DEVELOPMENT OF NEW PHOTOCROSSLINKING APPROACHES TO DISCOVER BINDING PARTNERS OF O-GLCNAC-MODIFIED PROTEINS

APPROVED BY SUPERVISORY COMMITTEE

Jennifer J. Kohler, Ph.D.

Nicholas K. Conrad, Ph.D.

Benjamin P. Tu, Ph.D.

Beatriz M. Fontoura, Ph.D.

DEDICATION

I would like to thank my family, loved ones, and friends for their support and encouragement to continue my education. I dedicate this work to my parents for their love and support to me. My grandparents were both affected by human diseases that lead me to become interested in science. I would like to thank my siblings for the amazing childhood memories and being my first set of best friends. Also, I would like to thank my best friends from high school, Marissa and Ashley, for their support and fun memories. My friends from San Diego State, the Grotjahn lab, Karina Kangas, and Devin Wakefield for their support in continuing my education in research. Also, the Keene family as they have adopted me as one of their own. I acknowledge the Kohler lab members, past and current, as they have assisted in my development as a scientist. I would also like to thank the members of my Graduate Committee for their guidance of my development as a scientist, and discussions.

DEVELOPMENT OF NEW PHOTOCROSSLINKING APPROACHES TO DISCOVER BINDING PARTNERS OF O-GLCNAC-MODIFIED PROTEINS

by

ANDREA CHRISTINE RODRIGUEZ

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

December, 2015

DEVELOPMENT OF NEW PHOTOCROSSLINKING APPROACHES TO DISCOVER BINDING PARTNERS OF O-GLCNAC-MODIFIED PROTEINS

Publication No._____

Andrea Christine Rodriguez, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2015

Supervising Professor: Jennifer Jean Kohler, Ph.D.

O-linked β-N-acetylglucosamine (O-GlcNAc) is an abundant post-translational modification that is regulated by two enzymes, O-GlcNAc transferase (OGT) and O-GlcNAc hydrolase (OGA). While it is elusive how O-GlcNAc alters protein function, altered O-GlcNAc levels are associated with human diseases. To gain insight into the functional consequences of O-GlcNAc-ylation, we reported a method to incorporate the diazirine photocrosslinking group onto O-GlcNAc residues in cellular proteins. Photocrosslinking O-GlcNAc, O-GlcNDAz, yields covalent crosslinking

between O-GlcNAc-ylated proteins and their binding partners, which further analysis can confirm these interactions.

I applied the GlcNDAz technology to a heavily O-GlcNAc-modified nucleoporin NUP98 and NUP98 leukemogenic fusions, produced under chromosomal translocation, to gain insight into the mechanism of NUP98 fusion-mediated cell transformation in leukemia. The wild-type nucleoporins are associated with nuclear trafficking. In **chapter 2**, I demonstrated both NUP98 and NUP98 fusions are O-GlcNAc-modified. Additionally, evidence suggested O-GlcNAc is near the site of interaction based on crosslinking experiments. While a powerful approach, the utility of in-cell O-GlcNDAz crosslinking was restricted by several limitations.

To solve this challenge, I first engineered a mutant OGA that is better able to remove GlcNDAz from proteins in order to facilitate homeostasis of O-GlcNDAz modification in cells and for potential use as an enzyme to release crosslinked material (**chapter 3**). Next, I constructed a mutant OGT to preferentially add GlcNDAz to proteins, in order to maximize possible crosslinking material (**chapter 4**). Finally, I initiated the development of a chemo-enzymatic synthesis of O-GlcNDAz-ylated peptides using three enzymes: A bifunctional enzyme consisting of *Bifidobacterium longum* N-acetylhexosamine 1-kinase (NahK) and *E. coli* N-acetylglucosamine-1-phosphate uridyltransferase (GlmU) to generate UDP-GlcNDAz, then added GlcNDAz to peptides by human OGT (**chapter 5**). Together, these strategies allow for the development of a complementary *cell free* O-GlcNDAz

crosslinking approach. O-GlcNDAz-modified peptides generated from this reaction can be crosslinked to molecules from cell lysates. This technology can be used to identify binding partners of O-GlcNAc-modified proteins, including both normal NUP98 and leukemogenic NUP98 fusions and may reveal functional roles of O-GlcNAc on NUPs and NUP fusions.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Wands AM*, Fujita A*, McCombs JE, Nygren J, Dedic B, **Rodriguez AC**, Nischan N, Bond MR, Mettlen M, Trudgian DC, Lemoff A, Quiding-Järbrink, Gustavsson B, Steentoft C, Clausen H, Mirzaei H, Teneberg S, Yrlid U, Kohler JJ. Fucosylation and Protein Glycosylation Create Functional Receptors for Cholera Toxin. *eLIFE* (2015) 10.7554./eLIFE09545. *Authors contributed equal work.

Rodriguez AC*, Yu S-H*, Li B, Zegzouti H, and Kohler JJ. Enhanced Transfer of a Photocrosslinking O-GlcNAc Analog by an O-GlcNAc Transferase Mutant with Converted Substrate Specificity. *Journal of Biological Chemistry.* (2015) 290: 22638-22648. *Authors contributed equal work.

Pham ND, Fermaintt CF*, **Rodriguez AC***, McCombs JE, Nischan N, and Kohler JJ. Cellular Metabolism of Unnatural Sialic Acid Precursors and Analogs. *Glycoconjugate Journal* (2015) 32: 515-529. *Authors contributed equal work.

Rodriguez AC and Kohler JJ. Recognition of diazirine-modified O-GlcNAc by human O-GlcNAcase. *MedChemComm* (2014), 5 (8): 1227-1234.

LIST OF FIGURES

CHAPTER ONE FIGURES

CHAPTER TWO FIGURES

CHAPTER THREE FIGURES

CHAPTER FOUR FIGURES

CHAPTER FIVE FIGURES

LIST OF TABLES

LIST OF DEFINITIONS

GlcNAc – N-acetylglucosamine

O-GlcNAc – O-linked β-N-acetylglucosamine

GalT1 – bovine milk galactosyltransferase T1

PNGaseF – peptide N-glycosidase F

HBP – hexosamine biosynthetic pathway

UDP-GlcNAc – uridine diphosphate N-acetylglucosamine

OGT – O-GlcNAc transferase

OGA – O-GlcNAc hydrolase/ O-GlcNAcase

UAP1 – UDP-N-acetylglucosamine pyrophosphorylase

NUP – nucleoporin

NPC – nuclear pore complex

FG repeats – phenylalanine-glycine repeat regions/domains

PTM – post-translational modification

WGA – wheat germ agglutinin

ETD – electron transfer dissociation

ECD – electron capture dissociation

GlcNDAz – N-acetylglucosamine photocrosslinking analog containing diazirine

O-GlcNDAz – O-linked β-N-acetylglucosamine photocrosslinking analog

GlcNDAz-1P – N-acetylglucosamine-1-phosphate photocrosslinking analog

Ac_3-GlcNDAz-1P$(Ac\text{-}SATE)_2$ – cell-permeable GlcNDAz-1P

UDP-GlcNDAz – uridine diphosphate N-acetylglucosamine photocrosslinking analog

HPAEC – high performance anion exchange chromatography

CBP – CREB-binding protein

HAT – histone acetyltransferase

PUGNAc – *O*-(2-acetamido-2-deoxy-D-glycopyranosylidene)amino-N-phenylcarbamate

NButGT – 1,2-dideoxy-2′-propyl-α-D-glucopyranoso-[2,1-D]-Δ2′-thiazoline

GT – N-acetyl-D-glucosamine-thiazoline

*p*NP-GlcNAc – 4-nitrophenol-N-acetylglucosamine

*p*NP-GlcNDAz- 4-nitrophenol-N-acetylglucosamine photocrosslinking analog

GalNAc – N-acetylgalactosamine

HPLC – high performance liquid chromatography

CKII – casein kinase II

TPR – tetratricopeptide repeats

wtOGT – wild-type OGT

GlmU – uridinyl transferase

NahK – N-acetylhexosamine 1-kinase

NahK/GlmU – bifunctional fusion enzyme of NahK fused to catalytic domain of GlmU

IRS1 – insulin receptor substrate 1

+P2 – human α A-crystallin

CHAPTER ONE:

O-GlcNAc is the most essential modification you never heard of

INTRODUCTION: Why are sugars important?

Sugars are fundamental building blocks found in organisms. In mammals, sugars form glycans composed of monosaccharide building blocks. N-acetylglucosamine (GlcNAc) is one of nine essential sugars found in humans. These essential sugars are only a subset of carbohydrates found in mammals. In particular, GlcNAc is found both in extracellular glycosylation, where it is present in an array with many other sugars, and intracellular glycosylation, where it is a single sugar post-translational modification (PTM). O-linked β-N-acetylglucosamine (O-GlcNAc) is an intracellular PTM found on serine and threonine residues. O-GlcNAc is not well studied and its functions remain poorly defined.

Discovery of the O-GlcNAc modification

In 1983, O-GlcNAc was discovered by accident in murine thymocytes (1). A graduate student radiolabeled terminal GlcNAc residues on glycans using UDP-[^{3}H]-galactose and bovine milk galactosyltransferase (GalT1). She expected to radiolabel extracellular glycoconjugates containing GlcNAcs. After removal of O-linked glycosylation with alkali-induced β-elimination and

removal of N-linked glycosylation through Peptide N-Glycosidase F (PNGase F) treatment, she discovered radiolabeling of single GlcNAcs on proteins in the nuclear and cytosolic compartments. There are several characteristics of O-GlcNAc that delayed the discovery.

Four main traits of O-GlcNAc assist to the late finding of O-GlcNAc. Unlike other modifications, O-GlcNAc is uncharged and has a low molecular weight. These traits explain why O-GlcNAc on proteins isn't detected by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Upon lysing of cells, hydrolases remove O-GlcNAc during protein isolation, and suggests O-GlcNAc is labile. O-GlcNAc-ylation is found in substoichiometric amounts on proteins (2). Hart and coworkers demonstrated 0.1 GlcNAc per protein by radiolabeling. Finally, traditional mass spectrometry conditions are unable to detect O-GlcNAc because the modification is lost during harsh ionization conditions. Since the O-GlcNAc modification is a recent recognition, relatively little is known about its functional roles.

The enzymes that regulate O-GlcNAc

UDP-GlcNAc, a nucleotide donor sugar, is essential for production of O-GlcNAc-ylation. UDP-GlcNAc is synthesized through two pathways. From glucose, the hexosamine biosynthetic pathway (HBP) generates UDP-GlcNAc (Fig. 1.1A). An alternative pathway to produce UDP-GlcNAc is to recycle

GlcNAc and use the GlcNAc salvage pathway (Fig. 1.1B). Because UDP-

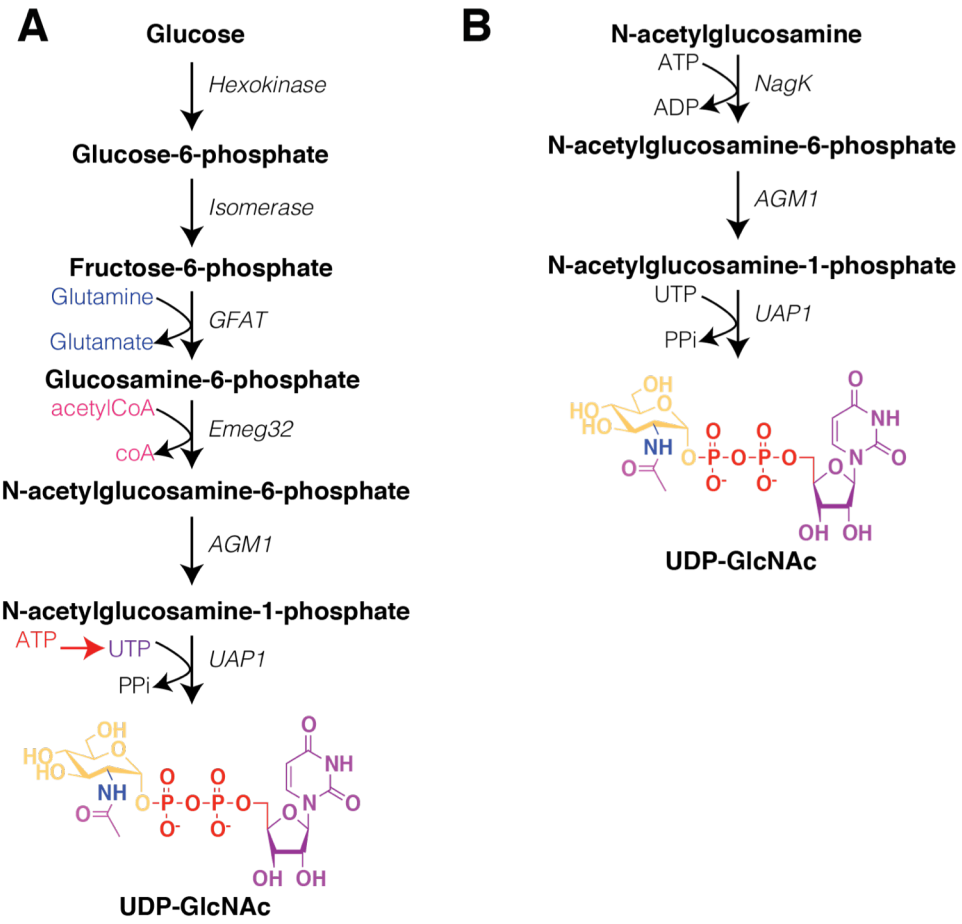

Figure 1.1. Two pathways for UDP-GlcNAc production. *A*, Glucose is taken up into cells and incorporated into the hexosamine biosynthetic pathway (HBP). A hexokinase converts glucose into glucose-6-phosphate. This 6-phosphate sugar can be isomerized into fructose-6-phosphate. Then, glutamine: fructose-6-phosphate aminotransferase (GFAT) alters the sugar to glucosamine with the use of glutamine. Glucosamine-6-phosphate is converted to N-acetylglucsamine-6-phosphate by glucosamine-6-phosphate acetyltransferase (Emeg32). N-acetylglucosamine mutase 1 (AGM1) isomerizes the phosphate from the carbon 6 position to the carbon 1 position. Then UDP-N-acetylglucosamine pyrophosphorylase 1 (UAP1) transfers UDP onto GlcNAc. *B*, An alternative method is that UDP-GlcNAc can be generated through the GlcNAc salvage pathway. GlcNAc can be converted to GlcNAc-6-phosphate by N-acetylhexosamine 1-kinase (NagK). Then AGM1 can isomerize the phosphate to the carbon 1 position. As in the HBP, UAP1 converts the GlcNAc-1P to UDP-GlcNAc.

GlcNAc biosynthesis requires glucose, acetyl-CoA, glutamine, and UTP production, this nucleotide sugar provides an integrated report on the metabolic state of a cell.

Similar to phosphorylation, O-GlcNAc is a reversible modification (Fig. 1.2), but while multiple kinases and phosphatases exist, mammalian cells harbor only a single enzyme capable of producing O-GlcNAc and a single enzyme capable of removing it. The O-GlcNAc transferase (OGT) catalyzes the addition of a single GlcNAc from the donor nucleotide sugar, UDP-GlcNAc (3), onto serine and threonine residues. As mentioned previously, O-GlcNAc is dynamic, and cycles on and off a protein much faster than the lifetime of the protein. The O-GlcNAc hydrolase (OGA) functions to remove GlcNAc from the

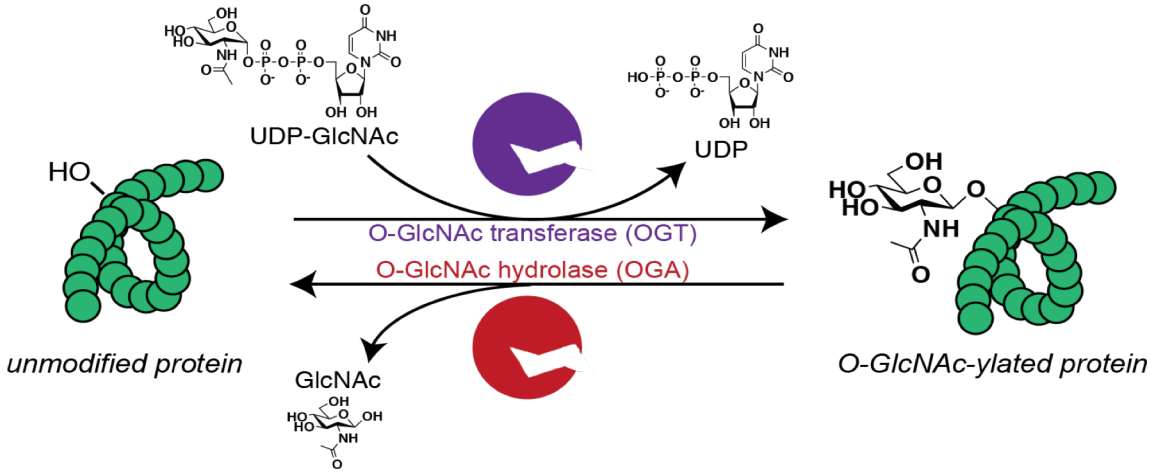

Figure 1.2. O-GlcNAc is a dynamic PTM. Two enzymes regulate O-GlcNAc. The O-GlcNAc transferase (OGT) transfers a single GlcNAc onto proteins from the nucleotide donor sugar, UDP-GlcNAc. This PTM is hydrolytically removed by the O-GlcNAc hydrolase (OGA).

modified proteins (4). O-GlcNAc integrates multiple metabolic signals and plays critical roles in cells' ability to respond to changes in nutritional state and other cues (5).

O-GlcNAc-ylation is found on many proteins, but at particularly high levels on nucleoporins

Typically, glycosylation on proteins occurs in the secretory pathway.

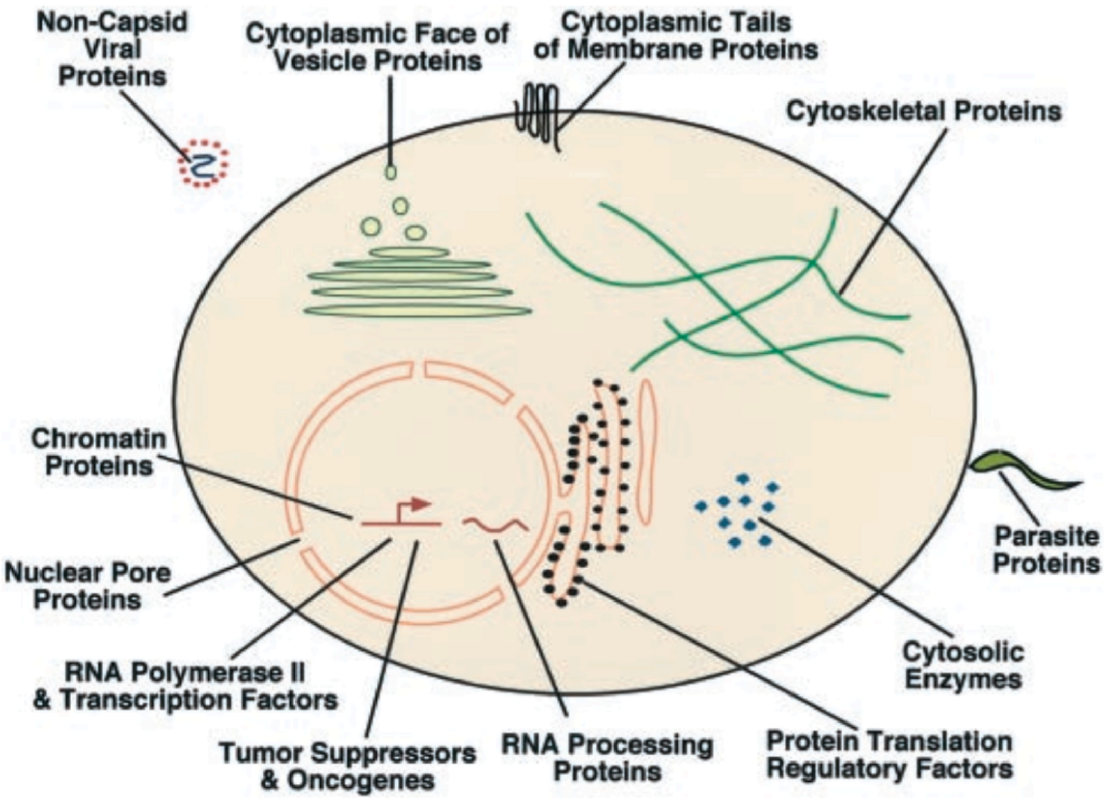

Figure 1.3. O-GlcNAc is widespread in multi-cellular organisms. O-GlcNAc is a ubiquitous modification. It is found in many compartments of the cell. These O-GlcNAc-modified proteins belong to different functional classes. This figure was originally published in Journal of Biological Chemistry. Comer, FI & Hart, GW. O-Glycosylation of Nuclear & Cytosolic Proteins. *Journal of Biological Chemistry.* 2000; 275: 29179-29182. © the American Society for Biochemistry and Molecular Biology.

However, O-GlcNAc is mostly synthesized in the nuclear and cytosolic compartments of the cells. Abundant amounts of O-GlcNAc-ylation are reported on nuclear and cytoplasmic proteins (including chromatin components and nuclear pore proteins). Also, evidence reveals O-GlcNAc on mitochondrial and extracellular proteins in most eukaryotes (Fig. 1.3) (5-7). In addition, viral proteins, like Plum pox potyvirus, are O-GlcNAc-modified (8, 9).

Nucleoporins (NUPs) are some of the first proteins identified to be O-GlcNAc-modified (10). NUPs form a nuclear pore complex (NPC) within the nuclear membrane to facilitate nuclear trafficking. Phenylalanine-glycine repeat regions (FG repeats) is found in one category of NUPs, known as FG NUPs. These FG repeats line the nuclear pore and interact directly with nuclear transport factors as they pass through the NPC. These FG repeats are interspersed with serines and threonines that are known to be heavily O-GlcNAc-modified. In this dissertation, I examine FG NUPs, like NUP153 and NUP98. NUP153 facilitates trafficking on the nuclear side of the NPC, while NUP98 assists in trafficking at both sides of the NPC (11-13). Both NUP153 and NUP98 are O-GlcNAc-modified (14, 15).

Tools to identify O-GlcNAc-ylated proteins and residues modified with O-GlcNAc

The occurrence of O-GlcNAc-ylation was confirmed by using a carbohydrate-binding lectin named wheat germ agglutinin (WGA), which is known to recognize GlcNAc and sialic acids (16, 17). Blobel and co-workers

demonstrated the WGA lectin interacted with NUP62, a FG repeat NUP, and inhibit nuclear trafficking. They confirmed that components of the nuclear pore complex bear O-GlcNAc moieties (18). Although, WGA lectin can be used to enrich for GlcNAc-modified proteins, WGA has low affinity to O-GlcNAc. also spurred demand for antibodies that specifically recognize O-GlcNAc. Since the discovery of O-GlcNAc, tools to advance the field have been developed to identify O-GlcNAc-modified proteins and O-GlcNAc sites.

RL2 is the first antibody that recognizes O-GlcNAc moieties. This antibody was raised against components of the nuclear pore. RL2 aided in the localization of nuclear and cytosolic proteins, such as NUP63, NUP54, NUP180 and NUP100, from rat liver nuclear envelopes (10, 19). As of now, other commercially available antibodies are available, such as CTD110.6 and peptide-based antibodies. CTD110.6 was raised against the O-GlcNAc-modified carboxyl-terminal domain (CTD) of RNA polymerase II (RNA pol II). CTD110.6 exhibits cross-reactivity toward other nuclear and cytosolic O-GlcNAc-ylated proteins (20). Another sets of antibodies recognize a broader range of O-GlcNAc-modified proteins because they were developed using synthetic O-GlcNAc-modified peptides as antigens (21). These antibodies allow for recognition of O-GlcNAc-modified proteins.

Another tool that aids in the identification of O-GlcNAc-modified proteins was inhibition of O-GlcNAc removal. Utilization of OGA inhibitors, such as

PUGNAc and Thiamet G, demonstrate presence of a higher level of O-GlcNAc-modified proteins at the time of protein isolation (22). However, deregulation of O-GlcNAc levels through use of these compounds may obscure true O-GlcNAc regulation. Therefore, the identified targets may not be biologically relevant.

Additionally, chemo-enzymatic tagging techniques assist in the identification of O-GlcNAc-modified proteins. The tagging techniques overcome the issues of lability and substoichiometric. One method of chemo-

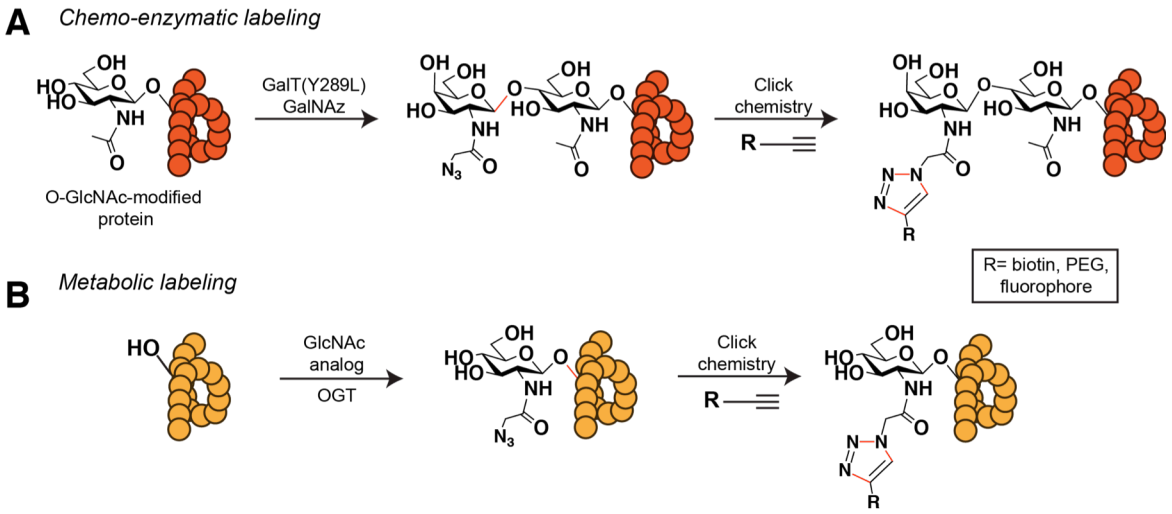

Figure 1.4. Chemo-enzymatic labeling of O-GlcNAc-modified proteins. *A*, O-GlcNAc-modified proteins can be labeled outside of the cells (*A*) or metabolically through GlcNAc salvage pathway (*B*). Chemo-enzymatic labeling of O-GlcNAc-modified proteins in lysate. The O-GlcNAc-modified proteins in lysate are incubated with Ac_4-GalNAz and GalT1(Y289L) to transfer the GalNAz onto the O-GlcNAc-modified protein. *B*, O-GlcNAc-modified substrates can be metabolically labeled with a GlcNAc analog (GlcNAz, GalNAz) using the GlcNAc salvage pathway. The azide can be reacted with an alkyne for the click chemistry reaction. Through click chemistry with an alkyne handle O-GlcNAc-modified proteins can be labeled with biotin, PEG or fluorophores.

enzymatic tagging is to transfer a GalNAc analog onto the O-GlcNAc-modified proteins. Hsieh-Wilson and co-workers developed this method by using a mutant GalT1 (GalT1(Y289L)) and a N-acetylgalactosamine analog containing an azide (GalNAz) (23-25). The GalT1(Y289L) transfers GalNAz onto terminal GlcNAcs, like O-GlcNAc-modified proteins, in lysates similar to the protocol that discovered O-GlcNAc (1). The azide handle can react with an alkyne conjugated to a probe, such as biotin or polyethylene glycol, using the Huisgen cycloaddition known as click chemistry, allowing for enrichment or identification of O-GlcNAc-modified proteins (Fig. 1.4A). Hsieh-Wilson and co-workers used this chemo-enzymatic tagging of O-GlcNAc-modified proteins to show that PFK1 is O-GlcNAc-modified at serine 529 (26). An advantage of this strategy is that it can be applied to many types of sample, including tissues or primary cells.

Metabolic chemo-enzymatic tagging is an alternative method established by Bertozzi, Pratt, and co-workers to identify O-GlcNAc-modified substrates. This method utilizes the GlcNAc salvage pathway of living cells to metabolically label O-GlcNAc-modified proteins. Bertozzi and co-workers synthesized cell-permeable sugar analogs, GalNAz and GlcNAz. These analogs incorporated into the GlcNAc salvage pathway to label O-GlcNAc-modified proteins (27, 28). These sugar analogs contain the azide handle, and can perform the Huisgen cycloaddition (Fig. 1.4B). Also, Pratt and co-workers

synthesized a different cell-permeable GlcNAc analog, GlcNAlk, to metabolically label O-GlcNAc-modified proteins (29). GlcNAlk contains the alkyne handle and can react with an azide-biotin reagent under the Huisgen cycloaddition condition.

These unnatural GlcNAc analogs demonstrate incorporation into various cells and conversion to the respective donor nucleotide sugars (UDP-GlcNAz, UDP-GalNAz and UDP-GlcNAlk) through hijacking of the GlcNAc salvage pathway. In cells, GalNAz converts to UDP-GalNAz, then epimerizes to UDP-GlcNAz through the UDP-galactose 4-epimerase. Once the sugar is activated with UDP, OGT can transfer it to target proteins. These chemo-enzymatic tags assist in the identification of O-GlcNAc-modified proteins and the O-GlcNAc site through proteomics analysis.

In complement with chemo-enzymatic tagging approaches, ionization conditions of mass spectrometry like electron transfer dissociation (ETD) and electron capture dissociation (ECD) enable softer ionization of peptides without the loss of the O-GlcNAc modification. Using the chemo-enzymatic methods, in combination with LC-MS/MS approaches contribute to the identification of O-GlcNAc-ylated sites (30-32). To date, more than 4000 proteins have been identified to be O-GlcNAc modified, and this number is still increasing (6).

The roles of O-GlcNAc modification

O-GlcNAc may exert biological effects at the cellular and organismal levels.

Disruption of O-GlcNAc homeostasis has catastrophic effects. Mutation of OGT or OGA, or increasing glucose uptake demonstrated a perturbation in the cell cycle and resulted in cell death (33). Additionally, the importance of O-GlcNAc at the organismal level revealed embryonic lethal phenotype in OGT knockout mice, and defects in the dauer formation and insulin-like signaling in OGT knockout *C.elegans* (4, 34). Altered O-GlcNAc levels are associated with several human diseases, such as diabetes, Alzheimer's and cancers (35-37).

A subset of evidence supports a relationship between O-GlcNAc and phosphorylation. Since both modifications occur on serine and threonine residues, there are reports that O-GlcNAc and phosphorylation can sometimes compete for the same sites (15, 38-40). For example, O-GlcNAc-ylation occurs on the CTD of RNA pol II, which is also known to be phosphorylated (41). Lewis and co-workers demonstrated in cell free and cell-based systems with mutagenesis and immunoblots that O-GlcNAc-ylation occurs at the same sites as phosphorylation on the CTD (42). They further revealed through mutagenesis in the two systems that O-GlcNAc-ylation of the CTD of RNA pol II halts transcription activity. This evidence demonstrates O-GlcNAc has functional relevance.

While O-GlcNAc-ylation has been demonstrated to inhibit RNA pol II activity, in other cases O-GlcNAc increased a protein's activity (43). For example, Yi and co-workers demonstrated O-GlcNAc-ylation of glucose-6-phosphate dehydrogenase (G6PD) increased activity (43). Glycosylation at serine 84 of G6PD increased

amounts of products generated from the pentose phosphate pathway. These results revealed insight into the mechanism of tumor survival for human lung cancer. These two examples, RNA pol II and G6PD are just two proteins out of the thousands of proteins known to be O-GlcNAc-modified, propose O-GlcNAc is important in regulating a protein's function.

Another suggested function of O-GlcNAc is mediating interactions between proteins (44, 45). Groner and co-workers demonstrated through immunoprecipitation that CREB-binding protein (CBP) preferably interacts with the O-GlcNAc-modified form of Stat5 (signal transducer and activator of transcription) (44). O-GlcNAc may also mediate interactions between nucleoporins; Yoneda and co-workers revealed through immunoprecipitation and OGT knockdowns that NUP88 binds to O-GlcNAc-modified NUP62 (45). In winter wheat, Chong and co-workers disclosed that O-GlcNAc-ylation of threonine 17 of the RNA-binding protein TaGRP2 was necessary for interaction with the jacalin lectin Ver2 (46). Although these reports suggested O-GlcNAc might assist protein-protein interactions, no O-GlcNAc-dependent interactions have been characterized structurally in mammals. Furthermore, the chemo-enzymatic tagging technologies do not assist in identifying the O-GlcNAc-modified protein-protein interactions.

A photocrosslinking sugar to identify binding partners of O-GlcNAc-modified proteins

O-GlcNAc may mediate protein-protein interactions as supported by literature (45, 46). We hypothesize that O-GlcNAc, like other PTMs, may promote protein-protein interactions at the interface or near the site of interaction. To determine if a PTM is necessary for interaction, ideally the residue modified is mutated to confirm the PTM-dependence of the interaction. As mentioned previously, O-GlcNAc modifications are labile, dynamic, and occur in substoichiometric levels, resulting in the failure of standard methods to confirm this mediator function of protein-protein interactions. As an example, immunoprecipitation is a tool that pulls down all the proteins associated with the protein of interest, regardless of the nature of that interaction, such as PTM-dependence, or amino acid-dependence. As of yet, we can't state that O-GlcNAc promotes protein-protein interactions in mammals.

Previously the Kohler lab reported a technology to introduce a diazirine photocrosslinking group on the N-acyl position of O-GlcNAc residues. This unnatural PTM, which they term O-GlcNDAz, can covalently crosslink to the molecules that surround an O-GlcNAc-modified protein. The diazirine is activated by UV irradiation (365 nm) causing the release of nitrogen gas and leaving behind a highly reactive carbene. The carbene reacts within microseconds to any nearby molecule, and forms a covalent crosslink between an O-GlcNAc-modified protein and its neighbor (Fig. 1.5A). In this

way, O-GlcNDAz can provide a photochemical “snapshot” of the binding

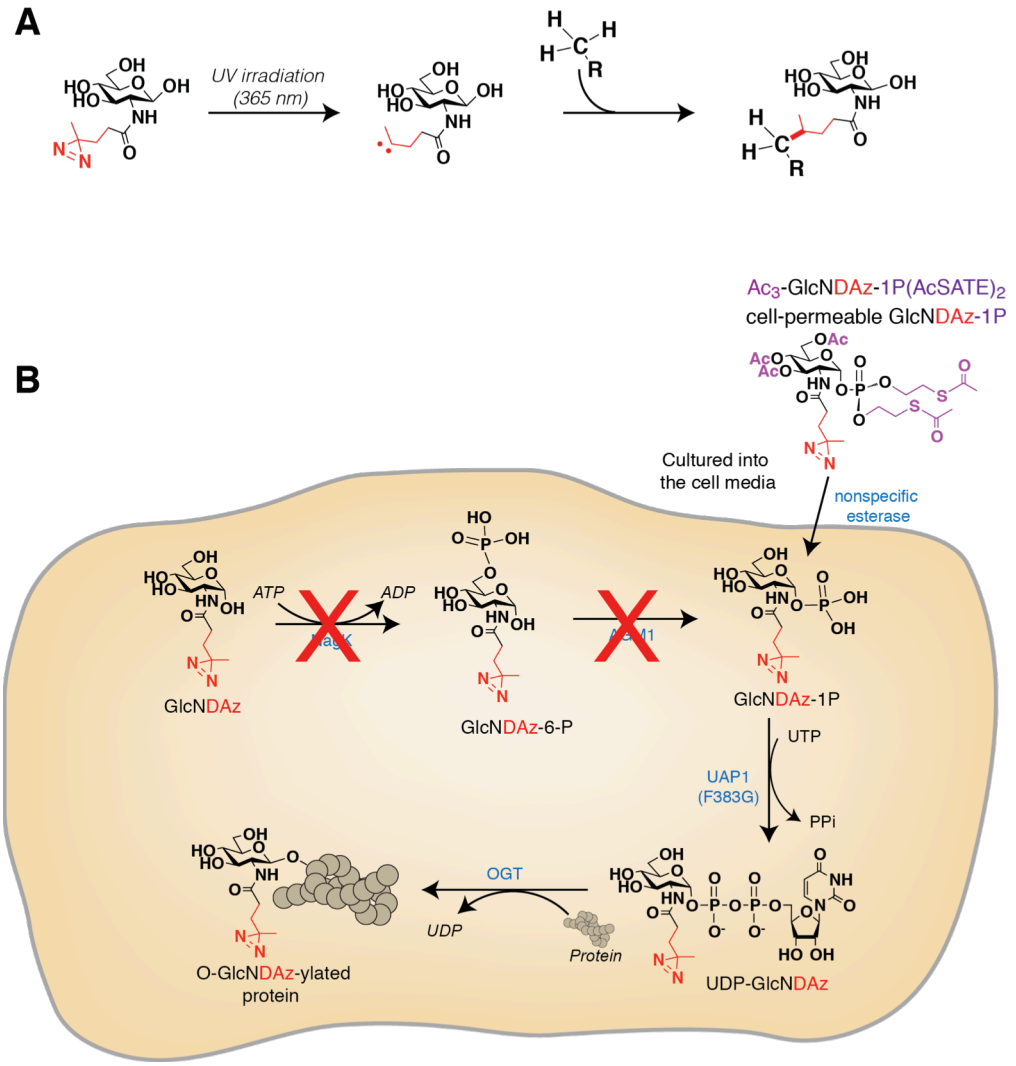

Figure 1.5. Engineering cells to produce O-GlcNDAz. *A*, The diazirine functional group is synthetically added to the N-acyl position of GlcNAc. Through UV irradiation, the diazirine is activated. This initiates the release of nitrogen gas that leaves behind a highly reactive carbene. The carbene reacts with any neighboring molecule to create a crosslinked complex of GlcNDAz covalently bound to the molecule. *B*, Cells are cultured with a cell-permeable GlcNDAz-1P analog, Ac_3-GlcNDAz-1P(Ac-SATE)$_2$, which is deprotected inside of the cells to yield GlcNDAz-1P. A mutant form of UAP1, UAP1(F383G), converts GlcNDAz-1P to UDP-GlcNDAz, the donor nucleotide sugar. OGT accepts UDP-GlcNDAz, and transfers GlcNDAz onto proteins, creating O-GlcNDAz.

interactions of an O-GlcNAc-modified protein (47). The diazirine handle addresses two issues of the O-GlcNAc characteristic. Since O-GlcNAc-ylation is dynamic, the interaction must be occurring at the time of irradiation is selected. Additionally, only one O-GlcNDAz is necessary to capture the interaction, addressing the issue of substoichiometry. Because O-GlcNDAz crosslinking is irreversible, non-specific binding partners can be removed through harsh stringent washing. After isolation of the crosslinked complex, LC-MS/MS analysis is employed to identify crosslinking partners. The Kohler lab proposes using GlcNDAz to identify the "proximity interactions" (48) involved in O-GlcNAc-dependent interactions for O-GlcNAc-modified proteins in order to provide insight into the functional roles of the O-GlcNAc modification (49).

To identify binding partners of O-GlcNAc-modified proteins, the Kohler lab developed a cell-based O-GlcNDAz crosslinking assay. A cell-permeable GlcNDAz-1P analog, Ac_3-GlcNDAz-1P(Ac-SATE)$_2$, was synthesized and added to the cell culture media. This fully protected GlcNDAz-1P analog is hydrophobic and can ready cross the plasma membrane. Once inside cells, the protecting groups are removed, releasing GlcNDAz-1P. Unfortunately, GlcNDAz-1P is a poor substrate for UDP-N-acetylglucosamine pyrophosphorylase 1 enzyme (UAP1), the enzyme that converts GlcNAc-1P to UDP-GlcNAc. Through a single point mutation to UAP1 at phenylalanine 383

to glycine (UAP1(F383G)), confers activity toward the diazirine-modified analog, GlcNDAz-1P. Once UDP-GlcNDAz is produced, OGT transfers GlcNDAz in place of the normal O-GlcNAc modification (Fig. 1.5B).

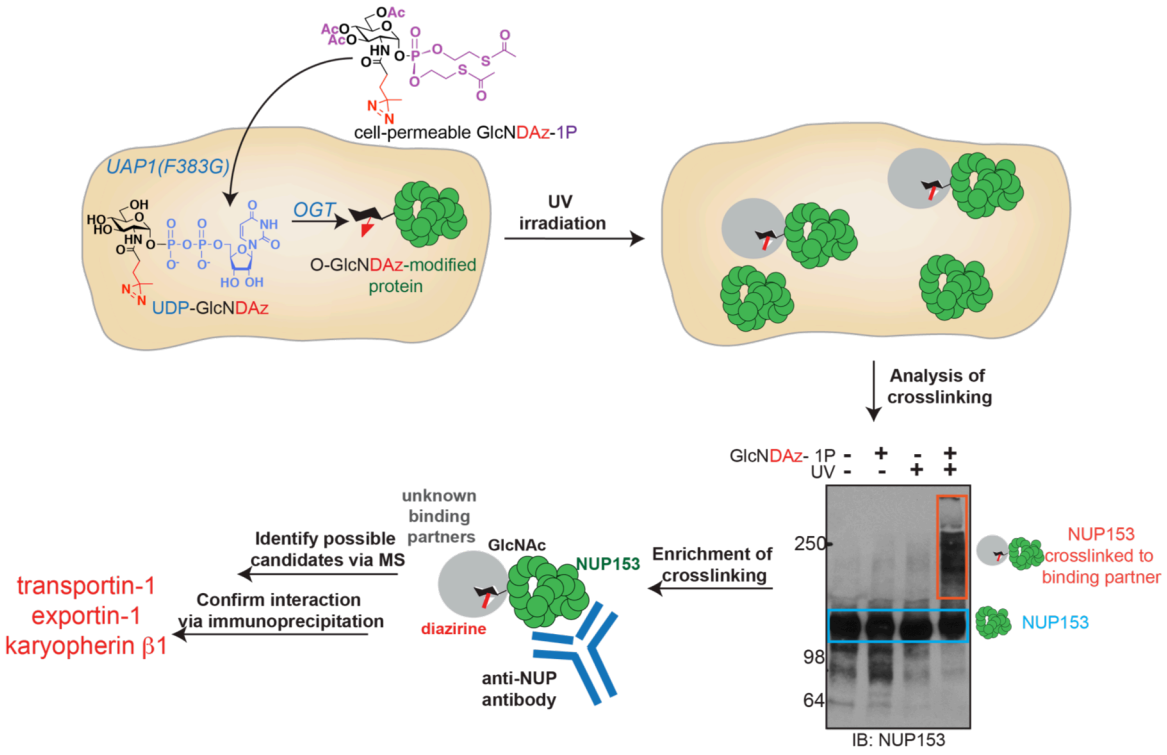

Figure 1.6. O-GlcNDAz-ylation assists in identifying binding partners of O-GlcNAc-modified proteins. Cells engineered to express the UAP1(F383G) are cultured with vehicle or sugar. The cells are then placed in ice or under UV irradiation. The O-GlcNAc-modified NUP153 was examined by immunoblot to detect crosslinked complexes, which appear at a higher molecular weight (red box). The crosslinked complexes are enriched using antibodies that recognize the O-GlcNAc-modified protein of interest, such as NUP153. The samples are submitted for proteomics analysis with in-gel trypsin digest. The candidates are verified by immunoprecipitation of the candidate and immunoblot for the O-GlcNAc-modified protein of interest.

To validate the utility of this technology, the Kohler lab identified interaction partners of the heavily O-GlcNAc-modified nucleoporin NUP153. The engineered cells demonstrated crosslinking of the heavily O-GlcNAc-modified NUP153 (higher molecular bands of NUP153, red box, Fig. 1.6) in the samples treated with both cell-permeable GlcNDAz-1P and UV irradiation. From living cells, crosslinked complexes were enriched through mAB414 antibody that recognizes several NUPs including NUP153, and proteomics analysis revealed several potential binding partners of NUP153.

To examine these candidates, the crosslinking experiments were repeated as previously described. The potential binding partners were immunoprecipitated with an antibody that recognizes the binding partners, followed by immunoblotting for NUP153. Kohler and co-workers revealed that transportin-1, exportin-1, and karyopherin β1 interact with O-GlcNAc-modified NUP153 (Fig. 1.6). These results suggest that through O-GlcNDAz binding partners of O-GlcNAc-modified proteins can be identified.

Exploiting GlcNDAz for other O-GlcNAc-modified proteins, NUP98 and leukemogenic NUP98 fusions

During my studies, I used the cell-based O-GlcNDAz crosslinking assay to examine other O-GlcNAc-modified proteins, including NUP98 (50). My interest in studying NUP98 stemmed from reports that chromosomal translocations involving this gene are associated with several types of leukemias (51). These chromosomal translocations lead to the production of fusion proteins with NUP98's N-terminus FG repeats fused to nucleic acid recognizing domains from other proteins, such as transcription factors (Fig. 1.7). The activity of these leukemogenic fusions is unknown, but there is speculation that they may act as transcription factors. Since the leukemogenic fusions include the FG repeats, I speculated that they would also be O-GlcNAc-modified, an idea that was supported by my observation that the fusion proteins could engage in O-GlcNDAz crosslinking. My goal was to identify the binding partners interacting with the O-GlcNAc-modified FG regions of NUP98 and leukemogenic NUP98 fusions, and study their

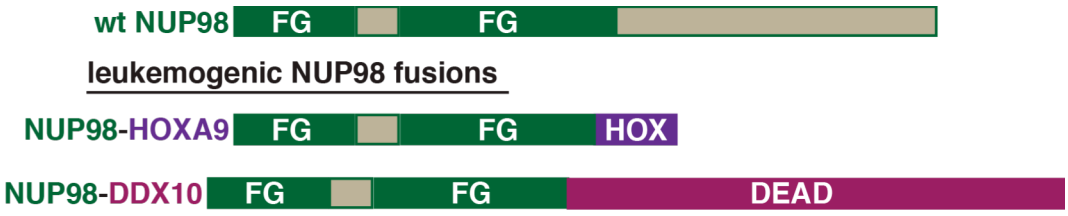

Figure 1.7. Leukemogenic NUP98 fusion proteins. In leukemia, chromosomal translocations occur with the FG repeats of NUP98 and nucleic acid binding proteins, such as transcription factors. NUP98-HOXA9 and NUP98-DDX10 are two well-studied leukemogenic fusions of NUP98.

importance in transformation of the cell. However, difficulties associated with purification of the crosslinked complexes caused me to switch my focus toward developing a cell free O-GlcNDAz crosslinking method to complement the cell-based approach.

Transiting towards an cell free approach

As discussed earlier, O-GlcNAc may play a role as a mediator in protein-protein interactions. Through immunoprecipitation and deregulation of O-GlcNAc-ylation, NUP88 was suggested to interact with O-GlcNAc-modified NUP62 (45). Also, the cell-based O-GlcNDAz crosslinking demonstrated that O-GlcNAc-modified NUP153 interacts with several karyopherins (47). I hypothesize that O-GlcNAc mediates protein-protein interactions. However, several challenges with the cell-based O-GlcNDAz crosslinking method became apparent in attempting to determine if O-GlcNAc provided structurally support for protein-protein interactions.

These issues set the tone for a new direction in my graduate studies. To state that an O-GlcNAc-modified residue is responsible for mediating protein-protein interaction, I will need to identify the site and confirm by mutation. To surmount this issue, a peptide substrate encoded from O-GlcNAc-modified proteins provides the sequence and possible residue necessary for interaction. The responsible residue will be confirmed by mutation to abolish the interaction. Another issue of the cell-based crosslinking method is that low

yields of crosslinking material are generated. This yield leads to difficulties to identify the binding partners through proteomics analysis. To transcend this issue, a cell free O-GlcNDAz crosslinking method using large mg scale peptide crosslinking that may resolve this issue. Also, in personal communications with lab members of Dr. Kohler's group, they observed it was crucial to employ stringent washing methods for a more narrowly focused candidate list generated from proteomics. This challenge can be addressed by biotin-tagging the peptides, allowing for hasher washing methods. Lastly, the Kohler lab has previously reported O-GlcNDAz-ylation is a static modification, and likely to affect the dynamics of O-GlcNAc-ylation. This condition may generate false positive candidates of O-GlcNAc-modified proteins. To restore normal dynamics and improve the cell free crosslinking method, OGA will be mutated to accommodate the diazirine of GlcNDAz. These improvements will provide for the development of a cell free O-GlcNDAz crosslinking method to identify the binding partners of O-GlcNAc-modified proteins.

SUMMARY

The function of O-GlcNAc remains elusive. Employment of GlcNDAz-based crosslinking may improve our understanding of the function of this PTM and expand the field. The Kohler lab has previously reported the use of O-GlcNDAz to identify nuclear transport factors interacting with NUP153. In this thesis, I attempted to employ the cell-based O-GlcNDAz approach to identify

the binding partners of NUP98 and leukemogenic NUP98 fusions discussed in **chapter 2**. However, several challenges became apparent.

My concentration was diverted to designing a complementary cell free methodology for O-GlcNDAz crosslinking. I identified an O-GlcNAc hydrolase mutant that released GlcNDAz, which is discussed in **chapter 3**. This OGA mutant can benefit by hydrolyzing the GlcNDAz crosslinked to the binding partners before trypsin digest. This step may achieve a more specific candidate list. Additionally, I examined an O-GlcNAc transferase mutant that preferred transfer of GlcNDAz onto substrates, which is discussed in **chapter 4**. This OGT mutant can increase crosslinking material to generate high yields. Finally, in efforts to create a versatile technology and understanding in whether O-GlcNAc assists in protein interactions, I initiated the development of a chemo-enzymatic synthesis of O-GlcNAc-modified biotin-tagged peptides, (discussed in **chapter 5**). The peptides will be designed based on O-GlcNAc-modified proteins and will be exploited in crosslinking experiments. Future work will identify the binding partners associated with O-GlcNAc-modified peptides using the complementary cell free approach. The candidates would be confirmed through the cell-based assay discussed earlier, and studied for their functional importance.

CHAPTER TWO:

Discovery of O-GlcNAc modifications on NUP98 and leukemogenic NUP fusions

INTRODUCTION

The presence of a nucleus distinguishes eukaryotic cells from prokaryotic cells. The nuclear envelope is a double membrane that maintains genetic material in the nucleus. Each nuclear envelope contains thousands of portals known as nuclear pore complexes (NPCs). To move between the nucleus and cytoplasm, generally molecules traverse the nuclear membrane through the NPC. Small molecules including ions, sugars, water and proteins smaller than 40 kDa freely diffuse through the pore of the NPC, while larger cargos are actively transported through the NPC by escorted proteins known as karyopherins. The cargos shuttled into the nucleus include inner membrane proteins and all nuclear resident proteins. Proteins exiting the nucleus include RNA-associated molecules (52). Transporting through the NPC is essential for cellular homoeostasis, and inhibition of nuclear transport has dramatic and pleiotropic effects (53, 54).

NPCs are immense macromolecular assemblies. Mammalian NPCs are more than 100 MDa in size and composed of multiple copies of more than 30 different individual proteins known as nucleoporins (NUPs) (55). Each NPC comprises of

eight spokes surrounding a central hole. This central space is about 30 nm in diameter and 50 nm in length. Filaments extend on the cytosolic face to form a basket-like structure on the nuclear side (Fig. 2.1A). The NUPs that form the NPC fall into different classes. Structural NUPs provide the scaffold for generating the pore, as well as the baskets. Linker NUPs form a connection between the structural NUPs and the final set of NUPs, the FG NUPs. The FG NUPs, named for their phenylalanine-glycine repeat regions, line the NPC. The FG regions are hydrophobic and unstructured. The FG regions project into the pore and interact directly with karyopherins as they pass through the NPC. These FG-karyopherin interactions are critical to the process of nuclear transport. In

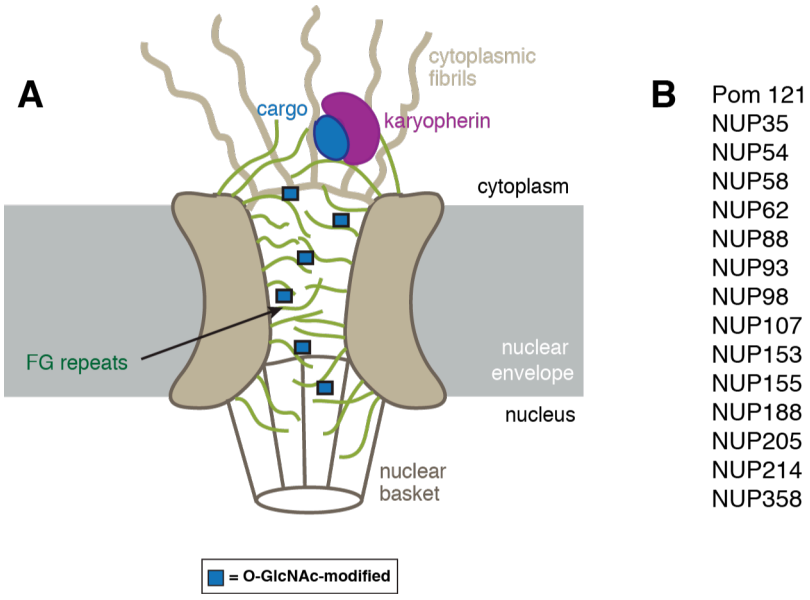

Figure 2.1. NUPs are O-GlcNAc modified. *A,* Several nucleoporin proteins are the scaffolds for creating the nuclear pore complex. Other nucleoporin proteins containing phenylalanine-glycine (FG) repeats facilitate the transport of molecules through the pore. *B*, A list of several NUPs reported to be O-GlcNAc modified.

addition, recent reports indicate that NUPs play roles beyond nuclear trafficking, including regulation of transcription and cell cycle activation (56).

The FG NUP regions that play critical and direct roles in nuclear transport are also heavily glycosylated (Fig. 2.1B) (49, 57). The FG NUP regions contain many serines and threonines that are modified with O-linked N-acetylglucosamine (O-GlcNAc). The functional role of O-GlcNAc on NUPs remains unknown. Hanover and co-workers demonstrated in *C.elegans* that deletion of the O-GlcNAc transferase (OGT), which catalyzes the addition of O-GlcNAc onto proteins, did not affect nuclear trafficking. Instead the OGT knockout in *C. elegans* led to defects in the dauer formation and insulin-like signaling pathway (34). In contrast, Gölrich and co-workers showed that non-glycosylated NUP98, a FG NUP, from *Xenopus*, were incapable of interacting with importins and exportins in a cell free system. Upon glycosylation of NUP98 hydrogels, recombinant importin β interacted with NUP98 at the surface, but did not result in penetration of the protein into the hydrogel (50). Furthermore, Yoneda and co-workers revealed in cervical cancer cells that O-GlcNAc-modified NUP62, a FG NUP, interacted with NUP88. They confirmed this result by deletion of OGT to remove O-GlcNAc-ylation and by incubation with wheat germ agglutinin (WGA) to bind weakly to O-GlcNAc moieties to inhibit the interaction between NUP62 and NUP88 (45). Evidence suggests O-GlcNAc on NUPs plays a role in substrate preference for nuclear trafficking, and perhaps protein-protein interaction.

NUP98 is a unique FG NUP because it is known as a mobile nucleoporin. NUP98 is found on both cytosolic and nuclear sides of the nucleus (58). Interestingly, in several reports of leukemias, chromosomal translocations occur with the FG repeat region of the FG NUPs, including NUP98 and NUP214. In fact there are over 30 translocations reports of NUP98 and NUP214 (51, 59-62). The translocations result in fusion of the FG repeat regions of NUPs to nucleic acid binding domains of other proteins, such as transcription factors. NUP98-HOXA9 and NUP98-DDX10 fusion proteins are the most well characterized leukemogenic fusion proteins (Fig. 2.2) (51, 63, 64).

Literature suggests these fusion proteins are transforming cells to the leukemogenic state (51). The NUP98 leukemogenic fusion proteins are relocated from the NPC to the nuclei (63, 65). The FG regions of the leukemogenic fusions are essential for inducing gene transcription as supported through mutagenesis experiments (64, 66). Additionally, NUP98-HOXA9 interacts with both the transcriptional coactivator CREB-binding

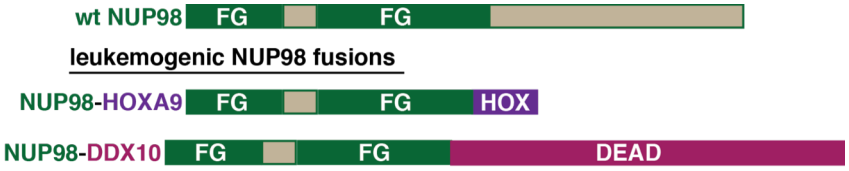

Figure 2.2. NUP98 is involved in leukemia. NUP98 is a nucleoporin that encodes for FG repeat regions in the N-terminus. In several leukemia patients, chromosomal translocations occur with the FG repeat domain of NUP98 fused to other nucleic acid recognizing domains, such as transcription factors. Two well-studied leukemogenic fusions are NUP98-HOXA9 and NUP98-DDX10.

protein (CBP) and the nuclear exportin protein CRM1. For example, Van Deursen and co-workers provided evidence that the FG repeats of NUP98-HOXA9 interact with CBP in cell free and cell-based systems through immunoprecipitation. They demonstrated that FG repeats of NUP98 are transactivating components by employment of a chimera fusion with GAL4 DNA and luciferase activity to measure transcription (66). Additionally, Yaseen and co-workers demonstrated that NUP98 fusion proteins co-immunoprecipitated and co-localized with endogenous CRM1 in the nucleus. They further revealed the NUP98 fusions affected two transcription factors, NFAT and NFκB, which are known to interact with CRM1 (65). The detainment of NFAT and NFκB in the nucleus led to enhanced transcriptional activity, which was measured by luciferase assay.

Furthermore, NUP98-DDX10 fusion protein demonstrated deregulation of gene transcription. The DDX10 portion encodes for conversed helicase motifs that are necessary for transformation of normal cells to the leukemia state (63). Overall, it is speculated in the field that these fusion proteins may act as transcriptional deregulators.

Since both NUP98 and NUP98 fusion proteins contain the FG repeat regions that are known to be heavily O-GlcNAc-modified, I hypothesized that the leukemogenic fusion proteins may be O-GlcNAc modified in the FG regions. NUPs are known to interact with karyopherins in a subcomplex (47, 65). As mentioned previously, O-GlcNAc may mediate protein-protein interactions between NUP62 and

NUP88 (45). I speculate that the O-GlcNAc modification is essential regulation of substrate preference, and may play a functional role as a mediator in protein-protein interaction. O-GlcNAc may be at or near the site of protein-protein interactions.

My initial biological question was to identify the binding partners associated with the O-GlcNAc modifications at the FG regions of NUP98 and the NUP98 leukemogenic fusion proteins. The identification of these binding partners would provide insight into the understanding of the transition that occurs from the normal state to the leukemogenic state of a cell. To identify the binding partners of O-GlcNAc-ylated proteins, the Kohler lab previously reported on a photocrosslinking GlcNAc analog, GlcNDAz. Using this technology, Kohler and co-workers identified the binding partners of the heavily O-GlcNAc-modified nucleoporin protein NUP153, including several karyopherins (47). In this approach, cells were engineered with a mutation in the GlcNAc salvage pathway, a pathway to regenerate O-GlcNAc (Fig. 1.1B), to metabolize. This mutation allowed the pathway to metabolize Ac_3-GlcNDAz-1P(Ac-SATE)$_2$, a cell-permeable GlcNDAz-1P sugar, into the unnatural photocrosslinking sugar to UDP-GlcNDAz. GlcNDAz was transferred onto O-GlcNAc-modified proteins with OGT. Ideally, this cell-based O-GlcNDAz crosslinking method would have assisted in the identification of the binding partners of NUP98 and leukemogenic fusion proteins.

In summary, the identification of the binding partners of NUP98 and NUP98 fusions proteins may elucidate the mechanism of transformation of the cells to a

leukemogenic state. Here I described my efforts to apply the cell-based O-GlcNDAz crosslinking method to identify binding interactions of both NUP98 and NUP98 fusion proteins. I demonstrated O-GlcNDAz-dependent crosslinking of both NUP98 and NUP98 fusions. This data suggested that NUP98 and the NUP98 fusions are O-GlcNAc modified and interact with binding partners through the O-GlcNAc modification. While robust crosslinking was observed, attempts to identify the binding partners were ineffective, due to challenges associated with the isolation of crosslinked material. However, these findings provided me with a key motivation for the development of a complementary cell free O-GlcNDAz crosslinking method, as described in subsequent chapters.

EXPERIMENTAL PROCEDURES

Cell lines and cell culture

HeLa cells, obtained from the ATCC, were cultured in DMEM, 10 % fetal bovine serum (FBS), and 1 % penicillin/streptomycin. K562 cells were obtained from ATCC and cultured in DMEM, 10 % FBS, and 1 % penicillin/ streptomycin for routine culturing. For crosslinking experiments, cells were cultured in low glucose DMEM (Invitrogen, 11885-084), 10% FBS, and 1% penicillin/streptomycin. Hek293T cells were obtained from ATCC and cultured in DMEM, 10% FBS, and 1% penicillin/streptomycin.

To produce cells stable expressing UAP1(F383G), K562 cells were infected with the pSIN4-UAP1 (F383G) lentivirus (47). Briefly, 2.0 x 10^5 K562 cells in a 1 mL volume of fresh media were mixed with 1 mL of the pSIN4-UAP1 (F383G) lentivirus. Cells were centrifuged at 514*g*, 32 °C for 30 min. After gentle mixing, centrifugation was repeated at 514*g*, 32 °C for 30 min. After centrifugation, media was removed by aspiration and cells were suspended in 2 mL of fresh media and transferred to a well of a 6 well plate. Selection with 1.5 μg/mL puromycin for cells expressing UAP1(F383G) was performed for 2 weeks.

Production of UDP-GlcNDAz in UAP1(F383G)-expressing K562 cells was analyzed by high performance anion exchange chromatography (HPAEC) as described (47). Briefly, K562 cells were cultured with or without Ac_3-GlcNDAz-1P(Ac-SATE)$_2$ (>95 % purity by ^{1}H NMR) for 6 h. Cells were harvested by centrifugation, washed once with PBS, and pelleted. Lysing cells released metabolites with 80 % "super cold" methanol on dry ice. Cellular debris was removed by centrifugation at 20,817*g*, 4 °C for 10 min. The supernatant containing metabolites was transferred to a clean vessel and evaporated under vacuum. The metabolite pellet was suspended in 40 mM sodium phosphate buffer (pH 7.0), and analyzed by high performance anion exchange chromatography (HPAEC) using conditions reported previously (47).

Stable cell lines expressing NUP98 or NUP98 fusion

Retrovirus plasmids encoding NUP98 and NUP98 fusions (NUP98-HOXA9 and NUP98-DDX10) were gifted to us (63, 64). Hek293T cells were cultured for

expression of virus particles. Cells engineered to produced UDP-GlcNDAz (HeLa and K562) were infected with MSCV-HA-NUP98, MSCV-HA-NUP98-HOXA9, or MSCV-HA-NUP98-DDX10 retroviruses to create three stable cells lines expressing the tagged NUP or NUP fusion. One million Hek293T cells were plated, and allowed to grow until 30 – 40% confluent. Retrovirus plasmid was incubated at room temperature for 15 minutes with Lipofectamine LTX at a ratio of 3:1 with pCL-ampho retrovirus packaging plasmid. The medium was aspirated. Fresh Opti-MEM media was added to the plate slowly, so as not to perturb the adherence of the cells from the plate. The mixture of retrovirus and transfection reagent was added drop wisely. Cells were incubated with the transfection material for ~20 hours. The medium was aspirated, and warm 10 mL of HI-BSA medium (DMEM, 5% FBS, 3.2 g BSA; to make 500 mL) medium was added. Cells incubated for 24 hours in the HI-BSA medium. The medium was removed and aliquot for usage of infection to generate stable cell lines.

Two thousand cervical cancer cells (HeLa) expressing the UAP1(F383G) mutant were plated in one well of a six-well plate per plasmid. Cells were incubated with 1 mL MSCV-HA-NUP98 or 1:1 MSCV-HA-NUP98-HOXA9/ MSCV-HA-NUP98-DDX10 (600 μL virus + 600 μL medium). After 24 Twenty-four hours later, medium was removed and fresh HeLa medium was added. Cells were sorted due to their expression of GFP for selection of the NUP98 or NUP98 fusion proteins on the MSCV plasmids. Leukemia (K562) cells expressing UAP1(F383G) were infected

using a protocol described previously for the UAP1(F383G) infection, but using the same dilutions as done for the HeLa cells expressing HA-NUP98 or HA-NUP98 fusions. Leukemia cells were also sorted for selection of cells only expressing the NUP98 or NUP98 fusion proteins.

A construct encoding the NUP98 gene tagged with myc in the pALTER-MAX vector was gifted to me from Beatriz Fontoura, UT Southwestern (13). The avi-tag nucleotide sequence (forward primer 5'-ATGTCCGGCCTGAACGACATCTTCGAGGCTCAGAAAATCGAATGGCACGAA- 3', and reverse primer 3'-TTCGTGCCATTCGATTTTCTGAGCCTCGAAGATGTCGTTCAGGCCGGACAT- 5') was generated during sub-cloning mycNUP98 gene into a pRRL lentivirus vector that contains GFP as a selection marker. The production of lentivirus was conducted in Hek293T cells, in a 3 to 1 volume ratio of X-treme gene 9 (Roche, cat. no. 06365787001) to DNA with OPTI-MEM. Lentivirus was packaged with the following vectors psPAX2 and pMD2G. Lentivirus was constructed in the same manner as the previously described protocol for retrovirus. Three mLs of lentivirus was tittered to 150,000 HeLa cells expressing UAP1(F383G). The media was aspirated and fresh HeLa media was added. Cells were grown until 20 million cells were collected and submitted for flow cytometry to sort using the GFP expression.

O-GlcNDAz crosslinking in leukemia cells

One million K562 cells stably expressing UAP1(F383G) were seeded into 10 cm plates to confirm O-GlcNDAz crosslinking. Forty-eight hours and 72 hours after seeding, cells were cultured with low glucose media containing 20 μL of 50 mM Ac_3-GlcNDAz-1P(Ac-SATE)$_2$ in DMSO, to achieve a final concentration of 100 μM Ac_3-GlcNDAz-1P(Ac-SATE)$_2$, or 20 μL DMSO alone. After 24 h of final sugar addition, cells were harvested by centrifugation and suspended in PBS before irradiation with 365 nm light for 15 min. Cells were lysed in 100 μL RIPA buffer (50 mM TrisHCl pH 8.0, 150 mM NaCl, 1.0% NP-40, 0.5 % sodium deoxycholate, 2.0 mM EDTA, and 1 mM DTT) with protease inhibitors (1 mM PMSF and protease inhibitor cocktail (Sigma, S8830)) per 2 million cells, then centrifuged at 20,817*g,* for 10 min at 4 °C to remove insoluble material. Lysates were separated by 5 % SDS-PAGE, transferred to nitrocellulose and probed for NUP153 (SA-1 Abcam, ab96462).

In cell crosslinking of NUP98 and NUP98 fusions

One million HeLa/K562 cells stably expressing UAP1(F383G) and indicated NUP plasmid (HA-NUP98, HA-NUP98-HOXA9, or HA-NUP98-DDX10) were seeded into 10 cm plates. Forty-eight hours and 72 hours after seeding, cells were cultured with low glucose media containing 20 μL of 50 mM Ac_3-GlcNDAz-1P(Ac-SATE)$_2$ in DMSO, to achieve a final concentration of 100 μM Ac_3-GlcNDAz-1P(Ac-SATE)$_2$, or 20 μL DMSO alone. After 24 h of final sugar addition, cells were harvested by

centrifugation and suspended in phosphate-buffered saline (PBS) before irradiation with 365 nm light for 15 min. Cells were lysed in 100 μL RIPA buffer per 2 million cells, then centrifuged at 20,817*g,* for 10 min at 4 °C to remove insoluble material. Lysates were separated by 5% SDS-PAGE, transferred to nitrocellulose and probed for anti-HA (Roche, 11583816001).

Purification of NUP98 and NUP98 fusions

Immunoprecipitation was performed with leukemia expressing HA-NUP98-HOXA9 cell culture conditions; 200 μg of total protein lysate was incubated with 10 μL of anti-HA antibody (0.5, 1, or 2 μg), as above, in total volume of 200 μL for 14 h at 4 °C. Samples were incubated with 25 μL Protein G Sepharose ® Fast Flow (Sigma Aldrich, P3296) for 3 h at 4 °C. The beads were washed three times with RIPA buffer without protease inhibitors. Proteins were eluted by addition of 2X loading dye (100 mM TrisHCl pH 6.8, 25 mM EDTA, 0.04 g/mL SDS, 0.4 mg/mL bromophenol blue) containing 20 mM DTT, and incubation for 10 min at 90 °C. Samples were analyzed by immunoblot using the HA antibody, as above.

An alternative approach to purify the NUP98 and leukemogenic fusions was conducted with avi-mycNUP98 proteins in cells (67). The avi tag was cloned into the pRRL-mycNUP98 lentivirus vector as described above. Five million cells expressing avi-mycNUP98 were harvested and lysed in RIPA buffer or the recommended avidity buffer provided. Cell free biotinylation reaction was conducted with 60 μg total protein including avi-mycNUP98 and 2 μg avi-MBP as a control using the Avidity

biotinylation kit (cat.no. BirA500). The reaction was incubated at 30 °C for 2 hours or 4 hours. The samples were run on a gel and immunobloted using biotin-HRP (Sigma, A0185) to detect biotinylation.

Localization of NUPs

500,000 HeLa cells were plated on glass coverslips in a 24-well plate in regular medium. Cells were allowed to grow to 80% confluence. Medium was aspirated and cells were briefly washed three times in PBS. Cells were fixed with pre-warmed 4% paraformaldehyde for 30 minutes. Following fixture, cells were briefly washed three times in PBS. Saponin was utilized to permeabilize the cells for 15 minutes. The cells were incubated with a non-specific binding blocking buffer (PBS, BSA, and lysine at pH 7.4) for 30 minutes. Cells were incubated with mAB414 (Covance, MMS-120R) in a 1:100 dilution 16 h at 4 °C. Antibody was removed and cells were washed briefly with blocking buffer (six times). HeLa cells were incubated with Alexa 488 and phalloidin antibody (Life technologies, A34055) for 1 h. Cells were washed six times in PBS and mounted on slides (Life technologies, P36966).

RESULTS

Establishing GlcNDAz crosslinking in leukemia cells

The O-GlcNDAz technology was previously established in a cervical cancer cell line (HeLa). For the purpose of this study, I desired to use the technology in a

leukemia cell line, K562, for a more biologically relevant assessment. To achieve this, leukemia cells were infected to stably express a mutant form of the enzyme

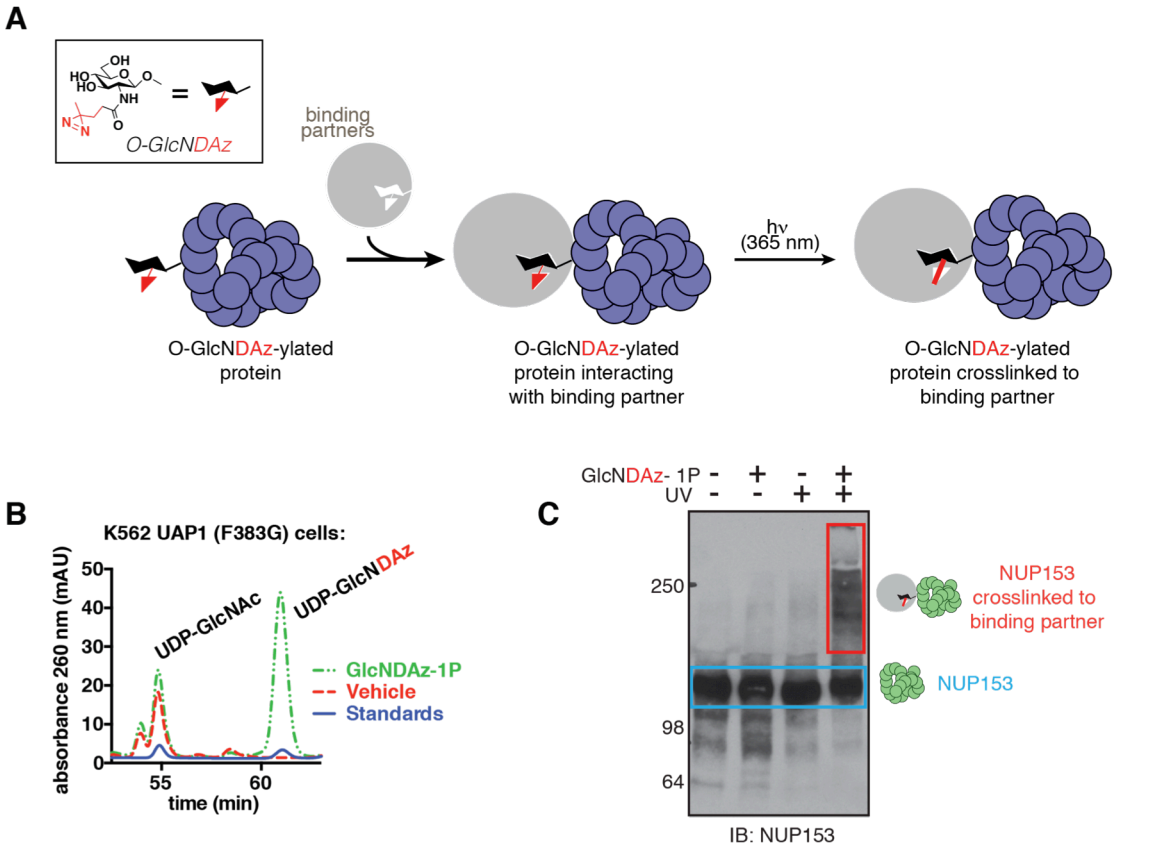

Figure 2.3. Application of O-GlcNDAz-ylation to identify binding partners. *A,* Photocrosslinking GlcNAc, GlcNDAz, is substituted onto O-GlcNAc-modified proteins. The diazirine is small and does not perturb natural interactions. Through UV irradiation, we create a snapshot of the interactions occurring at that moment. The irradiation leads to an irreversible bond generated between O-GlcNDAz-ylated protein and the binding partner. *B,* Leukemia cells (K562) engineered to express UAP1(F383G) were cultured with vehicle or cell-permeable GlcNDAz-1P analog. Through high performance anion exchange column, HPAEC, we detected UDP-GlcNDAz production in the sugar treated cells. *C*, Crosslinking analysis of the cells from 1B was immunoblotted for a heavily O-GlcNAc-modified nucleoporin protein, NUP153. The three negative controls (no treatment, GlcNDAz-1P treatment, and irradiation treatment) reveal a band of the appropriate molecular weight of NUP153 (150 kDa, blue box). In the sample treated with GlcNDAz-1P and irradiation, bands at higher molecular weight were detected, which correspond to crosslinking of NUP153 to unknown binding partners (red box).

UDP-N-acetylglucosamine pyrophosphorylase 1, UAP1, with a phenylalanine to glycine mutation at position 383 (68). Cells expressing the UAP1(F383G) mutant were cultured with or without the cell-permeable GlcNDAz-1P analog, and analyzed by HPAEC for production of UDP-GlcNDAz, which is the donor nucleotide sugar necessary for O-GlcNDAz-ylation. The leukemia cells expressing the mutated enzyme and cultured with the cell-permeable GlcNDAz-1P demonstrated a robust production of UDP-GlcNDAz (Fig. 2.3B). After confirmation of the nucleotide sugar production, cells were exploited for crosslinking ability. The cells were cultured as previously described (47). Crosslinking was analyzed by immunoblot for heavily O-GlcNAc-modified NUP153 (Fig. 2.3C), because the Kohler lab previously reported karyopherins interacting to NUP153. The result confirmed bands at the appropriate molecular weight of 150 kDa for NUP153 in all of the conditions. Higher molecular weight bands were detected in cells treated with both the GlcNDAz-1P and UV irradiation, which corresponded to crosslinking of NUP153 to unknown binding partners (Fig. 2.3C, red box). Overall, leukemia cells exhibited robust metabolic incorporation of the cell-permeable GlcNDAz-1P sugar and labeling of the heavily O-GlcNAc-modified protein, nucleoporin NUP153.

NUP98 and Leukemogenic NUP98 fusions are O-GlcNDAz-modified

Literature reported nucleoporin NUP98 being modified with O-GlcNAc-ylation, but the site of O-GlcNAc-ylation is unknown. Next, I examined the full-length protein of NUP98 to identify if O-GlcNAc on NUP98 resided at the interface of protein-protein

interactions. For this experiment, I used the cell-based O-GlcNDAz crosslinking assay in two cell lines. Cervical cancer (HeLa) and leukemia (K562) cells overexpressing UAP1(F383G) were infected with a HA-tagged form of wild-type NUP98 in order to examine the crosslinking of NUP98 since a good antibody of NUP98 is unavailable. Crosslinking experiments were conducted as previously described with the addition of cell-permeable GlcNDAz-1P and exposure to UV irradiation. Immunoblots of NUP98 revealed that a covalent bond was exhibited between the tagged form of NUP98 and unknown binding partners (Fig. 2.4, green boxes) in both leukemia and cervical cancer cell lines. This evidence suggested the confirmation that NUP98 is O-GlcNAc-modified. Also, the O-GlcNAc may be utilized to interact to neighboring molecules (Fig. 2.4, green boxes). In addition, the cervical cancer and leukemia cells have distinct crosslinking patterns from each other. I speculate the

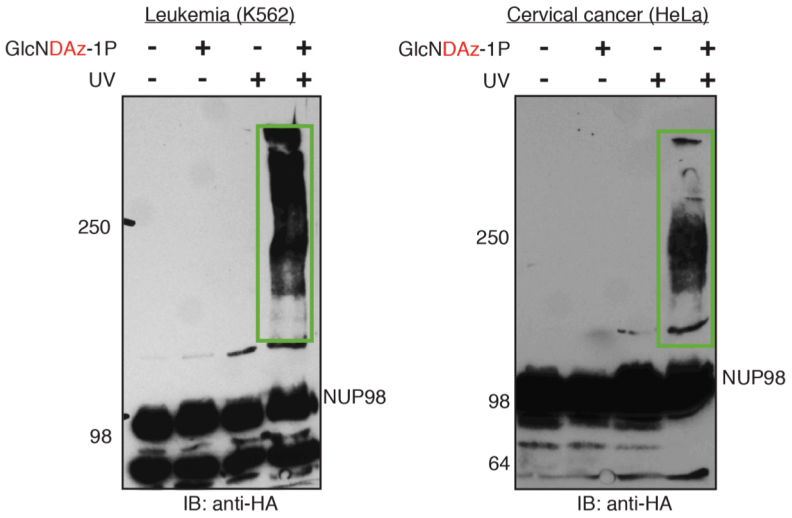

Figure 2.4. NUP98 exhibits O-GlcNDAz-ylation and crosslinking. UDP-GlcNDAz engineered cells (K562 and HeLa) were infected to express HA-NUP98. Cells were treated with vehicle or cell-permeable GlcNDAz-1P, and then placed under UV irradiation or in ice. Immunoblot of the samples reveal crosslinking of NUP98 modified by GlcNDAz to unknown proteins (green box).

different crosslinking patterns may be due to NUP98 interacting with different binding partners in the two different cell lines.

Next, I tested the cell-based O-GlcNDAz crosslinking method on the

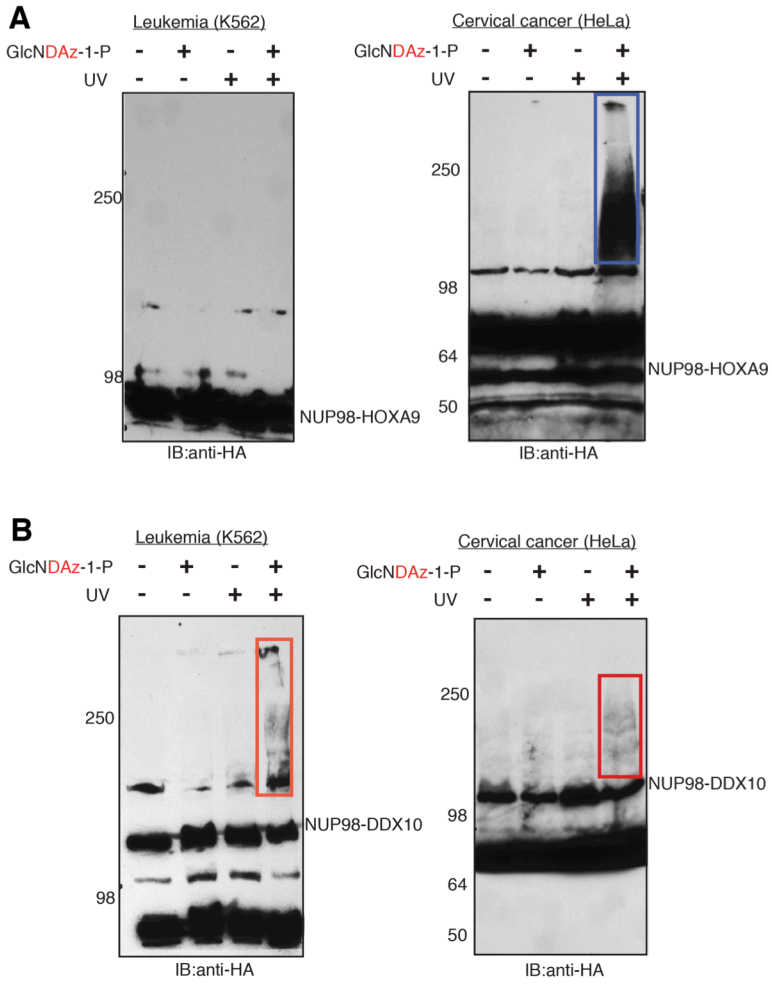

Figure 2.5. NUP98 fusions exhibit O-GlcNDAz-ylation and crosslinking. *A*, The UDP-GlcNDAz engineered cells (K562 and HeLa) were infected with HA-NUP98-HOXA9 (leukemogenic fusion) and treated similarly as the HA-NUP98. Immunoblot for the HA-tagged protein reveals crosslinking of the cervical cancer cells (blue box). *B*, Cells engineered to synthesize UDP-GlcNDAz and produce HA-NUP98-DDX10 were treated similarly as in *A*. Upon treatment with sugar and irradiation, distinct crosslinking is revealed in the K562 cells compared to the HeLa cells for NUP98-DDX10 (red boxes).

leukemogenic fusion proteins to gain insight into their mechanisms of transformation. I established cervical cancer (HeLa) and leukemia (K562) cells expressing a HA-tagged-form of the NUP98-HOXA9 or NUP98-DDX10 fusion because there is no good antibody available for the fusion proteins. As mentioned previously, NUP98-HOXA9 and NUP98-DDX10 are two well-studied leukemogenic NUP98 fusion proteins (63, 64). However no evidence suggest the leukemogenic NUP98 fusions are O-GlcNAc-modified. In the no treatment, GlcNDAz-1P treatment only, and UV irradiation treated only samples, bands at the appropriate molecular weight corresponding to the NUP98-HOXA9 (62 kDa) and NUP98-DDX10 (120 kDa) fusions were detected. The samples treated with both GlcNDAz-1P and UV irradiation, revealed higher molecular weight bands, which I hypothesized was the leukemogenic fusion proteins crosslinking to unknown binding partners (Fig. 2.5A, blue boxes and 2.5B, red boxes). This was surprised because O-GlcNAc modification has not been reported on the NUP98 fusions. These results suggested a novel finding of the NUP98 fusion proteins to be O-GlcNAc modified. Also, O-GlcNAc on the fusion proteins may reside at the interface of protein-protein interactions responsible for the leukemogenic activity.

Purification of NUP98 and Leukemogenic NUP98 fusions to identify binding partners

In effort to determine the unknown binding partners of NUP98 and NUP98 fusion proteins, I enriched for these crosslinked complexes using an antibody that recognizes the HA tag. The lysate expressing the HA-tagged NUP98-HOXA9 was

incubated with different amounts of antibody for immunoprecipitation. The purified proteins were boiled off and subjected to immunoblot with an HA antibody to recognize the tagged NUP98 fusion protein. My results demonstrated the different amounts of antibody used were not efficient to pull down HA-NUP98-HOXA9 (Fig. 2.6A). The antibody may not recognize the HA-tagged fusion protein due to steric. This approach to perform the crosslinked complexes was aborted as the cervical cancer cells express hemagglutinin which is recognized by the HA antibody raised against hemagglutinin (Fig. 2.4 and 2.5).

An alternative approach to purify the crosslinked complexes was to utilize the avi-tag technology. The avi-tag is a small amino acid sequence, GLNDIFEAQKIEWHE, which can be biotinylated through the *E.coli* biotin ligase, BirA. Cells expressing avi-tagged NUP98 were lysed in RIPA or the lysing buffer

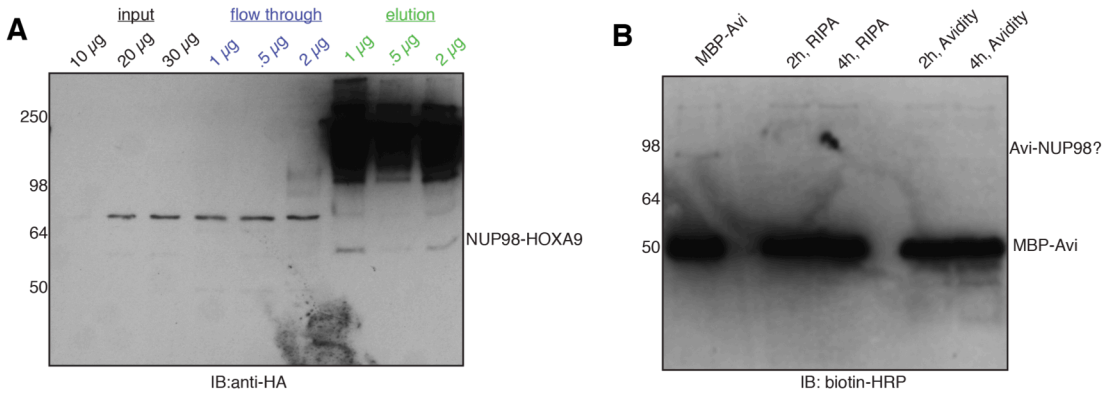

Figure 2.6. Attempts to purify NUP98 fusions and NUP98 for identification of binding partners. *A,* Leukemia cells expressing HA-NUP98-HOXA9 were immunopurified with HA antibody in 1 µg, 0.5 µg, and 2 µg incubation. Immunoblot of the pull-down proteins reveals little fusion protein extracted. *B,* Cells expressing avi-mycNUP98 were incubated with 2 µg avi-MBP in the biotinylation reaction (Avidity). Immunoblot reveals biotinylation of the control (avi-MBP), but a band at the molecular weight of NUP98 was not detected.

provided by the biotinylation kit. The addition of a control avi-tag protein, avi-MBP, was introduced into the lysate as a control. The proteins were incubated with the BirA ligase and biotin for different durations of time. Detection of the biotinylation was revealed through immunoblot. I detected biotinylation for the MBP control but no biotinylation of the tagged NUP98 (Fig. 2.6B). In personal communication with an Avidity scientist, the biotinylation reacts best with avi tags at the C-terminus of the protein. However, tagging the C-terminus of NUP98 affects its function. Alternative methods to purify the crosslinked complexes, which will be presented in the following section, will need to be conducted to identify these unknown binding partners of NUP98 and leukemogenic NUP98 fusions.

DISCUSSION

O-GlcNAc-ylation may be involved in the functional roles of both NUP98 and NUP98 fusion proteins. Evidence supports the FG repeat regions of the leukemogenic fusions are important for transformation of the cell and transcriptional activity. My hypothesis was that O-GlcNAc on the FG regions assist in the interactions for the leukemogenic fusions with other proteins.

GlcNDAz technology to identify binding partners of NUP98 and NUP98 fusions

Through the cell-based O-GlcNDAz crosslinking method, I confirmed that NUP98 is O-GlcNAc-modified. With the O-GlcNDAz modification, NUP98 crosslinked to unknown binding partners. This result suggested that O-GlcNAc was near or at the

site of interaction between NUP98 and its binding partners. However, I am unsure of the exact location of the O-GlcNAc site. Also based on the cell-based O-GlcNDAz crosslinking method, the NUP98 fusions, NUP98-HOXA9 and NUP98-DDX10, were likely O-GlcNAc-modified. Additionally, these NUP98 fusions demonstrated interaction to other molecules with O-GlcNAc at or near the interacting site (Fig. 2.5). The crosslinking data suggested that O-GlcNDAz resided near the site of protein-protein interactions of NUP98 and NUP98 fusions in two cell lines. Yet, the exact location of O-GlcNAc on the NUP98 fusions was undetermined. Two challenges were presented in the work. The first challenge was the failed attempts to identify the binding partners. The second challenge was the uncertainty of the O-GlcNAc site on NUP98 and NUP98 fusions.

Unsuccessful attempts to purify crosslinked complexes of NUP98 and leukemogenic NUP98 fusions

The efforts to identify these unknown candidates via immunoprecipitation of NUP98 and NUP98 fusion proteins were ineffective. I tried to immunoprecipitate the HA-tag form of the NUP98-HOXA9 (Fig. 2.6A). However, the antibody did not recognize the fusion protein and resulted in no detection of the protein. The HA-tag on the leukemogenic fusion protein may be unavailable for interaction with the antibody. While the Kohler lab identified NUP153 to interact with several karyopherins, (47) they did so by utilizing an antibody that has high affinity to recognize several NUPs including NUP153 without using harsh stringent washing

methods. However in personal communications with lab members from Dr. Kohler's lab, it was crucially identified to use stringent washes to obtain a more specific candidate list for proteins like cholera toxin, (69) if the antibody that recognizes the protein of interest does not have high affinity. As an alternative purification method, direct biotinylation of NUP98 and NUP98 fusions could allow for the use of a stringent washing method. This washing method will achieve an improved candidate list because of the high affinity interaction between biotin and streptavidin.

An alternative method to pull-down the NUP98 and leukemogenic NUP98 fusions used chemo-enzymatic approach for biotinylation. Through manipulation by polymerase chain reaction (PCR), the sequence that encodes for a substrate of biotinylation was added onto the gene of interest. I achieved the addition of the avi-tag sequence (GLNDIFEAQKIEWHE), at the N-terminus of a myc-tagged NUP98. The chemo-enzymatic biotinylation reaction demonstrated success in biotinylation of the avi-tag MBP protein, but not of the avi-myc NUP98 (Fig. 2.6B). The avi-tag on the MBP protein was at the C-terminus, while the NUP98 protein was tagged at the N-terminus. The C-terminus of NUP98 cannot be tagged as it will affect the protein's function. The biotinylation of avi-tag is known to be more successful with tags at the C-terminus. One reason why the reaction was ineffective may be that the avi-tag is inaccessible to the BirA enzyme as the tag could be buried in the NUP98 protein. This technology was examined with another nucleoporin NUP62, which avi tag was encoded at the N-terminus, to test if the protein difficulty was due to NUP98 alone.

However, these reactions were unproductive as well (data not shown). The ineffective approach may be due to the FG regions of the nucleporins, which are unstructured and hydrophobic, and prevent the success of the biotinylation.

Future work to attach biotin onto the NUP98 and leukemogenic NUP98 fusions is with the SNAP tag. SNAP-biotin is a protein-based chemical approach. The technology utilizes the human O^6-alkylguanine-DNA-alkyltransferase (hAGT) with the SNAP-tag substrates, which are derivatives of benzylpurines and benzylchloropyrmidines. The reaction specifically and covalently labels the fused protein with biotin or fluorophores.

I attempted to generate a mammalian cell line to express the SNAP-tag labeled NUP98 then for the NUP98 fusions. Unfortunately I was unable to subclone the NUP98 gene into the mammalian expression vector with the SNAP-tag. These experiments could be revisited.

Once the vectors are produced, the SNAP technology may be more successful because the protein may help with folding compared to the avi- tag, which is a small peptide. One potential issue with the SNAP tag is that it is a 20 kDa protein and may affect the structure of the NPC and NUP98 in living cells.

Other methods to study the functional importance of O-GlcNAc on NUP98 and leukemogenic NUP98 fusions

A better understanding of the mechanism of transformation for the leukemogenic fusions could advance the field. My efforts demonstrated that NUP98

and the leukemogenic fusions were likely O-GlcNAc modified. I believe that O-

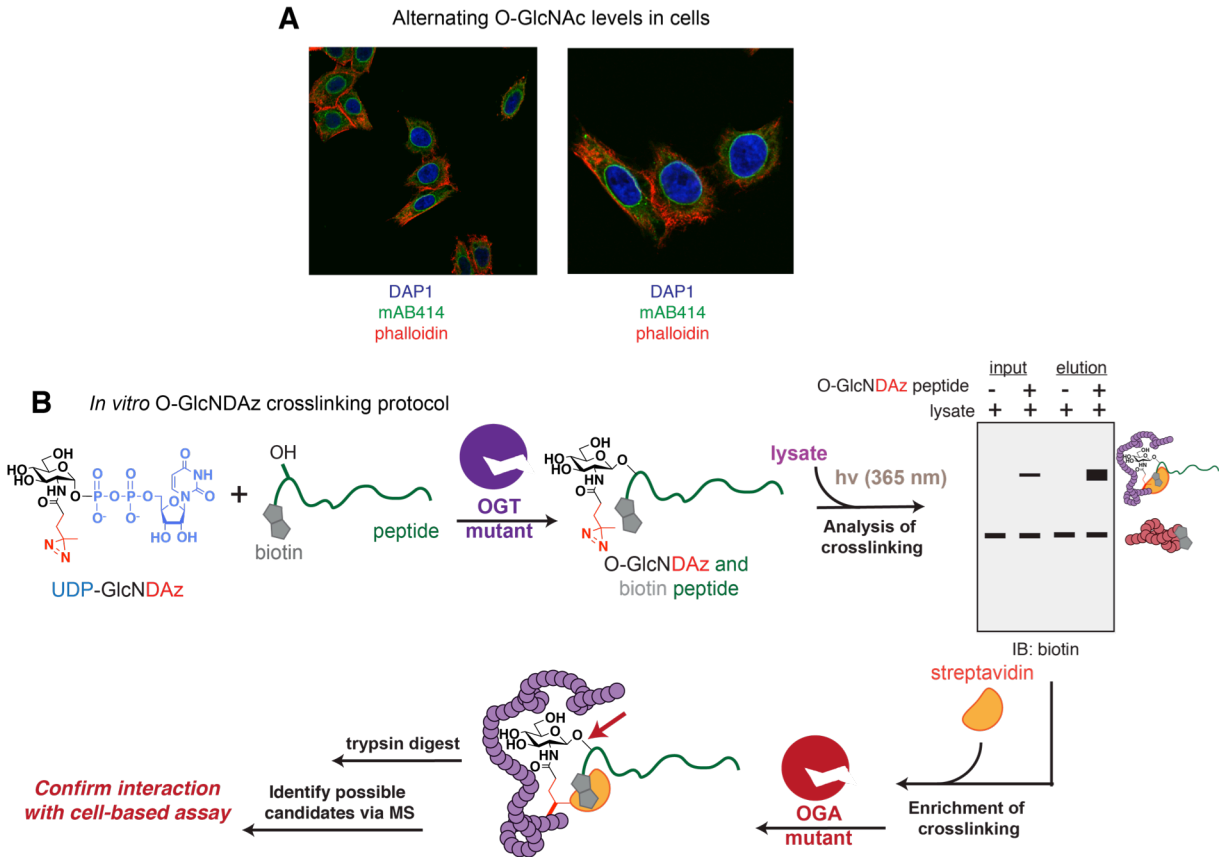

Figure 2.7. Understanding of the functional role of O-GlcNAc on NUP98 and leukemogenic NUP98 fusions. *A,* Effects of NUPs localization under no treatment HeLa cells. mAB414, an antibody that recognizes several NUPs including NUP153, NUP98, and NUP62, is a green fluorescent antibody. DAP1 detects the nucleus shown in blue. Phalloidin, represented in red fluorescent, is a toxin that detects F-actin. Experiments were performed in duplicate with a single representative trial shown *B,* Scheme of a cell free approach to identify the binding of O-GlcNAc-modified FG repeat region of NUP98. Biotin peptides are glycosylated with O-GlcNDAz using UDP-GlcNDAz and OGT mutant. This modified peptide is incubated with lysate and crosslinked. The confirmation of crosslinking is be conducted by immunoblot for biotin. Binding partners involved in the interaction with NUP98 are enriched using streptavidin purification. A band only in the glycosylated treated sample confirms crosslinking via immunoblot of biotin to recognize the biotin-peptide. After enrichment of crosslinking complexes, the material is incubated with an OGA mutant to hydrolyze O-GlcNDAz crosslinked with binding partners (red arrow indicates site of hydrolysis). The identification of binding partners are analyzed through proteomics and verified through the cell-based O-GlcNDAz crosslinking assay and mutagenesis.

GlcNAc may play an important role in nuclear trafficking not only on NUP98, but also on the leukemogenic fusions in transcriptional regulation. An alterative approach to study O-GlcNAc on NUP98 and leukemogenic NUP98 fusions is to alter O-GlcNAc levels and study the localization and function of NUP98 and the leukemogenic NUP98 fusions. Several compounds can alter O-GlcNAc levels: free GlcNAc, inhibitors of O-GlcNAc transferase (OGT), and inhibitors of O-GlcNAc hydrolase (OGA). The localization of the proteins can be determined by immunofluorescence (Fig. 2.7A). Nuclear trafficking and gene transcription can be studied to determine if altered O-GlcNAc levels in leukemia cells can vary function. I believe studying the nuclear trafficking and gene transcription under altered O-GlcNAc levels may also provide insight into whether O-GlcNAc plays an important role in the transformation of these cells.

I hypothesize the wild-type NUP98 and leukemogenic NUP98 fusion proteins may interact with similar or different proteins within each of these cell lines. There could be similar binding partners of NUP98 and NUP98 fusions since both proteins encode the NUP98 FG region structures. However, their distinct different locations may provide a different set of binding partners that are more essential for the transformation of the cell.

A Complementary Approach for the O-GlcNDAz crosslinking technology

Identification of the binding partners of NUP98 and NUP98 fusion proteins that are mediated by O-GlcNAc at the FG region is crucial to achieve a better

understanding of the mechanism of action that leads to the leukemogenic state of the cell. My crosslinking data suggested NUP98 and leukemogenic fusions were O-GlcNAc-modified and interact with binding partners. However several challenges became apparent while using the cell-based O-GlcNDAz crosslinking assay. These challenges were the low yield of crosslinking material, weak washing methods, and unknown O-GlcNAc site responsible for mediating protein-protein interaction. An alternative method to identify the potential binding partners is necessary.

The development of a complementary cell free O-GlcNDAz crosslinking assay can overcome the challenges of the cell-based assay (Fig. 2.7B). In later chapters, I will discuss the development of tools that may be used for the complementary cell free O-GlcNDAz crosslinking method. I will describe a mutant OGA that has the ability to remove O-GlcNDAz (**chapter 3**), and a mutant OGT that increases O-GlcNDAz-ylation (**chapter 4**). The OGA mutant could assist in the development of a more specific candidate list. The OGT mutant could increase crosslinking by increasing O-GlcNDAz to overcome the limit on crosslinking material. Also, I will review on a chemo-enzymatic synthesis protocol to O-GlcNDAz-modified biotin tagged peptides derived from O-GlcNAc-modified proteins (**chapter 5**). This peptide approach would allow the recognition of the peptide sequence and possibly O-GlcNAc site responsible for protein-protein interactions mediated by O-GlcNAc. This cell free O-GlcNDAz crosslinking assay will provide a complementary route to identifying binding partners of O-GlcNAc-modified proteins.

ACKNOWLEDGEMENTS

I acknowledge the generous gift of NUP98 and NUP98 fusion plasmids from Dr. Nabeel Yaseen at Washington University, St. Louis. I thank Dr. Beatriz Fontoura at UT Southwestern for plasmids and discussions.

CHAPTER THREE:

Recognition of O-GlcNDAz by a mutant form of human O-GlcNAcase (OGA)

INTRODUCTION

O-linked β-N-acetylglucosamine (O-GlcNAc) is a unique and promiscuous post translational modification (PTM) found in most eukaryotes on a wide range of nuclear and cytoplasmic proteins (6). O-GlcNAc is a dynamic modification, in which the addition of a single GlcNAc residue onto serine and threonine residues of a polypeptide chain is catalyzed by the enzyme O-GlcNAc transferase (OGT) (3) and removed by a single O-GlcNAc hydrolase (OGA) (Fig. 3.1) (70).

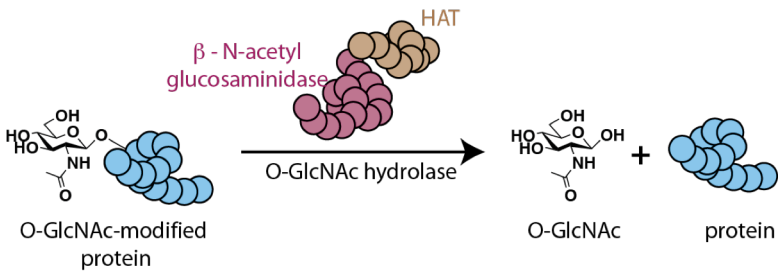

Figure 3.1. OGA is a bifunctional enzyme. O-GlcNAc is hydrolysed from the modified protein with the glucosaminidase domain of the O-GlcNAc hydrolase. This results in free GlcNAc and the unmodified protein.

Like O-GlcNAc, OGA has been linked to human diseases. The OGA gene is also known as MGEA5, and was previously annotated as a putative hyaluronidase. The OGA gene is mapped to the late-onset of Alzheimer's disease locus in humans (10q24.1), and is adjacent to the insulin-degrading

enzyme gene (71, 72). The result of translation for OGA leads to a bifunctional enzyme. The N-terminus of OGA encodes for a glycosidase domain, while the C-terminus is homologous to the histone acetyltransferase (HAT) family of enzymes (73). Human OGA is classified under the category of glycoside hydrolase family GH84 (74).

There are several bacterial OGA homologs of the human OGA. A reported demonstrated that O-GlcNAc-ylation could occur in *Thermobaculum terrenum* (75). These bacterial OGA homologs contain the glycosidase domain. Additionally, there have been reports of the bacterial OGA homologs complexed with OGA inhibitors to gain insight to the structure of human OGA (76, 77).

The levels of human OGA mRNA and protein expression are sensitive to the levels of O-GlcNAc modification in a cell. For example, treatment of cervical cancer and neuroblastoma cells with a highly selective O-GlcNAcase inhibitor, Thiamet G, resulted in increased O-GlcNAc levels, thereby caused OGA protein levels to increase and OGT protein levels to decrease (78). Deregulation of O-GlcNAc disturbed homeostasis of the cell and been demonstrated to be associated with several human diseases, such as cancer, diabetes and Alzheimer's (35-37). In OGA knockout *C. elegans*, OGA production was disrupted and led to deregulation in O-GlcNAc cycling and

metabolism (79). Additionally, Suh and co-workers revealed that a null OGA knockout mouse led to a perinatal lethality (80).

Previously, our lab described O-GlcNDAz, a photocrosslinking O-GlcNAc analog. Using O-GlcNDAz, Seok-Ho Yu revealed covalent crosslinking between NUP153 and specific karyopherins. This evidence suggested that O-GlcNAc residues were intimately associated with the nucleoporin-karyopherin interactions that occur during nuclear transport (47). However, during the course of his studies, he discovered that O-GlcNDAz is not a dynamic modification (47). For example, cell lysates overexpressing OGA were unable to hydrolyze an O-GlcNDAz mimic, *p*-nitrophenol-GlcNDAz. With this discovery, we were concern that accumulation of O-GlcNDAz might be toxic to cells, and adversely impact the utility of our technology.

Our concerns were based on previously observed toxicities attributed to OGA inhibition. For instance, Hart and co-workers studied a non-selective OGA inhibitor *O*-(2-acetamido-2-deoxy-D-glycopyranosylidene)amino-N-phenylcarbamate (PUGNAc) in cervical cancer cells. They revealed treatment with PUGNAc delayed G_2/M phase progression by flow cytometry and resulted in toxicity (81). However, Vocadlo and co-workers examined a more selective OGA inhibitor, 1,2-dideoxy-2′-propyl-α-D-glucopyranoso-[2,1-D]-Δ2′-thiazoline (NButGT), and discovered similar increased levels of O-GlcNAc dependent on dosage and time in adipocytes. However, NButGT treatment did not cause the

same adverse effects that PUGNAc treatment did (82). Hart and co-workers also demonstrated that OGA inhibition by N-acetyl-D-glucosamine-thiazoline (GT) led to increased O-GlcNAc levels and OGA protein levels. However, GT did not perturb OGT expression or location, nor did it prevent cells from progressing through the cell cycle (83). Based on these reports, I examined whether O-GlcNDAz production would be toxic or might affect homeostasis of O-GlcNAc.

In this chapter I demonstrated that recombinant human OGA is unable to remove a small molecule O-GlcNDAz mimic or to remove O-GlcNDAz from a peptide. Accumulation of O-GlcNDAz in cells resulted in no overt harmful effects on cellular function, but likely impacts O-GlcNAc dynamics. A static modification may affect cellular homeostasis. Thus, I developed a strategy to restore normal O-GlcNAc dynamics to cells producing O-GlcNDAz. I produced a model of the active site of human OGA based on homology to bacterial N-acetyl-β-glucosaminidases and identified residues in potential conflict with processing of diazirine-modified substrate. I discovered that mutations at C215 resulted in mutant OGA enzymes that could hydrolytically release GlcNDAz from a model substrate. I found that the C215A mutant of OGA could deglycosylate an O-GlcNDAz-modified peptide, offering an approach to restore normal glycosylation dynamics to cells producing O-GlcNDAz. In addition, the results presented here could be later used to improve a cell free

O-GlcNDAz crosslinking approach. In which OGA(C215A) is employed to specifically release O-GlcNDAz crosslinks, and thereby yield a narrower focused candidate list of binding partners.

EXPERIMENTAL PROCEDURES

Cell culture

Parental HeLa cells were obtained from ATCC. HeLa cells stably expressing UAP1(F383G) cells were described previously (47).

Chemicals

Diazirine-modified 4-nitrophenyl-N-acetylglucosamine (*p*NP-GlcNDAz) and diazirine-modified UDP-GlcNAc (UDP-GlcNDAz) were prepared as described previously (47) and dissolved in DMSO. 4-nitrophenol-N-acetylglucosamine (*p*NP-GlcNAc) was obtained from Sigma (N9376).

Toxicity analysis

Parental HeLa cells or HeLa cells expressing UAP1(F383G) were cultured in DMEM containing 5% FBS and 1% penicillin/streptomycin. Initial conditions were 2 million cells in 10 mL of media in a 10-cm tissue culture plate. After 24 h and 48 h, 100 μM of Ac_3GlcNDAz-1P(Ac-SATE)$_2$ in DMSO or DMSO alone was added to the cell media.(47) For the short time course, cell number and viability were determined at 72 h, using the Invitrogen Countess

automated cell counter according to the manufacturer's instructions. For the long time course, cells were diluted at 72 h to 2 million cells in 10 mL of media (containing Ac_3-GlcNDAz-1P(Ac-SATE)$_2$ in DMSO or DMSO alone) in a 10-cm tissue culture plate. Additional Ac_3-GlcNDAz-1P(Ac-SATE)$_2$ in DMSO or DMSO alone was added at 96 h and 120 h. At 144 h, cell number and viability were determined using the Invitrogen Countess automated cell counter according to the manufacturer's instructions.

Expression and purification of recombinant human OGA

A pBAD/HisA plasmid encoding OGA (84) was used to transform TOP-10 *E. coli*. A single colony was added to a 5-mL starter culture of LB media containing with 50 μg/mL ampicillin and cultured overnight at 37 °C with shaking. The entire starter culture was used to inoculate a 500-mL culture of LB media containing with antibiotics, which was then cultured at 37 °C with shaking until the OD_{600} reached 0.8. The culture was cooled to room temperature, then induced at 0.02% arabinose by culturing overnight at 20 °C with shaking. Bacteria were lysed in 50 mM sodium phosphate buffer pH 7.8, containing 500 mM NaCl, 10 mM imidazole, 0.1% Tween-20, and 0.2 mM PMSF. Sonication was performed at 4 °C for 10 min with alternate cycling between 5 sec of sonication and 8 sec recovery. Insoluble material was removed by centrifugation (15,000*g*, 4 °C, 30 min). For in-lysate assays, OGA

mutant overexpression was assessed by immunoblot using an antibody that recognizes the hexahistidine affinity tag (Sigma Aldrich, H1029). For experiments with purified protein, the supernatant was applied to a Profinia IMAC column. Purification and desalting were performed using the Bio-Rad Profinia system with recommended buffers. Recombinant OGA mutant proteins were prepared by the same method.

Expression and purification of recombinant human OGT

A pET24b plasmid encoding OGT (85) was used to transform BL21(DE3) *E. coli*. A single colony was added to a 50-mL culture of LB media with 50 μg/mL kanamycin, which was incubated at 37 °C overnight without shaking. Eight mL of this culture was used to inoculate a 1-L culture of LB containing ampicillin, which was incubated at 37 °C with shaking until the OD_{600} reached 0.6 - 0.8. After cooling to 20 °C for 1 h, cells were induced by adding 0.5 mM IPTG, then incubated overnight at 20 °C with shaking. After harvesting by centrifugation, cells were lysed in 50 mM TrisHCl pH 7.5 containing 500 mM NaCl, 1% Triton X-100, 20 mM β-mercaptoethanol, 0.3 mg/mL fresh lysozyme, and 1/50 dilution of protease inhibitor cocktail (Sigma, S8830; contains pancrease extract, thermolysin, chymotrypsin, trypsin, bestain HCl, pepstatin A and papain). Lysates were incubated on ice for 30 min, then sonicated at 4 °C for 5 min, alternating 10 sec sonication with 30 sec recovery.

After removing insoluble material by centrifugation (15,000*g*, 4 °C, 30 min), the supernatant was applied to Ni-NTA resin. Resin was washed with 300 mL of 20 mM TrisHCl buffer pH 7.5, containing 500 mM NaCl, 20 mM β-mercaptoethanol, 0.1 mM PMSF. Then, 50 mL of the same buffer, with the addition of 200 mM imidazole, was used to elute OGT from the resin.

Kinetic analysis

Recombinant OGA or OGA mutant was incubated in 50 mM sodium cacodylate buffer pH 6.5, containing 5 mg/mL BSA and 4 mM GalNAc. *p*NP-GlcNAc or *p*NP-GlcNDAz was added to achieve the indicated concentration. For *p*NP-GlcNAc reactions, the enzyme concentration was 50 nM; for *p*NP-GlcNDAz reactions, enzyme concentration was 250 nM. Reactions (100 μL total volume) were performed in a 96-well clear flat bottom plate. Reaction progress was monitored with a Synergy Neo plate reader at 405 nm. Hydrolytic release of *p*-nitrophenol yields an increase in absorbance at 405 nm. Because *p*-nitrophenolate has a stronger absorbance than *p*-nitrophenol, reactions are typically quenched with base to achieve maximal signal. However, I desired a continuous assay measuring enzyme activity at pH 6.5, so we could not perform measurements at a basic pH. Instead, I carefully controlled the pH of all reactions and calibrated values relative to a *p*-nitrophenol standard also measured at pH 6.5, in a manner similar to other reports describing the use of 4-nitrophenol reporters (86-89). Indeed Hayre *et*

al. demonstrated accurate determination of kinetics of sialidase-catalyzed reactions using a *p*-nitrophenol reporter and continuous absorbance measurements at pH 6.0 (87). Although the reaction buffer was at a pH lower than the pK_a of 4-nitrophenol, a substantial increase in absorbance was detectable when OGA activity was present. Changes in absorbance at 405 nm were linear with respect to time over the period between 5 and 35 minutes. These data were used to determine initial rates, which were plotted versus substrate concentration (Fig. 3.3D and 3.3.E). Michaelis-Menten parameters were calculated using GraphPad Prism and are presented in Table 3.1.

Enzymatic synthesis of O-GlcNAc-ylated and O-GlcNDAz-ylated peptides

CKII peptide (Ac-YPGGSTPVSSANMM-NH_2; 100 μM; obtained from GenScript) and UDP-GlcNAc or UDP-GlcNDAz (500 μM) were reacted with recombinant OGT (1 μM) in 20 mM TrisHCl buffer pH 8.0, containing 12.6 mM $MgCl_2$ and 20 mM β-mercaptoethanol. The reaction mixture was incubated for 16 h at 37 °C. Modified peptide was then purified by HPLC using an XDB-phenyl column, as described (47). Solvent was removed by rotary evaporation.

Peptide deglycosylation by OGA

Glycosylated (O-GlcNAc or O-GlcNDAz) peptide was incubated with 750 nM OGA (wild-type or C215A) in 50 mM sodium cacodylate buffer pH 6.5, containing 4 mM GalNAc for 20 h at room temperature. Reactions mixtures

were heated at 100 °C for 5 min to denature OGA, then analysed by HPLC using an XDB-phenyl column and UV detection at 220 nm, as described (47). Peaks were collected from the HPLC and analysed by intact mass measurement via electrospray ionization-quadruple time of flight (ESI-QTOF) to confirm peptide identities. From the reaction mixture containing O-GlcNDAz-modified CKII treated with wild-type OGA, the major peak had similar mobility to O-GlcNDAz-modified CKII and yielded an observed mass (1732 Da) corresponding to O-GlcNDAz-ylated CKII peptide (predicted 1728 Da). From the reaction mixture containing O-GlcNDAz-modified CKII treated with OGA(C215A), the major peak had mobility identical to unmodified CKII and yielded an observed mass (1485 Da) most consistent with deglycosylated CKII peptide (predicated 1440 Da).

RESULTS

OGA is inactive towards O-GlcNDAz

The Kohler group previously reported that lysate from HeLa cells overexpressing OGA was incapable of hydrolyzing an O-GlcNDAz mimic (47). I wished to confirm that OGA was unable to catalyze this reaction. To do this, I prepared purified recombinant human OGA for cell free assays (Fig. 3.2A).

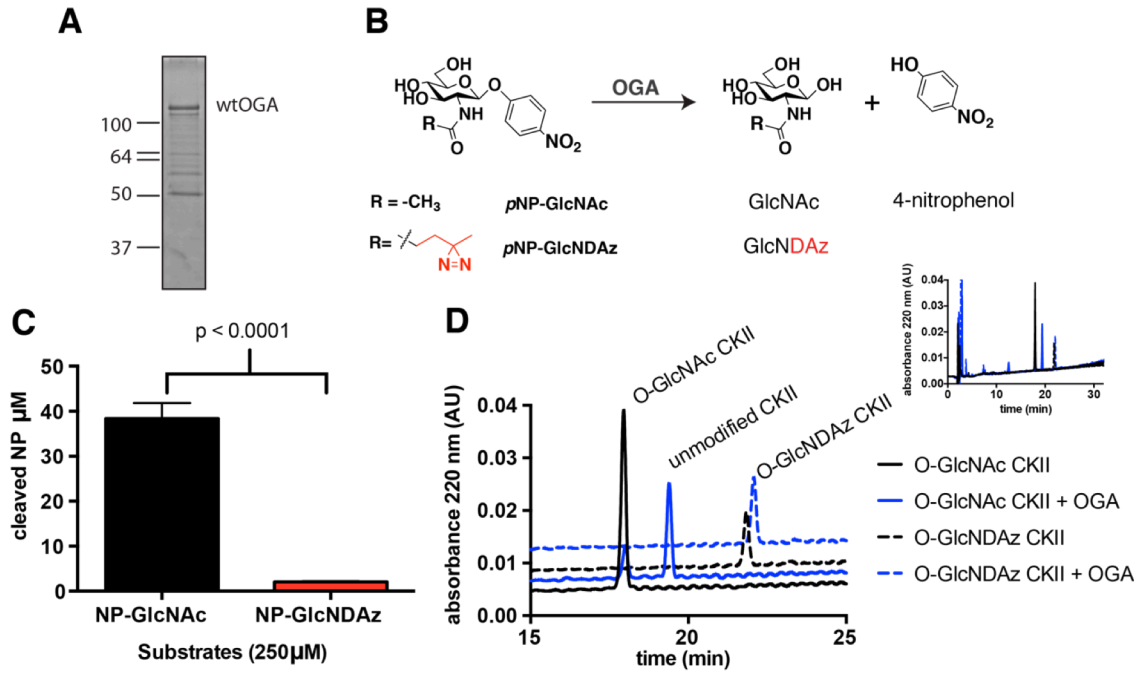

Figure 3.2. OGA is inactive toward substrates containing diazirine-modified GlcNAc. *A,* Purified recombinant wild-type OGA. *B,* Para-nitrophenol derivatives of GlcNAc (*p*NP-GlcNAc) and GlcNDAz (*p*NP-GlcNDAz) serve as chromogenic substrates to monitor OGA activity. *C*, Recombinant human OGA hydrolysed *p*NP-GlcNAc *p*NP-GlcNDAz. When 250 μM of substrate was incubated with OGA (50 nM). After 3 h, absorbance at 405 nm was measured to determine release of 4-nitrophenol. Error bars represent the standard deviation of three trials. *D*, O-GlcNAc- or O-GlcNDAz-modified CKII peptide was incubated with or without purified recombinant human OGA and reaction products were analysed by HPLC. Full HPLC traces are presented in top right corner.

The substrate tolerance of recombinant OGA was measured using *p*-nitrophenol derivatives of GlcNAc and GlcNDAz (47) (Fig. 3.2B). Para-nitrophenol N-acetylglucosamine (*p*NP-GlcNAc) is a mimic of O-GlcNAc that is hydrolyzed by OGA, yielding an increase in absorbance at 405 nm. Already, Seok-Ho Yu demonstrated that the diazirine-modified *p*NP-GlcNAc (*p*NP-GlcNDAz) served as an O-GlcNDAz mimic and reporter for O-GlcNDAz hydrolysis. As predicted, the recombinant OGA readily hydrolyzed *p*NP-GlcNAc, while showing no activity toward *p*NP-GlcNDAz (Fig. 3.2C).

To directly examine whether OGA is active towards O-GlcNDAz, I assessed the function of purified recombinant human OGA on a defined O-GlcNDAz-modified peptide. For this experiment, I employed a peptide derived from casein kinase II (CKII) containing a single serine residue that is readily O-GlcNAc-modified by OGT (2). O-GlcNAc- and O-GlcNDAz-modified forms of this peptide were prepared enzymatically. Glycosylated peptides were incubated with OGA and then analyzed by HPLC. I observed that OGA readily hydrolyzed the O-GlcNAc-modified CKII peptide, yielding unmodified CKII peptide, while displaying no activity toward the O-GlcNDAz-modified peptide, which remained intact (Fig. 3.2D).

O-GlcNDAz production is not toxic

While selective OGA inhibition is non-toxic,(82) other approaches have suggested that disruptions to O-GlcNAc regulatory enzymes can have deleterious effects on cellular homoeostasis (90). Due to the observation that OGA was ineffective towards O-GlcNDAz, I was concern that production of O-GlcNDAz could interfere with O-

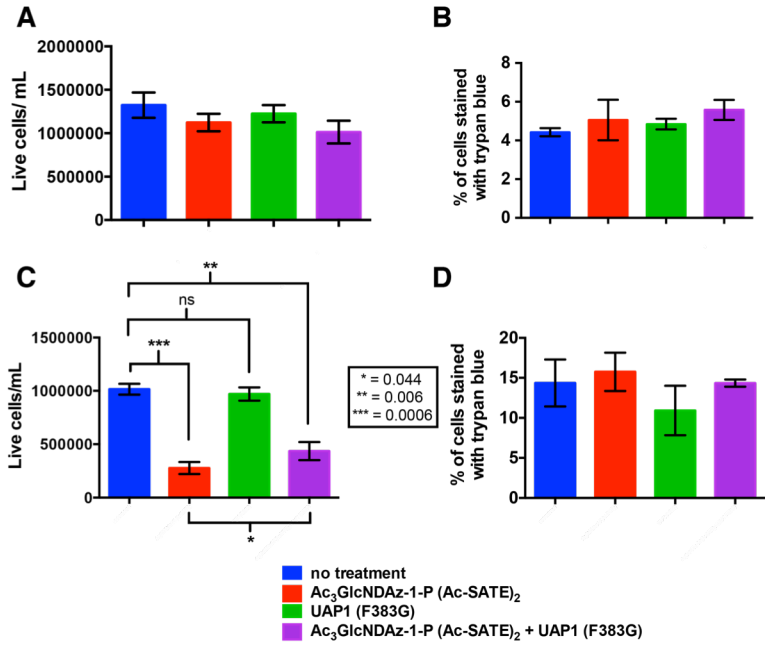

Figure 3.3. O-GlcNDAz production is non-toxic. *A*, Parental HeLa cells or HeLa cells expressing UAP1(F383G) (HeLa UAP1(F383G)) were cultured with or without Ac_3-GlcNDAz-1P(Ac-SATE)$_2$. After two days of O-GlcNDAz production, cells were counted. No significant differences among the conditions were detected. *B*, After two days, dead cells were stained with trypan blue and counted for viability. No significant differences among the conditions were detected. *C*, Parental HeLa cells or HeLa UAP1(F383G) were cultured with or without Ac_3-GlcNDAz-1P(Ac-SATE)$_2$. After two days of O-GlcNDAz production, cells were diluted and cultured with or without Ac_3-GlcNDAz-1P(Ac-SATE)$_2$ for three additional days. Then cells were counted. Addition of Ac_3-GlcNDAz-1P(Ac-SATE)$_2$ resulted in decreased proliferation, but this effect was partially reversed by simultaneous expression of UAP1(F383G). *D*, After five days, including one cell dilution, dead cells were stained with trypan blue and counted for viability. No significant differences among the conditions were detected. Error bars represent the standard deviation of three independent experiments for all conditions.

GlcNAc dynamics in the cell-based assay, and possibly lead to toxic effects. To address this problem, I examined whether production of O-GlcNDAz affected cell proliferation or viability. In our strategy for metabolic production of O-GlcNDAz, cells must express a mutant form of the UDP-GlcNAc pyrophosphorylase 1 (UAP1(F383G)) and be cultured in the presence of a cell-permeable GlcNDAz-1P (Ac_3-GlcNDAz-1P(Ac-SATE)$_2$) (47). I assessed control HeLa cells, UAP1(F383G)-expressing HeLa cells, HeLa cells cultured with cell-permeable GlcNDAz-1P, and UAP1(F383)-expressing HeLa cells cultured with GlcNDAz-1P (*i.e.* O-GlcNDAz-producing cells). When cells were cultured with cell-permeable GlcNDAz-1P for two days, corresponding to the time course of a typical photocrosslinking experiment, I observed no significant effects on cell proliferation (Fig. 3.3A) or viability (Fig. 3.3B). Cells were further cultured under these conditions, including a step where cells were diluted to allow more rapid proliferation. After five days, I observed decreased proliferation of the cells cultured with cell-permeable GlcNDAz-1P (Fig. 3.3C). However, this effect was not enhanced by expression of UAP1(F383G), which was required for O-GlcNDAz production. In addition, no effects on viability were observed during the longer time course (Fig. 3.3D). Thus, production of O-GlcNDAz does not have significant effects on the ability of cells to survive and divide during the course of typical photocrosslinking experiments.

However the cell-permeable GlcNDAz-1P caused slowed cell proliferation under extended conditions.

Modeling the interaction of OGA with O-GlcNDAz

Although O-GlcNDAz does not have dramatic effects on cell viability or proliferation, it likely alters the dynamics of intracellular protein glycosylation. Our data predict that O-GlcNDAz will accumulate in cells, mimicking the globally elevated O-GlcNAc levels that are observed in some disease states (5, 91). Based on this concern, my goal was to convert O-GlcNDAz to a dynamic modification. I hypothesized that a small change to the OGA active site might accommodate the diazirine-modified substrates.

While crystal structures of bacterial OGAs in complex with inhibitors have been solved, (76, 77) structural information about mammalian OGAs has come solely from modeling (92). As noted by others, (76, 92, 93) *Clostridium perfringens* OGA and *Bacteriodies thetaiotaomicron* OGA are closely related to human OGA: sequence alignment using Basic Local Alignment Search Tool (BLAST) (94) reveals that *Clostridium perfringens* OGA displays 51% similarity to human OGA, while *Bacteriodies thetaiotaomicron* is 55% similar to human OGA. Both bacterial enzymes have aspartates at positions corresponding to D174 and D175 of human OGA, identified as catalytic residues in the human enzyme (95).

I utilized the Protein Homology/analogY Recognition Engine V 2.0 (PHYRE^2; http://www.sbg.bio.ic.ac.uk/phyre2) remote homology modeling server (96) to create a model of human OGA. This model was aligned with the structure of *C. perfringens* OGA solved in complex with PUGNAc (PDB: 2CBJ) (76) (Fig. 3.4A) and with the structure of *B. thetaiotaomicron* solved in complex with Thiamet G (PDB: 2VVN) (77) (Fig. 3.4B). Using these alignments, I identified

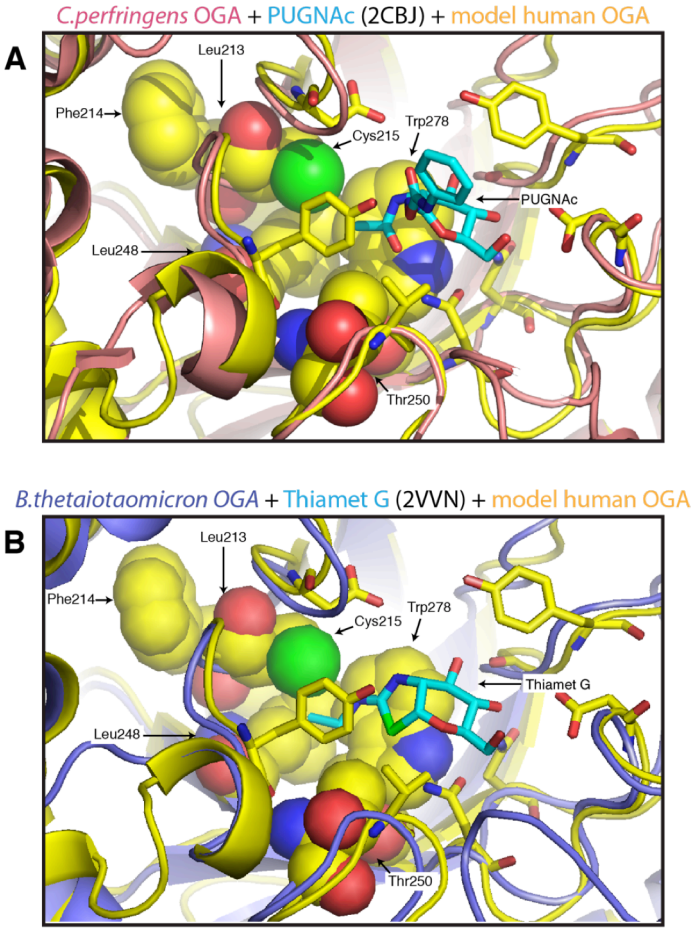

Figure 3.4. Mutations of OGA restore activity toward an O-GlcNDAz mimic. *A*, Structure of *C. perfringens* OGA complexed with PUGNAc (2CBJ; protein in pink, inhibitor in teal) aligned with a PHYRE^2-generated model of human OGA (in yellow). Residues selected for mutation are shown in space-filling representation. *B*, Structure of *B. thetaiotaomicron* OGA complexed with Thiamet G (2VVN; protein in blue, inhibitor in teal) aligned with the PHYRE^2-generated model of human OGA (in yellow). Residues selected for mutation are shown in space-filling representation.

OGA residues proximal to the nitrogen attached to the C-2 position of the pyranose ring of each inhibitor. Plasmids were constructed encoding OGA with each of the following mutations: F214A, C215A, L248G, T250A, W278A, and W278F, although the W278A mutant failed to express at detectable levels. *E. coli* lysates overexpressing each OGA mutant were incubated with *p*NP-GlcNAc (Fig. 3.5A) or *p*NP-GlcNDAz (Fig. 3.5B) and production of *p*-nitrophenol was detected by monitoring absorbance at 405 nm. Only lysate overexpressing the C215A mutant produced significant hydrolysis of both *p*NP-GlcNAc and *p*NP-GlcNDAz. Based on this result, an additional OGA mutant was made, changing C215 to glycine (C215G) based on the idea that additional space at this position might better accommodate hydrolysis of diazirine-modified substrate.

Mutations to OGA residue C215 confer activity toward an O-GlcNDAz mimic

Next, the activity of purified OGA(C215A) and OGA(C215G) was assessed (Fig. 3.5C) toward a diazirine-modified substrate. The kinetics of *p*NP-GlcNAc and *p*NP-GlcNDAz hydrolysis by wild-type OGA or mutant OGA was measured using a continuous assay that measured *p*-nitrophenol production. Wild-type OGA readily hydrolysed the *p*NP-GlcNAc substrate (Fig. 3.5D). Kinetic parameters for this reaction were similar to reported values, (86, 97) confirming the reliability of the continuous assay measurements. Both the C215A and C215G mutants retained activity toward *p*NP-GlcNAc, displaying

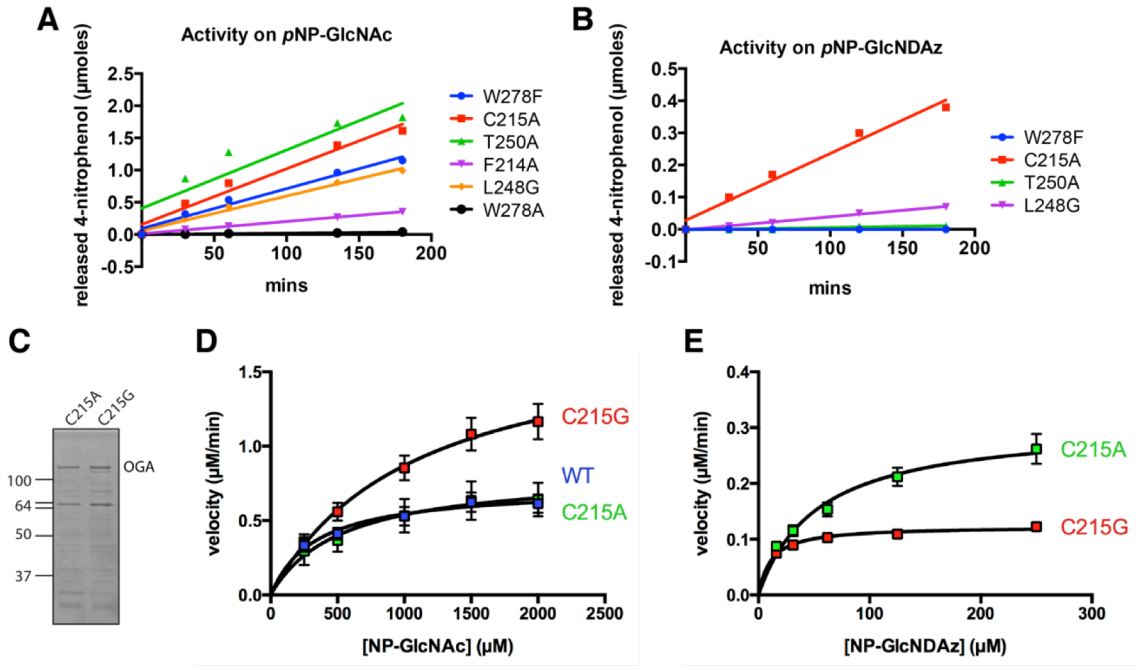

Figure 3.5. Mutations to C215 of OGA restore activity toward an O-GlcNDAz mimic. *A*, *E. coli* lysates overexpressing wild-type OGA or the indicated mutant were incubated with *p*NP-GlcNAc to measure release of nitrophenol. *B*, *E. coli* lysates overexpressing wild-type OGA or the indicated mutant were incubated with *p*NP-GlcNDAz to measure release of nitrophenol. *C*, Purified recombinant OGA(C215A) and OGA(C215G). *D*, Rates of hydrolysis of *p*NP-GlcNAc by wild-type OGA, C215A mutant of OGA, and C215G mutant of OGA. *E*, Rates of hydrolysis of *p*NP-GlcNDAz by the C215A and C215G mutants of OGA. Hydrolysis of *p*NP-GlcNDAz by wild-type OGA was too slow to be measured accurately by this approach.

catalytic efficiencies (k_{cat}/K_m) 2- to 3-fold lower than the wild-type enzyme (Table 3.1). In the case of the C215G mutant, this was achieved by modest increase in both k_{cat} and K_m. Both the C215A and C215G mutants also hydrolysed *p*NP-GlcNDAz (Fig. 3.5E). Remarkably, the catalytic efficiency for the C215A mutant hydrolysing *p*NP-GlcNDAz was comparable to the catalytic

efficiency for wild-type OGA acting on *p*NP-GlcNAc, although the k_{cat} and K_m values for the C215A/*p*NP-GlcNDAz pair were about 6-fold lower than those for the wild-type/*p*NP-GlcNAc pair. Overall, the kinetic analysis showed that both the C215A and C215G mutants of OGA are capable of efficient hydrolysis of *p*NP-GlcNDAz.

	***p*NP-GlcNAc**			***p*NP-GlcNDAz**		
Enzyme	**kcat (min^{-1})**	**Km (mM)**	**kcat/Km (mM^{-1} min^{-1})**	**kcat (min^{-1})**	**Km (mM)**	**kcat/Km (mM^{-1} min^{-1})**
Wild-type	14 (1)	0.17 (0.08)	82 (40)	no measurable activity		
C215A	17 (4)	0.54 (0.35)	32 (22)	2.4 (0.2)	0.03 (0.01)	86 (27)
C215G	36 (5)	0.86 (0.30)	42 (16)	1.5 (0.3)	0.04 (0.03)	38 (29)

Table 3.1. Michaelis-Menten parameters for hydrolysis of *p*NP-GlcNAc and *p*NP-GlcNDAz by purified OGA (wild-type, C215A, and C215G).

C215A mutant of OGA can remove O-GlcNDAz from a peptide substrate

Next, I examined whether the C215A mutation to OGA conferred the ability to hydrolytically remove O-GlcNDAz from a peptide. O-GlcNDAz-ylated CKII peptide was prepared enzymatically and purified by HPLC. The purified O-GlcNDAz-ylated peptide was incubated with wild-type OGA or the C215A mutant. Products of the reactions were analyzed by HPLC (Fig. 3.6A). Whereas wild-type OGA failed to hydrolyze O-GlcNDAz, the C215A mutant was active, as indicated by loss of the peak corresponding to O-GlcNDAz-ylated peptide. The appearance of a new peak corresponding to unmodified

peptide was verified by mass spectrometry. Under these conditions, approximately 37% of the O-GlcNDAz-modified peptide was deglycosylated, which compares favorably to the deglycosylation of O-GlcNAc-modified peptide by wild-type OGA, in which 53% of the peptide was deglycosylated (Fig. 3.1D). Thus, O-GlcNDAz became a reversible post-translational modification in the presence of the C215A mutant of OGA.

DISCUSSION

Here the activity of recombinant human OGA was examined toward a substrate in which the N-acyl side chain of GlcNAc was extended by modification with an alkyl diazirine (GlcNDAz). While Vocadlo and co-workers had shown that human OGA is relatively accommodating of steric bulk at the N-acyl position (97), the alkyl diazirine modification is larger than the substituents they evaluated. Therefore, it was perhaps unsurprising that I

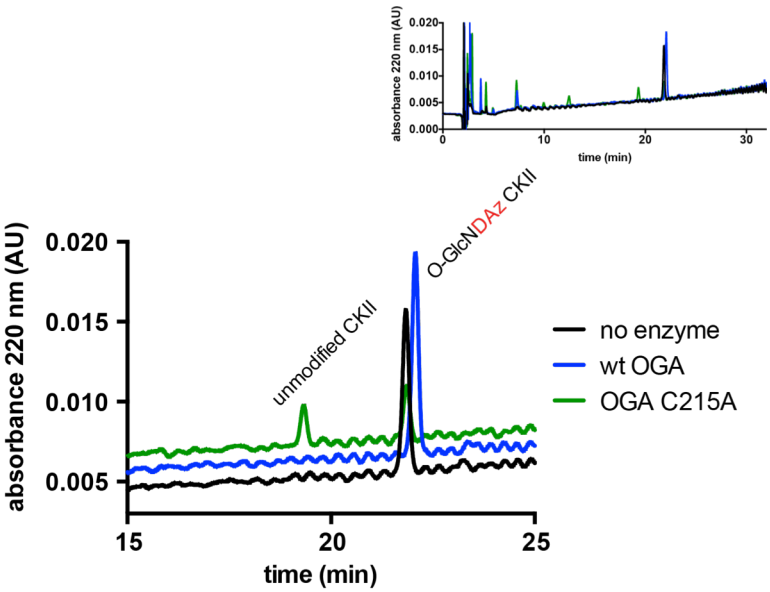

Figure 3.6. C215A mutant of OGA hydrolyzes an O-GlcNDAz-modified peptide. O-GlcNDAz-ylated CKII peptide was treated with no enzyme, with wild-type OGA, or with the C215A mutant of OGA. Treatment with the C215A mutant yielded a decrease in the peak corresponding to O-GlcNDAz-ylated CKII peptide and appearance of a peak corresponding to unmodified CKII peptide. Full HPLC traces are presented in top right corner.

found that wild-type OGA failed to hydrolytically remove GlcNDAz from either *p*-nitrophenol or a peptide.

To interpret these results molecularly, I turned to a structural model of human OGA constructed by homology modelling. The model was inspected in alignment with bacterial OGAs complexed with inhibitors that are structural analogs of O-GlcNAc. This analysis revealed the potential for steric clashes between the diazirine modification and a handful of amino acid side chains projecting into the enzyme active site. In particular, the N-acyl methyl group on the inhibitor PUGNAc projects toward the C215 side chain of OGA, with a C-S separation of about 4 Å. Mutations to C215 had little effect on the rate of *p*NP-GlcNAc hydrolysis, suggesting that this amino acid does not play a significant role in catalysis. As predicted, mutation of this position to a smaller side chain (alanine or glycine) enabled activity on diazirine-modified substrates. Mutations to other nearby amino acid positions resulted in dramatically reduced activity toward *p*NP-GlcNAc. The decreased activity suggested that these positions were critical for catalysis or proper enzyme folding. The model as a predictive tool permitted me to reengineer the enzyme active site to optimize enzyme-substrate recognition. In the absence of an available crystal structure of human OGA, homology modelling may also be employed in efforts to design new OGA inhibitors.

The observation that OGA was unable to remove O-GlcNDAz has important implications for our previous reported method for discovering interaction partners of O-GlcNAc-modified proteins. Since OGA fails to remove O-GlcNDAz, I predict that this modification will accumulate in cells, raising the overall level of intracellular protein glycosylation. Because inhibition of OGA was reported to cause slowed proliferation and cell cycle defects,(81) I measured the proliferation and viability of O-GlcNDAz-producing cells. I observed no dramatic effects attributable to O-GlcNDAz production, consistent with a growing recognition that selective OGA inhibition is non-toxic (82, 98). Nonetheless, the subtle effects of O-GlcNDAz may not be detected in proliferation and viability assays. Indeed, the half-life of O-GlcNAc modification varies for different proteins,(99, 100) with O-GlcNAc dynamics playing a critical role for some functions (39). Thus, OGA mutants capable of removing O-GlcNDAz were identified.

Future work for OGA(C215A) to improve O-GlcNDAz technology

As previously stated (**chapter 2**), O-GlcNDAz technology was employed in the identification of binding partners of NUP98 and leukemogenic fusions. However, challenges arose to identify the binding partners of the O-GlcNAc-modified NUP98 and NUP98 fusions. A cell free O-GlcNDAz crosslinking method to identify the potential candidates of O-GlcNAc-modified proteins will be discussed in later chapters. The discovery that mutations of C215 of OGA

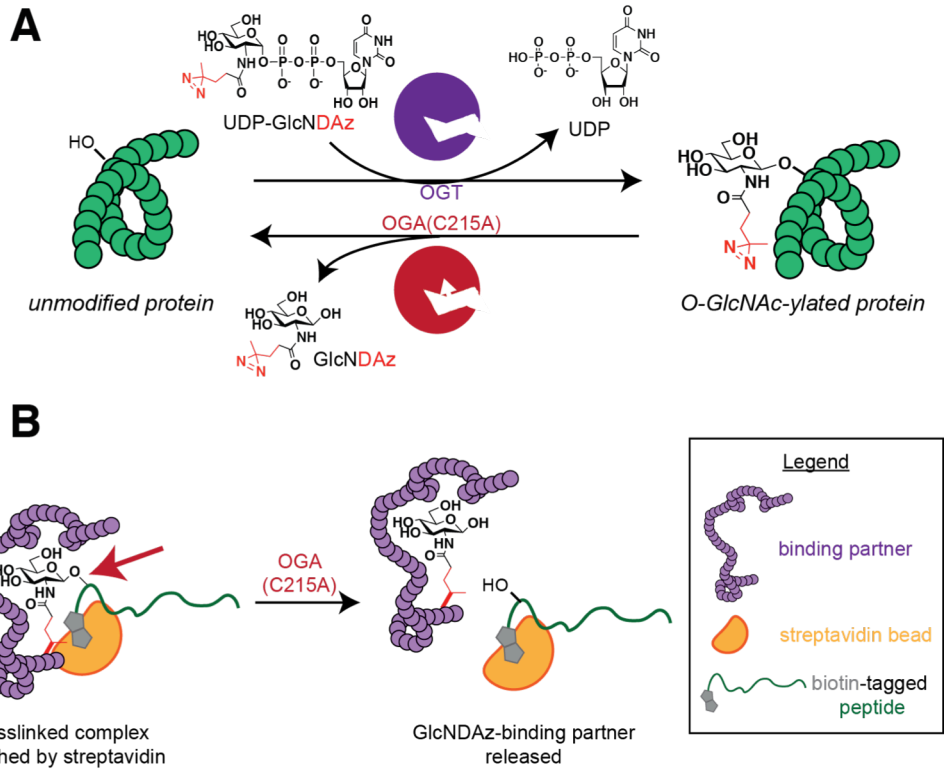

Figure 3.7. C215A mutant of OGA as an employment tool to improve the GlcNDAz technology. *A,* Cells overexpressing OGA(C215A) will ensure the dynamics of O-GlcNDAz modification, similar to that of O-GlcNAc. OGT, O-GlcNAc transferase, will transfer GlcNDAz onto O-GlcNAc-modified substrates, while OGA(C215A) will remove GlcNDAz. *B*, O-GlcNDAz-modified peptide crosslinked to unknown binding partners are enriched using streptavidin purification. OGA(C215A) is exploited to hydrolyze the O-GlcNDAz crosslink (red arrow site of hydrolysis). This will release molecules crosslinked to the immobilized O-GlcNDAz-modified peptide. Released molecules will be identified by mass spectrometry. The OGA(C215A)-catalyzed release is predicted to yield a more specific candidate list than other approaches.

confer activity toward diazirine-modified substrates offers a strategy for improving our O-GlcNDAz crosslinking method both in cells and cell free systems. In the cell-based O-GlcNDAz crosslinking, introduction of

OGA(C215A) into O-GlcNDAz producing cells will restore the normal dynamics of intracellular protein O-GlcNAc, and enable the photocrosslinking method to accurately report on physiologically relevant interactions in which O-GlcNAc plays a critical role (Fig. 3.7A). This OGA(C215A) would improve the cell-based O-GlcNDAz crosslinking assay.

Also, employment of OGA(C215A) in the cell free O-GlcNDAz crosslinking approach will be an advantage in obtaining a more specific candidate list. For example after enrichment of crosslinked complexes, the biotin-tagged O-GlcNAc-modified substrate will be interacting with streptavidin beads. Typically proteomics analysis uses trypsin digest to obtain proteins interacting with the substrate of interest on-beads or in-gels. However, these protein interactions may not be O-GlcNAc dependent. I propose that cleavage of the glycosidic bond of GlcNDAz-crosslinking binding partners before trypsin digest will produce a more narrow focus candidate list (Fig. 3.7B). This OGA mutant will assist in the advancement of cell free O-GlcNDAz crosslinking assay.

ACKNOWLEDGEMENTS

I thank Dr. John Hanover (NIDDK) for the pBAD/HisA OGA plasmid, Dr. Kate Harwood (NIDDK) for advice on OGA purification, Dr. Suzanne Walker (Harvard Medical School) for the pET24b OGT plasmid, Dr. Ilmin Kwon (UT Southwestern) for advice on OGT purification, Dr. Kenneth Westover (UT

Southwestern) for the use of the Profinia system, and Dr. David Vocadlo (Simon Frasier), Suong Nguyen (UT Southwestern), and Dr. Seok-Ho Yu (Complex Carbohydrate Research Center) for insightful discussions. A.C. Rodriguez and J.J. Kohler, *Med. Chem. Commun.,* 2014, 5, 1227 reproduced by permission of The Royal Society of Chemistry.

CHAPTER FOUR:

Improvement of O-GlcNDAz incorporation and crosslinking through mutation of O-GlcNAc transferase (OGT)

INTRODUCTION

O-linked β-*N*-acetylglucosamine (O-GlcNAc) is a dynamic single sugar post-translational modification (PTM), found on thousands of proteins (1, 101, 102). This reversible PTM occurs on proteins of diverse functional classes, but primarily those that reside in the cytosol and nucleus. The function of this PTM remains elusive. As presented in earlier chapters, O-GlcNAc is known to assist in regulation of cellular homeostasis. Additionally, this PTM has been suggested to inhibit or mediate protein-protein interactions of O-GlcNAc modified proteins (14, 15, 42).

Intracellular O-GlcNAc is produced by the enzyme O-GlcNAc transferase (OGT), which catalyzes the transfer of a single GlcNAc sugar from the nucleotide sugar donor UDP-GlcNAc to the hydroxyl groups of serine and threonine side chains of proteins through formation of a β-glycoside linkage (103, 104). However, a consensus sequence for substrate preference for OGT has not been determined, and is unlikely to exist due to the wide array of proteins that can be modified. The OGT gene, which resides on the X chromosome, is demonstrated to be vital (4). When cells are treated with siRNA targeting the OGT mRNA, the cells proliferate normally for up to 48 hours, but soon high levels of apoptosis occur (105). In a

mouse model, the knockout of OGT results in a lethal phenotype, in which the mice die during embryogenesis (4).

Since the discovery of OGT in the 1980s, there has been much speculation as to how the enzyme functions to transfer the GlcNAc to its peptide substrate. OGT belongs to the glycosyltransferase type-B (GT-B) fold with thirteen N-terminal tetratricopeptide repeats (TPRs) (Fig. 4.1A). Multiple aspects regulate the expression of the OGT activity and protein levels, such as PTMs, transcriptional regulation, mRNA splicing, and proteolytic processing (106). Importantly, these TPR domains suggest playing a role in substrate selectivity through protein-protein interactions (2, 107). Walker and co-workers reported the crystal structure of OGT, which provided insight into the mechanism of adding GlcNAc to peptides (85, 108). The group established that the enzyme coordinated with UDP-GlcNAc followed by the addition of the peptide substrate. OGT employs two histidine resides, H498 and H558, to transfer GlcNAc onto peptides before release of the O-GlcNAc-ylated peptide product followed by the UDP by-product (Fig. 4.1B).

The Kohler lab previously reported photocrosslinking O-GlcNDAz technology (47) could provide insight into the binding partners of O-GlcNAc-modified proteins (Fig. 4.1C). However, in cells, the unnatural photocrosslinking nucleotide donor sugar, UDP-GlcNDAz, competes with the endogenous nucleotide sugar, UDP-GlcNAc, for OGT-catalyzed transfer onto proteins. Limited production of the O-GlcNDAz modification in cells would reduce that amount of crosslinked material

available for analysis. Mutation of the binding pocket of OGT may increase production of the desired unnatural modification, O-GlcNDAz, and yield more crosslinked material for subsequent steps. Also, several challenges were revealed in **chapter 2** that led to the need for an improved method by mutagenesis of OGT for the GlcNDAz

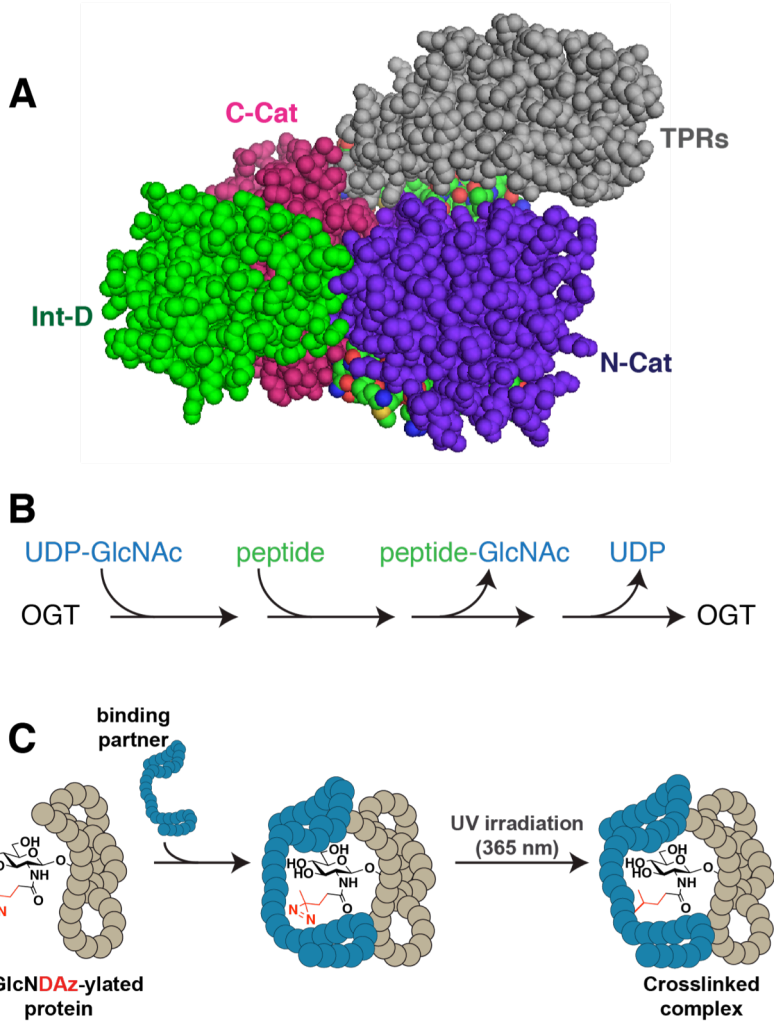

Figure 4.1. Methodology to improve O-GlcNDAz-ylation on proteins. *A,* Domains of the OGT enzyme include two catalytic domains, one Int-D domain, and TPR. *B,* The sequence of events in which OGT transfers GlcNAc onto the peptide substrate. *C,* O-GlcNAc-ylated proteins are metabolically labeled with O-GlcNDAz in place of the natural modification. UV irradiation of cells activates the diazirine for crosslinking, resulting in a covalent crosslinked complex between an O-GlcNDAz-ylated protein and its binding partner.

technology to allow me to expand to other O-GlcNAc-modified proteins.

In this chapter, I summarized my efforts to improve the O-GlcNDAz technology by examining the OGT enzyme. First, the ability of OGT to transfer natural GlcNAc and photocrosslinking GlcNDAz to peptide substrates was examined using cell free assays. I determined that wild-type OGT (wtOGT) exhibits a marked preference for the natural sugar UDP-GlcNAc, a feature that likely limited O-GlcNDAz production in cells. Next, an OGT mutant (C917A) was identified that exhibits inverted substrate specificity, preferentially transferring GlcNDAz rather than GlcNAc. Kinetic characterization of the activity of wtOGT and OGT(C917A) with UDP-GlcNAc and UDP-GlcNDAz confirmed the reciprocal substrate specificities of the two enzymes. Finally, OGT(C917A) revealed increased O-GlcNDAz-ylation, and enhanced O-GlcNDAz-mediated crosslinking both in cell free and in cell systems. Using OGT(C917A) should enable increased production of O-GlcNDAz both on peptides and proteins, and lead to increased O-GlcNDAz-mediated crosslinking.

EXPERIMENTAL PROCEDURES

Construction of OGT Plasmids

All experiments were conducted using the ncOGT isoform. The pET24b plasmid encoding OGT was provided by Suzanne Walker (Harvard Medical School) and used to express recombinant OGT in *E. coli* (85). The pcDNA/To-myc-His

plasmid encoding OGT was provided by Carolyn Bertozzi (UC Berkeley) and used to produce recombinant OGT in mammalian cells (109).

Point mutations were introduced into OGT plasmids using primers designed through the GenScript primer design algorithm. Mutagenesis was performed according to the QuikChange protocol (Stratagene). The following primers were used: M501G forward 5'-GTGCATCCTCATCATAGTGGGCTATATCCTCTTTCTCATGGC-3', and reverse 5'-GCCATGAGAAAGAGGATATAGCCCACTATGATGAGGATGCAC-3'; M501V forward 5'-GTGCATCCTCATCATAGTGTGCTATATCCTCTTTCTCATGGC-3', and reverse 5'-GCCATGAGAAAGAGGATATAGCACACTATGATGAGGATGCAC-3'; C917G forward 5'-GACACTCCACTCGGTAATGGGCACACCACAGGG-3', and reverse 5'-CCCTGTGGTGTGCCCATTACCGAGTGGAGTGTC-3'; C917A forward 5'-GACACTCCACTCGCTAATGGGCACACCACAGGG-3', and reverse 5'-CCCTGTGGTGTGCCCATTAGCGAGTGGAGTGTC-3'; A942G forward 5'-GGAGAGACTCTTGGTTCTCGAGTTGCAGCATCC-3', and reverse 5'-GGATGCTGCAACTCGAGAACCAAGAGTCTCTCC-3'. The entire coding sequence of each mutated plasmid was confirmed through sequencing at the UT Southwestern DNA sequencing core facility.

Cell lines and cell culture

HeLa cells, obtained from the ATCC, were cultured in DMEM, 10 % fetal bovine serum (FBS), and 1 % penicillin/streptomycin. K562 cells were obtained from

ATCC and cultured in DMEM, 10 % FBS, and 1 % penicillin/ streptomycin for routine culturing and for crosslinking experiments.

To produce cells stable expressing UAP1(F383G), K562 cells were infected with the pSIN4-UAP1(F383G) lentivirus.(47) Briefly, 2.0 x 10^5 K562 cells in a 1 mL volume of fresh media were mixed with 1 mL of the pSIN4-UAP1(F383G) lentivirus. Cells were centrifuged at 514*g*, 32 °C for 30 min. After gentle mixing, centrifugation was repeated at 514*g*, 32 °C for 30 min. After centrifugation, media was removed by aspiration and cells were resuspended in 2 mL of fresh media and transferred to a well of a 6 well plate. Selection with 1.5 *μ*g/mL puromycin for cells expressing UAP1(F383G) was performed for 2 weeks.

Production of UDP-GlcNDAz in UAP1(F383G)-expressing K562 cells was analyzed by high performance anion exchange chromatography (HPAEC) as described.(47) Briefly, K562 cells were cultured with or without Ac_3-GlcNDAz-1P(Ac-SATE)$_2$ (>95 % purity by ^{1}H NMR) for 6 h. Cells were harvested by centrifugation, washed once with PBS, and pelleted. Lysing cells released metabolites with 80 % “super cold” methanol on dry ice. Cellular debris were removed by centrifugation at 20,817*g*, 4 °C for 10 min. The supernatant containing metabolites was transferred to a clean vessel and evaporated under vacuum. The metabolite pellet was resuspended in 40 mM sodium phosphate buffer (pH 7.0), and analyzed by HPAEC using conditions reported previously (47).

Transient overexpression of mutant OGT enzymes in HeLa cells

OGT and OGT mutants were overexpressed in HeLa cells to identify mutants with improved activity toward UDP-GlcNDAz. HeLa cells (3 million) in Opti-MEM (Invitrogen, 11058-021) were transfected with 24 μg DNA and 60 μL Lipofectamine 2000 (Invitrogen, 11668-019). Four hours after transfection, Opti-MEM media was removed, and fresh DMEM media was added. After 24 h, cells were harvested and lysed in RIPA buffer (50 mM TrisHCl pH 8.0, 150 mM NaCl, 1.0 % NP-40, 0.5 % sodium deoxycholate, 2 mM EDTA, 1 mM DTT, protease inhibitor cocktail, and 1 mM phenylmethanesulfonylfluroide (PMSF)). Myc-His-tagged OGT (wild-type and mutants) were immunoprecipitated by incubating 1.2 mg of total protein with 1 μL of anti-myc antibody 9B11 (Cell Signaling, 22768) at 4 °C for 18 h. After overnight incubation, 25 μL of protein G beads (Sigma, P3296) were added and incubated for 3 h at 4 °C. Beads were washed three times with RIPA buffer lacking protease inhibitors and then used for the following glycosyltransferase activity assays.

Glycosylation reactions were performed to analyze the ability of the OGT (wild-type and indicated mutants) ability to release UDP via the Promega UDP-Glo™ Glycosyltransferase Assay (V6962). This assay measures UDP release via production of a luminescence signal. Beads with the immunopurified OGT enzymes (wild-type and indicated mutants) were resuspended in 25 μL of Transferase Buffer (25 mM HEPES pH 7.2, 10 mM $MnCl_2$, 1 mM EDTA with 1 x protease inhibitor cocktail (Santa Cruz, sc-29131) and 1 mM PMSF). Five μL of this resuspension was

mixed together with 20 μL of Transferase Buffer containing OGT substrates, to achieve final concentrations of 100 μM casein kinase II (CKII) peptide and 10 μM Ultra-Pure UDP-GlcNAc (provided by Promega) or UDP-GlcNDAz (synthesized as described previously;(19) >95 % purity by ^{1}H NMR). Reaction mixtures were incubated at room temperature for 0 min or 60 min. UDP-Glo Detection Reagent was added to quench the glycosylation reaction and develop luminescence. After 1 hour at room temperature, luminescence signal was measured using Synergy Neo BioTek Multi-mode plate reader with detection between 620 and 665 nm (Lance red emission, BioTek, part.no. 1035041).

Glycosylation reactions were performed to measure the ability of immunopurified OGT (wild-type and mutants) to transfer GlcNAc and GlcNDAz to the CKII peptide. Beads with the immunopurified OGT enzymes (wild-type and indicated mutants) were washed twice with Transferase Buffer. A 150 μL mixture of CKII peptide (final concentration 100 μM) and equimolar amounts of UDP-GlcNAc and UDP-GlcNDAz (final concentrations 500 μM each) was added to each bead sample. The reaction was allowed to incubate at room temperature for 1 d, followed by 4 °C for 2 d. The reaction was centrifuged to remove beads. The supernatant was analyzed by high performance liquid chromatography (HPLC) using a Dionex HPLC Ultimate 3000 with UV detection at 220 nm. Solvent systems and gradients were described previously (110).

Overexpression of recombinant OGT and OGT(C917A) in E. coli and purification

pET24b plasmid encoding OGT (20) or OGT(C917A) was used to produce recombinant enzyme as previously described (110). Briefly, BL21(DE3) competent cells were transformed with each plasmid. A single colony was added to a 50 mL of LB containing kanamycin and incubated at 37 °C, without shaking for 14 h. Eight mL of the starter culture was used to inoculated 1 L of LB, which was incubated at 37 °C with shaking at 250 rpm until OD_{600} reached 0.6-0.8. After cooling the culture to 20 °C for 30 minutes, cells were induced with 0.5 mM IPTG for 20 h. Bacteria were harvested by 30 min centrifugation at 1154*g* at 4 °C. Bacteria were lysed in 100 mM TrisHCl pH 7.5, 1 M NaCl, 1 mM Triton X-100, 0.3 mM lysozyme, 1 x protease inhibitor cocktail (Sigma, S8830) on ice for 30 minutes, followed by sonication for 3 minutes alternating 10 seconds pulse on and 30 seconds pulse off. Insoluble material was removed by centrifugation at 12,000*g*, 4 °C for 30 min. Cell lysates were incubated with nickel beads (Ni-NTA, Qiagen, 30210) for batch purification. Nickel beads were washed with 20 mM TrisHCl pH 7.5, 0.5 M NaCl, 20 mM β-mercaptoethanol, 20 mM imidazole, and 0.1 mM PMSF. OGT (wild-type or mutant) was eluted with 20 mM TrisHCl pH 7.5, 0.5 M NaCl, 20 mM β- mercaptoethanol, 200 mM imidazole, and 0.1 mM PMSF. As an alternative to batch purification, OGT was also purified used a Profinia instrument (Bio-Rad, 620-1010) with an IMAC column (Bio-Rad, 732-4610) according to the manufacturer's instructions and using the

same buffers as the batch method. Protein purity was estimated by Coomassie staining. wtOGT was > 80 % pure and OGT(C917A) was > 90 % pure.

Analysis of peptide glycosylation

wtOGT or OGT(C917A) recombinantly expressed in and purified from *E. coli* was used for analysis of peptide glycosylation. The CKII substrate peptide sequence was purchased from GenScript and had the sequence PGGSTPVSSANMMK with N-terminal acetylation and C-terminal amidation. Reactions contained 20 mM TrisHCl pH 8.0, 12.6 mM $MgCl_2$, 20 mM β-mercaptoethanol, 100 μM CKII peptide, and 1 μM wtOGT or OGT(C917A). UDP-GlcNAc (Sigma, U4375, > 98 % purity) and/or UDP-GlcNDAz were included at the concentrations indicated. After incubation at 37 °C for 16 h, the enzyme was removed by filtration through a 10 kDa MWCO Amicon unit (Millipore, UFC5010BK) that was centrifuged at 17000*g* for 5 min at room temperature. Peptide products were analyzed using a Dionex HPLC Ultimate 3000 with UV detection at 220 nm. Solvent systems and gradients were described previously (110).

Determination of Michaelis-Menten parameters

Initial rates of reactions catalyzed by wtOGT and OGT(C917A) were obtained using the Promega UDP-Glo™ Glycosyltransferase Assay (V6962) to measure UDP release. Reactions contained 20 mM TrisHCl pH 8.0, 12.6 mM $MgCl_2$, 20 mM β-mercaptoethanol, and 100 μM CKII peptide. The nucleotide sugar (UDP-GlcNAc

(provided by the Promega kit) or UDP-GlcNDAz) concentrations were 0.3, 1, 3, 10, and 30 μM. The enzyme concentration was 25 nM, except in the case of wtOGT with UDP-GlcNDAz where 50 nM wtOGT was needed to accurately measure activity. Reactions were incubated at 23 °C for a time course of 5, 10, 15, 30 or 45 min, then UDP-Glo Detection Reagent was added to quench the glycosylation reaction and develop luminescence. After 1 h at room temperature, luminescence signal was measured using Synergy Neo BioTek Multi-mode plate reader with detection between 620 and 665 nm (Lance red emission, BioTek, part no. 1035041). Each sample was measured twice and the averaged value was used. For each enzyme/substrate pair, five or six independent experiments were performed. Typically, the increase in luminescence was linear with respect to time between 5 to 45 min, except for the case of OGT(C917A) with UDP-GlcNDAz where the time-dependent increase in luminescence was linear only to 30 min. Linear data were used to determine initial rates, which were plotted versus substrate concentration. Michaelis–Menten parameters were calculated using GraphPad Prism.

Cell free peptide crosslinking and analysis

The biotin-tagged CKII peptide was purchased from GenScript and had the sequence biotin-PGGSTPVSSANMMK with C-terminal amidation. Glycosylation and HPLC analysis were performed as described above for the unmodified peptide. After removing the enzyme by filtration, the crude filtrate was incubated with an O-GlcNAc recognizing antibody (RL2, Thermo, MA1-027) in RIPA buffer for 15 h at 4

°C. Samples placed in an ice water bath under UV irradiation (365 nm, UVP, XX-20BLB lamp) for 15 min or kept on ice in the dark. After irradiation, the samples were immediately lysed in RIPA buffer and then analyzed by immunoblot, using an HRP-conjugated anti-biotin antibody (Cell Signaling, 7075S) to detect antibody crosslinked to biotinylated peptide. Quantification was performed using a ChemiDoc™ MP Imager (Bio-Rad, 170-8280). Blots were stripped and reprobed to detect the total amount of RL2 antibody. The amount of biotinylated (crosslinked) antibody in each lane was normalized to the total amount of antibody in the same lane. The amount of crosslinking in the sample containing wtOGT and 500 μM UDP-GlcNDAz was arbitrarily set to 1, and crosslinking in other samples was calculated relative to this value.

In cell crosslinking

Two million K562 cells stably expressing UAP1(F383G) were nucleofected with 1.5 μg DNA encoding an OGT plasmid (wtOGT or OGT(C917A), by following the protocol provided by the Lonza kit V (VCA-1003). After 5 h, cells were resuspended in DMEM. Twenty-four hours after nucleofection, cells were cultured with media containing 6 μL of 50 mM Ac_3-GlcNDAz-1P(Ac-SATE)$_2$ (cell-permeable GlcNDAz-1P) in DMSO, to achieve a final concentration of 100 μM Ac_3-GlcNDAz-1P(Ac-SATE)$_2$, or 6 μL DMSO alone. After 7 h, cells were harvested by centrifugation and resuspended in PBS before irradiation with 365 nm light for 15 min. Cells were lysed in 100 μL RIPA buffer per 2 million cells, then centrifuged at

20,817*g,* for 10 min at 4 °C to remove insoluble material. Lysates were separated by 7.5 % SDS-PAGE, transferred to polyvinylidene fluoride and probed for NUP153 (SA-1 Abcam, ab96462). The amount of crosslinked NUP153 was determined by quantifying the indicated region on the blot using a ChemiDoc™ MP Imager (Bio-Rad, 170-8280). Immunoblots were stripped and reprobed with an antibody against actin (Abcam, ab8227). The amount of crosslinked NUP153 in each sample was normalized to amount of actin in the same sample to control for sample loading. The amount of crosslinking in the sample containing overexpressed wtOGT was arbitrarily set to 1, and crosslinking in the sample containing OGT(C917A) was calculated relative to this value. Immunoblots were stripped and reprobed with an antibody against myc antibody 9B11 (Cell Signaling, 22768 (Abcam, ab8227) for OGT expression.

Immunoprecipitation was performed with the same cell culture conditions; 70 μg of total protein lysate was incubated with 1 μL of mAB414 antibody (Covance, MMS-120R) in total volume of 200 μL for 14 h at 4 °C. Samples were incubated with 10 μL Protein G Sepharose ® Fast Flow (Sigma Aldrich, P3296) for 3 h at 4 °C. The beads were washed three times with RIPA buffer. Proteins were eluted by addition of 2X loading dye (100 mM TrisHCl pH 6.8, 25 mM EDTA, 0.04 g/mL SDS, 0.4 mg/mL bromophenol blue) containing 20 mM DTT, and incubation for 10 min at 90 °C. Samples were analyzed by immunoblot using the SA-1 antibody, as above.

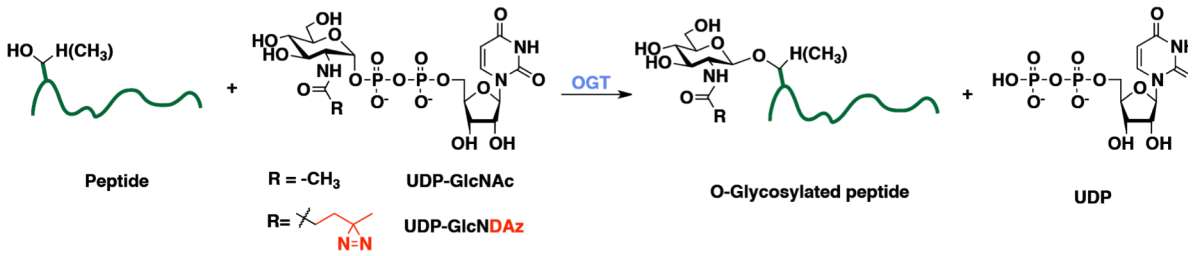

Figure 4.2. OGT glycosylation reactions with natural and unnatural substrates. OGT transfers GlcNAc from the natural substrate, UDP-GlcNAc, to serine or threonine residues of substrate peptides, producing O-GlcNAc-modified peptides. To incorporate O-GlcNDAz on peptides, OGT must transfer GlcNDAz from an unnatural nucleotide sugar, UDP-GlcNDAz.

RESULTS

Wild-type OGT prefers UDP-GlcNAc over UDP-GlcNDAz

Cell free assays were utilized to measure the ability of purified recombinant OGT to transfer GlcNAc and GlcNDAz to substrate peptides (Fig. 4.2). Recombinant wtOGT was expressed in *E. coli* and purified (Fig. 4.3A) (110). The enzyme was incubated with a synthetic CKII peptide (2, 110) and either UDP-GlcNAc or UDP-GlcNDAz, and production of the corresponding glycosylated peptides was evaluated by HPLC analysis. I observed efficient production of both O-GlcNAc-modified CKII (red trace, Fig. 4.3B) and O-GlcNDAz-modified CKII (green trace, Fig. 4.3B). Next, the nucleotide sugar specificity of wtOGT was evaluated using a competition assay in which the enzyme was incubated with an equimolar mixture of UDP-GlcNAc and UDP-GlcNDAz. This 1:1 ratio of UDP-GlcNAc to UDP-GlcNDAz was chosen because it is similar to the ratio was observed in cells that have been engineered for

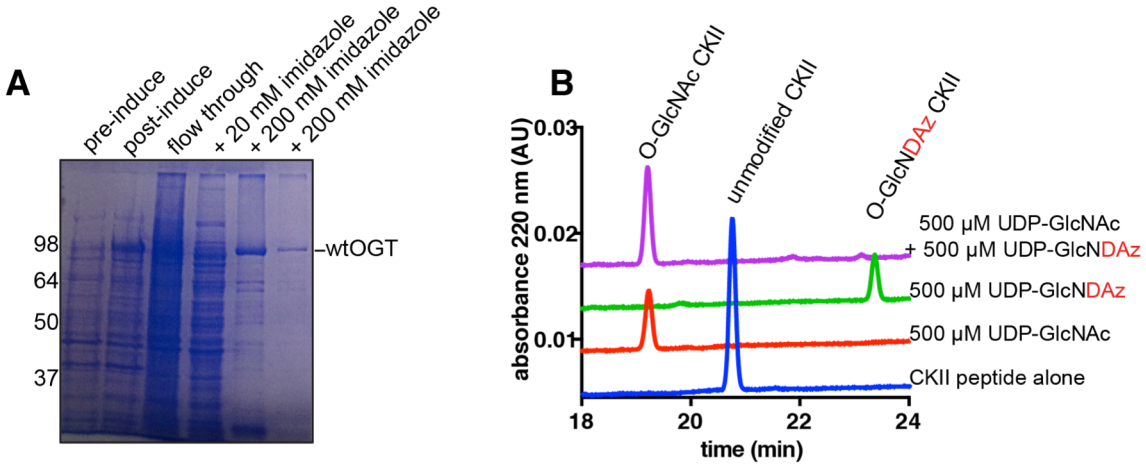

Figure 4.3. wtOGT prefers UDP-GlcNAc to UDP-GlcNDAz. *A,* Coomassie stain of the wtOGT expression and purification. *B,* Glycosylated CKII peptides produced using wtOGT. Wild-type OGT (1 μM) was incubated together with CKII peptide (100 μM) and 500 μM UDP-GlcNAc, 500 μM UDP-GlcNDAz, or a mixture of 500 μM UDP-GlcNAc and 500 μM UDP-GlcNDAz. Reaction products were separated by reverse-phase HPLC using UV absorbance (220 nm) for detection. Experiments were performed in triplicate with a single representative trial shown.

UDP-GlcNDAz production (47). Under these competitive conditions, wtOGT preferentially transferred GlcNAc, yielding primarily O-GlcNAc-modified CKII and very little O-GlcNDAz-modified CKII (purple trace, Fig. 4.3B). On average, about 92 % O-GlcNAc-ylated peptide was observed and about 8 % O-GlcNDAz-ylated peptide was produced under conditions of equimolar UDP-GlcNAc and UDP-GlcNDAz. This result indicated a significant limitation of the current O-GlcNDAz crosslinking protocol that depends on endogenous OGT for O-GlcNDAz production. I hypothesized that mutating the binding pocket of OGT could alter the substrate specificity might result in improved O-GlcNDAz production and crosslinking in cells. Ultimately a mutant could yield more material for proteomic identification of O-GlcNAc binding partners for both cell-based and cell free assays.

Expanding the OGT active site to accommodate the diazirine of UDP-GlcNDAz

Fortunately, several crystal structures of OGT in complex with substrates, including UDP-GlcNAc, have been reported (85, 111-113). The structure of OGT

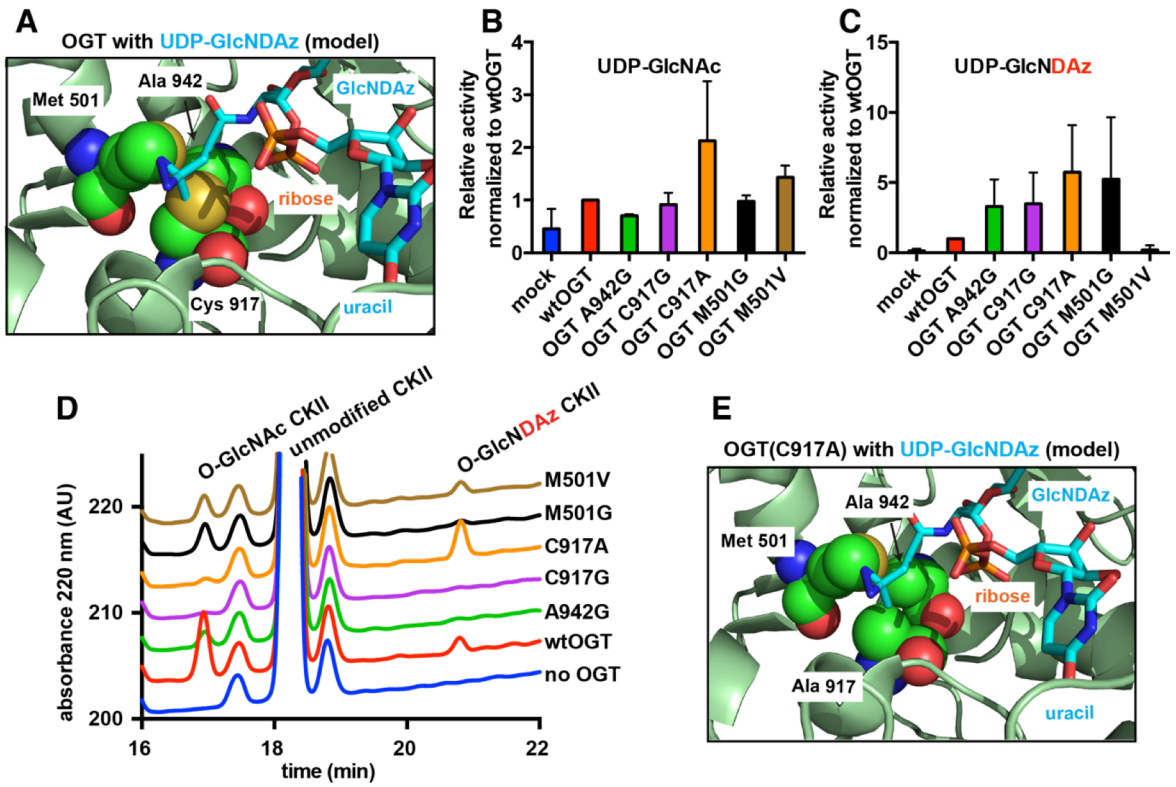

Figure 4.4. Mutagenesis of OGT to accommodate diazirine modification to UDP-GlcNAc. *A,* A model of human OGT (PBD:4GZ5) (24) complexed with UDP-GlcNDAz substrate shows the diazirine from the N-acyl group is in close proximity to positions 501, 917, and 942. Immunopurified OGT (wild-type or indicated mutant) was incubated together with 100 μM CKII peptide, and 10 μM UDP-GlcNAc (*B*) or 10 μM UDP-GlcNDAz (*C*). Release of UDP was analyzed by UDP-Glo assay. Averaged values from two independent experiments are shown, with error bars representing the standard error of the mean (SEM). *D,* Immunopurified OGT (wild-type or indicated mutant) was incubated with CKII peptide and an equimolar concentration of UDP-GlcNAc and UDP-GlcNDAz. Products were analyzed by HPLC. Data shown are a single trial representative of three independent experiments. *E*, A model of OGT(C917A) with UDP-GlcNDAz reveals space available to accommodate the extended diazirine functional group.

was examined in complex with UDP-GlcNAc (PDB: 4GZ5) (109) to identify residues that might sterically occlude UDP-GlcNDAz binding. The alkyl diazirine was modeled to extend off the N-acyl chain of GlcNAc. The side chains of Met501, Cys917, and Ala942 were in close proximity to the diazirine group (Fig. 4.4A). Therefore, plasmids encoding myc-tagged OGT mutants M501G, M501V, C917A, C917G, and A942G were prepared recombinantly using mammalian cells. wtOGT and OGT mutants were expressed in HeLa cells and enzymes were isolated by immunoprecipitation. First, I assessed whether the OGT mutants were active by incubating them with UDP-GlcNAc and CKII peptide. Release of UDP was measured by a coupled enzyme assay with luminescence detection (UDP-Glo). All mutants were capable of releasing UDP from UDP-GlcNAc (Fig 4.4B). The same assay was used to evaluate whether the mutants were capable of utilizing UDP-GlcNDAz as a donor. Four of the five mutants (M501G, C917A, C917G, A942G) demonstrated activity with UDP-GlcNDAz, with the M501G and C917A mutants yielding the most release of UDP (Fig. 4.4C). Next, a competition assay was employed to evaluate the ability of immunopurified wtOGT and each OGT mutant to transfer GlcNAc and GlcNDAz to a peptide substrate. Immunopurified wtOGT or OGT mutant was incubated with CKII peptide and an equimolar mixture of UDP-GlcNAc and UDP-GlcNDAz. Production of O-GlcNAc- and O-GlcNDAz-modified CKII was evaluated by HPLC. OGT(C917A) preferentially incorporated O-GlcNDAz on CKII, with little O-GlcNAc-modified CKII detected (Fig. 4.4D). This was

astonishing because I expected the mutant to at best demonstrate no discrimination. A model of OGT containing mutation of Cys917 to Ala was constructed. I observed that this mutant may accommodate the diazirine extension, while still providing structure to the protein and not introducing excess space in the binding pocket around the N-acyl group of GlcNDAz (Fig. 4.4E).

OGT(C917A) prefers UDP-GlcNDAz over UDP-GlcNAc

The ability of purified OGT(C917A) to transfer GlcNAc and GlcNDAz to a peptide was evaluated. Recombinant OGT(C917A) was expressed in *E. coli* and purified (Fig. 4.5A). The enzyme was incubated with CKII peptide and either UDP-GlcNAc or UDP-GlcNDAz. Production of the corresponding O-GlcNAc- or O-

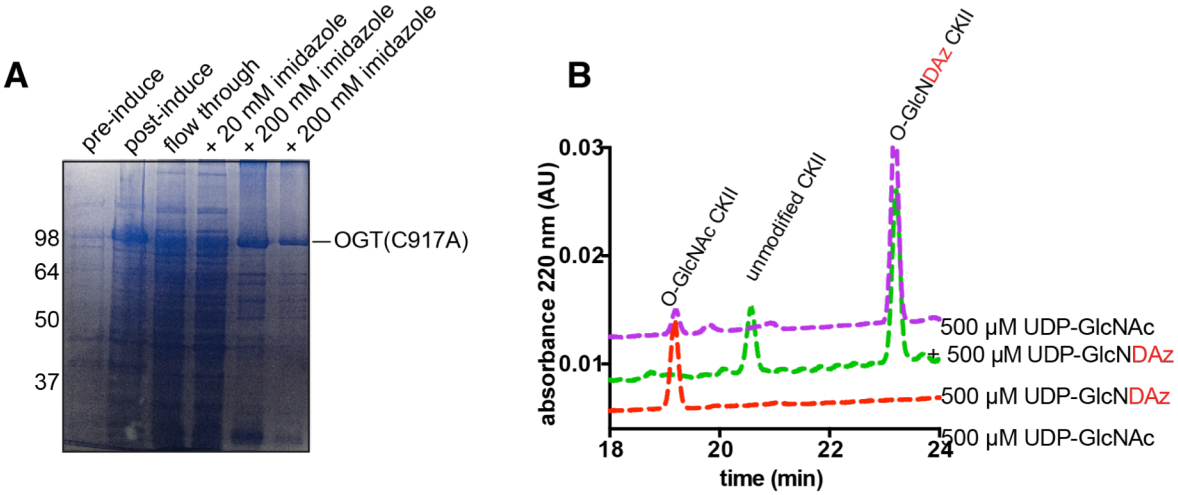

Figure 4.5. OGT(C917A) prefers UDP-GlcNDAz to UDP-GlcNAc. *A,* Coomassie stain of the OGT(C917A) expression and purification. *B,* Glycosylated CKII peptides produced using OGT(C917A). OGT(C917A) (1 μM) was incubated together with CKII peptide (100 μM) and 500 μM UDP-GlcNAc, 500 μM UDP-GlcNDAz, or a mixture of 500 μM UDP-GlcNAc and 500 μM UDP-GlcNDAz. Reaction products were separated by reverse-phase HPLC using UV absorbance (220 nm) for detection. Experiments were performed in triplicate with a single representative trial shown.

GlcNDAz-modified peptides was analyzed by HPLC analysis (red and green traces, Fig. 4.5B). As expected, OGT(C917A) was capable of transferring both GlcNAc and GlcNDAz. Next, I performed a competition experiment, incubating OGT(C917A) with CKII peptide and an equimolar mixture of UDP-GlcNAc and UDP-GlcNDAz. Recombinant bacterially-expressed OGT(C917A), like the immunopurified OGT(C917A), transferred primarily GlcNDAz to the peptide, with very little O-GlcNAc-modified peptide detected (purple trace, Fig. 4.5B). On average, about 8 % O-GlcNAc-ylated peptide was observed and about 92 % O-GlcNDAz-ylated peptide under conditions of equimolar UDP-GlcNAc and UDP-GlcNDAz. This result confirmed that OGT(C917A) preferred UDP-GlcNDAz to UDP-GlcNAc.

Kinetic parameters reveal wtOGT and OGT(C917A) have different substrate specificities

To further characterize the nucleotide sugar preferences of wtOGT and OGT(C917A), the Michaelis-Menten parameters were determined for each enzyme with UDP-GlcNAc and with UDP-GlcNDAz (Fig. 4.6 and Table 4.1). The CKII peptide was used as the acceptor substrate in this analysis. Values obtained for wtOGT with UDP-GlcNAc were similar to those reported by others (20). As predicted from the competition analysis, GlcNDAz was transferred less efficiently than GlcNAc (Fig. 4.6A, 4.6B, 4.6C), resulting in a 5-fold decrease in k_{cat} and a greater than 2-fold increase in K_m for the modified substrate. The kinetics of OGT(C917A) were examined with UDP-GlcNAc and UDP-GlcNDAz substrates.

OGT(C917A) exhibited reduced activity with the natural substrate, UDP-GlcNAc, as compared to wtOGT (Fig. 4.6A and 4.6D). This difference in activity could be attributed to a 3-fold reduction in k_{cat}, as well as greater than 7-fold increase in K_m (Table 4.1). To my delight, OGT(C917A) was demonstrated to transfer GlcNDAz

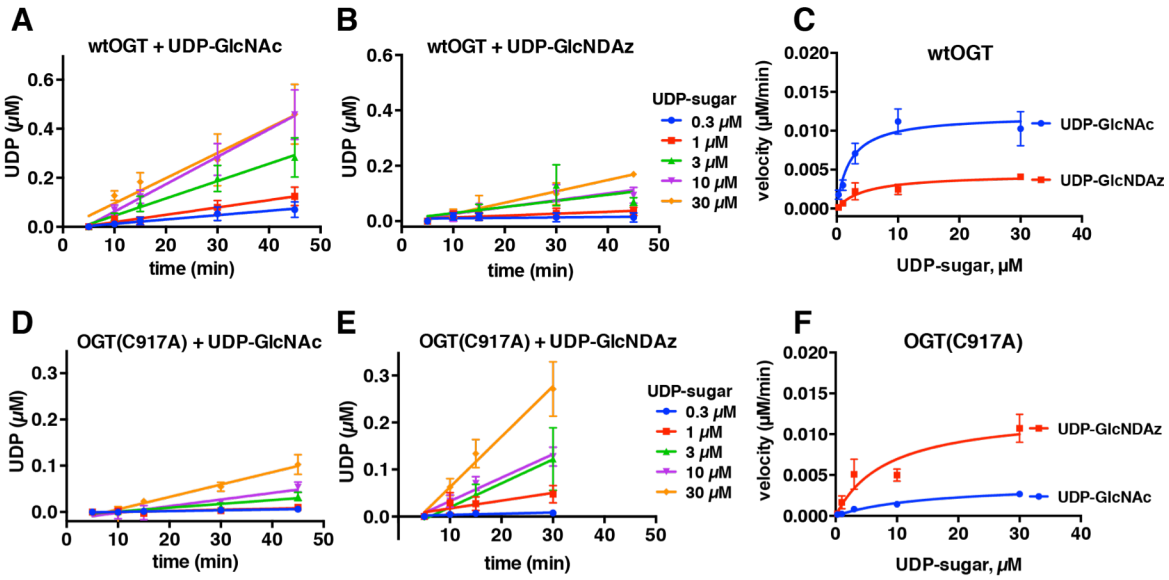

Figure 4.6. wtOGT and OGT(C917A) have reversed nucleotide sugar specificities. *A,* Initial rates of UDP production by wtOGT with various concentrations of UDP-GlcNAc (0.3, 1, 3, 10, and 30 μM). *B,* Initial rates of UDP production by wtOGT with various concentrations of UDP-GlcNDAz (0.3, 1, 3, 10, and 30 μM). *C,* The velocities of wtOGT with UDP-GlcNAc (blue circles) and UDP-GlcNDAz (red squares). *D,* Initial rates of UDP production by OGT(C917A) with various concentrations of UDP-GlcNAc (0.3, 1, 3, 10, and 30 μM). *E,* Initial rates of UDP production by OGT(C917A) with various concentrations of UDP-GlcNDAz (0.3, 1, 3, 10, and 30 μM). *F,* The velocities of OGT(C917A) with UDP-GlcNAc (blue circles) and UDP-GlcNDAz (red squares). Averaged values for Michaelis-Menten analysis were obtained from at least five independent experiments. Error bars represent standard error of the mean (SEM).

efficiently, with both an increased k_{cat} and decreased K_m as compared to the natural UDP-GlcNAc substrate (Fig. 4.6D, 4.6E, and 4.6F). Indeed, introduction of the

	UDP-GlcNAc			UDP-GlcNDAz		
Enzyme	**k_{cat} (min^{-1})**	**K_m (μM)**	**k_{cat}/K_m (μM^{-1} min^{-1})**	**k_{cat} (min^{-1})**	**K_m (μM)**	**k_{cat}/K_m (μM^{-1} min^{-1})**
Wild-type	0.48 (0.05)	2.1 (0.8)	0.22 (0.09)	0.09 (0.01)	4.7 (2.3)	0.019 (0.009)
C917A	0.16 (0.02)	15.4 (4.6)	0.010 (0.003)	0.51 (0.14)	8.5 (6.3)	0.060 (0.047)

Table 4.1. Kinetics of wtOGT and OGT(C917A) with UDP-GlcNAc and UDP-GlcNDAz. Initial rates of reaction (shown in Fig. 4.6) were analyzed to determine enzyme velocities. Kinetic parameters were obtained by non-linear regression analysis, fitting the data to the Michaelis-Menten equation. Parameters are expressed as the mean (standard error) as determined from at least five independent experiments.

C917A mutation resulted in an enzyme whose catalytic efficiency (k_{cat}/K_m) with UDP-GlcNDAz was about four times the value observed for wtOGT with the natural substrate, UDP-GlcNAc. Overall, the kinetic analysis shows that OGT(C917A) transfers GlcNDAz more efficiently than GlcNAc.

Increased GlcNDAz transfer leads to enhanced crosslinking in a cell free system

Encouraged by the observation that OGT(C917A) preferred UDP-GlcNDAz to UDP-GlcNAc, I investigated whether the altered substrate preference led to enhanced O-GlcNDAz-mediated crosslinking. For the initial analysis, a cell-free model system was exploited in which O-GlcNDAz-mediated crosslinking was evaluated between glycosylated CKII peptide and an O-GlcNAc recognizing antibody, RL2. To facilitate detection of the crosslinked complex, a biotin tag in the acceptor CKII peptide was incorporated. Glycosylation of the biotinylated CKII

peptide was similar to what was observed for the unmodified CKII peptide (Fig. 4.3B and 4.5B), with wtOGT preferentially transferred GlcNAc and OGT(C917A) preferentially transferred GlcNDAz (Fig. 4.7A). Indeed, when the biotinylated CKII

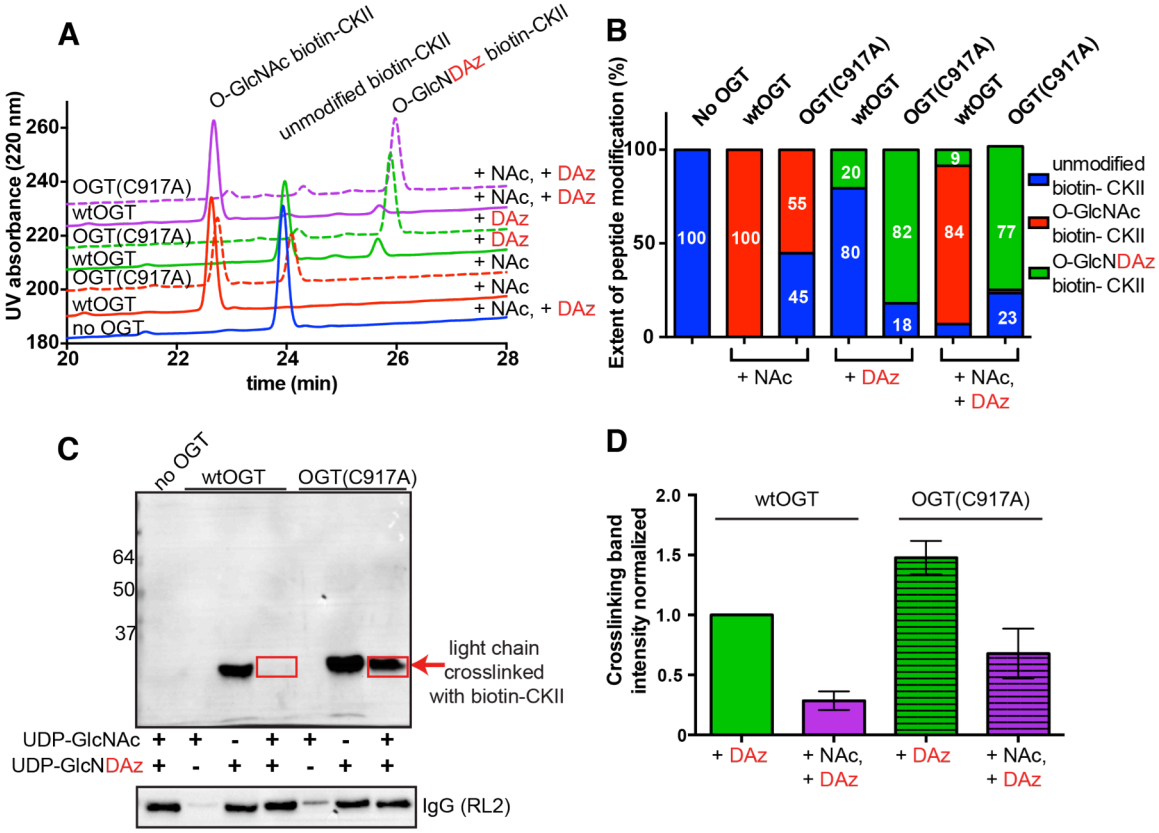

Figure 4.7. OGT(C917A) increased peptide crosslinking in a cell free system. *A,* Biotinylated CKII peptide was glycosylated by wtOGT or OGT(C917A) using UDP-GlcNAc alone (+ NAc), UDP-GlcNDAz alone (+ DAz) or an equimolar mixture of UDP-GlcNAc and UDP-GlcNDAz (+ NAc + DAz). Reaction products were analyzed by HPLC. *B,* Quantification of the ratios of unmodified, O-GlcNAc-modified, and O-GlcNDAz-modified peptides produced in panel *A*. *C,* Peptides produced in panel *A* were incubated with RL2 antibody, UV irradiated, and analyzed by immunoblot using an anti-biotin antibody. *D,* Average RL2-peptide crosslinking intensities for replicate experiments as in panel *C.* Crosslinking intensity in the sample containing wtOGT and UDP-GlcNDAz was arbitrarily set to 1, and other samples were normalized relative to this value. Three independent experiments were performed, with error bars representing SEM.

peptide was incubated with wtOGT and a 1:1 mixture of UDP-GlcNAc and UDP-GlcNDAz, about 84 % of the peptide was glycosylated with O-GlcNAc. In contrast, incubation of the peptide with OGT(C917A) and a 1:1 mixture of UDP-GlcNAc and UDP-GlcNDAz resulted in about 77 % of the peptide glycosylated with O-GlcNDAz (Fig. 4.7B). After removing wtOGT or OGT(C917A) from the reaction mixture, the resulting crude mixtures were incubated with RL2 antibody, and crosslinked by application of UV radiation. Crosslinking between the peptide and RL2 antibody was revealed by immunoblot analysis. The molecular weight corresponding to the light chain of RL2 antibody, as the Kohler lab has observed previously (47), was where anti-biotin reactivity appeared. As expected, reaction mixtures containing only UDP-GlcNAc yielded no O-GlcNDAz-modified peptide and no crosslinking to RL2 antibody. However, the reaction mixtures containing UDP-GlcNDAz generated O-GlcNDAz-ylation that crosslinks to RL2 (Fig. 4.7C). Notably, when the reaction mixtures contained equimolar UDP-GlcNAc and UDP-GlcNDAz, crosslinking was observed when OGT(C917A) was used, but not when wtOGT was used (red boxes, Fig. 4.7C). Furthermore, even when UDP-GlcNDAz was the only nucleotide sugar used, more crosslinking was observed from reactions with OGT(C917A) than from reactions with wtOGT (Fig. 4.7D). In a cell free assay, these results demonstrated the C917A mutation led to increased O-GlcNDAz-modified peptide, which in turn produced higher levels of O- GlcNDAz-mediated crosslinking. The OGT(C917A) will

benefit the cell free assay as it will decrease reaction time compared to wtOGT and also increase O-GlcNDAz-ylation yield compared to wtOGT.

Increased GlcNDAz transfer leads to enhanced crosslinking in cells

Next, I wished to determine if the C917A mutation to OGT could lead to increased O-GlcNDAz crosslinking in cells, which could improve our cell-based assay. Our O-GlcNDAz crosslinking method relies on the F383G mutant of UAP1 to catalyze in-cell production of UDP-GlcNDAz from cell-permeable GlcNDAz-1P (Fig. 4.8A). I prepared K562 cells stably expressing UAP1(F383G) and then cultured them with cell-permeable GlcNDAz-1P. The lysate was analyzed by HPAEC, revealing robust production of UDP-GlcNDAz (Fig. 2.3B). The UAP1(F383G)-expressing K562 cells were nucleofected with plasmid DNA encoding wtOGT or OGT(C917A) and then cultured with cell-permeable GlcNDAz-1P for 7 h. The intact cells were subjected to UV irradiation and then lysed. The lysates were analyzed by immunoblot, probing for NUP153, a protein that is known to be heavily O-GlcNAc-modified and for which we have previously observed O-GlcNDAz-mediated crosslinking (47). In cells transfected with both wtOGT and OGT(C917A), an increase of anti-NUP153 reactivity at higher apparent molecular weight was observed (red boxes, Fig. 4.8B). For both wild-type and mutant OGT, appearance of these high molecular weight species depended on both UV irradiation and inclusion of cell-permeable GlcNDAz-1P. Therefore I believed that these bands correspond to O-GlcNDAz-crosslinked forms of NUP153. The analysis was conducted four times,

each time quantifying of the reactivity in the crosslinked region, and normalizing

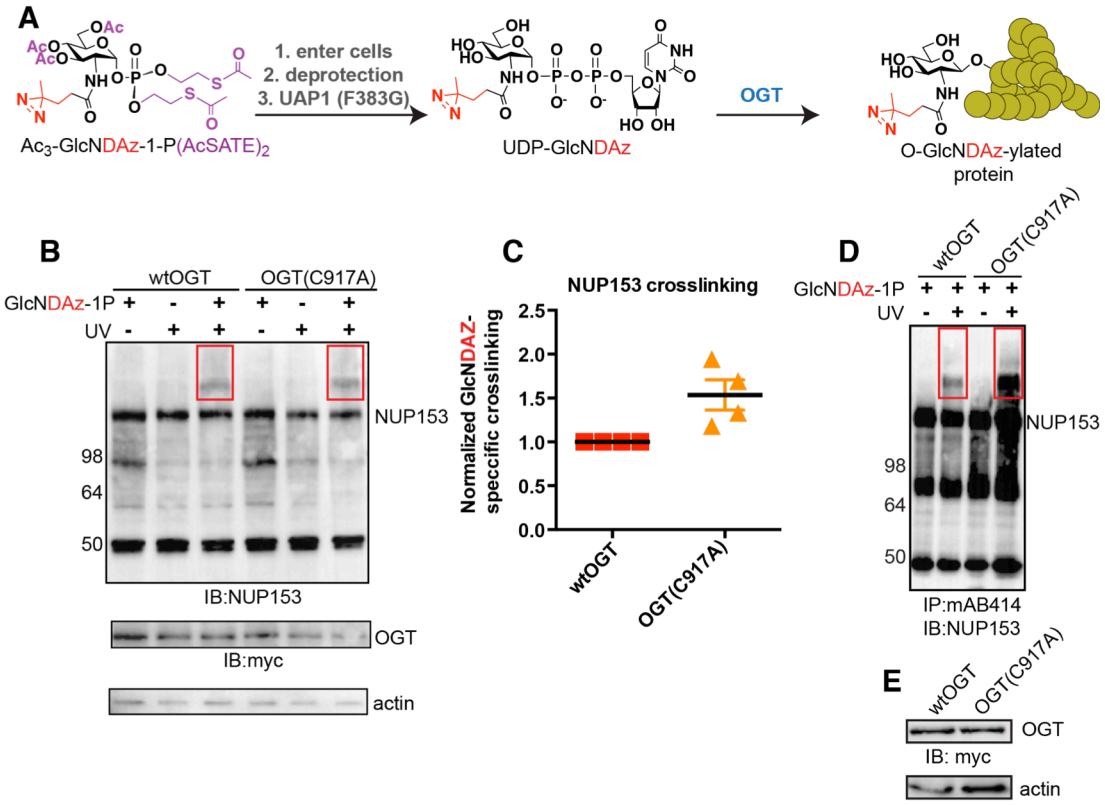

Figure 4.8. Expression of OGT(C917A) in cells results in enhanced O-GlcNDAz-dependent crosslinking. *A,* K562 cells expressing UAP1(F383G) (K562(UAP1(F383G))) were cultured with cell-permeable GlcNDAz-1P, which is deprotected inside of cells. This yields GlcNDAz-1P, and is activated to UDP-GlcNDAz by UAP1(F383G). OGT transfers GlcNDAz to produce O-GlcNDAz-modified proteins. *B,* K562 UAP1(F383G) were nucleofected with plasmid encoding wtOGT or OGT(C917A) were cultured with or without cell-permeable GlcNDAz-1P for 7 h and then irradiated for 15 min. Immunoblot analysis was used to detect crosslinking of NUP153. Red boxes highlight putative crosslinked complexes. *C,* The extent of NUP153 crosslinking was analyzed in four independent experiments, as in *B*. The amount of crosslinking observed in cells nucleofected with wtOGT was arbitrarily set to 1 and the amount of crosslinking in cells nucleofected with OGT(C917A) was normalized relative to this control. Error bar represents standard error of mean (SEM). *D,* K562 UAP1(F383G) was same condition as *B*. After lysis, nucleoporins were purified using mAb414. The immunoprecipitate was probed by immunoblot using anti-NUP153. Red boxes highlight putative crosslinked complexes. Data presented are representative results from two independent trials. *E,* Anti-myc immunoblot revealed comparable expression of wtOGT and OGT(C917A) in nucleofected K562 cells used for immunoprecipitation experiment shown in *D*.

these values relative to an actin loading control. In each experiment, more crosslinked material was observed in the cells transfected with OGT(C917A) as compared to cells transfected with wtOGT. The increases ranged from 18 % to 94 % (Fig. 4.8C). Additionally, I investigated if overexpression of OGT(C917A) would allow for isolation of more crosslinked material from cells. Again, UAP1(F383G)-expressing K562 cells were nucleofected with DNA encoding wtOGT or OGT(C917A), then cultured with cell-permeable GlcNDAz-1P and UV irradiated. I immunoprecipitated nucleoporins using mAb414 antibody, which recognizes multiple NUPs, including NUP153. Immunoblot analysis using anti-NUP153 revealed substantially more crosslinked NUP153 in the immunoprecipitate from cells transfected with OGT(C917A) as compared to cells transfected with wtOGT (Fig. 4.8D). Because the overexpression levels of wtOGT and OGT(C917A) were comparable (Fig. 4.8E), it was concluded that increased O-GlcNDAz production by OGT(C917A) enhanced O-GlcNDAz-mediated crosslinking in cells and enables isolation of more crosslinked material for downstream analysis.

DISCUSSION

C917A mutation to OGT reverses UDP-GlcNAc/UDP-GlcNDAz substrate specificity.

Here I described the use of site-directed mutagenesis to improve the O-GlcNDAz photocrosslinking method for discovering O-GlcNAc-mediated protein-protein interactions (47). Using a *cell free* competition assay, wtOGT demonstrated a

preference to modify the CKII peptide with O-GlcNAc, rather than O-GlcNDAz. This result was confirmed by kinetic analysis showing that UDP-GlcNAc is the preferred nucleotide sugar substrate for OGT. Indeed, inspection of a model of wtOGT complexed with UDP-GlcNDAz revealed a potential for steric conflict between the diazirine modification and the amino acid side chains in wtOGT (Fig. 4.4A). I speculated that wtOGT discrimination against UDP-GlcNDAz is an important factor limiting how much O-GlcNDAz crosslinked material can be isolated for mass spectrometry analysis.

OGT mutants were screened to identify an enzyme capable of efficient O-GlcNDAz production. While most mutants showed increased activity with UDP-GlcNDAz, the M501V mutant had very little activity with the modified nucleotide sugar. This result is consistent with a previous study in which the M501V mutation was introduced into the short isoform of OGT (sOGT). Ma *et al.* prepared sOGT(M501V) with the goal of enhancing activity toward UDP-GlcNAc analogs with extended acyl chains, but found that this mutant exhibited reduced activity toward the analogs (111). While the M501V mutation did not appear to be beneficial for GlcNDAz transfer, I found that OGT(C917A) showed enhanced O-GlcNDAz production, which in turn enabled increased O-GlcNDAz-mediated crosslinking both in cell free (Fig. 4.7) and in cell-based assays (Fig. 4.8C, and 4.8E). While the OGT(C917A) mutant had previously been shown to be active (112), I was thrilled and amazed to find that it transferred GlcNDAz better than GlcNAc. I speculated that

the diazirine extension of UDP-GlcNDAz engaged in productive interactions with OGT(C917A) that led to one or both of the following: improved substrate binding and promotion of a catalytically active conformation.

OGT(C917A) may also aid in the transfer of other GlcNAc analogs. Indeed, a variety of unnatural UDP-GlcNAc analogs with extended N-acyl side chains have been prepared, but wild-type OGT tended to discriminate against those that add steric bulk close to the N-acyl group (111, 113, 114). The C917A mutation might facilitate transfer of these bulkier GlcNAc analogs. Variation in the N-acyl side chain of UDP-GlcNAc also occurs in nature: Vocadlo and co-workers showed evidence for O-GlcNGc production, suggesting that OGT accepts UDP-GlcNGc (86). Similarly, small unnatural N-acyl substituents, such as the azide and alkyne, seem to be tolerated by OGT,(27, 29, 115, 116) but it is possible that the C917A mutation could also facilitate discrimination between smaller analogs and the natural substrate. Overall, the observed utility of OGT(C917A) for O-GlcNDAz production adds to other studies showing that metabolic oligosaccharide engineering can be augmented through mutagenesis of carbohydrate transforming enzymes to enhance metabolism of the unnatural sugar analogs (47, 110).

Extension of the GlcNDAz technology

For future work of the OGT(C917A), I desire to utilize the mutant in conjugation with a chemo-enzymatic approach for O-GlcNDAz production. I hypothesize that using OGT(C917A), rather than wtOGT, will allow me to prepare O-

GlcNDAz-ylated efficiently and in good yield. Additionally, an increase in the amount of glycosylated peptide will provide the ability to overcome the challenge of not having enough crosslinking material available for proteomic analysis. Moreover, the ability to glycosylate peptides in a chemo-enzymatic method will advance our knowledge in knowing the specific sequence necessary for interactions, as well as providing a more versatile technology for others.

Another possible way to improve the cell-based assay is to utilize the OGT(C917A) in cells to obtain more crosslinking. The cell-based crosslinking data reveals that expression of OGT(C917A) results in enhanced O-GlcNDAz crosslinking, which was demonstrated both cell free and in cells. In cells, the naturally occurring O-GlcNAc modification is dynamic and is hydrolyzed by OGA. As described previously, the wild-type OGA is inactive towards GlcNDAz (**chapter 2**) (47, 110). Thus, O-GlcNDAz likely persists in cells and expression of OGT(C917A) will act to enhance its accumulation, which may be toxic to the cells and/or interfere with the homeostasis. It may be desirable to convert O-GlcNDAz from a static modification to a dynamic one, to better mimic the natural O-GlcNAc modification. This could be accomplished by expressing an O-GlcNDAz-hydrolyzing mutant of OGA, OGA(C215A) (110) along with the O-GlcNDAz-producing form of OGT, OGT(C917A). I envision that creating cell lines that are optimized for O-GlcNDAz metabolism will improve the cell-based assay.

ACKNOWLEDGEMENTS

I thank Drs. Suzanne Walker (Harvard Medical School) and Carolyn Bertozzi (UC Berkeley) for plasmids that encode OGT, Dr. Ilmin Kwon (UT Southwestern Medical Center) for assistance with OGT purification, Dr. Kenneth Westover (UT Southwestern Medical Center) for equipment and advice with OGT purification, and Drs. Suong Nguyen and Oleg Volkov (UT Southwestern Medical Center) for assistance with enzyme kinetics analysis. I acknowledge support from the NIH (R01GM090271) and the Welch Foundation (I-1686). This research was originally published in Journal of Biological Chemistry. Rodriguez AC, Yu S-H, Li B, Zegzouti H & Kohler JJ. Enhanced transfer of a photocrosslinking O-GlcNAc analog by an O-GlcNAc transferase mutant with converted substrate specificity. *Journal of Biological Chemistry*. 2015; 290. 22638-22648.

CHAPTER FIVE:

Development of a chemo-enzymatic synthesis for glycosylating peptides in cell free system

INTRODUCTION

O-linked β-N-acetylglucosamine-ylation (O-GlcNAc-ylation) is a unique post-translational modification (PTM) that has multiple functional roles, such as transcriptional regulation and cell signaling (5). Since O-GlcNAc is transient, it is difficult to study. Therefore, novel technologies are required to advance the field and to study the roles of O-GlcNAc (49).

One under-studied function of O-GlcNAc is that it may mediate protein-protein interactions, like CBP and Stat5 (44) and NUP62 and NUP88 (45). Due to the issues in capturing these interactions, the Kohler lab developed a photocrosslinking GlcNAc, GlcNDAz, to capture binding partners of O-GlcNAc-modified proteins. The photocrosslinking handle addresses several characteristic issues of O-GlcNAc, like substoichiometric and dynamics. For example, Hart and co-workers revealed low levels of O-GlcNAc on proteins. They determined the substoichiometric amount of 0.1 GlcNAc/protein by radiolabeling (2). With O-GlcNDAz, only one modification is required to capture the interaction. This advantage allows for low abundance O-GlcNAc-modified proteins to be studied. Since O-GlcNAc cycles much faster than protein turnover, O-GlcNDAz captures the dynamic interactions between O-GlcNAc-modified proteins and

binding partners occurring at the time of irradiation. Therefore, O-GlcNDAz assists in the identification of binding partners interacting with an O-GlcNAc-modified protein of interest (47).

Previously, the Kohler lab developed a cell-based O-GlcNDAz crosslinking method to substitute O-GlcNDAz onto O-GlcNAc-modified proteins using a cell-permeable GlcNDAz-1P, Ac_3-GlcNDAz-1P(Ac-SATE)$_2$. Cells overexpressing the UDP-N-acetylglucosamine pyrophosphorylase 1 (UAP1), in which phenylalanine 383 is mutated to glycine (F383G), UAP1(F383G), converted GlcNDAz-1P into the nucleotide donor sugar, UDP-GlcNDAz. GlcNDAz was transferred onto O-GlcNAc-modified proteins employing O-GlcNAc transferase (OGT). Upon UV irradiation, the diazirine was activated and reacted to any neighboring molecule. This result led to a covalent bond between O-GlcNDAz-modified protein and the binding partner, which was coined as the crosslinked complex. The complexes were immunoprecipitated with an antibody that recognizes the O-GlcNAc-modified nucleoporin NUP153, a phenylalanine-glycine repeat nucleoporin (FG NUP). Through proteomic analysis, NUP153 was discovered to interact with several karyopherins, including exportin-1, transportin-1, and karyopherin β1 (47).

Two reports suggested O-GlcNAc mediates protein-protein interactions (44, 45). My initial biological question was to utilize the cell-based O-GlcNDAz crosslinking method, and identify the binding partners of another FG NUP, NUP98 and NUP98 leukemogenic fusions, as discussed in **chapter 2**. In cervical cancer and leukemia cells, the data suggested a novel finding that the leukemogenic fusions were O-GlcNAc-modified. Additionally, evidence proposed NUP98 and NUP98 fusions exploited O-

GlcNAc-ylation at or near the site of protein-protein interactions (Fig. 2.4, and 2.5). I desired to identify and study of these interactions to gain insight into how cells transform from normal homoeostasis to disease state through leukemogenic fusions. However, a challenge in identifying the binding partners became evident, as isolation of crosslinked complexes proved difficult. I attempted to enrich for the crosslinked complexes with HA-tagged or avi-tagged of the NUP98 or NUP98 fusions. Yet, the experimental results determined the attempts to purify the complexes were unsuccessful (Fig. 2.6). The purification of the crosslinked complexes is crucial to identifying the binding partners of O-GlcNAc-modified proteins.

Aside from the challenge of enriching crosslinked complexes, other obstacles arose when utilizing the cell-based O-GlcNDAz crosslinking method. The cell-based O-GlcNDAz crosslinking method allows for the identification of binding partners, but the exact O-GlcNAc site responsible for the mediation remains uncertain. Additionally in personal communication with Kohler lab members, the cell-based method requires a large amount of crosslinked material for proteomic analysis. This became problematic because the cell-based O-GlcNDAz crosslinking method didn't yield high amounts of crosslinked complexes (47). Moreover, in personal communications with Kohler lab members, it became clear that the employment of a highly stringent washing method was crucial to assist in the identification of the potential binding partners for other proteins of interest in the lab, like cholera toxin (69). Attempts to biotin tag NUP98 and the NUP98 fusions were unsuccessful, using avi-tag technology (Fig. 2.6B). Lastly, the Kohler lab previously demonstrated that O-GlcNDAz is a static modification (47, 110).

This issue may led to toxicity and deregulation of O-GlcNAc homeostasis, that in turn can affect the O-GlcNAc-modified proteins' interactions.

In efforts to overcome these challenges, I developed tools that will be employed for a complementary cell free O-GlcNDAz crosslinking method. I examined the activity of a mutant form of O-GlcNAc transferase (OGT), which was discussed in **chapter 4**, to increase the yield of crosslinked material (68). Furthermore, I examined a mutant form of O-GlcNAc hydrolase to restore the dynamics of O-GlcNAc, which was discussed in **chapter 3** (110). These OGT and OGA mutants may address some hindrance from the cell-based O-GlcNDAz crosslinking method.

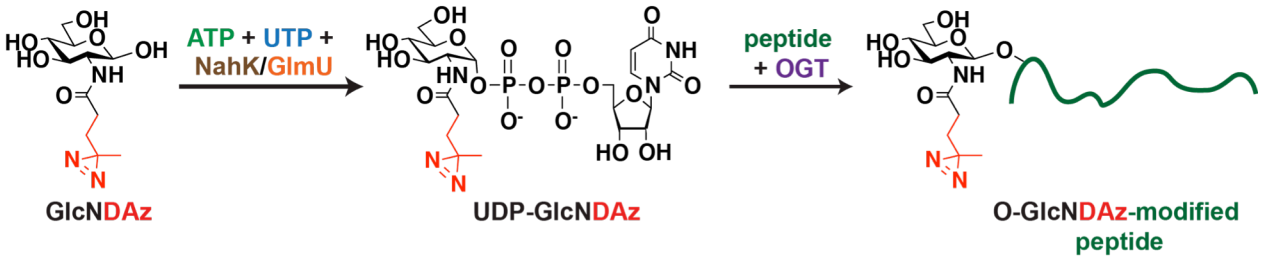

Figure 5.1. Chemo-enzymatic synthesis to develop O-GlcNDAz-modified peptides. Free GlcNDAz is incubated with ATP, UTP, and NahK/GlmU to generate UDP-GlcNDAz. The nucleotide donor sugar, UDP-GlcNDAz is incubated with OGT(C917A) to O-GlcNDAz-modified biotin-tagged peptides.

Another issue with the cell-based O-GlcNDAz crosslinking method is the uncertainty of whether the O-GlcNAc is necessary for mediating protein-protein interactions. In the cell-based method, endogenous proteins are crosslinked but the Kohler lab did not demonstrate the O-GlcNAc site responsible for the protein-protein interaction. As an alternative approach to proteins, peptides derived from O-GlcNAc-modified proteins will be employed in a chemo-enzymatic synthesis with UDP-GlcNDAz

and OGT(C917A), a mutant form of OGT that favors O-GlcNDAz (Fig. 5.1). The peptide approach allows for the recognition of the sequence necessary for interaction and possibly the specific residue that is O-GlcNAc modified. Furthermore, biotin will be incorporated onto the peptides to enable highly stringent washes for a more specific candidate list. However, one challenge with the chemo-enzymatic synthesis is the low amounts of UDP-GlcNDAz available.

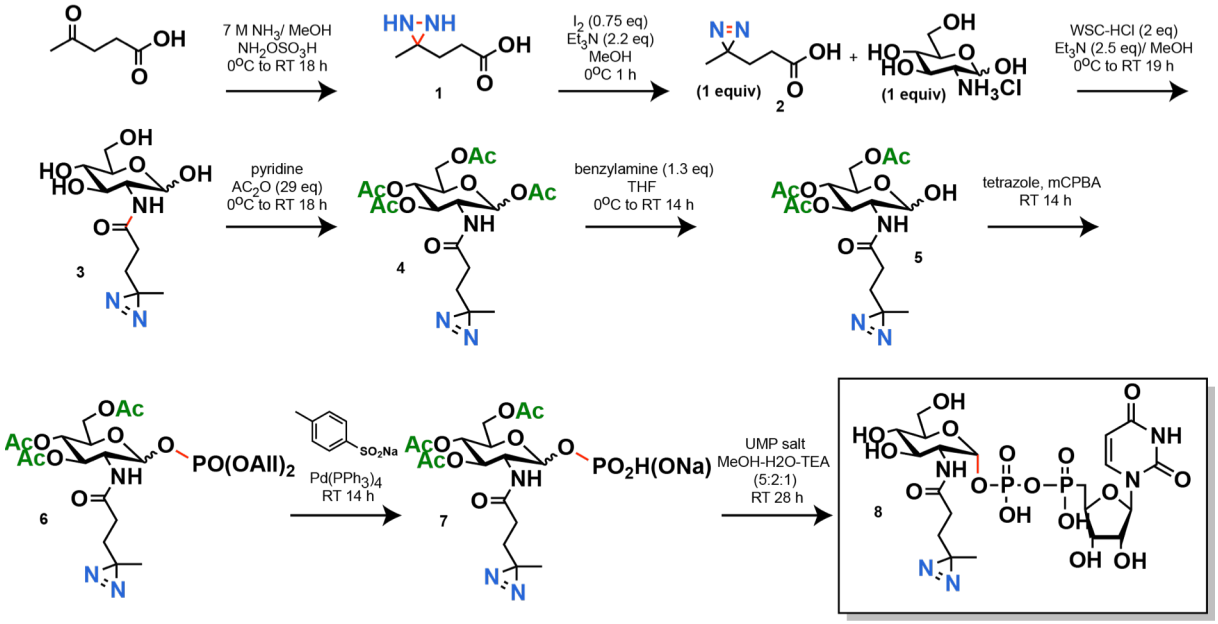

Figure 5.2. Chemical synthesis of UDP-GlcNDAz (8). There are eight chemical reactions necessary to generate the nucleotide donor sugar, UDP-GlcNDAz. This process is time consuming with severely difficult purification steps. Also, the resulting product (8) is generated in low yields. Additionally, it should be noted that the purity of the final product (8) is severe as the OGT reaction is inactive with impurities present.

Typically, the route of chemical synthesis of UDP-GlcNDAz, like many other nucleotide sugar syntheses, is difficult and results in low yields (Fig. 5.2) (47). However, an alternative approach to synthesize nucleotide sugars was established with enzymes via chemo-enzymatic synthesis (117, 118). Two enzymes involved in the UDP-GlcNAc

production synthesized UDP-GlcNAc in a cell free system with moderate activity and in reasonable yield (119). N-acetylhexosamine 1-kinase (NahK) derived from *Bifidobacterium longum* converts GlcNAc into GlcNAc-1P using adenosine triphosphate (ATP) (Fig. 5.3A). In conjunction with NahK, N-acetylglucosamine uridyltransferase (GlmU) from *Escherichia coli* can transfer uridine diphosphate (UDP) onto GlcNAc-1P. Liu, Wang, and co-workers demonstrated that NahK fused to the catalytic domain of GlmU was able to generate GlcNAc and several analogs to the UDP-sugar (119). With the OGT(C917A), NahK/GlmU will assist in the chemo-enzymatic synthesis of O-GlcNDAz-modified peptides.

The employment of O-GlcNDAz-modified peptides for crosslinking will surpass the apparent challenges. Here I described my chemo-enzymatic synthesis efforts to glycosylate peptides starting with free GlcNDAz and peptide. First, I exploited NahK/GlmU, a bifunctional enzyme to chemo-enzymatically synthesize UDP-GlcNDAz in large yield amounts and in fewer steps than chemical synthesis. In conjugation with OGT, I employed NahK/GlmU and OGT to generate O-GlcNAc-ylated, biotin-tagged peptide in a one-pot reaction. I attempted to generate a one-pot reaction for O-GlcNDAz-ylation of biotin-tagged peptides, but was unsuccessful. My efforts revealed that a multi-step synthesis method with NahK/GlmU and OGT(C917A) can O-GlcNDAz modify peptides in a cell free system. This multi-step chemo-enzymatic synthesis protocol was examined on other O-GlcNAc-modified proteins that could be potentially used to identify binding partners. The development of a cell free O-GlcNDAz

crosslinking method will overcome challenges of the cell-based method, and provide a versatile technology.

EXPERIMENTAL PROCEDURES

Chemicals and Peptides

GlcNAc was purchased from Sigma (cat. no. U4375). GlcNDAz was synthesized as previously described (47). UDP-GlcNDAz was previously synthesized and characterized (47). Peptides were purchased through GenScript. Casein Kinase II (CKII) was synthesized and had the sequence PGGSTPVSSANMMK with N-terminal acetylation or biotinylation and C-terminal amidation. Human α A-crystallin +P2 was synthesized and sequenced YAIPVSPREK with N-terminal acetylation and C-terminal amidation. Insulin Receptor Substrate 1 (IRS1) was synthesized and sequence left confident due to collaboration with Lauren Ball (Medical University of South Carolina).

Construction of NahK/GlmU mutant

The pET15b plasmid encoding trGlmU-NahK was provided by Xian-wi Liu (Shandong University) and used to express recombinant OGT in *E. coli (119)*. Point mutation was introduced into NahK/GlmU plasmid using primers designed through the Sigma primer design algorithm. Mutagenesis was performed according to the QuikChange protocol (Stratagene). The following PAGE purified primers were used: Y197A forward 5'-TAATGCTCAGGGCGAATACGCCATCACCGACATTATTGCG

-3', and reverse 5'-CGCAATAATGTCGGTGATGGCGTATTCGCCCTGAGCATTA-3'. The entire coding sequence of each mutated plasmid was confirmed through sequencing at the UT Southwestern DNA sequencing core facility.

Overexpression of recombinant NahK/GlmU in E. coli and purification

pET15b plasmid encoding NahK/GlmU (119) or NahK/GlmU(Y197A) was used to produce recombinant enzyme as previously described. Briefly, BL21(DE3) competent cells were transformed with each plasmid. A single colony was added to a 5 mL LB culture containing ampicillin and incubated at 37 °C, with shaking for 14 h. Three mL of the starter culture was inoculated in 1 L of LB, and incubated at 37 °C with shaking at 250 rpm until OD_{600} reached 0.8. After cooling the culture to 16 °C, cells were induced with 0.2 mM IPTG for 20 h. Bacteria was harvested by 30 min centrifugation at 1154*g*, 4 °C. Bacteria was lysed in 25 mM Na_2HPO_4, 100 mM NaCl, 14 mM β-mercaptoethanol, 1 mM EDTA and a tablet of protease inhibitor cocktail (Sigma, S8830). Bacteria were sonicated for 5 minutes alternating 10 seconds pulse on and 10 seconds pulse off. Insoluble material was removed by centrifugation at 12,000*g*, 4 °C for 30 min. Cell lysates were subjected for purification using a Profinia instrument (Bio-Rad, 620-1010) with an IMAC column (Bio-Rad, 732-4610) according to the manufacturer's instructions. Buffers for washes, elution, and desalting were used with the recommendations provided by Profinia.

OGT(C917A) was recombinant generated as previously described (68).

Chemo-enzymatic reaction for nucleotide sugar synthesis

Chemo-enzymatic reactions were performed to analyze the ability of the NahK/GlmU (wild-type and indicated mutant) ability to generate UDP-GlcNDAz. The bifunctional enzyme was incubated with 25 mM GlcNDAz, 30 mM ATP, 30 mM UTP with 0.5 mg/mL NahK/GlmU in 50 mM TrisHCl pH 8.0, and 10 mM $MgCl_2$. The reaction was incubated at 40 °C for 1 d (Fig. 5.3). The activity was analyzed by high performance anion exchange chromatography (HPAEC, ICS-3000, Dionex) with a CarboPac™PA1 column (Dionex). The HPAEC analysis was subjected to the following gradient using buffer A (1 mM NaOH) and buffer B (1 mM NaOH, and 1 M NaOAc): T_0= 95% A, 5% B; T_5= 85% A, 15% B; T_{15}= 70% A, 30% B; T_{20}= 60% A, 40% B; T_{45}= 60% A, 40% B; T_{50}= 0% A, 95% B; T_{60}= 0% A, 100% B; T_{65}= 95% A, 5% B; T_{75}= 95% A, 5% B. To determine the concentration of UDP-GlcNDAz production, several concentrations of UDP-GlcNDAz were injected onto the HPAEC in duplicates. The peak area of the assigned UDP-GlcNDAz peak was determined by integration analysis of Chromeleon 7, then quantification of the concentration was determine by a standard curve of increasing concentrations of UDP-GlcNDAz. AU stands for arbitrary units.

Multi-step reaction to generate O-GlcNDAz-modified peptides

An initial glycosylation reaction was conducted in two-step process to determine if NahK/GlmU generates UDP-GlcNDAz and OGT produces an O-GlcNDAz-ylation peptide. The reaction was set up with 25 mM GlcNDAz, 30 mM ATP, 30 mM UTP with 0.6 mg/mL NahK/GlmU and allowed to incubate at 40 °C for 3 d in 50 mM TrisHCl pH

8.0 and 10 mM $MgCl_2$ (Fig. 5.4). Mock reactions mean no enzyme was added, instead the volume equivalence of buffer. The enzyme was filtered through a 10 kDa MWCO Amicon unit (Millipore, UFC5010BK) that was centrifuged at 17000*g* for 10 min at room temperature. Five μL of filtrate was subjected to HPAEC for UDP-GlcNDAz analysis (Fig. 5.4A). The HPAEC solution and gradient program is described above. The remaining filtrate was incubated with 100 μM O-GlcNAc-modified CKII peptide and 1 μM OGT(C917A) at 37 °C for 1 d in 20 mM TrisHCl pH 8.0, 12.6 mM $MgCl_2$, and 20 mM β-mercaptoethanol. The reaction was filtered as described from the previous reaction. The filtrate was subjected to HPLC for O-GlcNDAz-ylation production (Fig. 5.4B). Peptide products were analyzed using a Dionex HPLC Ultimate 3000 with UV detection at 220 nm. The HPLC solvent and gradient program is described previously (68).

This two-step synthesis for O-GlcNDAz-modifying peptides was further analyzed for later experiments. The NahK/GlmU reaction was incubated for 1 d, 2 d, or 3 d with 0.25 mg/mL enzyme at 37 °C (Fig. 5.7). Mock reactions mean no enzyme was added, instead the volume equivalence of buffer. The two-step reaction of 3 d incubation with NahK/GlmU, followed by 1 d incubation with OGT(C917A) was utilized to glycosylate CKII, IRS1, and +P2 peptides with some modifications (Fig. 5.7 and 5.8).

Optimizing for one-pot condition for O-GlcNAc and O-GlcNDAz

The one-pot reaction was optimized for O-GlcNAc-ylation using several buffer conditions. Glycosylation reactions were performed to measure the ability of both NahK/GlmU and OGT in one step to generate UDP-GlcNAc and transfer GlcNAc to

CKII. Purified enzymes were incubated in indicated buffer with 25 mM GlcNAc, 30 mM ATP, 30 mM UTP, 100 μM CKII, 0.5 mg/mL NahK/GlmU and 1 μM wild-type OGT (wtOGT) (Fig. 5.5A). Mock reactions mean no enzyme was added, instead the volume equivalence of buffer. Reactions were also conducted with GlcNDAz and OGT(C917A) in a one-pot condition and without (Fig. 5.6A) or with (Fig. 5.6D) 1 U inorganic pyrophosphatase. The reaction was allowed to incubate at 37 °C for 1 d. After incubation for 24 h, the enzymes were removed by filtration through a 10 kDa MWCO Amicon unit (Millipore, UFC5010BK) that was centrifuged at 17000*g* for 10 min at room temperature. Peptide products were analyzed using a Dionex HPLC Ultimate 3000 with UV detection at 220 nm. Solvent systems and gradients were described previously (68).

As an alternative approach, reactions were performed to measure the ability of glycosylation of CKII when NahK/GlmU and OGT were added in a sequential matter. Free sugar (25 mM) GlcNDAz, 30 mM ATP, 30 mM UTP, and 0.5 mg/mL NahK/GlmU was incubated for 24 h at 37 °C without (Fig. 5.6B) or with (Fig. 5.6E) 1 U inorganic pyrophosphatase present. Then, 100 μM CKII, and 1 μM OGT(C917A) was added to the reaction. The reaction was allowed to incubate at 37 °C for 1 d. After incubation for 24 h, the enzymes were removed by filtration through a 10 kDa MWCO Amicon unit (Millipore, UFC5010BK) that was centrifuged at 17000*g* for 10 min at room temperature. Peptide products were analyzed by HPLC as mentioned above.

RESULTS

Mutation of NahK/GlmU to increase UDP-GlcNDAz production

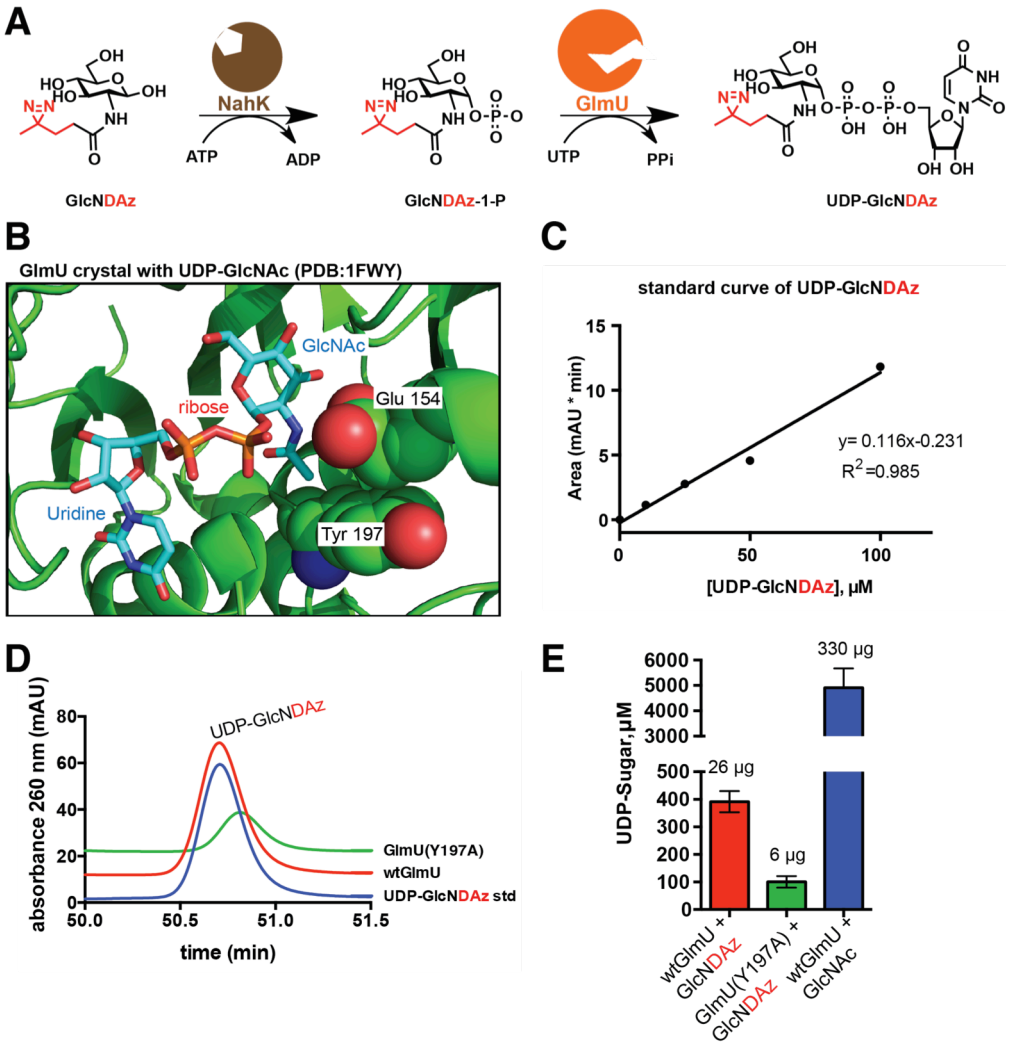

Figure 5.3. NahK/GlmU generates UDP-GlcNDAz. *A,* Reaction scheme of the bifunctional NahK/GlmU enzyme to generate UDP-GlcNDAz through chemo-enzymatic synthesis. Synthesized GlcNDAz is converted to GlcNDAz-1-P with ATP and NahK. UTP and GlmU convert GlcNDAz to UDP-GlcNDAz. *B,* Structure of GlmU co-crystallized with UDP-GlcNAc (PDB:1FWY). The tyrosine 197 is in close proximity to the N-acetyl group of GlcNAc. *C*, Standard curve generated to determine the concentration of UDP-GlcNDAz yielded from the bifunctional enzyme reaction based on the peak area. *D*, HPAEC trace of the bifunctional enzyme reaction with wild-type GlmU or GlmU(Y197A). Data shown is a representation of a single trial. *E*, Quantification of the amounts of UDP-GlcNAc (blue) and UDP-GlcNDAz (red and green) generated from wild-type GlmU or GlmU(Y197A).

The bifunctional enzyme, NahK/GlmU, was recombinantly generated through expression with *E. coli*, and followed by purification for His-tagged proteins. The enzyme was exploited to generate UDP-GlcNDAz (Fig. 5.3A). In efforts to increase the production of UDP-GlcNDAz, the crystal structure of GlmU complexed with UDP-GlcNAc (PDB:1FWY) was examined (Fig. 5.3B) (120). Based on the location of the N-acyl group of GlcNAc, I believe there may be steric hindrance for the diazirine extension within the binding pocket of GlmU. Mutation of the tyrosine 197 to alanine may accommodate the diazirine, and result in increased yields of UDP-GlcNDAz. High performance anion exchange chromatography (HPAEC) analysis was performed to measure production of UDP-GlcNDAz. As a control wild-type GlmU was incubated with the natural substrate, GlcNAc, and generated 330 μg UDP-GlcNAc (purple, Fig. 5.3E). While GlmU(Y197A) employment demonstrated low levels of UDP-GlcNDAz assembly of 6 μg, the wild-type GlmU exposed higher yield of product of 26 μg (Fig. 5.3E). The low generation of UDP-GlcNDAz may be due to GlmU(Y197A) being misfolded because tyrosine is necessary for structure fold or activity. Circular dichroism can be employed to determine if both wild-type and mutant enzymes contain the same protein structure and fold properties. Additionally, the highest UDP-GlcNDAz generated was 1/10 compared to the UDP-GlcNAc generated from wild-type NahK/GlmU.

Preliminary study to suggest two-step process for O-GlcNDAz-ylation

In efforts to demonstrate the ability of O-GlcNDAz-ylation with NahK/GlmU and OGT, a high school summer student under my supervision performed the following

experiments. Synthesized GlcNDAz was incubated with ATP, UTP and wild-type NahK/GlmU for 3 days in the bifunctional reaction buffer (50 mM TrisHCl pH 8.0 and 10 mM $MgCl_2$) (119). After filtering out the NahK/GlmU, the production of UDP-GlcNDAz was determined by HPAEC. UDP-GlcNDAz was generated with the addition of the bifunctional enzyme compared to control (purple trace, Fig. 5.4A). The peak

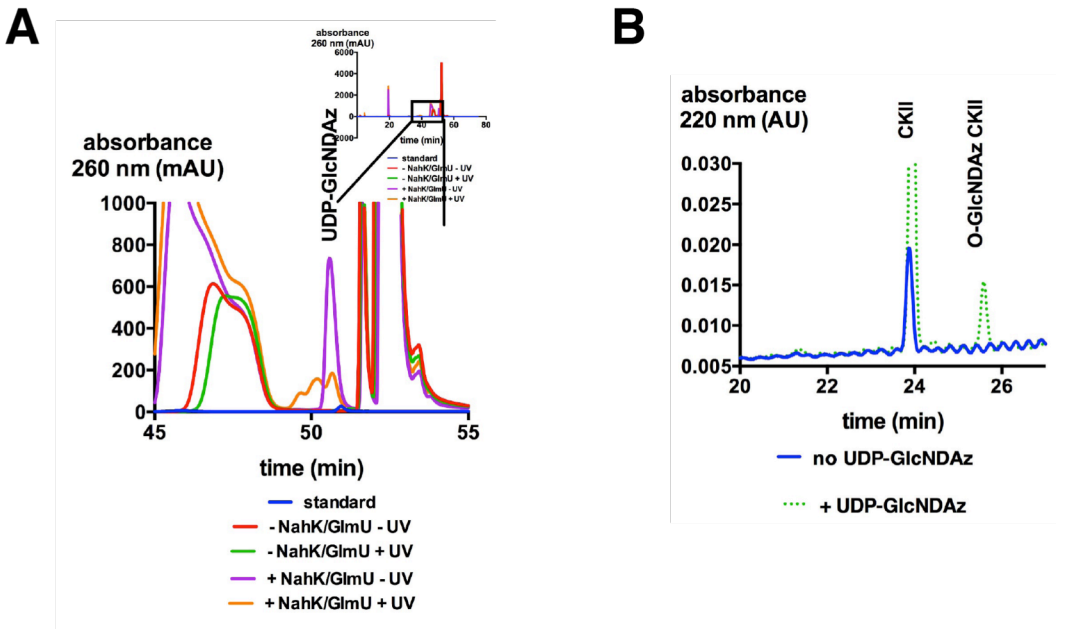

Figure 5.4. Preliminary experiment suggests capability of O-GlcNDAz-ylation with NahK/GlmU and OGT. *A,* Synthesized GlcNDAz was incubated for 3 d at 37 °C with ATP, UTP, and with or without NahK/GlmU. The enzyme was filtered out. The samples were split into two. Half of the reaction was subjected to UV irradiation (365 nm) for 15 minutes. The samples were subjected to HPAEC analysis. *B*, The filtrate from the chemo-enzymatic reaction of no NahK/GlmU or plus NahK/GlmU was incubated with CKII and wtOGT for 1 d at 37 °C. OGT and then filtered. The peptide products were analyzed by HPLC.

corresponding to UDP-GlcNDAz was confirmed to contain a diazirine as the peak decreased after UV irradiation treatment (orange trace, Fig. 5.4A). Next, the crude filtrate was exposed to a glycosylation reaction with a biotinylated, well-studied O-

GlcNAc-modified substrate, casein kinase II (CKII), and OGT(C917A) in the OGT reaction buffer (20 mM TrisHCl pH 8.0, 12.6 mM $MgCl_2$, and 20 mM β-mercaptoethanol) (68). The production of O-GlcNDAz-ylation was analyzed by HPLC. The preliminary result revealed 21.5% glycosylation of CKII peptide, but the transfer reaction didn't go to completion (Fig. 5.4B). These data suggested NahK/GlmU generated enough UDP-GlcNDAz for glycosyltransferase activity of GlcNDAz onto CKII.

One-pot reaction for O-GlcNAc-ylation of CKII with NahK/GlmU and OGT

Chen and co-workers reported the ability to construct a one-pot enzyme reaction to glycosylate glycan substrates with terminal sialic acids, another essential sugar found in humans (121). Currently to-date, there is no evidence of a one-pot O-GlcNAc-ylation reaction developed. I desired to develop a similar approach as Chen by using GlcNDAz as the starting material incubated with peptide, NahK/GlmU, and OGT to generate O-glycosylated peptides. NahK/GlmU and OGT, in the wild-type form, were examined for activity in a one-pot reaction. I chose to examine this one-pot condition with GlcNAc because this is the preferred substrate for both enzymes (Fig. 5.5A). NahK/GlmU was reported to be active in 50 mM TrisHCl, pH 8.0 buffer containing no reducing agent (119). OGT was active in 20 mM TrisHCl, pH 8.0, containing a high amount of reducing agent (20 mM β-mercaptoethanol) (68). The enzymes were examined in the ideal buffers for each along with increasing concentrations of β-mercaptoethanol (2.5 mM, 5 mM and 10 mM). GlcNAc was incubated with ATP, UTP, CKII and NahK/GlmU and wtOGT for 24 h at 37 °C in the indicated buffer, and enzymes were filtered. The

production of O-GlcNAc-ylation was investigated by HPLC. The buffer with no reducing agent revealed the highest yield of O-GlcNAc-ylation in the one-pot condition ranging from 40 % to 50 % (Fig. 5.5B). While in its preferred buffer condition, OGT demonstrated no glycosylation, but OGT demonstrated similar activity to glycosylate CKII in the other β-mercaptoethanol concentrations (2.5, 5, and 10 mM) (Fig. 5.5B). Lower concentrations of β-mercaptoethanol allowed NahK/GlmU and OGT to generate their respective products UDP-GlcNAc and O-GlcNAc. A possible thought is that NahK/GlmU inhibited OGT's

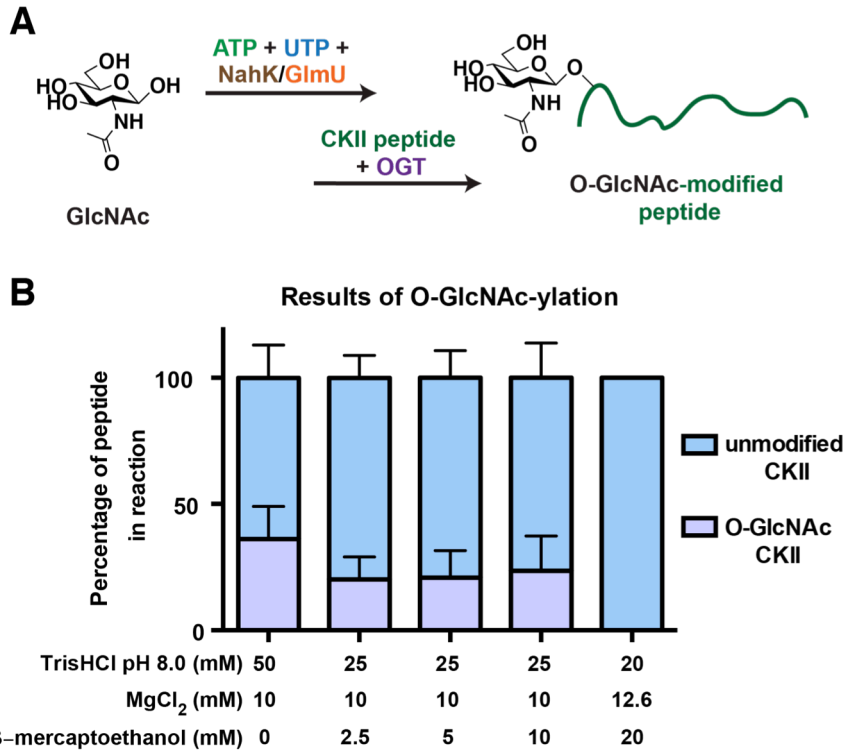

Figure 5.5. Production of O-GlcNAc-ylated CKII peptide through a one-pot reaction with NahK/GlmU and OGT. *A,* Scheme of one-pot reaction conducted with GlcNAc, ATP, UTP, CKII and both NahK/GlmU and OGT enzymes. *B*, The one-pot reaction was performed on CKII peptide under increasing concentrations of β-mercaptoethanol with wild-type forms of NahK/GlmU and OGT. The enzymes were filtered and the peptide products were analyzed by HPLC. The percentage of peptide from the reaction was quantified by peak area under different buffer conditions. Three independent experiments were performed, with error bars representing standard error mean.

activity in the 20 mM β-mercaptoethanol buffer. To examine this possibility, the NahK/GlmU reaction can be conducted first, followed by removal of the enzyme. Then in the same buffer as the first reaction, OGT and CKII peptide can be incubated with the filtrate from the first reaction to generate O-GlcNAc-modified CKII. I expect the glycosylation percentage of CKII to be similar in all five buffers (0, 2.5, 5, 10 and 20 mM β-mercaptoethanol) with this new filtration addition to the protocol.

Optimizing the one-pot reaction for O-GlcNDAz using OGT(C917A)

The promising results from the one-pot O-GlcNAc-ylation led me to believe the one-pot conditions may be utilized for O-GlcNDAz-ylation. The synthesized GlcNDAz was incubated in a similar one-pot reaction as used for GlcNAc, replacing wild-type OGT with OGT(C917A), the mutant form of OGT that favors O-GlcNDAz-ylation production (68). Free GlcNDAz was incubated with ATP, UTP, CKII, NahK/GlmU and OGT(C917A) for 24 h at 37 °C in the increasing β-mercaptoethanol buffer conditions (0, 2.5, 5, 10, and 20 mM). The enzymes were filtered, and the peptide product was examined by HPLC. In the different conditions, there was no evidence of O-GlcNDAz-ylation (Fig. 5.6A). Unsure if the glycosylation didn't work because OGT(C917A) was inactive in the one-pot or if NahK/GlmU didn't produce UDP-GlcNDAz, the one-pot reactions were examined for generation of UDP-GlcNDAz by HPAEC. NahK/GlmU was capable of chemo-enzymatically synthesizing UDP-GlcNDAz, generating about 250 μM

UDP-GlcNDAz (Fig. 5.6B), which is about ⅔ of what was generated in just the NahK/GlmU reaction along (Fig. 5.3E). The amount of GlcNDAz, ATP, and UTP used to

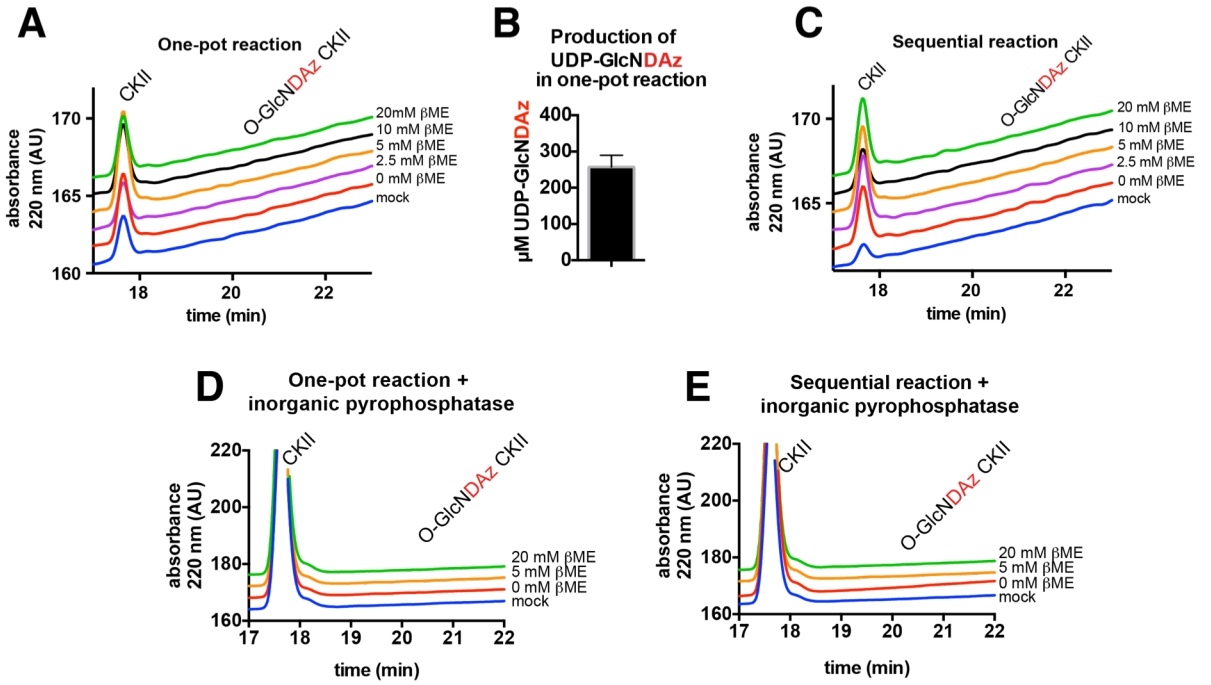

Figure 5.6. Troubleshooting for O-GlcNDAz-ylation of CKII peptide. *A,* The CKII peptide was incubated in the one-pot reaction to test O-GlcNDAz-ylation under increasing β-mercaptoethanol concentration. The reactions were filtered after 24 h incubation and analyzed by HPLC. There was no generation of O-GlcNDAz detected on the peptide. *B,* The production of UDP-GlcNDAz was detected by HPAEC analysis for the one-pot reaction. There is about 250 *μ*M UDP-GlcNDAz generated in the one-pot reaction after 24 h. *C,* The chemo-enzymatic reaction to generate UDP-GlcNDAz was incubated for 24 h, then CKII and OGT were added for another 24 h incubation. The enzymes were filtered. The products were analyzed by HPLC. There was no glycosylation produced in this sequential method. *D,* The one-pot reaction similar to *A*, but with the addition of inorganic pyrophosphatase to generate more UDP-GlcNDAz, was conducted. HPLC analysis revealed there was no glycosylation of CKII with the addition of the inorganic pyrophosphatase. *E,* A sequential method was conducted with a 24 h chemo-enzymatic reaction with inorganic pyrophosphatase to produce UDP-GlcNDAz, followed by addition of CKII and OGT(C917A) for another 24 h incubation. The peptide was analyzed by HPLC. There was no detection of glycosylation using the sequential method with inorganic pyrophosphatase. Experiments were conducted in duplicates. Traces shown are a representation of one trial.

start the reaction is 25 mM GlcNDAz and 30 mM nucleotides, and is 100-fold higher than the amount of UDP-GlcNDAz produced. Though the bifunctional enzyme can generate UDP-GlcNDAz in the one-pot reaction, I questioned if the amount of UDP-GlcNDAz generated was enough for the OGT reaction. Next, I examined if the NahK/GlmU reaction needed a longer time to generate more UDP-GlcNDAz. Therefore in a sequential method and increasing amounts of β-mercaptoethanol, the starting material components of GlcNDAz, ATP, UTP, and NahK/GlmU were incubated for 24 h, then CKII and OGT(C917A) were added and incubated for another 24 h. The enzymes were filtered. The production of O-GlcNDAz from this sequential method was analyzed by HPLC. The reaction condition under multiple buffer conditions did not generate O-GlcNDAz-ylation (Fig. 5.6C). While NahK/GlmU generated more than 1.5 mM UDP-GlcNAc (data not shown) after 24 hours, the bifunctional enzyme produced 250 μM after 24 hours. It may be plausible that there is not enough UDP-GlcNDAz generated in the one-pot or sequential method. This hypothesis was tested by adding inorganic pyrophosphatase, which regenerates ATP from ADP, to increase the production of UDP-GlcNDAz, as it has been done previous with UDP-GlcNAc (119). With the addition of inorganic pyrophosphatase in the one-pot (Fig. 5.6D) or sequential method (Fig. 5.6E), there was no glycosylation of the CKII peptide detected by HPLC. Even though the NahK/GlmU generated UDP-GlcNDAz, it is possible the UDP impurity from UTP may be interfering with OGT(C917A). The UDP-GlcNDAz could be purified from the reaction using HPAEC purification. Also, NahK/GlmU may be interacting with OGT(C917A) and preventing O-GlcNDAz-ylation of CKII. To address this issue

NahK/GlmU can be removed as conducted earlier (Fig. 5.4). The method to combine both enzymes together to generate O-GlcNDAz may not be feasible as others have also demonstrated the need for multi-step synthesis for assembly of glycoconjugates (122).

A two-step process to O-GlcNDAz-modified peptides

In **chapter 4**, I demonstrated that CKII was about 85% glycosylated when incubated with chemically synthesized UDP-GlcNDAz and OGT(C917A) alone (Fig. 4.5B). Previously mentioned, the O-GlcNDAz was generated when the chemo-enzymatic reaction was performed first followed by the glycosylation reaction (Fig. 5.4B). It may be key that NahK/GlmU needs to be removed from the reaction before

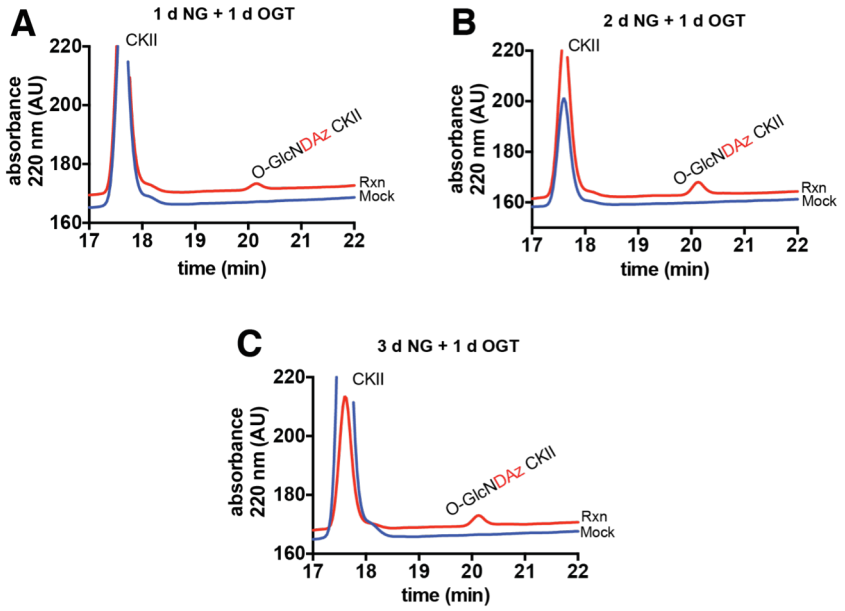

Figure 5.7. O-GlcNDAz-ylation requires a two-step process. *A,* Free GlcNDAz was incubated with NahK/GlmU, ATP, and UTP for 1 d (*A*), 2 d (*B*) or 3 d (*C*). The enzyme was filtered and incubated with CKII and OGT(C917A) for 1 d. OGT was filtered and the peptide products were analyzed by HPLC. The transfer activity represented a range of 4-14% for 1 d NahK/GlmU, 2-16% for 2 d NahK/GlmU, and 4- 37% NahK/GlmU for 3 d. The experiment was conducted in triplicates, but traces shown are a representation of one trial.

proceeding to the transferase reaction with OGT(C917A). Synthesized GlcNDAz was incubated with ATP, UTP, and NahK/GlmU for 1 d (Fig. 5.7A), 2 d (Fig. 5.7B) or 3 d (Fig. 5.7C) at 37 °C. The enzyme was filtered from the reaction. The filtrate was incubated with CKII and OGT(C917A) for another 24 h at 37 °C. The glycosylation of O-GlcNDAz on the CKII peptide was examined and confirmed by HPLC. The amount of O-GlcNDAz generated from 1 d ranged from 4.5 – 14.7%. The 2 d incubation of NahK/GlmU produced 2 – 16% O-GlcNDAz-ylation on CKII. Finally, NahK/GlmU incubated for 3 d constructed 4 – 37% O-GlcNDAz modification on CKII. The filtration step between NahK/GlmU and OGT(C917A) reactions was necessary for transferase activity. These data suggested NahK/GlmU is inhibiting OGT(C917A) activity. Although the one-pot condition is ideal to generate glycosylated peptides in a cell free system, the two-step synthesis is an alternative approach to generate the O-GlcNDAz-ylated peptides.

I further examined the ability of the NahK/GlmU and OGT(C917A) to glycosylate O-GlcNAc-modified peptides. The studies of this two-step glycosylation reaction were extended on additional O-GlcNAc-modified peptides to demonstrate the versatility of this chemo-enzymatic synthesis. Two peptides chosen were +P2 from human α A crystallin and IRS1 from insulin receptor substrate 1. Ball and co-workers demonstrated insulin receptor substrate 1 (IRS1) to be O-GlcNAc-modified (123). Bovine A crystallin was demonstrated to be O-GlcNAc-modified (124), and the sequence known to be O-GlcNAc-modified demonstrated conservation to the human α A crystallin (+P2). Peptides at 100 μM were incubated with 500 μM UDP-GlcNDAz and 1 μM

OGT(C917A), which were similar to conditions used in **chapter 4.** A former post-doc in lab demonstrated the ability of OGT(C917A) to glycosylate +P2 peptide with about 55%

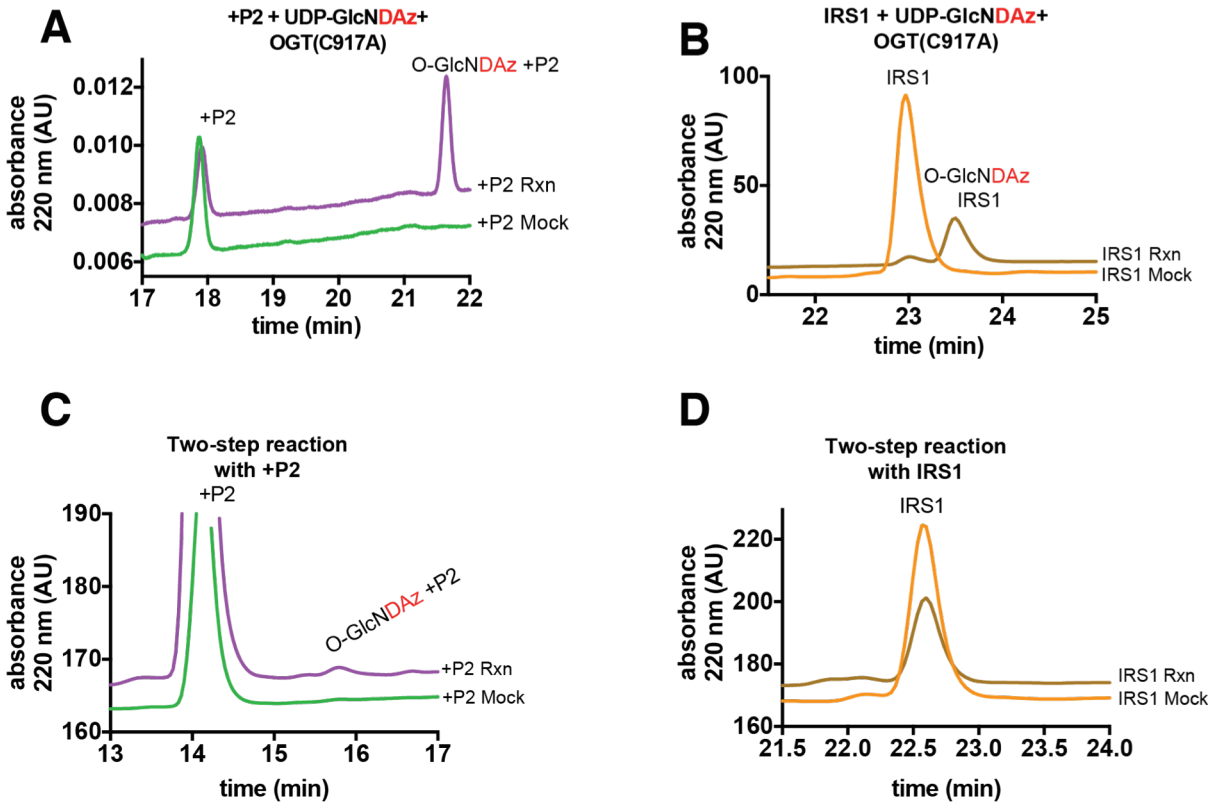

Figure 5.8. Versatility of two-step reaction to O-GlcNDAz-modified additional peptides. *A,* The glycosylation reaction was conducted with +P2, chemically synthesized UDP-GlcNDAz, and OGT(C917A) or no enzyme for 1 d at 37 °C. The enzyme was filtered and the peptide products were analyzed by HPLC. While the +P2 Mock reaction (green trace) demonstrated no glycosylation, the addition of OGT(C917A) exhibited about 55% O-GlcNDAz-ylation. *B*, IRS1 was incubated for 1 d at 37 °C with UDP-GlcNDAz and with or without OGT(C917A) as reaction and mock samples. Though IRS1 Mock reaction (orange trace) revealed no glycosylation, the addition of OGT(C917A) demonstrated O-GlcNDAz-ylation ranging from 10- 70%. Chemo-enzymatic reactions were performed with GlcNDAz, ATP, UTP, and NahK/GlmU for 3 d, and filtered. The filtrate containing UDP-GlcNDAz was incubated with +P2 (*C*) or IRS1 (*D*) and OGT(C917A) for 1 d. The peptide products were examined by HPLC. +P2 exhibited 1 % O-GlcNDAz-ylation, and IRS1 revealed no glycosylation. Experiment was performed in duplicate, and data presented are representative results.

O-GlcNDAz modification (Fig. 5.8A). Additionally, IRS1 peptide exhibited 10 – 70% O-GlcNDAz-ylation under the UDP-GlcNDAz and OGT(C917A) reaction (Fig. 5.8B). These results suggested the OGT(C917A) had a high selectivity to transfer GlcNDAz onto +P2 and IRS1. Next, I examined the ability of the peptides to be glycosylated under the two-step process (Fig. 5.7). The chemo-enzymatic reaction was conducted with GlcNDAz, ATP, UTP, and NahK/GlmU for 3 d. NahK/GlmU was filtered, and the filtrate was incubated with indicated peptide and OGT(C917A) for 1 d. OGT(C917A) was filtered, and the peptide products were examined by HPLC. While +P2 demonstrated 1% O-GlcNDAz-ylation (Fig. 5.8C), IRS1 exhibited no glycosylation (Fig. 5.8D). Though +P2 and IRS1 peptides proved to be moderate substrates for O-GlcNDAz-ylation (Fig. 5.8A and 5.8B respectively), they didn't reveal high glycosylation. While OGT has demonstrated selectivity on substrate preference, through radiolabeling of purified O-GlcNAc-modified proteins (125).

One explanation as to why CKII, IRS1 and +P2 do not achieve the same percentage of O-GlcNDAz-ylation could be that the leftover GlcNDAz, ATP, and UTP from the NahK/GlmU reaction. There are two possibilities that could be explored to determine if the leftover starting materials are inhibiting transfer of GlcNDAz. First, adding inorganic pyrophosphatase or more NahK/GlmU may increase the amount of UDP-GlcNDAz generated in the first reaction. Another test would be to purify UDP-GlcNDAz to remove the leftover starting material. Another explanation as to why the peptides were not glycosylated in the same percentage could be caused by differences in enzyme efficiency for the different O-GlcNAc- modified peptide substrates. Vocadlo

and co-workers demonstrated through kinetic analysis that OGT and OGA have different enzyme efficiency towards different O-GlcNAc-modified protein substrates (126). There is still troubleshooting to be considered before this chemo-enzymatic route to glycosylate peptides can be utilized for the cell free O-GlcNDAz crosslinking approach.

DISCUSSION

Chemo-enzymatic Studies for UDP-GlcNDAz production

The development of a chemo-enzymatic synthesis for O-GlcNDAz-modified biotin-tagged peptides addresses several challenges surfaced in the cell-based O-GlcNDAz crosslinking assay. The use of UDP-GlcNDAz was the first setback, as a large amount was needed for this project. Generally an eight-step chemical synthesis (Fig. 5.2) has been a standard way to generate nucleotide donor sugars, such as UDP-GlcNDAz (47, 110). In efforts to bypass the severe synthesis route, I employed the bifunctional enzyme, NahK/GlmU involved in UDP-GlcNAc production. The data revealed wild-type NahK/GlmU generated higher amounts of UDP-GlcNDAz compared to the NahK/GlmU(Y197A) (Fig. 5.3E). The tyrosine residue may provide structural integrity to GlmU or be involved in the enzyme's activity. Circular dichroism can use employed to determine if wild-type and mutant GlmU share the same protein structure and fold characteristics.

Generation of O-GlcNDAz-modified peptides via cell free synthesis

NahK/GlmU and OGT(C917A), a form of OGT that preferred GlcNDAz, were examined in a two-step process to test their abilities on GlcNDAz. The chemo-enzymatic reaction was exploited with synthesized GlcNDAz, ATP, UTP, and NahK/GlmU, and the production of UDP-GlcNDAz was confirmed by HPAEC and UV irradiation (Fig. 5.4A). The irradiation of samples with and without NahK/GlmU confirmed the peak contained a diazirine modification. The filtrate from the NahK/GlmU reaction, OGT(C917A) and CKII demonstrated ability to glycosylate CKII with the crude filtrate of about 20% (Fig. 5.4B). The data suggested that the leftover starting materials or impurities from the NahK/GlmU did not fully interfere with OGT(C917A)'s activity.

A method to generate O-GlcNDAz-ylated peptides was to establish a one-pot reaction. First, I examined if both NahK/GlmU and OGT were functionally in a one-pot reaction and active toward the natural substrate GlcNAc. Both enzymes demonstrated the ability to generate their respective product, UDP-GlcNAc and O-GlcNAc in several different buffer conditions (Fig. 5.5B). The buffer with 0 mM β-mercaptoethanol revealed the highest production of O-GlcNAc modification about 40 – 50%, while the other concentrations of β-mercaptoethanol (2.5, 5, 10 and 20 mM) experienced lower glycosylation activities. NahK/GlmU generated enough UDP-GlcNAc from the one-pot reaction to glycosylate CKII peptide with OGT. Future work to increase O-GlcNAc-ylation would be to add inorganic pyrophosphatase to increase the amount of UDP-GlcNAc generated.

Based on the positive findings from the O-GlcNAc-ylation one-pot reaction, I attempted to construct an O-GlcNDAz-ylation one-pot reaction. The one-pot reaction was conducted with free GlcNDAz under increasing β-mercaptoethanol concentration, and without (Fig. 5.6A) or with (Fig. 5.6D) inorganic pyrophosphatase. By HPLC, none of the reaction conditions appeared to transfer GlcNDAz onto CKII. The production of UDP-GlcNDAz was analyzed by HPAEC and revealed 250 μM UDP-GlcNDAz consumption (Fig. 5.6B). I believed there was not enough UDP-GlcNDAz generated in the one-pot reaction. Therefore, I expected NahK/GlmU would need more time to generate UDP-GlcNDAz. In sequential process without (Fig. 5.6C) and with (Fig. 5.6E) inorganic pyrophosphatase, I explored the ability of OGT to transfer O-GlcNDAz to CKII. However under several different reaction conditions, OGT(C917A) was unable to transfer GlcNDAz (Fig. 5.6B). In the one-pot reaction, NahK/GlmU proved to be active and generated half of the amount of UDP-GlcNDAz (250 μM) of a typical OGT reaction. The addition of inorganic pyrophosphatase did not assist in the production of O-GlcNDAz-ylation of CKII. Removal of NahK/GlmU, ATP, and UTP may be necessary before the OGT reaction to overcome OGT inactivity.

From the positive findings of the two-step methodology, I further explored the variability of this methodology on other O-GlcNAc-modified peptides. I examined the efficiency of O-GlcNDAz-ylation with NahK/GlmU reacting under different time duration (1 d, 2 d, and 3 d) (Fig. 5.7). All of the two-step reactions exhibited variable low range of 4 – 37 % glycosylation independent of the NahK/GlmU time duration. It might be necessary to examine the production of UDP-GlcNDAz by HPAEC analysis to determine

how much is generated before continuing onto the transferase reaction. If UDP-GlcNDAz production is lower than 200 μM, I believe the transferase reaction would not work because there is not enough nucleotide donor sugar compared to the GlcNDAz and nucleotide starting materials. If more UDP-GlcNDAz is required, I suggest the employment of inorganic pyrophosphatase to increase production.

Other O-GlcNAc-modified peptides, such as +P2 and IRS1, were studied for further versatility of the two-step process for O-GlcNDAz-ylation. While +P2 revealed 1% O-GlcNDAz-ylation, IRS1 exhibited no glycosylation. When OGT(C917A) was given synthesized UDP-GlcNDAz and peptide, it demonstrated ability to glycosylate 85% CKII (Fig. 4.5B), 55% +P2 (Fig. 5.8A), and 10 – 70% IRS1 (Fig. 5.8B). However, the two-step process revealed low glycosylation for all three peptides, which may be due to low levels of UDP-GlcNDAz. As stated earlier, there are several methods to increase O-GlcNDAz-ylation, more NahK/GlmU and inorganic pyrophosphatase, as well as purification of UDP-GlcNDAz may assist in increasing OGT activity. Since OGT exhibits diverse enzyme efficiency on different substrates, I would believe not every O-GlcNAc-modified peptide would be fully O-GlcNDAz-modified. I suggest adjustment of the peptide, OGT(C917A) or UDP-GlcNDAz concentrations may assist to increase O-GlcNDAz-ylation production.

Future work for the development towards a versatile GlcNDAz technology

Once the glycosylation reaction is optimized to produce at least a 50% yield of O-GlcNDAz-ylated peptide, the complementary cell free O-GlcNDAz crosslinking method can be employed to identify binding partners of O-GlcNAc-modified proteins. The cell free O-GlcNDAz crosslinking assay will be conducted as follows (Fig. 5.9). Biotin-tagged peptides will be incubated with UDP-GlcNDAz and OGT(C917A) to create high yields of O-GlcNDAz-modified peptides. These glycosylated peptides will be incubated with mammalian cell lysate and crosslinked to capture potential binding partners.

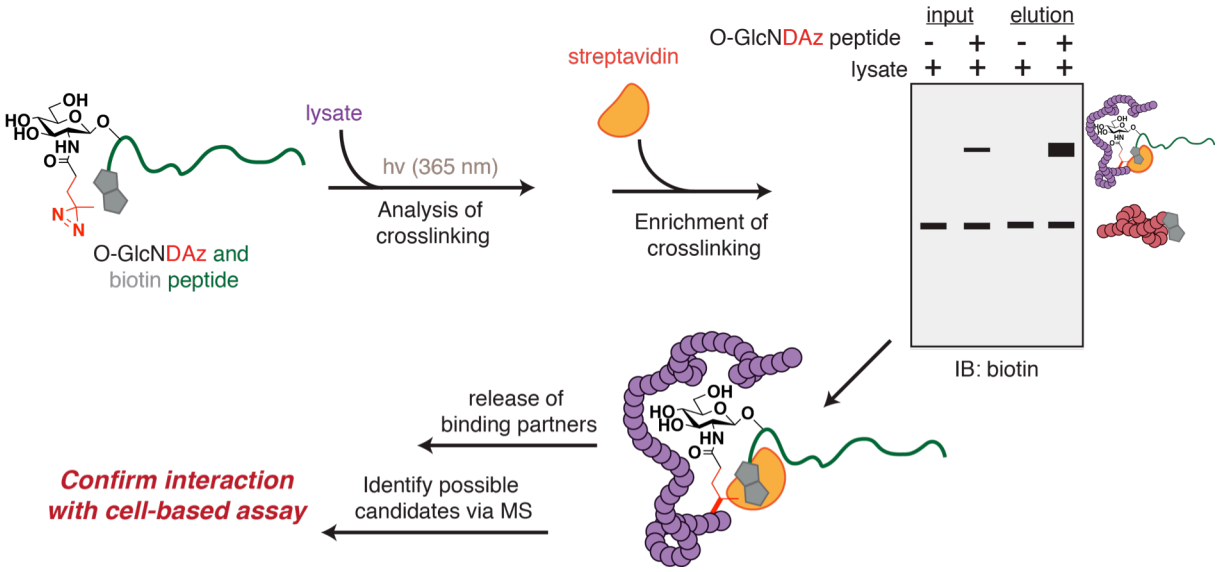

Figure 5.9. Methodology to improve O-GlcNDAz-ylation technology via cell free approach. Biotinylated peptides derived from O-GlcNAc modified proteins are glycosylated with O-GlcNDAz using OGT or no enzyme for control. The unmodified and modified peptides are incubated with lysate and crosslinked. Crosslinking of the peptide to proteins is confirmed by immunoblot for biotin. A new band in the glycosylated peptide-treated sample would appear which corresponds to the crosslinking of the peptide to an unknown protein. The crosslinked complex is enriched using streptavidin purification. The binding partners are released and identified by proteomics analysis. These potential candidates are confirmed through the cell-based assay.

Streptavidin purification will be employed to purify the crosslinked complexes containing the biotin-tagged, O-GlcNDAz-modified peptide. Analysis of crosslinking will be examined by immunoblot probing for biotin, which recognizes the tag on the peptide. Samples treated with glycosylated peptides compared to unmodified peptides will reveal a higher molecular weight band, which corresponds to crosslinking of the O-GlcNDAz-modified peptide to an unknown binding partner. This new band will be intensified in the streptavidin treated sample to demonstrate an enrichment of crosslinking. After enrichment of the purification, OGA(C215A) is incubated with the enriched material to hydrolyze GlcNDAz crosslinked to binding partners. The released samples will be digested and submitted for proteomic analysis to identify the potential binding partners of the O-GlcNAc-ylated peptide. These candidates corresponding to potential binding partners will be confirmed using the cell-based O-GlcNDAz crosslinking assay and studied for their functional importance in homeostasis or disease settings.

As mentioned in previous chapters, NUP98 and leukemogenic NUP98 fusions are an interesting set of proteins to study, as the functional importance of O-GlcNAc on these proteins has not been exclusively determined in protein-protein interactions. To study this problem, I suggest identifying O-GlcNAc-modified residues on both NUP98 and NUP98 fusions through proteomics and confirming by the GlcNAz labeling. Once the sequence is confirmed, biotin-tagged peptides will be generated and employed into the cell free O-GlcNDAz crosslinking assay. There may be similar binding partners identified because both proteins encoded the N-terminus FG repeat regions that are known to be heavily O-GlcNAc-modified. However, there may also be different partners

due to the locations of nuclear pore based NUP98 and the nuclei based NUP98 fusions. The study of these interactions may provide insight into the transformation of normal state to leukemia state of the cell.

ACKNOWLEDGEMENTS

I thank Dr. Liu for gifting me a plasmid encoding tr-GlmU-NahK. I thank Courtney Le, a former STARS summer student, for her hard work in trying to make the Y197A mutation and preliminary data on the O-GlcNDAz-ylation reaction condition. I thank Drs. Amberlyn Wands and Janet McCombs for their insightful discussions in troubleshooting this reaction.

BIBLIOGRAPHY

1. Torres CR, Hart GW. Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. The Journal of biological chemistry. 1984;259(5):3308-17. PubMed PMID: 6421821.
2. Kreppel LK, Hart GW. Regulation of a cytosolic and nuclear O-GlcNAc transferase. Role of the tetratricopeptide repeats. The Journal of biological chemistry. 1999;274(45):32015-22. PubMed PMID: 10542233.
3. Haltiwanger RS, Holt GD, Hart GW. Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-acetylglucosamine:peptide beta-N-acetylglucosaminyltransferase. The Journal of biological chemistry. 1990;265(5):2563-8. PubMed PMID: 2137449.
4. Shafi R, Iyer SP, Ellies LG, O'Donnell N, Marek KW, Chui D, Hart GW, Marth JD. The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(11):5735-9. doi: 10.1073/pnas.100471497. PubMed PMID: 10801981; PubMed Central PMCID: PMC18502.
5. Bond MR, Hanover JA. O-GlcNAc cycling: a link between metabolism and chronic disease. Annual review of nutrition. 2013;33:205-29. doi: 10.1146/annurev-nutr-071812-161240. PubMed PMID: 23642195.
6. Ma J, Hart GW. O-GlcNAc profiling: from proteins to proteomes. Clinical proteomics. 2014;11(1):8. doi: 10.1186/1559-0275-11-8. PubMed PMID: 24593906; PubMed Central PMCID: PMC4015695.
7. Love DC, Hanover JA. The hexosamine signaling pathway: deciphering the "O-GlcNAc code". Science's STKE : signal transduction knowledge environment. 2005;2005(312):re13. doi: 10.1126/stke.3122005re13. PubMed PMID: 16317114.
8. Perez Jde J, Udeshi ND, Shabanowitz J, Ciordia S, Juarez S, Scott CL, Olszewski NE, Hunt DF, Garcia JA. O-GlcNAc modification of the coat protein of the potyvirus Plum pox virus enhances viral infection. Virology. 2013;442(2):122-31. doi: 10.1016/j.virol.2013.03.029. PubMed PMID: 23639873.
9. Khidekel N, Ficarro SB, Peters EC, Hsieh-Wilson LC. Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(36):13132-7. doi: 10.1073/pnas.0403471101. PubMed PMID: 15340146; PubMed Central PMCID: PMC516536.
10. Holt GD, Snow CM, Senior A, Haltiwanger RS, Gerace L, Hart GW. Nuclear pore complex glycoproteins contain cytoplasmically disposed O-linked N-acetylglucosamine. The Journal of cell biology. 1987;104(5):1157-64. PubMed PMID: 3571327; PubMed Central PMCID: PMC2114481.

11. Walther TC, Fornerod M, Pickersgill H, Goldberg M, Allen TD, Mattaj IW. The nucleoporin Nup153 is required for nuclear pore basket formation, nuclear pore complex anchoring and import of a subset of nuclear proteins. The EMBO journal. 2001;20(20):5703-14. doi: 10.1093/emboj/20.20.5703. PubMed PMID: 11598013; PubMed Central PMCID: PMC125666.
12. Fontoura BM, Blobel G, Matunis MJ. A conserved biogenesis pathway for nucleoporins: proteolytic processing of a 186-kilodalton precursor generates Nup98 and the novel nucleoporin, Nup96. The Journal of cell biology. 1999;144(6):1097-112. PubMed PMID: 10087256; PubMed Central PMCID: PMC2150585.
13. Fontoura BM, Dales S, Blobel G, Zhong H. The nucleoporin Nup98 associates with the intranuclear filamentous protein network of TPR. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(6):3208-13. doi: 10.1073/pnas.061014698. PubMed PMID: 11248057; PubMed Central PMCID: PMC30632.
14. Favreau C, Worman HJ, Wozniak RW, Frappier T, Courvalin JC. Cell cycle-dependent phosphorylation of nucleoporins and nuclear pore membrane protein Gp210. Biochemistry. 1996;35(24):8035-44. doi: 10.1021/bi9600660. PubMed PMID: 8672508.
15. Wang Z, Udeshi ND, Slawson C, Compton PD, Sakabe K, Cheung WD, Shabanowitz J, Hunt DF, Hart GW. Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. Science signaling. 2010;3(104):ra2. doi: 10.1126/scisignal.2000526. PubMed PMID: 20068230; PubMed Central PMCID: PMC2866299.
16. Hanover JA, Cohen CK, Willingham MC, Park MK. O-linked N-acetylglucosamine is attached to proteins of the nuclear pore. Evidence for cytoplasmic and nucleoplasmic glycoproteins. The Journal of biological chemistry. 1987;262(20):9887-94. PubMed PMID: 3110163.
17. Finlay DR, Newmeyer DD, Price TM, Forbes DJ. Inhibition of in vitro nuclear transport by a lectin that binds to nuclear pores. The Journal of cell biology. 1987;104(2):189-200. PubMed PMID: 3805121; PubMed Central PMCID: PMC2114419.
18. Davis LI, Blobel G. Identification and characterization of a nuclear pore complex protein. Cell. 1986;45(5):699-709. PubMed PMID: 3518946.
19. Snow CM, Senior A, Gerace L. Monoclonal antibodies identify a group of nuclear pore complex glycoproteins. The Journal of cell biology. 1987;104(5):1143-56. PubMed PMID: 2437126; PubMed Central PMCID: PMC2114474.
20. Comer FI, Vosseller K, Wells L, Accavitti MA, Hart GW. Characterization of a mouse monoclonal antibody specific for O-linked N-acetylglucosamine. Analytical biochemistry. 2001;293(2):169-77. doi: 10.1006/abio.2001.5132. PubMed PMID: 11399029.

21. Teo CF, Ingale S, Wolfert MA, Elsayed GA, Not LG, Chatham JC, Wells L, Boons GJ. Glycopeptide-specific monoclonal antibodies suggest new roles for O-GlcNAc. Nature chemical biology. 2010;6(5):338-43. doi: 10.1038/nchembio.338. PubMed PMID: 20305658; PubMed Central PMCID: PMC2857662.
22. Horsch M, Hoesch L, Vasella A, Rast DM. N-acetylglucosaminono-1,5-lactone oxime and the corresponding (phenylcarbamoyl)oxime. Novel and potent inhibitors of beta-N-acetylglucosaminidase. European journal of biochemistry / FEBS. 1991;197(3):815-8. PubMed PMID: 2029909.
23. Khidekel N, Arndt S, Lamarre-Vincent N, Lippert A, Poulin-Kerstien KG, Ramakrishnan B, Qasba PK, Hsieh-Wilson LC. A chemoenzymatic approach toward the rapid and sensitive detection of O-GlcNAc posttranslational modifications. Journal of the American Chemical Society. 2003;125(52):16162-3. doi: 10.1021/ja038545r. PubMed PMID: 14692737.
24. Clark PM, Dweck JF, Mason DE, Hart CR, Buck SB, Peters EC, Agnew BJ, Hsieh-Wilson LC. Direct in-gel fluorescence detection and cellular imaging of O-GlcNAc-modified proteins. Journal of the American Chemical Society. 2008;130(35):11576-7. doi: 10.1021/ja8030467. PubMed PMID: 18683930; PubMed Central PMCID: PMC2649877.
25. Ramakrishnan B, Qasba PK. Structure-based design of beta 1,4-galactosyltransferase I (beta 4Gal-T1) with equally efficient N-acetylgalactosaminyltransferase activity: point mutation broadens beta 4Gal-T1 donor specificity. The Journal of biological chemistry. 2002;277(23):20833-9. doi: 10.1074/jbc.M111183200. PubMed PMID: 11916963.
26. Yi W, Clark PM, Mason DE, Keenan MC, Hill C, Goddard WA, 3rd, Peters EC, Driggers EM, Hsieh-Wilson LC. Phosphofructokinase 1 glycosylation regulates cell growth and metabolism. Science. 2012;337(6097):975-80. doi: 10.1126/science.1222278. PubMed PMID: 22923583; PubMed Central PMCID: PMC3534962.
27. Vocadlo DJ, Hang HC, Kim EJ, Hanover JA, Bertozzi CR. A chemical approach for identifying O-GlcNAc-modified proteins in cells. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(16):9116-21. doi: 10.1073/pnas.1632821100. PubMed PMID: 12874386; PubMed Central PMCID: PMC171382.
28. Boyce M, Carrico IS, Ganguli AS, Yu SH, Hangauer MJ, Hubbard SC, Kohler JJ, Bertozzi CR. Metabolic cross-talk allows labeling of O-linked beta-N-acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(8):3141-6. doi: 10.1073/pnas.1010045108. PubMed PMID: 21300897; PubMed Central PMCID: PMC3044403.
29. Zaro BW, Yang YY, Hang HC, Pratt MR. Chemical reporters for fluorescent detection and identification of O-GlcNAc-modified proteins reveal glycosylation of the ubiquitin ligase NEDD4-1. Proceedings of the National

Academy of Sciences of the United States of America. 2011;108(20):8146-51. doi: 10.1073/pnas.1102458108. PubMed PMID: 21540332; PubMed Central PMCID: PMC3100932.
30. Mikesh LM, Ueberheide B, Chi A, Coon JJ, Syka JE, Shabanowitz J, Hunt DF. The utility of ETD mass spectrometry in proteomic analysis. Biochimica et biophysica acta. 2006;1764(12):1811-22. doi: 10.1016/j.bbapap.2006.10.003. PubMed PMID: 17118725; PubMed Central PMCID: PMC1853258.
31. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(26):9528-33. doi: 10.1073/pnas.0402700101. PubMed PMID: 15210983; PubMed Central PMCID: PMC470779.
32. Zhao P, Viner R, Teo CF, Boons GJ, Horn D, Wells L. Combining high-energy C-trap dissociation and electron transfer dissociation for protein O-GlcNAc modification site assignment. Journal of proteome research. 2011;10(9):4088-104. doi: 10.1021/pr2002726. PubMed PMID: 21740066; PubMed Central PMCID: PMC3172619.
33. Zachara NE, Hart GW. O-GlcNAc modification: a nutritional sensor that modulates proteasome function. Trends in cell biology. 2004;14(5):218-21. doi: 10.1016/j.tcb.2004.03.005. PubMed PMID: 15130576.
34. Hanover JA, Forsythe ME, Hennessey PT, Brodigan TM, Love DC, Ashwell G, Krause M. A Caenorhabditis elegans model of insulin resistance: altered macronutrient storage and dauer formation in an OGT-1 knockout. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(32):11266-71. doi: 10.1073/pnas.0408771102. PubMed PMID: 16051707; PubMed Central PMCID: PMC1183534.
35. Lehman DM, Fu DJ, Freeman AB, Hunt KJ, Leach RJ, Johnson-Pais T, Hamlington J, Dyer TD, Arya R, Abboud H, Goring HH, Duggirala R, Blangero J, Konrad RJ, Stern MP. A single nucleotide polymorphism in MGEA5 encoding O-GlcNAc-selective N-acetyl-beta-D glucosaminidase is associated with type 2 diabetes in Mexican Americans. Diabetes. 2005;54(4):1214-21. PubMed PMID: 15793264.
36. Cole RN, Hart GW. Cytosolic O-glycosylation is abundant in nerve terminals. Journal of neurochemistry. 2001;79(5):1080-9. PubMed PMID: 11739622.
37. Ma Z, Vosseller K. O-GlcNAc in cancer biology. Amino acids. 2013;45(4):719-33. doi: 10.1007/s00726-013-1543-8. PubMed PMID: 23836420.
38. Hart GW, Greis KD, Dong LY, Blomberg MA, Chou TY, Jiang MS, Roquemore EP, Snow DM, Kreppel LK, Cole RN, et al. O-linked N-acetylglucosamine: the "yin-yang" of Ser/Thr phosphorylation? Nuclear and cytoplasmic glycosylation. Advances in experimental medicine and biology. 1995;376:115-23. PubMed PMID: 8597237.

39. Rexach JE, Clark PM, Mason DE, Neve RL, Peters EC, Hsieh-Wilson LC. Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. Nature chemical biology. 2012;8(3):253-61. doi: 10.1038/nchembio.770. PubMed PMID: 22267118; PubMed Central PMCID: PMC3288555.
40. Bauer C, Gobel K, Nagaraj N, Colantuoni C, Wang M, Muller U, Kremmer E, Rottach A, Leonhardt H. Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). The Journal of biological chemistry. 2015;290(8):4801-12. doi: 10.1074/jbc.M114.605881. PubMed PMID: 25568311; PubMed Central PMCID: PMC4335217.
41. Kelly WG, Dahmus ME, Hart GW. RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. The Journal of biological chemistry. 1993;268(14):10416-24. PubMed PMID: 8486697.
42. Ranuncolo SM, Ghosh S, Hanover JA, Hart GW, Lewis BA. Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription in vitro and in vivo. The Journal of biological chemistry. 2012;287(28):23549-61. doi: 10.1074/jbc.M111.330910. PubMed PMID: 22605332; PubMed Central PMCID: PMC3390630.
43. Rao X, Duan X, Mao W, Li X, Li Z, Li Q, Zheng Z, Xu H, Chen M, Wang PG, Wang Y, Shen B, Yi W. O-GlcNAcylation of G6PD promotes the pentose phosphate pathway and tumor growth. Nature communications. 2015;6:8468. doi: 10.1038/ncomms9468. PubMed PMID: 26399441.
44. Gewinner C, Hart G, Zachara N, Cole R, Beisenherz-Huss C, Groner B. The coactivator of transcription CREB-binding protein interacts preferentially with the glycosylated form of Stat5. The Journal of biological chemistry. 2004;279(5):3563-72. doi: 10.1074/jbc.M306449200. PubMed PMID: 14597631.
45. Mizuguchi-Hata C, Ogawa Y, Oka M, Yoneda Y. Quantitative regulation of nuclear pore complex proteins by O-GlcNAcylation. Biochimica et biophysica acta. 2013;1833(12):2682-9. doi: 10.1016/j.bbamcr.2013.06.008. PubMed PMID: 23777819.
46. Xiao J, Xu S, Li C, Xu Y, Xing L, Niu Y, Huan Q, Tang Y, Zhao C, Wagner D, Gao C, Chong K. O-GlcNAc-mediated interaction between VER2 and TaGRP2 elicits TaVRN1 mRNA accumulation during vernalization in winter wheat. Nature communications. 2014;5:4572. doi: 10.1038/ncomms5572. PubMed PMID: 25091017; PubMed Central PMCID: PMC4143922.
47. Yu SH, Boyce M, Wands AM, Bond MR, Bertozzi CR, Kohler JJ. Metabolic labeling enables selective photocrosslinking of O-GlcNAc-modified proteins to their binding partners. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(13):4834-9. doi: 10.1073/pnas.1114356109. PubMed PMID: 22411826; PubMed Central PMCID: PMC3323966.

48. Firat-Karalar EN, Rauniyar N, Yates JR, 3rd, Stearns T. Proximity interactions among centrosome components identify regulators of centriole duplication. Current biology : CB. 2014;24(6):664-70. doi: 10.1016/j.cub.2014.01.067. PubMed PMID: 24613305; PubMed Central PMCID: PMC4004176.
49. Li B, Kohler JJ. Glycosylation of the nuclear pore. Traffic. 2014;15(4):347-61. doi: 10.1111/tra.12150. PubMed PMID: 24423194; PubMed Central PMCID: PMC4001855.
50. Labokha AA, Gradmann S, Frey S, Hulsmann BB, Urlaub H, Baldus M, Gorlich D. Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. The EMBO journal. 2013;32(2):204-18. doi: 10.1038/emboj.2012.302. PubMed PMID: 23202855; PubMed Central PMCID: PMC3553378.
51. Gough SM, Slape CI, Aplan PD. NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. Blood. 2011;118(24):6247-57. doi: 10.1182/blood-2011-07-328880. PubMed PMID: 21948299; PubMed Central PMCID: PMC3236115.
52. Hetzer MW, Wente SR. Border control at the nucleus: biogenesis and organization of the nuclear membrane and pore complexes. Developmental cell. 2009;17(5):606-16. doi: 10.1016/j.devcel.2009.10.007. PubMed PMID: 19922866; PubMed Central PMCID: PMC3746554.
53. King MC, Raposo G, Lemmon MA. Inhibition of nuclear import and cell-cycle progression by mutated forms of the dynamin-like GTPase MxB. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(24):8957-62. doi: 10.1073/pnas.0403167101. PubMed PMID: 15184662; PubMed Central PMCID: PMC428454.
54. Wolff B, Willingham MC, Hanover JA. Nuclear protein import: specificity for transport across the nuclear pore. Experimental cell research. 1988;178(2):318-34. PubMed PMID: 3169133.
55. Strambio-De-Castillia C, Niepel M, Rout MP. The nuclear pore complex: bridging nuclear transport and gene regulation. Nature reviews Molecular cell biology. 2010;11(7):490-501. doi: 10.1038/nrm2928. PubMed PMID: 20571586.
56. Raices M, D'Angelo MA. Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. Nature reviews Molecular cell biology. 2012;13(11):687-99. doi: 10.1038/nrm3461. PubMed PMID: 23090414.
57. Zhu Y, Liu TW, Madden Z, Yuzwa SA, Murray K, Cecioni S, Zachara N, Vocadlo DJ. Post-translational O-GlcNAcylation is essential for nuclear pore integrity and maintenance of the pore selectivity filter. Journal of molecular cell biology. 2015. doi: 10.1093/jmcb/mjv033. PubMed PMID: 26031751.
58. Griffis ER, Altan N, Lippincott-Schwartz J, Powers MA. Nup98 is a mobile nucleoporin with transcription-dependent dynamics. Molecular biology of the cell.

2002;13(4):1282-97. doi: 10.1091/mbc.01-11-0538. PubMed PMID: 11950939; PubMed Central PMCID: PMC102269.
59. Graux C, Cools J, Melotte C, Quentmeier H, Ferrando A, Levine R, Vermeesch JR, Stul M, Dutta B, Boeckx N, Bosly A, Heimann P, Uyttebroeck A, Mentens N, Somers R, MacLeod RA, Drexler HG, Look AT, Gilliland DG, Michaux L, Vandenberghe P, Wlodarska I, Marynen P, Hagemeijer A. Fusion of NUP214 to ABL1 on amplified episomes in T-cell acute lymphoblastic leukemia. Nature genetics. 2004;36(10):1084-9. doi: 10.1038/ng1425. PubMed PMID: 15361874.
60. Ageberg M, Drott K, Olofsson T, Gullberg U, Lindmark A. Identification of a novel and myeloid specific role of the leukemia-associated fusion protein DEK-NUP214 leading to increased protein synthesis. Genes, chromosomes & cancer. 2008;47(4):276-87. doi: 10.1002/gcc.20531. PubMed PMID: 18181180.
61. Gorello P, La Starza R, Di Giacomo D, Messina M, Puzzolo MC, Crescenzi B, Santoro A, Chiaretti S, Mecucci C. SQSTM1-NUP214: a new gene fusion in adult T-cell acute lymphoblastic leukemia. Haematologica. 2010;95(12):2161-3. doi: 10.3324/haematol.2010.029769. PubMed PMID: 20851865; PubMed Central PMCID: PMC2995577.
62. Quentmeier H, Schneider B, Rohrs S, Romani J, Zaborski M, Macleod RA, Drexler HG. SET-NUP214 fusion in acute myeloid leukemia- and T-cell acute lymphoblastic leukemia-derived cell lines. Journal of hematology & oncology. 2009;2:3. doi: 10.1186/1756-8722-2-3. PubMed PMID: 19166587; PubMed Central PMCID: PMC2636835.
63. Yassin ER, Abdul-Nabi AM, Takeda A, Yaseen NR. Effects of the NUP98-DDX10 oncogene on primary human CD34+ cells: role of a conserved helicase motif. Leukemia. 2010;24(5):1001-11. doi: 10.1038/leu.2010.42. PubMed PMID: 20339440; PubMed Central PMCID: PMC2868946.
64. Ghannam G, Takeda A, Camarata T, Moore MA, Viale A, Yaseen NR. The oncogene Nup98-HOXA9 induces gene transcription in myeloid cells. The Journal of biological chemistry. 2004;279(2):866-75. doi: 10.1074/jbc.M307280200. PubMed PMID: 14561764.
65. Takeda A, Sarma NJ, Abdul-Nabi AM, Yaseen NR. Inhibition of CRM1-mediated nuclear export of transcription factors by leukemogenic NUP98 fusion proteins. The Journal of biological chemistry. 2010;285(21):16248-57. doi: 10.1074/jbc.M109.048785. PubMed PMID: 20233715; PubMed Central PMCID: PMC2871492.
66. Kasper LH, Brindle PK, Schnabel CA, Pritchard CE, Cleary ML, van Deursen JM. CREB binding protein interacts with nucleoporin-specific FG repeats that activate transcription and mediate NUP98-HOXA9 oncogenicity. Molecular and cellular biology. 1999;19(1):764-76. PubMed PMID: 9858599; PubMed Central PMCID: PMC83933.

67. Li Y, Sousa R. Expression and purification of E. coli BirA biotin ligase for in vitro biotinylation. Protein expression and purification. 2012;82(1):162-7. doi: 10.1016/j.pep.2011.12.008. PubMed PMID: 22227598; PubMed Central PMCID: PMC3288220.
68. Rodriguez AC, Yu SH, Li B, Zegzouti H, Kohler JJ. Enhanced transfer of a photocrosslinking GlcNAc analog by an O-GlcNAc transferase mutant with converted substrate specificity. The Journal of biological chemistry. 2015. doi: 10.1074/jbc.M115.667006. PubMed PMID: 26240142.
69. Wands AM, Fujita A, McCombs JE, Cervin J, Dedic B, Rodriguez AC, Nischan N, Bond MR, Mettlen M, Trudgian DC, Lemoff A, Quiding-Jarbrink M, Gustavsson B, Steentoft C, Clausen H, Mirzaei H, Teneberg S, Yrlid U, Kohler JJ. Fucosylation and protein glycosylation create functional receptors for cholera toxin. eLife. 2015;4. doi: 10.7554/eLife.09545. PubMed PMID: 26512888.
70. Dong DL, Hart GW. Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol. The Journal of biological chemistry. 1994;269(30):19321-30. PubMed PMID: 8034696.
71. Heckel D, Comtesse N, Brass N, Blin N, Zang KD, Meese E. Novel immunogenic antigen homologous to hyaluronidase in meningioma. Human molecular genetics. 1998;7(12):1859-72. PubMed PMID: 9811929.
72. Bertram L, Blacker D, Mullin K, Keeney D, Jones J, Basu S, Yhu S, McInnis MG, Go RC, Vekrellis K, Selkoe DJ, Saunders AJ, Tanzi RE. Evidence for genetic linkage of Alzheimer's disease to chromosome 10q. Science. 2000;290(5500):2302-3. doi: 10.1126/science.290.5500.2302. PubMed PMID: 11125142.
73. Schultz J, Pils B. Prediction of structure and functional residues for O-GlcNAcase, a divergent homologue of acetyltransferases. FEBS letters. 2002;529(2-3):179-82. PubMed PMID: 12372596.
74. Henrissat B, Davies G. Structural and sequence-based classification of glycoside hydrolases. Current opinion in structural biology. 1997;7(5):637-44. PubMed PMID: 9345621.
75. Ostrowski A, Gundogdu M, Ferenbach AT, Lebedev AA, van Aalten DM. Evidence for a functional O-GlcNAc system in the thermophilic bacterium Thermobaculum terrenum. The Journal of biological chemistry. 2015. doi: 10.1074/jbc.M115.689596. PubMed PMID: 26491011.
76. Rao FV, Dorfmueller HC, Villa F, Allwood M, Eggleston IM, van Aalten DM. Structural insights into the mechanism and inhibition of eukaryotic O-GlcNAc hydrolysis. The EMBO journal. 2006;25(7):1569-78. doi: 10.1038/sj.emboj.7601026. PubMed PMID: 16541109; PubMed Central PMCID: PMC1440316.
77. Yuzwa SA, Macauley MS, Heinonen JE, Shan X, Dennis RJ, He Y, Whitworth GE, Stubbs KA, McEachern EJ, Davies GJ, Vocadlo DJ. A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in

vivo. Nature chemical biology. 2008;4(8):483-90. doi: 10.1038/nchembio.96. PubMed PMID: 18587388.
78. Zhang Z, Tan EP, VandenHull NJ, Peterson KR, Slawson C. O-GlcNAcase Expression is Sensitive to Changes in O-GlcNAc Homeostasis. Frontiers in endocrinology. 2014;5:206. doi: 10.3389/fendo.2014.00206. PubMed PMID: 25520704; PubMed Central PMCID: PMC4249489.
79. Forsythe ME, Love DC, Lazarus BD, Kim EJ, Prinz WA, Ashwell G, Krause MW, Hanover JA. Caenorhabditis elegans ortholog of a diabetes susceptibility locus: oga-1 (O-GlcNAcase) knockout impacts O-GlcNAc cycling, metabolism, and dauer. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(32):11952-7. doi: 10.1073/pnas.0601931103. PubMed PMID: 16882729; PubMed Central PMCID: PMC1567679.
80. Yang YR, Song M, Lee H, Jeon Y, Choi EJ, Jang HJ, Moon HY, Byun HY, Kim EK, Kim DH, Lee MN, Koh A, Ghim J, Choi JH, Lee-Kwon W, Kim KT, Ryu SH, Suh PG. O-GlcNAcase is essential for embryonic development and maintenance of genomic stability. Aging cell. 2012;11(3):439-48. doi: 10.1111/j.1474-9726.2012.00801.x. PubMed PMID: 22314054.
81. Slawson C, Zachara NE, Vosseller K, Cheung WD, Lane MD, Hart GW. Perturbations in O-linked beta-N-acetylglucosamine protein modification cause severe defects in mitotic progression and cytokinesis. The Journal of biological chemistry. 2005;280(38):32944-56. doi: 10.1074/jbc.M503396200. PubMed PMID: 16027160.
82. Macauley MS, Bubb AK, Martinez-Fleites C, Davies GJ, Vocadlo DJ. Elevation of global O-GlcNAc levels in 3T3-L1 adipocytes by selective inhibition of O-GlcNAcase does not induce insulin resistance. The Journal of biological chemistry. 2008;283(50):34687-95. doi: 10.1074/jbc.M804525200. PubMed PMID: 18842583; PubMed Central PMCID: PMC3259902.
83. Slawson C, Lakshmanan T, Knapp S, Hart GW. A mitotic GlcNAcylation/phosphorylation signaling complex alters the posttranslational state of the cytoskeletal protein vimentin. Molecular biology of the cell. 2008;19(10):4130-40. doi: 10.1091/mbc.E07-11-1146. PubMed PMID: 18653473; PubMed Central PMCID: PMC2555957.
84. Kim EJ, Kang DO, Love DC, Hanover JA. Enzymatic characterization of O-GlcNAcase isoforms using a fluorogenic GlcNAc substrate. Carbohydrate research. 2006;341(8):971-82. doi: 10.1016/j.carres.2006.03.004. PubMed PMID: 16584714.
85. Lazarus MB, 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. doi: 10.1038/nature09638. PubMed PMID: 21240259; PubMed Central PMCID: PMC3064491.

86. Macauley MS, Chan J, Zandberg WF, He Y, Whitworth GE, Stubbs KA, Yuzwa SA, Bennet AJ, Varki A, Davies GJ, Vocadlo DJ. Metabolism of vertebrate amino sugars with N-glycolyl groups: intracellular beta-O-linked N-glycolylglucosamine (GlcNGc), UDP-GlcNGc, and the biochemical and structural rationale for the substrate tolerance of beta-O-linked beta-N-acetylglucosaminidase. The Journal of biological chemistry. 2012;287(34):28882-97. doi: 10.1074/jbc.M112.363721. PubMed PMID: 22692202; PubMed Central PMCID: PMC3436527.
87. Hayre JK, Xu G, Borgianni L, Taylor GL, Andrew PW, Docquier JD, Oggioni MR. Optimization of a direct spectrophotometric method to investigate the kinetics and inhibition of sialidases. BMC biochemistry. 2012;13:19. doi: 10.1186/1471-2091-13-19. PubMed PMID: 23031230; PubMed Central PMCID: PMC3483245.
88. Paritala H, Carroll KS. A continuous spectrophotometric assay for adenosine 5'-phosphosulfate reductase activity with sulfite-selective probes. Analytical biochemistry. 2013;440(1):32-9. doi: 10.1016/j.ab.2013.05.007. PubMed PMID: 23711725; PubMed Central PMCID: PMC3714338.
89. Tabachnikov O, Shoham Y. Functional characterization of the galactan utilization system of Geobacillus stearothermophilus. The FEBS journal. 2013;280(3):950-64. doi: 10.1111/febs.12089. PubMed PMID: 23216604.
90. Sekine O, Love DC, Rubenstein DS, Hanover JA. Blocking O-linked GlcNAc cycling in Drosophila insulin-producing cells perturbs glucose-insulin homeostasis. The Journal of biological chemistry. 2010;285(49):38684-91. doi: 10.1074/jbc.M110.155192. PubMed PMID: 20926386; PubMed Central PMCID: PMC2992301.
91. Ruan HB, Singh JP, Li MD, Wu J, Yang X. Cracking the O-GlcNAc code in metabolism. Trends in endocrinology and metabolism: TEM. 2013;24(6):301-9. doi: 10.1016/j.tem.2013.02.002. PubMed PMID: 23647930; PubMed Central PMCID: PMC3783028.
92. de Alencar NA, Sousa PR, Silva JR, Lameira J, Alves CN, Marti S, Moliner V. Computational analysis of human OGA structure in complex with PUGNAc and NAG-thiazoline derivatives. Journal of chemical information and modeling. 2012;52(10):2775-83. doi: 10.1021/ci2006005. PubMed PMID: 22937904.
93. Dennis RJ, Taylor EJ, Macauley MS, Stubbs KA, Turkenburg JP, Hart SJ, Black GN, Vocadlo DJ, Davies GJ. Structure and mechanism of a bacterial beta-glucosaminidase having O-GlcNAcase activity. Nature structural & molecular biology. 2006;13(4):365-71. doi: 10.1038/nsmb1079. PubMed PMID: 16565725.
94. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nucleic acids research. 2008;36(Web Server issue):W5-9. doi: 10.1093/nar/gkn201. PubMed PMID: 18440982; PubMed Central PMCID: PMC2447716.

95. Cetinbas N, Macauley MS, Stubbs KA, Drapala R, Vocadlo DJ. Identification of Asp174 and Asp175 as the key catalytic residues of human O-GlcNAcase by functional analysis of site-directed mutants. Biochemistry. 2006;45(11):3835-44. doi: 10.1021/bi052370b. PubMed PMID: 16533067.
96. Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. Nature protocols. 2009;4(3):363-71. doi: 10.1038/nprot.2009.2. PubMed PMID: 19247286.
97. Macauley MS, Whitworth GE, Debowski AW, Chin D, Vocadlo DJ. O-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors. The Journal of biological chemistry. 2005;280(27):25313-22. doi: 10.1074/jbc.M413819200. PubMed PMID: 15795231.
98. Macauley MS, Vocadlo DJ. Increasing O-GlcNAc levels: An overview of small-molecule inhibitors of O-GlcNAcase. Biochimica et biophysica acta. 2010;1800(2):107-21. doi: 10.1016/j.bbagen.2009.07.028. PubMed PMID: 19664691.
99. Roquemore EP, Chevrier MR, Cotter RJ, Hart GW. Dynamic O-GlcNAcylation of the small heat shock protein alpha B-crystallin. Biochemistry. 1996;35(11):3578-86. doi: 10.1021/bi951918j. PubMed PMID: 8639509.
100. Chou CF, Smith AJ, Omary MB. Characterization and dynamics of O-linked glycosylation of human cytokeratin 8 and 18. The Journal of biological chemistry. 1992;267(6):3901-6. PubMed PMID: 1371281.
101. Matsuura A, Ito M, Sakaidani Y, Kondo T, Murakami K, Furukawa K, Nadano D, Matsuda T, Okajima T. O-linked N-acetylglucosamine is present on the extracellular domain of notch receptors. The Journal of biological chemistry. 2008;283(51):35486-95. doi: 10.1074/jbc.M806202200. PubMed PMID: 18948267.
102. Bond MR, Hanover JA. A little sugar goes a long way: the cell biology of O-GlcNAc. The Journal of cell biology. 2015;208(7):869-80. doi: 10.1083/jcb.201501101. PubMed PMID: 25825515; PubMed Central PMCID: PMC4384737.
103. Kreppel LK, Blomberg MA, Hart GW. Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. The Journal of biological chemistry. 1997;272(14):9308-15. PubMed PMID: 9083067.
104. Lubas WA, Frank DW, Krause M, Hanover JA. O-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. The Journal of biological chemistry. 1997;272(14):9316-24. PubMed PMID: 9083068.
105. Kazemi Z, Chang H, Haserodt S, McKen C, Zachara NE. O-linked beta-N-acetylglucosamine (O-GlcNAc) regulates stress-induced heat shock protein expression in a GSK-3beta-dependent manner. The Journal of biological

chemistry. 2010;285(50):39096-107. doi: 10.1074/jbc.M110.131102. PubMed PMID: 20926391; PubMed Central PMCID: PMC2998145.
106. Comer FI, Hart GW. O-Glycosylation of nuclear and cytosolic proteins. Dynamic interplay between O-GlcNAc and O-phosphate. The Journal of biological chemistry. 2000;275(38):29179-82. doi: 10.1074/jbc.R000010200. PubMed PMID: 10924527.
107. Jinek M, Rehwinkel J, Lazarus BD, Izaurralde E, Hanover JA, Conti E. The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha. Nature structural & molecular biology. 2004;11(10):1001-7. doi: 10.1038/nsmb833. PubMed PMID: 15361863.
108. Lazarus MB, Jiang J, Kapuria V, Bhuiyan T, Janetzko J, Zandberg WF, Vocadlo DJ, Herr W, Walker S. HCF-1 is cleaved in the active site of O-GlcNAc transferase. Science. 2013;342(6163):1235-9. doi: 10.1126/science.1243990. PubMed PMID: 24311690; PubMed Central PMCID: PMC3930058.
109. Lazarus MB, Jiang J, Gloster TM, Zandberg WF, Whitworth GE, Vocadlo DJ, Walker S. Structural snapshots of the reaction coordinate for O-GlcNAc transferase. Nature chemical biology. 2012;8(12):966-8. doi: 10.1038/nchembio.1109. PubMed PMID: 23103939; PubMed Central PMCID: PMC3508357.
110. Rodriguez AC, Kohler JJ. Recognition of diazirine-modified O-GlcNAc by human O-GlcNAcase. MedChemComm. 2014;5(8):1227-34. doi: 10.1039/C4MD00164H. PubMed PMID: 25068034; PubMed Central PMCID: PMC4109824.
111. Ma X, Liu P, Yan H, Sun H, Liu X, Zhou F, Li L, Chen Y, Muthana MM, Chen X, Wang PG, Zhang L. Substrate specificity provides insights into the sugar donor recognition mechanism of O-GlcNAc transferase (OGT). PloS one. 2013;8(5):e63452. doi: 10.1371/journal.pone.0063452. PubMed PMID: 23700425; PubMed Central PMCID: PMC3660302.
112. Jiang JY, Lazarus MB, Pasquina L, Sliz P, Walker S. A neutral diphosphate mimic crosslinks the active site of human O-GlcNAc transferase. Nature chemical biology. 2012;8(1):72-7. doi: Doi 10.1038/Nchembio.711. PubMed PMID: WOS:000298414100014.
113. Gross BJ, Kraybill BC, Walker S. Discovery of O-GlcNAc transferase inhibitors. Journal of the American Chemical Society. 2005;127(42):14588-9. doi: 10.1021/ja0555217. PubMed PMID: 16231908.
114. Schimpl M, Zheng X, Borodkin VS, Blair DE, Ferenbach AT, Schuttelkopf AW, Navratilova I, Aristotelous T, Albarbarawi O, Robinson DA, Macnaughtan MA, van Aalten DM. O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis. Nature chemical biology. 2012;8(12):969-74. doi: 10.1038/nchembio.1108. PubMed PMID: 23103942; PubMed Central PMCID: PMC3509171.

115. Leavy TM, Bertozzi CR. A high-throughput assay for O-GlcNAc transferase detects primary sequence preferences in peptide substrates. Bioorganic & medicinal chemistry letters. 2007;17(14):3851-4. doi: 10.1016/j.bmcl.2007.05.008. PubMed PMID: 17531489; PubMed Central PMCID: PMC3225185.
116. Bateman LA, Zaro BW, Chuh KN, Pratt MR. N-Propargyloxycarbamate monosaccharides as metabolic chemical reporters of carbohydrate salvage pathways and protein glycosylation. Chemical communications. 2013;49(39):4328-30. doi: 10.1039/c2cc37963e. PubMed PMID: 23235740.
117. Chen Y, Thon V, Li Y, Yu H, Ding L, Lau K, Qu J, Hie L, Chen X. One-pot three-enzyme synthesis of UDP-GlcNAc derivatives. Chemical communications. 2011;47(38):10815-7. doi: 10.1039/c1cc14034e. PubMed PMID: 21863157.
118. Zhao G, Guan W, Cai L, Wang PG. Enzymatic route to preparative-scale synthesis of UDP-GlcNAc/GalNAc, their analogues and GDP-fucose. Nature protocols. 2010;5(4):636-46. doi: 10.1038/nprot.2010.3. PubMed PMID: 20224564; PubMed Central PMCID: PMC2915437.
119. Zhai Y, Liang M, Fang J, Wang X, Guan W, Liu XW, Wang P, Wang F. NahK/GlmU fusion enzyme: characterization and one-step enzymatic synthesis of UDP-N-acetylglucosamine. Biotechnology letters. 2012;34(7):1321-6. doi: 10.1007/s10529-012-0910-y. PubMed PMID: 22456903.
120. Brown K, Pompeo F, Dixon S, Mengin-Lecreulx D, Cambillau C, Bourne Y. Crystal structure of the bifunctional N-acetylglucosamine 1-phosphate uridyltransferase from Escherichia coli: a paradigm for the related pyrophosphorylase superfamily. The EMBO journal. 1999;18(15):4096-107. doi: 10.1093/emboj/18.15.4096. PubMed PMID: 10428949; PubMed Central PMCID: PMC1171487.
121. Yu H, Chokhawala HA, Huang S, Chen X. One-pot three-enzyme chemoenzymatic approach to the synthesis of sialosides containing natural and non-natural functionalities. Nature protocols. 2006;1(5):2485-92. doi: 10.1038/nprot.2006.401. PubMed PMID: 17406495; PubMed Central PMCID: PMC2586341.
122. Lukose V, Whitworth G, Guan Z, Imperiali B. Chemoenzymatic Assembly of Bacterial Glycoconjugates for Site-Specific Orthogonal Labeling. Journal of the American Chemical Society. 2015. doi: 10.1021/jacs.5b07146. PubMed PMID: 26352466.
123. Klein AL, Berkaw MN, Buse MG, Ball LE. O-linked N-acetylglucosamine modification of insulin receptor substrate-1 occurs in close proximity to multiple SH2 domain binding motifs. Molecular & cellular proteomics : MCP. 2009;8(12):2733-45. doi: 10.1074/mcp.M900207-MCP200. PubMed PMID: 19671924; PubMed Central PMCID: PMC2816021.
124. Haynes PA, Aebersold R. Simultaneous detection and identification of O-GlcNAc-modified glycoproteins using liquid chromatography-tandem mass

spectrometry. Analytical chemistry. 2000;72(21):5402-10. PubMed PMID: 11080893.
125. Lazarus BD, Love DC, Hanover JA. Recombinant O-GlcNAc transferase isoforms: identification of O-GlcNAcase, yes tyrosine kinase, and tau as isoform-specific substrates. Glycobiology. 2006;16(5):415-21. doi: 10.1093/glycob/cwj078. PubMed PMID: 16434389.
126. Shen DL, Gloster TM, Yuzwa SA, Vocadlo DJ. Insights into O-linked N-acetylglucosamine ([0-9]O-GlcNAc) processing and dynamics through kinetic analysis of O-GlcNAc transferase and O-GlcNAcase activity on protein substrates. The Journal of biological chemistry. 2012;287(19):15395-408. doi: 10.1074/jbc.M111.310664. PubMed PMID: 22311971; PubMed Central PMCID: PMC3346082.