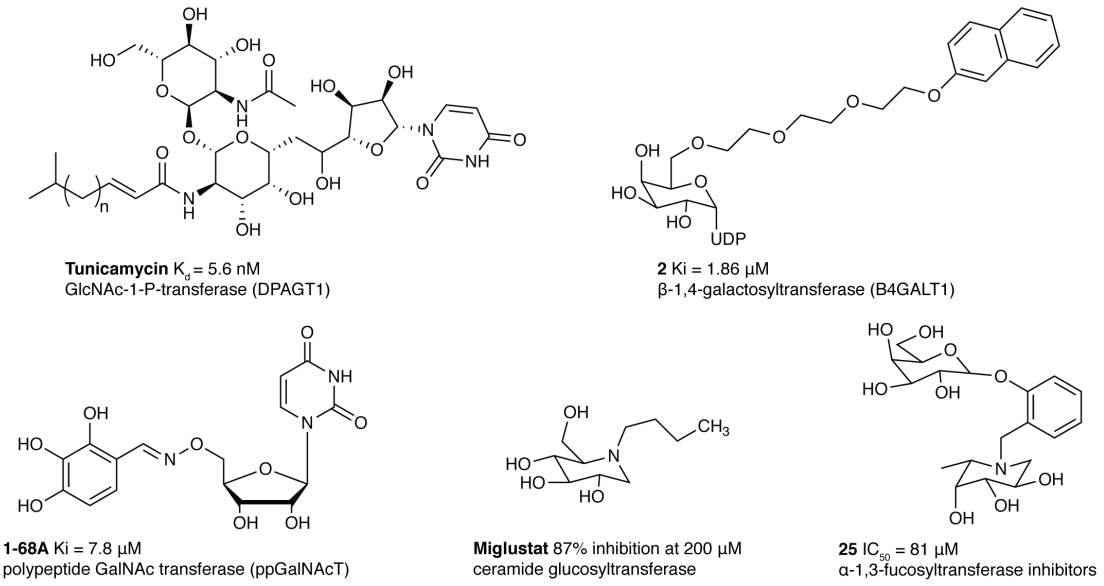


Fig 1.1.3. Examples of substrate-based inhibitors of GTs[27-31]

It is not trivial to develop inhibitors against GTs. Despite a long history of attempts to develop glycosyltransferase inhibitors, many of GTs' inhibitors are not usable in cells. The most common approach has been to develop substrate or bisubstrate analogs, in which the inhibitors are designed based on structural features found in the natural substrate(s) for the GTs [21, 32-37] . For example, many inhibitors made against β-(1-4)-galactosyltransferase, which is involved in the construction of sialyl Lewis X and important for inflammation, come mainly from UDP-GalNAc analogues [28, 38] (Figure 1.1.3). These substrate-based inhibitors have been used as tools for mechanistic studies of GTs with recombinant enzymes, and have provided valuable insights

into the biochemistry of GTs [39]. Through biochemical studies, it has been found that achieving good binding affinities with substrate-based inhibitors requires incorporating a phosphate or diphosphate element that mimics part of the leaving group. However, the inclusion of these moieties limits the cellular application of these inhibitors due to poor cellular penetration of phosphates. Moreover, because numerous GTs in cells share the same sugar-nucleotide substrates, there are concerns about whether substrate-based inhibitors would be adequately selective. These limitations have hampered use of substrate-based inhibitors as templates for further drug development [40-42]. Therefore, non-substrate based inhibitors of GTs, which possess drug-like properties, are highly sought after [43-44].

There are two main challenges that have hindered the development of non-substrate based inhibitors. First, despite the availability of a multitude of crystal structures of GTs, none of these crystal structures were of the GT bound to a non-substrate-based inhibitor, meaning there was no guideline for designing inhibitors before our recent work, which will be discussed in Chapter 2 [45]. Second, in order for GTs to bind to the highly polar and charged sugar-nucleotide substrates, the active sites of most GTs need to be large, highly hydrophilic, and non-hydrophobic (Figure 1.1.4) [46-47]. This characteristic makes developing small molecule inhibitors against GTs particularly challenging because binding of small molecules to proteins is mostly driven by hydrophobic interactions [48].

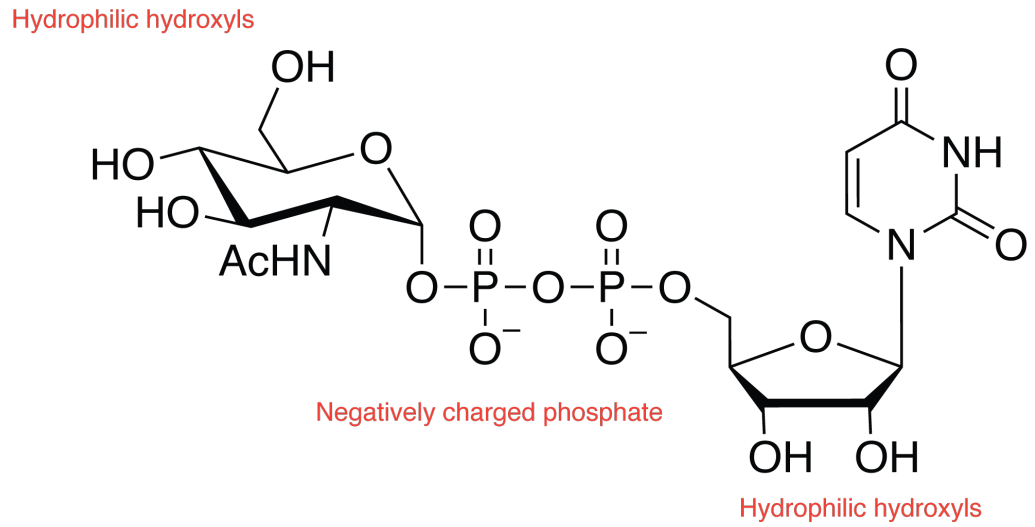


Fig 1.1.4. Representation of nucleotide-sugar donor. Negatively charged and hydrophilic character give rise to charged and hydrophilic binding pocket

1.2.1 OGT is an unusual glycosyltransferase

Among the GTs, O-GlcNAc transferase (OGT) is a unique and non-canonical GT that belongs to the GT-B family [49]. The catalytic region of OGT comprises three domains: the N-terminal domain (N-Cat), the C-terminal domain (C-Cat), and the intervening domain (Int-D), with both the N-Cat and C-Cat domains containing Rossman-like folds that span the catalytic site [50] (Figure 1.2.1). What makes OGT unusual is presence of the Int-D, which is not observed in other GT-B enzymes. The Int-D adopts a unique fold that contains several alpha helixes flanking a seven-stranded beta sheet core. Many of the Int-D residues are surface exposed and the function of Int-D is still unknown.

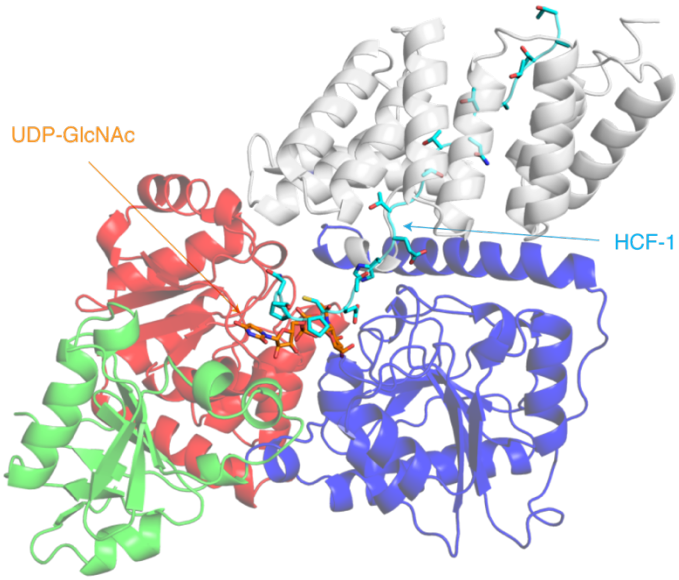


Fig 1.2.1. Overall structure of OGT complex with UDP-GlcNAc and Host cell factor 1 (HCF-1). PDB code: 4N3C. TPR domain is colored in grey, N-cat domain in blue, C-cat in red and Int-D in green.

OGT is also unique in its function. OGT uses UDP-GlcNAc as the donor substrate to catalyze the formation of a single *O*-linked *N*-acetylglucosamine (O-GlcNAcylation) onto cytoplasmic, nuclear, and mitochondria proteins, mainly at Ser and Thr residues (Figure 1.2.2) [51]. Compared with other mammalian GTs, OGT bears no sequence similarity to any described GTs and is the only GT that is known to glycosylate short polypeptide chain. In addition, OGT does not modify or elongate existing O-GlcNAc, nor is the O-GlcNAc moiety further extended or modified by other GTs. This is unlike other glycosylation steps where the glycan chains are normally extended or further modified to form more complex structures.

In mammalian cells, OGT is the only enzyme that can catalyze O-GlcNAcylation in the cytoplasm and nuclear compartment, and until now, there have been more than 1000 substrates that have been reported to be glycosylated by OGT [51-54], a level of promiscuity not yet observed in other GTs. O-GlcNAcylation is a reversible and dynamic process, with this modification being removed by O-GlcNAcase (OGA), which is also the only enzyme in mammalian cells that can

remove O-GlcNAcylation. Both OGT and OGA work together to maintain the dynamic equilibrium of O-GlcNAcylation in cells.

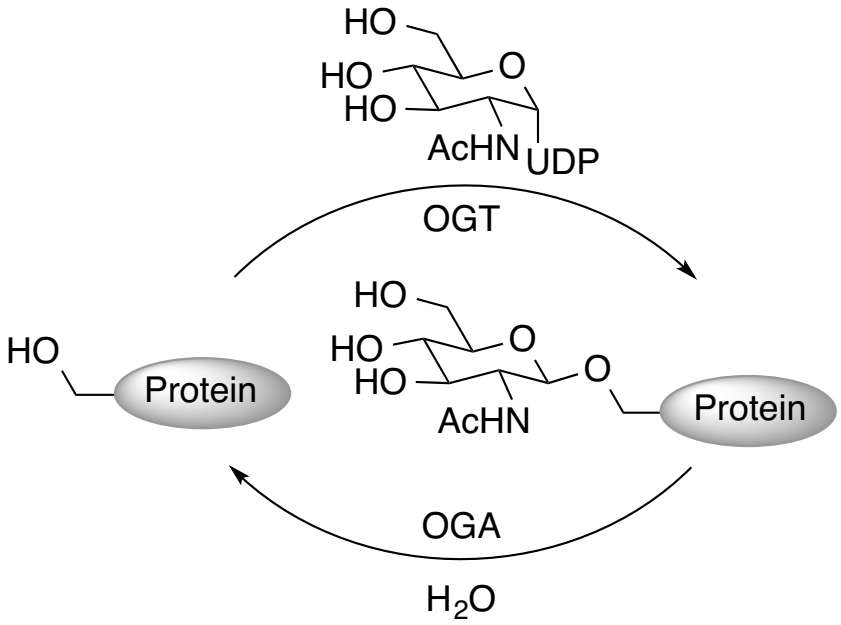


Fig 1.2.2. OGT and OGA reversibly regulates O-GlcNAcylation. OGT uses UDP-GlcNAc as a substrate to add onto Ser/Thr residues on protein and OGA removes the O-GlcNAc modification.

1.2.2 OGT regulates a wide range of biology

Because of the large number of substrates, O-GlcNAcylation has been linked to a multitude of cellular functions [55-57]. Some of the biological functions that O-GlcNAcylation regulates are transcription, epigenetic regulation, metabolism, stress response, dynamic cellular signaling, and cellular differentiation (Table 2). For example, in T_{reg} cells, O-GlcNAcylation of FOXP3, a transcription factor and master regulator of the development and maintenance of Treg cells, was found to increase the stability of FOXP3 possibly through decreased ubiquitination, underscoring the importance of O-GlcNAcylation in immune regulation [58]. Another example is the stress-induced increase of O-GlcNAc on eIF4G, which causes the release of mRNA encoding

stress-response genes from stress granules and pauses global protein translation through the disruption of the closed mRNA loop formed by PABP1-eIF4G in MEF cells [59].

On top of all that, perhaps the most commonly invoked putative function, O-GlcNAcylation is also proposed to be a nutrient sensor in cells [57, 60]. This hypothesis is supported by two observations. First, O-GlcNAc levels respond readily to glucose availability and second, metabolic flux is rewired by O-GlcNAc to maintain glucose levels [61]. For example, it was reported that increased O-GlcNAc levels by OGT overexpression or OGA inhibition was found to redirect glucose flux from the glycolytic pathway to oxidative pentose phosphate pathway through the O-GlcNAcylation of phosphofructokinase1 (PFK1) [62]. Because of the diversity in biological roles that O-GlcNAc takes on, aberrant O-GlcNAcylation has been implicated in a number of diseases including cancer, diabetes and neurodegenerative disorders [63-67].

Table 2

Biological functions of O-GlcNAcylation

Function	Biology	Pathway	Proteins	Ref
Transcription	Immune	T cell activation	NFATc1	[68]
	Immune	T cells maintenance	FOXP3	[58]
	Metabolism	gluconeogenesis	CRTC2	[69]
	Metabolism	Insulin signaling	FOXO1	[70]
Epigenetic	Development	DNA demethylation	TET1	[71]
	Development	Histone methylation	EZH2	[72]
Stress response	Cellular damages recovery	Stress granules	eIF4G	[59]
	Protein folding stress	Unfolded protein response	XBP1	[73]
Dynamic cellular signaling	Nutrient sensing	Cellular growth	AKT1	[65]
	Nutrient sensing	Cellular growth	PFK1	[62]
Cellular differentiation	Development	Neuronal stem cells maintenance	unknown	[74]
	Development	Embryonic stem cells maintenance	OCT4	[75]

1.3.1 OGT inhibitors can provide insights into OGT's diverse functions

While earlier work has revealed some of the fascinating and complex roles that O-GlcNAc regulates, most of the biology discovered so far has come from knockdown or tissue-specific knockout studies, which take between 24-48 h for efficient protein depletion [76]. However, O-GlcNAcylation is a dynamic process and occurs on a time frame of minutes to hours [77-80]. Moreover, because OGT is essential in mammals, OGT knockout organisms are not viable,

and all of the *in vivo* studies have been restricted to tissue specific knockouts of OGT [81-82]. To note, we still do not understand the reason for the essentiality of OGT and OGA in mammals and the non-essentiality of OGA in cells [83-85]. Hence, inhibitors are required to further uncover the biological roles of O-GlcNAcylation. To further dissect and decipher the biology of O-GlcNAcylation, chemical probes that can rapidly and specifically affect the level of O-GlcNAcylation in cells and *in vivo* will be needed.

There is a need for specific and acute OGT inhibitors in the field. There have been a number of reported inhibitors for OGA, and their application has provided valuable insights into the biology of increased O-GlcNAcylation [86-88]. Yet those inhibitors have also yielded results that are different from genetic approaches because both OGT and OGA can have non-catalytic and catalytic roles [89]. For example, OGT was shown to be part of the HCF-1/VP16/OCT transactivation complex in which OGT most probably functions as a scaffolding protein in that complex [90]. Because of the important non-catalytic roles of OGT, knockdown or knockout studies have the potential to convolute those roles with the effects that are mediated by O-GlcNAc. Hence, inhibitors that are fast acting, potent, and specific for OGT will be valuable for further elucidating the biological role of OGT.

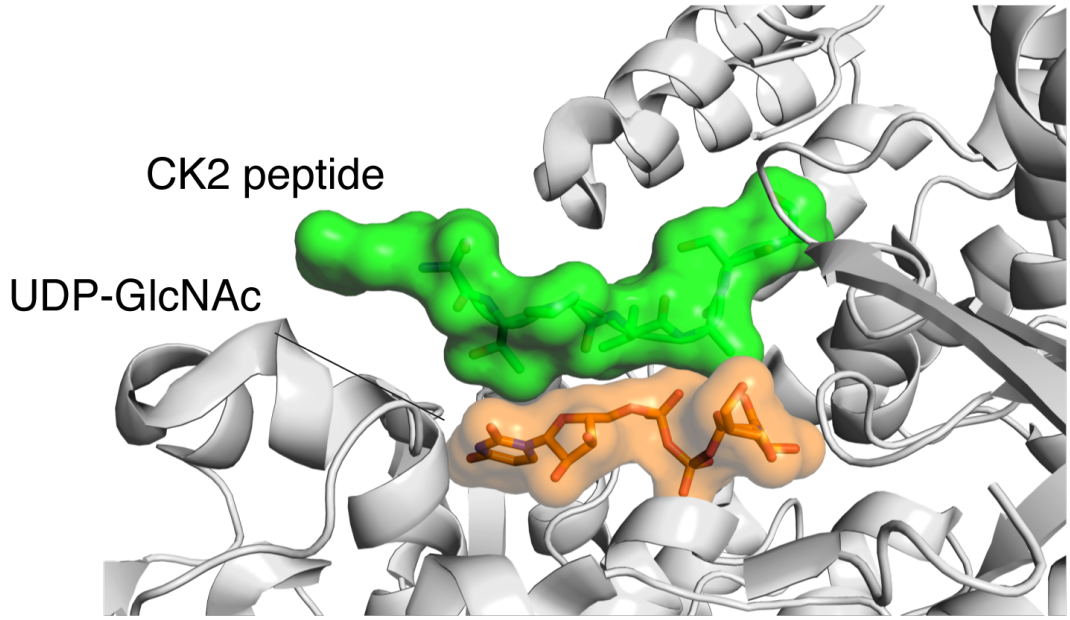


Fig 1.3.1. Ribbon representation of OGT and its substrates CK2 and UDP-GlcNAc. Stacked conformation between UDP-GlcNAc and CK2 needs to be mimicked for inhibitors to be effective. Because of how CK2 peptide is stacked above UDP-GlcNAc (facing the open solvent), substrate selectivity is not determined by the active site contacts.

As mentioned above, however, developing inhibitors for OGT, a glycosyltransferase, is not trivial. The active site of OGT is particularly challenging to inhibit. The nucleotide-sugar substrate, UDP-GlcNAc, lies directly underneath the peptide that is glycosylated, so filling the active site requires molecules that can mimic this stacked geometry (Figure 1.3.1) [50, 91]. Complicating matters further, OGT's active site is hydrophilic and accommodates a wide range of peptide sequences, with substrate selectivity being determined not by specific contacts to OGT side chains, but by binding of proteins to the tetratricopeptide repeat (TPR) domain distal from the active site itself [92-93]. Hence, most inhibitors of OGT, like many other GTs, are designed based on the sugar-nucleotide or protein/peptide substrates of OGT.

1.3.2 Reported OGT inhibitors are not suitable for cellular studies

Tremendous efforts have been made to develop OGT inhibitors but they have not been very successful. The first known inhibitor of OGT is the reaction product, UDP, which inhibits OGT with an *in vitro* IC_{50} of 1.8 μM (Figure 1.3.2a) [94]. Still, UDP is unsuitable for use in cells because UDP does not penetrate into cells, and even if it did, UDP would inhibit a large number of GTs that use UDP-sugars as substrates as well as interfere with the cell's natural feedback systems that regulate UDP levels [95]. Still, this didn't stop efforts to develop a UDP-inspired inhibitor. The first reported cell penetrable inhibitor of OGT is an uracil analogue called alloxan [96]. As a uracil analogue, alloxan lacks specificity for OGT and so its effects in cells cannot be assumed to report on OGT. Furthermore, alloxan displayed off-target toxicity in cells through forming reactive oxygen species like superoxide radicals [97].

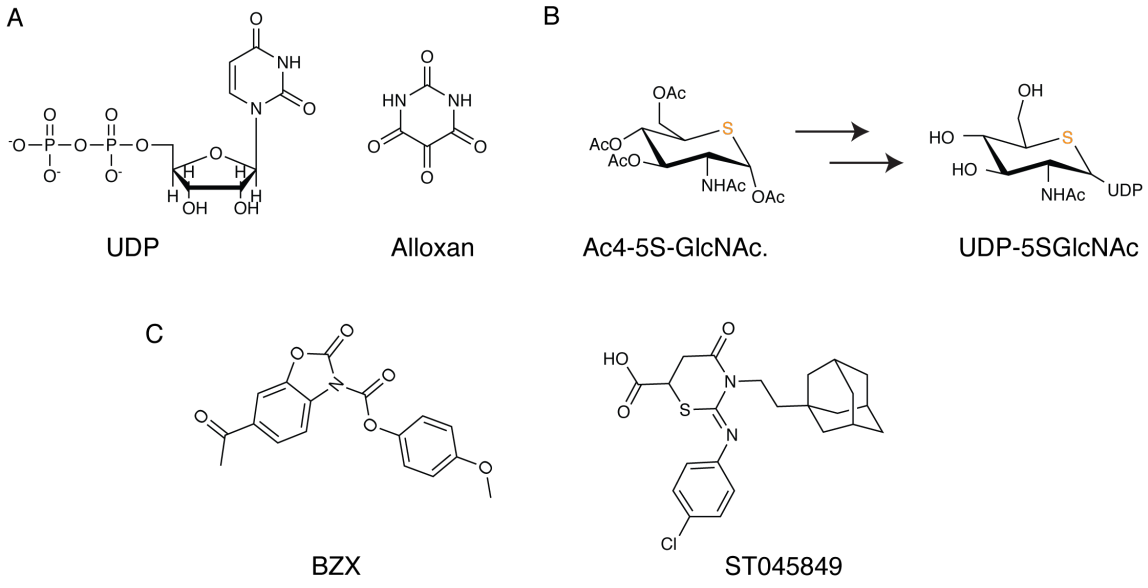


Fig 1.3.2. Structures of OGT inhibitors. A: Structures of early OGT inhibitors, UDP and Alloxan. B: Ac_4-5S-GlcNAc is converted enzymically in cells into UDP-5SGlcNAc, which inhibit OGT at micromolar potency. C: Structures of non-substrate inhibitor of OGT

Some successes have been achieved by employing strategies commonly used to develop inhibitors for GTs. For example, Ac_4-5S-GlcNAc, in which the endocyclic ring oxygen atom of GlcNAc is replaced with a sulphur atom, hijacks the hexosamine pathway to form UDP-5SGlcNAc, a donor substrate mimic that inhibits OGT at micromolar potency *in vitro* and is comparable to the binding affinity of native UDP-GlcNAc substrate (Figure 1.3.2b) [98]. The fact that Ac_4-5S-GlcNAc works is surprising because most cells contains on average ~100 μM of UDP-GlcNAc, which is ~10 fold higher than the potency of UDP-5SGlcNAc [99]. We believe Ac_4-5S-GlcNAc can inhibit OGT in cells because when Ac_4-5S-GlcNAc is converted to the active form UDP-5SGlcNAc, UDP-5SGlcNAc would accumulate to very high level as the negatively charged phosphate group of UDP-5SGlcNAc would prevent the UDP-5SGlcNAc from escaping outside the cells. Ac_4-5S-GlcNAc has been used successfully, in combination with genetic knockdown, in a number of studies to probe the effect of decreased O-GlcNAcylation [100-102]. For example, Ac_4-5S-GlcNAc was used to confirm that lowering O-GlcNAc levels by either OGT knockdown or inhibition increases AMPK activation and decreases mTORC activation, which resulted in a lower flux through the glycolytic pathway [66].

However, substrate-based inhibitors have the same problems as any other inhibitors that are modelled after the native substrate. We identified three problems associated with Ac_4-5S-GlcNAc. First, metabolic incorporation via the hexosamine biosynthetic pathway by Ac_4-5S-GlcNAc will reduce the intracellular pool of UDP-GlcNAc, which is also synthesized through the same hexosamine pathway. Second, UDP-5S-GlcNAc, which is a substrate analogue of UDP-GlcNAc, will inhibit other GTs that use UDP-GlcNAc as a substrate [103]. Third, intracellular

epimerases that convert UDP-GlcNAc to UDP-GalNAc will convert UDP-5SGlcNAc to UDP-5SGalNAc, thus perturbing UDP-GalNAc pools and inhibiting enzymes that use UDP-GalNAc. Hence, Ac_4-5S-GlcNAc is not an ideal inhibitor of OGT, and the usage of Ac_4-5S-GlcNAc can potentially produce confounding biological effects.

In addition to substrate-based inhibitors, significant efforts have been made to develop non-substrate inhibitors of OGT (Figure 1.3.2c). Due to the nature of the active site of OGT that makes it arduous to design inhibitors, a high throughput screen using a library of drug-like compounds was performed to find compounds that can displace UDP-GlcNAc from OGT's active site [104]. This screen uncovered two non-substrate-like inhibitors, BZX and ST045849, that inhibit OGT at micromolar potency as well as a compound that carries a quinolinone-6-sulfonamide (Q6S) core, which culminated in OSMI-1. Details of OSMI-1 will be discussed in the latter part of this thesis. BZX functions as a covalent binder of OGT and blocks the active site of OGT through crosslinking active site residues K842 and C917 [105]. Although the mechanism of action of this compound is interesting, it is not useful in cells because it reacts too rapidly with adventitious nucleophiles. The binding mode of ST045849 is still unknown, but this compound was shown to displace UDP-GlcNAc from OGT's active site. Although both BZX and ST045849 displayed cellular potency against OGT and have also been used in a number of cellular studies, they are both toxic in cells and are not selective towards OGT. The lack of specificity of BZX and ST045849 casts serious doubts on the validity of the cellular readout with usage of BZX or ST045849. Considering all these factors, there is a need to develop non-toxic, potent and cell permeable inhibitors of OGT and will be part of the focus in this dissertation.

1.4 Focus of this thesis

Advances have been made to understand the biology of OGT and O-GlcNAcylation. However there remains a need to develop potent, rapid, and cell permeable inhibitors of OGT to serve as probes for further dissecting the biology of OGT. What other novel strategies can we employ to develop inhibitors for OGT that can be used in cells? Considering that there are no binding assays to quantify the binding affinity of OGT's inhibitors, how can we fully characterize the inhibitors *in vitro* and in cells? Because cellular responses are typically fast acting and O-GlcNAcylation is a rapid process, we still do not know those responses when O-GlcNAc levels are rapidly perturbed. In this dissertation, I will attempt to address those questions.

Chapter 2

Contributions: The chapter is mainly a collaboration between Sara Martin and myself. We are both co-first authors on the paper that came out from this chapter. She performed crystallization of OGT:inhibitors and, together with collaborators from the National Institute of Health (NIH), synthesized the inhibitors. I performed molecular docking, which led to the next generation of OGT inhibitors, and the *in vitro* and cellular assays to determine the potency of the compounds. Global energy minimization was performed in a collaboration with Lisa Töerk and myself.

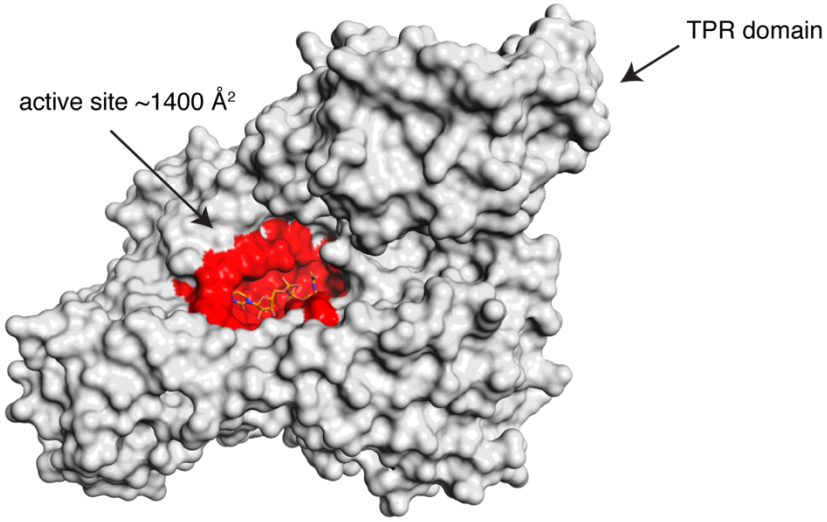
2.1 Introduction

OGT is an essential mammalian enzyme that modifies myriad nuclear and cytoplasmic proteins with O-linked N-acetylglucosamine (O-GlcNAc), affecting their stability, localization, activity, and interactions with other proteins [49]. Evidence points to a crucial role for O-GlcNAc in metabolic homeostasis, and elevated O-GlcNAc levels have been linked to metabolic adaptations associated with several disease phenotypes, including the abnormal proliferative capacity of cancer cells [57]. To better understand OGT function, small molecule OGT inhibitors are required. OGT inhibitors with some cellular activity have been reported, but most are substrate analogs that offer limited opportunities for modifications to improve potency or selectivity [106-108]. This chapter will focus on the strategy used to develop new generation of OGT inhibitors.

2.2 Probing OGT's active site with SiteMap highlights the difficulty in inhibiting OGT

We wondered if we can develop non-substrate binders for OGT's active site. To provide further insights into the active site of OGT, we performed SiteMap [47], which calculates the druggability of the active site through quantifying the size, hydrophobicity and hydrophilicity of a pocket, on OGT's active site where UDP-GlcNAc resides (Figure 2.2). UDP-GlcNAc's binding site was chosen because OSMI-1 was hypothesized to bind in the same binding pocket as UDP-GlcNAc. We inferred this from our earlier findings that OSMI-1, discovered from a high throughput screen that utilizes the displacement of a fluorescently labeled UDP-GlcNAc, was a competitive inhibitor of OGT [103]. We found that the active site of OGT has a high hydrophilicity score (1.49) and very low hydrophobicity score (0.17), consistent with the notion that GTs'

pockets tend to be more hydrophilic in order to bind nucleotide-sugars. Because of the highly unfavorable characteristics of the active site, we expected the pocket to be classified as "undruggable" because most small molecule requires a hydrophobic cleft for binding. Fortunately, SiteMap characterize the pocket to be druggable, albeit in the "difficult" category. We reasoned that OGT is still considered druggable because the active site has an attractive binding pocket of substantial size and good enclosure, and that compensates for the other unfavorable properties. This classification illustrates the difficulty of designing inhibitors against OGT, as is consistent with the literature. From the druggability classification, we concluded that we can further improve the binding of OSMI-1.



Calculated Druggability with SiteMap

Dscore[a]	size	hydrophilic	hydrophobic
0.94	119	1.49	0.17

[a]Dscore < 1.0 sites belong to the difficult category

Fig 2.2. Sitemap of OGT's active site. The active site of OGT is large and hydrophilic as calculated by SiteMap,[109] making it hard to design inhibitors for OGT.

2.3 Molecular docking of OGT inhibitor:OSMI-1 gave rise to OSMI-2

To further improve the binding of OSMI-1, we needed a molecular picture on how OSMI-1 binds to OGT. OSMI-1 was first discovered through a combined effort between our lab and the NIH. The original high throughput screen mentioned in chapter 1 resulted in weak binders of OGT that carry a Q6S core. A biased screen of a library of compounds carrying the Q6S scaffold was then performed, and hits from the biased screen were further optimized by the NIH using standard medicinal chemistry. These enabled us to define the structural activity relationship (SAR) of the Q6S binders that culminated in OSMI-1 (Figure 2.3.1). In short, we found that the Q6S binders require a disubstituted amide and a phenyl glycine and prefer R stereochemistry. Despite the SAR analysis, we were not able to further improve the potency of OSMI-1, which inhibits OGT with an IC_{50} of 2.7 μM and globally reduces O-GlcNAcylation in a number of cell lines. To enable further improvement of OSMI-1, we focused on obtaining a crystal structure of OGT bound to OSMI-1. After trying a number of screening conditions to obtain crystals of OGT bound with OSMI-1 (data not shown here), we were not able to obtain any suitable crystals, most likely due to the insolubility of OSMI-1 in the absence of DMSO. Using DMSO would not work because OGT does not crystalize well in the presence of DMSO.

To obtain a binding structure of OSMI-1 to OGT, we turned to molecular docking. We reasoned that molecular docking, which predicts the orientation of a ligand in a pocket by calculating the binding energy between the ligand and the amino side chains in the binding pocket [110], could be an alternative and parallel approach to access the binding of OSMI-1 to OGT. Moreover, we could utilize the SAR data that we had previously obtained from Q6S binders to cross check the docked pose of OSMI-1.

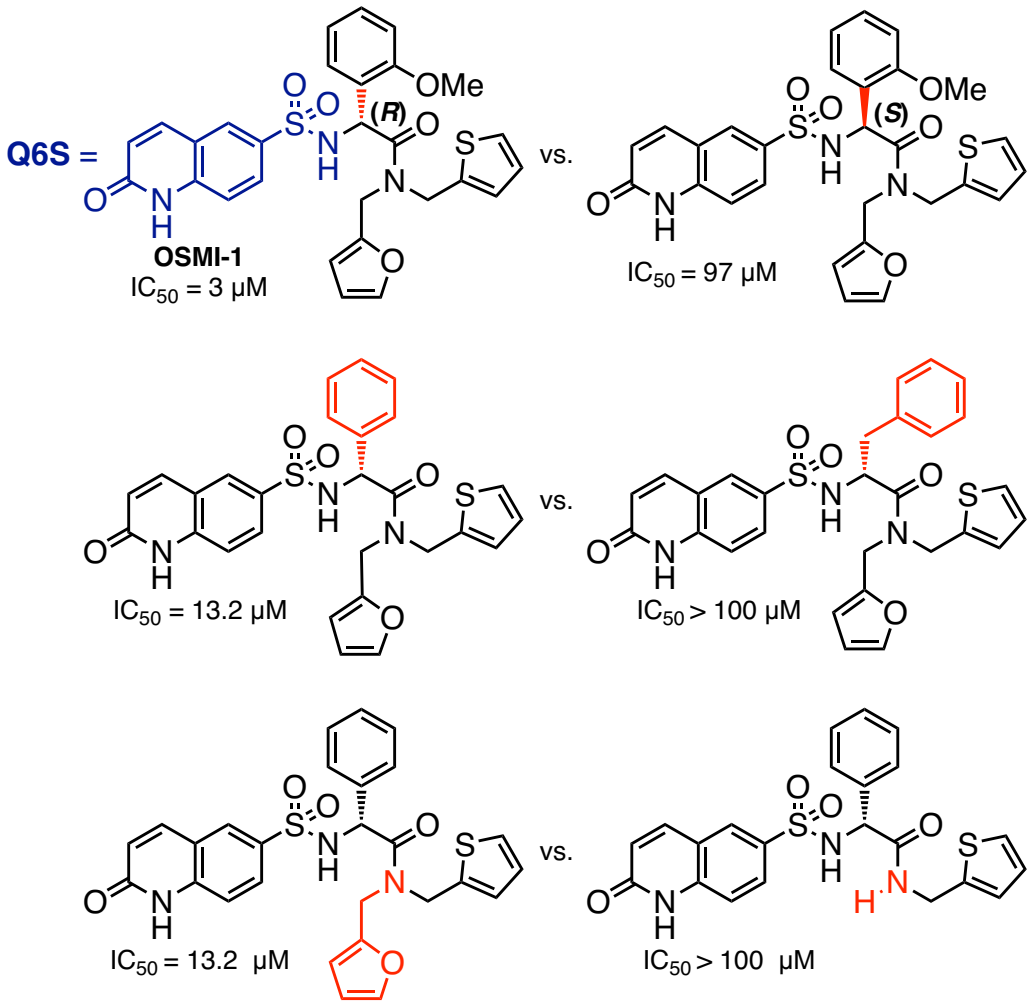


Fig 2.3.1 SAR of OSMI-1. OSMI-1 prefers R stereochemistry, requires phenylglycine and disubstituted amide. IC_{50} was determined with UDP-Glo™ assay from Promega using CK2 peptide as the acceptor substrate.

To obtain a reasonable starting binding model, molecular docking of OSMI-1 was performed on OGT with Schrodinger Maestro suit [111-113]. Since OSMI-1 was discovered from the displacement of UDP-GlcNAc from the active site of OGT, we expected that OSMI-1 would bind in the UDP-GlcNAc's binding pocket. We first tried standard docking protocols with Glide XP, which uses the most sophisticated scoring function in Schrodinger [114]. We first docked UDP-GlcNAc into OGT to validate the GlideXP method. We were able to redock UDP-GlcNAc successfully into the active site, suggesting GlideXP as a suitable docking method for OGT. But

we were not able to find any docked pose of OSMI-1 that is consistent with the previously described SAR (Fig 2.3.1). For example, the Q6S does not overlap with uridine from UDP-GlcNAc as we have expected. Because GlideXP is a widely use protocol for molecular docking, we believe that the failure to find the right pose for OSMI-1 comes from the rigid nature of the amino acids in the binding pocket during the docking process. Understanding that the active site of a protein complex is flexible and changes according to the bound ligand, a more computational exhaustive protocol called Schrodinger's induced fit docking (IFD) was explored. IFD allows the movement of protein side chains during the docking process but is also much more computational expensive than Glide XP and has to be performed with a computational cluster.

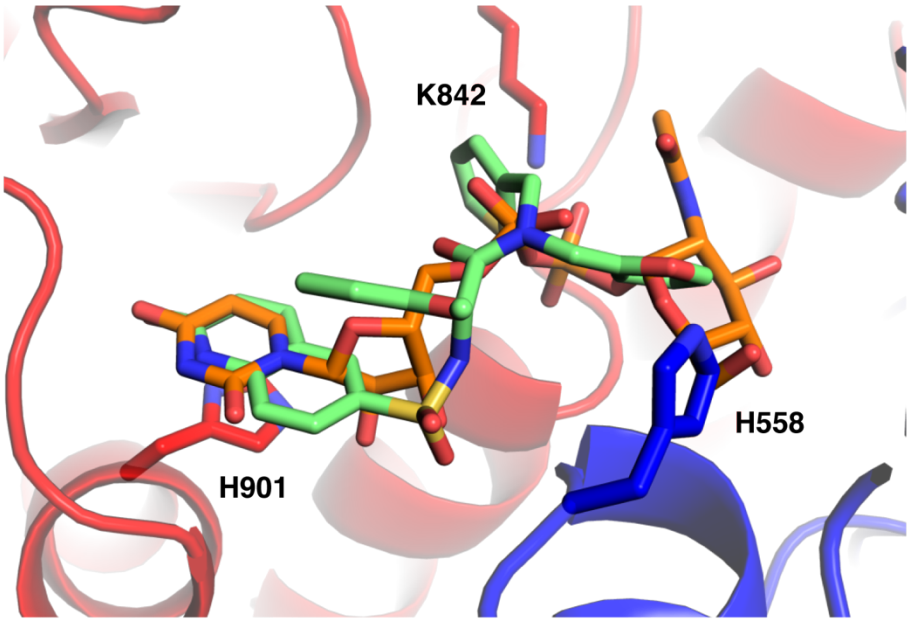


Fig 2.3.2 OSMI-1 binds as predicted. Overlay of IFD docked OSMI-1 (green) with UDP-GlcNAc (orange, PDB code: 4N3C). Disubstituted amide are needed for OSMI-1; with furan forming a π-π interaction with H558 and thiophene forming π-charged interaction with K842. Q6S mimics uridine.

We were able to find binding poses of OSMI-1 that agree well with our SAR data. The lowest energy docked pose of OGT:OSMI-1 showed that OSMI-1 binds in a manner that is consistent with the SAR (Figure 2.3.2). First, the Q6S core mimicked uridine, as seen from the π stacking interaction observed between quinolinone and His^{901}. Second, the π-charged interaction of thiophene and Lys^{842}, which also interacts with the phosphates of UDP-GlcNAc, and π-π interaction between furan and His^{558} supported the requirement of disubstituted amide. Third, the phenylglycine that made contacts with Phe^{878} would have steric clash with Phe^{878} if the phenylglycine were to be swapped with a longer chain phenylalanine. The agreement with SAR confirms the binding mode of OSMI-1 and gives us confidence to use this docked OSMI-1 as a guide to further optimize our inhibitor.

We next made modifications to OSMI-1 to increase its solubility and potency. The reason for wanting a more soluble analog comes from our crystallization observation that OGT does not crystalize well even in the presence of minute DMSO, which OSMI-1 requires for solubility. Because both the Q6S core and the phenylglycine moiety have been shown to be important for the potency of OSMI-1, we focused on the portion of the disubstituted amide and wonder if we can replace the thiophene or furan group, while still maintaining its overall molecular size. Molecular size is important for cell permeability and OSMI-1 already has large molecular size of ~500 Da. Upon examining the docked OGT:OSMI-1 structure, we believe that we can replace the thiophene group, which forms a π-charged interaction with Lys^{842}, with a carboxylic acid group. This substitution will allow us to meet the criteria that we have set. First, the negatively charged carboxylate group would be able to form a charged-charged interaction with Lys^{842}, which we believe was important for the binding of OSMI-1. Next, the carboxylate group will greatly increase OSMI-1 solubility due to the negative charge from carboxylate ion at

physiological pH. Third, this new OSMI-1 analogue would decrease the molecular size by ~11 Da. Because furan and thiophene are interchangeable, i.e. either furan or thiophene can interact with Lys[842] from our docking model, and our SAR data suggests that thiophene is more preferred over furan, we decided to replace the furan instead. This new OSMI-1 analogue is termed OSMI-2 acid (OSMI-2a) (Figure 2.3.3).

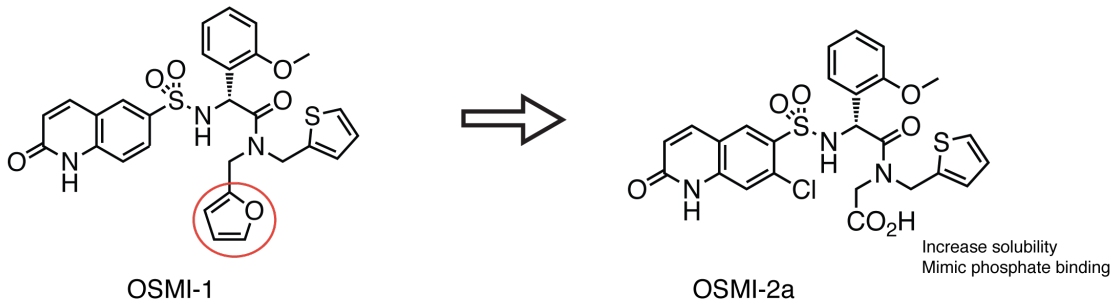


Fig 2.3.3 Design of OSMI-2a based from docked OSMI-1. Swapping furan for carboxylic acid would increase solubility and possibly still maintain contacts with Lys[842]

2.4 Structures of OGT:OSMI-2a revealed close overlap of uridine and the Q6S core

OSMI-2a was synthesized and the potency was determined by UDP-Glo™ assay, which measures the conversion of UDP-GlcNAc to UDP in the presence of a GT that uses UDP-GlcNAc as a substrate. Consistent with our prediction, OSMI-2a showed comparable potency compared to OSMI-1 with an IC_{50} of ~1 μM (data not shown). We also achieved a much higher solubility of this analogue as compared to OSMI-1; OSMI-2a is soluble in aqueous buffer at pH 8.0 even without DMSO. This increased solubility allowed us to obtain the first crystal structure of OGT bounded to an analogue containing a Q6S core: OGT:OSMI-2a (Figure 2.4.1a).

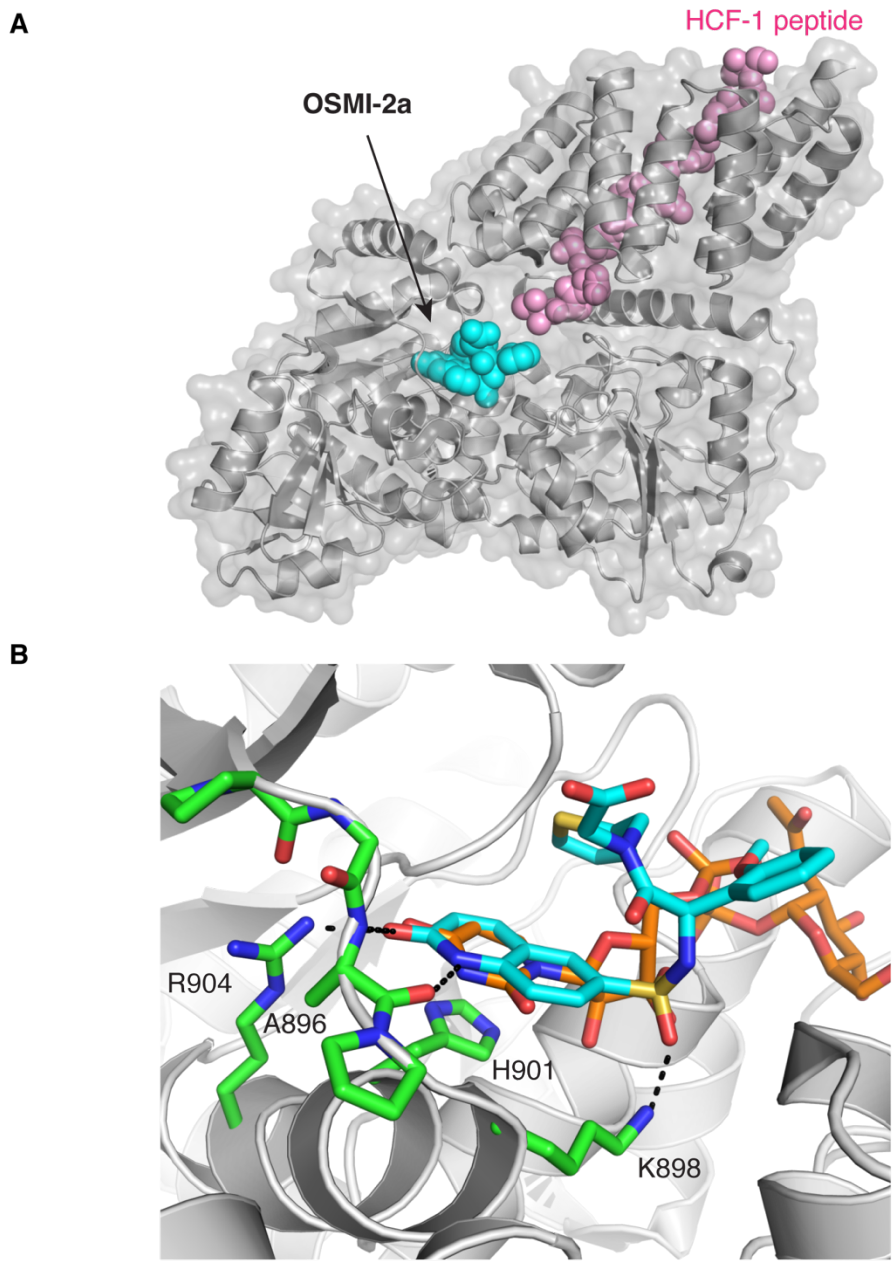


Fig 2.4.1 Structure of OGT:OSMI-2a complex. A: Overview of OGT:OSMI-2a structure (gray) showing OSMI-2a (cyan) bound in the active site. All crystals were obtained using a TPR-binding peptide derived from HCF-1 (pink) to improve resolution. B: Overlay of OSMI-2a (cyan) and UDP-GlcNAc (orange, PDB:4N3C) showed that Q6S mimics uridine.

The structure of OGT:OSMI-2a revealed that Q6S in a faithful mimic of uridine (Figure 2.4.1b). Like uracil, the quinolinone ring stacks directly over the imidazole of His^{901}; moreover, the nitrogen and adjacent carbonyl of the heterocycle make the same contacts to Arg^{904} and

Ala^{896} as N3 and O4 of uracil. In addition, one of the sulfonamide oxygens forms an apparent hydrogen bond with the Lys^{898} side chain, mimicking an interaction between the ribose hydroxyls and that same residue. The remarkable overlap between the quinolinone and uridine suggests this structural motif may serve as a privileged fragment for designing inhibitors against other GTs that use similar nucleotide-sugar substrates.

2.5 OSMI-2a architecture complements OGT's active pocket

The structure showed that OSMI-2a has a U-shaped architecture that helps to explain ability of OSMI-2a to inhibit OGT. The S-N bond veers up from the plane of the quinolinone ring and the backbone of the molecule then folds back over it, positioning the substituents on the disubstituted amide directly over the quinolinone (Figure 2.5.1a). The inhibitor's U-shape allows it to fully occupy a space that accommodates the uridine and the segment of peptide that lies over it. Indeed, the thiophene substituent on the disubstituted amide penetrates so deeply into the active site that Gln^{839}, Leu^{866}, and Phe^{868} must rotate to make room (Figure 2.5.1b). Because this U-shape structure is unusual, we wonder if there is an entropic penalty for OSMI-2a binding when it adopts this U-shape conformation. To address that, we performed density functional theory on OSMI-2a to evaluate the most stable conformers of OSMI-2a. To our surprise, density functional theory calculations show that the conformer observed in the crystal structures is also the most stable conformer in aqueous solvent. This may explain why OSMI-2a inhibits OGT at micromolar potency despite not filling up the whole UDP-GlcNAc binding site. We believe the presence of sulfonamide in the OSMI series helps confer this unusual U-shape conformation and enables OSMI-2a to jam into OGT, complementing the shape of the pocket (Figure 2.5.2).

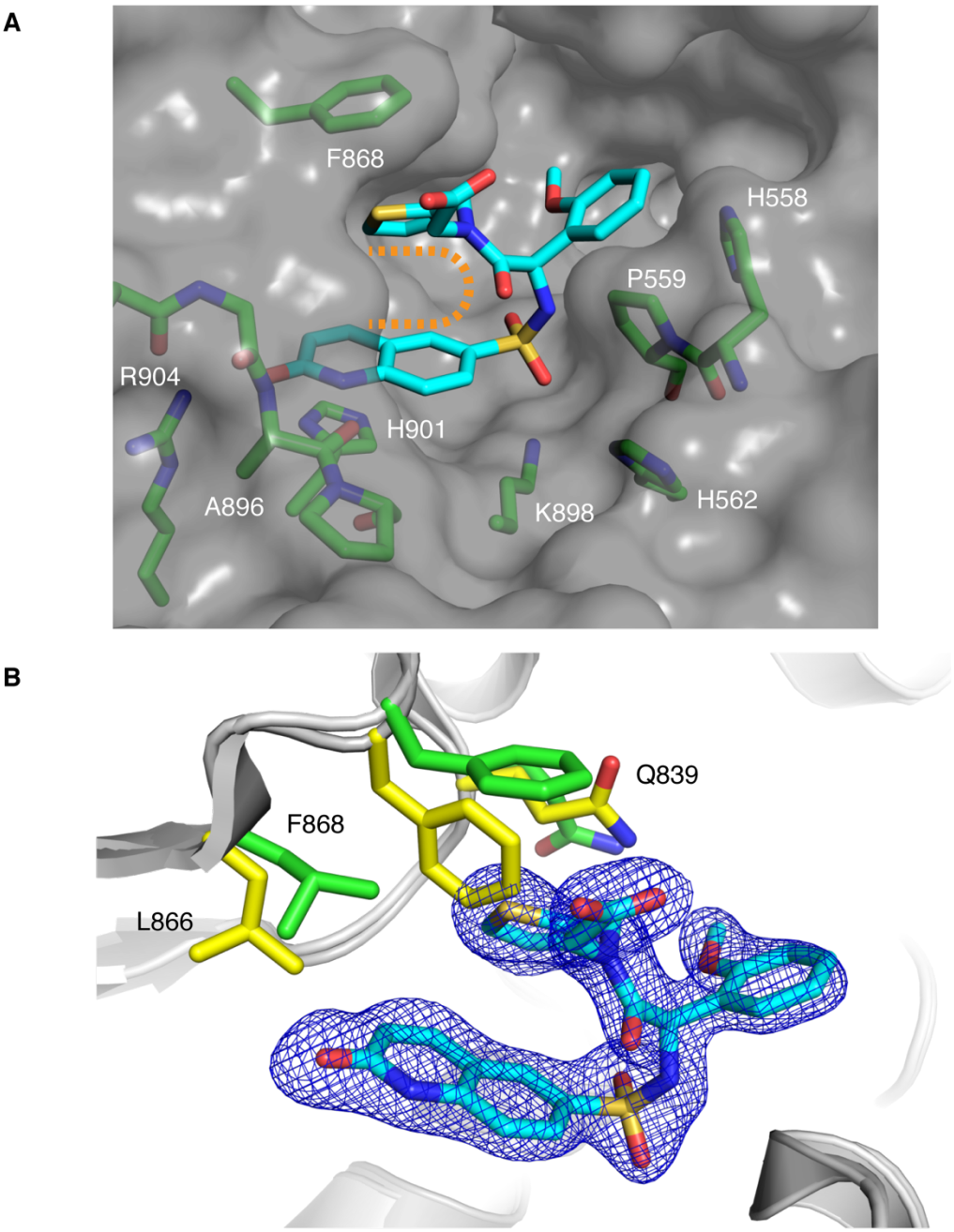


Fig 2.5.1 Structure of OGT:OSMI-2a complex showing the U-shape conformation. A: The U-shape conformation of OSMI-2a allows the amide substituents to fill the space above quinolinone. B: Positions of those side chains in an OGT:UDP-GlcNAc cocrystal are shown in yellow (PDB: 4N3C). The difference density map, F0-Fc, of OSMI-2a is shown in blue mesh, contoured to 3σ**.** A 90° rotation shows an alternate view of the crystal.

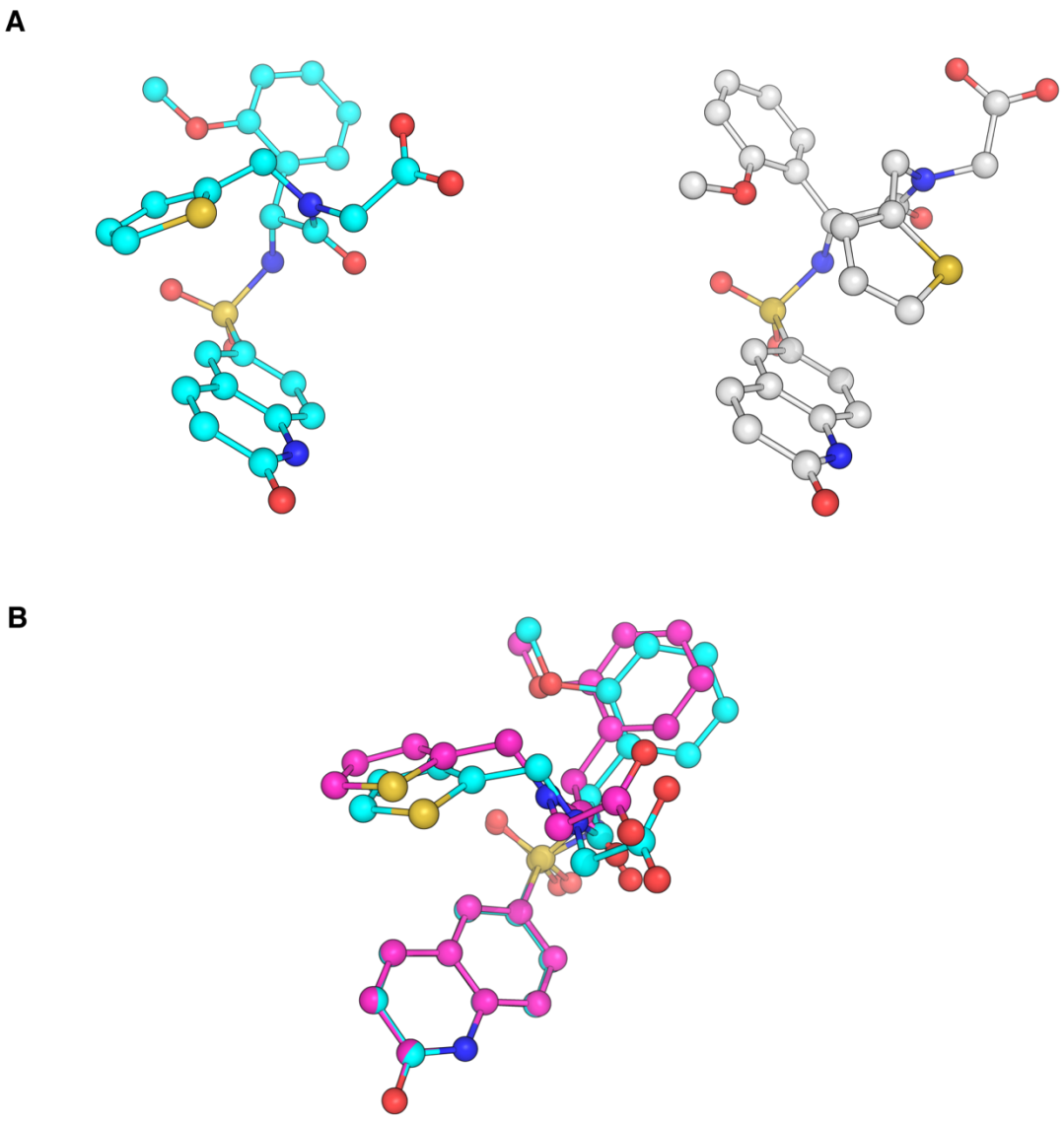


Fig 2.5.2 The most stable conformer of OSMI-2a matches that observed in the OGT:OSMI-2a structure. (**A**) The conformation of OSMI-2a when bound to OGT (cyan) shown next to a representative of the five lowest energy conformers (gray), calculated by MacroModel, shows a different starting conformer before geometry optimization. (**B**) An overlay of the conformation of OSMI-2a in the OGT:OSMI-2a crystal (cyan) with the lowest energy conformer as determined by Density-Functional Theory calculations (magenta) binding shows remarkably close agreement between the bound and calculated conformations.

2.6 OGT:OSMI-2a structure confirm SAR analysis

The structure of OGT:OSMI-2a revealed that OSMI-2a binds in a different mode than what we expected from the OSMI-1 docked structure (Figure 2.6.1). The Q6S core from the docked model and the crystal structure overlap with each other but the sulfonamide in

OGT:OSMI-2a has a different conformation than the docked OGT:OSMI-1. We believe this difference is mainly caused by the inability of the Schrodinger suite to model the large shifts in the three residues Gln^{839}, Leu^{866}, and Phe^{868} upon OSMI-2a binding, even with the IFD protocol. Without the movement of the three residues, there will be steric clash with the thiophene and the three residues if the thiophene were to stick into the same cleft seen in OGT:OSMI-2a structure. For OSMI-1 to still fit into the pocket, IFD adopt a different conformer of the sulfonamide, which resulted in the sulfonamide substituents pointing inwards into the pocket, with the thiophene group forming an interaction with Lys^{842}. Nevertheless, this docking model allowed us to design a molecule that proved to be a good inhibitor even if some details of how it binds were not correctly predicted by the docking model.

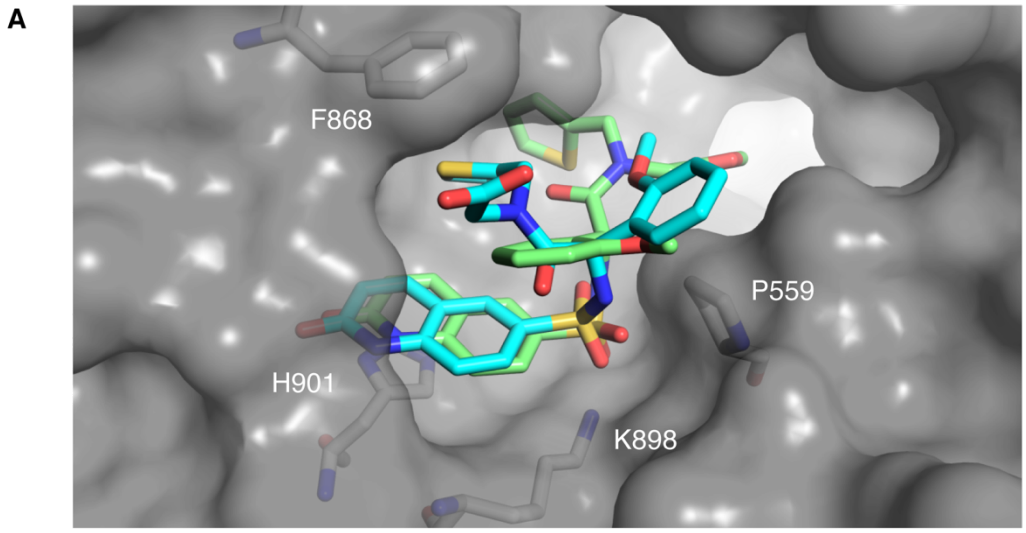


Fig 2.6.1 Overlay of docked OSMI-1 and crystalized OSMI-2a (**A**) Overlay of OSMI-2a (cyan) when bound to OGT (represented in grey surface) compared to docked OSMI-1 (green).

We next attempt to explain the SAR observed in the OSMI series with our OGT:OSMI-2a structures. There were mainly 3 observations in the SAR: the preference of phenylglycine, the requirement of disubstituted amide and the preference of R stereochemistry. Phenylglycine is

preferred over phenylalanine because a longer linker would have caused the phenyl group to exit the hydrophobic groove and move into the aqueous solvent, which will be highly unfavorable for a non-polar group like phenylglycine. The requirement of disubstituted amide can be explained by the necessary interactions of Phe^{868} and Gln^{839} with the carboxylic acid or thiophene respectively. To address the preference of R stereochemistry of the OSMI series, we also obtained a structure of OGT bound to the S-enantiomer of OSMI-2a (*ent*-OSMI-2a), which we expected to bind more weakly to OGT than the R-enantiomer. This observed weaker inhibition of the opposite stereoisomer is consistent with the SAR analysis from OSMI-1. From the structure, we noticed that the Q6S element in *ent*-OSMI-2a binds exactly as in OSMI-2a, confirming the importance of this fragment in binding (figure. 2.6.2). The switch in chirality of the substituted phenylglycine means that the substituents on the disubstituted amide project away from, rather than into, the deeper recesses of the active site. The weaker interactions of the amide substituents with active site residues undoubtedly drive the lower affinity of the S-configured compounds.

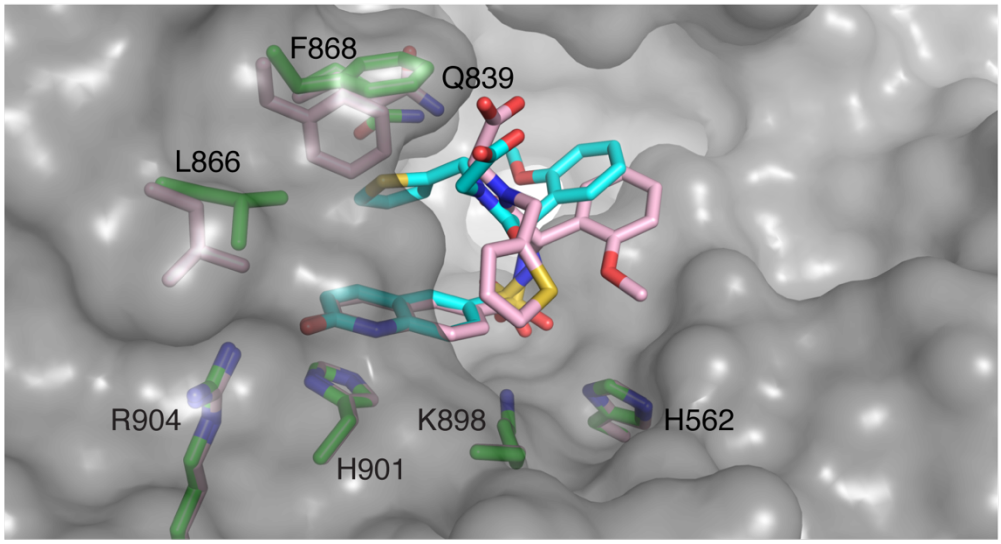


Fig 2.6.2. The crystal structure of the S-enantiomer of OSMI-2a confirms the importance of the Q6S group. An overlay of OSMI-2a (cyan) with the enantiomer of OSMI-2a a (*ent*-2a, pink) confirms that the Q6S group binds identically to OGT regardless of inhibitor stereochemistry. However, fewer contacts are made by *ent*- OSMI-2a to the active site as indicated by the lack of shifts in the side chains (yellow) of Gln839, Leu866, and Phe868 in the *ent*-OSMI-2a:OGT structure.

2.7 Development of an *in vitro* binding assay to quantify the potency of different inhibitors

We next wanted to develop a binding assay that can directly quantify the potency of OGT's inhibitors. Traditionally, radioactive assays have been used to measure the activity of OGT. This conventional assay utilizes radioactive substrates like UDP-[^{14}C] GlcNAc or UDP-[^{3}H] GlcNAc. The potency of inhibitors can then be determined by measuring the amount of transferred radioactive GlcNAc to an acceptor substrate [93]. However, this assay produces toxic radioactive waste, and radioactive UDP-GlcNAc is expensive. Thus, inhibitors of OGT have been mostly characterized with a coupled enzymatic assay such as the UDP-Glo assay, which measures the amount of UDP produced when GlcNAc is transferred to a protein/peptide acceptor

in the presence of OGT [103]. But coupled kinetic assays are limited in their sensitivity, and inhibitory potency depends on the concentration of UDP-GlcNAc, OGT and acceptor, making comparisons across sets of experiments challenging. To more accurately determine the binding potency of the inhibitors, we sought to develop an assay that can directly measure the dissociation constant (K_d).

We adopted Microscale Thermophoresis (MST), among other established methods, as our binding assay to measure the K_d of the inhibitors. MST is a facile and quick assay to quantify the K_d of binders. We reasoned that MST would be a better fit than other methods, which include surface plasmon resonance (SPR), isothermal titration calorimetry and biolayer interferometry [115-118]. Surface plasmon resonance has been adopted successfully using truncated OGT and the K_d of UDP-GlcNAc and UDP was determined to be 15 μM and 1 μM respectively [119]. Because SPR was developed on a truncated form of OGT, the K_d that was measured might not reflect the true value of the binding constant if the other domain (TPR region) effects the catalytic domain. To more accurately determine the K_d, we sought to develop a method that can measure K_d with full length OGT. Full length OGT, which is prone to aggregation, would not work with SPR, which is known to be incompatible with proteins that tend to aggerate. On the other hand, MST is a powerful and recently developed binding assay that can quantify the interaction of protein-small molecules or protein-protein with low sample consumption. This technique is highly sensitive to changes in size, charge and hydration shell, which is the main driver for most small molecule binding to proteins [120]. MST is also compatible in complex solutions like cell lysate, detergents or liposomes, which will enable us to rigorously optimize buffer conditions to minimize the aggregation of OGT. Hence, we chose MST as our binding assay to compare the binding of our inhibitors.

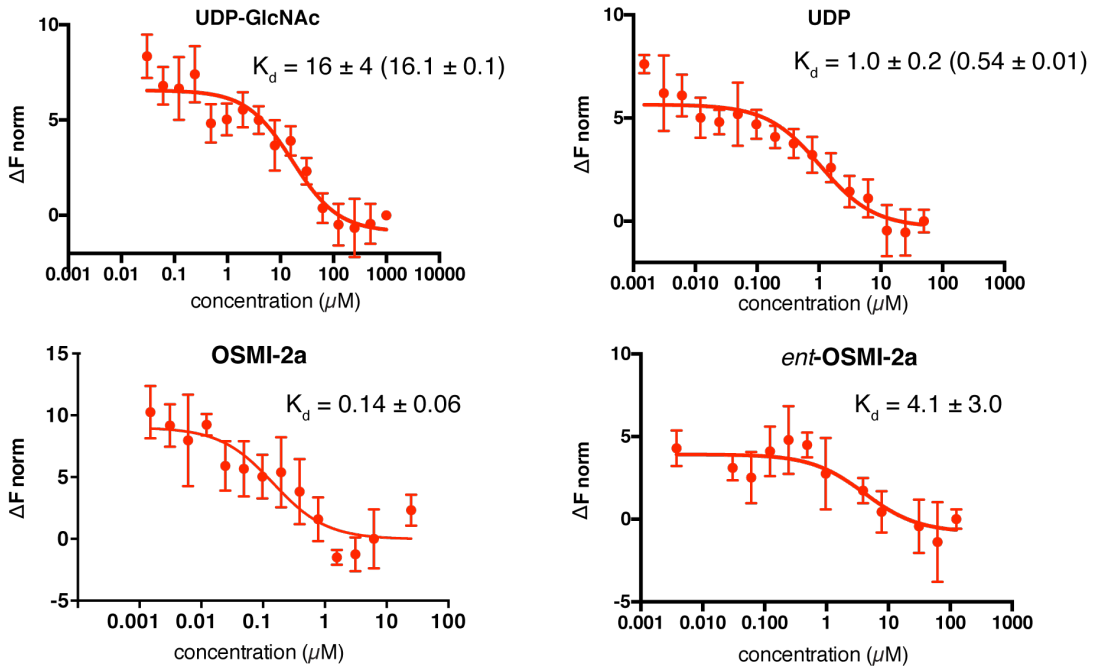


Fig 2.7. Binding curves of inhibitors with wild-type OGT as determined by MST. Data shown as normalized change in fluorescence (DF_{norm}) relative to highest concentration used vs. inhibitor concentration (μM). Error bars of graphs represent standard deviation (s.d.) at least three replicates. K_d is in μM. K_ds from SPR are given in parenthesis

To determine the validity of MST, we first quantify and compare the K_d of UDP and UDP-GlcNAc. Encouragingly, we obtained K_{ds} of UDP and UDP-GlcNAc that are similar to the ones reported from SPR (Figure 2.7). The fact that we did not see huge changes in the binding constant of UDP or UDP-GlcNAc with full length OGT or truncated OGT suggest that the upper region of the TPR domain does not affect active site binding of the inhibitor. We then determined the binding constant of OSMI-2a and *ent*-OSMI-2a. Agreeing with the drop in potency when the stereo configuration of OSMI-compounds (R-enantiomers) are reversed to the S-enantiomers, we observed a ~100 fold decrease in binding affinity for *ent*-OSMI-2a [103]. Taken together, we conclude that MST is suitable for determining the binding constant of OGT's inhibitors.

We next attempted to make a series of mutants to validate the structure of OGT:OSMI-2a. Upon inspection of the active site, we made a series of mutants, K898A, K898M, R904A and H562M, that we speculated are important to the binding of OSMI-2a. Because those residues also make contact with UDP-GlcNAc, we also measured the binding constant of UDP-GlcNAc and UDP (Table 3). R914A mutant displayed similar K_d to UDP-GlcNAc and UDP as wtOGT but showed a ~20-fold loss in binding to OSMI-2a, highlighting the importance of hydrogen bonding between Arg^{904} and quinolinone in OSMI-2a. The lack of changes in the binding of UDP-GlcNAc and UDP to OGT when Arg is mutated to Ala suggested that Arg^{914} mainly serves to anchor uracil in the right position for catalysis, rather than contribute to the binding energy. Confirming this anchoring role, when R914A mutant was assessed for its ability to glycosylate CKII peptide or NUP62 proteins, we did not observe any visible glycosylated bands (data not shown here). K898A mutant was not stable and causes OGT to aggregate, which makes MST measurement impossible. Both K898M and H572M mutants resulted in a ~10-fold decrease in binding of OSMI-2a and abrogated binding for UDP-GlcNAc. Overall, this results further confirm the importance of the Q6S core as a mimic for uridine in the active site.

Table 3

Binding of OGT mutants

	Wild type	**H562M**	**K898M**	**R904A**
UDP	1.0 ± 0.2	n.d.	n.d.	0.6 ± 0.2
UDP-GlcNAc	16 ± 4	No binding	No binding	17 ± 8
OSMI-2a	0.14 ± 0.06	1.6 ± 1.0	2.4 ± 1.0	5 ± 3

Binding affinity (K_d ± s.d.) of various OGT mutants determined by MST in μM. n.d. means not determined

2.8 Structural optimization of OSMI-2a yields nanomolar binders

We next asked if we can further improve OSMI-2a by using the OGT:OSMI-2a structure. Analysis of the OGT complex with OSMI-2a suggested two strategies to modify these inhibitors to make additional contacts. One strategy exploited the observation that the *ortho*-methoxy group points toward an unoccupied region of the active site where the diphosphate of UDP-GlcNAc would bind (Figure 2.8.1). Previous structural studies of the OGT:UDP-GlcNAc complex showed that one of the anomeric phosphate oxygens forms hydrogen bonds to the backbone amides at the N-terminus of a proximal helix [121], while another contacts the side chain amine of catalytically essential Lys[842] [121-123]. The binding pose of OSMI-2a suggested that it would be possible to mimic these interactions by attaching a carboxylate to the phenyl ring via a sufficiently long linker to bridge the distance to the phosphate binding site. We prepared compounds OSMI-3a and OSMI-3b with linkers containing three or four methylene and solved structures of OGT bound to both inhibitors. The second strategy focused on including a small substituent on the quinolinone *ortho* to the sulfonamide because analysis of the crystal structures suggested it would be possible to achieve a tighter fit in the uridine pocket (Figure 2.8.2). We hypothesized that a larger atom like chlorine could fit into the groove and prepared chlorine derivative OSMI-4a and solved the structure of OGT:OSMI-4a.

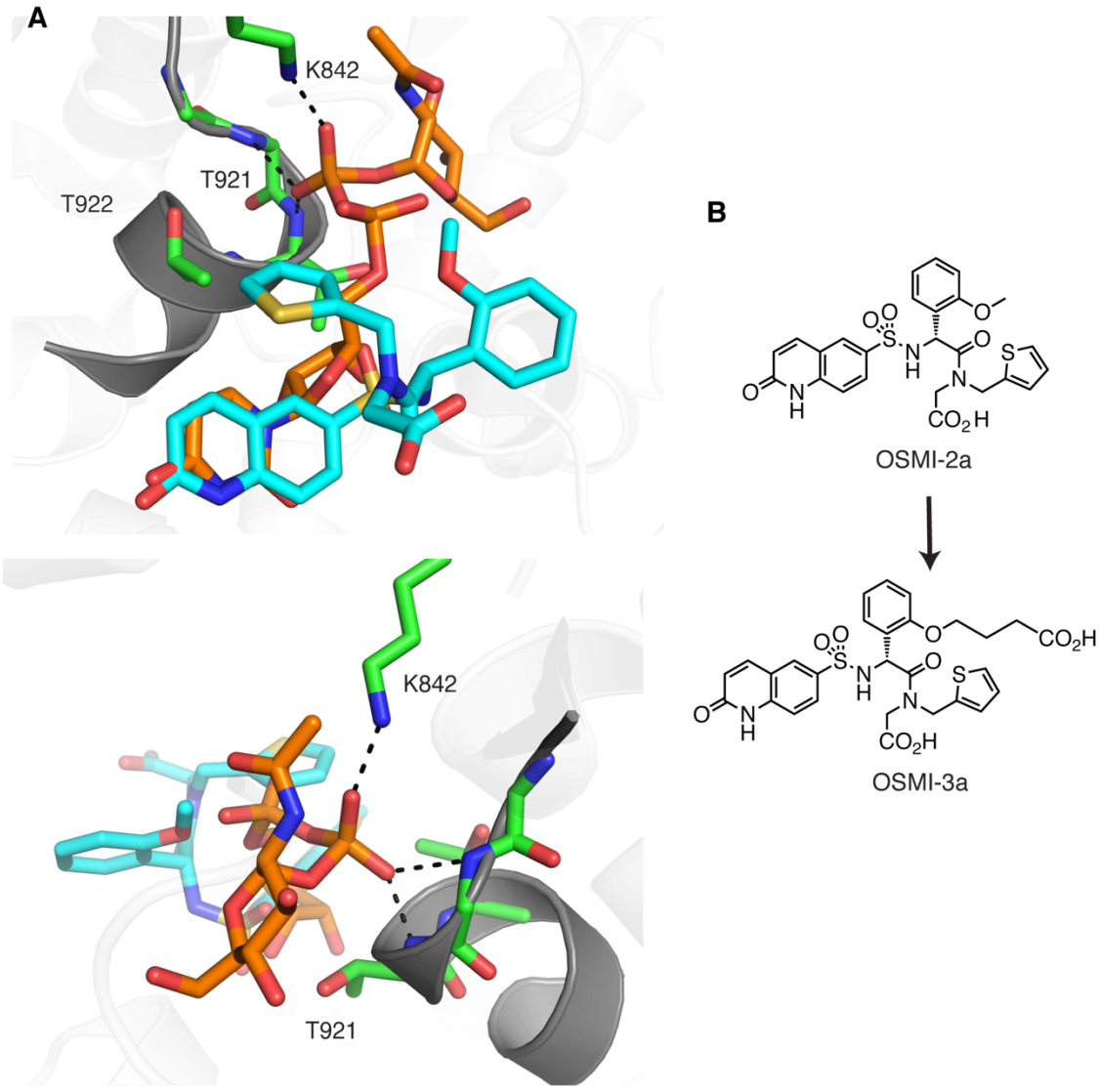


Fig 2.8.1 The methoxy group of OSMI-2a points toward an unoccupied region of the active site. (A) An overlay of OSMI-2a (cyan) with UDP-GlcNAc (orange, PDB: 4N3C) shows that a large portion of the OGT active site remains unsampled by 1a. Phosphate oxygens make contacts to residues Lys^{842} and Thr^{921}. **(B)** Proposed modification to OSMI-2a by extending the ortho-methoxy group.

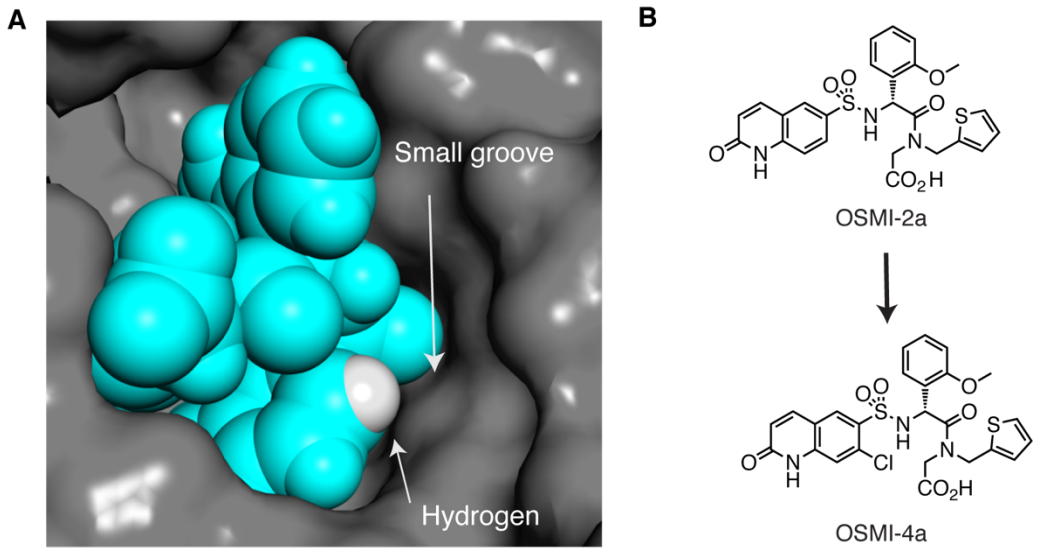


Fig 2.8.2 Structure of OGT:OSMI-2a showed a small groove for additional atoms. (**A**) Space filling view of OSMI-2a showing the ortho-hydrogen atom in white. (**B**) Proposed modification to OSMI-2a by adding a chlorine atom to fill up the groove

All proposed compounds make additional contacts to OGT in the expected region of the active site (Figure 2.8.3 and 2.8.4). Strikingly, the pendant carboxylate of compound OSMI-3b overlaps almost perfectly with the UDP-GlcNAc phosphate and makes the same interactions with the proximal helix (Figure 2.8.3a). While the carboxylate on OSMI-3a is not within H-bonding distance of the N-terminal amides, one of its oxygens interacts with the side chain of Thr^{921} while the other is oriented towards Lys842, which is 3.7 Å away (Figure 2.8.3). Structure of OGT:OSMI-4a confirms the tighter and better fit of OSMI-4a with the chlorine tugging into the groove (Figure 2.8.4). Importantly, all the OSMI compounds show a clear overlap of the Q6S core, which further confirms the importance of having the Q6S core.

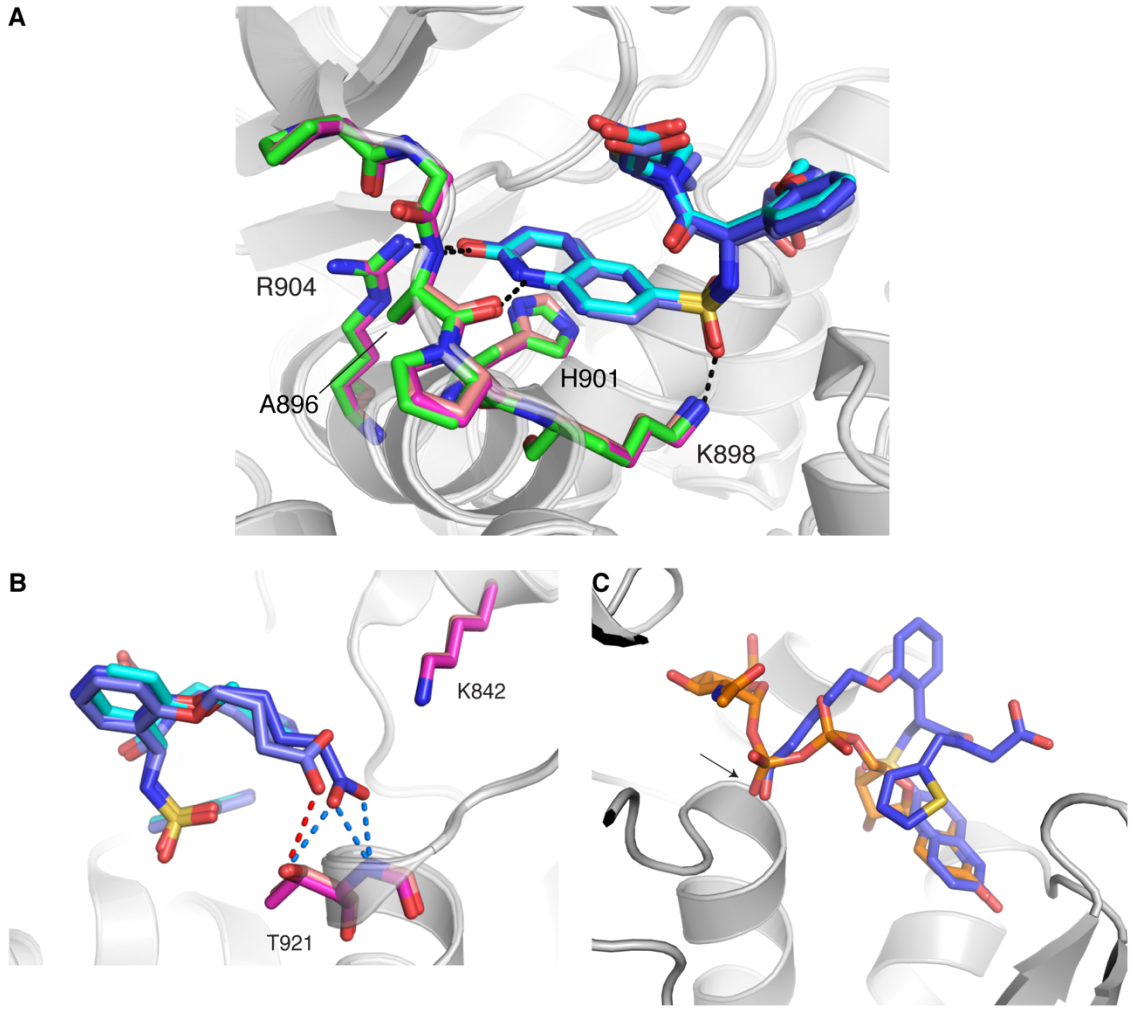


Fig 2.8.3 Structure of OGT:OSMI-3a showed the mimic of phosphate and the carboxylate. (A) Overlays of OSMI-2a (cyan), OSMI-3a (light purple), and OSMI-3B (dark blue) with residues to which inhibitors are expected to make contacts overlaid in green, dark pink, and beige, respectively. Hydrogen bond contacts to Ala^{896}, Lys^{898}, and Arg^{904} are shown with dashed lines. His^{901} engages in a pi-stacking interaction with the Q6S group. (**B**) Overlay of OSMI-2a (cyan), OSMI-3a (light purple), and OSMI-3b (dark blue). Dashed lines indicate hydrogen bond contacts to Thr^{921} from OSMI-3a (red) and OSMI-3b (blue), and to backbone amides from OSMI-3b. **(C)** A superposition of OSMI-3b with UDP-GlcNAc (PDB: 4N3C) shows close overlap of a carboxylate oxygen of OSMI-3b with a phosphate oxygen of UDP that is anchored to the N-terminus of a nearby a- helix.

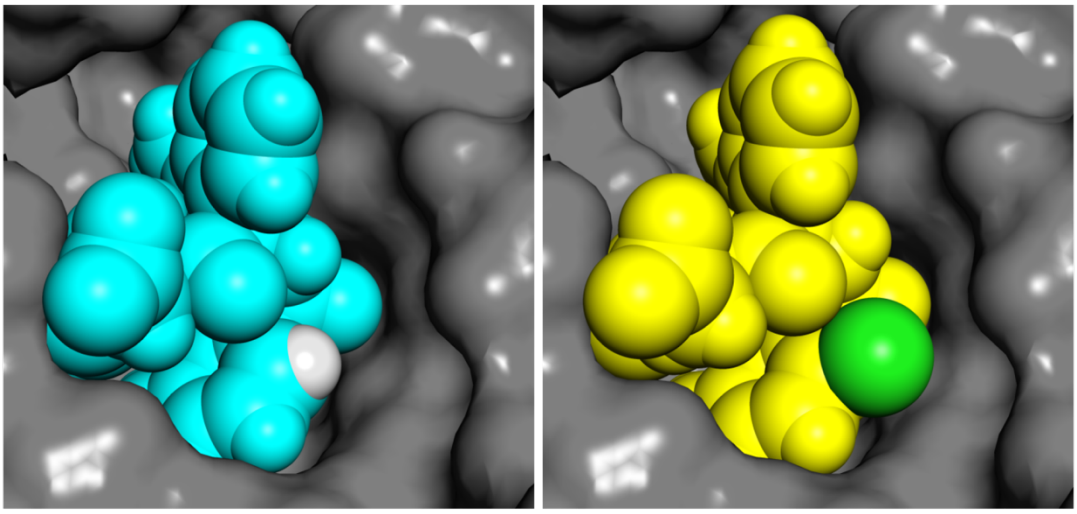


Fig 2.8.4 Structure of OGT:OSMI-4a showed the snugger fit of the chlorine atom (A) Space filling view of OSM-2a (cyan) and OSMI-4a (yellow) with the hydrogen and chlorine shown in white and green respectively.

We next asked how those new compounds, with extra binding contacts, would compare in binding affinity. We measured the K_d of the new compounds with MST. Agreeing with our structures showing a better fit and interaction in the pocket, compounds OSMI-3a and OSMI-4a bound OGT with K_ds of ~5 and ~8 nM, respectively, while the K_ds of OSMI-2a and OSMI-3b were an order of magnitude higher (Table 4). The tighter binding of OSMI-3a compared with OSMI-3b is likely explained by a stronger interaction of the negatively-charged carboxylate of OSMI-3a with the positively charged Lys^{842} amine combined with a smaller loss in conformational entropy due to immobilization of the shorter linker. The tighter binding of OSMI-4a compared with OSMI-2a is attributed to its snugger fit in the binding pocket.

Table 4. Binding of compounds to OGT[a]

UDP	1.0 ± 0.2 (0.54 ± 0.01)[b]
UDP-GlcNAc	16 ± 4 (16.1 ± 0.1)[b]
OSMI-2a	0.14 ± 0.06
ent-OSMI-2a[c]	4.1 ± 3.0
OSMI-3a	0.005 ± 0.001
ent-OSMI-3a	0.14 ± 0.05
OSMI-3b	0.045 ± 0.02
OSMI-4a	0.008 ± 0.02

[a]Binding affinity (K_d ± s.d.) in μM, [b]determined by SPR[16], [c]*ent* = enantiomer

2.9 Picomolar binder of OGT: OSMI-5a

Because OSMI-3a and OSMI-4a bind in the nanomolar range and contain modifications that are unique from each other, we hypothesize that we can further increase the binding affinity of the OSMI-series by appending the unique features from OSMI-3a and OSMI-4a to OSMI-2a (Figure 2.9.1a). Therefore, we prepared OSMI-5a, which contains the methyl carboxylate from the phenyl ring from OSMI-3a and the chlorine derivative from OSMI-4a. OSMI-5a was synthesized by collaborators from NIH and the K_d was determined by MST. Notably, we broke into the picomolar potency and obtained a binding constant of 0.3 nM for OSMI-5a (Figure 2.9.1b). This increase in potency of ~400 fold from OSMI-2a aligned with our prediction, considering that OSM-3a and OSMI-4a gave ~20-fold increase in potency from OSMI-2a. We are currently trying to obtain a crystal structure of OGT:OSMI-5a.

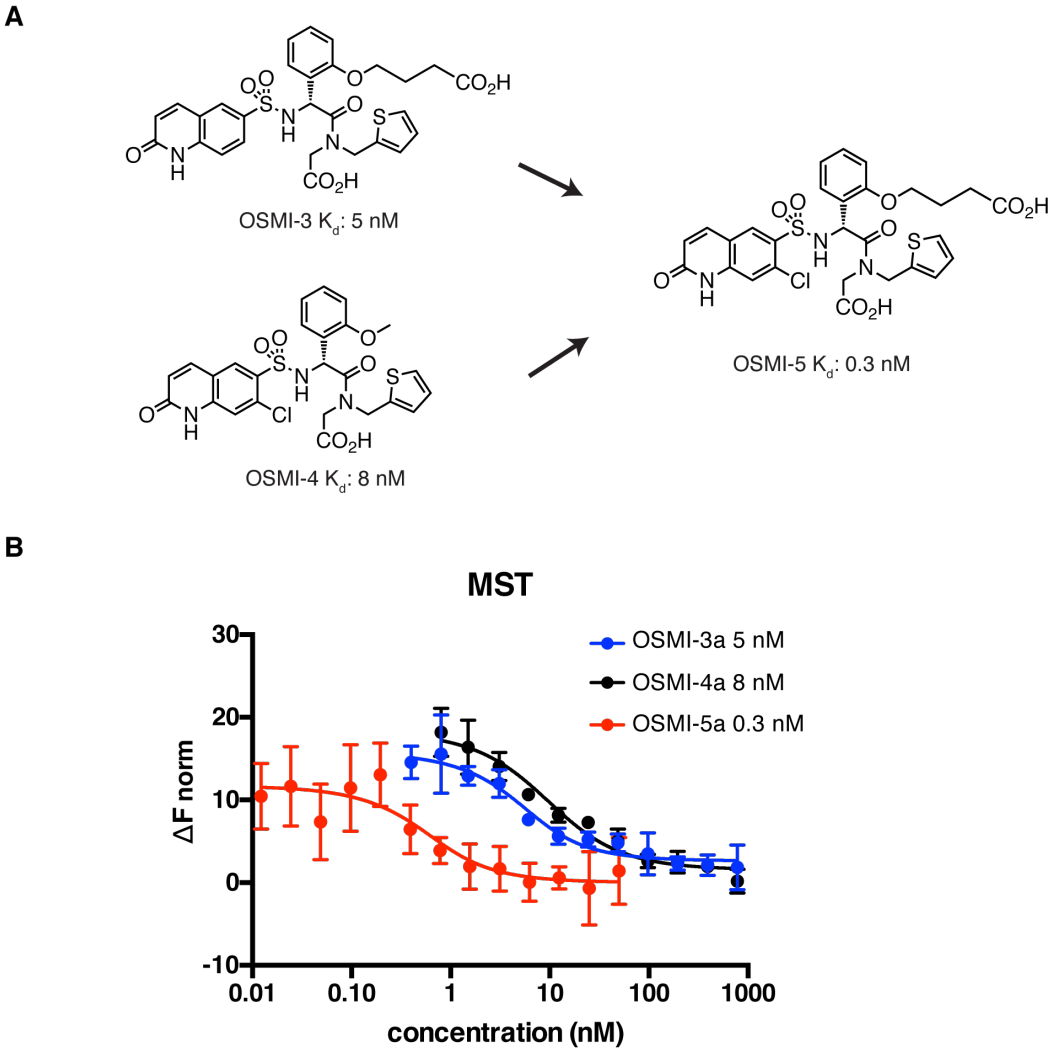


Fig 2.9.1 OSMI-5a breaks into the picomolar potency. (A) OSM-5a is derived from the features of OSMI-3a and OSMI-4a. **(B)** MST data shown as normalized change in fluorescence (ΔF_{norm}) relative to highest concentration used vs. inhibitor concentration (nM). Error bars of graphs represent standard deviation (s.d.) at least three replicates. K_d is in nM.

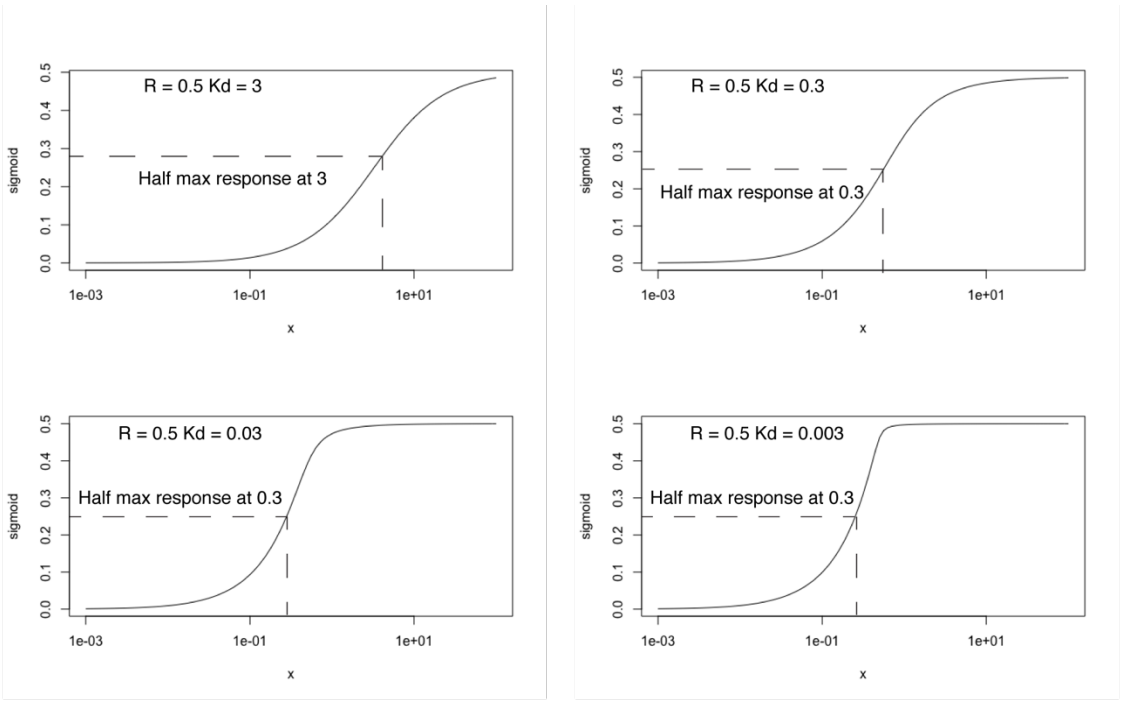


Fig 2.9.2 Simulation of sigmoidal K_d curves show similar half max value if K_d << receptor concentration. R represents the receptor concentration. Units are in arbitrary unit.

One concern with K_d determination for many binding assays is the quantification of tight binders. For example, if K_d << concentration of receptor, the K_d, which is usually calculated from half the maximum signal, will fall in the range of half the concentration of the receptor and the K_d measured will thus be an overestimation of the real value (Figure 2.9.2). Because the concentration of the receptor usually determines the sensitivity of the assay, it is typically difficult to use a concentration that is much less than the K_d for a tight binder. One concern with OSMI-5a is that OSMI-5a binds OGT in the picomolar range and that the concentration of OGT used in the MST experiment was 0.5 nM, which is the lowest concentration of OGT that can be reliably detected by MST. Because the K_d of OSMI-5a is 0.3 nM, which is about half the concentration of OGT, we wondered if the exact K_d might be lower than 0.3 nM. To address that possibility, we performed simulation of K_d curves with the mass equation used in our study to see if mass equation would be

able to correctly predict the K_d with the simulated data points. At the same concentration of 0.5 nM OGT, the mass equation fit is able to accurately estimate the K_d even when the simulated K_d is 0.0003 nM, which is ~1000-fold less than the concentration of OGT (Figure 2.9.3). We reasoned that the mass equation works extremely well because the equation is independent on the max signal used and does not estimate K_d from half the maximum signal. From this result, we conclude that MST can be used to determine the K_d of tight binders.

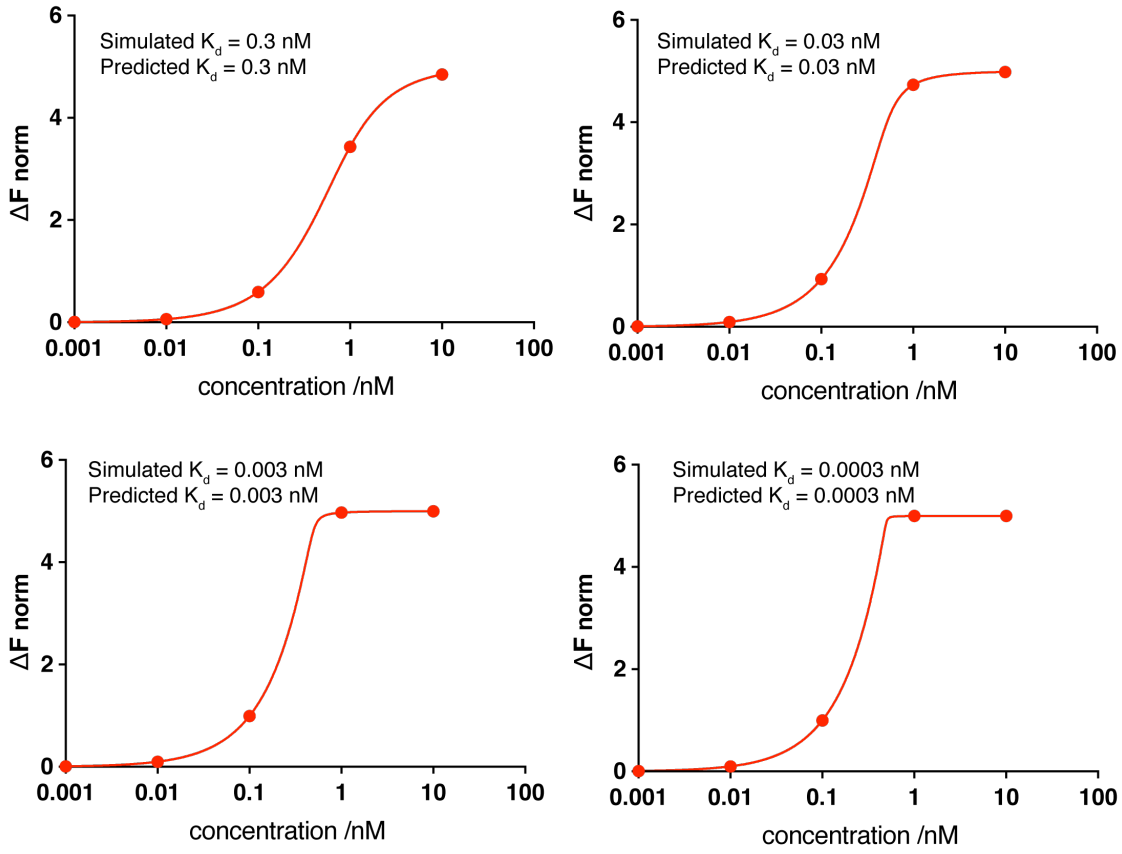


Fig 2.9.2 Simulated curve fitting shows that MST can be used for tight binders. Concentration of OGT was kept at 0.5 nM and the K_d was varied from 0.0003 nM to 0.3 nM, which is the K_d of OSMI-5a.

2.10 Conclusions

In this chapter, we first modelled the binding of OSMI-1 into the active site of OGT and obtained a binding mode that is consistent with SAR analysis. From this model, we designed OSMI-2a, which allowed us to obtain the first crystal structure of OGT. This is the first crystal structure of OGT with any small molecule inhibitors bound. Though the crystal structure of OGT:OSMI-2a differs from the docked OGT:OSMI-1 structure, the docking model helped us to formulate a strategy that resulted in the replacement of the furan with a carboxylic acid group. This replacement is not obvious from the structure of OSMI-1 and highlights the importance of docking in inhibitor development.

To our knowledge, there has never been a crystal structure of any GTs bound to a cell permeable small molecule inhibitor that can serve as a guide, until this discovery. The crystal structure, OGT:OSMI-2a, is a significant breakthrough in the field. Using OGT:OSMI-2a, we then further evolved OSMI-2a to OSMI-3a/4a/5a. OSMI-3a and OSMI-4a break into the nanomolar potency, and OSMI-5a, which combines features from OSMI-3a and OSMI-4a has an impressive K_d of 0.3 nM, making OSMI-5a the best in class inhibitor of OGT. In fact, OSMI-5a is also the first picomolar inhibitor in the field of GT inhibitors. There are still unanswered questions regarding the inhibitors, however. Do these inhibitors work in cells? How can we accurately determine the potency of the inhibitors in cells? How fast do they act in cells and how do cells respond to inhibitor treatment? Next chapter will focus on the characterization of the inhibitors in cells.

2.10 Methods

Molecular Modelling

Structures of OGT complexes with bound UDP-GlcNAc (PDB IDS: 4N3C) were prepared using Protein Prep Wizard Schrödinger.[124] The SiteMap package within the Schrödinger suite was used to calculate the Dscore, size, hydrophilicity and hydrophobicity in the UDP-GlcNAc binding pocket, using the default parameters [109].

IFD and XP were performed using the module by Schrödinger suite with the both docking methods. Protein Prep Wizard were used to perform energy minimization of OGT:UDP-GlcNAc structure (PDB IDS: 4N3C).

For the superposition of OSMI-2a into the other Gtfs, the uridine ring from OGT (PDB IDS:4N3C) was first superimposed onto the uridine ring from the respective complexes with the Superposition panel in Maestro within the Schrödinger suite. Then OSMI-2a**:**OGT complexes were aligned onto OGT (PDB IDS:4N3C) in PyMOL to generate the overlay.

Density functional theory calculations

To find the lowest energy conformers, a solution phase conformational search using MacroModel was performed, [125] with OPLS3 [126] as the force field with the default settings, on **1a** extracted from the crystal structure. A representative of the five lowest energy conformers were geometry optimized using density functional theory (DFT) calculations with B3LYP as the functional and the 6-31+G(d) basis set [127]. The vibrational frequencies were computed at the same level to verify that the optimized structure is an energy minima and to evaluate zero-point vibrational energies (ZPVE) and thermal corrections at 298 K. The calculations were performed

using SMD solvent continuum as the solvent method [128]. All calculations were performed using Gaussian09 software [129].

Protein expression and purification

Full length human OGT (ncOGT) and human $OGT_{4.5}$ ($hOGT_{4.5}$) were prepared as previously described with minor variations [122]. Cultures were grown at 37 °C after diluting an overnight culture 1 to 100 in fresh LB media. Cells were grown to an A_{600} of 1, at which point they were cooled to a temperature of 16 °C. After letting the cells grow at 16 °C for 1 hour, they were induced with 0.2 mM IPTG and grown at 16 °C for 16 h. Cells were pelleted, resuspended in TBS (20 mM Tris, pH 7.4, 250 mM NaCl) supplemented with 1mM PMSF and 0.1 mg/mL lysozyme, lysed, and the lysate was centrifuged at 5,000xg for 20 min to remove unbroken cells. Imidazole was then added to the supernatant to a final concentration of 40 mM before the lysate was incubated with Ni-NTA agarose superflow resin (Qiagen) which was prewashed with TBS + 40 mM imidazole for batch nickel affinity purification. After incubating the lysate and the resin with gentle rocking at 4 °C, the flowthrough was removed, and the resin was washed with 10 column volumes of TBS + 50 mM imidazole. The protein was then eluted with 4 column volumes of TBS + 250 mM imidazole. The eluate was supplemented with 1 mM THP (Novagen) to prevent aggregation and then concentrated with centrifugal concentrators (Millipore). After protein concentration determination, for $hOGT_{4.5}$ the N-terminal tags were cleaved by adding HRV3C protease (EMD, M02905) to the concentrated purified protein at a ratio of 1 unit/150 mg of protein (determined by NanoDrop, MW = 80876 Da, $\varepsilon = 77240\ M^{-1}cm^{-1}$) and incubating at 4 °C for 16 h. Following cleavage, the protein was further purified by gel filtration on a Superdex 200 column (GE Healthcare) in TBS (20mM Tris, pH 8.0, 150mM NaCl) and fractions were supplemented with 1 mM THP after gel filtration. The fractions were collected and again

concentrated using centrifugal concentrators. The protein was monomeric in solution as determined by gel filtration.

Measuring dissociation constants using microscale thermophoresis

Microscale thermophoresis was performed with a NanoTemper monolith NT.115Pico instrument.[130] Purified full length OGT was fluorescently labeled with Alexa Fluor™ 647 NHS Ester (ThermoFisher) (NHS-ester:OGT = 1.5:1 mole ratio) in labeling buffer (PBS pH 8.5, 1 mM DTT) for 1 hour at room temperature with end to end rotation in the dark. Excess NHS-ester was removed with a Zeba™ Spin desalting column (ThermoFisher). A range of concentrations of the ligand were prepared in which the concentration of labeled OGT was kept constant at 5 nM, and all experiments were performed in MST running buffer (PBS, pH 7.4, 0.05% Tween-20, 1 mM DTT, 2% glycerol, 0.5% DMSO). Mixtures were incubated for at least 10 minutes to facilitate binding before MST experiments. MST experiments were carried out using 60% LED power and 20 % MST power in standard capillaries (from NanoTemper Technologies). Experiments were performed with at least three biological replicates except for *ent*-OSMI-2 (two replicates). K_d values were calculated using a user-defined mass action equation in the GraphPad Prism 6 software (GraphPad Software, Inc.).

Chapter 3

Contributions: The chapter involves a close collaboration with Joao Paulo, who has been a great mentor and help for the proteomics preparation. All of the HEK cell line work was performed by me. Harri Itkonen helped to characterize of the compounds in other cell lines by running, for example, the cellular viability assays.

Characterization of OGT inhibitors in cells

3.1 Introduction

Despite advances in inhibitors development for OGT, there has been a lack of methods that can accurately and quantitatively determine the potency of OGT's inhibitors in cell, not least due to the wide variety of substrates that OGT bears which further increases the complexity. For example, because each substrate has different binding affinity to OGT, should we then measure the efficacy of an inhibitor against a single substrate, a subset of substrates or all the substrates? Coupled with a lack in potent and specific cell based inhibitors for OGT, there have not been much focus in methods development for measuring OGT inhibition. The most widely used method for measuring cellular OGT inhibition is immunoblotting with O-GlcNAc antibodies, which can detect a specific set of O-GlcNAcylated proteins [131-133] (Table 5). Among them, the most widely used O-GlcNAc antibodies are RL-2 and CTD110.6. RL-2 was raised against O-GlcNAcylated nucleoporin proteins and CTD110.6 was raised against a synthetic O-GlcNAcylated peptide that corresponds to a sequence from RNA pol II. Interestingly, both antibodies also recognized additional O-GlcNAcylated proteins other than what they were raised against, and because of that, they have been widely used as surrogates for quantifying global O-GlcNAc levels [61, 134]. However,

antibodies are not perfect. They are epitope dependent and could also cross react with other GlcNAcylated species, which might result in erroneous detection of O-GlcNAc levels. For instance, it was found that CTD110.6, which cross recognized terminal β-linked-GlcNAc (β-GlcNAc) and other N-glycan cores, provided a wrong readout that decreased glucose levels increases O-GlcNAc levels [135]. Therefore, methods that can reliably quantify O-GlcNAc levels in cells will be useful to the field.

Table 5

Specificity of OGT antibodies

Name	Epitope	Specific to O-GlcNAc?	Source	Ref
RL-2	Nuclear pore	Yes	Abcam	[136]
CTD110.6	RNA Pol II	No	Sigma	[137]
HGAC85	Streptococci glycoconjugate	Yes	Thermo	[138]
18B10.C7	CKII	Yes	N.A.	[139]
9D1.E4	CKII	No	N.A.	[139]
1F6.D6	CKII	Yes	N.A.	[139]

Chemoenzymatic detection has emerged as an orthogonal method to measure O-GlcNAc changes. There have been a number of other tools that have been developed for specific and selective O-GlcNAc detection in the recent years. For example, chemoenzymatic methods like metabolic feeding of the azido sugar, GalNAz, which gets converted into UDP-GlcNAz in cells, or attachment of GalNAz to terminal GlcNAc *in vitro* using an engineered B-1,4-galactosyltransferase, have been successfully performed to detect and measure O-GlcNAc levels [140-141]. Despite successes with O-GlcNAc detections, the abovementioned methods have yet been reliably tested and validated with OGT inhibitors.

The first section will focus on the adoption and optimization of chemoenzymatic methods to develop assays that can quantitatively measure O-GlcNAc levels in cells. As mentioned in

chapter 1, there are still many areas in OGT biology that have yet to be discovered. In order to access OGT as a therapeutic target, it is crucial to better understand its biology. Thus far, no reports have reported on the global effects of lowering O-GlcNAc levels (in the 24 hour timeline) mainly because 1) genetic knockdown usually takes > 2 days to reduce protein levels (average half-life of protein ~1 day) and 2) other inhibitors are not potent or selective to OGT [98, 142]. Here we report on selective and permeable inhibitors of OGT that are fast acting and potent in a number of different cells. Moreover, we found that the inhibitors were able to reduce O-GlcNAc by ~50% within 2 h of treatment. To explore the effects on decreased O-GlcNAcylation, we performed quantitative proteomics to access how cells respond to inhibitor treatment, i.e. lowering O-GlcNAc levels for < 6 h (rapid) or 18-24 h (intermediate). We found that reduction of O-GlcNAc has a profound effect on sterol metabolic enzymes and decreases cholesterol levels. We conclude that the inhibitors are active in cells and can be used to further delineate the biology of OGT in a way that has not been possible before.

3.2 Western blotting that detects O-GlcNAc levels are semi-quantitative

To reliably quantify O-GlcNAcylation in cells, a two-step chemoenzymatic method, which can fluorescently label all O-GlcNAc modified proteins independent of the sequences, was optimized (Figure 3.2.1a). The first step involves addition of mutant B-1,4-galactosyltransferase GalT (Y289L) and UDP-GalNAz, which carries an azido group, to cell lysates. Because GalT Y289L transfers modified GalNAc (GalNAz) to the O-GlcNAc moiety, this step ensures that all O-GlcNAc residues carry an azido group. The second step involves a copper click reaction with a fluorophore-alkyne, which will then label O-GlcNAc modified proteins with a fluorophore. After

this step, detection and quantification of O-GlcNAc modified proteins can be performed quantitatively on a protein

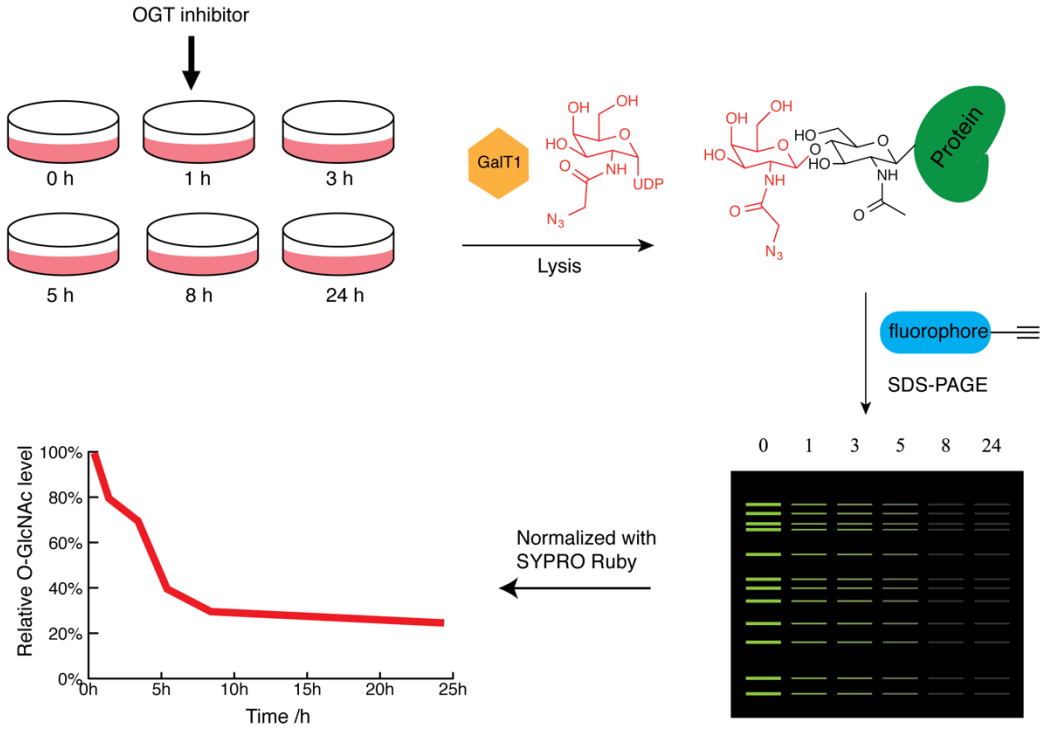


Fig 3.2.1 Schematic of fluorescence gel assay to quantify O-GlcNAcs. First, GalT1 are added to cell lysate to selectively tag O-GlcNAc with GalNAz, which carries an azido group. An alkyne carrying a biotin or fluorophore can then be added on using click chemistry, which selectively reacts with the azido moiety. O-GlcNAc levels can then be visualized with fluorescence after protein gel electrophoresis.

gel by fluorescence imaging. After screening a number of conditions, we determined that 500 µM of alkyne at 1h room temperature gives the best signal and highest signal to noise ratio.

We asked if western blotting and gel fluorescence would give similar results. We first compared the potency of OSMI-1 and OSMI-2a with the gel fluorescence assay. Because OSMI-2a contains a negatively charged acidic group, OSMI-2a would not penetrate into cells due to the negatively charged cellular membrane. To enable cellular penetration, we prepared corresponding ester derivatives for all the OSMI compounds described in chapter 2 (Figure 3.2.2a). This strategy

uses the ubiquitous and promiscuous intracellular esterase to unmask the carboxylic acid from OSMI-2 [143]. We found that OSMI-2 is better than OSMI-1, reaching ~50% reduction in global O-GlcNAc levels after 4 h treatment with OSMI-2 (Figure 3.2.2b). We also observed similar level in reduction with western blotting. For example, both methods gave similar level of reduction in global O-GlcNAc levels with OSMI-2 at both 4 h and 8 h treatment (Figure 3.2.2c). This finding is not unexpected. First, RL-2, which was raised against a mixture of nucleoporin proteins in pore complex-lamina fraction, can potentially recognize a range of other O-GlcNAcylated proteins [136]. Next, because the signal from the chemoenzymatic method is skewed towards heavily O-GlcNAcylated proteins and nucleoporins are one of the most O-GlcNAcylated proteins, the gel fluorescence method could be mainly reporting changes on nucleoporins' O-GlcNAcylation, which is what Rl-2 was raised against [144-147]. These results suggest that western blotting with RL-2 is semi-quantitative and can be used to compare the potency of OGT inhibitors in cells.

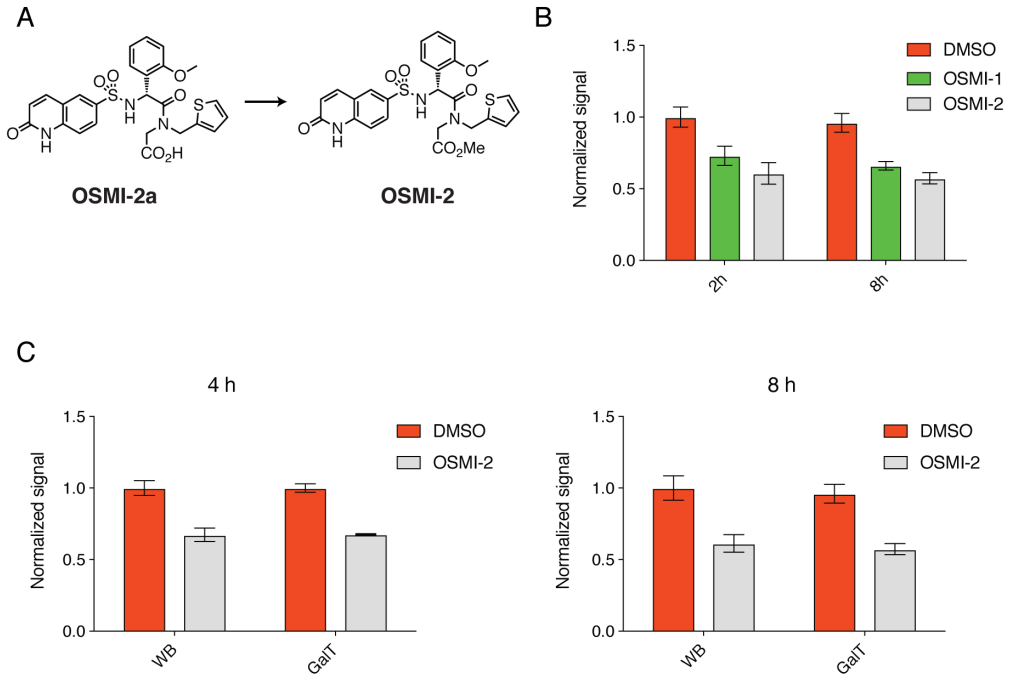


Fig 3.2.2 Western blotting can be used to quantify changes in cellular O-GlcNAc levels. (A) OSMI-2a is esterified to form cell permeable OSMI-2. **(B)** Gel fluorescence with GalT method showed that OSMI-2 is more potent than OSMI-1 in cells at both short and long time point. Signals are normalized to total protein, stained by SYPRO ruby, and compared to DMSO treatment. **(B)** Western blotting with Rl-2 gives similar results as Gel fluorescence method at both 4 h (left) and 8 h (right).

3.3 OSMI-4 is the best in class inhibitor for OGT

We then explored the cellular activity of the other OSMI compounds. We first examined OSMI-2/3/4 ability to inhibit OGT in a number of cell lines, including HEK293T, HCT116 and PC3 cells. All compounds reduced protein O-GlcNAc levels at the highest concentration tested and as fast as 2 h with just 10 uM (Figure 3.3.1). All of the compounds showed similar level of inhibition at 20 uM at short time point. Consistent with *in vitro* data that the S-enantiomeric version of OSMI-2a binds ~100x weaker to OGT, *ent*-OSMI-2/3 displayed no cellular activity against OGT (Figure 3.3.2).

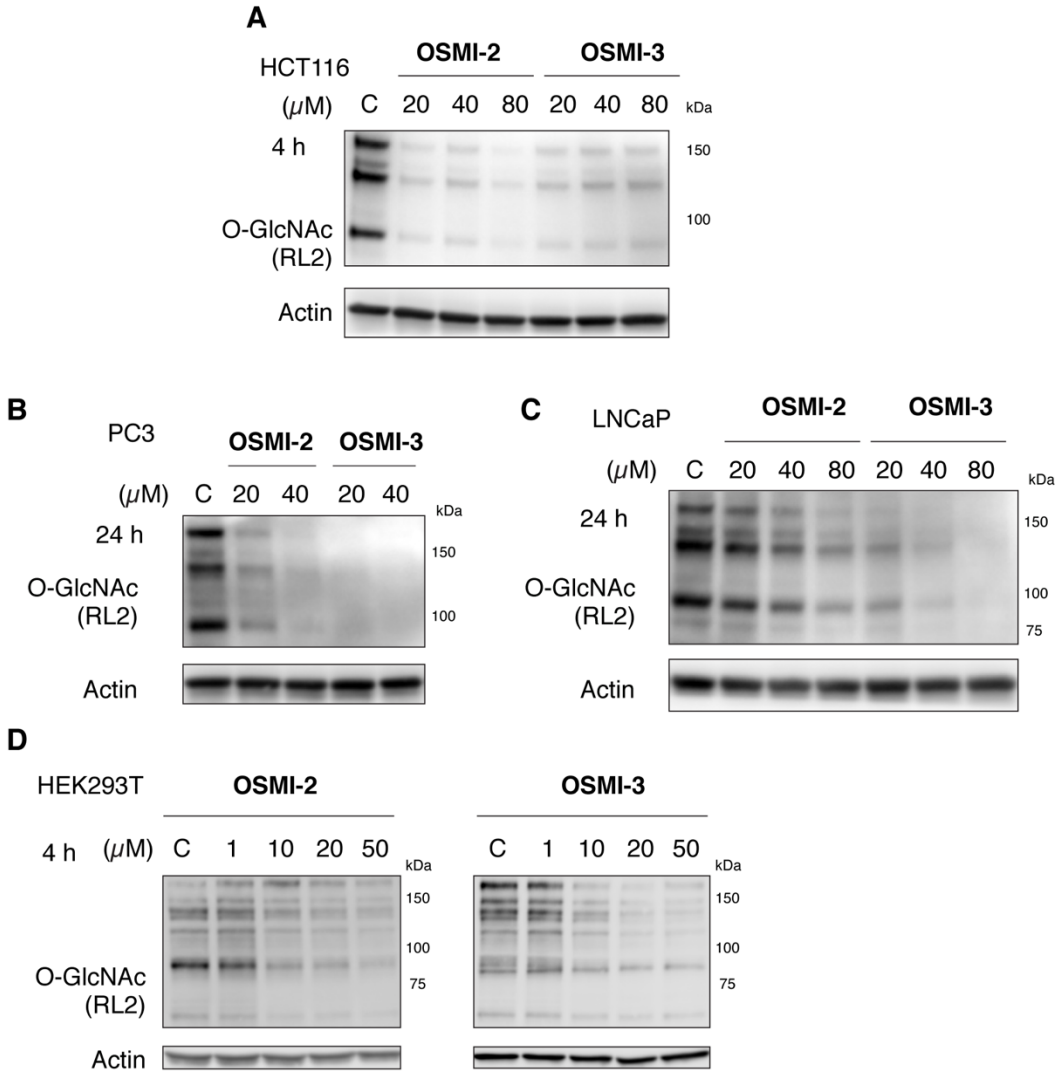


Fig 3.3.1 OSMI-2 and OSMI-3 inhibit OGT in cells. Both OSMI-2 and OSIMI-3 inhibit OGT in a number of cell lines, including **(A)** HCT116, **(B)** PC3, **(C)** LNCaP and **(D)** HEK293T cells.

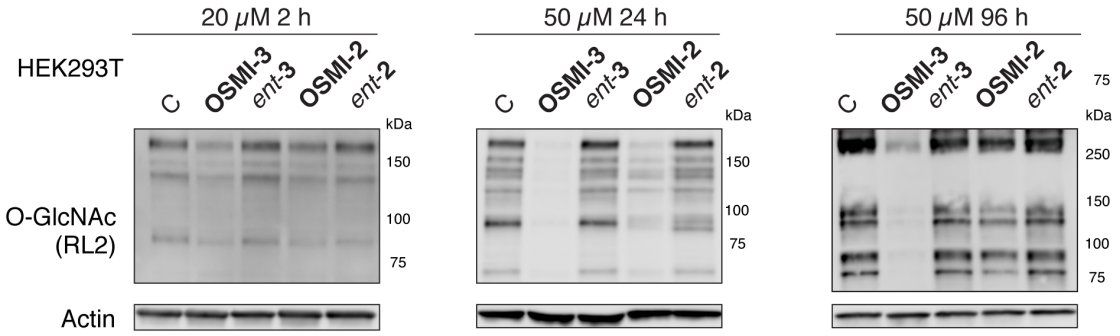


Fig 3.3.2 Enantiomers of OSMI-2/3 can be used as negative controls. Western blotting of HEK293T cells with OSMI-2/3 or their enantiomer version (*ent*-2/3 respectively) showed that the enantiomers do not inhibit OGT at all time point tested.

At longer incubation times (24h), O-GlcNAc levels began to recover for OSMI-2 and this phenomenon is further explained in chapter 3.4. In addition, we observed that the level of reduction in O-GlcNAc is in the order of OSMI-4 > OSMI-3 > OSMI-2 (Figure 3.3.3a). The greater sustained effect of OSMI-3 compared with OSMI-2 is due to the greater affinity of OSMI-3a than OSMI-2a for OGT. Despite having similar *in vitro* potency, the greater effect of OSMI-4 compared to OSMI-3 might be due to the lower molecular weight of OSMI-4 resulting in increased cellular penetration. Notably, OSMI-4 was able to reduce O-GlcNAc levels by 80% at 5 uM and with very little detectable O-GlcNAc levels at 20 uM (Figure 3.3.3b). To examine the potency of OSMI-5, we did a direct comparison against OSMI-4, our best in cells inhibitor. At 24h, OSMI-4 showed higher level of inhibition against OSMI-5 at all concentrations tested except at 20 uM, where the potency is similar. The greater inhibitory effect of OSMI-4 than OSMI-5 is attributed to the greater cellular penetration of OSMI-4, similar to what we observed for OSMI-4 vs OSMI-3 (Figure 3.3.3c).

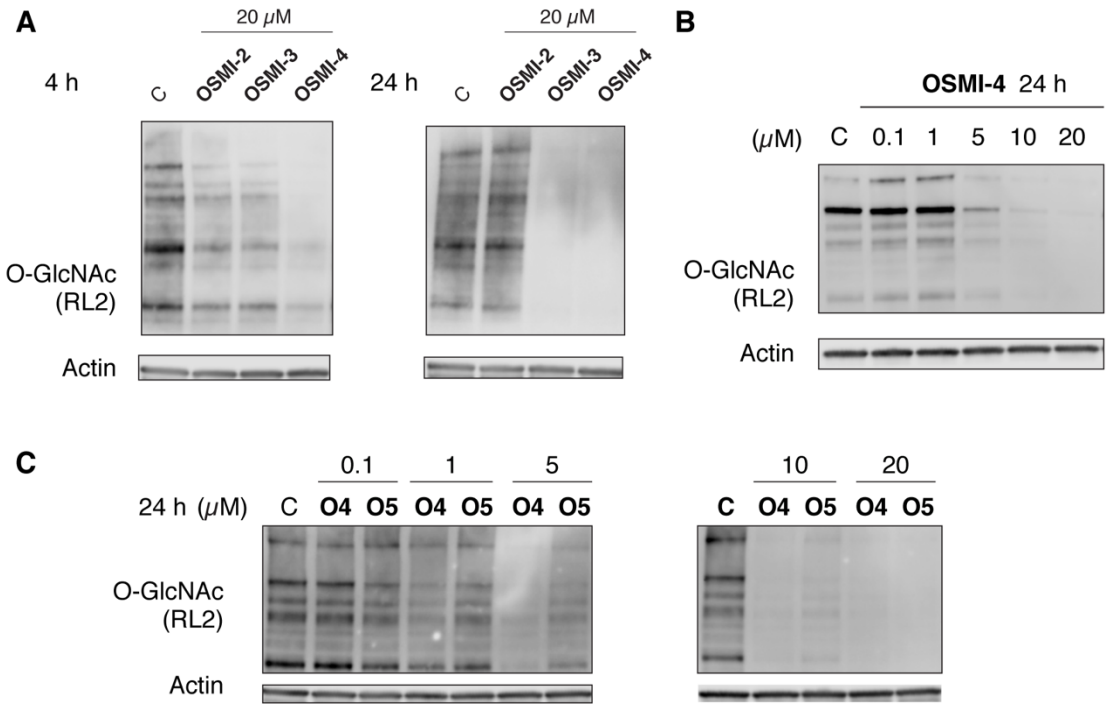


Fig 3.3.3 OSMI-4 is the best in class inhibitor. (A) Western blotting of HEK293T cells with OSMI-2/3/4 showed that OSMI-4 decreases O-GlcNAc to the greatest extent. **(B)** OSMI-4 is able to reduce O-GlcNAc with as low as 5 uM concentration. **(C)** OSMI-4 (O4) is better than OSMI-5 (O5) in cells at all concentration tested.

3.4 OSMI-2/3/4 shows on-target activity and selectivity towards OGT.

In addition to glycosylating Ser/Thr residues of nuclear and cytoplasmic proteins, OGT catalyzes cleavage of the cell-cycle regulator HCF-1 by glycosylating a glutamate in the HCF-1 cleavage sequence [148-149]. An OGT inhibitor would be expected to block cleavage. Indeed, we observed a decrease in HCF-1 cleavage products and the appearance of uncleaved HCF-1 in cells treated with OSMI-2/3/4 even after 3 days of inhibition. (Figure 3.4.1a and b).

A concern with inhibitors is target promiscuity, and it has been previously shown that OSMI-1 has an unknown off target that led to cellular cytotoxicity. Although it is possible to profile kinase inhibitors' selectivity against a panel of kinases, there are no

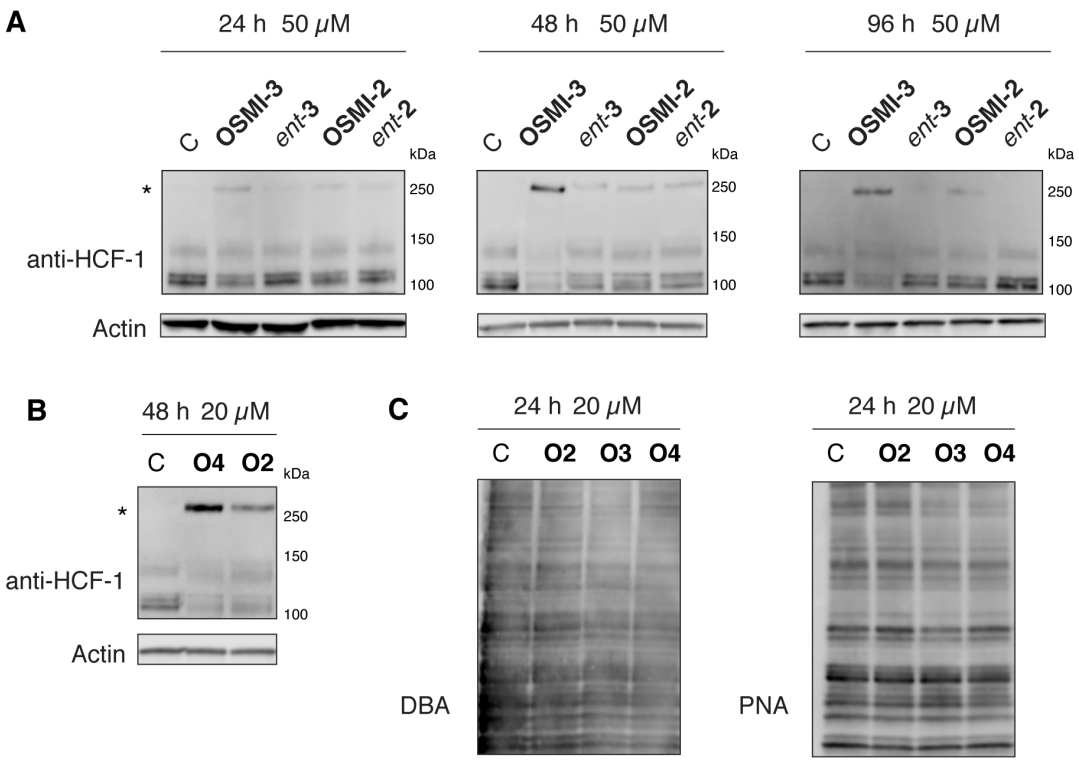


Fig 3.4.1 OSMI compounds are selective inhibitors of OGT. (A) OSMI-2 and OSMI-3 blocks HCF-1 cleavage (uncleaved HCF-1 indicated by asterisk) but not their enantiomeric forms. **(B)** OSMI-4 blocks HCF-1 cleavage to a larger extent than OSMI-2. **(C)** OSMI-2/3/4 do not inhibit other GTs as indicated by DBA and PNA blots.

panel of GTs developed yet. One possible method would be to probe the changes in cell surfaces glycan because OGT belongs to the family of GTs and it was previously shown that UDP-5s-GlcNAc may block some GTs involved in mucin type O-glycan synthesis [103]. To test the selectivity of the OSMI series, we tested against 2 different lectins (PNA and DBA) that recognizes β-GalNAc and a series of GTs *in vitro* in collaboration with Kelley Moreman's lab [150]. Significantly, we did not notice any inhibition against the panel of GTs that was tested (data not shown) and we saw minimal changes with PNA and DBA, suggesting that our OSMI-2/3/4 are selective towards OGT and do not inhibit other GTs (Figure 3.4.1c).

Previously, we observed that OSMI-1 is cytotoxic towards cells in 24h of treatment, presumably due to an off-target effect because this cytotoxicity was also observed with PG34, which does not inhibit OGT. Because OSMI-2/3/4 contain the same scaffold as OSMI-1, we wanted to probe whether this new series of OSMI compounds would be cytotoxic to cells at 24h. We first compared cell viability, cytotoxicity, and apoptosis with ApoTox-Glo triplex assay on OSMI-1 and OSMI-2 [151]. This assay quantifies cell viability by measuring the activity of a live cell protease with a fluorogenic cell permeable substrate, quantifies cytotoxicity with a non-cell permeable fluorogenic substrate that gets cleaved when cells release proteases upon loss of membrane integrity, and quantifies apoptosis with a luciferase that measures the level of cleaved Caspase-3/7, a hallmark of apoptosis. This assay was also used in a previous OSMI-1 study. Consistent with previous results, OSMI-1 has a huge and significant decrease in cell viability (~50%) and showed no effect on apoptosis (Figure 3.4.2a). Interestingly, we did not observe an increase in cytotoxicity, as was previously observed. This discrepancy in cytotoxicity can be explained by the different cell line used (HEK293T vs CHO cells). For example, HEK293T cells might not have a very responsive protease upon membrane disruption or the off-target of OSMI-1 might not be present in HEK293T cells. Nevertheless, we did not observe changes in cell viability/toxicity/apoptosis with OSMI-2 treatment, further supporting the selectivity of this new OSMI series.

To further evaluate and confirm cytotoxic effects with the OSMI inhibitors, we used a different cytotoxic assay that measures disruption of membrane integrity by quantifying the amount of DNA from dead cells [152-153]. We found no significant changes at all concentrations across the OSMI series (Figure 3.4.2b). These results suggest that there is no disruption or ruffling of the cell's membrane with the OSMI series (no cytotoxic effect) and the decrease in cell viability

by OSMI-1 is due to growth inhibition. These results collectively indicate that the new OSMI series are not cytotoxic nor do they inhibit growth in the 24 h time point.

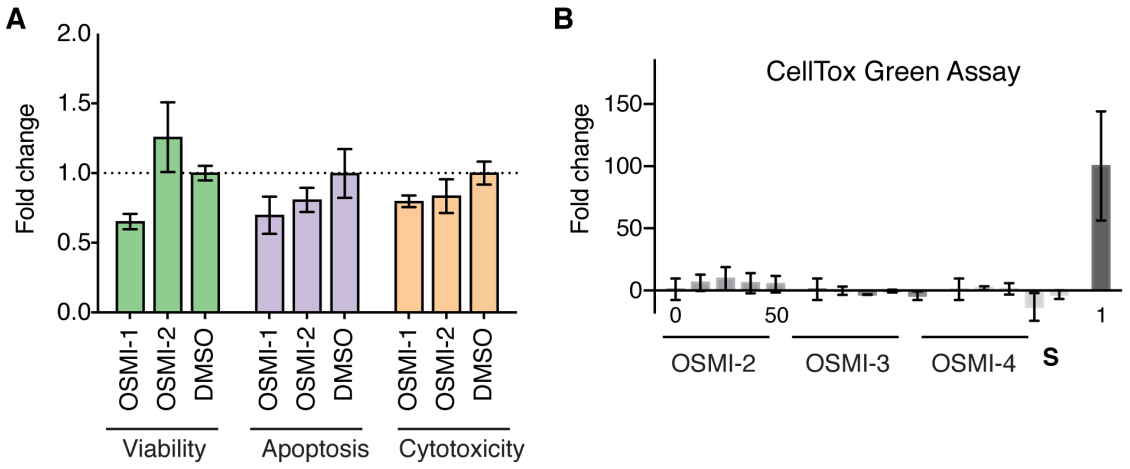


Fig 3.4.2 OSMI compounds are not cytotoxic. (A) OSMI-2 unlike OSMI-1 inhibits cell growth as measured by viability, apoptosis and cytoxicity. **(B)** Treating HEK293T cells for 24 hours with increasing concentrations of OSMI-2, OSMI-3 or OSMI-4 (0-50 μM) did not result in cytotoxic cell death. "S" denotes staurosporine**.** Error bars represent s.d. of three replicates.

Because OGT knockdown is known to decrease cell proliferation [154], we also visually monitored the effects of OSMI-2/3/4 on cell growth in culture over a 96 h period on HCT116 cells. We could not use HEK293T cells because HEK293T cells do not grow well at low confluency and do not adhere well during the necessary wash step. We observed reduced growth of cells over time (Figure 3.4.3), as expected from the knockdown results. Notably, we only began to observe a reduction in cell viability with OSMI-2 and OSMI-4 after 2 days of treatment, with OSMI-4 having a greater effect than OSMI-2 after 96 h. Consistent with the effect of OSMI-3 on cell viability in HEK293T cells, we also noticed a reduction in cell viability with OSMI-3 by 24 h.

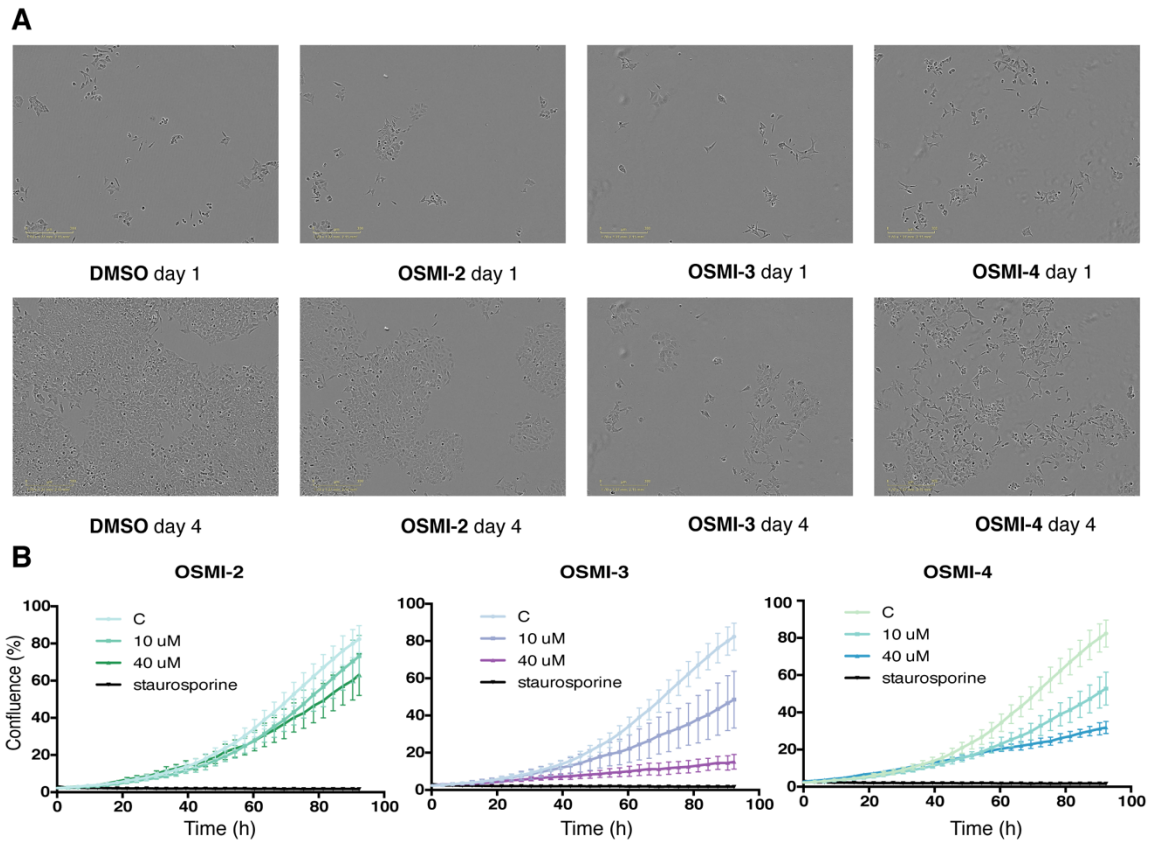


Fig 3.4.3 OSMI compounds are not cytotoxic. (**A**) Treatment with OGT inhibitors reduces cellular confluency. Images of HCT116 cells in culture at 1 and 4 days after treatment with DMSO, or 40 μM of OSMI-2, 3 or 4. At 24 hours, confluency is similar to DMSO control for all compounds. At 96 hours, confluency is reduced for all compounds relative to DMSO control, with the greatest reductions observed in cells treated with OSMI-3. (**B**) Graphical representations of confluence after treatment with DMSO (negative control), 10 or 40 μM of **1b**, **2b**, or **4b**, or 10 μM staurosporine (positive control).

3.5 Quantitative proteomics with OGT inhibition reveals strong connection between OGT inhibition and sterol metabolism.

It has been known that O-GlcNAc can affect protein stability, like EZH2, p53, and some nascent polypeptide chains [72, 155-156]. Despite the increasing evidence that O-GlcNAc is involved in the stability of proteins, there have been no global studies to investigate how widespread this stabilization extends to. We have now established that OSMI-2/3/4 are acute, potent and selective

inhibitors of OGT. With those inhibitors, we can now explore how changing O-GlcNAc can affect protein stability at the whole proteome level with quantitative proteomics without affecting the level of OGT proteins. Using inhibitors will also allow us to delineate the biology of O-GlcNAc vs the scaffolding property of OGT. To provide deeper insights into protein stability, we performed quantitative proteomics with OSMI-2/3/4 and used *ent*-OSMI-2/3 as negative controls (Figure 3.5.1a).

To perform the quantitative proteomics experiments, we used the Tandem Mass Tag (TMT) labeling method. TMT reagents are NHS-ester-based isobaric tags that would react with free amine from digested peptide [157]. Each TMT-tag has the same total mass, but with different mass deviation in the reporter and balancer regions of the molecule. Under MS^2 or MS^3 fragmentation conditions, the reporter ion would fragment out from the labeled TMT-peptide (Figure 3.5.1b). The mass of the reporter ion would allow us to differentiate the samples and the intensity of the reporter ion will report on the intensity of the peptide and proteins.

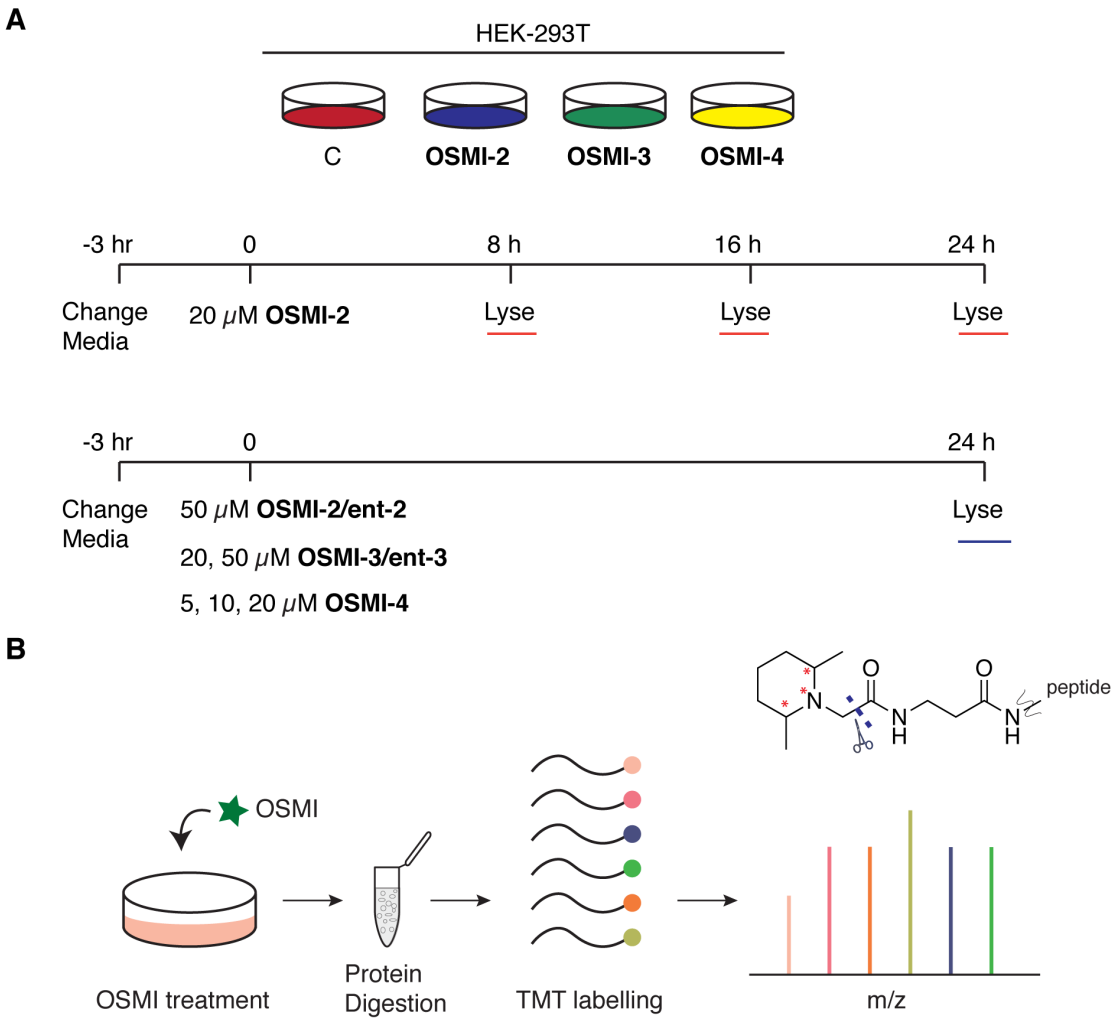


Fig 3.5.1 Schematic of quantitative proteomics with OGT inhibitors. (**A**) Schematic of proteomics performed with OSMI-2, 3, 4. Media was changed 3 h before treatment. **(B)** TMT labeling on proteins allow quantitation up to ~11 samples. Each samples has a different TMT tag that has a different isotopic mass, which can be detected on MS^3. The intensity of the tag quantifies the intensity of the peptide. Protein intensity is measured by the sum of all the peptides corresponding to the protein.

To determine an optimal time frame, we first performed quantitative proteomics with 20 μM OSMI-2, a concentration that lowers O-GlcNAcylation by ~50% at 2 h, over 8, 16 and 24 h. At all the time points, we noted < 10 proteins that have changed significantly in abundance from ~6000 proteins detected (Figure 3.5.2a). Notably, we observed reciprocal changes in the abundance of OGT and OGA, which may explain the recovery of O-GlcNAcylation with OSMI-2.

Compensatory changes in OGT and OGA abundance have been observed previously when levels or activity of these proteins are perturbed, supporting on-target inhibition of OGT by OSMI-2. These proteomics results additionally suggest that OSMI-2 is specific and selective towards OGT as we would expect to see more changes in protein abundance if there are multiple off targets. Nevertheless, the lack of global changes could also be explained by the almost full recovery of O-GlcNAcylation at 24 h with 20 μM OSMI-2. Next, the fact that OGT protein already increased by 8 h (we further confirmed that OSMI-2 increased OGT abundance by 4 h) suggests that OSMI-2 is a very fast acting inhibitor and can be used as a tool to probe cellular responses upon acute changes in O-GlcNAcylation.

Among the increased proteins, we observed a significant increase in the abundance of Baf60c at both 16 h and 24 h. Because *ent*-OSMI-2 did not elicit an increase in the level of Baf60c protein, the increase in Baf60c most likely arises from OGT inhibition (Figure 3.5.2b). Baf60c forms part of the major SWI/SNF chromatin remodeling complex (BAF complex) and is important for mediating interaction between transcription factors and the BAF complex to activate target genes [158-159]. OGT has been previously linked with chromatin remodeling and transcription [160]. The upregulation of Baf60c could represent a novel mechanism in which O-GlcNAc modification can selectively remodel chromatin and activate transcription.

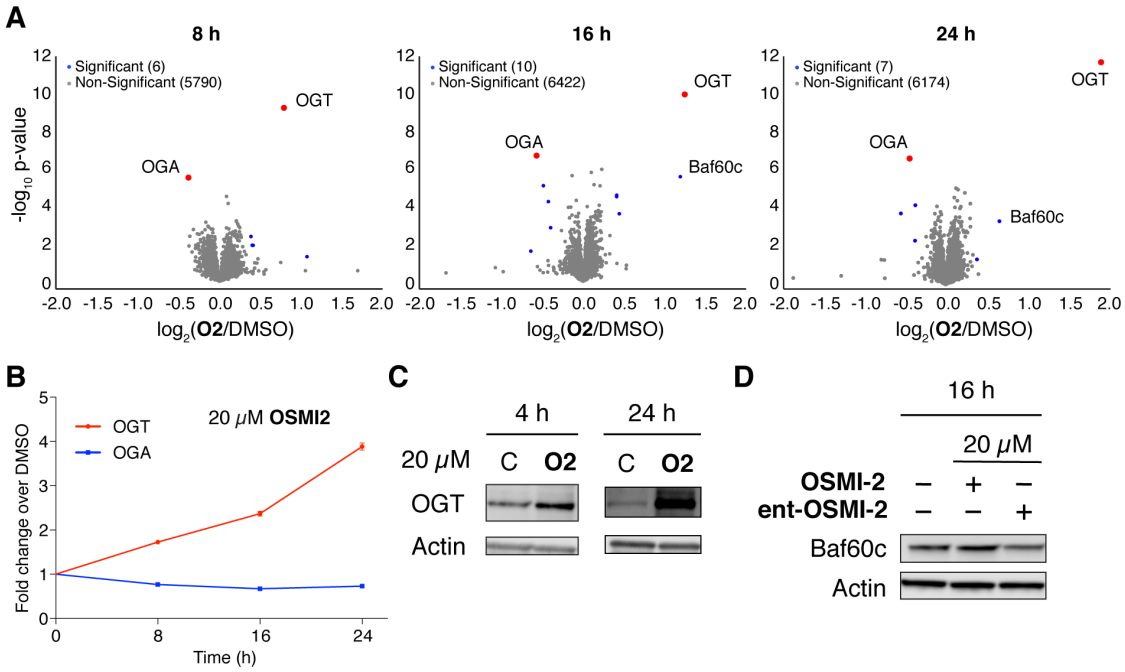


Fig 3.5.2 Proteomic analysis of OGT inhibition by OSMI-2 . (**A**) Volcano plots of quantitative proteomic analysis showing log2 fold change versus –log10 *P* value of OSMI-2 at 20 μM for 8, 16, and 24 hours. Blue circles represent proteins that exhibited greater than 1.3-fold change in abundance and a *p*-value < 0.05 relative to DMSO. (**B**) OGT and OGA showed a time- dependent change in their protein abundance with 20 μM OSMI-2. Error bars represent s.d. of five replicates. **(C)** Western blotting shows an increase in OGT at 4 and 24 hours after treatment of HEK293T cells with 20 μM OSMI-2, confirming the proteomic studies. **(D)** Western blotting shows an increase in Baf60c 16 hour after treatment of HEK293T cells with 20 μM OSMI-2 but not *ent*-OSMI-2.

After establishing that 24 h as an optimal time point for proteomics study, we performed quantitative proteomics with OSMI-3 and 4 at concentrations where reduced O-GlcNAc levels are maintained over 24 h. OSMI-3 and *ent*-OSMI-3 were performed at 50 μM; OSMI-4 were used at 5, 10 and 20 μM. As expected, we observed significantly more changes in protein abundance at concentrations where sustained reduction in O-GlcNAcylation was obtained. (Figure 3.5.3a to c). We also observed reciprocal changes in OGT and OGA in all OSMI compounds, supporting on-target inhibition by OSMI-3 and OSMI-4. Consistent with the loss in inhibition when the stereochemistry is reversed, we did not observe changes in abundance in OGT or OGA with *ent*-OSMI-3.

From the ~6000 proteins quantified, we observed 38 proteins that changed >1.6 fold with OSMI-3 and <1.3 fold with *ent*-OSMI-3. From these 38 proteins, we saw an enrichment in pathway in sterol metabolism and autophagy. With OSMI-4, 86 proteins, most showing increased abundance, changed significantly (Figure 3.5.3b). These proteins included SQSTM1 (sequestosome-1), a protein involved in autophagy that was shown to increase upon conditional deletion of OGT in the liver, as well as additional proteins involved in autophagy. Proteins involved in other processes previously linked to OGT, including transcription and ER stress (Figure 3.5.3d), also changed [55]. We also observed increased abundance of key enzymes involved in the biosynthesis of cholesterol, including 3-Hydroxy-3-methylglutaryl-CoA reductase (HMGCR), which catalyzes the rate-limiting step in cholesterol synthesis, squalene synthase (FDFT1), which catalyzes the first committed step of cholesterol synthesis, and squalene epoxidase (SQLE). This finding suggests a strong connection between OGT activity and sterol homeostasis that led us to further investigate the effects on cholesterol with OGT inhibition.

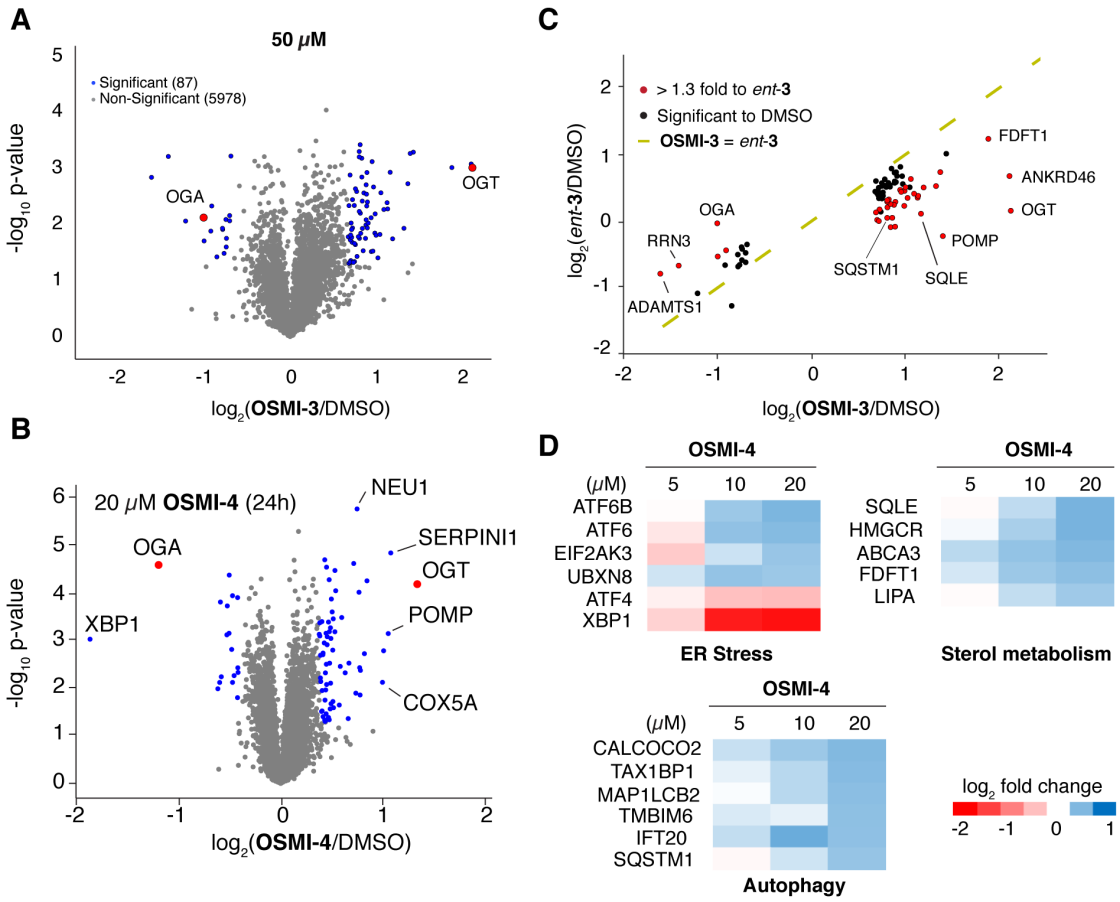


Fig 3.5.3 Proteomic analysis of OGT inhibition by OSMI-3 and OSMI-4 . (**A**) Volcano plots of quantitative proteomic analysis showing log2 fold change versus –log10 *P* value of OSMI-3 or OSMI-4 **(B)** at 50 μM for 24 hours. Blue circles represent proteins that exhibited greater than 1.3-fold change in abundance and a *p*-value < 0.05 relative to DMSO. **(C)** Log_2 Fold change signal for every protein for OSMI-3 (x-axis) and *ent*-3 (y-axis). Dashed line represent equivalent fold change between OSMI-3 and *ent*-OSMI-3 samples **(D)** Clusters of hits from OSMI-4 samples were found centering around ER stress, sterol metabolism, and autophagy.

3.6 Cholesterol levels are potentially regulated by O-GlcNAcylation

All of the proteomics data collectively point to a strong connection of OGT inhibition with sterol metabolism or lipid metabolism. Supporting this connection, a previous study showed that OGT inhibition with 5S-GlcNAc or OGT knockdown resulted in decreased lipid metabolism through reduction in the level of SREBP-1 protein, a key transcription factor that regulates lipid synthesis [161]. HMGCR and SQLE are both key cholesterol metabolic enzymes in which their

protein stabilities are regulated by lanosterol, which is one of the intermediates along cholesterol synthetic pathway, or cholesterol level respectively through well documented feedback mechanisms [162-164]. For example, if there is an increase in cholesterol flux, the subsequent increase in lanosterol or cholesterol level will act to downregulate the cholesterol flux through increased degradation of HMGCR or SQLE proteins to restore cholesterol homeostasis. Because we observed increased abundance in both HMGCR and SQLE proteins, we speculate that OGT inhibition might decrease cholesterol flux and cholesterol levels, which then upregulates the level of SQLE and HMGCR.

To test this hypothesis, we first confirmed the changes in the levels of sterol metabolic enzyme, FDFT1, when OGT is inhibited. We performed immunoblot with anti-FDFT1 and saw a ~4 fold increase in FDFT1 with OSMI-3 and minimal increase with *ent*-OSMI-3 (Figure 3.6a). This suggests that the increase in sterol metabolic enzymes, as observed from the proteomic data, is due to OGT inhibition and not an off-target effect from the OSMI series. We then measured the change in the level of cholesterol. We extracted both the free cholesterol and its esterified form after treating cells with OSMI-3, *ent*-OSMI-3 and lovastatin, which inhibits HMGCR, as a positive control. After 24 h of treatment, we observed that OSMI-3 lowered total cholesterol levels (free cholesterol and cholesterol esters) to the same extent as lovastatin (Figure 3.6b). In addition, *ent*-OSMI-3 did not change total cholesterol levels, suggesting that the decreased in cholesterol levels comes from the inhibition of OGT. As a next step, we plan to verify this result with a different OGT inhibitor like OSMI-4, to confirm this decrease. Nonetheless, this preliminary finding shows that OGT's inhibition decreases cholesterol flux and that the upregulation in the levels of cholesterol metabolic enzymes is probably due to a homeostasis response by cells to maintain cholesterol levels.

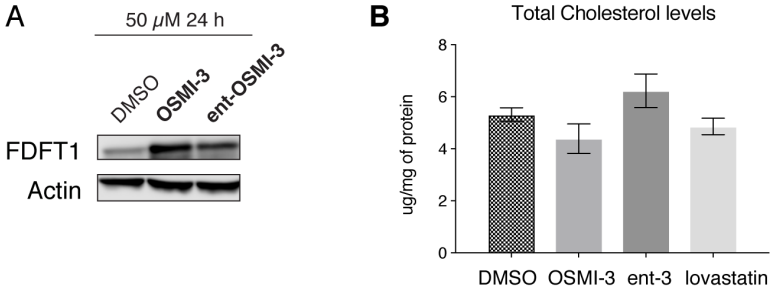


Fig 3.5.3 OGT inhibition affects cholesterol metabolism . (**A**) Western blotting showed an increase in FDFT1 levels with OSMI-3. **(B)** OSMI-3 but not *ent*-OSMI3 decreases total cholesterol levels to the same level as positive control lovastatin.

3.7 Conclusions

In this chapter, we developed an assay that can quantify global O-GlcNAc levels in cells and compared that with western blotting method. We adopted and optimized a two-step chemoenzymatic assay that uses mutant GalT Y289L to tag O-GlcNAc moiety with azido functional group, which can then be labeled with a fluorophore and imaged by in-gel fluorescence. With this assay, we found that western blotting, which is cheaper and less time consuming, agreed very well with the chemoenzymatic method. This prompted us to conclude that western blotting with O-GlcNAc antibodies is semi quantitative and can be used to determine the potency of OGT inhibitors in cells. Nevertheless, the chemoenzymatic tool developed here could have other uses. For example, by changing the fluorophore to a biotin, we could adopt this assay to unbiasedly enrich for O-GlcNAclated proteins. This assay could also be used to detect O-GlcNAcylation on specific proteins where O-GlcNAc antibodies fail to detect.

Using the corresponding ester derivatives to enhance cellular penetration, we showed that OSMI-2/3/4/5 is able to inhibit OGT in cells. The cellular potency is in the order of OSMI-4 > OSMI-5 > OSMI-3 > OSMI-2 and this correlates very well with our *in vitro* binding data.

Additionally, we found that this new series of OSMI compounds, unlike OSMI-1, are not cytotoxic to cells. We believe the cytotoxicity effect seen in OSMI-1 might have come from an off-target binding that requires the inclusion of both the furan and thiophene group in OSMI-1. Because the enantiomeric form of OSMI-2 and OSMI-3 did not inhibit OGT in cells, we reasoned that *ent*-OSMI-3 can be used as a control for possible off-targets. Additionally, all OSMI compounds reduce cell viability and proliferation over 96 h, consistent with previous OGT knockdown studies [165-166].

We next queried changes on the global proteome upon OGT inhibition. We quantified ~6000 proteins for every samples. Notably, we observed an increase in the levels of OGT and decrease in OGA as early as 8 h, presumably to maintain O-GlcNAc homeostasis. Furthermore, we saw an increase in the levels of sterol biosynthetic enzymes with OSMI-2, OSMI-3 and OSMI-4 but not with their inactive enantiomeric forms. We also confirmed that cholesterol levels are decreased under OGT inhibition. Because of a negative feedback mechanism by cholesterol and its metabolites, we speculated that the decrease in cholesterol might have accounted for the upregulation in sterol biosynthetic enzymes. These results led us to speculate whether O-GlcNAc could regulate other forms of lipid biosynthesis other than just sterol metabolism. Indeed, there has been a recent report linking O-GlcNAc to changes in lipid metabolism with OGT knockdown [161]. To address that question, we plan to perform lipidomics in the presence of our inhibitors.

We now provided inhibitors that are not only fast acting in cells but are also the best in class inhibitors for OGT. We revealed an intriguing link between sterol metabolism with OGT inhibition after 24 h of inhibition. Some questions still remain. 1) Can the inhibitors be further optimized for use in animals? Possibly-- There is an ongoing collaboration with the NIH to further optimize the OSMI series so that the inhibitors can be used in animals. One possible solution would

be to modify or replace the metabolically reactive thiophene group with a more stable substituted thiophene. 2) Is there a general decrease in cholesterol flux and why does cholesterol flux decrease with OGT inhibition? One way to address a decrease in cholesterol flux would be to perform a time course targeted metabolomic experiment that focuses on cholesterol metabolites. 3) We have now shown the effects on the level of proteins at the proteome scale. How about other signaling effects that usually happen in the timeline of minutes to hours when OGT is inhibited? Next chapter will attempt to cover that.

3.8 Methods

Cells and Reagents

HEK293T, LNCaP, PC3 and HCT116 were purchased from American Type Culture Collection (ATCC). HEK293T cells were grown in DMEM media supplemented with 10% FBS and 1X Penicillin-Streptomycin solution (Corning). LNCaP and PC3 were maintained in RPMI media supplemented with 10% FBS, and HCT116 cells were cultured in McCoy's 5A media supplemented with 10% FBS. Antibodies against OGT were obtained from Cell Signaling Technology (24083S), against O-GlcNAc (Rl2) and Actin from Abcam (ab2739, ab49900, respectively) and against HCF-1 from Bethyl Laboratories (A301-400A).

Cell culture

Cells were plated and grown in their respective media until they reached a confluency of 60-80%. For HEK293T cells, media was changed 3 hours before compounds were added; compounds were added directly to each well at the indicated concentrations. For all other cell lines, compounds were dissolved in 0.2 mL of fresh media and then added directly to the cells.

Preparation of cell lysates and western blotting

Cell lysates for from HEK293T, LNCaP, HCT116 and PC3 cells were prepared for western blotting in the following manner, and all the steps were conducted at 4°C. Cells were washed once with PBS, collected in PBS, centrifuged and 100 µL of cell lysis buffer was added (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1% Triton X-100, 0.1% SDS (2% for HEK293T cells), 150 mM NaCl supplemented with Complete protease inhibitor mixture (Sigma), complete phosphatase inhibitor mixture (Sigma), and 50 µM PUGNAc (Sigma). After this, samples were lysed using a Bioruptor sonicator, centrifuged at 14,000xg for 5-10 minutes, and supernatants were collected. Protein concentration was determined using the BCA assay and 10-25 µg of sample was loaded on 4-15% or 4-20% SDS-PAGE (Bio-Rad). Samples were transferred to a nitrocellulose membrane (Bio-Rad). Membranes were blocked with 5% bovine serum albumin in TBST and probed with primary antibodies overnight at 4ºC and for 1h at RT with peroxidase conjugated seocondary antibodies, as indicated in each figure. Blots were developed using enhanced chemiluminescence (Pierce) and visualized with an Amersham Imager 600.

GalT gel fluorescence

Cell lysates from PC3 cells were used for GalT click labeling. GalT click labeling were performed according to manufacturer's instructions. Briefly, cell lysate were chloroform/methanol precipitated and redissolved in 1% SDS 20 mM HEPES buffer to a concentration of ~5 mg/ml. After this, GalT mixture was added and incubated overnight at 4ºC. On the next day, PNGase F (around 1 ul for 50 ul reaction) was added for an hour at room temperature to remove N-linked oligosaccharide. Click chemistry was performed by adding Alexa-647 alkyne to a final concentration of 500 µM, Ascorbate to 5 mM and premixed CuTHPTA (1:5 ratio) to concentration of 300 µM (based on the concentration of Cu), shake on a thermomixer at 600 rpm for an hour at room temperature. Proteins were then further precipitated by performing trichloroacetic acid

protein precipitation (see http://www.its.caltech.edu/~bjorker/TCA_ppt_protocol.pdf). Samples were transferred to 4-15% SDS-page and visualized on a Typhoon gel imager (Typhoon FLA 9500).

Measurement of cell viability and live-cell imaging

HCT116 cells were plated into 384-well plate one day prior to treatment with compounds (300 cells per well). Compounds were dispensed using a D300e Digital Dispenser (Tecan). Cells were imaged every 3 hours using an IncuCyte ZOOM instrument (Sartorius) to generate growth curves. After four days of treatment, cell viability was assessed using the CellTiter-Glo® Luminescent Cell Viability Assay (Promega). Luminescence signal was recorded using the GloMax® Explorer System (Promega). For Cytotoxicity, HEK293T cells were plated into a 96-well plate two days prior to treatment with compounds at 80% confluency. The number of dead cells were assessed by CellTox™ Green Cytotoxicity Assay (Promega).

Quantitative proteomics

For quantitative analysis of the proteome, HEK293T cells in fresh media (changed 3 hours before treatment) were treated with 20 μM OSMI-2 or DMSO in five biological replicates (5 treatment and 5 DMSO at each time point) for 8 hours, 16 hours or 24 hours; DMSO, 50 μM OSMI-3, 50 μM ent-3 in triplicates; DMSO, 5 and 20 μM OSMI-4 in triplicates, and 10 μM in duplicates (11-plex). At the indicated time point, cells were lysed in lysis buffer (2 % SDS 50 mM Tris-HCl and 150 mM NaCl), sonicated (BioDisruptor) and protein concentration determined with the BCA assay. Samples were reduced with 5 mM DTT for 45 mins at 60ºC, then alkylated with 14 mM iodoacetamide for 45 mins at room temperature in the dark. Then 100 μg of protein was precipitated using chloroform/methanol [167]. Protein pellets were resuspended in 200 mM HEPES pH 8.5 to 1 mg/mL. Proteins were digested with LysC (Wako) (substrate:enzyme=100) overnight

at 37ºC and then with sequencing grade Trypsin (Promega) (substrate:enzyme= 100) for 6 hours at 37ºC. The resulting peptide solutions were then labelled with TMT 10/11-plex reagents (Thermo Scientific) for 1.5 hours at room temperature. Reactions were stopped by addition of 5% hydroxylamine for 30 minutes. Equal amounts of peptide samples were combined, dried by vacuum centrifugation and desalted on a Waters C18 solid phase extraction Sep-Pak. TMT-labeled peptide samples were fractionated via basic-pH reverse-phase (BPRP) HPLC to 96 fractions and then consolidated to 12 fractions. These fractions were subsequently acidified with 1% formic acid, vacuum centrifuged to near dryness and desalted with C18 stagetips [167]. Dried peptides were resuspended in 5% acetonitrile/5% formic acid for LC-MS/MS processing.

Mass-Spectrometry Analysis

Our mass spectrometry data were collected using an Orbitrap Fusion Lumos mass spectrometer (ThermoFisher Scientific, San Jose, CA) coupled to a Proxeon EASY-nLC 1200 liquid chromatography (LC) pump (ThermoFisher Scientific). Peptides were separated on a 100 μm inner diameter microcapillary column packed with 35 cm of Accucore C18 resin (2.6 μm, 150 Å, ThermoFisher). For each analysis, we loaded ~2 μg onto the column.

Separation was in-line with the mass spectrometer and was performed using a 3 hr gradient of 6 to 26% acetonitrile in 0.125% formic acid at a flow rate of ~450 nL/min. Each analysis used an TMT-based TMT method [168-169] which has been shown to reduce ion interference compared to MS2 quantification [170]. The scan sequence began with an MS1 spectrum (Orbitrap analysis; resolution 120,000; mass range 400−1400 *m/z*; automatic gain control (AGC) target 5×10^5; maximum injection time 100 ms). Precursors for MS2/MS3 analysis were selected using a Top10 method. MS2 analysis consisted of collision-induced dissociation (CID); AGC 2.0×10^4; normalized collision energy (NCE) 35; maximum injection time 120 ms; and isolation window of 0.4 Da.

Following acquisition of each MS2 spectrum, we collected an MS3 spectrum using our recently described method in which multiple MS2 fragment ions were captured in the MS3 precursor population using isolation waveforms with multiple frequency notches [169]. MS3 precursors were fragmented by high energy collision-induced dissociation (HCD) and analyzed using the Orbitrap (NCE 65; AGC 1.5×10^5; maximum injection time 150 ms, resolution was 50,000 at 400 Th, isolation window 0.7 Da).

Mass spectra were processed using a SEQUEST-based pipeline [171] Spectra were converted to mzXML using a modified version of ReAdW.exe. Database searching included all entries from the human UniProt database. This database was concatenated with one composed of all protein sequences in the reversed order. Searches were performed using a 50 ppm precursor ion tolerance for total protein level analysis. The product ion tolerance was set to 0.9 Da. These wide mass tolerance windows were chosen to maximize sensitivity in conjunction with Sequest searches and linear discriminant analysis [171-172]. TMT tags on lysine residues and peptide N termini (+229.163 Da) and carbamidomethylation of cysteine residues (+57.021 Da) were set as static modifications, while oxidation of methionine residues (+15.995 Da) was set as a variable modification.

Peptide-spectrum matches (PSMs) were adjusted to a 1% false discovery rate (FDR) [173-174] PSM filtering was performed using a linear discriminant analysis, as described previously, [171] while considering the following parameters: XCorr, ΔCn, missed cleavages, peptide length, charge state, and precursor mass accuracy. For TMT-based reporter ion quantitation, we extracted the signal-to-noise (S:N) ratio for each TMT channel and found the closest matching centroid to the expected mass of the TMT reporter ion. PSMs were identified, quantified, and collapsed to a 1% peptide false discovery rate (FDR) and then collapsed further to a final protein-level FDR of 1%.

Moreover, protein assembly was guided by principles of parsimony to produce the smallest set necessary to account for all observed peptides.

Peptide intensities were quantified by summing reporter ion counts across all matching PSMs, as described previously [169, 175]. Briefly, a 0.003 Th window around the theoretical m/z of each reporter ion was scanned for ions, and the maximum intensity nearest the theoretical m/z was used. PSMs with poor quality, MS3 spectra with TMT reporter summed signal-to-noise ratio less than 100, or no MS3 spectra were excluded from quantitation, and isolation specificity ≥ 0.7 was required [175].

Chapter 4

Contributions: The chapter involves a close collaboration with Paul Boutz and Joao Paulo. Paul Boutz has been a great mentor who helped and guided me on how to perform the RNA-seq and analysis, and the RT-qPCR experiments. Joao Paulo, as mentioned in chapter 3, helped me on the phosphoproteomics preparation. Chapter 4 is mainly adapted from the paper I recently submitted to Nuclei Acid Research.

O-GlcNAc regulates gene expression by controlling detained intron splicing

4.1 Introduction

O-GlcNAc transferase (OGT), a glycosyltransferase that catalyzes the post-translational addition of O-linked N-acetylglucosamine (O-GlcNAc) to serine and threonine side chains of proteins, is required for viability of all mammalian cells [83-84]. Four key features distinguish OGT as a unique glycosyltransferase that is distinct from canonical eukaryotic glycosyltransferases, which act exclusively in the secretory pathway to assemble the oligosaccharides found on cell surface glycoproteins. First, OGT and its substrates are found in the nucleus and cytoplasm and not in compartments of the secretory pathway such as the endoplasmic reticulum and the Golgi body [176]. Second, OGT has over a thousand protein substrates that belong to virtually all classes of proteins in the cell, including structural and trafficking proteins, transcription factors, epigenetic regulators, components of the ribosome, the proteasome, and the kinome [54, 133, 177-178]. Third, the modifications installed by OGT are dynamic due to the presence of a dedicated glycosidase, O-GlcNAcase (OGA), which removes O-GlcNAc modifications [179]. In mammals, OGT and OGA are solely responsible for dynamic cycling of O-GlcNAc modifications [60, 180]. Fourth, O-GlcNAc levels change with nutrient conditions and are especially responsive to glucose levels [181]. Because the concentration of OGT's substrate, UDP-GlcNAc, increases with nutrient abundance, OGT is proposed to transduce overall nutrient levels, as reflected in UDP-GlcNAc concentration, into a signaling response that affects myriad cellular pathways [55, 182-183]. These pathways include phosphorylation networks that control metabolism, cell cycle regulation, and transcriptional regulation [57].

Many signaling processes act at short time scales. Because compounds that rapidly and specifically reduce O-GlcNAc through inhibition of OGT were not available until recently, studies that interrogate OGT function have primarily focused on changes that occur long after an initial perturbation has been applied (~24-48 h) [77, 98]. Much has been learned about OGT, but the

long time frame of typical studies is a limitation: some of the observed pathway changes may be downstream of the initial changes, and important effects may have been missed due to compensatory changes that occur prior to interrogation.

We have developed a class of OGT inhibitors useful for interrogating acute cellular changes that occur when OGT function is disrupted [45, 103]. Here we probe system-wide effects of OGT inhibition at short time points. We initially asked how the phosphoproteome changes because numerous studies have shown changes in protein phosphorylation when O-GlcNAc levels are reduced, suggesting extensive crosstalk between phosphorylation and O-GlcNAc signaling pathways [49, 184-185]. Quantitative phosphoproteomics showed changes in a large number of RNA splicing factors upon acute inhibition. Deep sequencing of polyadenylated RNA to probe how O-GlcNAc affects splicing showed distinct effects on detained introns, a specific form of alternative splicing.

Most introns are removed co-transcriptionally—that is, they are completely spliced prior to transcriptional termination and polyadenylation. In contrast, detained introns (DIs) are individual introns that remain in otherwise completely spliced, polyadenylated (post-transcriptional) messages, causing them to be detained in the nucleus rather than exported to the cytoplasm for translation [186]. Many human gene transcripts contain detained introns, and incomplete splicing of these transcripts serves to control levels and timing of expression of productive (protein-coding) mRNAs. Based primarily on analysis of *Clk1* and *Clk4*, it was initially proposed that detained intron transcripts serve as a nuclear reservoir of potentially productive pre-mRNA that can be spliced if needed to increase protein abundance rapidly; indeed this mode of action was recently shown to be important for neuronal activity responses [187-189]. However, evidence now shows that not all detained introns function in this way. For

example, some detained intron transcripts represent a "dead-end" because they cannot be spliced to functional message. They undergo decay in the nucleus or are spliced to a non-functional transcript in response to an environmental signal [186, 190-193]. Regardless of how detained intron pathways function for any given gene, it is believed that detained intron splicing serves as a mechanism to fine tune mRNA abundance without changing basal levels of transcription. By fine tuning gene expression, detained intron splicing provides an important mechanism by which cells carry out cell state transitions or homeostatic responses [194]. Consistent with this idea, detained introns are frequently found in genes expected to be sensitive to environmental changes, including genes important in metabolism, stress responses, and differentiation [186, 190].

A central hypothesis is that O-GlcNAc serves as a signal to integrate metabolism with cell state, but how it does so is not known [195]. We found two major, temporally distinct responses to OGT inhibition. After a 2 hour inhibitor treatment, cells responded to O-GlcNAc perturbation with highly specific changes in *OGT* and *OGA* splicing pathways, and we have shown that the resultant changes in splicing inversely regulate their productive mRNA levels. These splicing changes alter OGT and OGA protein levels to buffer changes in O-GlcNAc. Second, after 6 hours of OGT inhibitor treatment, over 80% of detained introns decreased, a remarkable response suggesting a coordinately regulated program for cell state transition. We conclude that when the initial rapid buffering response does not return cells to O-GlcNAc homeostasis, there are widespread changes in mRNA levels for almost all genes subject to detained-intron splicing control. Because we did not alter nutrient levels in these studies, our studies establish that O-GlcNAc is the direct signal for a nutrient-dependent response that changes gene expression by altering splicing pathways.

4.2 Acute OGT inhibition changes splicing factor phosphorylation

We hypothesized that changes in phosphorylation would identify processes rapidly affected by acute OGT inhibition. OSMI-2, as mentioned in chapter, is a cell-permeable OGT inhibitor with a submicromolar K_d and a well-defined binding mode characterized by X-ray crystallography (chapter 2 and 3) [45]. This compound was shown previously to reduce global O-GlcNAc levels by 4 h of treatment in a number of cell lines, including HEK-293T cells [45, 196]. Because shorter treatment times were not examined, we first asked whether OSMI-2 reduced O-GlcNAc at earlier time points.

We assessed O-GlcNAc levels after OSMI-2 treatment by immunoblotting with a commonly used antibody and observed a substantial decrease by 0.5 h (Figure 4.2.1a). We also performed a pulldown of O-GlcNAc-modified proteins after a 0.5 and 1 h incubation with OSMI-2 and observed a ~30% and ~50% depletion in protein abundance, respectively, in good agreement with the immunoblotting data (Figure 4.2.1b). Although changes in global O-GlcNAc levels may underestimate the extent of changes in O-GlcNAc for some proteins, and overestimate them in others, we were satisfied that there was a sufficient perturbation by 0.5 h of inhibitor treatment to use this time point in further studies.

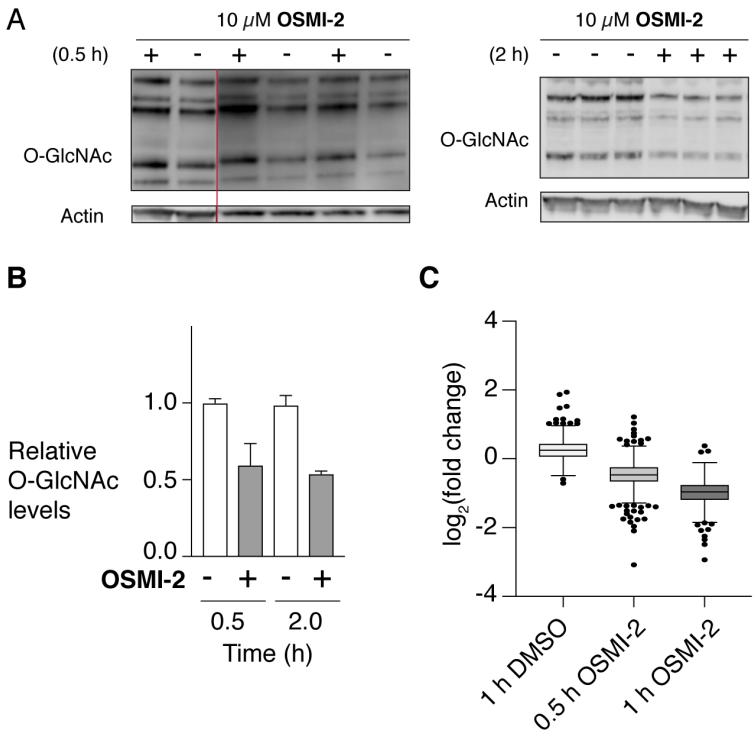


Fig 4.2.1 OSMI-2 inhibits OGT in cells as fast as 0.5 h . (**A**) **)** Western blot for O-GlcNAc levels (α-RL2) after treatment of cells with inhibitors at 10 μM OSMI-2 for 0.5 h (left) or 2.0 h (right). Triplicate lanes for each treatment condition represent biological replicates. **(B)** Barplot derived from densitometry from **(A)**, errors bars are shown in s.d. **(C)** Box plot of protein abundance, enriched with a combination of RL2 and CTD110.6, after treatment with OSMI-2 for various time point. 1 h DMSO indicates the protein abundance of 1 h DMSO compared to 0.5 h DMSO. DMSO does not decrease global O-GlcNAcylation.

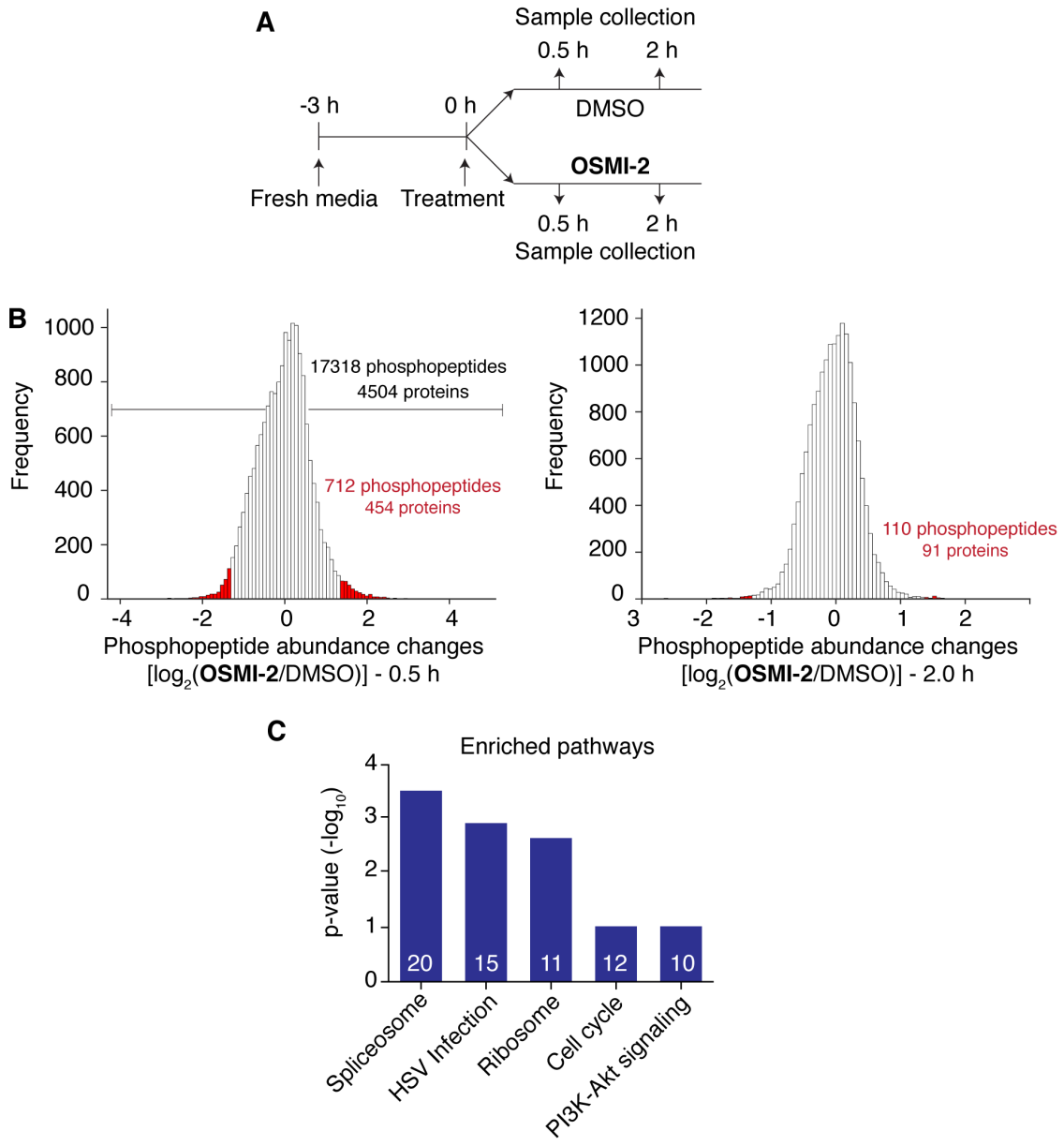


Fig 4.2.2 Splicing pathway is affected by OGT inhibition . (**A**) Scheme depicting treatment conditions for quantitative phosphoproteomics. **(B)** Distribution of phosphopeptide abundances after 0.5 h (left) or 2.0 h treatment (right) with OSMI-2, relative to DMSO control. Hits showed greater than a 2.5-fold change (2 STD from the median; shown in red) **(C)** KEGG pathway analysis of phosphoproteins that changed significantly by 30 minutes after OGT inhibitor treatment showed that the spliceosome was the most affected pathway.

We asked how inhibiting OGT for 0.5 and 2 h with OSMI-2 affected global phosphorylation networks (Figure 4.2.2a). Using an isobaric-tagging approach (Tandem Mass Tagging, or TMT) for quantitative phosphoproteomics, we identified 17,318 phosphopeptides

belonging to 4,504 proteins, of which 712 phosphopeptides showed at least a 2.5-fold change, corresponding to >2 standard deviations from the mean, at 0.5 h in the compound-treated sample compared with the DMSO control (Figure 4.2.2b). We analyzed the KEGG pathways that were enriched at 0.5 h and found some that were expected, as well as some that were not (Figure 4.2.2c). For example, O-GlcNAc modifications have been shown previously to regulate the cell cycle and play a crucial role in PI3K-Akt signaling, and both of these pathways were enriched [197]. Ribosomal components were also enriched and several studies have reported a link between O-GlcNAc and ribosome biogenesis and function [198-200]. However, the top two enriched pathways, the spliceosome and HSV infection, have received little to no attention as pathways regulated by OGT. One previous study reported that OGT inhibition interferes with HSV replication through an undefined mechanism [90]. Consistent with the fact that HSV replication relies on co-opting the splicing machinery of the host [201-202], when we examined the KEGG pathway genes associated with HSV infection, we found that half of them also belong to the spliceosome pathway. These results implicated RNA splicing as a process that changes rapidly upon OGT inhibition.

We asked whether the phosphorylation changes observed at 2 h recapitulated our observations at 0.5 h, highlighting RNA splicing as an affected pathway. Indeed, there was a good correlation between phosphorylation changes at 0.5 h and 2.0 h (Pearson's correlation r = 0.42; $P \sim 0$, Figure 4.2.3a). We identified a total of 55 splicing-related proteins that contained phosphopeptides showing at least a 2.5-fold change relative to the DMSO control at either 0.5 h (52 proteins), 2 h (14 proteins), or both time points (11 proteins) (Figure 4.2.3b). Analysis of only those phosphopeptides that showed significant fold-changes in the same direction at both

time points identified RNA splicing as the only enriched process (Figure 4.2.3c). The correlation analysis therefore confirmed splicing as the process to examine further.

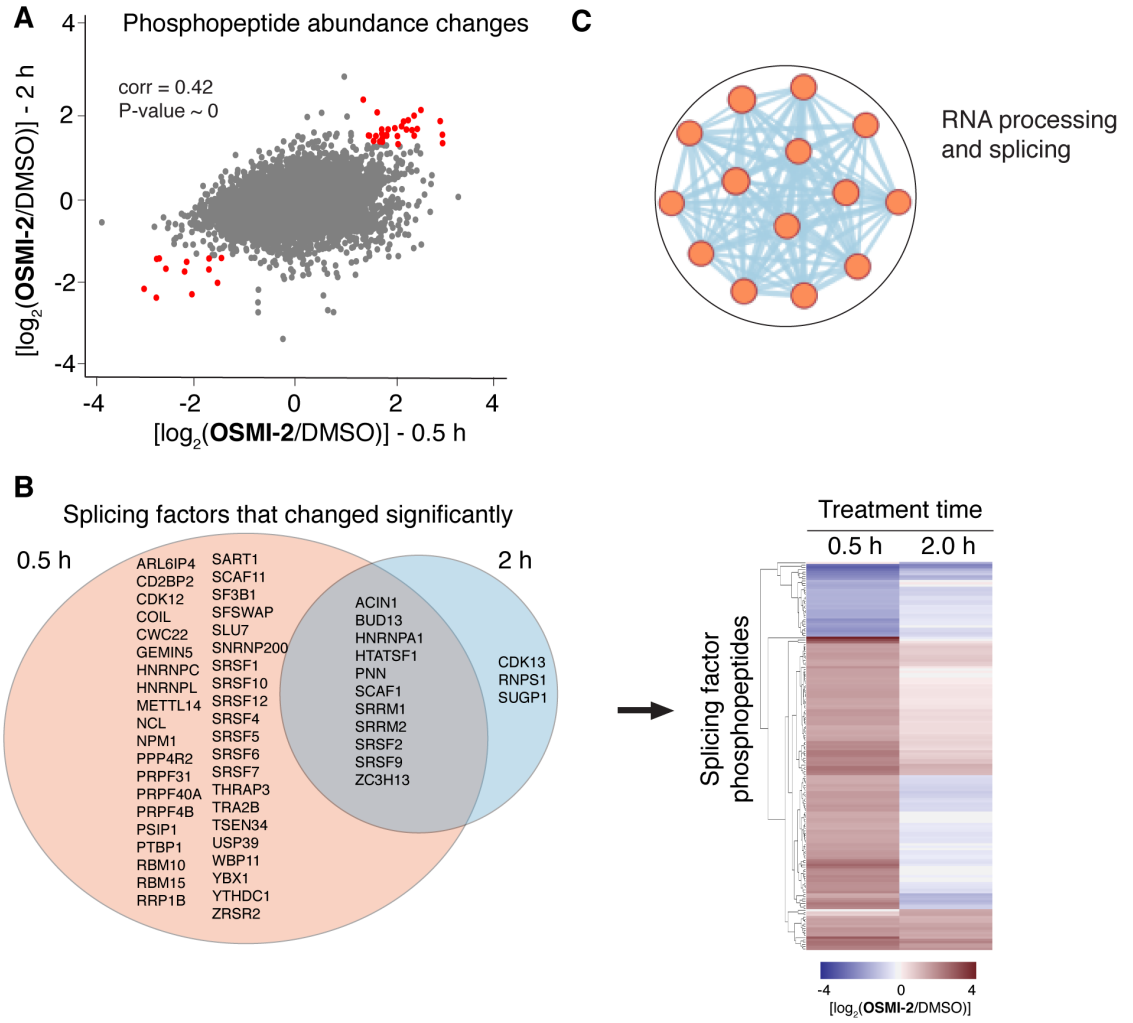


Fig 4.2.3 OGT inhibition predominantly regulates splicing . (**A**) Scatterplot of phosphopeptides showing changes at 0.5 h and 2 h treatment with OGT inhibitor (Pearson's correlation coefficient = 0.42). Red dots represent hits meeting the cutoffs at both 0.5 h and 2 h. **(B)** Gene list and heatmap showing splicing factors with phosphosites that changed significantly at either 0.5 h or 2.0 h treatment with OGT inhibitor. **(C)** Gene Ontology analysis of proteins displaying significant differential phosphorylation at both 0.5 h and 2 h time points. No other module resulted from the analysis

4.3 RNA-seq showed that acute OGT inhibition specifically affects detained introns in transcripts for *OGT* and *OGA*

Prior to this work, connections between RNA splicing and O-GlcNAc were limited. It was known that some splicing factors are O-GlcNAc-modified, but the functional significance of O-GlcNAc on splicing factors has not been examined [54]. Previous studies showed that the *OGT* gene contains a detained intron, and it was known that levels of this intron respond to O-GlcNAc perturbation, indicating that *OGT* transcript splicing is subject to feedback regulation by O-GlcNAc (Figure 4.3.1) [186, 203]. Despite the evidence that O-GlcNAc modification is involved in the splicing regulation of *OGT*, global transcriptome analysis has not been performed to assess whether detained introns in other genes are under O-GlcNAc control, nor have other splicing modalities been examined for O-GlcNAc-regulated changes.

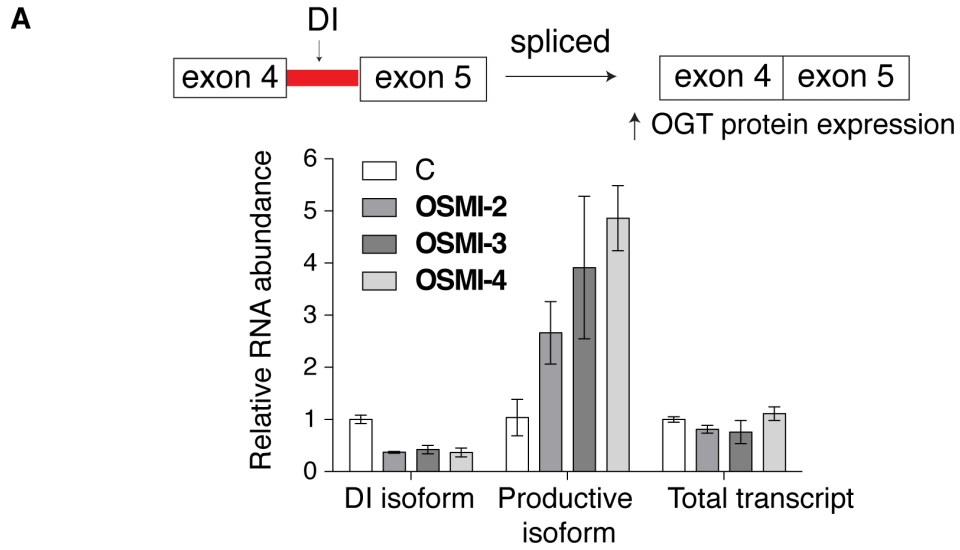


Fig 4.3.1 OGT inhibition affects DI splicing. Bar graph summarizing quantitative PCR results using primers to OGT's detained intron (intron 4) and spliced exons (exon 4 and 5) after 2 h treatment of HEK293T cells with 10 μM OSMI-2, 3, 4. Error bars represent s.d. (n=3).

To test whether any detained introns other than those of *OGT* are affected by acute OGT inhibition, we performed RNA-seq on poly(A)-selected RNA after 2 h or 6 h of treatment with

OSMI-2. The 2 h timepoint was chosen because DI levels for *OGT* respond within this timeframe [203]. At this early timepoint after inhibitor treatment, we observed changes in DI splicing for only three genes: *OGT*, as expected; *OGA*, which encodes the glycosidase that together with OGT controls cellular O-GlcNAc levels; and *QTRTD1*, which encodes a subunit of tRNA-guanine transglycosylase, an enzyme that incorporates the hypermodified guanine analog queuosine into the wobble position of some tRNAs (Figure 4.3.2a). Although we did not follow-up on the latter observation, we point out that the queuosine tRNA modification is proposed to serve as a nutrient-controlled mechanism to fine-tune protein translation [204]. We found it remarkable that the DI splicing changes we observed at 2 h were so selective, with *OGT* and *OGA* among the three genes showing changes. This finding suggests a highly specific, O-GlcNAc-regulated splicing mechanism.

One might have expected the detained introns in *OGT* and *OGA* to show opposite changes in abundance, but instead the DI levels decreased for both genes. It is increasingly appreciated, however, that DIs control gene expression in different ways. For example, some studies have shown that when cellular conditions change, the fraction of nascent transcripts that undergo co-transcriptional splicing versus intron detention can be dynamically controlled to rapidly adjust protein levels [187, 190, 205]. For some genes, co-transcriptional splicing evidently competes with intron detention that, in turn, leads to nuclear decay (Figure 4.3.2b, pathway I). For other genes, the detained-intron pathway leads to productive messages. In these cases, cellular signals may affect co-transcriptional splicing to favor the detained intron pathway and post-transcriptional splicing to productive message may also increase (Figure 4.3.2b, pathway II).

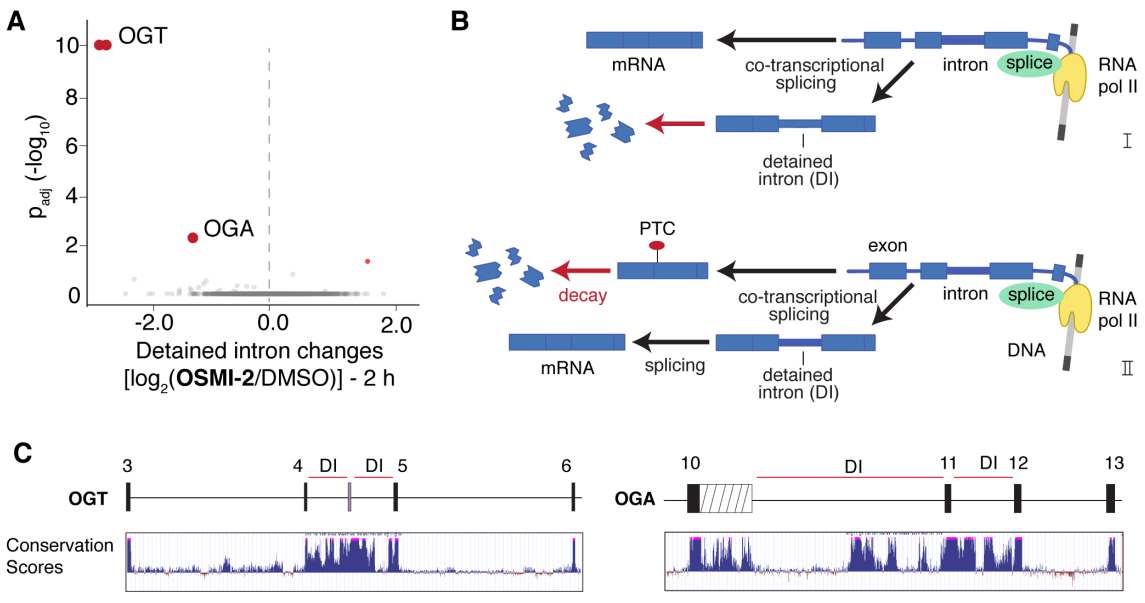


Fig 4.3.2 Acute OGT inhibition specifically regulates the highly conserved DIs of *OGT* and *OGA*. (A) Volcano plot showing fold changes (log_2) in DIs after treatment with 10 μM OSMI-2 for 2 h. Inhibiting OGT specifically decreases DIs in *OGT* and *OGA*. The small red dot shows increased DI levels in *QTRTD1*. **(B)** Scheme showing two different models for formation of productive mRNA through DI splicing. In pathway I, co-transcriptional splicing to make productive message competes with DI retention. In this model, the DI transcript is degraded. In pathway II, co-transcriptional splicing generates unproductive mRNA in competition with formation of a DI transcript that leads to productive mRNA. **(C)** Vertebrate conservation track (UCSC Genome Browser) showing the conservation of DIs in both *OGT* and *OGA*. The blue histogram indicates nucleotide-level conservation across 100 vertebrates, with pink caps indicating regions of 100% identity. Black boxes denote exons, and striped box denotes a short isoform of OGA produced by alternative polyadenylation.

We examined the gene structure of *OGA* and found that it, like *OGT*, contains DIs characterized by several regions of highly phylogenetically conserved intronic sequence (Figure 4.3.2c). As introns typically exhibit low sequence conservation, this finding, combined with our RNA-seq results suggested that a highly-specific, O-GlcNAc-responsive mechanism, which is conserved across vertebrates, controls splicing of both the *OGT* and *OGA* transcripts to regulate protein abundance.

4.4 O-GlcNAc-responsive DI splicing regulates *OGT* and *OGA* mRNA transcript abundance

We examined both the gene sequences and our RNA-seq data for clues to how the *OGT* and *OGA* DIs may function. *OGT* was reported to contain a detained intron between exons 4 and 5, but closer inspection revealed that this region actually consists of two DIs flanking a small cassette exon (Figure 4.4.1a) [186, 203]. This same cassette exon was recently described as belonging to the class of 'decoy exons', exons upon which pre-spliceosome assembly stalls, resulting in two flanking DIs [206]. This decoy exon is also termed a 'poison' exon due to the presence of a premature termination codon (PTC) within its sequence. Fully-spliced mRNA transcripts that contain poison exons undergo nonsense-mediated decay (NMD) upon export to the cytoplasm and translation [207]. *OGA* also contains two conserved DIs that flank an exon, but in this case the exon (exon 11) is found in the productive, protein-coding *OGA* transcript. The RNA-seq data revealed an exon 11-skipped transcript containing a premature stop codon introduced by the reading-frame shift resulting from joining exon 10 and exon 12 (Figure 4.4.1b). The structure of *OGA's* DIs flanking an exon that can be included to produce functional mRNA or skipped to produce non-functional mRNA is reminiscent of the DIs found in *Clk1/Clk4* [187]. The structure of *OGT's* DIs around a poison exon that, when included, produces non-functional mRNA is reminiscent of the DIs found in virtually all of the genes encoding the canonical SR proteins [208]. Thus, *OGT* and *OGA* share a common structure comprising a central cassette exon flanked on both sides by DIs, with the splicing outcome of inclusion versus skipping producing opposite effects on the production of coding mRNA for each gene.

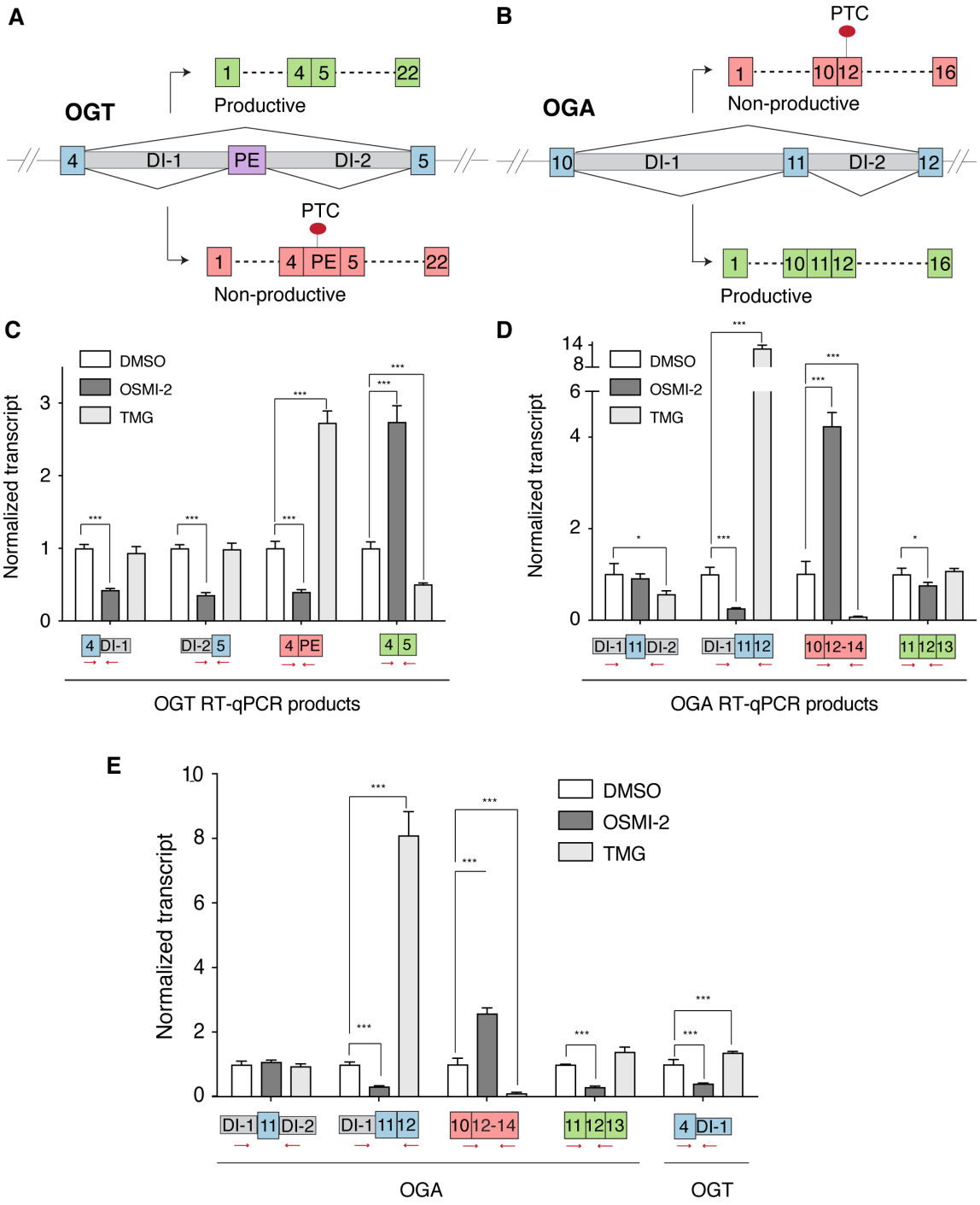


Fig 4.4.1 O-GlcNAc levels regulate *OGT* and *OGA* DI splicing so that exon inclusion/skipping results in opposite effects on productive message. (A) Splice junctions derived from RNA-seq for *OGT* and *OGA* show two predominant fully-spliced isoforms, a functional one for *OGT* **(A)** and a non-functional one for *OGA* **(B)**. HEK-293T cells were treated with DMSO (white bars), OGT inhibitor (10 μM OSMI-2, dark grey) or OGA inhibitor (5 μM TMG, light grey) for 2 h and RT-qPCR was used to measure levels of the indicated products for OGT **(C)** or OGA **(D)**, which report on the presence of DIs as well as skipped or included exons. Red arrows depict location of primers with respect to detected products. Products levels were determined relative to a housekeeping gene (actin)

and were normalized to the DMSO control (n ≥ 3 biological replicates; mean ± s.d, *$P \leq 0.05$, ***$P \leq 0.001$, two-tailed Student's *t*-test). **(E)**. RT-qPCR measurement for OGA and OGT's DI but for 6 h of treatment.

To confirm O-GlcNAc regulation of DI splicing for *OGT* and *OGA*, we designed PCR primers to quantify levels of the different transcripts and splice products following treatment with either the OGT inhibitor OSMI-2, or thiamet-G (TMG), a well-validated OGA inhibitor [209]. With OGT inhibition, both *OGT* DIs decreased and levels of the productive (protein coding) mRNA isoform increased (Figure 4.4.1c). These changes corroborate northern blot data obtained previously for *OGT* DI splicing [203]. With OGA inhibition, the spliced, productive mRNA transcript decreased, and the non-productive mRNA transcript containing the poison exon increased. Our results show that the RNA processing pathways to productive versus non-productive *OGT* transcripts oppose each other in an O-GlcNAc-dependent manner, confirming that O-GlcNAc regulates *OGT* DI splicing. We also identify poison exon inclusion as an important mechanism for downregulating productive *OGT* mRNA expression when OGA is inhibited.

We next probed O-GlcNAc regulation of *OGA* splicing, using primers that reported on transcripts containing both DIs, or only DI-1, and on the exon 11-skipped (non-productive) mRNA as well as the protein-coding mRNA containing exon 11. In the presence of OGT inhibitor, we observed a decrease in both DI-containing transcripts, but with a substantially larger decrease for transcripts containing only DI-1 (Figure 4.4.1d). We also observed a large increase in the exon 11-skipped transcript and a statistically significant decrease in productive transcript, although not to an extent commensurate with the increase in the exon-skipped product. However, we note that productive mRNAs comprise ~80% of *OGA* transcripts under steady-state conditions and the mRNA half-life is relatively long (~8.5 h) [142]; both of these factors would dampen the impact of splicing changes on productive mRNA levels at short time

points. Indeed, with longer treatment time (6 h), we observed a substantial decrease in the productive transcript of *OGA* (Figure 4.4.1e). When OGA was inhibited, we again observed large changes in the amount of transcript containing DI-1 and in the exon 11-skipped transcript, but they were in the opposite direction from those observed under OGT inhibition. These results indicate that splicing of *OGA*, similar to *OGT*, is regulated by O-GlcNAc.

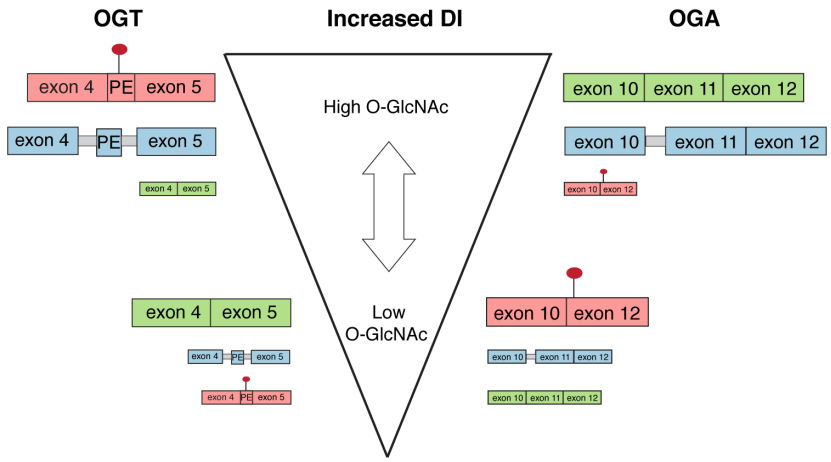


Fig 4.4.2 Proposed model on how O-GlcNAc uses DI to regulate productive transcript. Low O-GlcNAc favors the non-DI pathway to increase functional *OGT* mRNA and non-functional *OGA* mRNA. High O-GlcNAc favors formation of DIs, the splicing of which, increases non-functional *OGT* and functional *OGA*.

From our RT-qPCR analysis of different transcripts under O-GlcNAc perturbation, we propose a working model for how O-GlcNAc affects *OGT* and *OGA* splicing pathways (Figure 4.4.2). For this model, we assume that co-transcriptional splicing produces fully spliced mRNA in competition with the detained intron pathway, and that nascent transcripts are directed to one of the two outcomes in a proportion determined by O-GlcNAc levels. For both *OGT* and *OGA*, high O-GlcNAc results in increased transcript flux into the detained intron pathway. In both cases, intron detention appears to require commitment to the inclusion of the central exon, consistent with the 'decoy exon' hypothesis [206]. For OGT, this results in the observed higher levels of the exon 4-poison exon spliced product in TMG-treated cells. For *OGA*, the detained

intron pathway evidently results in functional product. We infer this based on the observation that when OGA is inhibited, there is a very large increase in a DI transcript that contains DI-1 but not DI-2. Because exons 11 and 12 have been joined, this transcript is a direct precursor to functional mRNA. Further, because the joining of exon 11 and 12 obviates the skipping of exon 11 through the removal of the 3' splice site, formation of this transcript competes with the formation of the non-productive isoform. As a result of their inverted, symmetrical gene structures, high O-GlcNAc levels – through the single mechanism of increasing intron detention – lead to the downregulation of OGT by increasing splicing to the non-productive NMD isoform, while increasing productive OGA mRNA through exon 11 inclusion. The resulting enzymatic activity would be directed toward removal of O-GlcNAc modifications to restore homeostasis.

When cellular O-GlcNAc levels are low, the balance shifts toward increased co-transcriptional splicing and away from intron detention. In both loci, increased co-transcriptional splicing appears to promote reduced inclusion of the central DI-flanked exon. In the case of *OGT*, low O-GlcNAc results in more functional mRNA by promoting the skipping of the poison exon, while in the case of *OGA*, low O-GlcNAc enhances splicing of the frame-shifted NMD substrate (Figure 4.4.2). The rapid, reciprocal, O-GlcNAc-controlled changes we observed in *OGT* and *OGA* splicing pathways provide an explanation for a well-known observation, namely, that OGT and OGA levels change in opposite directions when either protein is blocked [81, 196]. These O-GlcNAc-regulated splicing changes provide an elegant mechanism for how a single signal can be read out to produce opposing gene expression outcomes through the structural configuration of exons and introns, enabling OGT and OGA levels to be precisely tuned to buffer nutrient fluctuations.

4.5 O-GlcNAc is a master regulator of detained intron splicing

Because DI-containing transcripts typically represent a small portion of the RNA pool due to their short half-lives (< 1 h), DI changes can be observed at early time points following a perturbation. Changes in other alternative splicing modes require more time to observe, however, because the stable mRNAs produced must be turned over before a change in isoform composition becomes apparent. We sought to answer whether other changes in splicing occur after a longer period of OGT inhibition. Because the average mRNA half-life is about ~8.5 h, we chose 6 h as our extended time point where we might detect substantive changes not only in DI splicing, but also in other forms of alternative splicing (AS). At 6 h post-treatment with the OGT inhibitor, we observed a global change in DIs, with 3658 out of 4460 DI changes reaching statistical significance (Figure 4.5.1a). Strikingly, the vast majority of these (3651, or 99%) decreased in abundance. We randomly selected for validation six genes with substantial DI decreases during OGT inhibition. RT-qPCR showed that DI abundance in all six genes was substantially reduced in response to OGT inhibition, confirming the global effect on DI splicing (Figure 4.5.1b). The global decrease in DIs observed after 6 h of OGT inhibition suggested that O-GlcNAc affects another layer of DI splicing regulation that is more systemic than the highly specific response observed at 2 h.

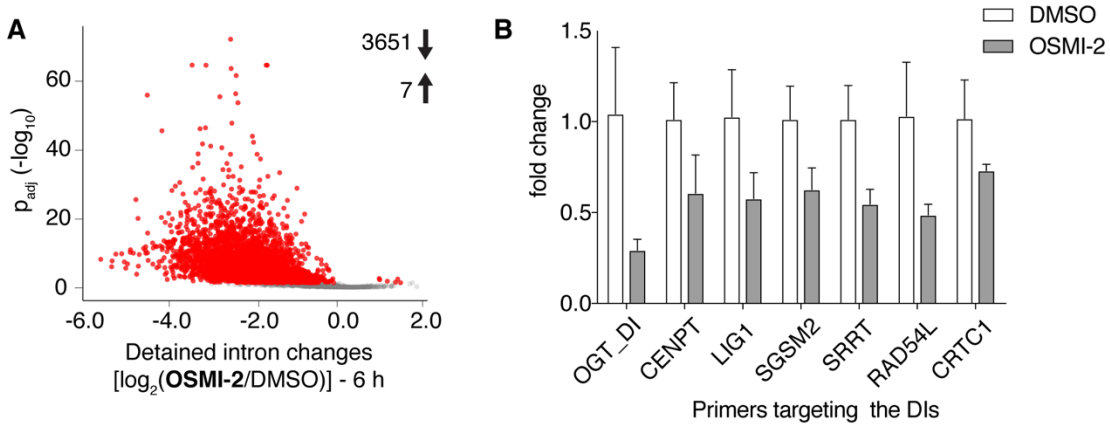


Fig 4.5.1 OGT inhibition globally decreases DI at long time point. (A) Volcano plot of fold changes (log_2) in DIs of polyadenylated transcripts isolated from cells treated with OGT inhibitor for 6 h (10 μM OSMI-2 relative to DMSO) showing a global decrease in DI abundance. **(B)** RT-qPCR results validate DI splicing analysis of RNA-seq data. The panel of genes with DI-containing transcripts were chosen based on those whose DIs displayed significant changes in abundance by RNA-seq. cDNA generated from poly(A)-mRNA isolated from either DMSO control or OSMI-2 treated cells for 6 h were used for the RT-qPCR. Levels for each qPCR product were normalized to actin in a given treatment condition, and fold changes are calculated relative to DMSO control. (n = 3 biological replicates; mean ± s.d)

We asked whether the global decrease in DIs could be due to decreases in total transcript levels for the affected genes. Several findings have suggested that O-GlcNAc levels may regulate the basal transcription machinery, DNA methylation, or nucleosome modification and remodeling complexes [57]. For example, the C-terminus of RNA Pol II can be transiently modified with O-GlcNAc, and some evidence has suggested that these modifications may be important for formation of the pre-initiation complex [210]. It has also been observed that O-GlcNAc-modified proteins are enriched at promoters, and that numerous transcription factors are heavily modified [196, 211]. We analyzed changes at the total gene transcript level using RNA-seq data and found that 635 genes showed at least a 2-fold decrease in transcription, while 350 genes showed at least a 2-fold increase, with >0.95 posterior probability of differential expression being considered significant (Figure 4.5.2). However, among all genes for which total transcript levels decreased, only 34 contained DIs that decreased when OGT was inhibited. We therefore concluded that changes in total transcript quantities cannot account for the effect of reduced O-GlcNAc on DI splicing.

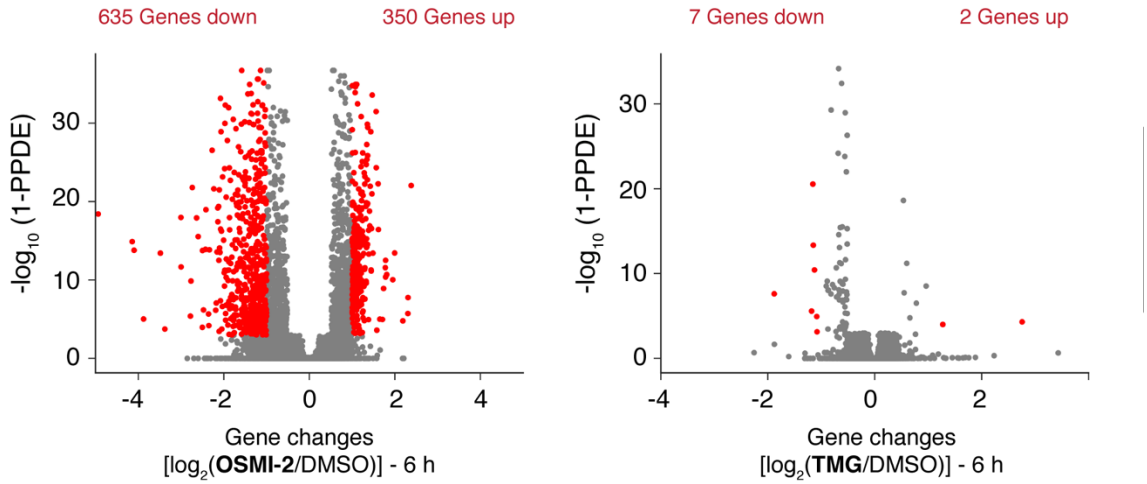


Fig 4.5.2 RNA pol II activity is not directly regulated by O-GlcNAc levels. Volcano plots for gene-level transcript abundances are shown for cells treated with 10 μM OSMI-2 (left) or 5 μM TMG (right). Red dots indicate genes that meet the statistical significance cutoff (PPDE > 0.95) and the fold change cutoff (>2-fold). Of the 635 significantly downregulated genes, only 34 contain a DI.

We reasoned that if the global effect we observed on DI levels is dependent on O-GlcNAc, then OGA inhibition should cause an increase in DI abundance. We performed another round of RNA-seq on poly(A)-selected RNA after TMG treatment. At both 2 h and 6 h, we observed a systematic increase in DI abundance that paralleled the systematic decrease observed upon OGA inhibition (Figure 4.5.3a and b). Fold changes were smaller, and fewer DIs met statistical significance when OGA was inhibited compared to OGT inhibition, possibly because O-GlcNAc levels under steady-state conditions were sufficiently high that the system was resistant to further change. Nevertheless, there was a clear global trend toward increased intron detention as O-GlcNAc levels increased. Taken together, the opposing effects of OGT and OGA inhibition on DI abundance show that cells effect a global DI splicing switch in response to changing O-GlcNAc levels.

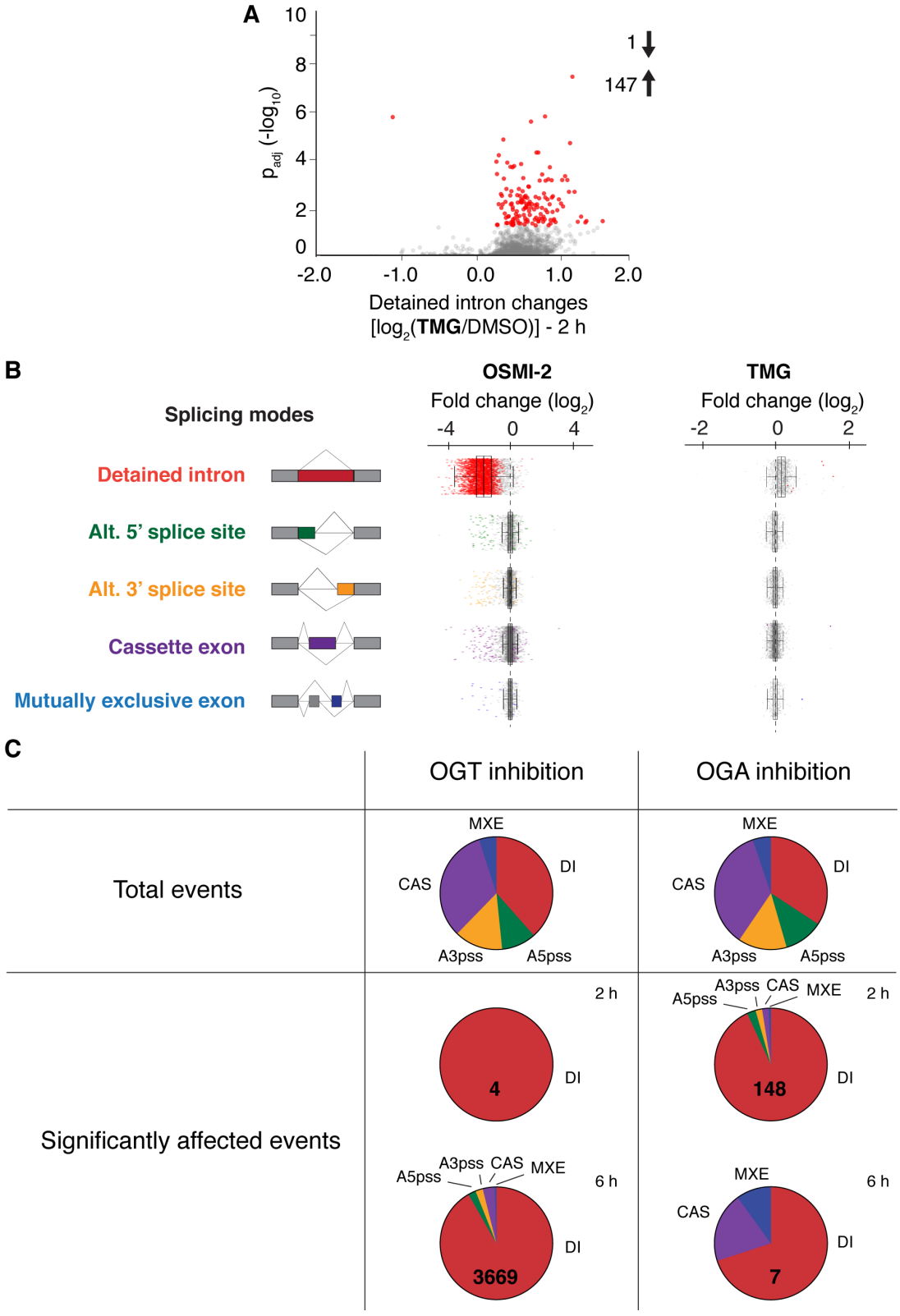


Fig 4.5.3 O-GlcNAc specifically regulates DI levels. (A) A similar volcano plot as in 4.5.1a but generated from cells treated with OGA inhibitor for 2 h (5 uM TMG), showing a global shift toward increased intron detention. **(B)** Changes in alternative splicing events after 6 h treatment with OGT inhibitor (10 uM OSMI-2, left) or OGA

inhibitor (5 uM TMG, right) represented as boxplots. Color coded dots represent splicing events that meet the statistical cutoff. The black dotted line denotes no change compared to the DMSO control. **(C)** Comparison of different alternative splicing events following either OSMI-2 or TMG treatment. Total events refer to all canonical alternative splicing events (including DIs) that are detected. Significantly affected events refer to only those that meet the statistical significance cutoff of FDR < 0.05.

The striking impact of O-GlcNAc on DI splicing pathways prompted us to ask whether there were global changes in other forms of alternative splicing. We examined four canonical forms of alternative splicing after OGT inhibition for 6 h, and compared the median fold changes to that of DIs (Figure 4.5.3b). DIs displayed a median 3-fold decrease [$\log_2$(fold-change) of -1.6], but the median fold-changes for other alternative splicing modes were negligible. Moreover, of all non-DI alternative splicing events, only 5% (330 events) reached statistical significance (Figure 4.5.3c). Similar results were observed after OGA inhibition in that the $\log_2$ median fold change was > 0 for DI splicing, but remained essentially unchanged for other alternative splicing modes. Therefore, the global change in DI splicing observed when O-GlcNAc levels were perturbed was not paralleled by system-wide changes in other forms of alternative splicing. We conclude that O-GlcNAc serves as a global regulator of DI-containing isoform production versus co-transcriptional splicing, thereby linking nutrient levels to mRNA abundance through OGT/OGA activity (Figure 4.5.4).

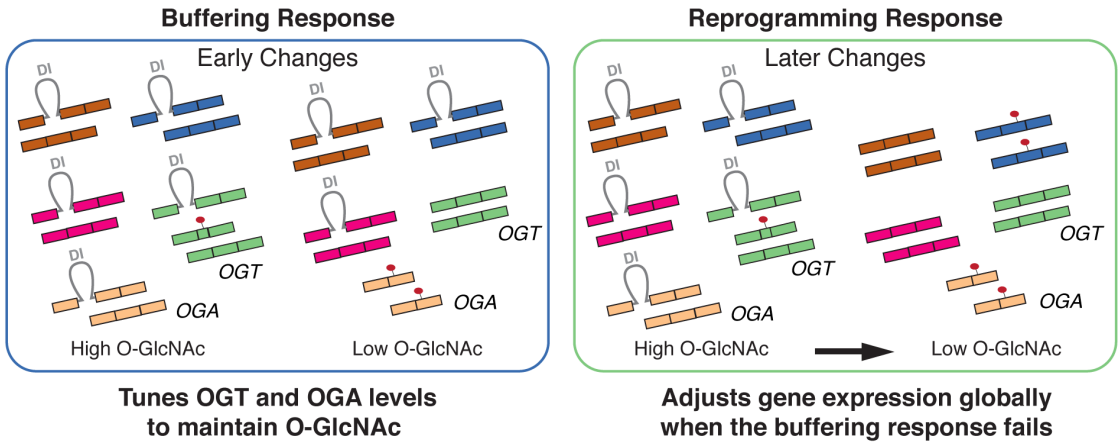


Fig 4.5.3 Models summarizing early and later responses to perturbations in O-GlcNAc levels. Left panel: O-GlcNAc-dependent feedback regulation of DI splicing in *OGT* and *OGA* buffers fluctuations in nutrient availability by tuning levels of productive and non-productive mRNA isoforms to control enzyme abundance. Right panel: Changes in O-GlcNAc levels that exceed the buffering capacity trigger a global DI splicing response to increase (high O-GlcNAc) or decrease (low O-GlcNAc) DI levels to promote an adaptive cell state transition.

4.6 Conclusion

There are two important findings from this study. First, we found that OGT and OGA levels, which are known to change reciprocally when either protein is blocked [81, 196], are regulated by detained intron splicing in an O-GlcNAc-dependent manner. These rapid splicing changes serve to maintain O-GlcNAc homeostasis by tuning levels of the O-GlcNAc cycling enzymes as cellular conditions change. The productive versus non-productive forms of the fully-spliced mRNAs are dictated by alternative exon inclusion/skipping, which enables the switch between DI versus co-transcriptional splicing to be regulated through the same mechanism, even though the alternative exon splicing accomplishes opposite effects on the productivity of *OGT* and *OGA* mRNAs. Second, a reduction in O-GlcNAc that is not rescued by reciprocal, rapid alterations in OGT and OGA causes a global reduction in detained introns, leading to changes in levels of functional mRNA for genes under detained intron splicing control. This global response may serve to reestablish a new setpoint for gene expression under low nutrient conditions. The

temporal difference between these two responses suggests that two distinct splicing processes control them.

The OGT and OGA detained introns contain regions that are extremely conserved across vertebrates [212], which is unusual within introns and suggests a conserved mechanism of regulating mRNA levels through DIs in response to changes in O-GlcNAc. This observation highlights the importance of maintaining O-GlcNAc homeostasis even as conditions change. Detained introns are found in other genes involved in metabolism, including *MAT2A*, required for S-adenosylmethionine (SAM) synthesis. The *MAT2A* detained intron is involved in a feedback loop that adjusts *MAT2A* mRNA levels in response to changes in methionine [213]. Under low SAM conditions, the RNA methyltransferase METTL16, which uses SAM as a substrate, binds to the *MAT2A* RNA and induces splicing of the *MAT2A* intron. Although molecular details of *MAT2A* DI splicing are still lacking, MAT2A provides the best understood example of how an enzyme can regulate its own DI.

The molecular mechanism underlying O-GlcNAc regulation of DI splicing of *OGT* or *OGA* is a complete mystery. It is safe to say that it must involve a factor possessing an O-GlcNAc-dependent activity that influences splicing specifically at the *OGT* and *OGA* DIs. It is likely that the extreme nucleotide sequence conservation found within regions of these DIs is required for the specific regulation of the introns. Indeed, many so-called 'ultraconserved' intronic elements occur in RNA binding proteins that auto-regulate the splicing of their own transcripts [214]. The respective *OGT* and *OGA* splicing configurations are reminiscent of those seen in the genes encoding SR proteins and CLK kinases, which undergo similar splicing and detained intron regulation and contain ultraconserved elements [186-187, 208]. While SR proteins directly bind their pre-mRNA to affect splicing, the CLK kinases are not known to do so, and so

presumably act through their kinase activity to affect the assembly of prespliceosome components, similar to the presumed action of OGT. However, it remains a mystery how this mechanism could accomplish the extreme selectivity observed in the splicing events examined here and how the ultraconserved elements contribute to the regulation of these feedback loops.

The global effect on DI splicing seen at later time points is presumably the result of a different mechanism. One possibility is that a core spliceosome component or peripheral RNA-binding splicing regulator responsible for the DI regulation is O-GlcNAc modified but is inaccessible to rapid turnover through OGA activity. In this case the splicing factor, a number of which are known to be O-GlcNAc modified, may have to undergo protein degradation to turn over the pool of modified protein and thereby affect the splicing of many DIs. This would account for the delay in DI splicing changes between 2h and 6h we observed here. A second possibility is that the modified factor is involved in chromatin modification, nucleosome remodeling, DNA methylation, or the basal transcription machinery, as effectors of all of these pathways are known to be O-GlcNAc modified [54]. Again, there could be a lag in the reconfiguration of these factors necessary to either directly affect splicing, or change the expression of some necessary component. Indeed, differences in O-GlcNAc resistance to removal by OGA have been observed at different loci undergoing chromatin modification [215]. A third intriguing possibility comes from recent findings that many RNA-binding factors are involved in liquid-liquid phase transitions that partition RNA and its processing factors into membraneless subcellular compartments [216-218]. It is not difficult to imagine how sequestration of DI containing or NMD-sensitive isoforms within such a compartment could contribute to changes in the stability or processing of transcripts, and that O-GlcNAc could modulate these phase transitions by altering multivalent interaction dynamics between the proteins and RNAs

involved [199]. The complexity of the O-GlcNAc-modified proteome constitutes the highest barrier to dissecting the exact factors controlling this striking example of splicing-dependent gene regulation [53, 133, 219].

4.7 Methods

Cell lines and cell culture

HEK-293T cells were purchased from American Type Culture Collection (ATCC). HEK-293T cells were grown in Dulbecco's Modified Eagle Medium (Gibco, USA) supplemented with 10% FBS and 1X Penicillin-Streptomycin solution (Corning) at 37ºC in 5% CO_2. OSMI-2 was synthesized as previously described and Thiamet-G (SML0244) was purchased from Sigma-Aldrich [45].

Western blotting

Cells were prepared for western blotting in the following manner, and all the steps were conducted at 4°C. Cells, treated with inhibitors or DMSO in fresh media, (changed 3 h before treatment) were washed once with cold PBS, collected in cold PBS, centrifuged and lysed in RIPA buffer (25 mM Tris, pH 8.0, 1% NP40, 0.5% DOC, 0.1% SDS, 150 mM NaCl) supplemented with cOmplete™ protease inhibitor cocktail (Sigma), PhosSTOP™ (Sigma), and 50 μM thiamet-G (Sigma). After this, samples were loaded on an SDS-PAGE gel and transferred to nitrocellulose membrane for immunoblotting. Antibodies used in this study include: anti-OGT (24083S, CST), anti-O-GlcNAc (RL2, ab2739, Abcam) and anti-actin (ab49900, Abcam).

Immunoprecipitation

Cells cultured in 100 mM plates were treated either with 10 μM OSMI-2 or DMSO for 0.5 h or 1.0 h. Cells were washed and collected as indicated in the Western blotting section. At the indicated time point, cells were lysed in IP buffer (1% NP-40, 50 mM Tris, 2 mM EDTA and 150 mM NaCl) supplemented with cOmplete™ protease inhibitor cocktail, PhosSTOP™, and 50 μM thiamet-G, and homogenized by 10 passes through a 21-guage needle, and protein concentration was determined with the Pierce BCA protein assay kit (ThermoFisher). Insoluble materials were removed by centrifugation at 21,000 g for 30 mins at 4ºC. The supernatant was precleared by incubation with Dynabeads protein G magnetic beads and anti-mouse IgG (sc-2025, Santa Cruz) for 1 h at 4ºC. Protein concentrations of precleared lysates were determined by BCA assay and normalized before adding a mixture of two O-GlcNAc antibodies (Rl2 and CTD110.6). Precleared lysates were then incubated overnight at 4ºC. The next day, Dynabeads protein G magnetic beads were added and samples were incubated for 2 h at room temperature. Beads were washed thrice with IP buffer and eluted with 0.5 M NH_4OH solution. Samples were vacuum centrifuged to dryness and redissolved in 8 M urea, 100 mM EPPS pH 8.5. Samples were reduced with 5 mM TCEP for 15 mins at room temperature in the dark, then alkylated with 14 mM iodoacetamide for 45 mins at room temperature in the dark. Samples were then diluted 10-fold with 100 mM EPPS pH 8.5. Proteins were then digested with 1 μg LysC (Wako) overnight at room temperature and then with 1 μg sequencing grade trypsin (Promega) for 6 hours at 37ºC. The resulting peptide solutions were then labelled with TMT 10/11-plex reagents (ThermoFisher). These fractions were subsequently acidified with 1% formic acid, vacuum centrifuged to near dryness and desalted with C18 stagetips. Dried peptides were resuspended in 5 % acetonitrile/5% formic acid for LC-MS^3 processing.

RNA sequencing

Cells in fresh media (changed 3 h before treatment) were treated either with compounds or DMSO for 2.0 or 6.0 h. Cells were washed and collected as indicated in the western blotting section. At the indicated time point, total RNA from OSMI-2 /DMSO samples (DMSO in triplicates and OSMI-2 in duplicates) and TMG/DMSO samples (DMSO in triplicates and TMG in triplicates) were isolated using Trizol (ThermoFisher) or RNeasy Plus kit (Qiagen) respectively, according to manufacturer's directions. In Trizol isolated total RNA, Turbo DNase (ThermoFisher) was added to remove residual DNA. Agilent Bioanalyzer was used to determine RNA quality and integrity before isolation of poly(A)$^+$ RNA. A total of 2 μg of RNA was used for subsequent steps. The TruSeq Stranded mRNA library preparation kit (Illumina) was used according to manufacturer's instructions for a poly(A)$^+$-based mRNA enrichment. The fragmentated mRNA samples were subjected to cDNA synthesis and library generation using the TruSeq mRNA library preparation kit (Illumina). Sequencing was performed with a NextSeq sequencer (Illumina) for paired end reads of ~75 bp. Control, OSMI-2, and TMG-treated RNA-seq data is available from the Sequence Read Archive (SRA) under the accession number .

Bioinformatics and statistical analyses

RNA-seq analysis were performed as previously described [186, 190]. Briefly, for OSMI-2/DMSO or TMG/DMSO samples, an average of ~42 M (OSMI-2) reads or ~31 M (TMG) reads per sample were generated with an average successful alignment of ~96% (OSMI-2) or ~91% (TMG) using Star Aligner [220]. For gene expression analysis, reads were aligned to hg19 reference genome using RSEM, and EBseq was used to perform subsequent differential expression analysis [221-222]. Detained introns were performed as previously described [186]. Using reads from OSMI-2/DMSO, reads were mapped using Bowtie and filtered with Bedtools. DESeq was then used to normalize

intronic read counts [223-228]. Alternative and constitutive intron classifications were performed using custom Python scripts, and are agnostic with regard to existing annotations other than known gene boundaries. Annotated reads were then used as an input to generate an 'exon part' gtf that was compatible with DEXSeq [229]. The same generated gtf from OSMI-2/DMSO samples is used for TMG/DMSO analysis. Reads were counted from mapped .bam files using the counting script included with DEXSeq to generate count tables for each exon part. Differential expression of the alternative splicing events and detained introns was then determined using standard DEXSeq analysis with a padj. <0.05 as the cutoff for significant changes. One of the OSMI-2 treated samples appeared to have a dosing error because it clustered with the DMSO samples, and was excluded in subsequent analysis.

Phosphopeptide enrichment

Cells were grown in a 150 mm dish in fresh media (changed 3 h before treatment) and were treated either with OSMI-2 or DMSO for 0.5 or 2.0 h. Cells were washed and collected as indicated in the western blotting section. At the indicated time point, cells were lysed in lysis buffer (2 % SDS 50 mM Tris and 150 mM NaCl), sonicated (BioDisruptor) and protein concentration determined with the BCA assay. Samples were reduced with 5 mM DTT for 45 mins at 60ºC, then alkylated with 14 mM iodoacetamide for 45 mins at room temperature in the dark. Then 1 mg of protein was precipitated using chloroform/methanol. Protein pellets were resuspended in 200 mM HEPES pH 8.5 to 1 mg/mL. Proteins were digested with LysC (Wako) (substrate:enzyme=100) overnight at 37ºC and then with sequencing grade trypsin (Promega) (substrate:enzyme= 100) for 6 hours at 37ºC. Equal amounts of peptide samples were desalted on a Waters C18 solid phase extraction Sep-Pak. The resulting peptide solutions were resuspended in 2 M lactic acid 50% ACN to a concentration of 2 mg/ml. Titansphere TiO_2 beads (Gl

Sciences) were washed twice with 2 M lactic acid 50% ACN and added into the peptide mixture to enrich for phosphopeptides. Phosphopeptides were washed twice with 1% TFA in 50% ACN and eluted with elution buffer (50 mM KH_2PO_4 pH 10). Formic acid was then added to a final concentration of 4% to acidify the solution. Phosphopetides were desalted and then labelled with TMT 10/11-plex reagents (ThermoFisher) for subsequent quantitative proteomics. TMT-labeled peptide samples were fractionated via basic-pH reversed-phase (BPRP) HPLC to 96 fractions and then consolidated to 12 fractions. These fractions were subsequently acidified with 1% formic acid, vacuum centrifuged to near dryness and desalted with C18 stagetips [157]. Dried peptides were resuspended in 5 % acetonitrile/5% formic acid for LC-MS^3 processing.

Mass spectrometry and LC-MS/MS measurement

Our mass spectrometry data were collected using an Orbitrap Fusion Lumos mass spectrometer (ThermoFisher Scientific, San Jose, CA) coupled to a Proxeon EASY-nLC 1200 liquid chromatography (LC) pump (ThermoFisher). Peptides were separated on a 100 μm inner diameter microcapillary column packed with 35 cm of Accucore C18 resin (2.6 μm, 150 Å, ThermoFisher). For each analysis, we loaded ~2 μg onto the column.

Separation was in-line with the mass spectrometer and was performed using a 3 h gradient of 6 to 26% acetonitrile in 0.125% formic acid at a flow rate of ~450 nL/min. Each analysis used a MS3-based TMT method, which has been shown to reduce ion interference compared to MS2 quantification [230-231]. The scan sequence began with an MS1 spectrum (Orbitrap analysis; resolution 120,000; mass range 400−1400 *m/z*; automatic gain control (AGC) target 5×10^5; maximum injection time 100 ms). Precursors for MS2/MS3 analysis were selected using a Top10 method. Following acquisition of each MS2 spectrum, we collected an MS3 spectrum using our

recently described method in which multiple MS2 fragment ions were captured in the MS3 precursor population using isolation waveforms with multiple frequency notches.

Mass spectra were processed using a SEQUEST-based pipeline [232]. Spectra were converted to mzXML using a modified version of ReAdW.exe. Database searching included all entries from the human UniProt database. This database was concatenated with one composed of all protein sequences in the reversed order. Searches were performed using a 50 ppm precursor ion tolerance for total protein level analysis. The product ion tolerance was set to 0.9 Da. These wide mass tolerance windows were chosen to maximize sensitivity in conjunction with Sequest searches and linear discriminant analysis [233]. TMT tags on lysine residues and peptide N-termini and carbamidomethylation of cysteine residues were set as static modifications, while oxidation of methionine residues was set as a variable modification. Relative quantitation was performed as previously described [234].

RT-qPCR analysis

Total RNA was isolated employing the RNeasy Plus kit (Qiagen) according to manufacturer's protocol. Extracted RNA was reversed transcribed into cDNA using random hexamer priming and Superscript III (ThermoFisher) according to the protocol supplied by the manufacturer. For qPCR, cDNA templates were amplified, and the C_T values were quantified with PowerUp SYBR Green Master Mix (ThermoFisher), and normalized to actin as an internal control. Experiments were performed using a StepOnePlus Real-Time PCR System (Applied Biosystems). poly(A)-mRNA was isolated using Ambion® Poly(A)Purist™ MAG Kit (ThermoFisher) according to manufacturer's protocol. The list of primers is provided in Supplementary Data.

5 Future work and directions

In this thesis, I explored the biological effects of O-GlcNAc perturbation. Many studies have examined the effects of O-GlcNAc perturbation by knockdown or overexpression of OGT or OGA. Those studies revealed that maintaining proper levels of O-GlcNAc is important for many cellular processes including cellular viability, metabolism, transcription and epigenetics regulation [57]. Because those studies were performed under conditions in which O-GlcNAc was perturbed over a prolonged period (> 48 h), and O-GlcNAc turnover can occur on the timescale of minutes to a few hours for some proteins [77], some of the important functions that can only be observed at short time points may have been overlooked. The lack of tools to acutely and rapidly decrease O-GlcNAc has been the major impediment to fully understanding the biological roles of O-GlcNAc. To address this issue, we sought to accomplish two tasks: 1) Develop a small-molecule inhibitor that can rapidly decrease O-GlcNAc levels in cells, and 2) Perform systems-wide studies with the inhibitors to assess global changes in cellular state.

Through an interdisciplinary approach that combines previous SAR analyses of OSMI-1, molecular docking, crystallization, and biochemical and cellular assays, we discovered the next generation of OGT inhibitors: OSMI-2/3/4/5. All the OSMI compounds inhibit OGT rapidly in cells (< 4 h) and are selective. We demonstrated that our best inhibitor, OSMI-4, is also the best in class inhibitor for OGT in cells and abrogates O-GlcNAc levels at a concentration as low as 5 μM. All the other reported inhibitors of OGT have to be used at a minimum concentration of ~50 μM to obtain a level of decrease comparable with OSMI-4. Furthermore, OSMI-5a broke into the picomolar binding potency with a K_d of ~0.3 pM. To our knowledge, this is the first reported inhibitor that inhibits GT with picomolar potency. However, OSMI-5 is less effective than OSMI-4 in cells probably due to the larger molecular size, which decreases cellular penetration.

We conclude that the inhibitors reported here are adequate to investigate OGT's function in cells, whether at short time points or long time points. We are also confident that the OGT:OSMI structures we have solved will serve as a useful guide for development of inhibitors for other GTs.

One of the major needs in the OGT field is to uncover OGT inhibitors that are compatible with animal studies. We are currently working towards this goal in collaboration with the NIH. Most of the functions attributed to O-GlcNAc have originated from cellular or biochemical studies. However, not all the functions observed in those studies can be extrapolated or translated into animal physiology due to the complex cell-to-cell communication and heterogeneity of cells in animal models, which is not present in tissue cultures or biochemical assays. There have been a number of studies that used tissue-specific knockdown/knockout of OGT to explore the functions of O-GlcNAc *in vivo* but those studies are limited and expensive to perform [82, 235-236]. Additionally, there have been no studies on the influence of decreased O-GlcNAc on whole animal physiology due to the essentiality of O-GlcNAc. We believe that the reported OSMI series can be further optimized for animal work. We have identified a major metabolic liability, namely the thiophene group, which will be readily oxidized into reactive and toxic thiophene S-oxides [237]. To overcome thiophene's metabolic liability, we could either add a nitrogen to the thiophene or replace thiophene with a different bioisostere [238]. Those studies are currently being performed by the NIH and a new graduate student, Sarah Potter, in the lab.

Using OSMI-2/3/4, we evaluated the effects of inhibiting OGT in cells on a short timescale. We performed quantitative mass spectrometry experiments, which included O-GlcNAcmics, phosphoproteomics, and proteomics. With proteomics, we observed an increase in OGT and decrease in OGA levels after 8 h of inhibition. We believe the rapid change in the

amounts of O-GlcNAc cycling enzymes exists to buffer changes in O-GlcNAc levels, highlighting the importance of maintaining proper O-GlcNAc levels in cells. In fact, we already observed a significant increase in OGT levels with just 2 h of inhibition. With longer inhibition (24 h), we observed an increase in a number of key enzymes involved in sterol biosynthesis. Moreover, we noted a decrease in cholesterol level when OGT is inhibited, confirming that O-GlcNAc regulates sterol metabolism. We speculated that O-GlcNAc might also regulate other lipid biosynthesis pathways considering the role of O-GlcNAc as a nutrient sensor. To probe changes in lipid metabolites, we intend to perform lipid metabolomics with OSMI-4. We are also currently investigating this intriguing link between O-GlcNAc and cholesterol biosynthesis

Our O-GlcNAcmics and phosphoproteomics data strongly suggest that one of the major functions of O-GlcNAc is regulation of pre-mRNA splicing. Indeed, through RNA-seq with OGT and OGA inhibitors, we found that O-GlcNAc is intricately connected to DI splicing. First, we observed that after only two hours of perturbation, O-GlcNAc specifically affects transcripts corresponding to the two O-GlcNAc cycling enzymes that install and remove O-GlcNAc. Both genes contain extremely conserved detained introns, and we describe how O-GlcNAc regulation of their splicing provides an elegant feedback loop to change productive mRNA levels when O-GlcNAc is perturbed. Our findings explain previous observations that OGT and OGA protein abundance changes reciprocally when O-GlcNAc is perturbed. We conclude that O-GlcNAc-regulated detained intron splicing for these genes serves to buffer cells against changes in O-GlcNAc levels. To our knowledge, such a high degree of specificity in splicing regulation has not previously been observed. Second, we noted a remarkable response to O-GlcNAc perturbation that occurred by six hours. If the initial buffering response to restore O-GlcNAc levels fails, under OGT inhibition, 80% of detained intron transcripts decrease. Under OGA

inhibition, DI levels show a global increase. Typically, other modes of alternative splicing do not undergo such comprehensive, global, and unidirectional changes in splicing patterns in response to a single stimulus. As far as we know, this study provides the first example of a posttranslational modification that rapidly, globally, and unidirectionally controls detained intron levels.

Despite the importance of O-GlcNAc in regulating detained introns, we still do not know the mechanism underlying this phenomenon. Our recent proteomic-wide study called Proteome Integral Stability Alteration assay (PISA) suggests one plausible mechanism. PISA is a variation of thermal proteome profiling (TPP) [239]. In a typical TPP experiment, both drug-treated and untreated biological samples are incubated at ~8-10 different temperature points, after which the insoluble proteins aggregates (due to heating) are separated by ultracentrifugation. Each different temperature point is then labeled with mass tag like TMT and the abundance of remaining soluble proteins can be determined by quantitative proteomics (Figure 5a) [240-241]. Because compound binding usually stabilizes its binding partner, an increase in Tm (melting temperature) or a right shift in the sigmoidal curve can be observed upon target engagement. PISA makes use of the shift in the sigmoidal curve in TPP by comparing the integral under the curve (see Figure 5b for explanation). Instead of labeling each different temperature points, all the temperature points are pooled into a single sample and labeled with one tag. Proteins that are stabilized by the drug would be expected to exist in greater abundance (due to a right shift in sigmoidal curve). Notably, PISA and TPP do not only measure changes in protein stabilization due to drug binding. Because protein stabilization and destabilization can arise from changes in post-translational modifications, protein compartmentalization, or interaction partners, data generated from PISA needs to be carefully interpreted, especially if PISA/TPP was performed in live cells.

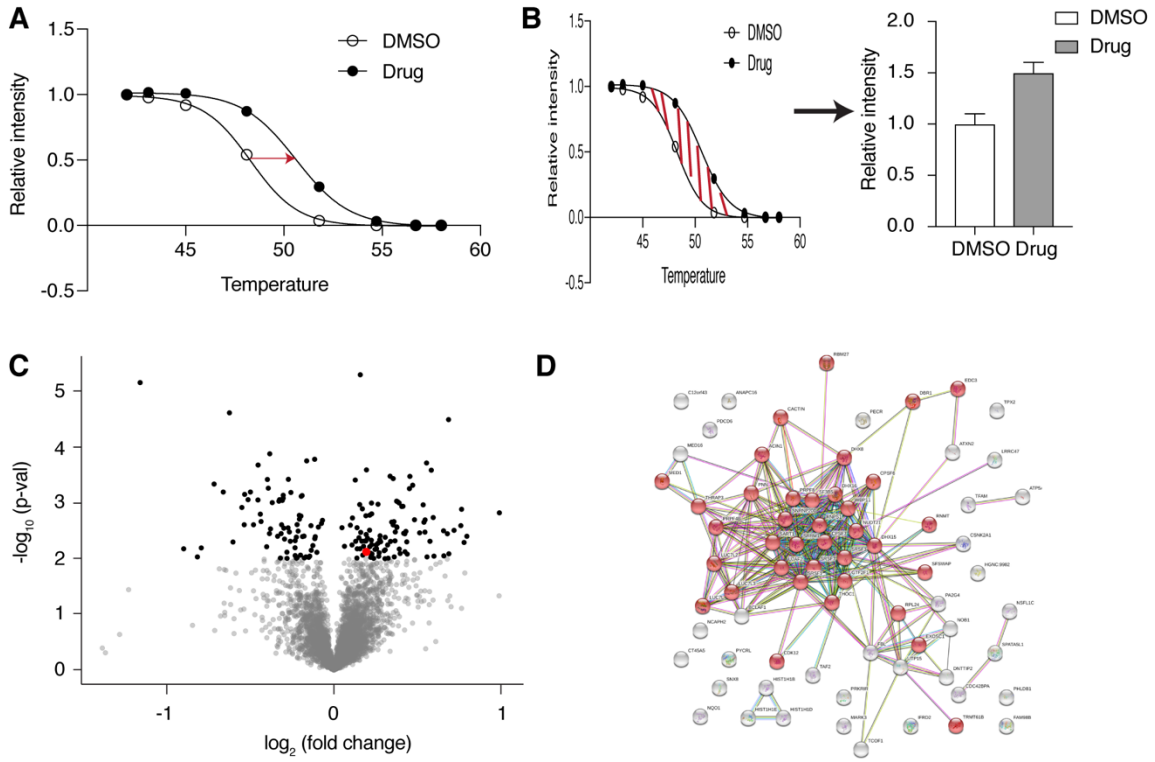


Fig 5 PISA experiments suggest O-GlcNAc affects splicing condensates. (A) Temperature scanning data for drug-treated (black dots) vs DMSO control (white circle) samples. **(B)** PISA measurements showed that drug-treated samples have a larger integral under the curve (red lines) than control sample. A larger integral area indicates greater protein abundance. **(C)** Volcano plot for protein abundance changes as measured by PISA. Log_2fold change refers to changes in protein abundance of OSMI-4-treated vs DMSO samples in 3 biological replicates. Red dot indicates OGT protein. **(D)** Strings-db analysis of proteins that were destabilized by OSMI-4 showed an enrichment of splicing factors (in red).

To further investigate OSMI-4 selectivity, we recently performed PISA using OSMI-4 in K562 cells. K562 cells were chosen because TPP tends to work well in suspension cells and K562 cells have been widely used in TPP experiments. After 1 h of incubation with OSMI-4, we observed stabilization of OGT. Moreover, we found that a majority of the proteins that decreased in abundance are splicing factors (Figure 5c and d). One potential explanation for this

observation is the destabilization of proteins when their O-GlcNAc modifications are removed. While this hypothesis might be true for some splicing factors, most of the observed destabilized splicing factors are not OGT substrates [77, 242]. Another hypothesis is that OGT inhibition causes the splicing factors to undergo a phase change, presumably from a splicing condensate state to a non-condensate state. Supporting this notion, we found that most of the destabilized splicing factors are involved in condensates formation. To address this hypothesis, we have started to collaborate with Eric Guo, who is a world expect in splicing and transcriptional condensates, from Rick Young's lab. This project is currently helmed by another graduate student, George Fei, who will be studying the effects of OGT inhibition on splicing condensates.

It is interesting that O-GlcNAc selectively regulates the splicing of DIs but not other forms of alternative splicing. Because splicing factors are usually promiscuous, one would imagine other alternative splicing pathways to be also affected if O-GlcNAc were to regulate splicing factors to tune DI levels. For O-GlcNAc to only regulate DIs, there must be a DI-specific pathway or compartment that separates DIs from other alternative splicing events. One tempting hypothesis could be that O-GlcNAc causes DI to accumulate in condensates, and depletion of O-GlcNAc causes dissolution of condensates, releasing DIs to be spliced or degraded. It is still unknown whether DIs are concentrated in condensates and it will be interesting to explore this possibility using FISH experiments. It will also be interesting to test if nutrient levels govern DI splicing as well, considering that O-GlcNAc is thought to be a nutrient sensor.

References Cited

1. Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G., Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **2008,** *77*, 521-55.

2. Rini, J. M.; Esko, J. D., Glycosyltransferases and Glycan-Processing Enzymes. In *Essentials of Glycobiology*, rd; Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H., Eds. Cold Spring Harbor (NY), 2015; pp 65-75.

3. Kornberg, S. R.; Zimmerman, S. B.; Kornberg, A., Glucosylation of deoxyribonucleic acid by enzymes from bacteriophage-infected Escherichia coli. *J. Biol. Chem.* **1961,** *236*, 1487-93.

4. Mengin-Lecreulx, D.; Texier, L.; Rousseau, M.; van Heijenoort, J., The murG gene of Escherichia coli codes for the UDP-N-acetylglucosamine: N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. *J. Bacteriol.* **1991,** *173* (15), 4625-36.

5. Wandall, H. H.; Hassan, H.; Mirgorodskaya, E.; Kristensen, A. K.; Roepstorff, P.; Bennett, E. P.; Nielsen, P. A.; Hollingsworth, M. A.; Burchell, J.; Taylor-Papadimitriou, J.; Clausen, H., Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3. *J. Biol. Chem.* **1997,** *272* (38), 23503-14.

6. Chiu, C. P.; Watts, A. G.; Lairson, L. L.; Gilbert, M.; Lim, D.; Wakarchuk, W. W.; Withers, S. G.; Strynadka, N. C., Structural analysis of the sialyltransferase CstII from Campylobacter jejuni in complex with a substrate analog. *Nat. Struct. Mol. Biol.* **2004,** *11* (2), 163-70.

7. Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B., An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* **2003,** *328* (2), 307-17.

8. Breton, C.; Snajdrova, L.; Jeanneau, C.; Koca, J.; Imberty, A., Structures and mechanisms of glycosyltransferases. *Glycobiology* **2006,** *16* (2), 29R-37R.

9. Martins, L. O.; Empadinhas, N.; Marugg, J. D.; Miguel, C.; Ferreira, C.; da Costa, M. S.; Santos, H., Biosynthesis of mannosylglycerate in the thermophilic bacterium Rhodothermus marinus. Biochemical and genetic characterization of a mannosylglycerate synthase. *J. Biol. Chem.* **1999,** *274* (50), 35407-14.

10. Hurtado-Guerrero, R.; Davies, G. J., Recent structural and mechanistic insights into post-translational enzymatic glycosylation. *Curr. Opin. Chem. Biol.* **2012,** *16* (5-6), 479-87.

11. Charnock, S. J.; Davies, G. J., Structure of the nucleotide-diphospho-sugar transferase, SpsA from Bacillus subtilis, in native and nucleotide-complexed forms. *Biochemistry* **1999,** *38* (20), 6380-5.

12. Vrielink, A.; Ruger, W.; Driessen, H. P.; Freemont, P. S., Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J.* **1994,** *13* (15), 3413-22.

13. Liu, J.; Mushegian, A., Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci.* **2003,** *12* (7), 1418-31.

14. Strahl-Bolsinger, S.; Immervoll, T.; Deutzmann, R.; Tanner, W., PMT1, the gene for a key enzyme of protein O-glycosylation in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **1993,** *90* (17), 8164-8.

15. Wiggins, C. A.; Munro, S., Activity of the yeast MNN1 alpha-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95* (14), 7945-50.

16. Hu, Y.; Chen, L.; Ha, S.; Gross, B.; Falcone, B.; Walker, D.; Mokhtarzadeh, M.; Walker, S., Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100* (3), 845-9.

17. Lizak, C.; Gerber, S.; Numao, S.; Aebi, M.; Locher, K. P., X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **2011,** *474* (7351), 350-5.

18. Marth, J. D.; Grewal, P. K., Mammalian glycosylation in immunity. *Nat. Rev. Immunol.* **2008,** *8* (11), 874-87.

19. Berg, S.; Kaur, D.; Jackson, M.; Brennan, P. J., The glycosyltransferases of Mycobacterium tuberculosis - roles in the synthesis of arabinogalactan, lipoarabinomannan, and other glycoconjugates. *Glycobiology* **2007,** *17* (6), 35-56R.

20. Dube, D. H.; Bertozzi, C. R., Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nat Rev Drug Discov* **2005,** *4* (6), 477-88.

21. Brown, J. R.; Crawford, B. E.; Esko, J. D., Glycan antagonists and inhibitors: a fount for drug discovery. *Crit. Rev. Biochem. Mol. Biol.* **2007,** *42* (6), 481-515.

22. Tedaldi, L.; Wagner, G. K., Beyond substrate analogues: new inhibitor chemotypes for glycosyltransferases. *MedChemComm* **2014,** *5* (8), 1106-1125.

23. Cui, H.; Yang, S.; Jiang, Y.; Li, C.; Zhao, Y.; Shi, Y.; Hao, Y.; Qian, F.; Tang, B.; Yu, P., The glycosyltransferase ST6Gal-I is enriched in cancer stem-like cells in colorectal carcinoma and contributes to their chemo-resistance. *Clin. Transl. Oncol.* **2018,** *20* (9), 1175-1184.

24. Britain, C. M.; Dorsett, K. A.; Bellis, S. L., The Glycosyltransferase ST6Gal-I Protects Tumor Cells against Serum Growth Factor Withdrawal by Enhancing Survival Signaling and Proliferative Potential. *J. Biol. Chem.* **2017,** *292* (11), 4663-4673.

25. Lehle, L.; Tanner, W., The specific site of tunicamycin inhibition in the formation of dolichol-bound N-acetylglucosamine derivatives. *FEBS Lett.* **1976,** *72* (1), 167-70.

26. Oslowski, C. M.; Urano, F., Measuring ER stress and the unfolded protein response using mammalian tissue culture system. *Methods Enzymol.* **2011,** *490*, 71-92.

27. Yoo, J.; Mashalidis, E. H.; Kuk, A. C. Y.; Yamamoto, K.; Kaeser, B.; Ichikawa, S.; Lee, S. Y., GlcNAc-1-P-transferase-tunicamycin complex structure reveals basis for inhibition of N-glycosylation. *Nat. Struct. Mol. Biol.* **2018,** *25* (3), 217-224.

28. Takaya, K.; Nagahori, N.; Kurogochi, M.; Furuike, T.; Miura, N.; Monde, K.; Lee, Y. C.; Nishimura, S., Rational design, synthesis, and characterization of novel inhibitors for human beta1,4-galactosyltransferase. *J. Med. Chem.* **2005,** *48* (19), 6054-65.

29. Hang, H. C.; Yu, C.; Ten Hagen, K. G.; Tian, E.; Winans, K. A.; Tabak, L. A.; Bertozzi, C. R., Small molecule inhibitors of mucin-type O-linked glycosylation from a uridine-based library. *Chem. Biol.* **2004,** *11* (3), 337-45.

30. Butters, T. D.; van den Broek, L. A. G. M.; Fleet, G. W. J.; Krulle, T. M.; Wormald, M. R.; Dwek, R. A.; Platt, F. M., Molecular requirements of imino sugars for the selective control of N-

linked glycosylation and glycosphingolipid biosynthesis. *Tetrahedron: Asymmetry* **2000,** *11* (1), 113-124.

31. Mitchell, M. L.; Tian, F.; Lee, L. V.; Wong, C. H., Synthesis and evaluation of transition-state analogue inhibitors of alpha-1,3-fucosyltransferase. *Angew. Chem. Int. Ed. Engl.* **2002,** *41* (16), 3041-4.

32. Compain, P.; Martin, O. R., Carbohydrate mimetics-based glycosyltransferase inhibitors. *Bioorg. Med. Chem.* **2001,** *9* (12), 3077-92.

33. Wang, R.; Steensma, D. H.; Takaoka, Y.; Yun, J. W.; Kajimoto, T.; Wong, C. H., A search for pyrophosphate mimics for the development of substrates and inhibitors of glycosyltransferases. *Bioorg. Med. Chem.* **1997,** *5* (4), 661-72.

34. Compain, P.; Martin, O. R., Design, synthesis and biological evaluation of iminosugar-based glycosyltransferase inhibitors. *Curr. Top. Med. Chem.* **2003,** *3* (5), 541-60.

35. Kellenberger, C.; Hendrickson, T. L.; Imperiali, B., Structural and functional analysis of peptidyl oligosaccharyl transferase inhibitors. *Biochemistry* **1997,** *36* (41), 12554-9.

36. Lee, L. V.; Mitchell, M. L.; Huang, S. J.; Fokin, V. V.; Sharpless, K. B.; Wong, C. H., A potent and highly selective inhibitor of human alpha-1,3-fucosyltransferase via click chemistry. *J. Am. Chem. Soc.* **2003,** *125* (32), 9588-9.

37. Hinou, H.; Sun, X. L.; Ito, Y., Systematic syntheses and inhibitory activities of bisubstrate-type inhibitors of sialyltransferases. *J. Org. Chem.* **2003,** *68* (14), 5602-13.

38. Pesnot, T.; Jørgensen, R.; Palcic, M. M.; Wagner, G. K., Structural and mechanistic basis for a new mode of glycosyltransferase inhibition. *Nat. Chem. Biol.* **2010,** *6* (5), 321-323.

39. Frantom, P. A.; Coward, J. K.; Blanchard, J. S., UDP-(5F)-GlcNAc acts as a slow-binding inhibitor of MshA, a retaining glycosyltransferase. *J. Am. Chem. Soc.* **2010,** *132* (19), 6626-7.

40. Boltje, T. J.; Buskas, T.; Boons, G. J., Opportunities and challenges in synthetic oligosaccharide and glycoconjugate research. *Nat. Chem.* **2009,** *1* (8), 611-22.

41. Ye, X. Y.; Lo, M. C.; Brunner, L.; Walker, D.; Kahne, D.; Walker, S., Better substrates for bacterial transglycosylases. *J. Am. Chem. Soc.* **2001,** *123* (13), 3155-6.

42. Derouaux, A.; Sauvage, E.; Terrak, M., Peptidoglycan glycosyltransferase substrate mimics as templates for the design of new antibacterial drugs. *Front. Immunol.* **2013,** *4*, 78.

43. Xu, Y.; Smith, R.; Vivoli, M.; Ema, M.; Goos, N.; Gehrke, S.; Harmer, N. J.; Wagner, G. K., Covalent inhibitors of LgtC: A blueprint for the discovery of non-substrate-like inhibitors for bacterial glycosyltransferases. *Bioorg. Med. Chem.* **2017,** *25* (12), 3182-3194.

44. Hu, Y.; Helm, J. S.; Chen, L.; Ginsberg, C.; Gross, B.; Kraybill, B.; Tiyanont, K.; Fang, X.; Wu, T.; Walker, S., Identification of selective inhibitors for the glycosyltransferase MurG via high-throughput screening. *Chem. Biol.* **2004,** *11* (5), 703-11.

45. Martin, S. E. S.; Tan, Z. W.; Itkonen, H. M.; Duveau, D. Y.; Paulo, J. A.; Janetzko, J.; Boutz, P. L.; Tork, L.; Moss, F. A.; Thomas, C. J.; Gygi, S. P.; Lazarus, M. B.; Walker, S., Structure-Based Evolution of Low Nanomolar O-GlcNAc Transferase Inhibitors. *J. Am. Chem. Soc.* **2018,** *140* (42), 13542-13545.

46. Cheng, A. C.; Coleman, R. G.; Smyth, K. T.; Cao, Q.; Soulard, P.; Caffrey, D. R.; Salzberg, A. C.; Huang, E. S., Structure-based maximal affinity model predicts small-molecule druggability. *Nat. Biotechnol.* **2007,** *25* (1), 71-5.

47. Halgren, T. A., Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009,** *49* (2), 377-89.

48. Mobley, D. L.; Dill, K. A., Binding of small-molecule ligands to proteins: "what you see" is not always "what you get". *Structure* **2009,** *17* (4), 489-98.

49. Hart, G. W.; Slawson, C.; Ramirez-Correa, G.; Lagerlof, O., Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu. Rev. Biochem.* **2011,** *80*, 825-58.

50. Lazarus, M. B.; Nam, Y.; Jiang, J.; Sliz, P.; Walker, S., Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature* **2011,** *469* (7331), 564-7.

51. Gao, Y.; Wells, L.; Comer, F. I.; Parker, G. J.; Hart, G. W., Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic beta-N-acetylglucosaminidase from human brain. *J. Biol. Chem.* **2001,** *276* (13), 9838-45.

52. Ortiz-Meoz, R. F.; Merbl, Y.; Kirschner, M. W.; Walker, S., Microarray discovery of new OGT substrates: the medulloblastoma oncogene OTX2 is O-GlcNAcylated. *J. Am. Chem. Soc.* **2014,** *136* (13), 4845-8.

53. Levine, Z. G.; Fan, C.; Melicher, M. S.; Orman, M.; Benjamin, T.; Walker, S., O-GlcNAc Transferase Recognizes Protein Substrates Using an Asparagine Ladder in the Tetratricopeptide Repeat (TPR) Superhelix. *J. Am. Chem. Soc.* **2018,** *140* (10), 3510-3513.

54. Woo, C. M.; Lund, P. J.; Huang, A. C.; Davis, M. M.; Bertozzi, C. R.; Pitteri, S. J., Mapping and Quantification of Over 2000 O-linked Glycopeptides in Activated Human T Cells with Isotope-Targeted Glycoproteomics (Isotag). *Mol. Cell. Proteomics* **2018,** *17* (4), 764-775.

55. Bond, M. R.; Hanover, J. A., A little sugar goes a long way: the cell biology of O-GlcNAc. *J. Cell Biol.* **2015,** *208* (7), 869-80.

56. Slawson, C.; Hart, G. W., O-GlcNAc signalling: implications for cancer cell biology. *Nat. Rev. Cancer* **2011,** *11* (9), 678-84.

57. Yang, X.; Qian, K., Protein O-GlcNAcylation: emerging mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* **2017,** *18* (7), 452-465.

58. Liu, B.; Salgado, O. C.; Singh, S.; Hippen, K. L.; Maynard, J. C.; Burlingame, A. L.; Ball, L. E.; Blazar, B. R.; Farrar, M. A.; Hogquist, K. A.; Ruan, H. B., The lineage stability and suppressive program of regulatory T cells require protein O-GlcNAcylation. *Nat Commun* **2019,** *10* (1), 354.

59. Zhang, X.; Shu, X. E.; Qian, S.-B., O-GlcNAc modification of eIF4GI acts as a translational switch in heat shock response. *Nat. Chem. Biol.* **2018,** *14* (10), 909-916.

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

61. Liu, T. W.; Zandberg, W. F.; Gloster, T. M.; Deng, L.; Murray, K. D.; Shan, X.; Vocadlo, D. J., Metabolic Inhibitors of O-GlcNAc Transferase That Act In Vivo Implicate Decreased O-GlcNAc

Levels in Leptin-Mediated Nutrient Sensing. *Angew. Chem. Int. Ed. Engl.* **2018,** *57* (26), 7644-7648.

62. Yi, W.; Clark, P. M.; Mason, D. E.; Keenan, M. C.; Hill, C.; Goddard, W. A., 3rd; Peters, E. C.; Driggers, E. M.; Hsieh-Wilson, L. C., Phosphofructokinase 1 glycosylation regulates cell growth and metabolism. *Science (New York, N.Y.)* **2012,** *337* (6097), 975-980.

63. Zhu, Y.; Shan, X.; Yuzwa, S. A.; Vocadlo, D. J., The emerging link between O-GlcNAc and Alzheimer disease. *J. Biol. Chem.* **2014,** *289* (50), 34472-81.

64. Yuzwa, S. A.; Shan, X.; Macauley, M. S.; Clark, T.; Skorobogatko, Y.; Vosseller, K.; Vocadlo, D. J., Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation. *Nat. Chem. Biol.* **2012,** *8* (4), 393-9.

65. 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., Phosphoinositide signalling links O-GlcNAc transferase to insulin resistance. *Nature* **2008,** *451* (7181), 964-9.

66. Ferrer, C. M.; Lynch, T. P.; Sodi, V. L.; Falcone, J. N.; Schwab, L. P.; Peacock, D. L.; Vocadlo, D. J.; Seagroves, T. N.; Reginato, M. J., O-GlcNAcylation regulates cancer metabolism and survival stress signaling via regulation of the HIF-1 pathway. *Mol. Cell* **2014,** *54* (5), 820-31.

67. Itkonen, H. M.; Minner, S.; Guldvik, I. J.; Sandmann, M. J.; Tsourlakis, M. C.; Berge, V.; Svindland, A.; Schlomm, T.; Mills, I. G., O-GlcNAc transferase integrates metabolic pathways to regulate the stability of c-MYC in human prostate cancer cells. *Cancer Res.* **2013,** *73* (16), 5277-87.

68. Golks, A.; Tran, T. T.; Goetschy, J. F.; Guerini, D., Requirement for O-linked N-acetylglucosaminyltransferase in lymphocytes activation. *EMBO J.* **2007,** *26* (20), 4368-79.

69. Dentin, R.; Hedrick, S.; Xie, J.; Yates, J., 3rd; Montminy, M., Hepatic glucose sensing via the CREB coactivator CRTC2. *Science* **2008,** *319* (5868), 1402-5.

70. Housley, M. P.; Rodgers, J. T.; Udeshi, N. D.; Kelly, T. J.; Shabanowitz, J.; Hunt, D. F.; Puigserver, P.; Hart, G. W., O-GlcNAc regulates FoxO activation in response to glucose. *J. Biol. Chem.* **2008,** *283* (24), 16283-92.

71. Hrit, J.; Goodrich, L.; Li, C.; Wang, B. A.; Nie, J.; Cui, X.; Martin, E. A.; Simental, E.; Fernandez, J.; Liu, M. Y.; Nery, J. R.; Castanon, R.; Kohli, R. M.; Tretyakova, N.; He, C.; Ecker, J. R.; Goll, M.; Panning, B., OGT binds a conserved C-terminal domain of TET1 to regulate TET1 activity and function in development. *Elife* **2018,** *7*.

72. Chu, C. S.; Lo, P. W.; Yeh, Y. H.; Hsu, P. H.; Peng, S. H.; Teng, Y. C.; Kang, M. L.; Wong, C. H.; Juan, L. J., O-GlcNAcylation regulates EZH2 protein stability and function. *Proc. Natl. Acad. Sci. U. S. A.* **2014,** *111* (4), 1355-60.

73. Wang, Z. V.; Deng, Y.; Gao, N.; Pedrozo, Z.; Li, D. L.; Morales, C. R.; Criollo, A.; Luo, X.; Tan, W.; Jiang, N.; Lehrman, M. A.; Rothermel, B. A.; Lee, A. H.; Lavandero, S.; Mammen, P. P. A.; Ferdous, A.; Gillette, T. G.; Scherer, P. E.; Hill, J. A., Spliced X-box binding protein 1 couples the unfolded protein response to hexosamine biosynthetic pathway. *Cell* **2014,** *156* (6), 1179-1192.

74. Andres, L. M.; Blong, I. W.; Evans, A. C.; Rumachik, N. G.; Yamaguchi, T.; Pham, N. D.; Thompson, P.; Kohler, J. J.; Bertozzi, C. R., Chemical Modulation of Protein O-GlcNAcylation via OGT Inhibition Promotes Human Neural Cell Differentiation. *ACS Chem. Biol.* **2017,** *12* (8), 2030-2039.

75. Jang, H.; Kim, T. W.; Yoon, S.; Choi, S. Y.; Kang, T. W.; Kim, S. Y.; Kwon, Y. W.; Cho, E. J.; Youn, H. D., O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. *Cell Stem Cell* **2012,** *11* (1), 62-74.

76. Bartlett, D. W.; Davis, M. E., Insights into the kinetics of siRNA-mediated gene silencing from live-cell and live-animal bioluminescent imaging. *Nucleic Acids Res.* **2006,** *34* (1), 322-333.

77. Qin, W.; Lv, P.; Fan, X.; Quan, B.; Zhu, Y.; Qin, K.; Chen, Y.; Wang, C.; Chen, X., Quantitative time-resolved chemoproteomics reveals that stable O-GlcNAc regulates box C/D snoRNP biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2017,** *114* (33), E6749-E6758.

78. Roquemore, E. P.; Chevrier, M. R.; Cotter, R. J.; Hart, G. W., Dynamic O-GlcNAcylation of the small heat shock protein alpha B-crystallin. *Biochemistry* **1996,** *35* (11), 3578-86.

79. Khidekel, N.; Ficarro, S. B.; Clark, P. M.; Bryan, M. C.; Swaney, D. L.; Rexach, J. E.; Sun, Y. E.; Coon, J. J.; Peters, E. C.; Hsieh-Wilson, L. C., Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics. *Nat. Chem. Biol.* **2007,** *3* (6), 339-48.

80. Wang, X.; Yuan, Z. F.; Fan, J.; Karch, K. R.; Ball, L. E.; Denu, J. M.; Garcia, B. A., A Novel Quantitative Mass Spectrometry Platform for Determining Protein O-GlcNAcylation Dynamics. *Mol. Cell. Proteomics* **2016,** *15* (7), 2462-75.

81. Kazemi, Z.; Chang, H.; Haserodt, S.; McKen, C.; Zachara, N. E., O-linked beta-N-acetylglucosamine (O-GlcNAc) regulates stress-induced heat shock protein expression in a GSK-3beta-dependent manner. *J. Biol. Chem.* **2010,** *285* (50), 39096-107.

82. Lagerlof, O.; Slocomb, J. E.; Hong, I.; Aponte, Y.; Blackshaw, S.; Hart, G. W.; Huganir, R. L., The nutrient sensor OGT in PVN neurons regulates feeding. *Science* **2016,** *351* (6279), 1293-6.

83. Levine, Z. G.; Walker, S., The Biochemistry of O-GlcNAc Transferase: Which Functions Make It Essential in Mammalian Cells? *Annu. Rev. Biochem.* **2016,** *85*, 631-57.

84. Shafi, R.; Iyer, S. P.; Ellies, L. G.; O'Donnell, N.; Marek, K. W.; Chui, D.; Hart, G. W.; Marth, J. D., The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc. Natl. Acad. Sci. U. S. A.* **2000,** *97* (11), 5735-9.

85. Yang, Y. R.; Song, M.; Lee, H.; Jeon, Y.; Choi, E. J.; Jang, H. J.; Moon, H. Y.; Byun, H. Y.; Kim, E. K.; Kim, D. H.; Lee, M. N.; Koh, A.; Ghim, J.; Choi, J. H.; Lee-Kwon, W.; Kim, K. T.; Ryu, S. H.; Suh, P. G., O-GlcNAcase is essential for embryonic development and maintenance of genomic stability. *Aging Cell* **2012,** *11* (3), 439-48.

86. Macauley, M. S.; Vocadlo, D. J., Increasing O-GlcNAc levels: An overview of small-molecule inhibitors of O-GlcNAcase. *Biochim. Biophys. Acta* **2010,** *1800* (2), 107-21.

87. Macauley, M. S.; Whitworth, G. E.; Debowski, A. W.; Chin, D.; Vocadlo, D. J., O-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors. *J. Biol. Chem.* **2005,** *280* (27), 25313-22.

88. Kim, E. J.; Kang, D. O.; Love, D. C.; Hanover, J. A., Enzymatic characterization of O-GlcNAcase isoforms using a fluorogenic GlcNAc substrate. *Carbohydr. Res.* **2006,** *341* (8), 971-82.

89. Knight, Z. A.; Shokat, K. M., Chemical genetics: where genetics and pharmacology meet. *Cell* **2007,** *128* (3), 425-30.

90. Angelova, M.; Ortiz-Meoz, R. F.; Walker, S.; Knipe, D. M., Inhibition of O-Linked N-Acetylglucosamine Transferase Reduces Replication of Herpes Simplex Virus and Human Cytomegalovirus. *J. Virol.* **2015,** *89* (16), 8474-83.

91. Lazarus, M. B.; Jiang, J.; Kapuria, V.; Bhuiyan, T.; Janetzko, J.; Zandberg, W. F.; Vocadlo, D. J.; Herr, W.; Walker, S., HCF-1 is cleaved in the active site of O-GlcNAc transferase. *Science (New York, N.Y.)* **2013,** *342* (6163), 1235-1239.

92. Jinek, M.; Rehwinkel, J.; Lazarus, B. D.; Izaurralde, E.; Hanover, J. A.; Conti, E., The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha. *Nat. Struct. Mol. Biol.* **2004,** *11* (10), 1001-7.

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

94. Dorfmueller, H. C.; Borodkin, V. S.; Blair, D. E.; Pathak, S.; Navratilova, I.; van Aalten, D. M., Substrate and product analogues as human O-GlcNAc transferase inhibitors. *Amino Acids* **2011,** *40* (3), 781-92.

95. Anderson, C. M.; Parkinson, F. E., Potential signalling roles for UTP and UDP: sources, regulation and release of uracil nucleotides. *Trends Pharmacol. Sci.* **1997,** *18* (10), 387-92.

96. Konrad, R. J.; Zhang, F.; Hale, J. E.; Knierman, M. D.; Becker, G. W.; Kudlow, J. E., Alloxan is an inhibitor of the enzyme O-linked N-acetylglucosamine transferase. *Biochem. Biophys. Res. Commun.* **2002,** *293* (1), 207-212.

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

98. Gloster, T. M.; Zandberg, W. F.; Heinonen, J. E.; Shen, D. L.; Deng, L.; Vocadlo, D. J., Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells. *Nat. Chem. Biol.* **2011,** *7* (3), 174-81.

99. Murphy, G.; Ariyanayagam, A. D.; Kuhn, N. J., Progesterone and the metabolic control of the lactose biosynthetic pathway during lactogenesis in the rat. *Biochem. J.* **1973,** *136* (4), 1105-16.

100. Donovan, K.; Alekseev, O.; Qi, X.; Cho, W.; Azizkhan-Clifford, J., O-GlcNAc modification of transcription factor Sp1 mediates hyperglycemia-induced VEGF-A upregulation in retinal cells. *Invest. Ophthalmol. Vis. Sci.* **2014,** *55* (12), 7862-73.

101. Perez-Cervera, Y.; Dehennaut, V.; Aquino Gil, M.; Guedri, K.; Solorzano Mata, C. J.; Olivier-Van Stichelen, S.; Michalski, J. C.; Foulquier, F.; Lefebvre, T., Insulin signaling controls the expression of O-GlcNAc transferase and its interaction with lipid microdomains. *FASEB J.* **2013,** *27* (9), 3478-86.

102. Chen, P. H.; Smith, T. J.; Wu, J.; Siesser, P. F.; Bisnett, B. J.; Khan, F.; Hogue, M.; Soderblom, E.; Tang, F.; Marks, J. R.; Major, M. B.; Swarts, B. M.; Boyce, M.; Chi, J. T., Glycosylation of KEAP1 links nutrient sensing to redox stress signaling. *EMBO J.* **2017,** *36* (15), 2233-2250.

103. Ortiz-Meoz, R. F.; Jiang, J.; Lazarus, M. B.; Orman, M.; Janetzko, J.; Fan, C.; Duveau, D. Y.; Tan, Z. W.; Thomas, C. J.; Walker, S., A small molecule that inhibits OGT activity in cells. *ACS Chem. Biol.* **2015**.

104. Gross, B. J.; Kraybill, B. C.; Walker, S., Discovery of O-GlcNAc transferase inhibitors. *J. Am. Chem. Soc.* **2005,** *127* (42), 14588-9.

105. Jiang, J.; Lazarus, M. B.; Pasquina, L.; Sliz, P.; Walker, S., A neutral diphosphate mimic crosslinks the active site of human O-GlcNAc transferase. *Nat. Chem. Biol.* **2011,** *8* (1), 72-7.

106. Gloster, T. M.; Zandberg, W. F.; Heinonen, J. E.; Shen, D. L.; Deng, L.; Vocadlo, D. J., Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells. *Nat. Chem. Biol.* **2011,** *7* (3), 174-181.

107. Ortiz-Meoz, R. F.; Jiang, J.; Lazarus, M. B.; Orman, M.; Janetzko, J.; Fan, C.; Duveau, D. Y.; Tan, Z. W.; Thomas, C. J.; Walker, S., A Small Molecule that Inhibits OGT Activity in Cells. *ACS Chem. Biol.* **2015,** *10* (6), 1392-1397.

108. Hu, C. W.; Worth, M.; Fan, D.; Li, B.; Li, H.; Lu, L.; Zhong, X.; Lin, Z.; Wei, L.; Ge, Y.; Li, L.; Jiang, J., Electrophilic probes for deciphering substrate recognition by O-GlcNAc transferase. *Nat Chem Biol* **2017,** *13* (12), 1267-1273.

109. Halgren, T. A., Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model* **2009,** *49* (2), 377-389.

110. Meng, X. Y.; Zhang, H. X.; Mezei, M.; Cui, M., Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comput. Aided Drug Des.* **2011,** *7* (2), 146-57.

111. Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J.; Honig, B.; Shaw, D. E.; Friesner, R. A., A hierarchical approach to all-atom protein loop prediction. *Proteins* **2004,** *55* (2), 351-67.

112. Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R., Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006,** *49* (2), 534-53.

113. Sherman, W.; Beard, H. S.; Farid, R., Use of an Induced Fit Receptor Structure in Virtual Screening. *Chem. Biol. Drug Des.* **2006,** *67* (1), 83-84.

114. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006,** *49* (21), 6177-96.

115. Wear, M. A.; Patterson, A.; Malone, K.; Dunsmore, C.; Turner, N. J.; Walkinshaw, M. D., A surface plasmon resonance-based assay for small molecule inhibitors of human cyclophilin A. *Anal. Biochem.* **2005,** *345* (2), 214-26.

116. Leavitt, S.; Freire, E., Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr. Opin. Struct. Biol.* **2001,** *11* (5), 560-6.

117. Concepcion, J.; Witte, K.; Wartchow, C.; Choo, S.; Yao, D.; Persson, H.; Wei, J.; Li, P.; Heidecker, B.; Ma, W.; Varma, R.; Zhao, L.-S.; Perillat, D.; Carricato, G.; Recknor, M.; Du, K.; Ho, H.; Ellis, T.; Gamez, J.; Howes, M.; Phi-Wilson, J.; Lockard, S.; Zuk, R.; Tan, H., Label-Free Detection of Biomolecular Interactions Using BioLayer Interferometry for Kinetic Characterization. *Comb. Chem. High Throughput Screen.* **2009,** *12* (8), 791-800.

118. Seidel, S. A.; Dijkman, P. M.; Lea, W. A.; van den Bogaart, G.; Jerabek-Willemsen, M.; Lazic, A.; Joseph, J. S.; Srinivasan, P.; Baaske, P.; Simeonov, A.; Katritch, I.; Melo, F. A.; Ladbury, J. E.; Schreiber, G.; Watts, A.; Braun, D.; Duhr, S., Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods* **2013,** *59* (3), 301-15.

119. 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., O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis. *Nat. Chem. Biol.* **2012,** *8* (12), 969-74.

120. Jerabek-Willemsen, M.; Wienken, C. J.; Braun, D.; Baaske, P.; Duhr, S., Molecular interaction studies using microscale thermophoresis. *Assay and drug development technologies* **2011,** *9* (4), 342-353.

121. Lazarus, M. B.; Jiang, J.; Gloster, T. M.; Zandberg, W. F.; Whitworth, G. E.; Vocadlo, D. J.; Walker, S., Structural snapshots of the reaction coordinate for O-GlcNAc transferase. *Nat. Chem. Biol.* **2012,** *8* (12), 966-968.

122. Lazarus, M. B.; Nam, Y.; Jiang, J.; Sliz, P.; Walker, S., Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature* **2011,** *469* (7331), 564-567.

123. Schimpl, M.; Zheng, X.; Borodkin, V. S.; Blair, D. E.; Ferenbach, A. T.; Schüttelkopf, A. W.; Navratilova, I.; Aristotelous, T.; Albarbarawi, O.; Robinson, D. A.; Macnaughtan, M. A.; van Aalten, D. M. F., O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis. *Nat. Chem. Biol.* **2012,** *8* (12), 969-974.

124. Schrödinger Release 2018-1: Maestro, Schrödinger, LLC, New York, NY, 2018

125. Schrödinger Release 2018-1: MacroModel, Schrödinger, LLC, New York, NY, 2018

126. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A., OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016,** *12* (1), 281-296.

127. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988,** *37* (2), 785-789.

128. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G., Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B.* **2009,** *113* (18), 6378-6396.

129. Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato,

A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.

130. Wienken, C. J.; Baaske, P.; Rothbauer, U.; Braun, D.; Duhr, S., Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* **2010,** *1*, 100.

131. Liu, C.; Li, J., O-GlcNAc: A Sweetheart of the Cell Cycle and DNA Damage Response. *Front. Endocrinol. (Lausanne)* **2018,** *9*, 415.

132. Tashima, Y.; Stanley, P., Antibodies that detect O-linked beta-D-N-acetylglucosamine on the extracellular domain of cell surface glycoproteins. *J. Biol. Chem.* **2014,** *289* (16), 11132-42.

133. Ma, J.; Hart, G. W., O-GlcNAc profiling: from proteins to proteomes. *Clin. Proteomics* **2014,** *11* (1), 8.

134. Liu, Y.; Ren, Y.; Cao, Y.; Huang, H.; Wu, Q.; Li, W.; Wu, S.; Zhang, J., Discovery of a Low Toxicity O-GlcNAc Transferase (OGT) Inhibitor by Structure-based Virtual Screening of Natural Products. *Sci. Rep.* **2017,** *7* (1), 12334.

135. Isono, T., O-GlcNAc-specific antibody CTD110.6 cross-reacts with N-GlcNAc2-modified proteins induced under glucose deprivation. *PLoS One* **2011,** *6* (4), e18959.

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

137. Comer, F. I.; Vosseller, K.; Wells, L.; Accavitti, M. A.; Hart, G. W., Characterization of a mouse monoclonal antibody specific for O-linked N-acetylglucosamine. *Anal. Biochem.* **2001,** *293* (2), 169-77.

138. Turner, J. R.; Tartakoff, A. M.; Greenspan, N. S., Cytologic assessment of nuclear and cytoplasmic O-linked N-acetylglucosamine distribution by using anti-streptococcal monoclonal antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **1990,** *87* (15), 5608-12.

139. Teo, C. F.; Ingale, S.; Wolfert, M. A.; Elsayed, G. A.; Not, L. G.; Chatham, J. C.; Wells, L.; Boons, G. J., Glycopeptide-specific monoclonal antibodies suggest new roles for O-GlcNAc. *Nat. Chem. Biol.* **2010,** *6* (5), 338-43.

140. Boyce, M.; Carrico, I. S.; Ganguli, A. S.; Yu, S. H.; Hangauer, M. J.; Hubbard, S. C.; Kohler, J. J.; Bertozzi, C. R., Metabolic cross-talk allows labeling of O-linked beta-N-acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108* (8), 3141-6.

141. Clark, P. M.; Dweck, J. F.; Mason, D. E.; Hart, C. R.; Buck, S. B.; Peters, E. C.; Agnew, B. J.; Hsieh-Wilson, L. C., Direct in-gel fluorescence detection and cellular imaging of O-GlcNAc-modified proteins. *J. Am. Chem. Soc.* **2008,** *130* (35), 11576-7.

142. Schwanhausser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M., Global quantification of mammalian gene expression control. *Nature* **2011,** *473* (7347), 337-42.

143. Kim, Y.; Choi, Y.; Weissleder, R.; Tung, C.-H., Membrane permeable esterase-activated fluorescent imaging probe. *Bioorg. Med. Chem. Lett.* **2007,** *17* (18), 5054-5057.

144. Li, B.; Kohler, J. J., Glycosylation of the nuclear pore. *Traffic* **2014,** *15* (4), 347-61.

145. Lubas, W. A.; Smith, M.; Starr, C. M.; Hanover, J. A., Analysis of nuclear pore protein p62 glycosylation. *Biochemistry* **1995,** *34* (5), 1686-94.

146. Davis, L. I.; Blobel, G., Nuclear pore complex contains a family of glycoproteins that includes p62: glycosylation through a previously unidentified cellular pathway. *Proc. Natl. Acad. Sci. U. S. A.* **1987,** *84* (21), 7552-6.

147. Holt, G. D.; Snow, C. M.; Senior, A.; Haltiwanger, R. S.; Gerace, L.; Hart, G. W., Nuclear pore complex glycoproteins contain cytoplasmically disposed O-linked N-acetylglucosamine. *J. Cell Biol.* **1987,** *104* (5), 1157-64.

148. Lazarus, M. B.; Jiang, J.; Kapuria, V.; Bhuiyan, T.; Janetzko, J.; Zandberg, W. F.; Vocadlo, D. J.; Herr, W.; Walker, S., HCF-1 is cleaved in the active site of O-GlcNAc transferase. *Science* **2013,** *342* (6163), 1235-9.

149. Capotosti, F.; Guernier, S.; Lammers, F.; Waridel, P.; Cai, Y.; Jin, J.; Conaway, J. W.; Conaway, R. C.; Herr, W., O-GlcNAc transferase catalyzes site-specific proteolysis of HCF-1. *Cell* **2011,** *144* (3), 376-88.

150. Watanabe, M.; Muramatsu, T.; Shirane, H.; Ugai, K., Discrete Distribution of binding sites for Dolichos biflorus agglutinin (DBA) and for peanut agglutinin (PNA) in mouse organ tissues. *J. Histochem. Cytochem.* **1981,** *29* (7), 779-80.

151. Shultz, S. Determining the Predictive Mechanism of Toxicity Using a Single-Well Multiplexed Assay. http://www.promega.com/resources/pubhub/determining-the-predictive-mechanism-of-toxicity-with-apotox-glo/.

152. Moravec, T. R. a. R. Introducing the CytoTox-ONE™ Homogeneous Membrane Integrity Assay. https://www.promega.com/resources/pubhub/cellnotes/introducing-the-cytotox-one-homogeneous-membrane-integrity-assay/.

153. Mark Bratz, B. H. a. T. S. Measuring cAMP Levels and Cytotoxicity in a Single Plate Well. https://www.promega.com/resources/pubhub/measuring-camp-levels-and-cytotoxicity-in-a-single-plate-well-article/.

154. Itkonen, H. M.; Minner, S.; Guldvik, I. J.; Sandmann, M. J.; Tsourlakis, M. C.; Berge, V.; Svindland, A.; Schlomm, T.; Mills, I. G., O-GlcNAc Transferase Integrates Metabolic Pathways to Regulate the Stability of c-MYC in Human Prostate Cancer Cells. *Cancer Res.* **2013,** *73* (16), 5277-5287.

155. Yang, W. H.; Kim, J. E.; Nam, H. W.; Ju, J. W.; Kim, H. S.; Kim, Y. S.; Cho, J. W., Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. *Nat. Cell Biol.* **2006,** *8* (10), 1074-83.

156. Zhu, Y.; Liu, T. W.; Cecioni, S.; Eskandari, R.; Zandberg, W. F.; Vocadlo, D. J., O-GlcNAc occurs cotranslationally to stabilize nascent polypeptide chains. *Nat. Chem. Biol.* **2015,** *11* (5), 319-25.

157. Navarrete-Perea, J.; Yu, Q.; Gygi, S. P.; Paulo, J. A., Streamlined Tandem Mass Tag (SL-TMT) Protocol: An Efficient Strategy for Quantitative (Phospho)proteome Profiling Using Tandem Mass Tag-Synchronous Precursor Selection-MS3. *J. Proteome Res.* **2018,** *17* (6), 2226-2236.

158. Lickert, H.; Takeuchi, J. K.; Von Both, I.; Walls, J. R.; McAuliffe, F.; Adamson, S. L.; Henkelman, R. M.; Wrana, J. L.; Rossant, J.; Bruneau, B. G., Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. *Nature* **2004,** *432* (7013), 107-12.

159. Zhang, P.; Li, L.; Bao, Z.; Huang, F., Role of BAF60a/BAF60c in chromatin remodeling and hepatic lipid metabolism. *Nutr. Metab. (Lond.)* **2016,** *13*, 30.

160. Dehennaut, V.; Leprince, D.; Lefebvre, T., O-GlcNAcylation, an Epigenetic Mark. Focus on the Histone Code, TET Family Proteins, and Polycomb Group Proteins. *Front. Endocrinol. (Lausanne)* **2014,** *5*, 155.

161. Sodi, V. L.; Bacigalupa, Z. A.; Ferrer, C. M.; Lee, J. V.; Gocal, W. A.; Mukhopadhyay, D.; Wellen, K. E.; Ivan, M.; Reginato, M. J., Nutrient sensor O-GlcNAc transferase controls cancer lipid metabolism via SREBP-1 regulation. *Oncogene* **2018,** *37* (7), 924-934.

162. Song, B. L.; Javitt, N. B.; DeBose-Boyd, R. A., Insig-mediated degradation of HMG CoA reductase stimulated by lanosterol, an intermediate in the synthesis of cholesterol. *Cell Metab.* **2005,** *1* (3), 179-89.

163. Yang, C.; McDonald, J. G.; Patel, A.; Zhang, Y.; Umetani, M.; Xu, F.; Westover, E. J.; Covey, D. F.; Mangelsdorf, D. J.; Cohen, J. C.; Hobbs, H. H., Sterol intermediates from cholesterol biosynthetic pathway as liver X receptor ligands. *J. Biol. Chem.* **2006,** *281* (38), 27816-26.

164. Gill, S.; Stevenson, J.; Kristiana, I.; Brown, A. J., Cholesterol-dependent degradation of squalene monooxygenase, a control point in cholesterol synthesis beyond HMG-CoA reductase. *Cell Metab.* **2011,** *13* (3), 260-73.

165. Slawson, C.; Zachara, N. E.; Vosseller, K.; Cheung, W. D.; Lane, M. D.; Hart, G. W., Perturbations in O-linked beta-N-acetylglucosamine protein modification cause severe defects in mitotic progression and cytokinesis. *J. Biol. Chem.* **2005,** *280* (38), 32944-56.

166. Caldwell, S. A.; Jackson, S. R.; Shahriari, K. S.; Lynch, T. P.; Sethi, G.; Walker, S.; Vosseller, K.; Reginato, M. J., Nutrient sensor O-GlcNAc transferase regulates breast cancer tumorigenesis

through targeting of the oncogenic transcription factor FoxM1. *Oncogene* **2010,** *29* (19), 2831-42.

167. Navarrete-Perea, J.; Yu, Q.; Gygi, S. P.; Paulo, J. A., Streamlined Tandem Mass Tag (SL-TMT) Protocol: An Efficient Strategy for Quantitative (Phospho)proteome Profiling Using Tandem Mass Tag-Synchronous Precursor Selection-MS3. *J. Proteome Res.* **2018,** *17* (6), 2226-2236.

168. Ting, L.; Rad, R.; Gygi, S. P.; Haas, W., MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat. Methods* **2011,** *8*, 937.

169. McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wühr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P., MultiNotch MS3 Enables Accurate, Sensitive, and Multiplexed Detection of Differential Expression across Cancer Cell Line Proteomes. *Anal. Chem.* **2014,** *86* (14), 7150-7158.

170. Paulo, J. A.; O'Connell, J. D.; Gygi, S. P., A Triple Knockout (TKO) Proteomics Standard for Diagnosing Ion Interference in Isobaric Labeling Experiments. *J. Am. Soc. Mass Spectrom.* **2016,** *27* (10), 1620-1625.

171. Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villén, J.; Haas, W.; Sowa, M. E.; Gygi, S. P., A Tissue-Specific Atlas of Mouse Protein Phosphorylation and Expression. *Cell* **2010,** *143* (7), 1174-1189.

172. Beausoleil, S. A.; Villén, J.; Gerber, S. A.; Rush, J.; Gygi, S. P., A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* **2006,** *24*, 1285.

173. Elias, J. E.; Gygi, S. P., Target-Decoy Search Strategy for Mass Spectrometry-Based Proteomics. In *Proteome Bioinformatics*, Hubbard, S. J.; Jones, A. R., Eds. Humana Press: Totowa, NJ, 2010; pp 55-71.

174. Elias, J. E.; Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **2007,** *4*, 207.

175. McAlister, G. C.; Huttlin, E. L.; Haas, W.; Ting, L.; Jedrychowski, M. P.; Rogers, J. C.; Kuhn, K.; Pike, I.; Grothe, R. A.; Blethrow, J. D.; Gygi, S. P., Increasing the Multiplexing Capacity of

TMTs Using Reporter Ion Isotopologues with Isobaric Masses. *Anal. Chem.* **2012,** *84* (17), 7469-7478.

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

177. Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C., Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101* (36), 13132-7.

178. Zaro, B. W.; Yang, Y. Y.; Hang, H. C.; Pratt, M. R., Chemical reporters for fluorescent detection and identification of O-GlcNAc-modified proteins reveal glycosylation of the ubiquitin ligase NEDD4-1. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108* (20), 8146-51.

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

180. Joiner, C. M.; Li, H.; Jiang, J.; Walker, S., Structural characterization of the O-GlcNAc cycling enzymes: insights into substrate recognition and catalytic mechanisms. *Curr. Opin. Struct. Biol.* **2019,** *56*, 97-106.

181. Liu, K.; Paterson, A. J.; Chin, E.; Kudlow, J. E., Glucose stimulates protein modification by O-linked GlcNAc in pancreatic beta cells: linkage of O-linked GlcNAc to beta cell death. *Proc. Natl. Acad. Sci. U. S. A.* **2000,** *97* (6), 2820-5.

182. Hardiville, S.; Hart, G. W., Nutrient regulation of signaling, transcription, and cell physiology by O-GlcNAcylation. *Cell Metab.* **2014,** *20* (2), 208-13.

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

184. Leney, A. C.; El Atmioui, D.; Wu, W.; Ovaa, H.; Heck, A. J. R., Elucidating crosstalk mechanisms between phosphorylation and O-GlcNAcylation. *Proc. Natl. Acad. Sci. U. S. A.* **2017,** *114* (35), E7255-E7261.

185. Zhong, J.; Martinez, M.; Sengupta, S.; Lee, A.; Wu, X.; Chaerkady, R.; Chatterjee, A.; O'Meally, R. N.; Cole, R. N.; Pandey, A.; Zachara, N. E., Quantitative phosphoproteomics reveals crosstalk between phosphorylation and O-GlcNAc in the DNA damage response pathway. *Proteomics* **2015,** *15* (2-3), 591-607.

186. Boutz, P. L.; Bhutkar, A.; Sharp, P. A., Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev.* **2015,** *29* (1), 63-80.

187. Ninomiya, K.; Kataoka, N.; Hagiwara, M., Stress-responsive maturation of Clk1/4 pre-mRNAs promotes phosphorylation of SR splicing factor. *J. Cell Biol.* **2011,** *195* (1), 27-40.

188. Mauger, O.; Lemoine, F.; Scheiffele, P., Targeted Intron Retention and Excision for Rapid Gene Regulation in Response to Neuronal Activity. *Neuron* **2016,** *92* (6), 1266-1278.

189. Gill, J.; Park, Y.; McGinnis, J. P.; Perez-Sanchez, C.; Blanchette, M.; Si, K., Regulated Intron Removal Integrates Motivational State and Experience. *Cell* **2017,** *169* (5), 836-848 e15.

190. Braun, C. J.; Stanciu, M.; Boutz, P. L.; Patterson, J. C.; Calligaris, D.; Higuchi, F.; Neupane, R.; Fenoglio, S.; Cahill, D. P.; Wakimoto, H.; Agar, N. Y. R.; Yaffe, M. B.; Sharp, P. A.; Hemann, M. T.; Lees, J. A., Coordinated Splicing of Regulatory Detained Introns within Oncogenic Transcripts Creates an Exploitable Vulnerability in Malignant Glioma. *Cancer Cell* **2017,** *32* (4), 411-426 e11.

191. Braunschweig, U.; Barbosa-Morais, N. L.; Pan, Q.; Nachman, E. N.; Alipanahi, B.; Gonatopoulos-Pournatzis, T.; Frey, B.; Irimia, M.; Blencowe, B. J., Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res.* **2014,** *24* (11), 1774-86.

192. Bresson, S. M.; Hunter, O. V.; Hunter, A. C.; Conrad, N. K., Canonical Poly(A) Polymerase Activity Promotes the Decay of a Wide Variety of Mammalian Nuclear RNAs. *PLoS Genet* **2015,** *11* (10), e1005610.

193. Pirnie, S. P.; Osman, A.; Zhu, Y.; Carmichael, G. G., An Ultraconserved Element (UCE) controls homeostatic splicing of ARGLU1 mRNA. *Nucleic Acids Res.* **2017,** *45* (6), 3473-3486.

194. Jacob, A. G.; Smith, C. W. J., Intron retention as a component of regulated gene expression programs. *Hum. Genet.* **2017,** *136* (9), 1043-1057.

195. Bond, M. R.; Hanover, J. A., O-GlcNAc cycling: a link between metabolism and chronic disease. *Annu. Rev. Nutr.* **2013,** *33*, 205-29.

196. Itkonen, H. M.; Urbanucci, A.; Martin, S. E.; Khan, A.; Mathelier, A.; Thiede, B.; Walker, S.; Mills, I. G., High OGT activity is essential for MYC-driven proliferation of prostate cancer cells. *Theranostics* **2019,** *9* (8), 2183-2197.

197. Whelan, S. A.; Dias, W. B.; Thiruneelakantapillai, L.; Lane, M. D.; Hart, G. W., Regulation of insulin receptor substrate 1 (IRS-1)/AKT kinase-mediated insulin signaling by O-Linked beta-N-acetylglucosamine in 3T3-L1 adipocytes. *J. Biol. Chem.* **2010,** *285* (8), 5204-11.

198. Zeidan, Q.; Wang, Z.; De Maio, A.; Hart, G. W., O-GlcNAc cycling enzymes associate with the translational machinery and modify core ribosomal proteins. *Mol. Biol. Cell* **2010,** *21* (12), 1922-36.

199. Ohn, T.; Kedersha, N.; Hickman, T.; Tisdale, S.; Anderson, P., A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. *Nat. Cell Biol.* **2008,** *10* (10), 1224-31.

200. Li, X.; Zhu, Q.; Shi, X.; Cheng, Y.; Li, X.; Xu, H.; Duan, X.; Hsieh-Wilson, L. C.; Chu, J.; Pelletier, J.; Ni, M.; Zheng, Z.; Li, S.; Yi, W., O-GlcNAcylation of core components of the translation initiation machinery regulates protein synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2019,** *116* (16), 7857-7866.

201. Hardy, W. R.; Sandri-Goldin, R. M., Herpes simplex virus inhibits host cell splicing, and regulatory protein ICP27 is required for this effect. *J. Virol.* **1994,** *68* (12), 7790-9.

202. Tang, S.; Patel, A.; Krause, P. R., Hidden regulation of herpes simplex virus 1 pre-mRNA splicing and polyadenylation by virally encoded immediate early gene ICP27. *PLoS Pathog.* **2019,** *15* (6), e1007884.

203. Park, S. K.; Zhou, X.; Pendleton, K. E.; Hunter, O. V.; Kohler, J. J.; O'Donnell, K. A.; Conrad, N. K., A Conserved Splicing Silencer Dynamically Regulates O-GlcNAc Transferase Intron Retention and O-GlcNAc Homeostasis. *Cell Rep* **2017,** *20* (5), 1088-1099.

204. Tuorto, F.; Legrand, C.; Cirzi, C.; Federico, G.; Liebers, R.; Muller, M.; Ehrenhofer-Murray, A. E.; Dittmar, G.; Grone, H. J.; Lyko, F., Queuosine-modified tRNAs confer nutritional control of protein translation. *EMBO J.* **2018,** *37* (18).

205. Shalgi, R.; Hurt, J. A.; Lindquist, S.; Burge, C. B., Widespread inhibition of posttranscriptional splicing shapes the cellular transcriptome following heat shock. *Cell Rep* **2014,** *7* (5), 1362-1370.

206. Parra, M.; Booth, B. W.; Weiszmann, R.; Yee, B.; Yeo, G. W.; Brown, J. B.; Celniker, S. E.; Conboy, J. G., An important class of intron retention events in human erythroblasts is regulated by cryptic exons proposed to function as splicing decoys. *RNA* **2018,** *24* (9), 1255-1265.

207. Baker, K. E.; Parker, R., Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr. Opin. Cell Biol.* **2004,** *16* (3), 293-9.

208. Lareau, L. F.; Inada, M.; Green, R. E.; Wengrod, J. C.; Brenner, S. E., Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* **2007,** *446* (7138), 926-9.

209. Yuzwa, S. A.; Macauley, M. S.; Heinonen, J. E.; Shan, X.; Dennis, R. J.; He, Y.; Whitworth, G. E.; Stubbs, K. A.; McEachern, E. J.; Davies, G. J.; Vocadlo, D. J., A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo. *Nat. Chem. Biol.* **2008,** *4* (8), 483-90.

210. Ranuncolo, S. M.; Ghosh, S.; Hanover, J. A.; Hart, G. W.; Lewis, B. A., Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription in vitro and in vivo. *J. Biol. Chem.* **2012,** *287* (28), 23549-61.

211. 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., TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J.* **2013,** *32* (5), 645-55.

212. Bejerano, G.; Pheasant, M.; Makunin, I.; Stephen, S.; Kent, W. J.; Mattick, J. S.; Haussler, D., Ultraconserved elements in the human genome. *Science* **2004,** *304* (5675), 1321-5.

213. Pendleton, K. E.; Chen, B.; Liu, K.; Hunter, O. V.; Xie, Y.; Tu, B. P.; Conrad, N. K., The U6 snRNA m(6)A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. *Cell* **2017,** *169* (5), 824-835 e14.

214. Ni, J. Z.; Grate, L.; Donohue, J. P.; Preston, C.; Nobida, N.; O'Brien, G.; Shiue, L.; Clark, T. A.; Blume, J. E.; Ares, M., Jr., Ultraconserved elements are associated with homeostatic control

of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev.* **2007,** *21* (6), 708-18.

215. Akan, I.; Love, D. C.; Harwood, K. R.; Bond, M. R.; Hanover, J. A., Drosophila O-GlcNAcase Deletion Globally Perturbs Chromatin O-GlcNAcylation. *J. Biol. Chem.* **2016,** *291* (19), 9906-19.

216. Gueroussov, S.; Weatheritt, R. J.; O'Hanlon, D.; Lin, Z. Y.; Narula, A.; Gingras, A. C.; Blencowe, B. J., Regulatory Expansion in Mammals of Multivalent hnRNP Assemblies that Globally Control Alternative Splicing. *Cell* **2017,** *170* (2), 324-339 e23.

217. Guo, Y. E.; Manteiga, J. C.; Henninger, J. E.; Sabari, B. R.; Dall'Agnese, A.; Hannett, N. M.; Spille, J.-H.; Afeyan, L. K.; Zamudio, A. V.; Shrinivas, K.; Abraham, B. J.; Boija, A.; Decker, T.-M.; Rimel, J. K.; Fant, C. B.; Lee, T. I.; Cisse, I. I.; Sharp, P. A.; Taatjes, D. J.; Young, R. A., Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **2019,** *572* (7770), 543-548.

218. Ying, Y.; Wang, X. J.; Vuong, C. K.; Lin, C. H.; Damianov, A.; Black, D. L., Splicing Activation by Rbfox Requires Self-Aggregation through Its Tyrosine-Rich Domain. *Cell* **2017,** *170* (2), 312-323 e10.

219. Joiner, C. M.; Levine, Z. G.; Aonbangkhen, C.; Woo, C. M.; Walker, S., Aspartate Residues Far from the Active Site Drive O-GlcNAc Transferase Substrate Selection. *J. Am. Chem. Soc.* **2019,** *141* (33), 12974-12978.

220. Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **2013,** *29* (1), 15-21.

221. Li, B.; Dewey, C. N., RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **2011,** *12*, 323.

222. Leng, N.; Dawson, J. A.; Thomson, J. A.; Ruotti, V.; Rissman, A. I.; Smits, B. M.; Haag, J. D.; Gould, M. N.; Stewart, R. M.; Kendziorski, C., EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* **2013,** *29* (8), 1035-43.

223. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S. L., Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **2009,** *10* (3), R25.

224. Quinlan, A. R.; Hall, I. M., BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010,** *26* (6), 841-2.

225. Harrow, J.; Frankish, A.; Gonzalez, J. M.; Tapanari, E.; Diekhans, M.; Kokocinski, F.; Aken, B. L.; Barrell, D.; Zadissa, A.; Searle, S.; Barnes, I.; Bignell, A.; Boychenko, V.; Hunt, T.; Kay, M.; Mukherjee, G.; Rajan, J.; Despacio-Reyes, G.; Saunders, G.; Steward, C.; Harte, R.; Lin, M.; Howald, C.; Tanzer, A.; Derrien, T.; Chrast, J.; Walters, N.; Balasubramanian, S.; Pei, B.; Tress, M.; Rodriguez, J. M.; Ezkurdia, I.; van Baren, J.; Brent, M.; Haussler, D.; Kellis, M.; Valencia, A.; Reymond, A.; Gerstein, M.; Guigo, R.; Hubbard, T. J., GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* **2012,** *22* (9), 1760-74.

226. Yates, A.; Akanni, W.; Amode, M. R.; Barrell, D.; Billis, K.; Carvalho-Silva, D.; Cummins, C.; Clapham, P.; Fitzgerald, S.; Gil, L.; Giron, C. G.; Gordon, L.; Hourlier, T.; Hunt, S. E.; Janacek, S. H.; Johnson, N.; Juettemann, T.; Keenan, S.; Lavidas, I.; Martin, F. J.; Maurel, T.; McLaren, W.; Murphy, D. N.; Nag, R.; Nuhn, M.; Parker, A.; Patricio, M.; Pignatelli, M.; Rahtz, M.; Riat, H. S.; Sheppard, D.; Taylor, K.; Thormann, A.; Vullo, A.; Wilder, S. P.; Zadissa, A.; Birney, E.; Harrow, J.; Muffato, M.; Perry, E.; Ruffier, M.; Spudich, G.; Trevanion, S. J.; Cunningham, F.; Aken, B. L.; Zerbino, D. R.; Flicek, P., Ensembl 2016. *Nucleic Acids Res.* **2016,** *44* (D1), D710-6.

227. Kent, W. J.; Sugnet, C. W.; Furey, T. S.; Roskin, K. M.; Pringle, T. H.; Zahler, A. M.; Haussler, D., The human genome browser at UCSC. *Genome Res.* **2002,** *12* (6), 996-1006.

228. Anders, S.; Huber, W., Differential expression analysis for sequence count data. *Genome Biol.* **2010,** *11* (10), R106.

229. Anders, S.; Reyes, A.; Huber, W., Detecting differential usage of exons from RNA-seq data. *Genome Res.* **2012,** *22* (10), 2008-17.

230. Ting, L.; Rad, R.; Gygi, S. P.; Haas, W., MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat Methods* **2011,** *8* (11), 937-40.

231. Paulo, J. A.; O'Connell, J. D.; Gygi, S. P., A Triple Knockout (TKO) Proteomics Standard for Diagnosing Ion Interference in Isobaric Labeling Experiments. *J. Am. Soc. Mass Spectrom.* **2016,** *27* (10), 1620-5.

232. Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villen, J.; Haas, W.; Sowa, M. E.; Gygi, S. P., A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **2010,** *143* (7), 1174-89.

233. Beausoleil, S. A.; Villen, J.; Gerber, S. A.; Rush, J.; Gygi, S. P., A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* **2006,** *24* (10), 1285-92.

234. Elias, J. E.; Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **2007,** *4* (3), 207-14.

235. Wang, A. C.; Jensen, E. H.; Rexach, J. E.; Vinters, H. V.; Hsieh-Wilson, L. C., Loss of O-GlcNAc glycosylation in forebrain excitatory neurons induces neurodegeneration. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113* (52), 15120-15125.

236. Lagerlof, O.; Hart, G. W.; Huganir, R. L., O-GlcNAc transferase regulates excitatory synapse maturity. *Proc. Natl. Acad. Sci. U. S. A.* **2017,** *114* (7), 1684-1689.

237. Harrison, S. D., Jr.; Bosin, T. R.; Maickel, R. P., Metabolism of 3-(2-dimethylaminoethyl)benzo(b)thiophene in vitro and in vivo in the rat. *Drug Metab. Dispos.* **1974,** *2* (3), 228-36.

238. St Jean, D. J., Jr.; Fotsch, C., Mitigating heterocycle metabolism in drug discovery. *J. Med. Chem.* **2012,** *55* (13), 6002-20.

239. Gaetani, M.; Sabatier, P.; Saei, A. A.; Beusch, C.; Yang, Z.; Lundström, S.; Zubarev, R. A., Proteome Integral Stability Alteration assay dramatically increases throughput and sensitivity in profiling factor-induced proteome changes. *bioRxiv* **2018**.

240. Savitski, M. M.; Reinhard, F. B.; Franken, H.; Werner, T.; Savitski, M. F.; Eberhard, D.; Martinez Molina, D.; Jafari, R.; Dovega, R. B.; Klaeger, S.; Kuster, B.; Nordlund, P.; Bantscheff, M.; Drewes, G., Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **2014,** *346* (6205), 1255784.

241. Martinez Molina, D.; Jafari, R.; Ignatushchenko, M.; Seki, T.; Larsson, E. A.; Dan, C.; Sreekumar, L.; Cao, Y.; Nordlund, P., Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **2013,** *341* (6141), 84-7.

242. Lund, P. J.; Elias, J. E.; Davis, M. M., Global Analysis of O-GlcNAc Glycoproteins in Activated Human T Cells. *J. Immunol.* **2016,** *197* (8), 3086-3098.