THE O-GLCNAC TRANSFERASE TETRATRICOPEPTIDE REPEAT DOMAIN INTERACTOME OFFERS INSIGHT INTO CELLULAR FUNCTIONS AND DISEASE MECHANISMS

by

HANNAH MICHELLE STEPHEN

(Under the Direction of Lance Wells)

ABSTRACT

The O-GlcNAc Transferase (OGT) is a nucleocytoplasmic glycosyltransferase that modifies nuclear and cytosolic proteins with a single β-N-Acetylglucosamine, creating a modification called O-GlcNAc. O-GlcNAc performs a myriad of cellular functions, including regulation of transcription, nutrient sensing, and modulation of cellular stress response. The O-GlcNAc modification is often considered analogous to protein phosphorylation, but while there are hundreds of protein kinases, there is only one mammalian OGT. Given that there are thousands of OGT substrates, a predominant question in the field has been: How does OGT select its substrates? It is believed that OGT substrate selection is mediated by its N-terminal tetratricopeptide repeat (TPR) domain, but this hypothesis has only been explored for a limited number of substrates. Adding additional impetus for the study of OGT protein-protein interaction is the recent discovery of several mutations in the TPR domain of OGT that are causal for X-Linked Intellectual Disability (XLID). Therefore, using the BioID and TurboID methods, we set out

to define a global OGT TPR interactome and to determine how this interactome is altered in XLID. We first defined a basal OGT TPR interactome in HeLa cells, identifying both known and novel OGT interactors with roles in transcriptional regulation, nuclear export, and chromatin remodeling. We then examined how the interactome changes in different cell types and under different cellular states. The OGT TPR interactome varies significantly between HeLa cells and the neuroblastoma cell line, SHSY5Y. Furthermore, the interactome can be altered by depolarization and serum starvation but is not affected by glucose availability in HeLa cells. Finally, we examined the interactomes for OGT XLID TPR variants. We identified four candidate interactors exhibiting significantly reduced interaction with OGT XLID variants as compared to wild-type: Tet2, Piccolo, Dim1, and ZC3H1. These proteins represent candidate mechanistic interactors underlying the OGT XLID phenotype. Future work will focus on describing the molecular consequence of these altered interactions and describing the function of novel OGT interactions.

INDEX WORDS: Protein glycosylation, O-GlcNAc, O-GlcNAc Transferase, Protein Interaction, X-Linked Intellectual Disability

THE O-GLCNAC TRANSFERASE TETRATRICOPEPTIDE REPEAT DOMAIN INTERACTOME OFFERS INSIGHT INTO CELLULAR FUNCTIONS AND DISEASE

by

HANNAH MICHELLE STEPHEN

BS, Lipscomb University, 2016

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2021

THE O-GLCNAC TRANSFERASE TETRATRICOPEPTIDE REPEAT DOMAIN INTERACTOME OFFERS INSIGHT INTO CELLULAR FUNCTIONS AND DISEASE

by

HANNAH MICHELLE STEPHEN

Major Professor:	Lance Wells
Committee:	Renee Barber
	Robert Haltiwanger
	Michael Tiemeyer

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2021

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my mentor, Lance Wells, for the support, freedom, and patience he has offered throughout my graduate studies. I would not be the independent scientist I am today without his mentorship. I would also like to thank my committee members, Michael Tiemeyer, Robert Haltiwanger, and Renee Barber, for their guidance and support throughout my dissertation work.

To all of my labmates, past and present: Chelsea Desbiens, Stephanie Halmo, Nithya Selvan, Rob Bridger, Jeremy Praissman, Linda Zhao, Osman Sheikh, Trevor Adams, Johnathan Mayfield, David Steen, thank you for your presence, support, guidance, and commiseration. To my undergraduate mentees: Jenny Lee, Elizabeth Cobb, Sandeep Yerraguntla, and Philip Schroeder, thank you for letting me learn with you.

To the memory of Brent Weatherly, thank you for teaching me everything I know about how to work with, understand, and be patient with proteomic data analysis.

To all of my colleagues at the CCRC, thank you for your scientific guidance and perpetual willingness to share equipment, reagents, and words of encouragement.

To my friends in the veterinary medical scientist training program, thank you for allowing me to nerd out about glycans and reminding me that I am 50% clinician too. To Dr. Kelsey Hart, thank you for your support and the many times you've gone to bat for me over the years. To Kate Sabey, thank you for being my statistician-on-call and dear friend.

To my family and friends, far and wide, thank you for taking an interest, believing in me, and never letting me give up. Especially to my mom, Becky, and my sister, Rachael, thank you for supporting me in too many ways to list.

To those who have reminded me what it is to be both a scientist and an artist: to Jaime Webb, Sarah Donnell, and all my classmates at the Alliance Theater, thank you for allowing me to stretch my right brain amidst all the logic of science.

And I would not be a veterinarian-to-be if I didn't say: thank you to my dog, Lady. Your love, sense of humor, and insistence that I go to bed at a reasonable hour have truly kept me sane over the last three years.

Finally, I am truly the product of all the educators and scientists who have taken the time and interest to teach me, both about science and life. Mr. Patterson, Mr. Gibson, Mr. Dodd, Dr. Enstminger, Dr. Conway, Dr. Gallaher, Dr. Mhlanga – I am a scientist today because of you. Thank you.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION TO DISSERTATION

Introduction to Dissertation

This chapter serves as an overview of the structure of this dissertation and a brief literature review covering the topics discussed therein. Generally, this dissertation covers the topic area of the mammalian O-GlcNAc transferase (OGT) and the study of its interactors on a global scale. Chapter two is a literature review previously published in Glycobiology in 2021, consisting of a deep literature review on the mechanisms underlying OGT substrate selection, as well as the substrate selection of its counterpart enzyme, the O-GlcNAcase (OGA). Chapters 3-5 are research chapters, with chapter 3 being a manuscript originally published in The Journal of Proteome Research in 2021 discussing the basal OGT interactome in HeLa cells. Chapters 4 and 5 are unpublished work and describe, respectively, various perturbations of the OGT interactome and the OGT interactome in X-Linked Intellectual Disability. Chapter 6 discusses conclusions from the results of the previous chapters and makes suggestions for future research building off this work.

The O-GlcNAc Transferase and the O-GlcNAc modification

The O-GlcNAc transferase (OGT) is a nuclear and cytoplasmic glycosyltransferase that modifies substrate proteins with a single β-N-acetylglucosamine on serine and threonine residues, resulting in a modification termed O-GlcNAc[1–3]. OGT and O-GlcNAc serve a vast breadth of cellular functions: chromatin remodeling, transcriptional regulation, nutrient sensing, cellular response to stress, etc.[4–8]. The O-GlcNAc modification is type of protein glycosylation, but is quite unique from “canonical” protein glycosylation which occurs in the secretory pathway[9].

Most protein glycosylation, forming sugar structures referred to as “glycans”, occurs on cell-surface and secreted proteins. These glycans are large, complex, and branched, and are also relatively static in that they are typically not turned over separately from their corresponding protein. The O-GlcNAc modification is different from this paradigm in almost every way: the modification consists of a single, non-extended monosaccharide; it occurs on nuclear and cytosolic proteins; and it is both inducible and dynamic on its protein substrates[10]. Given these factors, the O-GlcNAc modification is often considered more analogous to protein phosphorylation. In fact, O-GlcNAc and protein phosphorylation exhibit extensive crosstalk, often competing for the same residues on substrates[11–14]. Both modifications occur on thousands of proteins. However, one major difference divides these two post-translational modifications: the enzymes that create the modification. Hundreds of protein kinases exist with fairly tight substrate specificity; this is how evolution “decided” to handle the need for protein phosphorylation to occur across thousands of substrates under different conditions[15]. In contrast, only one mammalian O-GlcNAc transferase exists.

Considering the number and functional diversity of O-GlcNAc modified substrates, and the fact that OGT is essential for mammalian development[4], a major question in the O-GlcNAc field for some time has been: How does OGT select its substrates? It is believed that the answer lies in the domain structure of OGT. OGT consists of two primary domains: the N-terminal tetratricopeptide repeat (TPR) domain, which consists of 13.5 34 amino-acid repeats for the full-length OGT, and the C-terminal catalytic domain, which contains the transferase activity[16]. TPRs are known to be involved in protein-protein interaction in other proteins where they occur[17], so it has long been hypothesized that the TPR domain of OGT may be responsible for its substrate selection. Previous research has demonstrated that the TPR domain is essential for OGT multimerization[18] and modification of peptide substrates despite possessing no catalytic activity itself[18]. A limited number of OGT substrates have been shown to depend on the TPR domain for

OGT interaction and modification[5,19–21], but little research has been done on OGT protein-protein interaction, despite the establishment of a number of O-GlcNAc-omes[22]. In order to fully understand how OGT selects from among thousands of substrates, a deeper understanding of the role of the TPR domain in OGT protein-protein interaction must be obtained.

Additionally, our current limited knowledge of the mechanisms of OGT protein interaction hampers our understanding of the molecular mechanisms underlying OGT's function. While it is true that thousands of O-GlcNAc-modified proteins have been identified, and that these inevitably must interact with OGT, direct substrate interaction alone cannot describe the full breadth of OGT substrate selection. The possible mechanisms of OGT substrate selection are fully detailed in Chapter 2, but important to note here is the existence of "partner proteins"; that is, proteins that interact with OGT that may or may not be O-GlcNAc modified themselves that target OGT to other substrates or subcellular regions. These interactors may represent an essential missing piece of how OGT substrate selection is regulated and describing global OGT interactors is essential for identifying these proteins.

OGT and O-GlcNAc in disease

Given the wide scope of processes O-GlcNAc plays roles in, and the ubiquitous expression of OGT[2], it makes sense that O-GlcNAc would be implicated in many disease states. O-GlcNAc is a major player in type 2 diabetes by virtue of its nutrient sensing function[8]. The amount of O-GlcNAc modification in the cell directly correlates to the amount of glucose available to the cell, by virtue of the hexosamine biosynthetic pathway that synthesizes UDP-GlcNAc. In type 2 diabetes, excess extracellular glucose leads to increased O-GlcNAc and results in insulin resistance and β-cell apoptosis in the pancreas. O-GlcNAc has also been implicated in a variety of cancers, in both deleterious and protective roles[23]. O-GlcNAc is also thought to be a major contributor to the formation of neurofibrillary tangles in Alzheimer's Disease (AD). Reduced glucose metabolism in the brain in AD leads to reduced O-GlcNAc, contributing to

hyperphosphorylation of tau and promoting tau aggregation[24]. O-GlcNAc has also been suggested to play a cardioprotective role in myocardial infarction[25]. However, despite awareness of O-GlcNAc contributing to many disease states, these connections in many cases are only correlative. Understanding what proteins OGT interacts with and the molecular mechanisms underlying their function are an essential first step in being able to understand what mechanisms O-GlcNAc serves in these disorders and to allow for research into possible therapeutics.

Until recently, OGT had not been linked to developmental disorders, likely because OGT is essential for life so any null mutations would not result in living patients[4]. However, in recent years, several mutations in *OGT* have emerged as causal for X-Linked Intellectual Disability (XLID)[26–29]. XLID is an umbrella term for any intellectual disability disorder resulting from abnormalities on a gene located on the X-chromosome, and is diagnosed based off of patient presenting with an IQ of less than 70 and defects in adaptative functioning[30]. By nature of these abnormalities occurring on the X-chromosome, XLID predominately occurs in males. The first *OGT* mutations identified as causal for XLID were all localized to the TPR domain[26,27], with additional mutations localized to the catalytic domain being discovered later[28,29]. Since their discovery, a major question in the O-GlcNAc field has been: How do these mutations in OGT lead to XLID? While the catalytic domain mutations lead to reduced OGT activity *in vitro*, the XLID TPR variants are catalytically comparable to the wild-type OGT. Additionally, the TPR variant enzymes are thermodynamically stable, and patient lymphoblasts do not exhibit significant alterations in global O-GlcNAc. In the absence of a catalytic defect, we hypothesized that TPR mutations in OGT lead to altered protein-protein interaction, possibly with neurodevelopment-specific proteins, and that this defect in protein interaction leads to XLID.

Our approach to defining the OGT interactome

At the start of this research, the OGT interactome had never been defined. Therefore, we set out to define 1) the basal OGT interactome and 2) determine how the OGT interactome

changes with alterations in cellular state and 3) determine if the interactome is altered with OGT TPR XLID variants. We also sought to lend credence to the hypothesis that the TPR domain of OGT is responsible for substrate selection, and therefore focused our interactome studies on the TPR domain.

One challenge inherent to studying the interactome of an enzyme is that protein interactors are likely to be transient and thus difficult to identify with traditional affinity purification methods. To avoid this pitfall, we utilized the BioID method[31], and its derivative, TurboID[32]. BioID utilizes the promiscuous biotin ligase BirA*. When fused to a protein of interest, BirA* becomes a powerful tool to study protein-protein interaction. In our case, we utilized a TPR-BirA* fusion protein to capture OGT TPR interactors, and eGFP-BirA* as a negative control. BirA* catalyzes biotin to form a reactive biotin intermediate which then ligates to lysine residues of proteins in close proximity (interacting proteins). This biotin tag remains stable even if the interaction dissociates, thus making it a valuable tool for isolating even transient interactors. Interacting proteins can then easily be extracted with biotin affinity approaches and identified using mass spectrometry. Method details are described in chapters 3, 4, and 5. Our initial OGT TPR interactome studies utilized BioID, which requires a 24-hour labeling time. For later studies, we turned to TurboID, a BioID analogue that has a much faster labeling time, enabling us to study rapid alterations in the interactome on a time scale of 10 minutes.

Our basal OGT TPR interactome was defined in HeLa cells. From there, we expanded our study to the neuroblastoma cell line SHSY5Y, to examine OGT interactions in a neural cell system and as a cell line often used to study intellectual disability[33–35]. We explored several perturbations of cellular state to see how the OGT interactome would change: various glucose concentrations, depolarization, and serum starvation. Finally, we compared the wild-type TPR interactome to the interactome of the OGT XLID TPR variants. In doing so, we identified a number of candidate proteins representing interesting novel interactors for future study of OGT interactions, several

promising interactors that may underlie the mechanism of OGT XLID, and provided insight into the global landscape of the OGT interactome.

CHAPTER 2

REGULATING THE REGULATORS: MECHANISMS OF SUBSTRATE SELECTION OF THE O-GLCNAC CYCLING ENZYMES OGT AND OGA[1]

---

[1] Stephen, H.M, Adams, T.M., and Wells, L. 2021 *Glycobiology.* Online ahead of print.

Reprinted here with permission of the publisher

Abstract

Thousands of nuclear and cytosolic proteins are modified with a single β-N-Acetylglucosamine on serine and threonine residue in mammals, a modification termed O-GlcNAc. This modification is essential for normal development and plays important roles in virtually all intracellular processes. Additionally, O-GlcNAc is involved in many disease states, including cancer, diabetes, and X-Linked Intellectual Disability. Given the myriad of functions of the O-GlcNAc modification, it is therefore very interesting to note that in mammals that O-GlcNAc cycling is mediated by only two enzymes: the O-GlcNAc Transferase (OGT), which adds O-GlcNAc, and the O-GlcNAcase (OGA), which removes it. A significant outstanding question in the O-GlcNAc field is how it is possible for these two enzymes to solely mediate such an abundant and dynamic modification. In this review, we explore the current understanding of mechanisms of substrate selection for the O-GlcNAc cycling enzymes. These mechanisms include direct substrate interaction with specific domains of OGT or OGA, selection of interactors via partner proteins, post-translational modification of OGT or OGA, nutrient sensing, and localization alteration. Altogether, current research paints a picture of an exquisitely regulated and complex system by which OGT and OGA select substrates. We also make recommendations for future work, toward the goal of identifying interaction mechanisms for specific substrates which may be able to be exploited for various research and medical treatment goals.

Introduction

Thousands of proteins are post-translationally modified with O-GlcNAc, consisting of a single β-N-Acetylglucosamine attached to serine or threonine residues. This modification, discovered over 30 years ago[36], has since been implicated in nearly every intracellular biological process and many disease states. O-GlcNAc is involved in the regulation of cellular nutrient sensing, transcription, translation, and many other fundamental cellular processes[37]. Furthermore, O-GlcNAc has been shown to play important roles in a myriad of diseases that affect millions worldwide, including many forms of cancer, neurological disorders including Alzheimer's Disease, X-linked intellectual disability, Parkinson's Disease, and diabetes[8,38]. Such a modification with this many substrates and roles must be exquisitely regulated. This leads to one of the great mysteries of this field: How does regulation of the O-GlcNAc modification occur?

In the mammalian cell, the focus of this review, there are only two enzymes responsible for O-GlcNAc cycling: the O-GlcNAc Transferase (OGT), which adds the modification (**Fig. 2.1A)**, and the O-GlcNAcase (OGA), which removes it (**Fig. 2.2A)**. This is in direct contrast to the other highly prevalent intracellular post-translational modification, phosphorylation, which modifies a similar number of proteins as O-GlcNAc (and in fact often directly competes for O-GlcNAc sites), but is instead mediated by hundreds of protein kinases and dozens of protein phosphatases. In order for OGT and OGA to single-handedly modulate such an abundant and finely-tuned protein modification, the enzymes must exhibit complex mechanisms of substrate selection.

Significant work has attempted to define various mechanisms of OGT and OGA substrate interaction, revealing a complex network of both global and specific methods by which O-GlcNAc substrates are selected. In this review, we explore biochemical evidence for various hypotheses that comprise our current understanding of how protein selection for O-GlcNAc modification occurs. We also suggest future avenues for research to continue defining these mechanisms. For the reader interested in information on global cellular processes and disease mechanisms

affected by the O-GlcNAc modification, we refer to the excellent review published by the founder of the O-GlcNAc field, Dr. Gerald Hart[39].

The O-GlcNAc Transferase

Enzyme features

The O-GlcNAc Transferase (OGT) is encoded by a single gene found on the X-chromosome[4]. It consists of two primary domains: The N-terminal tetratricopeptide repeat (TPR) domain, and the C-terminal catalytic domain, which contains the glycosyltransferase activity (**Fig. 2.1B)**. The TPR domain consists of individual tetratricopeptide repeats, each 34 amino acids in length. The *OGT* gene encodes three different isoforms differing only in the length of their TPR domains. The full-length OGT (ncOGT) consists of 13.5 tetratricopeptide repeats[2,40]. The mitochondrial OGT (mOGT), localized only to the mitochondria and of debated importance/catalytic activity[41–43], has 9.5 TPRs. Finally, short OGT (sOGT) consists of 3.5 TPRs, and has very limited research on its role and function[44]. Since the vast majority of work on OGT substrate selectivity has used the full-length ncOGT, and since ncOGT seems to be responsible for the majority of intracellular O-GlcNAc over the other isoforms, this review will focus on this isoform.

A variety of hypotheses, not mutually exclusive, on how OGT is able to select from among thousands of substrates have arisen over several decades of research. Each of the major hypotheses is discussed below and summarized in **Fig. 2.1C**.

Substrate Selection: Consensus Sequence

The potential for identifying a consensus sequence that dictates where the O-GlcNAc modification can be made is appealing, considering that much of secretory pathway N-linked glycosylation is dictated by consensus sequence. Characterization done upon the discovery of the OGT enzyme suggested that high enrichment of S/T residues, along with a proline and acidic residues adjacent to the glycosylation site, might be necessary for glycosylation to occur[45]. Early

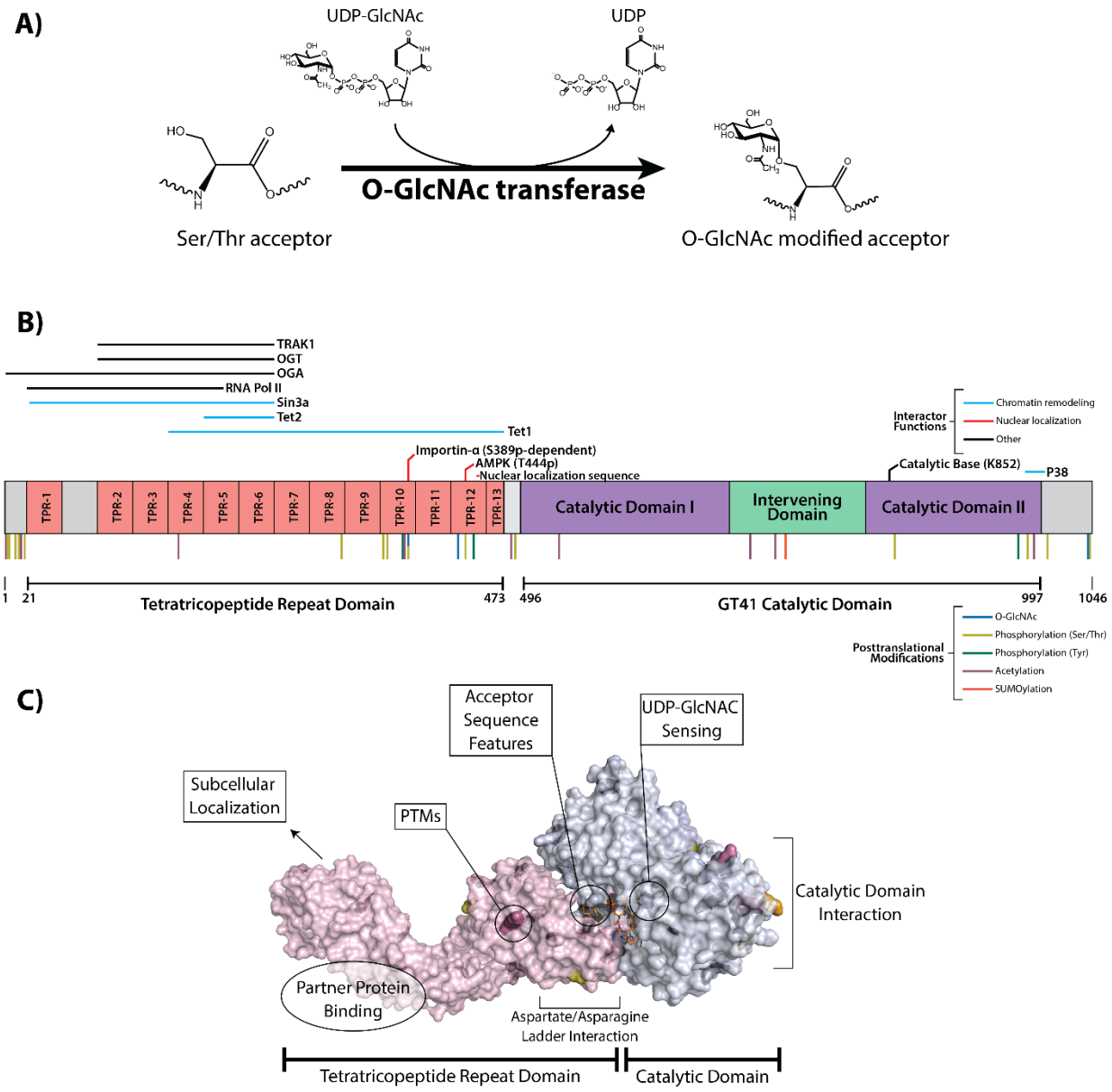


Figure 2.1: The O-GlcNAc Transferase. ***A,*** Schematic of the addition of O-GlcNAc to an acceptor substrate by OGT. ***B,*** Linear representation of the structure and domains of OGT. Interacting proteins and their location of interaction are shown above the structure. Known post-translational modifications are shown below. ***C,*** Possible mechanisms of OGT substrate selection, shown on a surface and cartoon representation of OGT in complex with a substrate peptide. A structure for the N-terminal TPR domain (PDB: 1W3B (Jinek et al. 2004)) was aligned with an overlapping structure containing the remaining C-terminal TPR and catalytic domains (PDB: 4GYY (Lazarus et al. 2012)) using PyMOL 2.4.1.

characterization identified prevalent “PVST” and “TTA” motifs near the site of glycosylation on O-GlcNAc modified peptides[48] somewhat reminiscent of “mucin-domains” for O-GalNAc modification in the secretory pathway[49]. Later work identified a degenerate consensus sequence, [TS][PT][VT][ST][RLV][ASY] based on structural evidence obtained by crystallizing OGT with synthetic peptides[50]. Additional research based on comparative sequence analysis and mutational evidence found that many OGT substrates have positively-charged residues (R/K) 7-10 residues upstream of the glycosylation site that contact Asp residues in the TPR domain[51]. Structural work performed later attempted to identify consistent secondary structures of OGT substrates and was unable to identify any consistent pattern to structure, order, or disorder of known OGT substrates[52].

A few themes can be identified from work attempting to define a consensus sequence, including S/T and acidic residue enrichment near the modification site, and a proline residue being near the glycosylation site. However, current research has been unable to identify one consistent consensus sequence for the O-GlcNAc modification. It is very likely that no universal consensus sequence exists, but rather specific features related to amino acid sequence or structural aspects may make a substrate more likely to be O-GlcNAc modified. These consensus features may play more of a role in determining whether or not a substrate is capable of ever being O-GlcNAc modified, rather than determining the temporal and spatial selection of a specific O-GlcNAc substrate.

Substrate Selection: Selection by OGT’s catalytic domain

One logical suggestion of how OGT may achieve substrate selectivity is by utilizing its two domains: the TPR domain and the catalytic domain. While the TPR domain remains the more popular choice for substrate selectivity (see below), work has also been done regarding the role of the catalytic domain itself in selecting substrates. It was suggested that OGT had a unique PIP3 binding domain on its C-terminus, which was capable of binding phosphoinositides, and that

this interaction during insulin signaling caused OGT to modify proteins involved in insulin signaling[53]. However, this binding activity has not been able to be recapitulated in future work[16].

One intriguing piece of evidence has found that a specific OGT interactor, p38, interacts exclusively with the catalytic domain of OGT[54]. This interaction was found not to change what proteins OGT interacts with, but induced OGT to O-GlcNAc modify the constitutive interactor neurofilament-H. To our knowledge, this is the only protein interactor that has been shown to interact only with the catalytic domain of OGT without involvement from the TPR domain.

Substrate Selection: Selection by OGT's TPR domain

By far the most attractive and most well-studied hypothesis for OGT's substrate selection mechanism is that the TPR domain of OGT is responsible for substrate binding and selection. TPRs, which occur in a variety of proteins, are known to be involved in protein-protein interaction[17]. This hypothesis is particularly appealing because it consists of several layers of potential regulation even beyond simply binding to substrates. The subhypotheses (that are not mutually exclusive!) are discussed below.

Subhypothesis 1: The TPR domain directly modulates substrate interaction

Early characterization of OGT revealed that the TPR domain, despite it lacking any catalytic activity itself, was essential for OGT's glycosylation activity. The catalytic domain of OGT alone, with no TPRs, has no catalytic activity even toward peptide substrates, and 3 TPRs are required for a peptide substrate to be glycosylated[18,40]. Furthermore, any truncation mutations reduced OGT activity toward protein substrates. Intriguingly, biochemical assays also revealed that the TPR domain, when co-expressed with full-length OGT and a known OGT substrate, is capable of competitively inhibiting substrate glycosylation, indicating that the TPR domain is a major player in substrate interaction[46]. This work pointed strongly toward the TPR domain of OGT being responsible for global substrate binding, and suggested a role for substrate specificity.

Significant structural work has been done to identify important residues and structural features in the TPR domain involved in OGT substrate selection. We will discuss this work generally here. For the reader desiring a deeper look into specific residues and mutation studies performed, we refer to the excellent review from the group of Dr. Suzanne Walker, who has pioneered much of the structural work on OGT[55].

Early structural work on OGT consisted of crystal structures of partial OGTs or OGT homologs. One crystal structure of the TPR domain of OGT revealed that hydrophobic residues in the TPR domain are responsible for dimerization of the TPR domain. This study also identified an “asparagine ladder” similar to that found in another TPR-containing protein, importin-α, which is responsible for peptide binding on that protein[46]. This asparagine ladder was studied in further detail years later and found to indeed be necessary for the binding of many substrates, although this interaction is with the peptide backbone rather than R-groups, so it is likely a global interaction mechanism rather than a mechanism of substrate specificity[56]. Early full-length crystal structures of bacterial homologs of OGT further confirmed an essential role for the TPR domain in OGT substrate binding[57,58]. The first human OGT crystal structure was completed with 4.5 TPRs, along with a model for the full-length OGT[16]. Intriguingly, this crystal structure identified the existence of a “latch” between TPRs 10 and 11 which allows the TPR domain to be highly mobile and alters the conformation of OGT upon binding a substrate. The authors suggested that positional changes of the TPR domain may play a greater role in OGT substrate selection than direct residue-determined interaction, and that these positional changes could be mediated by other factors including adapter proteins (see section *Partner Proteins* below). It is likely that there is a synergy between this mechanism and direct substrate binding, however, since further work has continued to identify specific residues and regions of the TPR domain that are involved in direct substrate interaction. Recent work has identified an aspartate ladder in the TPR domain of OGT, which unlike the asparagine ladder, coordinates with Thr side chains on bound substrates,

indicating that its role in protein interaction will vary depending on the substrate[51]. Mutational studies of this aspartate ladder revealed that it plays an important role in OGT specificity and selectivity.

While a number of OGT interactors have been generally identified, the most intriguing studies in regards to the question of OGT substrate selectivity are those that perform truncation mutations to identify exact regions of binding. Collectively, these studies paint a very interesting picture of OGT substrates whose binding to OGT is mediated by the TPR domain, and whose interaction region is unique to that specific substrate. The majority of proteins that have been studied in this way demonstrate the necessity of the TPR domain, and usually a specific region of the TPR domain, in order to be glycosylated (**Fig 2.1B)**. For RNA Pol II, the binding region necessary for it to be glycosylated is contained within the N-terminal 5.5 TPRs[59]. mSin3A requires TPRs 1-6 to interact with OGT, and these TPRs alone are also sufficient for interaction to occur[5]. Another protein, OIP106 (now known as TRAK1[60]) interacts with OGT through TPRs 2-6[61]. Deletion studies of OGT for the protein Tet1 found that the entire TPR domain, except TPRs 1-3, is required for Tet1 interaction with OGT[62]. For the protein Tet2, deletion of TPRs 5-6 abrogates its binding to OGT[20], and deletion of either TPRs 9-10 or 11-12 reduce interaction but do not abrogate it. This suggests that substrates can interact with OGT via several different regions of the TPR domain, with different repeats having varying necessity for interaction to occur. Intriguingly, OGT interacts with OGA through TPRs 1-6 but also requires the N-terminal amino acids prior to the first TPR to interact[63]. Collectively, these studies strongly suggest that different proteins interact with OGT via specific regions of the TPR domain unique to that protein. However, a very limited number of interactors have been studied in this way, so it is necessary to do additional truncation mutations with known OGT interactors to determine which proteins use this selection mechanism and where on the OGT protein they interact.

Subhypothesis 2: Partner proteins assist in targeting the TPR domain to substrates

While we have clear evidence that the TPR domain directly selects OGT substrates, this mechanism alone is likely insufficient to explain how OGT is able to select from among thousands of individual substrates. Additional modulatory effects must be in play. One answer to this problem is the suggestion that other proteins assist in targeting the TPR domain to specific substrates. Under this hypothesis, a specific protein (called a partner or adapter protein) may bind to the TPR domain of OGT, and it may or may not be functionally O-GlcNAc modified itself. This interaction then induces OGT to interact with a substrate protein which is O-GlcNAc modified to have a functional outcome. Very limited examples in the literature exist of this phenomenon. The most predominant example is OGT's interaction with mSin3a and HDAC1[5]. In the model presented, OGT interacts with mSin3a, via specific TPRs as described above, and OGT modifies mSin3a. OGT is then recruited to interact with HDAC1, a mSin3A interactor, which is also O-GlcNAc modified, which promotes transcriptional silencing. Additionally, a more global study found that knockdown of the protein MYPT1 caused reduction in O-GlcNAc modification of several proteins[60]. This suggests that MYPT1 may be a partner protein for OGT that targets it to other proteins and induces their O-GlcNAc modification. To our knowledge, this observation has not been further studied. Additional research identifying the proteins whose modification by OGT is dependent on MYPT1, as well as the functional outcomes of these proteins losing O-GlcNAc modification, will be helpful in confirming this partner protein hypothesis and enhancing our understanding of how it occurs within the cell.

One other intriguing suggestion of an OGT partner protein comes from a study performed in *E. coli*[64]. *E. coli* has no natural OGT or OGT substrates. The authors of this study found that expressing OGT alongside the partner protein Sp1 induces OGT to glycosylate many bacterial proteins that are not normally substrates, and that this is dependent on the TPR domain of OGT.

The authors suggest that Sp1 may act as a partner protein that induces a conformational change in the TPR domain, allowing for many substrates to then interact with OGT.

Substrate Selection: OGT subcellular localization is altered to affect substrate access and selection

OGT typically localizes primarily to the nucleus with some presence in the cytoplasm[2]. It has been suggested that OGT substrate selection is influenced by its subcellular localization, which has been documented to be affected by a variety of factors. OGT contains a nuclear localization sequence, originally thought to be between the TPR and catalytic domains[3], and later found to be three consecutive amino acids contained in the 12th TPR[65]. Importin-α interacts with this "DFP" motif and induces OGT's nuclear transport. This research also suggested that this interaction is dependent on OGT being O-GlcNAc modified at S389, which may reveal the DFP motif and allow importin-α interaction and nuclear transport. To our knowledge, no follow-up research has been performed following this finding, but it suggests that other interactions or post-translational modifications may influence OGT's localization. Further studies exploring localization, O-GlcNAc modification, and importin-α interaction may reveal if this nuclear localization mechanism is used to modulate OGT localization and substrate access.

Additionally, partner proteins (described above) may play a role in OGT subcellular localization, thus affecting its substrate access and selection. One such example was described above, in that mSin3a promotes OGT's localization at gene promoters. This has also been observed with OGT's interaction with both Tet2 and Tet3, where in both cases OGT interacts with the Tet protein and is recruited to chromatin to modulate transcription[20,21]. Intriguingly, both Tet proteins are O-GlcNAc modified but this does not seem to have any effect on their function. Finally, during DNA damage, OGT O-GlcNAc modifies the protein H2AX, and is then recruited to the site of DNA damage where it is thought to restrain the DNA damage signaling to prevent its expansion[66].

This hypothesis may also have some overlap with the contribution of post-translational modification to OGT selection and localization (discussed further below). During cytokinesis, Chk1 phosphorylates OGT, inducing it to localize to the midbody, where it O-GlcNAc modifies vimentin and may contribute to the completion of cytokinesis in conjunction with vimentin phosphorylation[67].

Substrate Selection: Post-Translational Modifications

It has been suggested that OGT may be regulated by post-translational modifications. OGT is known to be extensively post-translationally modified with a variety of modifications, including O-GlcNAc, phosphorylation, and acetylation[68]. However, very limited work has been done at this time exploring how post-translational modification of OGT affects its substrate specificity. One published work demonstrates that OGT is phosphorylated by AMP-activated protein kinase (AMPK) at T444, and this phosphorylation induces OGT nuclear localization (Bullen et al. 2014). Furthermore, mutation of the phosphorylated residue causes changes in the profile of global O-GlcNAc modifications, indicating that the phosphorylation modification induces alterations in OGT substrate selectivity. Additionally, one study focusing on the short isoform of OGT (sOGT) found that six O-GlcNAc sites on OGT induce alterations in sOGT substrate binding and global O-GlcNAc modifications, indicating that these sites play a role in substrate specificity for this isoform[44]. Other post-translational modifications on OGT exist (**Fig. 2.1B)**, and additional work is needed in this area to identify how OGT post-translational modifications affect its activity in various cellular contexts.

Substrate Selection: UDP-GlcNAc Sensing

One final possible avenue for the modulation of OGT substrate selectivity is related to its nutrient sensing. It has been well-documented that O-GlcNAc levels are responsive to glucose levels in the cell given that synthesis of the donor molecule for O-GlcNAc, UDP-GlcNAc, is dependent on the hexosamine biosynthetic pathway which branches from glycolysis. It has also

been documented *in vitro* that OGT is inhibited by excess amounts of free UDP, suggesting a possible feedback mechanism[16]. Early work in an *in vitro* peptide glycosylation assay showed that O-GlcNAc modification of OGT substrate peptides was affected by UDP-GlcNAc levels. However, the connection between UDP-GlcNAc levels correlating with O-GlcNAc levels and how OGT substrate selectivity is altered by UDP-GlcNAc levels has not been explored to our knowledge. Global analysis of OGT interactors under various glucose status and UDP-GlcNAc levels would help uncover whether or not UDP-GlcNAc levels play a role in OGT substrate selectivity, and if so, by what mechanisms this occurs.

Summary of possible OGT substrate selection mechanisms

The nature of OGT as the sole enzyme modifying thousands of proteins with O-GlcNAc demands that the regulation of its substrate selection be elegant and complex. Many different possible mechanisms of substrate selection have been studied for a limited number of substrates. It is essential to note that the reality of OGT substrate selection is likely an intricate combination of those described above. The substrate selection mechanism is likely to be unique for each protein or protein class and exist as a combination of global interaction mechanisms, global cellular effects (from nutrient status), specific pathways designated for that protein (involving partner proteins and OGT post-translational modifications), and interaction features unique to that interactor (specific amino acid interactions). In order to better understand OGT substrate selection, a combination of targeted research, like studies described above, and global assays must be performed.

To the end of understanding OGT interaction on a global scale, a limited number of attempts to define the OGT interactome have been performed. Most have focused on the full-length OGT interactome, which is helpful to define what proteins interact with OGT, but cannot determine domain effects on protein-protein interaction. These include Co-IP and microarray interactomes[70,71]. One OGT TPR interactome, which identified 115 OGT-TPR interacting proteins,

has been recently defined by our group demonstrating that the TPR domain of OGT can select for substrates/partners[72].

The O-GlcNAcase

Enzyme Features

The O-GlcNAcase (OGA) is responsible for the removal of the O-GlcNAc modification added by OGT (**Fig. 2.2A**). Like OGT, it is the only mammalian OGA, but exists on somatic chromosome 10[73]. Also similar to OGT, it consists of two distinct domains connected by a stalk domain. The N-terminal domain is the catalytic domain containing the O-GlcNAcase activity, and the C-terminal domain is termed the HAT-like, or histone acetyltransferase-like, domain (**Fig. 2.2B**). Early characterization of the OGA domains suggested that this C-terminal domain contained histone acetyltransferase activity[74], but this was disproven in later studies and the function of this domain is currently unknown[75–77]. OGA also exists as two distinct isoforms, one being the full length OGA and a shorter isoform nvOGA, or “nuclear variant” OGA, so termed because the full length OGA localizes primarily to the cytoplasm and nucleolus[73,78,79], whereas the shorter variant localizes more dominantly to the nucleus[80]. Like OGT, the vast majority of work on OGA has focused on the full-length isoform, and so this review will focus on this isoform as well.

Comparatively much less work has been done characterizing OGA’s mechanisms of substrate selection. What is currently known is discussed below.

Mechanisms of Substrate Selection

Whether or not the O-GlcNAcase enzyme exhibits substrate selectivity is somewhat a subject of debate. One study suggested that OGA doesn’t significantly recognize protein structure or sequence and rather only recognizes the sugar moiety, based on biochemical data showing that the kinetic parameters for OGA are largely unaltered by different substrates, unlike OGT[81]. However, later structural studies of OGA contest this claim. One structural and biochemical study

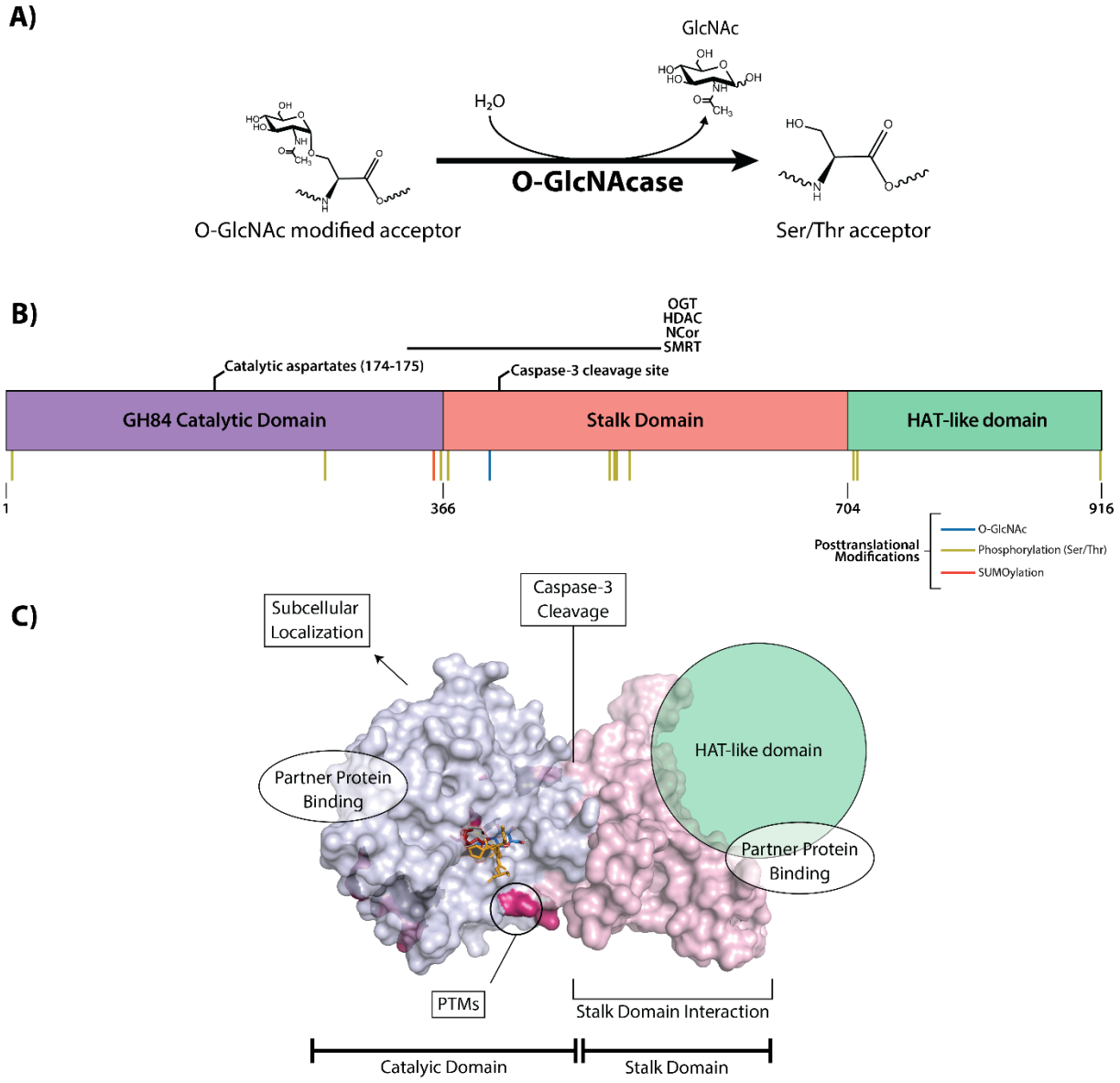


Figure 2.2: The O-GlcNAcase. ***A,*** Schematic of the removal of O-GlcNAc from an acceptor substrate by OGA. ***B,*** Linear representation of the structure and domains of OGA. Interacting proteins and their location of interaction are shown above the structure. Known post-translational modifications are shown below. ***C,*** Possible mechanisms of OGA substrate selection, shown on a surface and cartoon representation of OGA in complex with a substrate O-GlcNAc modified glycopeptide. A structure of OGA in complex with thiamet-G (PDB: 5UN93) was aligned with a structure of an OGA D175N-mutant in complex with a glycopeptide substrate (PDB: 5VVU (Li et al. 2017 a)) using PyMOL 2.4.1.

performed mutations in the stalk domain of a bacterial homolog of OGA[83]. The authors found that these mutations affected substrate sugar cleavage and did so uniquely for each substrate tested. This suggests that the stalk domain may play a role in substrate recognition and selection. Structural modeling data on human OGA further suggested that when dimerized, the stalk domain contributes to substrate interactions specifically by interacting with hydrophobic side chains of substrates[84]. Two concurrently published human OGA structures similarly identified this groove in the dimer as of potential importance to OGA substrate binding[82,85]. What remains unclear from this data is whether these substrate interaction mechanisms are simply mechanisms by which OGA binds to all substrates or if they also contribute to selection. It is also important to note that two of these structural studies were performed on truncation mutations of OGA lacking the HAT-like domain, and the third was performed on both domains of OGA but crystallized separately and modeled together. It is possible that the HAT-like domain also contributes to OGA substrate selection, but little is known on its potential roles in this area.

An intriguing potential mechanism for substrate recognition for OGA lies in a unique cleavage event identified early on in OGA characterization. OGA is cleaved between its two domains by the apoptosis-executor protease caspase-3, and this cleavage surprisingly has no effect on its catalytic activity[78]. A later study showed that this cleavage occurs during apoptosis and results in the N- and C-terminal fragments continuing to associate, thus retaining catalytic activity[75]. The authors of this study suggest that the cleavage event may have an effect on OGA substrate binding and selectivity. However, additional research is necessary to determine if and to what extent caspase-3 cleavage might alter OGA substrate selectivity.

Some specific interactors of OGA have been identified that may offer insights into mechanisms of substrate selection (**Fig. 2.2B**). In one structural study mentioned previously, several OGA mutations in the stalk domain eliminate OGA catalytic activity toward the glycoprotein substrates TAB1, FOXO, and CREB[83]. Additionally, different mutations in some

cases have different effects on OGA activity for specific substrates, suggesting a possible differential binding mechanism for this region of OGA. Further specific substrate studies corroborate the importance of the stalk domain for OGA substrate binding. One study found that OGA interacts with several substrates via amino acids 336-548[63]. These amino acids are located between the catalytic domain and stalk domain, as defined by one of the manuscripts first defining the structural characteristics of human OGA[82]. These amino acids were identified as essential for OGA to interact with OGT, HDAC, NCor, and SMRT[63]. This manuscript also intriguingly identified that OGA interacts with the corepressor Sin3A, but that this interaction requires OGT to also be present. This points to a potential "partner protein"-like mechanism for OGA as well as OGT.

Another recent manuscript that identified global OGA interactors supports this hypothesis as well[86]. This manuscript used the BioID approach[31] to identify OGA interactors under oxidative stress conditions, and found several intriguing results with implications for our understanding of OGA substrate selection. First, the authors found that inhibiting OGA did not significantly alter the profile of proteins that were labeled as interactors with biotin. This strongly supports the notion of partner proteins for OGA; that is, proteins that interact with OGA and are not necessarily substrates themselves, but may affect OGA's activity through a myriad of mechanisms. The authors also identified the protein FAS as an OGA interactor and found that under oxidative stress, FAS interaction with OGA reduces its global activity. A similar phenomenon was observed when OGT interacts with OGA in T-cell viral infection, where the OGT interaction reduces OGA activity[87]. This points to a potential unique mechanism for OGA regulation in which partner proteins may affect its activity globally to the end of increasing overall O-GlcNAc levels.

It is also possible that, like OGT, OGA substrate selectivity is regulated by altered localization. However, little work has been done in this area. As mentioned previously, OGA is known to localize to the cytoplasm and nucleolus[73,78,79], and one study demonstrated increased

nucleolar localization under nucleolar stress[79]. Additional studies are necessary to determine how OGA localization changes due to cellular state and what factors drive these localization changes. Additionally, OGA is post-translationally modified (**Fig 2.2B**), but any role that these post-translational modifications may play in substrate selection has yet to be elucidated. Possible mechanisms of substrate selection are summarized in **Figure 2C.**

Current Understandings and Recommendations for Future Study

Multiple studies have focused on the O-GlcNAc cycling enzymes since their discovery several decades ago. Much of this work has focused on the functional outcomes and disease roles that these enzymes and the modification play. However, one fundamental question still remains unanswered: How are only two enzymes capable of modulating such an abundant and finely-tuned protein modification with so many diverse roles? Most of the work in this area has focused on the transferase, OGT. Many probable and mutually inclusive hypotheses exist to explain how OGT can select substrates. Most evidence points strongly toward a significant role for the TPR domain of OGT in substrate selection, via a variety of mechanisms including direct substrate selection and/or partner protein interactions. However, other mechanisms have been identified and may play varying roles, including post-translational modification of OGT, alterations in OGT localization, and OGT nutrient sensing. OGA, to the contrary, has had comparatively much less research performed to identify how it selects substrates. Some evidence exists for selection via structural features of the OGA enzyme, and a few potential “partner proteins” have been identified. It is also possible that OGA regulation occurs uniquely on a more global scale; e.g. interacting proteins like those described above dampen its activity globally when necessary. However, many mysteries still exist regarding OGA substrate selection. One such large gap in knowledge is the function of the HAT-like domain.

Research into OGT’s substrate selection mechanism has reached a point where highly targeted studies are possible and will be beneficial. Several global interactor lists have been

identified, and many specific interacting substrates have been studied as well. The next logical step is to focus on specific interactors and identify several characteristics.

1) By what mechanism does the interactor bind OGT? What domain is responsible for the interaction? Do post-translational modifications on OGT or the interactor play a role? Is the interactor O-GlcNAc modified?

2) How does the interaction affect OGT? Is localization altered? Does the interaction induce interaction with other substrates?

3) Under what conditions does the interaction occur? What is the functional outcome of the interaction?

One challenge that will come with validating and identifying individual interactions is that by nature many OGT interactions are transient, making them difficult to identify. A combination of biochemical techniques, including truncation mutations and *in vitro* binding assays to determine binding regions, and *in cellulo* methods like FRET[88] and bimolecular fluorescence complementation[89] will be useful in determining where and when interactions occur in cells. Methods used to determine functional outcomes will vary with the interactor, but since OGT interacts with many proteins involved in well studied systems (e.g. transcription, chromatin regulation, heat shock), many assays are available for this purpose.

In order to understand how OGA selects its substrates, more preliminary work is needed. Truncation studies, like those performed for OGT, to determine what region (s) of OGA are responsible for its interaction with known substrates will be useful. With currently available proteomic and proximity labeling technology, it may also be extremely useful to use these approaches to identify interactors of the various domains of OGA. This approach can be used to distinguish between proteins that interact with the catalytic domain, the stalk domain, or the HAT-like domain. Structural studies focusing on the HAT-like domain may also be of use to determine if this domain plays a role in substrate selection and, if not, what function, if any, it serves. Once

a more unified theory of OGA substrate interaction is identified, studies for individual interactors like described above for OGT can take place. It may also be useful to determine OGA's interaction with OGT substrates concurrently, if the cycling of the O-GlcNAc modification on a given substrate is of particular interest.

The O-GlcNAc cycling enzymes present a fascinating story of an essential protein modification that plays a role in nearly every biological process within the cell but is regulated by only two enzymes. Uncovering the complicated and elegant mechanisms for how this regulation occurs is essential to understand how these enzymes and the resulting modification interplay with multiple biological processes related to diseases that affect millions of people worldwide.

Funding

This work was supported by a grant from the W.M. Keck foundation (L.W. Co-PI); an NICHD National Institutes of Health (NIH) [grant R21HD097652] (L.W); and an NICHD Grant [F30 HD098828] (H.S.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Acknowledgments

We would like to thank all authors of the works cited here. Although we have attempted to be thorough in our review of the literature, with thousands of papers published in the O-GlcNAc field, we may have missed a few relevant manuscripts. Any exclusion is unintentional.

CHAPTER 3

GENERATION OF AN INTERACTOME FOR THE TETRATRICOPEPTIDE REPEAT DOMAIN OF O-GLCNAC TRANSFERASE INDICATES A ROLE FOR THE ENZYME IN INTELLECTUAL DISABILITY[2]

---

[2] Stephen, H.M, Praissman, J.L, and Wells, L. 2021 *Journal of Proteome Research.* 20(2): 1229-1242

Reprinted here with permission of the publisher

Abstract

The O-GlcNAc transferase (OGT) modifies nuclear and cytoplasmic proteins with β-N-Acetyl-Glucosamine (O-GlcNAc). With thousands of O-GlcNAc modified proteins but only one OGT encoded in the mammalian genome, a prevailing question is how OGT selects its substrates. Prior work has indicated that the tetratricopeptide repeat (TPR) domain of OGT is involved in substrate selection. Furthermore, several variants of OGT causal for X-linked intellectual disability (XLID) occur in the TPR domain. Therefore, we adapted the BioID labeling method to identify interactors of a TPR-BirA* fusion protein in HeLa cells. We identified 115 interactors representing known and novel O-GlcNAc modified proteins and OGT interactors (Raw data deposited in MassIVE, Dataset ID MSV000085626). The interactors are enriched in known OGT processes (e.g. chromatin remodeling) as well as processes in which OGT has yet to be implicated (e.g. pre-mRNA processing). Importantly, the identified TPR interactors are linked to several disease states but most notably are enriched in pathologies featuring intellectual disability that may underlie the mechanism by which mutations in *OGT* lead to XLID. This interactome for the TPR domain of OGT serves as a jumping off point for future research exploring the role of OGT, the TPR domain, and its protein interactors in multiple cellular processes and disease mechanisms, including intellectual disability.

Keywords: OGT, TPR, O-GlcNAc, Proximity Proteomics, Mass Spectrometry, Biotin ligase, X-Linked Intellectual Disability

Introduction

The O-GlcNAc transferase (OGT) is a nucleocytoplasmic glycosyltransferase that modifies substrate proteins with a β-N-acetylglucosamine (O-GlcNAc) on serine and threonine residues. OGT is a unique mammalian glycosyltransferase in that it modifies intracellular proteins outside of the secretory pathway, and the O-GlcNAc modification it creates is non-extended, dynamic, and inducible[10,90,91]. The O-GlcNAc modification is often compared to phosphorylation, given their similar characteristics and the fact that both occur on thousands of nuclear and cytosolic proteins[12,13,92]. In fact, OGT and Ser/Thr kinases often compete for the same sites on certain protein substrates[92] and can regulate each other by post-translational modification[14,67]. However, unlike protein phosphorylation which is mediated by hundreds of kinases, there is only one gene encoding intracellular O-GlcNAc Transferase in mammals. Thus, not surprisingly, OGT is necessary for the development of mammalian life[4] and is involved in many intracellular processes including nutrient sensing, transcription, and cellular stress[39]. OGT also has been implicated in many diseases including cancer, Alzheimer's disease, diabetes, and more recently, in X-Linked Intellectual Disability (XLID) as identified originally by our team in collaboration with clinical partners[26] and further confirmed and expanded on by our group and others[27–29].

Given the wide diversity of OGT substrates and functions, and the existence of only one mammalian OGT, a prevailing question in the O-GlcNAc field is how OGT selects its substrates. Previous research suggests that the N-terminal tetratricopeptide repeat (TPR) domain of OGT (consisting of 13.5 repeats in the full-length protein), rather than its C-terminal catalytic domain, is responsible for OGT substrate selectivity[46,56,61]. However, the hypothesis of the TPR domain mediating protein-protein interactions has only been directly tested for a few select proteins[5,19,20,59]. A few early attempts to define the full-length OGT-interactome using co-immunoprecipitation have also been made[70,71]. An unbiased approach to identifying proteins that interact specifically with the TPR domain would lend further support to the hypothesis of the TPR domain mediating OGT substrate selectivity, and allow for the identification of new potential

substrates and “partner proteins”, which interact with the TPR domain of OGT to target it to specific subcellular regions and/or protein complexes.

An additional impetus for TPR interaction studies is the observation that the majority of reported missense mutations in *OGT* causal for XLID are localized to the TPR domain and do not grossly affect catalytic activity or stability, suggesting a potential protein-protein interaction-based mechanism[26,27]. Therefore, to demonstrate that the TPR domain of OGT is capable of substrate selection, and to capture endogenous OGT TPR interactors including transient interactors, we took advantage of the BioID method, utilizing a fusion protein consisting of the full-length OGT TPR domain with a modified biotin ligase in place of the catalytic domain of OGT.

BioID is a well-established proximity proteomic labeling method that utilizes a promiscuous biotin ligase (BirA*) to label nearby proteins with biotin, which allows them to easily be extracted and identified using mass-spectrometry based proteomics[31,93–95]. Using a TPR-BirA* fusion protein in a HeLa cell system (and a eGFP-BirA* fusion protein as a negative control), we have identified over 100 OGT TPR interactors, including both known and novel OGT substrates and interactors. This work strongly suggests that the TPR domain, through protein-protein interactions, plays a major role in OGT substrate selectivity. Exploiting these interactions may allow for fine-tuning of the modification of specific O-GlcNAc modified substrates which has been explored using other techniques[96,97]. This protein set also further supports OGT’s role in many cellular processes and reveals potential novel pathways in which O-GlcNAc may play an intricate role. Finally, the interactome is highly enriched in proteins involved in neurological disorders that present with intellectual disability. These proteins represent a set of candidate interactors to explore for future mechanistic studies of the functional role of OGT and the O-GlcNAc modification in XLID.

Experimental Procedures

Plasmid Constructs

Plasmids for proximity proteomics were constructed on a CMV promoter with a C-terminal BirA*. Fusion gene construction is as follows. For TPR BirA*: TPR - 3X GGGGS linker - BirA* - 2X FLAG. For eGFP-BirA*: eGFP - 3X GGGGS Linker - BirA* - 2XFLAG. The BirA* sequence was obtained from the original paper describing BioID[31]. For TPR-BirA*, residues 1-473 of OGT (consisting of the 13.5 TPR repeats of full-length OGT, Uniprot Accession O15294) were used. Full fusion protein sequences are in Supplementary Table 1.

Cell culture and BioID

HeLa cells were grown in DMEM with 10% FBS on 14.5cm plates, passaged for maintenance every 4-6 days (1:4-1:10). For expression of fusion proteins, cells were transfected at ~70% confluency using lipofectamine 2000 (Invitrogen) according to manufacturer recommended ratios with 60µg plasmid DNA for TPR-BirA* and 6µg plasmid DNA for eGFP-BirA* (eGFP-BirA* expresses at a much higher level than TPR-BirA* - see **Fig. 3.1 C/D**). After 24 hours, cell media was replaced with media containing 50uM biotin to induce labeling for 24 hours. Cells were then collected and subjected to nucleocytoplasmic lysis to collect protein. Briefly, cells were lysed in hypotonic buffer A (10mM Tris-HCl pH 5.5, 500uM DTT, 500uM EDTA, protease and phosphatase inhibitor cocktails (Sigma)) for 10 minutes, vortexed twice for 30 seconds, then lysed in hypertonic buffer B (10mM Tris-HCl pH 7.5, 800mM NaCl, protease and phosphatase inhibitors, and universal nuclease (ThermoFisher Scientific)) for 5 minutes, followed by 30 seconds of high-speed vortexing and an additional 5 minute incubation. Membranes were pelleted by centrifugation, and the supernatant containing cytosolic and nuclear proteins used in further analysis. Three biological replicates were performed, each consisting of one TPR-BirA* and one eGFP-BirA* sample, with TPR-BirA* representing the experimental condition and eGFP-BirA* representing the negative control. For supplementary figure 1, HeLa cells were transfected as described and lysed in RIPA buffer.

Quantification of FLAG-tagged proteins

FLAG-tagged fusion proteins were quantified from HeLa cell lysate using a FLAG ELISA (Cayman Chemical) according to the manufacturer's protocol.

Identification of biotinylated proteins

Biotinylated proteins were purified from cellular lysate using an anti-biotin immunoprecipitation protocol as previously described[98]. 25% of eluate was reserved for anti-biotin western blot, then remaining biotinylated proteins were then run out on a 10% SDS-PAGE gel using the Bio-Rad Mini-PROTEAN gel system. The gels were not stained. Each lane was cut into four equal fractions based on molecular weight markers, then extracted, reduced, alkylated, and digested with trypsin as previously described[98]. Dried peptides were resuspended in 40µL sample buffer (10% Buffer B (80% acetonitrile, 0.1% formic acid), 90% Buffer A (0.1% formic acid), and 15uL of this was injected for each instrument run. LC-MS/MS was performed on an Orbitrap Fusion Tribrid mass spectrometer (ThermoFisher Scientific) equipped with an Ultimate 3000 RSLCnano HPLC system (Thermofisher Scientific). Peptides were separated on an Acclaim™ PepMap™ RSLC C18 column (75 µm ID × 15 cm; 2 µm particle size) at a flow rate of 0.200µL/min over a 150min linear gradient of 1-99% Buffer B with a total run time of 180min. Precursor scans were collected using the Orbitrap mass analyzer with a scan range of 300-2000m/z and mass resolution of 60,000. Most intense ions were fragmented using 38% CID collision energy and detected in the Ion Trap with 1 microscan and dynamic exclusion for 15 seconds after one occurrence. Samples were run in randomized pairs within a replicate, with each pair consisting of a corresponding gel fraction in eGFP-BirA*, run first, then TPR-BirA*, run second, with one 90 min wash in between each sample and two 90 minute washes in between pairs (20uL 10% Buffer B injection for washes). The raw data for all 24 LC-MS/MS analyses (12 control and 12 experimental) has been deposited to the MassIVE database (https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp, Dataset ID: MSV000085626).

Data analysis

Raw files were searched with no prior peaklist selection by both ProteinMetrics Inc. Byonic (v3.8.13) and ThermoFisher Scientific Proteome Discoverer - SequestHT (2.2.0.338). The proteomic database consisted of all SwissProt annotated human protein sequences (obtained 09-2019 – 20,434 sequences), plus a list of common contaminants (trypsin, keratins, and serum albumins – 179 sequences, available in Supplementary Table 1) and the sequences for eGFP and BirA*. A concatenated database was generated for FDR calculations by including reversed protein sequences for all proteins in the database, creating a final database of 41230 sequences, all of which were searched unbiasedly. Tryptic cleavage was fully specific at Arg and Lys, with two missed cleavages allowed. For both search engines, precursor mass tolerance was 20ppm and fragment mass tolerance was 0.5Da. Carbamidomethylation on Cys was set as a fixed modification, and variable modifications were allowed: Oxidation of Met, HexNAc on Ser/Thr, and Biotin on Lys. Full peptide lists were exported from PMI-Byonic and SequestHT with no score cuts.

Peptide lists from each sample (four fractions each) were then combined using ProValt[99]. Peptide cuts were made based on peptide FDR, set at 2% for the eGFP-BirA* and 1% for TPR-BirA* - with the lower stringency of filtering for eGFP-BirA* samples selected to help decrease the incidence of potential false positives in the final protein set. Protein lists were generated by removing all peptides not identified by both Byonic and Proteome Discoverer (peptides required to have a positive Byonic Score and Sequest Xcorr), and by removing any proteins that were identified by fewer than two peptides. Only proteins identified as the top protein from among isoforms were kept in the final protein list. Spectral counts were converted to normalized spectral abundance factors (NSAFs) for analysis[100]. The normal logarithm (ln) of NSAF values was computed to create a more Gaussian/normal distribution, and the Student's T-Test was utilized to determine the significance of the difference in abundance between experimental and negative

control conditions for each potential interacting protein. Protein IDs with a pvalue of 0.05 or less and with at least a 3-fold higher NSAF were kept in the final TPR-BirA* interactome.

Further quantification was carried out by reconstructed ion chromatogram (RIC) analysis of peptide intensity performed in Xcalibur Qual Browser (v2.0.3.2). Peptides for analysis were selected based on their appearance in all three replicates in both TPR-BirA* and eGFP-BirA*. Peptide intensity was examined for all gel fractions in which that peptide occurred (For HCF1: fractions 1 and 2; for OGT: fraction 1 [rep 3] and fraction 2 [reps 1,2]; for KNL1: fraction 1). Time ranges for intensity analysis were selected to be the same in TPR-BirA* and eGFP-BirA*. The time range was selected based on the area of overlap between the corresponding peaks, or, in the case of peaks with a slight time offset leading to no overlap, the time range was selected so that the time was evenly split between the two peaks. Peptide intensity was determined by the normalization level (NL) of the monoisotopic peak. Peptide identity was validated by recorded retention time in ProteomeDiscoverer 2.2.

Western Blots and Antibodies

SDS-PAGE gels (4-15%) were run using the BioRad Mini-PROTEAN gel system. Gels were transferred onto Immobilin-FL PVDF membranes (Sigma) using the BioRad Trans-Blot SD Semi-Dry Transfer Cell. Membranes were blocked in 1% cold water fish skin gelatin (Sigma), then incubated with primary antibody at these ratios: anti-FLAG (1:2500, Sigma F3165), anti-biotin (1:1000, Jackson 200-002-211), anti-OGT (1:1000, Santa Cruz sc-74546), anti-β-actin (WB 1:2000, Cell Signaling 4970) Histone H3 (1:1000, Cell Signaling 14269), anti-GAPDH (WB, Cell Signaling 2118). Secondary antibodies were Li-Cor IRDye: 680RD donkey-anti mouse 680 (1:10000), 800CW Goat anti-rabbit (1:20000). Three washes in TBST (0.1% tween) were performed after each antibody incubation. All Western blots were imaged on a LiCor Odyssey Clx system. Densitometric measurements were made using Image Studio Lite v5.2

Localization studies (Nucleocytoplasmic Fractionation)

Protein localization was determined using UniProt[101]. Nuclear and cytoplasmic fractions from HeLa cells were obtained using subcellular fractionation as previously described[102] and analyzed via western blot as above.

Pathway Analysis

Gene ontology analysis was performed using The Gene Ontology Resource (geneontology.org)[103,104]. All GO lists were filtered at pvalue less than 0.01, FDR of less than 0.01, and a minimum of 5-fold enrichment over expected number of proteins found in that category in a random protein dataset. Biological process and molecular function analysis were performed using the GO Ontology Database Released 2019-12-09. ReViGo[105] was used to generate condensed lists of GO terms and CirGo[106] to generate plots from the condensed data. PANTHER pathway analysis was performed using PANTHER version 15 released 2020-02-14. Reactome data was also obtained from The Gene Ontology Resource, using Reactome version 65 released 2019-12-22, and parsed at FDR less than 5E-9. Condensed GO term lists and the full reactome pathway list are available in Supplementary Table 5.

Disease association for proteins was identified using the OMIM catalog[107]. Disorders were categorized manually, where “Intellectual Disability” refers to any disorder with the symptom intellectual disability (or several other related terms), “Immunodeficiency” refers to disorders causing immunodeficiency, “Malignancy” refers to any of several cancers, “Congenital, other” refers to congenital disorders not featuring intellectual disability, “Neurological, other” refers to non-congenital neurological disorders, and “Hormone” refers to disorders of the endocrine system.

Results

Defining the OGT TPR Interactome in HeLa Cells

To identify OGT TPR interactors, we utilized a fusion protein strategy using promiscuous biotin ligase BirA*. We generated a fusion protein TPR-BirA*, essentially replacing the catalytic

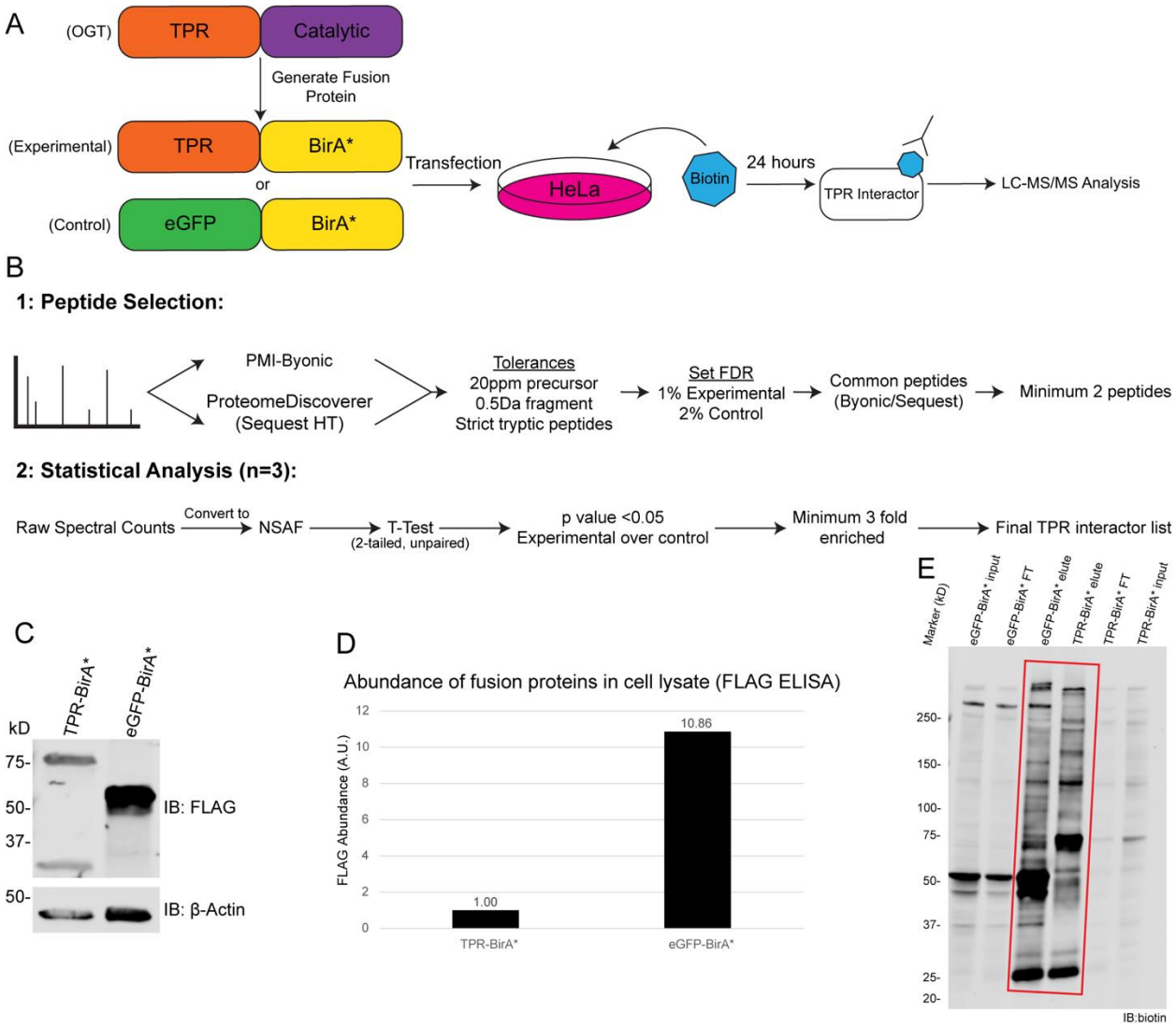


Figure 3.1: BioID approach to define OGT TPR interactome in HeLa cells. ***A***, Workflow of BioID process for identifying TPR interactors. ***B***, Workflow of MS analysis and protein validation. Samples were prepped in biological triplicate. ***C,*** Western blot showing expression of TPR-BirA* and eGFP-BirA* fusion proteins (anti-FLAG tag), β-actin as loading control. 20ug/lane. ***D***, ELISA quantification of FLAG-tagged fusion proteins (one representative replicate, A.U.=Arbitrary units) ***E,*** Representative western blot with anti-biotin antibody demonstrating enrichment of biotinylated proteins following biotin immunoprecipitation (representative blot from three replicates). Red box indicates lanes showing enrichment fo biotinylated proteins by biotin IP. 10ug/lane. FT=flow-through. For eluate and FT, 10ug/lane. For elute, 25% total eluate loaded.

domain of OGT with BirA*, and also created an eGFP-BirA* fusion protein to serve as a negative control for nonspecific protein interactions or promiscuous labeling (Supplementary Table 1). Each fusion protein was transiently overexpressed in HeLa cells and induced with biotin for 24 hours for labeling of proximal proteins (**Fig. 3.1A**). A comparison of the expression level of the TPR-BirA* fusion protein and endogenous OGT can be found in Supplementary Figure 1. Note that for transfections, 10x more TPR-BirA* plasmid was used than eGFP-BirA*, due to eGFP-BirA* expressing at a much higher level than TPR-BirA* (**Fig. 3.1C/D**). Following labeling, we isolated biotinylated proteins with a biotin immunopurification (**Fig. 3.1E**). Even though there is notable biotin labeling in both TPR-BirA* and eGFP-BirA*, we observe a difference in the band patterns, indicating a change in the specificity of biotinylation between eGFP-BirA* and TPR-BirA*. A sectioned SDS-PAGE gel was subjected to in-gel digestion and the resulting peptides separated by nanoflow reverse-phase liquid chromatography in-line to a tribrid mass spectrometer for protein identification (see methods). This entire procedure (transfection to LC-MS/MS analyses) was carried out in 3 independent biological replicates for both TPR-BirA* and eGFP-BirA*.

For analysis of the mass spectrometry data, we opted for a multi-algorithm search to increase the confidence of our protein IDs (**Fig. 3.1B**). Raw mass spectrometry data was searched using both PMI-Byonic and Sequest HT (through ProteomeDiscoverer 2.2) against the human database (Swissprot 09/2019), and only peptides identified by both algorithms were used to generate the final protein set. The negative control (eGFP-BirA*) protein set was searched at a looser peptide FDR (2%) than the TPR protein set (1%) to reduce the likelihood of proteins being kept in the final TPR interactors list due to barely being over the FDR threshold in the control condition. To generate proteins for a final TPR interactors list, we combined uniquely identified proteins in the TPR-BirA* analyses with those that were enriched in the TPR-BirA* protein lists as compared to the eGFP-BirA* (negative control) protein lists. This enrichment was required to be significant according to the Student's t-test with the cut-off for significance being a p-value of 0.05,

Table 3.1: Interactors identified as TPR-BirA* interactors in HeLa cells

Uniprot Accession	Sequence Name
ZHX3_HUMAN	Zinc fingers and homeoboxes protein 3
CCAR1_HUMAN	Cell division cycle and apoptosis regulator protein 1
JHD2C_HUMAN	*Probable JmjC domain-containing histone demethylation protein 2C*
ZN609_HUMAN	Zinc finger protein 609
QSER1_HUMAN	*Glutamine and serine-rich protein 1*
RBM26_HUMAN	RNA-binding protein 26
ZHX2_HUMAN	Zinc fingers and homeoboxes protein 2
PRR12_HUMAN	Proline-rich protein 12
TRPS1_HUMAN	Zinc finger transcription factor Trps1
BCOR_HUMAN	*BCL-6 corepressor*
OGA_HUMAN	***Protein O-GlcNAcase***[108]
CIC_HUMAN	Protein capicua homolog
KMT2C_HUMAN	Histone-lysine N-methyltransferase 2C
CEP85_HUMAN	Centrosomal protein of 85 kDa
NCOA6_HUMAN	*Nuclear receptor coactivator 6*
HIRA_HUMAN	***Protein HIRA***[109]
ARI5B_HUMAN	*AT-rich interactive domain-containing protein 5B*
ZN318_HUMAN	Zinc finger protein 318
RGPD8_HUMAN	*RANBP2-like and GRIP domain-containing protein 8*
RGPD5_HUMAN	RANBP2-like and GRIP domain-containing protein 5/6
ZFR_HUMAN	*Zinc finger RNA-binding protein*
RGPD3_HUMAN	*RanBP2-like and GRIP domain-containing protein 3*
ZC3HE_HUMAN	Zinc finger CCCH domain-containing protein 14
BAP1_HUMAN	***Ubiquitin carboxyl-terminal hydrolase BAP1***[110]
TET2_HUMAN	***Methylcytosine dioxygenase TET2***[20]
ZEP1_HUMAN	Zinc finger protein 40
RPRD2_HUMAN	*Regulation of nuclear pre-mRNA domain-containing protein 2*
PHC3_HUMAN	*Polyhomeotic-like protein 3*
CABIN_HUMAN	Calcineurin-binding protein cabin-1
SET1A_HUMAN	Histone-lysine N-methyltransferase SETD1A
TBL1X_HUMAN	F-box-like/WD repeat-containing protein TBL1X
MEF2D_HUMAN	*Myocyte-specific enhancer factor 2D*
NUP62_HUMAN	*Nuclear pore glycoprotein p62*
PCF11_HUMAN	Pre-mRNA cleavage complex 2 protein Pcf11
TAF4_HUMAN	*Transcription initiation factor TFIID subunit 4*
LIN54_HUMAN	Protein lin-54 homolog
UBN2_HUMAN	Ubinuclein-2
P121C_HUMAN	Nuclear envelope pore membrane protein POM 121C
RESF1_HUMAN	Retroelement silencing factor 1
TASO2_HUMAN	Protein TASOR 2
DOT1L_HUMAN	Histone-lysine N-methyltransferase, H3 lysine-79 specific
NUP58_HUMAN	Nucleoporin p58/p45
ASXL2_HUMAN	*Putative Polycomb group protein ASXL2*
RBM33_HUMAN	RNA-binding protein 33
SMRC1_HUMAN	SWI/SNF complex subunit SMARCC1
ICE1_HUMAN	Little elongation complex subunit 1
PRC2B_HUMAN	*Protein PRRC2B*
NUP54_HUMAN	Nucleoporin p54
SIN3A_HUMAN	***Paired amphipathic helix protein Sin3a***[5]
ARI1B_HUMAN	AT-rich interactive domain-containing protein 1B
CBP_HUMAN	*CREB-binding protein*
SMRC2_HUMAN	SWI/SNF complex subunit SMARCC2
LRIF1_HUMAN	Ligand-dependent nuclear receptor-interacting factor 1
KDM2B_HUMAN	Lysine-specific demethylase 2B
NCOA2_HUMAN	*Nuclear receptor coactivator 2*
UBN1_HUMAN	***Ubinuclein-1***[109]
KMT2D_HUMAN	Histone-lysine N-methyltransferase 2D
ZHX1_HUMAN	*Zinc fingers and homeoboxes protein 1*
PF21A_HUMAN	PHD finger protein 21A
ARI1A_HUMAN	AT-rich interactive domain-containing protein 1A
ZFHX3_HUMAN	Zinc finger homeobox protein 3
EP300_HUMAN	Histone acetyltransferase p300
SF3A1_HUMAN	Splicing factor 3A subunit 1
BCL9_HUMAN	B-cell CLL/lymphoma 9 protein
TRRAP_HUMAN	Transformation/transcription domain-associated protein
ACL6A_HUMAN	Actin-like protein 6A
TAF6_HUMAN	*Transcription initiation factor TFIID subunit 6*
TNR6A_HUMAN	*Trinucleotide repeat-containing gene 6A protein*
POGZ_HUMAN	*Pogo transposable element with ZNF domain*
AHDC1_HUMAN	AT-hook DNA-binding motif-containing protein 1
I2BP2_HUMAN	Interferon regulatory factor 2-binding protein 2
RREB1_HUMAN	*Ras-responsive element-binding protein 1*
NCOR2_HUMAN	*Nuclear receptor corepressor 2*
PHF12_HUMAN	PHD finger protein 12
OGT1_HUMAN	***UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit***[18]
ANR17_HUMAN	*Ankyrin repeat domain-containing protein 17*
RBM27_HUMAN	*RNA-binding protein 27*
NU153_HUMAN	*Nuclear pore complex protein Nup153*
KMT2A_HUMAN	Histone-lysine N-methyltransferase 2A
NUP50_HUMAN	Nuclear pore complex protein Nup50
TNR6B_HUMAN	Trinucleotide repeat-containing gene 6B protein
HCFC1_HUMAN	***Host cell factor 1***[19]
ZC3H1_HUMAN	Zinc finger C3H1 domain-containing protein
GANP_HUMAN	Germinal-center associated nuclear protein
UBAP2_HUMAN	*Ubiquitin-associated protein 2*
NCOR1_HUMAN	*Nuclear receptor corepressor 1*
EMSY_HUMAN	*BRCA2-interacting transcriptional repressor EMSY*
KDM1A_HUMAN	Lysine-specific histone demethylase 1A
PR40A_HUMAN	Pre-mRNA-processing factor 40 homolog A
MINT_HUMAN	*Msx2-interacting protein*
TBL1R_HUMAN	F-box-like/WD repeat-containing protein TBL1XR1
KDM3B_HUMAN	Lysine-specific demethylase 3B
EP400_HUMAN	*E1A-binding protein p400*
ILF2_HUMAN	Interleukin enhancer-binding factor 2
ZN281_HUMAN	Zinc finger protein 281
NU214_HUMAN	*Nuclear pore complex protein Nup214*
DIDO1_HUMAN	*Death-inducer obliterator 1*
SMCA4_HUMAN	Transcription activator BRG1
PRC2A_HUMAN	Protein PRRC2A
YLPM1_HUMAN	*YLP motif-containing protein 1*
NUP88_HUMAN	Nuclear pore complex protein Nup88
ANKH1_HUMAN	Ankyrin repeat and KH domain-containing protein 1
WNK1_HUMAN	*Serine/threonine-protein kinase WNK1*
TAF5_HUMAN	Transcription initiation factor TFIID subunit 5
PRC2C_HUMAN	*Protein PRRC2C*
RBM25_HUMAN	RNA-binding protein 25
RBP2_HUMAN	*E3 SUMO-protein ligase RanBP2*
BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF
CPSF1_HUMAN	Cleavage and polyadenylation specificity factor subunit 1
HNRPL_HUMAN	Heterogeneous nuclear ribonucleoprotein L
RAGP1_HUMAN	Ran GTPase-activating protein 1
UBP2L_HUMAN	*Ubiquitin-associated protein 2-like*
ZN106_HUMAN	Zinc finger protein 106
KNL1_HUMAN	Kinetochore scaffold 1

Legend:
Black Accession: ID'd only in TPR
Bold Accession: ID'd in both TPR and eGFP, p value <0.05
Italicized Protein Name: Known O-GlcNAc modified
Bold and Italicized Protein Name: Known OGT interactors (with citations)

and we also required the average NSAF[100] to be at least 3 times higher in the TPR-BirA* condition compared to the eGFP-BirA* condition. These proteins represent a stringent list of 115 OGT TPR interactors (**Table 3.1**, Supplementary Table 4).

46 of the proteins (indicated by an *italicized* protein name in **Table 1**) had already been identified as O-GlcNAc modified by previous -omic datasets[68,111]. This result suggests that the TPR domain alone can select OGT substrate proteins without the presence of the catalytic domain. In addition, 8 (indicated by an ***italicized and bold*** protein name in **Table 1**) of the identified proteins had previously been demonstrated to specifically interact with OGT, where "interaction" here is defined as a one or two directional co-immunoprecipitation[5,18–20,108–110]. Together, these factors lend confidence to the novel protein IDs in the dataset.

It is noteworthy that OGT itself is identified in the screen. Although many peptides attributed to OGT are due to the overexpression of the TPR-BirA* fusion protein, several high-confidence peptides in the catalytic domain of OGT were also identified. This suggests that endogenous OGT, which normally exists as a dimer[46], did complex with the fusion TPR protein. Several well-studied OGT interactors were also identified, including HCF1[19], mSin3a[5], and Tet2[20]. We also identify the O-GlcNAc hydrolase OGA, which OGT is known to regulate both pre- and post-translationally[26,87].

Many of the interactors identified here are members of protein complexes which may imply that OGT does not directly interact with all of the proteins in the complex. One such complex is the HIRA protein complex, previously shown to interact with OGT[109]. We identified all three members of this complex (HIRA, UBN1, and CABIN). We also identified a novel TPR interaction with the SWI/SNF complex. Six members of the SWI/SNF complex were identified (SMRC2, SMRC1, ACL6A, SMCA4, ARI1A, ARI1B). SWI/SNF proteins, like OGT[5,20,62,69,112–115], function in chromatin remodeling[116], but OGT has, to our knowledge, never been shown to interact with these proteins.

In addition to many protein interactors involved in known OGT functions, we also identified proteins with roles in RNA processing, an area of cellular biology for which limited research exists on the role of OGT and the O-GlcNAc modification. These interactors include proteins with known and putative roles in pre-mRNA splicing (SF3A1, PCF11, PRC2A, PR40A), polyadenylation (CPSF1), and RNA binding (ZN106, TNR6B, RBM33, RBM25, RBM26).

Validation of Proteins Identified in both TPR-BirA and eGFP-BirA**

72 of the 115 OGT TPR interactors were only observed in the TPR interactome. Several protein IDs (43 – indicated by a **bold** Uniprot accession in **Table 1**) were identified in both the TPR-BirA* and eGFP-BirA* samples, but were significantly enriched in TPR-BirA* at the level of average NSAF (Student's t-test p value < 0.05, fold enrichment of average NSAF >3) (**Fig. 3.2**). To further confirm the validity of the inclusion of these proteins in the final dataset, we examined MS1 reconstructed ion chromatograms for peptides identified in both TPR-BirA* and eGFP-BirA*. OGT itself was identified in both, although it is highly enriched in the TPR-BirA*, likely in part due to the overexpression of the TPR-BirA* fusion protein. To confirm that endogenous OGT labeling is enriched in the TPR-BirA* sample, we compared the intensity of a catalytic domain peptide between the TPR-BirA* samples and the eGFP-BirA* samples. The average intensity of this peptide in the TPR-BirA* samples is 9.1 (standard deviation of 2.3) times higher than in the eGFP-BirA* samples, supporting specific interaction of TPR-BirA* with full-length endogenous OGT (average mass error 2.17ppm) (**Fig. 3.3A**).

To further confirm the enrichment of relevant OGT interactors in TPR-BirA*, we next examined the intensity of peptides from HCF-1, a well-studied OGT interactor with roles in intellectual disability[19,117]. Average intensity for an HCF-1 peptide are 36.4-fold (standard deviation of 24.4) higher in TPR-BirA* than eGFP-BirA* (average mass error 0.749ppm) (**Fig. 3.3B**). Finally, we confirmed a fold increase in peptide intensity for the protein closest to our fold enrichment cutoff, kinetochore scaffold 1 (KNL1) (**Fig. 3.3C**). The peptide shown from KNL1 has an average

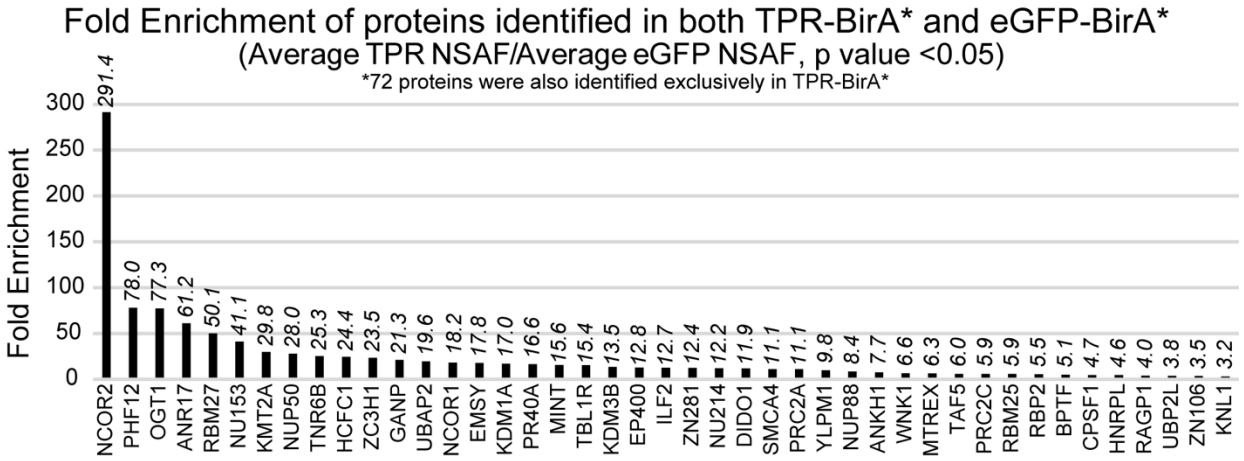


Figure 3.2: Proteins identified in both TPR-BirA* and eGFP-BirA* are enriched in TPR-BirA*. Fold enrichment values for all proteins identified in both TPR-BirA* and eGFP-BirA*. Fold enrichment values are average NSAF of TPR-BirA* over average NSAF of eGFP-BirA*. Note that 72 proteins were only observed in TPR-BirA*

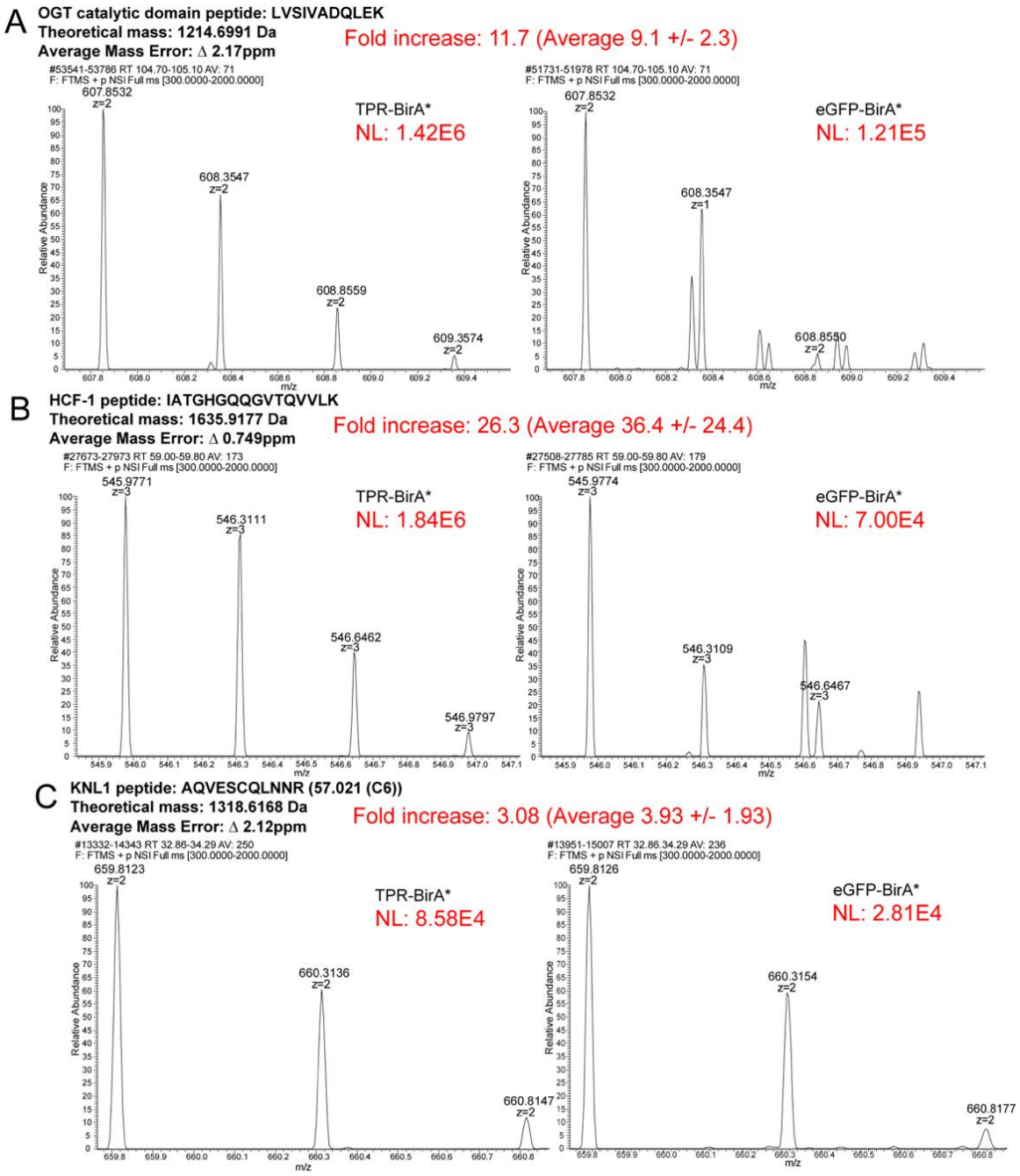


Figure 3.3: Reconstructed ion chromatograms confirm peptide-level enrichment of TPR-BirA* interactors. Fold increase is for the replicate shown. The average fold intensity for that peptide (averaged across all replicates and fractions in which that peptide appeared) is shown with the standard deviation. Average mass error is the absolute value of the mean across 3 replicates of both TPR-BirA* and eGFP-BirA*. NL=Normalization Level (Base Peak Intensity) ***A,*** Reconstructed ion chromatograms demonstrating greater enrichment of a catalytic OGT peptide in TPR-BirA* compared to eGFP-BirA*. ***B,*** Reconstructed ion chromatograms demonstrating greater enrichment of an HCF-1 peptide in TPR-BirA* compared to eGFP-BirA*. ***C,*** Reconstructed ion chromatograms demonstrating greater enrichment of a KNL1 peptide (novel interactor closest to cut-off for assignment) in TPR-BirA* compared to eGFP-BirA*

intensity in TPR-BirA* that is 3.93 (standard deviation of 1.93) higher than in eGFP-BirA* (average mass error 2.12ppm). Taken together, these results indicate that although some proteins were identified in both the experimental and negative control conditions, their inclusion in the final protein interactor list due to higher enrichment is supported by the raw mass spectrometry data.

Subcellular Localization of TPR-BirA Interactors*

OGT localizes primarily to the nucleus, but also to the cytoplasm, in the mammalian cell[2]. To confirm that the TPR-BirA* fusion protein also localized to both compartments, we examined the subcellular localization of OGT and our fusion proteins. By subcellular fractionation (**Fig. 3.4A/B)**, endogenous OGT localizes primarily to the nucleus with some expression in the cytoplasm, as expected[2]. In contrast, the TPR-BirA* fusion protein localizes more highly to the cytoplasm, although it is also present in the nucleus. This result is unsurprising as previous research has shown that overexpressed OGT localizes more highly to the cytoplasm than native OGT[2]. eGFP-BirA* also localizes to both the nucleus and cytoplasm. The subcellular localization profiles of TPR-BirA* and eGFP-BirA* are very similar, supporting the use of eGFP-BirA* as a sufficient negative control for nonspecific labeling by BirA* in both the nuclear and cytosolic compartments.

We expected most TPR-BirA* interactors to be primarily nuclear, as most recorded OGT interactors are as well[39]. Analysis of the subcellular localization of identified TPR-BirA* interactors supports this (Visualized in an UpsetR plot **Fig. 3.4C**, and in a pie chart in **Fig. 3.4D**). 67 of the 115 identified interactors are exclusively nuclear, with an additional 23 occurring in both the nucleus and the cytoplasm. Several interactors localize specifically to the nuclear pore. Only four proteins exclusively localize to the cytoplasm, all of which are novel OGT-TPR interactors (RGPD5, WNK1, TNR6B, and ANKH1). [39]

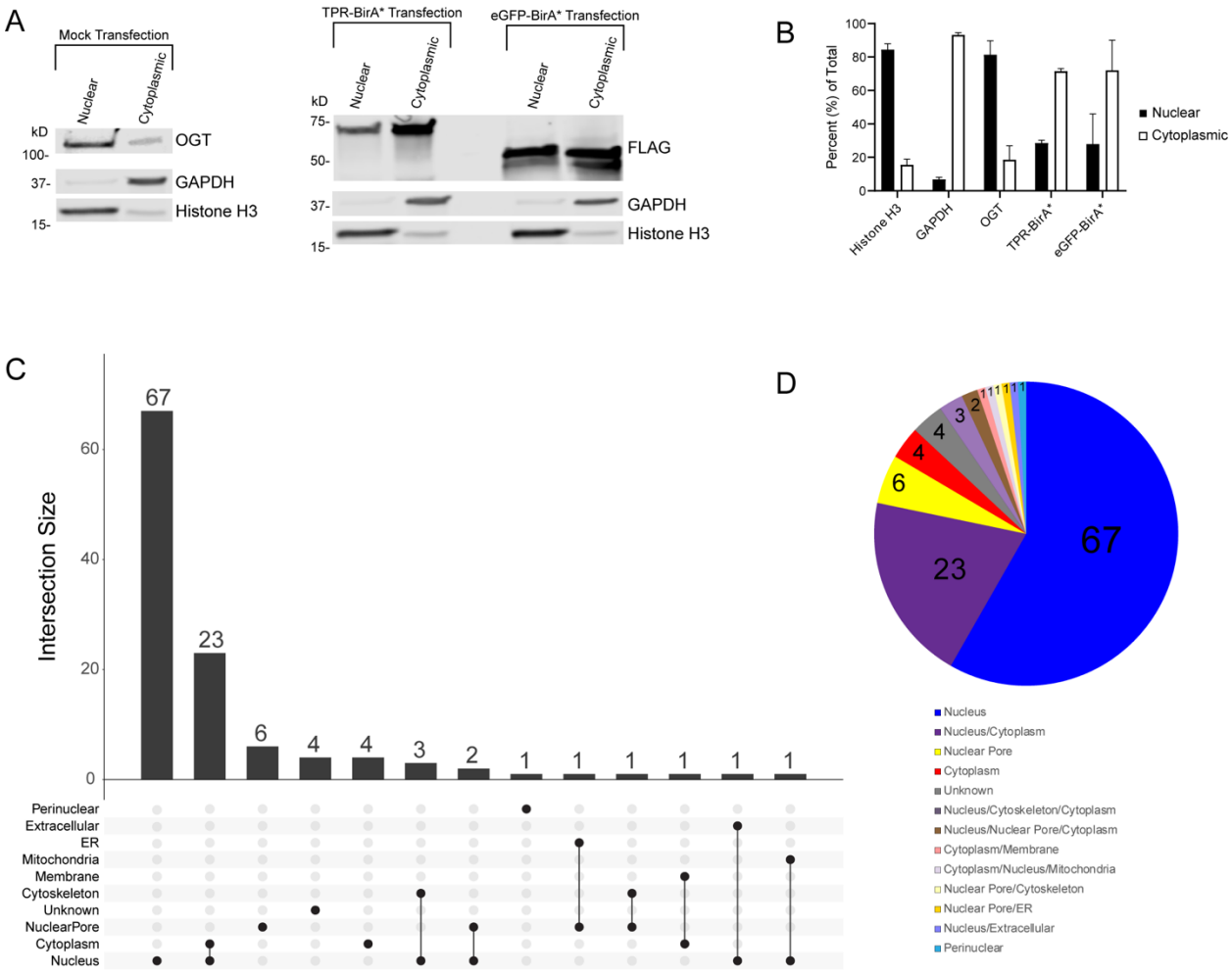


Figure 3.4: TPR interactors are primarily nuclear localized. ***A,*** Subcellular fractionation of HeLa cells demonstrating localization of OGT (anti-OGT F12) and BirA* fusion proteins (anti-FLAG tag). Cytoplasmic marker is GAPDH, nuclear marker is Histone H3. 10ug/lane, representative western blot of three biological replicates. ***B,*** Ratios of nuclear to cytoplasmic expression of marker proteins (Nuclear: Histone H3, Cytoplasmic: GAPDH) and fusion proteins. Averaged across three biological replicates. ***C,*** UpsetR plot showing the subcellular localization of TPR interactors ***D,*** Venn diagram showing the subcellular localization of TPR interactors. Numbers represent the total number of TPR interactors in that category. Localization determined using UniProt.

Ontology Analyses of the OGT TPR-Interactome

To further understand the enrichment of various processes in our OGT TPR interactome, we performed several different Gene Ontology (GO) analyses using the Gene Ontology Resource (GeneOntology.org) (**Fig 3.5**). Performance of a PANTHER Overrepresentation Test for biological processes in our interactome further confirms OGT-TPR interactors having frequent roles in transcriptional and chromatin regulation (**Fig 3.5A**). General chromatin organization is a strongly enriched category, along with the related peptidyl-lysine modification (indicative of histone modification). Of note is the enrichment of OGT-TPR interactors specifically involved in gene silencing. OGT is a Polycomb Group Protein (*sxc* in *Drosophila melanogaster),* which is responsible for the silencing of Hox genes during developmental patterning[112,113]. The TPR interactors identified here further corroborate OGT's previously-identified tendency toward roles in gene silencing and may reveal further avenues by which OGT regulates gene silencing[5]. Of additional note is the enrichment of interactors involved in the regulation of cellular response to heat, since previous work has demonstrated a role for OGT in coping with cellular heat shock[7]; however, limited work has been published exploring the specific OGT interactions that help it to perform this function. Finally, the enrichment of proteins involved in rhythmic process and circadian rhythm aligns with previous research demonstrating that OGT is involved in circadian rhythm regulation[118]. To confirm these enriched processes, we also performed a PANTHER Overrepresentation Test for molecular function pathways in our OGT TPR interactome (**Fig 3.5B**). Many molecular functions corroborate our findings of enriched biological processes, including consistent high enrichment in chromatin and transcriptional regulation. It is interesting to note that RNA Pol II transcription factor binding in particular is an enriched molecular function, since OGT is known to interact with and regulate RNA Pol II-mediated transcription[6,59,119]. Also enriched are processes specifically relating to histone modification, further confirming the enrichment of peptidyl-lysine modification of histones as identified in biological process enrichment and

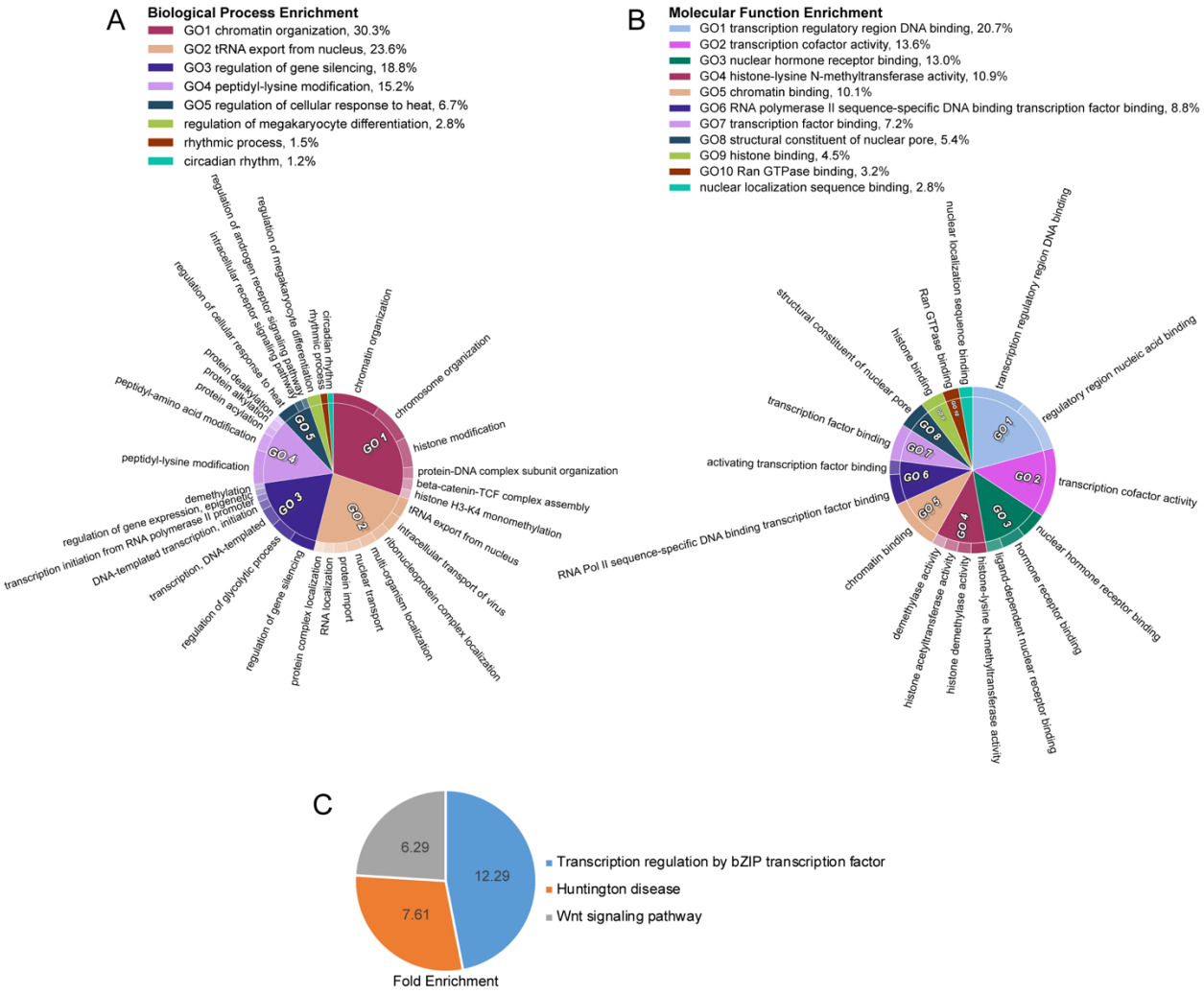


Figure 3.5: TPR interactors demonstrate enrichment in biological processes and disease states. ***A,*** CirGo plot showing enriched biological processes in the TPR protein interactor list ***B,*** CirGo plot showing enriched molecular functions in the TPR protein interactor list ***C,*** Venn diagram of PANTHER Pathway enrichment of TPR interactors. Numbers are the fold enrichment of the pathway process over expected enrichment.

consistent with the O-GlcNAc modification being part of the histone code[115]. Finally, enrichment of nuclear pore components and nuclear localization sequence binding confirms the long-standing role for OGT in nuclear pore structure and/or regulation[120].

We further examined enriched Reactome pathways (**Table 3.2**) among TPR interactors, specifically the mostly highly enriched pathways at an FDR of less than 5E-9. This pathway analysis validates our previous GO analyses that reveal roles for OGT in chromatin regulation, transcriptional regulation, and nuclear pore processes. Reactome pathway enrichment also uniquely reveals several roles for OGT TPR interactors in viral infection, nuclear import, and processing. OGT has been demonstrated to play a role in a limited number of specific viral infections[87,121] but these enriched pathways point to a potentially broader role for OGT and its interactors more generally in viral infection. Interactors are also enriched in the reactome pathway "regulation of glucokinase by glucokinase regulatory protein". OGT has already been shown to regulate glucokinase[122] as well as other proteins involved in glucose metabolism including phosphofructokinase 1[123]. Panther pathway enrichment analysis (**Fig 3.5C**) reveals OGT-TPR interactor involvement in basic leucine-zipped transcription factor mediated transcriptional regulation, the Wnt signaling pathway, and Huntington disease related processes. The interplay with basic leucine-zipped transcription factors points to another potential avenue for OGT's regulation of transcription. Furthermore, OGT has already been shown to interface with the Wnt pathway by modulating β-Catenin stability[124]. The interactors identified here involved in this pathway may point to other mechanisms by which OGT modulates Wnt signaling.

Pathophysiology Analyses of the OGT TPR-Interactome and Orthogonal validation of XLID-related Interactors

The identification of Huntington's disease (**Fig 3.5C**) as an enriched disease process among the TPR interactors prompted us to examine whether identified TPR interactors are involved in other disease processes. Unsurprisingly, as determined using the OMIM catalogue, many TPR

Table 3.2: Reactome pathways enriched in TPR interactors list (FDR <5E-9)

Reactome Pathway	Identifier	Protein Number	Fold enrichment	P value	FDR
Chromatin organization	R-HSA-4839726	27	20.4	9.70E-27	1.11E-23
Chromatin modifying enzymes	R-HSA-3247509	27	20.4	9.70E-27	2.22E-23
SUMO E3 ligases SUMOylate target proteins	R-HSA-3108232	15	17.32	3.26E-14	1.86E-11
SUMOylation	R-HSA-2990846	15	16.69	5.42E-14	2.47E-11
Transcriptional regulation by RUNX1	R-HSA-8878171	16	14.22	7.48E-14	2.85E-11
Rev-mediated nuclear export of HIV RNA	R-HSA-165054	9	50.99	9.89E-13	3.23E-10
HCMV Early Events	R-HSA-9609690	12	21.76	1.15E-12	3.29E-10
Interactions of Rev with host cellular proteins	R-HSA-177243	9	47.99	1.58E-12	4.00E-10
Transport of Mature mRNA Derived from an Intronless Transcript	R-HSA-159231	9	41.84	4.58E-12	1.05E-09
Transport of Mature mRNAs Derived from Intronless Transcripts	R-HSA-159234	9	40.8	5.59E-12	1.16E-09
SUMOylation of DNA replication proteins	R-HSA-4615885	9	37.95	9.87E-12	1.61E-09
Processing of Capped Intron-Containing Pre-mRNA	R-HSA-72203	15	11.43	9.06E-12	1.72E-09
SUMOylation of RNA binding proteins	R-HSA-4570464	9	37.95	9.87E-12	1.73E-09
HCMV Infection	R-HSA-9609646	12	17.55	1.18E-11	1.80E-09
Transport of Ribonucleoproteins into the Host Nucleus	R-HSA-168271	8	50.02	2.16E-11	2.60E-09
NEP/NS2 Interacts with the Cellular Export Machinery	R-HSA-168333	8	50.02	2.16E-11	2.75E-09
Defective TPR may confer susceptibility towards thyroid papillary carcinoma (TPC)	R-HSA-5619107	8	50.02	2.16E-11	2.91E-09
Regulation of Glucokinase by Glucokinase Regulatory Protein	R-HSA-170822	8	50.02	2.16E-11	3.09E-09
Export of Viral Ribonucleoproteins from Nucleus	R-HSA-168274	8	48.35	2.73E-11	3.12E-09
Vpr-mediated nuclear import of PICs	R-HSA-180910	8	46.79	3.42E-11	3.39E-09
Late Phase of HIV Life Cycle	R-HSA-162599	12	16	3.23E-11	3.51E-09
Nuclear import of Rev protein	R-HSA-180746	8	46.79	3.42E-11	3.55E-09
SUMOylation of SUMOylation proteins	R-HSA-4085377	8	45.33	4.25E-11	3.89E-09
Transport of the SLBP independent Mature mRNA	R-HSA-159227	8	45.33	4.25E-11	4.05E-09
Nuclear Pore Complex (NPC) Disassembly	R-HSA-3301854	8	43.95	5.26E-11	4.29E-09
Transport of the SLBP Dependant Mature mRNA	R-HSA-159230	8	43.95	5.26E-11	4.45E-09
SUMOylation of chromatin organization proteins	R-HSA-4551638	9	30.79	5.20E-11	4.57E-09

interactors are involved in disease processes with which OGT is already associated, including malignancy[23] and neurological[125] disorders (**Fig 3.6A/B**).

Strikingly, of the 44 TPR interactors involved in an OMIM-classified phenotype, 24 are linked to disorders which feature intellectual disability. Three interactors are linked to two different intellectual disability-related disorders each, bringing the total count of intellectual disability disorders associated to OGT-TPR interactors to 27 (**Fig 3.6A/B**) [26–29,26,27,38]. Therefore, these 24 protein interactors are of significant interest in the search for a mechanism underlying the OGT XLID phenotype.

Discussion

One of the prevailing mysteries in the O-GlcNAc field is how the O-GlcNAc Transferase (OGT) enzyme is able to select from among thousands of possible substrates given that there is only one gene encoding the protein in the mammalian cell. A prevailing hypothesis in the field is that post-translational modification of OGT and protein-protein interactors are responsible for OGT substrate selection. This mechanism may be similar to RNA PolII that transcribes all protein-coding genes but is exquisitely regulated by protein-protein associations including the basal transcriptional machinery and transcription factors[126]. The role of the TPR domain of OGT in promoting highly specific substrate selection has been demonstrated in part in structural biology studies[16,46,51,56], but has only been explored for a limited number of specific substrates. Here, we demonstrated that the TPR domain of OGT is capable of interacting with substrate proteins even without the presence of the catalytic domain, and using the BioID technique we have identified 115 TPR interactors, representing both known OGT interactors and novel OGT-TPR interactors.

OGT is known to localize to the nucleus and the cytoplasm, but usually primarily resides in the nucleus[2]. The fact that most interactors found here are exclusively localized to the nucleus despite the TPR-BirA* fusion protein being localized more dominantly to the cytoplasm is an intriguing observation, suggesting that OGT more strongly and/or frequently interacts with nuclear

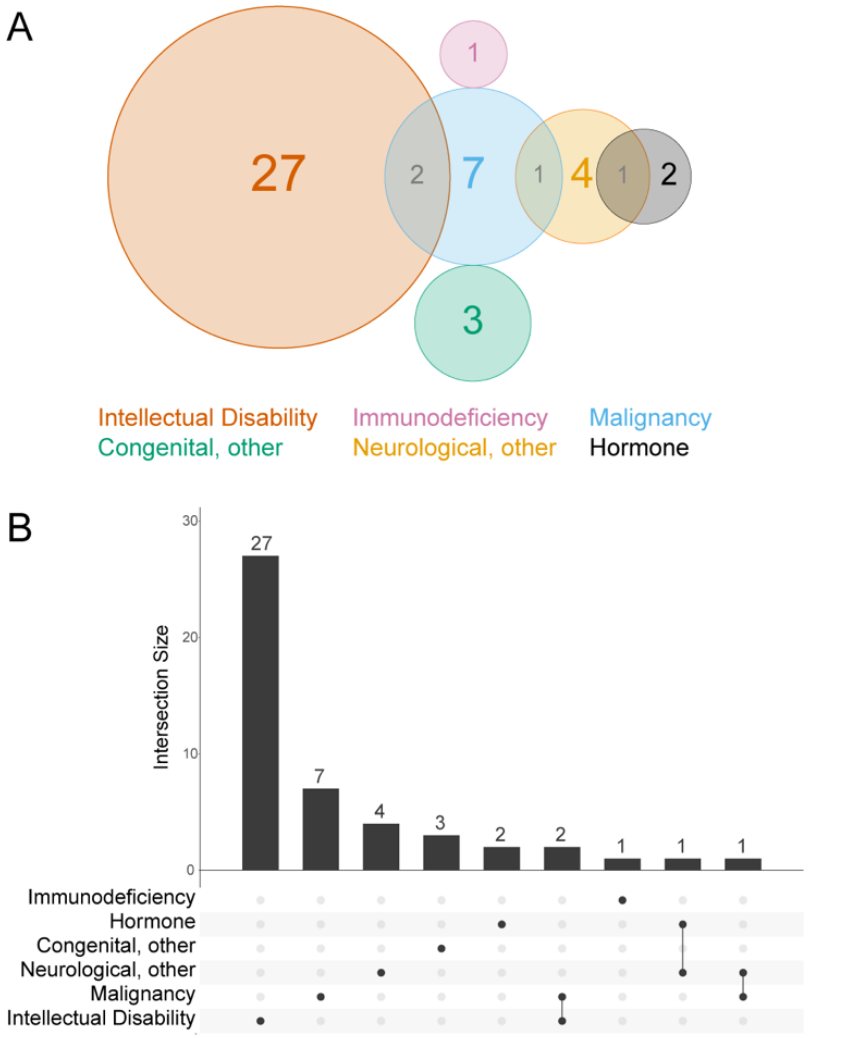


Figure 3.6: OGT interacts with proteins involved in intellectual disability in HeLa cells. *A,* Venn diagram showing disease states enriched in the TPR interactors. ***B,*** UpsetR plot showing disease states enriched in the TPR interactors. Disease associations determined using the OMIM resource.

proteins regardless of localization. It is important to note that the nature of the TPR-BirA* as a truncation of OGT and an overexpressed fusion protein opens the possibility for false-negative or false-positive protein IDs. While we have attempted to thoroughly mitigate the occurrence of false positives through stringent statistical analyses and the use of a negative control (eGFP-BirA*), future work to biochemically confirm each interaction with the full-length OGT is necessary. This will also allow for the determination of the degree to which given interactions are transient or stable, under what conditions they occur, and what functional roles they play in the cell.

The identified interactors likely fall into several different classes. Many of the identified proteins are O-GlcNAc modified, and these may be the effector substrates by which OGT modulates cellular status. Another group, not mutually exclusive to the first, may represent partner proteins; that is, proteins that interact with the TPR domain of OGT to target it to specific substrates or intracellular regions, thus affecting the substrates OGT can access. These proteins may or may not themselves be functionally O-GlcNAc modified. As an example, Tet2 binds to OGT and directs it to histones, but the O-GlcNAc modification on Tet2 has no observed effect on its function[20]. Finally, it is likely that some of the identified interactors do not directly interact with OGT but rather are members of a complex, a subset of which interact directly with OGT. The Swi/Snf complex is an example of this – we have identified 6 of at least 20 possible subunits in our TPR-BirA* interaction list. It is likely that the TPR domain does not directly interact with all 6 identified components and instead interacts with a subset, but additional complex members are labeled due to the diffusion of the reactive biotin intermediate released by the BirA* protein. This would also explain why we fail to identify the full complex; additional members of the complex may be too distant in space to be biotin labeled. Further work is required to identify direct versus indirect interactors.

In general, the TPR interactors identified here corroborate OGT's role as a high-level regulator of cellular function as identified in other studies[5,112,113]. OGT has previously been

characterized as a “rheostat” rather than a switch[39], and the data here supports this notion. Most of the TPR interactors we identified are “modulators” themselves, e.g. they are not enzymes with a direct effect on a given substrate, but rather affect cellular physiology at a global level by modulating transcription, protein stability, or transport. This gives a perspective of OGT as a modulator of the modulators; that is, OGT regulates cellular function by making many subtle changes in global regulators, adding up to a more significant functional outcome. One such global regulation avenue is chromatin remodeling, which is a previously known function of OGT and a function in which many of our TPR interactors are involved. While our data does not determine how OGT’s interaction with these chromatin remodelers affects their function, many of our identified interactors are involved in lysine modification of histones, pointing to a potential avenue for OGT’s regulation of chromatin remodeling. Indeed, OGT has already been noted to interact with histone modifying enzymes including HDACs[5]. We are unsure why we did not identify any HDACs in our screen – it is possible that they interact with OGT as a part of a protein complex but at a distance outside of the BirA* biotin labeling radius.

We have also identified interactors involved in biological processes in which OGT has yet to be implicated, most noteworthy in RNA processing and transport. Interestingly, OGA, which removes the O-GlcNAc modification, has previously been shown to localize to the nucleolus[79], indicating the presence and possible role of O-GlcNAc modified proteins in this subcellular structure involved in RNA processing. Future work will be necessary to determine the specific role OGT plays in these processes.

Finally, we have identified many OGT-TPR interactors that are involved in disease. The TPR interactors we present here may represent avenues for future research uncovering mechanistic proteins underlying OGT’s role in various disease states. Of current high interest is the mechanism underlying *OGT* mutations leading to X-Linked Intellectual Disability (XLID)[26,29,127]. One prominent hypothesis that we have previously suggested[26,27,38] is that mutations in the TPR

domain disrupt OGT protein interactions, leading to downstream developmental effects that lead to the XLID phenotype. Here, we have identified 24 OGT TPR interactors directly involved in disorders with intellectual disability. While it is possible that a novel interactor or set of interactors underlies the OGT-XLID mechanism, these interactors represent a strong set of candidate interactors that may contribute to the phenotype. The high number of interactors with connections to intellectual disability may also indicate that there may be a more global interruption in protein-protein interactions caused by XLID variants in OGT. Rather than one specific interactor failing to interact with OGT and leading to XLID, there may be a more subtle reduction in interaction with many proteins, leading to global neurodevelopmental abnormalities.

Our lab is currently undertaking BioID and immunoprecipitation studies to identify any perturbations in the OGT interactome in neural lines harboring XLID-linked OGT variants. The BioID method described here will be a valuable tool to identify potentially tissue-/cell type-specific TPR interactors that fail to interact with XLID-associated OGT variants. In a more directed approach, intellectual disability-related interactors identified here are being screened for protein interaction with XLID-linked OGT variants to determine if they may represent protein interactors underlying the XLID-OGT phenotype. Thus, the work presented here lays a groundwork for additional studies to understand OGT substrate selectivity and the role of OGT and the O-GlcNAc modification in a plethora of biological processes and human pathophysiology including XLID.

Acknowledgments

This article is dedicated to the memory of our beloved colleague Dr. Brent Weatherly, who developed the proteomic analysis workflows used in this manuscript. We thank Dr. Kelley Moremen for technical advice and plasmid constructs. This work was supported by a grant from the W.M. Keck foundation (L.W. Co-PI); an NICHD National Institutes of Health (NIH) grant R21HD097652 (L.W); and an NICHD Grant F30 HD098828 (H.S.). The content is solely the

responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Supporting Information

The following supporting information is available free of charge at ACS website http://pubs.acs.org

Supplementary Figure 1: Comparison of TPR-BirA* and endogenous OGT expression

Supplementary Figure 2: Full western blot membranes

Supplementary Table 1: Fusion protein sequences and common contaminants database

Supplementary Table 2: TPR Interactors list with statistical analyses

Supplementary Table 3: All peptide matches

Supplementary Table 4: All protein IDs

Supplementary Table 5: GO Term lists

Raw mass spectrometry data: Deposited in the MassIVE database (https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp, Dataset ID: MSV000085626)

Author Contributions

H.S. and L.W. conceived and coordinated the study and wrote the manuscript. H.S. performed all experiments and data searches. J.P. performed calculations and statistical analyses on final protein lists and assisted in revision of the manuscript.

CHAPTER 4

PERTURBING THE OGT TPR INTERACTOME

Introduction

The O-GlcNAc transferase (OGT) is a unique glycosyltransferase in that it resides in the nucleus and cytoplasm, modifies thousands of nuclear and cytosolic proteins, and creates a non-extended glycan modification that is both dynamic and inducible, called O-GlcNAc[10]. OGT and the O-GlcNAc modification have been linked to a myriad of different cellular functions, such as cell cycling[108,128], chromatin remodeling[5,115], and nutrient sensing[54,69], as well as many medically significant disease states, such as Alzheimer's disease[24,129,130], various types of cancer[23,131–133], diabetes[8], and X-Linked Intellectual Disability[26,28,29,38,127]. O-GlcNAc is often considered more analogous to protein phosphorylation than canonical secretory pathway glycosylation. However, unlike protein phosphorylation which is mediated by hundreds of protein kinases[15], the O-GlcNAc modification is only mediated by one OGT in the mammalian cell[4]. Therefore, there must be elegant and complex mechanisms in place for OGT to select from among thousands of substrates with such diverse functions.

The previous chapters have discussed possible mechanisms of OGT substrate selection, of which a popular hypothesis is substrate and partner protein interaction. We have also described the basal OGT TPR interactome in one cell type[134]. However, considering OGT's ubiquitous expression[2] and variety of functional roles, this interactome alone cannot describe the breadth of OGT's protein interaction. Therefore, we hypothesize that the OGT interactome is dynamic 1) across cell types and 2) under different cellular states. Following previous research that suggests that the TPR domain of OGT is primarily responsible for its substrate selection[16,18,46], we specifically focused on the interactors of the TPR domain. Here, we set out to describe a basal

OGT TPR interactome in a neuroblastoma cell line, SHSY5Y, and to perturb the OGT TPR interactome in these and HeLa cells.

Studying the dynamics of the OGT interactome has historically been a challenge due to the nature of OGT as an enzyme – most interactions are likely to be brief and difficult to detect with typical protein interaction methods such as immunoprecipitation. The BioID method is an ideal system for studying interactome dynamics due to its labeling of transient interactors[31]. In Chapter 3 and the first interactome described in this chapter, we utilize the original BioID method which requires a biotin labeling time of 24 hours. While this method is excellent for attaining a global picture of interactors over a period of time, it does not allow for effective study of more dynamic alterations to an interactome, because short-lived interaction differences may be suppressed by more abundant interactions. In order to study briefer alterations in cellular state and their effects on the OGT TPR interactome, we utilized the TurboID system[32]. TurboID is an analogue of BioID with additional mutations that cause it to label interacting proteins on a much shorter time scale of 10-30 minutes. This adapted version of BioID enabled us to probe whether the OGT TPR interactome rapidly responds to cellular stresses, specifically here the initiation of serum signaling and membrane depolarization.

Methods

Plasmid Constructs

TurboID plasmid constructs were built by replacing the BirA* in the original plasmids with TurboID. TurboID was PCR cloned from a construct that was gifted to us by Walter Schmidt's lab. For TPR-TurboID: TPR - 3X GGGGS linker - TurboID - 2X FLAG. For eGFP-TurboID: eGFP - 3X GGGGS Linker – TurboID - 2XFLAG.

HeLa glucose starvation BioID interactome

HeLa cell glucose starvation interactomes were performed essentially as described previously. In brief, cells were plated 24 hours before transfection in either high glucose DMEM,

DMEM containing 4mM glucose, or DMEM containing 1mM glucose. Transfection, biotin labeling, cell collection, biotin pulldown, mass spectrometry, and data analysis were performed as previously described[134].

HeLa serum starvation TurboID interactome

HeLa cells were plated on 10cm dishes with 2.7e6 cells per plate, then cultured for 24 hours. Cells were transfected with Lipofectamine 2000 24 hours after plating, using either eGFP-Turbo (6µg) or TPR-Turbo (12µg) at a 1:2 lipofectamine ratio. Following incubation at 37C for 24 hours, media was changed for serum starvation. For serum plates, media was changed to maintenance DMEM with 10% FBS and antibiotic-antimycotic (ThermoFisher Scientific). For serum starvation and serum pulse plates, plates were washed 2X with DMEM containing no FBS before culturing in serum-free media. Cells were cultured for 24 hours.

For serum pulse plates, media was replaced with serum-containing media for 10 minutes prior to labeling. For all plates, biotin was added to media to a final concentration of 500µM and incubated at 37C for 10 minutes. Cells were placed on ice to stop labeling, then washed with PBS. Cells were then scraped into PBS, centrifuged to pellet, and resuspended in RIPA lysis buffer (50mM Tris, 150mM NaCl, 1% IGEPAL-CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 1mM DTT, 1mM EDTA, protease and phosphatase inhibitors, and universal nuclease (ThermoFisher Scientific)) for 15 minutes on ice with vortexing for 30 seconds at 10 and 5 minutes. Three biological replicates were performed, each consisting of TPR-TurboID and eGFP each for serum, serum starve, and serum pulse.

SHSY5Y serum starvation TurboID interactome

SHSY5Y cells were grown in DMEM/F12 50:50 with 10% FBS, 20mM Glutagro (Corning) and antibiotic-antimycotic. SHSY5Y serum starvation was performed essentially as described above, with cell specific changes detailed here. SHSY5Y cells were plated on 10cm dishes with 5.7e6 cells per plate. Transfection occurred 48 hours after plating with 31.3µg eGFP-Turbo or 61.2µg

TPR-Turbo. Cells were serum starved for 12 hours, with no washes prior to serum starvation due to loose cell attachment.

SHSY5Y depolarization TurboID interactome

SHSY5Y cells were plated and transfected as described above. 24 hours after transfection, cells were either labeled with 500µM biotin or labeled and depolarized with 50mM KCl and 500µM biotin. Cell collection was described above.

Identification of biotinylated proteins from TurboID interactomes

Biotin pulldowns for TurboID interactomes were performed using the same method for all interactomes. 2mg of protein was added to 100uL neutravidin magnetic beads (Cytiva Life Sciences) and incubated overnight, rotating at 4C. Beads were washed 5X with RIPA buffer without inhibitors, then once with 1X TBS. Biotinylated proteins were eluted by adding 5% SDS in 50mM TEAB buffer and boiling at 95C for 10 minutes. Eluates were then digested on S-trap columns (Protifi) following the manufacturer's recommendation. For both serum starvation interactomes, a C18 peptide cleanup was performed following S-trap elution. Dried peptides were resuspended in 20uL sample buffer as described previously[134] for mass spectrometry.

For serum starvation interactomes, 7uL of resuspended peptides was injected for each instrument run. HeLa serum starvation interactome LC-MS/MS was performed on an Orbitrap Eclipse Tribrid mass spectrometer (ThermoFisher Scientific) equipped with an Ultimate 3000 RSLCnano HPLC system (Thermofisher Scientific). SHSY5Y Serum starvation interactome LC-MS/MS was performed on an Orbitrap Fusion Tribrid mass spectrometer (ThermoFisher Scientific) equipped with an Ultimate 3000 RSLCnano HPLC system (Thermofisher Scientific). For depolarization interactome, 3.5uL of resuspended peptides was injected for each instrument run. LC-MS/MS was performed on an Orbitrap Lumos Tribrid mass spectrometer (ThermoFisher Scientific) equipped with an Ultimate 3000 RSLCnano HPLC system (Thermofisher Scientific). In

all cases, gradient and instrument settings were identical to those previously described[134], excepting dynamic exclusion occurring for 20 seconds after 2 occurrences in 10 seconds.

Raw files were searched using ThermoFisher Scientific Proteome Discoverer (2.5.0.400) using a combined SequestHT and Mascot (2.6.2) search. The proteomic database consisted of all SwissProt annotated human protein sequences (obtained 03-2021) plus a list of common contaminants and the sequences for eGFP and TurboID. Tryptic cleavage was specific at Arg and Lys with two missed cleavages allowed. For both search engines, precursor mass tolerance was 10ppm and fragment mass tolerance was 0.6Da. Methylthio on Cys was a fixed modification, and variable modifications were allowed: Oxidation of Met, HexNAc on Ser/Thr, and biotin on Lys. For eGFP-TurboID, peptide and PSM FDR were set at 2%, and for TPR-TurboID, peptide and PSM FDR were set at 1%. Proteins were included in each final protein list were required to be the master protein of a protein group, have 2 or greater unique peptides, and a positive score for both Sequest and Mascot.

We utilized the proteomic data analysis software Simplifi (Protifi) for statistical analysis. Statistical significance among groups was determined using a combination of Wilcoxon Signed Rank Test and T-Test. Since Simplifi does not require data to be parametric we did not utilize NSAF of normal logarithmic transformation and instead performed statistical tests on raw spectral counts with 1 added to all spectral counts to avoid zero values. In order to be considered significant, a Wilcoxon p value of less than 0.05, a T-test p value of less than 0.05, and a 2-fold higher spectral count was required. This was required for both comparisons of experimental condition to negative control (TPR to eGFP) and across experimental conditions.

Western Blots and Antibodies

SDS-PAGE gels (4-15%) were run using the BioRad Mini-PROTEAN gel system and transferred onto Immobilin-FL PVDF membranes (Sigma-Aldritch) using the BioRad Trans-Blot SD Semi-Dry Transfer Cell system. Membrane blocking was in 1% cold water fish skin gelatin

(Sigma-Aldrich). Primary antibody ratios were as follows: FLAG (1:2500 Sigma F3165), anti-biotin (1:1000, Jackson 200-002-211), anti-β-actin (1:2000, Cell Signaling 4970). Secondary antibodies were LiCor IRDye: 680RD donkey-anti-mouse 680 (1:10000) and 800CW Goat anti-rabbit (1:20000). Three washes in TBST (0.1% tween) were performed after each antibody incubator. All Western blots were imaged on a Li-Cor Odyssey Clx system.

Gene Ontology Analysis

Gene ontology analysis was performed using The Gene Ontology Resource (geneontology.org)[104], with pathway reduction performed with Revigo[105], and visualized by CirGo[106]. All GO lists were filtered at a p-value and FDR of less than 0.01, and a minimum of 5-fold enrichment over expected number of proteins.

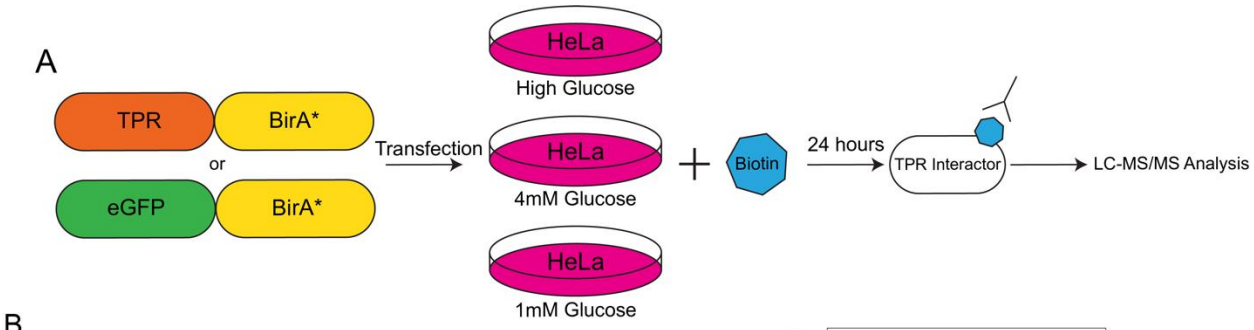
Results

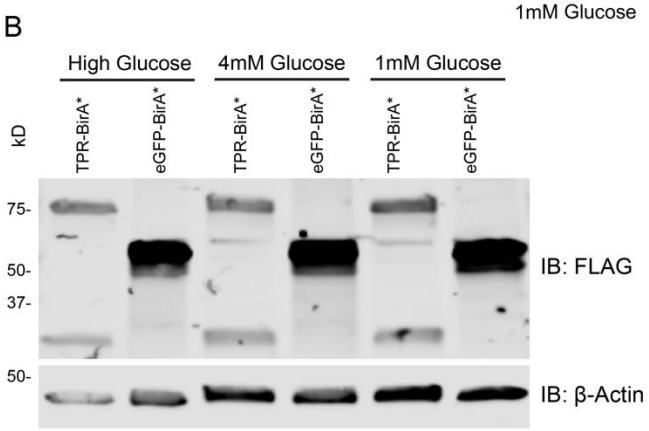
The OGT TPR interactome under various glucose conditions in HeLa cells

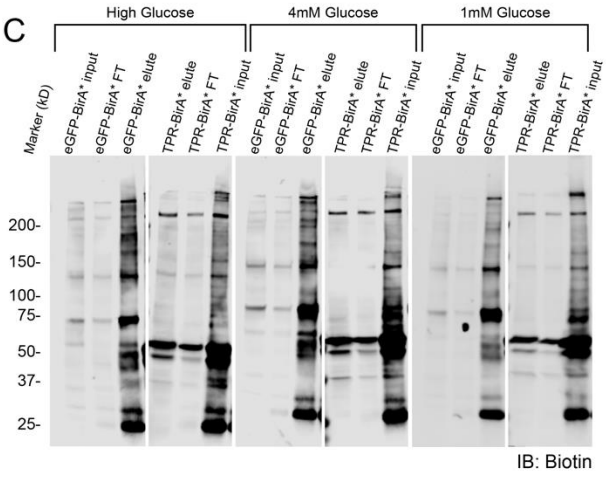
Building on our previous work identifying the basal OGT TPR interactome in HeLa cells, we next asked if this interactome could be altered by dynamic cellular states. OGT and the O-GlcNAc modification serve as a nutrient sensor for the cell by way of their connection to the hexosamine biosynthetic pathway, which synthesizes UDP-GlcNAc, the donor molecular for O-GlcNAc, as a pathway that branches from glycolysis[135]. Therefore, the amount of O-GlcNAc in the cell is directly proportional to the amount of glucose available to the cell. However, the exact mechanism underlying how this sensing occurs is unclear. To determine if OGT TPR interactors and the OGT TPR interactome play a role in this glucose sensing process, we performed BioID in HeLa cells under various glucose conditions (**Fig. 4.1A**). Cells were cultured in three different glucose conditions: “High” glucose (~25mM, standard cell culture glucose), 4mM glucose (representing euglycemia), and 1mM glucose (representing hypoglycemia). Following transfection of either TPR-BirA* or eGFP-BirA* (**Fig. 4.1B**) and biotin enrichment with a biotin IP

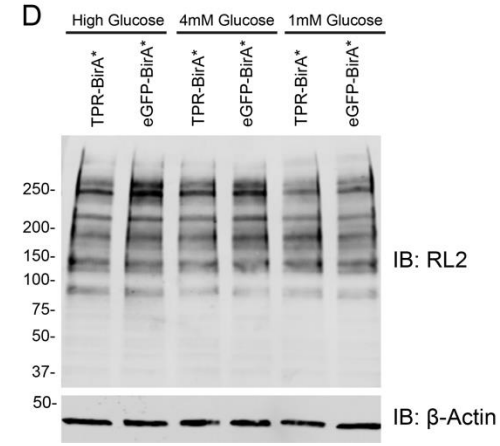
(**Fig. 4.1C)**, protein interactors were identified using mass spectrometry and interactomes compared to identify differential protein interactors.

Figure 4.1E shows volcano plots of all identified interactors in each pair of glucose condition, and full interactome lists are in **Table 4.1**. In order for an interactor to be considered significantly enriched in one glucose condition or the other, it had to pass several statistical metrics. First, an interactor had to be significantly enriched in a TPR-BirA* sample over its corresponding eGFP-BirA* sample, by a Wilcoxon p-value of 0.05, a t-test p-value of 0.05, and at least 2-fold spectral counts. At this point in the data analysis, it would be tempting to simply venn the final interactomes per condition and determine differential interactors by a simple presence/absence metric. However, this method fails to account for proteins that may have been identified in one sample but simply at an abundance too low to "pass" statistically over its own negative control. The variability inherent in spectral count-based relative quantification makes this method risky for creating an abundance of false positives in the differential interactors lists. Therefore, out of our desire to only report very high-confidence differential interactors, we also performed the same statistical testing on proteins across experimental conditions. In any paired analysis, all spectral counts for the experimental condition were used to determine if that protein is significantly enriched in one glucose condition or the other, and this logic follows for the rest of the differential interactomes described in this dissertation. For this reason, in **Table 4.1** (and future interactome tables), proteins identified in an experimental condition over its corresponding negative control are indicated by a gray square, and proteins identified as enriched in one experimental condition over another are indicated by a blue square. A protein being over its negative control in only one condition, but not considered enriched across experimental conditions, indicates that the statistical testing across experimental samples did not result in high confidence that that interaction occurs differentially across conditions.


A
TPR
BirA*
or
eGFP
BirA*
Transfection
HeLa
High Glucose
HeLa
4mM Glucose
HeLa
1mM Glucose
Biotin
24 hours
TPR Interactor
LC-MS/MS Analysis


B
High Glucose
4mM Glucose
1mM Glucose
kD
TPR-BirA*
eGFP-BirA*
TPR-BirA*
eGFP-BirA*
TPR-BirA*
eGFP-BirA*
75-
50-
37-
IB: FLAG
50-
IB: β-Actin


C
High Glucose
4mM Glucose
1mM Glucose
Marker (kD)
eGFP-BirA* input
eGFP-BirA* FT
eGFP-BirA* elute
TPR-BirA* elute
TPR-BirA* FT
TPR-BirA* input
200-
150-
100-
75-
50-
37-
25-
IB: Biotin


D
High Glucose
4mM Glucose
1mM Glucose
TPR-BirA*
eGFP-BirA*
TPR-BirA*
eGFP-BirA*
TPR-BirA*
eGFP-BirA*
250-
200-
150-
100-
75-
50-
37-
IB: RL2
50-
IB: β-Actin

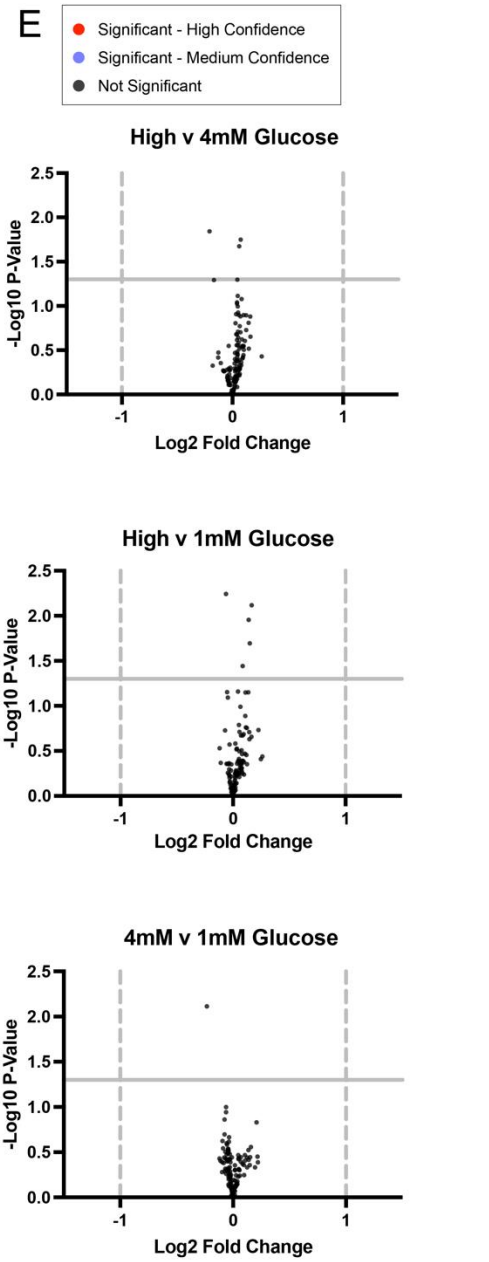

E
Significant - High Confidence
Significant - Medium Confidence
Not Significant
High v 4mM Glucose
-Log10 P-Value
Log2 Fold Change
High v 1mM Glucose
-Log10 P-Value
Log2 Fold Change
4mM v 1mM Glucose
-Log10 P-Value
Log2 Fold Change

Figure 4.1: In a HeLa cell system, glucose starvation has no effect on the OGT TPR interactome. ***A,*** Workflow of HeLa BioID interactomes under different glucose conditions. ***B,*** Expression of BirA* fusion proteins ***C,*** Biotin western blot showing enrichment of biotinylated proteins following biotin immunoprecipitation ***D,*** O-GlcNAc levels across glucose conditions (Western blot-RL2) ***E,*** Volcano plot showing abundance and significance of proteins across glucose conditions. Vertical dashed lines indicate log2 fold change cutoff of 1 (fold change of 2) and horizontal solid line indicates T-test p-value cutoff of 0.05. Red points are "high confidence", meaning that at least one of the two conditions had higher than 10 total spectral counts. Purple points are "medium confidence", meaning that while statistical and fold metrics were achieved, both conditions had less than ten spectral counts.

We were surprised to find that no interactors were significantly enriched in any glucose condition, as is shown in **Figure 4.1E**. In order to investigate why this might be the case, we performed a western blot for O-GlcNAc on cell lysates grown in differential glucose (**Fig. 4.1D)** and found that O-GlcNAc abundance was not significantly changed across glucose conditions. This is counter to what has been described in other cell lines where reducing glucose increases O-GlcNAc abundance[54,136]. It is important to note that while it is possible that this indicates that the OGT TPR interactome isn't dynamic under glucose conditions, there are several caveats to this data set. Our cells were only in differential glucose conditions for 48 hours after being grown in high glucose media. HeLa cells primarily express the low-affinity glucose transporter GLUT1, so their responsiveness to varying glucose conditions especially over such a short time scale may be reduced[137]. This combined with HeLa cells' adaptation to grow in high glucose media may meant that HeLa cells do not have the machinery to rapidly adapt to lower glucose conditions. In order to fully address whether the OGT TPR interactome is dynamic under different glucose conditions, it would be beneficial to explore 1) cells that are grown constitutively under different glucose conditions, or at least cultured for a longer period of time, and 2) cell lines that are more dynamically responsive to glucose, such as pancreatic β-cell lines[138], liver cells[139], or neural cells[140]. It may also be interesting to more directly target the hexosamine biosynthetic pathway by starving cells of L-glutamine.

The OGT TPR interactome under serum starvation in HeLa cells

A majority of lab-cultured cells are adapted to grow in fetal bovine serum, a protein-rich supplement that contains many growth factors that encourage cell proliferation[141]. Many of these growth-promoting pathways, such as the MAPK/ERK axis, are mediated by protein phosphorylation. The O-GlcNAc modification has significant crosstalk with protein phosphorylation, including directly competing for phosphorylation sites in many cases[12]. Given this and that O-GlcNAc is involved in cellular processes which are affected by serum signaling, we set out to see if the OGT TPR interactome is dynamic under different serum conditions in HeLa cells. For this experiment, cells were cultured in serum, serum starved for 24 hours, or serum starved and then pulsed with serum for 10 minutes prior to biotin labeling **(Fig. 4.2A)**..The differential interactors that would be identified by this scheme can be envisioned to fall into several categories: interactions that only occur under constitutive serum signaling, interactions that only occur under serum starvation, and interactions that occur briefly upon the re-addition of serum but do not occur under constitutive serum signaling.

Cells were transfected with TurboID fusion proteins (**Fig 4.2B)**, and biotin-labeled proteins enriched with neutravidin (**Fig 4.2C)**. Like the HeLa glucose interactome, we did not observe noteworthy changes in the global O-GlcNAc profile of cells under various serum conditions (**Fig 4.2D).** However, as observed in **Figure 4.2E** and **Table 4.2**, unlike the glucose interactome, many significant changes in the OGT TPR interactome were observed with alterations in serum state. A small number of differential interactors can be observed when comparing serum and starved cells, in both directions of enrichment. Perhaps most surprisingly, the largest number of significant changes were observed when the cells were pulsed, regardless of whether comparing to serum or starved cells. Curiously, all of these differential interactions are loss of interaction upon re-addition of serum. This phenomenon does not have an immediately obvious explanation, but may suggest that the OGT TPR interactome is fairly stable under any “steady-state” condition, but then

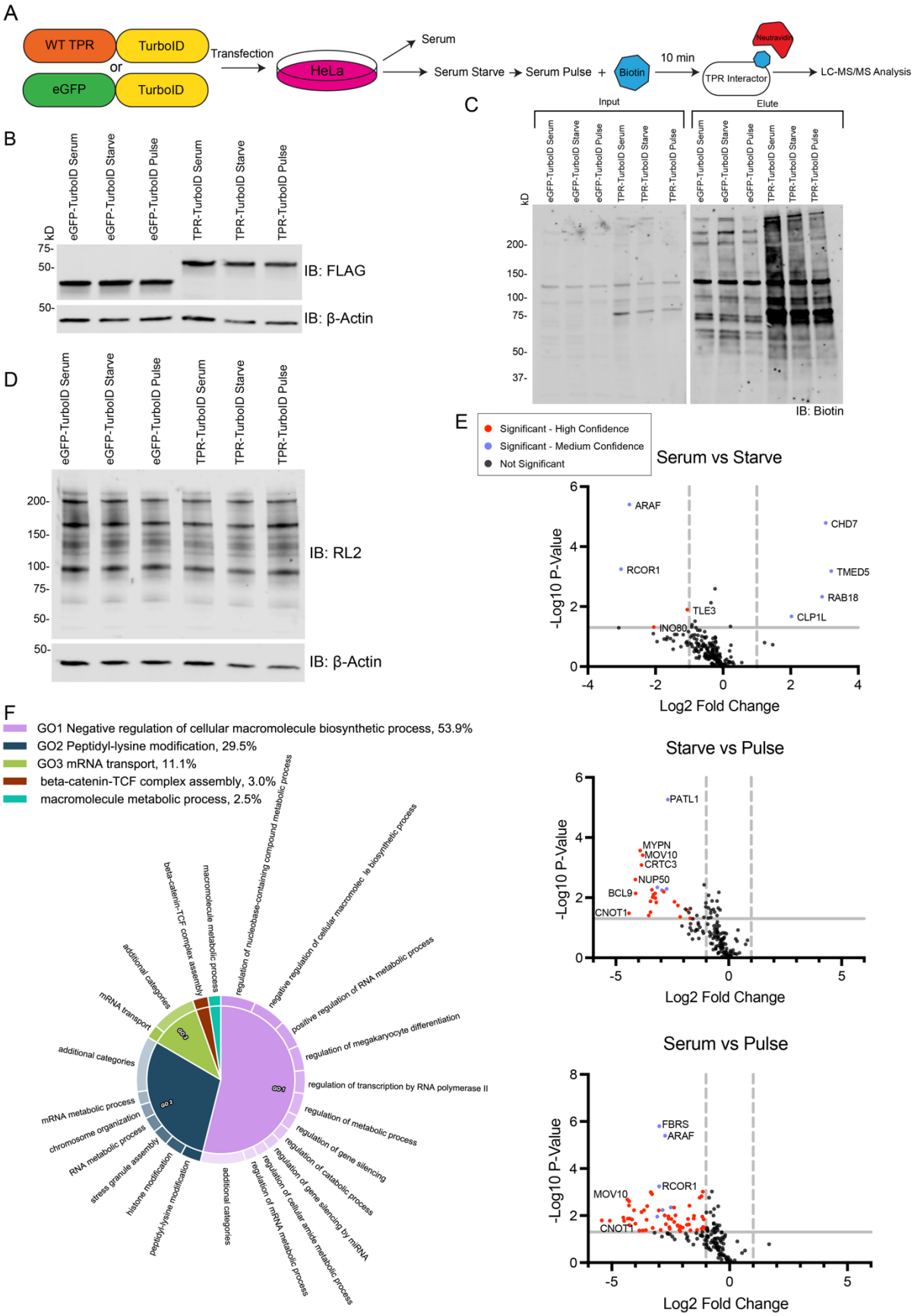

A
WT TPR
TurboID
or
eGFP
TurboID
Transfection
HeLa
Serum
Serum Starve
Serum Pulse
Biotin
10 min
Neutravidin
TPR Interactor
LC-MS/MS Analysis
B
eGFP-TurboID Serum
eGFP-TurboID Starve
eGFP-TurboID Pulse
TPR-TurboID Serum
TPR-TurboID Starve
TPR-TurboID Pulse
kD
75-
50-
IB: FLAG
50-
IB: β-Actin
C
Input
Elute
200-
150-
100-
75-
50-
37-
IB: Biotin
D
200-
150-
100-
75-
50-
IB: RL2
50-
IB: β-Actin
E
Significant - High Confidence
Significant - Medium Confidence
Not Significant
Serum vs Starve
-Log10 P-Value
Log2 Fold Change
ARAF
CHD7
RCOR1
TMED5
RAB18
TLE3
CLP1L
INO80
Starve vs Pulse
PATL1
MYPN
MOV10
CRTC3
NUP50
BCL9
CNOT1
Serum vs Pulse
FBRS
ARAF
RCOR1
MOV10
CNOT1
F
GO1 Negative regulation of cellular macromolecule biosynthetic process, 53.9%
GO2 Peptidyl-lysine modification, 29.5%
GO3 mRNA transport, 11.1%
beta-catenin-TCF complex assembly, 3.0%
macromolecule metabolic process, 2.5%
regulation of nucleobase-containing compound metabolic process
negative regulation of cellular macromolec le biosynthetic process
positive regulation of RNA metabolic process
regulation of megakaryocyte differentiation
regulation of transcription by RNA polymerase II
regulation of metabolic process
regulation of gene silencing
regulation of catabolic process
regulation of gene silencing by miRNA
regulation of cellular amide metabolic process
regulation of mRNA metabolic process
additional categories
peptidyl-lysine modification
histone modification
stress granule assembly
RNA metabolic process
chromosome organization
mRNA metabolic process
additional categories
mRNA transport
additional categories
beta-catenin-TCF complex assembly
macromolecule metabolic process

Figure 4.2: The OGT TPR interactome is dynamic under brief serum stimulation in HeLa cells. ***A***, Workflow of HeLa TurboID interactomes under serum, serum starvation, or pulse of serum after starvation. ***B,*** Expression of TurboID fusion proteins ***C,*** Biotin western blot showing enrichment of biotinylated proteins following neutravidin pulldown. ***D***, O-GlcNAc levels across serum conditions (Western blot – RL2). ***E***, Volcano plots showing abundance and significance across serum conditions. Vertical dashed lines indicate log2 fold change cutoff of 1 (fold change of 2) and horizontal solid line indicates T-test p-value cutoff of 0.05. Red points are "high confidence", meaning that at least one of the two conditions had higher than 10 total spectral counts. Purple points are "medium confidence", meaning that while statistical and fold metrics were achieved, both conditions had less than ten spectral counts. ***F,*** Biological process enrichment of protein whose interactions are decreased in the serum pulse condition

is significantly altered upon a stress to the cell such as the sudden re-instatement of serum signaling.

In order to better understand what cellular function might be served by reduced interaction with OGT during initial serum signaling, we assessed biological process enrichment of all proteins whose interaction is downregulated during serum pulse (**Fig 4.2F**). One very interesting enriched process is beta-catenin-TCF complex assembly. OGT has in fact previously been shown to interact with β-catenin with serum stimulation of HeLa cells[142]. These authors overexpressed β-catenin in order to detect its interaction with OGT, which may be why we don't detect β-catenin in our interactome. It's unclear here why we observe reduced OGT interaction with proteins associated with β-catenin-TCF complex assembly. TCF is a transcriptional activator downstream of Wnt signaling[143], so it is possible that this reduction in OGT interaction is capturing a function of OGT that occurs later in the signaling process. Another interesting enriched process group is negative regulation of cellular macromolecule biosynthetic process, which covers a variety of cellular synthetic processes. Since initial serum signaling would be expected to promote many biosynthetic processes, it is possible that OGT is enhancing negative regulation of synthesis by interacting with proteins belonging to this biological process group, and then the loss of interaction upon serum signaling promotes biosynthesis to occur. However, it is impossible to know the

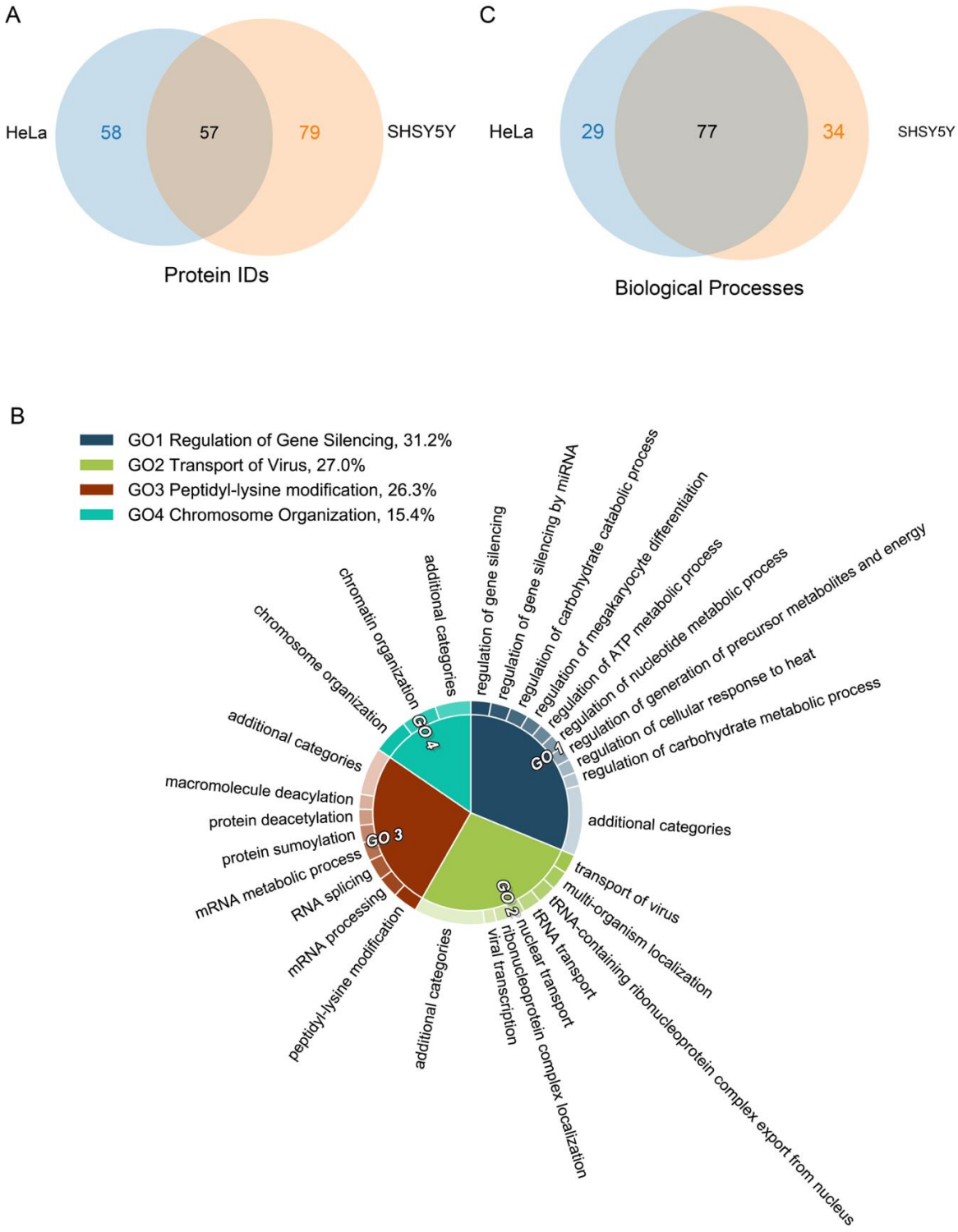


Figure 4.3: Comparison of basal HeLa and SHSY5Y TPR interactomes. ***A,*** Venn of interacting protein IDs from HeLa and SHSY5Y cells. ***B,*** CirGo plot of biological processes enriched in SHSY5Y interactome. ***C,*** Venn of enriched biological processes of interacting proteins from HeLa and SHSY5Y cells.

function of these interactions from this study and further research needs to be performed to identify what function OGT's interaction serves.

The OGT TPR Interactome in HeLa cells versus SHSY5Y cells

One important aspect of how OGT regulates its substrate selection is how its interactome differs among cell types. Since OGT is ubiquitously expressed, it likely interacts with different proteins across cell types in order to mediate the specific functional needs of that cell type. In order to examine how the OGT TPR interactome differs as compared to our HeLa cell interactome, we used TurboID to define an OGT TPR interactome in SHSY5Y cells, a neuroblastoma cell line.

Table 4.3 details the proteins identified in HeLa or SHSY5Y cells, and **Figure 4.3A** shows a venn of the protein identifications. 57 proteins are common to both interactomes, and these proteins represent "basal" OGT interactors between these cell lines. 58 and 79 proteins from HeLa and SHSY5Y, respectively, are unique interactors to that cell line. These interactors likely represent a mix of proteins that uniquely interact with OGT in that cell line and those that are uniquely expressed in that cell type. Several well-known OGT interactors, such as HCF-1, HIRA, nuclear pore proteins, sin3a, and tet2 are common to both interactomes. Interestingly, OGA is identified as an OGT interactor in the HeLa cell interactome, but not SHSY5Y. We also uniquely identify classes of proteins known to interact with OGT in the SHSY5Y interactome, specifically HDACs and protein phosphatases. This points to the importance of understanding OGT interaction as a variable, cell-type-specific phenomenon rather than assuming an interaction observed in one cell type will be universally found or functional.

Considering the diversity of protein interactors across cell type, it is interesting to note that the biological processes enriched among interactors are much more common. **Figure 4.3B** shows biological processes enriched in SHSY5Y cells, for comparison to HeLa cell shown in **Figure 3.5A**. **Figure 4.3C** shows a venn of all enriched biological process terms across HeLa and

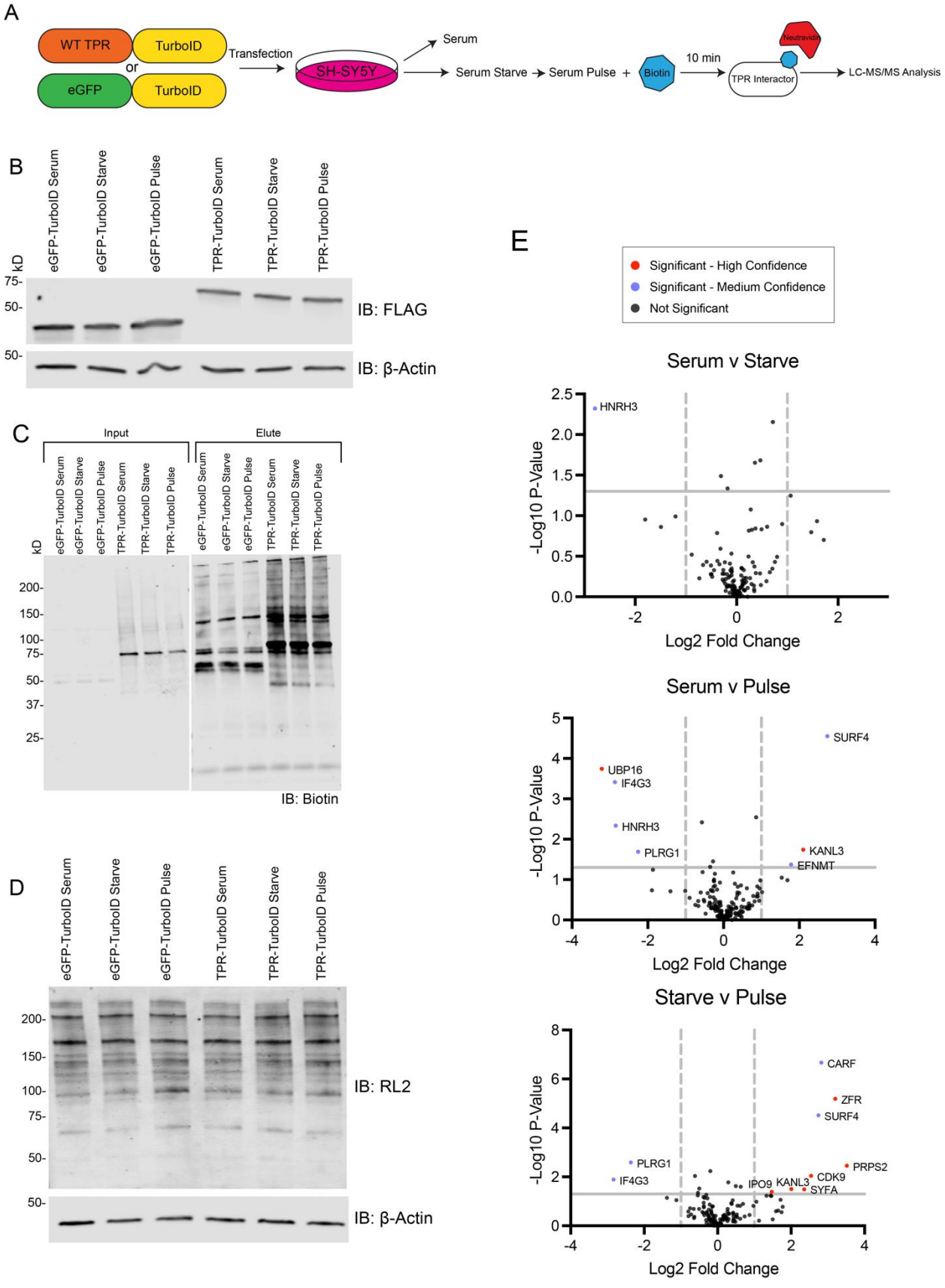

A
WT TPR
TurboID
or
eGFP
TurboID
Transfection
SH-SY5Y
Serum
Serum Starve
Serum Pulse +
Biotin
10 min
Neutravidin
TPR Interactor
LC-MS/MS Analysis
B
eGFP-TurboID Serum
eGFP-TurboID Starve
eGFP-TurboID Pulse
TPR-TurboID Serum
TPR-TurboID Starve
TPR-TurboID Pulse
kD
75-
50-
IB: FLAG
50-
IB: β-Actin
C
Input
Elute
kD
200-
150-
100-
75-
50-
37-
25-
IB: Biotin
D
200-
150-
100-
75-
50-
IB: RL2
50-
IB: β-Actin
E
Significant - High Confidence
Significant - Medium Confidence
Not Significant
Serum v Starve
-Log10 P-Value
Log2 Fold Change
HNRH3
Serum v Pulse
SURF4
UBP16
IF4G3
HNRH3
PLRG1
KANL3
EFNMT
Starve v Pulse
CARF
ZFR
SURF4
PLRG1
IF4G3
IPO9
KANL3
CDK9
SYFA
PRPS2

Figure 4.4: The SHSH5Y OGT TPR interactome is largely stable under various serum conditions. ***A,*** Workflow of SHSY5Y TurboID interactomes under serum, serum starvation, or pulse of serum after starvation. ***B,*** Expression of TurboID fusion proteins **C,** Biotin western blot showing enrichment of biotinylated proteins following neutravidin pulldown ***D,*** O-GlcNAc levels across serum conditions (Western blot - RL2) ***E,*** Volcano plots showing abundance and significance of proteins across serum conditions. Vertical dashed lines indicate log2 fold change cutoff of 1 (fold change of 2) and horizontal solid line indicates T-test p-value cutoff of 0.05. Red points are "high confidence", meaning that at least one of the two conditions had higher than 10 total spectral counts. Purple points are "medium confidence", meaning that while statistical and fold metrics were achieved, both conditions had less than ten spectral counts.

SHSY5Y cells. In contrast to protein interactors, of which less than 30% are in common, over 50% of biological processes are in common between the cell lines. This indicates that OGT may use different protein interactors to accomplish the same biological functions across cell lines.

The OGT TPR interactome under serum starvation in SHSY5Y cells

SHSY5Y cells, like HeLa cells, grow in media supplemented with fetal bovine serum. We performed a similar serum starvation experiment as described above to see if the OGT TPR interactome is similarly dynamic under various serum conditions in SHSY5Y cells (**Fig 4.4A,B,C**). Similarly to HeLa cells, global O-GlcNAc levels are largely unaltered by various serum conditions in SHSY5Y cells (**Fig. 4.4D**). Surprisingly, however, the profile of differential interactors is quite different (**Fig. 4.4E**). The SHSY5Y OGT TPR interactome is largely stable under serum or serum starve conditions, with only one interaction being significantly upregulated under serum. Like HeLa cells, the serum pulse condition has the largest number of differential interactors, although the total number of differential interactors is much smaller. Interestingly, interactions both newly occur and disappear in the pulse condition in SHSY5Y cells. This further corroborates the suggestion that the OGT TPR interactome may be fairly stable under steady-state conditions, and becomes more dynamic in rapid response to stress. However, the much smaller number of differential interactors, and the fact that the differential interactors do not overlap between SHSY5Y and HeLa cells under the same condition, suggests that the degree of this interaction response is highly variable depending on cell type.

Several interactors of interest can be parsed from this smaller list of differential interactors. One such protein is HNRH3, an RNA binding protein involved in mRNA splicing and stress-induced splicing arrest[144], that only interacts with OGT under constitutive serum. While it is impossible to know from this dataset what the functional outcome of this interactions is, one possible explanation is that OGT interacts with HNRH3, restraining its function in a "non-stressed" cellular state, and upon stress the interaction is released, allowing HNRH3 to perform its stress-related function. Further research is necessary to determine the exact function of these interactions, however. Other proteins of interest are IF4G3, which is involved in translation initiation[145], and PLRG1, which is involved in mRNA splicing[146], which both interact with OGT only during constitutive serum or starvation. Again, from this data we cannot determine what the functional outcome of these interactions are but the loss of this interaction upon immediate re-addition of serum may point to some functional role for these proteins during early serum-induced cellular signaling. Finally, the protein interactions that are gained when serum is pulsed after starvation are of high interest in determining what functional roles OGT might play during serum signaling. Of these 6 proteins, several play roles in transcription (CARF, CDK9) or nucleotide synthesis (PRPS2), which may point toward the reactivation of cell cycling following the addition of serum. Further research must be performed to determine what role OGT's interaction with these proteins plays in these processes.

The OGT TPR interactome under depolarization in SHSY5Y cells

OGT is known to play roles in many neuronal processes. It is localized to the synapse[147], involved in the formation of dendritic spines on excitatory neurons[148], and plays important roles in the formation of long-term memory[149]. However, the specific functional roles of OGT and the proteins it interacts with in neural systems have not been well studied. Considering what is known about OGT in the nervous system, it stands to reason that OGT may interact with proteins during important neuronal processes, such as depolarization. SHSY5Y cells, despite not forming

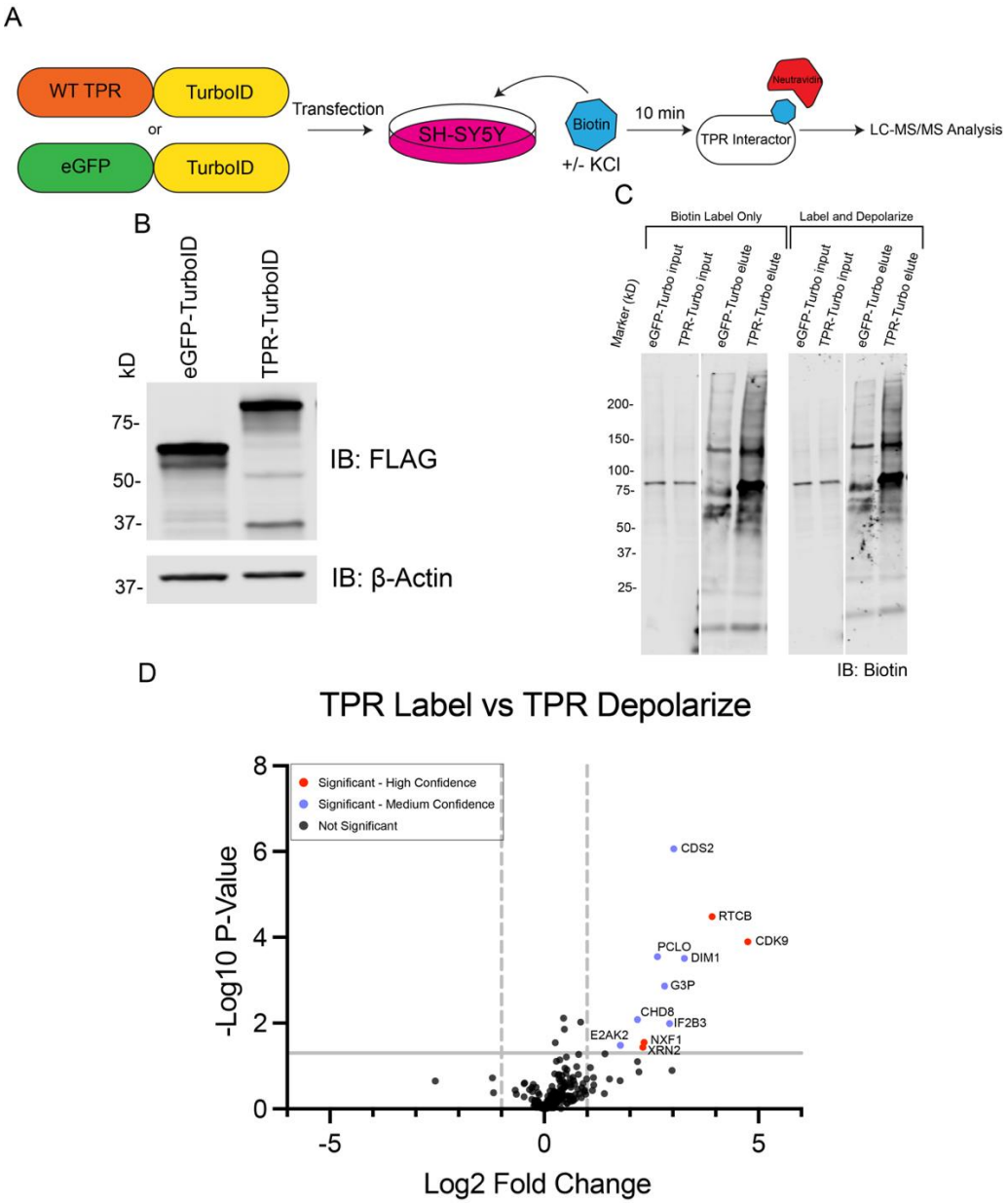


Figure 4.5: OGT gains protein interactions upon neuroblastoma cell depolarization. ***A,*** Workflow of SHSY5Y cell TurboID interactome. Cells were either "label" only with only the addition of biotin, or "label and depolarize" with the addition of both biotin and KCl to depolarize. ***B,*** Western blot showing expression of eGFP-TurboID or TPR-TurboID. ***C,*** Biotin western blot showing enrichment of biotinylated proteins following neutravidin pulldown in both label and depolarize conditions. ***D,*** Volcano plot showing abundance and significance of proteins in label versus depolarize condition. Vertical dashed lines indicate log2 fold change cutoff of 1 (fold change of 2) and horizontal solid line indicates T-test p-value cutoff of 0.05. A positive fold change indicates higher expression in depolarize, a negative fold change indicates a higher expression in label only. Red points are "high confidence", meaning that at least one of the two conditions had higher than 10 total spectral counts. Purple points are "medium confidence", meaning that while statistical and fold metrics were achieved, both conditions had less than ten spectral counts.

functional synapses as immature cells[150], are capable of membrane depolarization[151]. To begin to probe how OGT interactions change in neural cell processes, we used TurboID to determine the OGT TPR interactome under depolarization of SHSY5Y cells (**Fig 4.5A**). After transfection and expression of fusion proteins (**Fig. 4.5B)**, cells were either labeled with biotin (label only condition) or simultaneously labeled and depolarized (depolarize condition). Biotinylated interacting proteins were enriched and processed as before (**Fig. 4.5C**).

As seen in the previous SHSY5Y interactome, the OGT TPR interactome remains largely stable under depolarization (**Fig. 4.5D**). However, several proteins newly interact with OGT upon depolarization. These proteins are largely ubiquitously expressed and may not represent an interaction unique to neuronal cells, except for the protein Piccolo (PCLO), which is part of presynaptic vesicle release and is known to be heavily O-GlcNAc modified[152]. The observation that OGT's interaction with piccolo may be dependent on membrane depolarization prior to synaptic signaling is intriguing. Additional study on the functional role of this interaction and how it alters O-GlcNAc modification on Piccolo may help to further elucidate the roles that OGT plays in synaptic signaling.

It is worth noting that one protein, CDK9, interacts with OGT under depolarization and serum pulse in SHSY5Y cells. This protein may represent a more labile OGT interactor and possibly one that contributes to generic stress response since the interaction occurs under two very different conditions. Proteins belonging to the GO molecular function category "RNA binding" (XNRN2, E2AK2, IF2B3, CDK9, DIM1, NXF1, RTCB) are overrepresented in the depolarization interactome as well. Since these proteins are not neuronal specific, and RNA processing is an important part of cellular stress response[153], it's possible that these OGT interactors are part of a generic stress response rather than a specific response to depolarization. Both studying stress responses in different cell types and looking at OGT interactions under synaptic signaling in functional neurons will help to elucidate these differences.

Discussion

While the identify of O-GlcNAc modified proteins has been extensively studied[154], the OGT interactome remains more of a mystery, and especially how the interactome changes to facilitate OGT's functional roles across cell types and functions. Here, we have explored two cell types and several cellular states to begin to probe what proteins the TPR domain of OGT interacts with and how those interactions change with cellular state.

The OGT TPR interactome across different cell types (here, a cervical cancer and neuroblastoma cell line) differs significantly, although a core set of interactors remain the same, many of which are previously studied. It's likely that these interactors represent the more frequent, abundant, and universal interactions, which explains why they have previously been captured by interaction study methods that are more limited (like immunoprecipitation). The novel and differential interactions may represent interactions that are more transient, infrequent, and cell-type specific. Identifying these proteins interactors is the first step in understanding how OGT's functional roles differ among cell types.

Under different cellular states, the OGT TPR interactome appears to be dynamic, but to a smaller degree than may have been suggested by the wide diversity of functions OGT is involved in. It also appears that the degree of interactome alteration in response to a specific stimulus may vary by cell type, e.g. how the OGT TPR interactome changes vastly under changes in serum signaling in HeLa cells, but only exhibits more subtle changes in SHSY5Y cells. It is interesting to note that specific interactor changes for this condition do not overlap between cell types despite differential interactors belonging to both the common and unique interactome. Furthermore, whether changes in the interactome occur at all may be very cell-type dependent. The OGT TPR interactome in HeLa cells under various glucose conditions appears to be extremely stable, but this may not be the case in a more glucose-sensitive cell type.

It is also interesting to note that when the OGT TPR interactome changes, under the conditions we've studied, larger changes occur under brief, intense stimulus (pulsing with serum, depolarization) than long-term, "steady-state" stimulus (serum, serum starvation, altered glucose conditions). OGT has previously been characterized as a "rheostat"[39], pointing toward it as a subtle tuner of cellular function rather than one that performs broad, sweeping changes (like protein kinases tend to). The interactome data here presents a potential new and intriguing angle in this discussion – that the OGT interactome remains largely stable except under significant cellular stress, where it will be altered to bring the cell back to homeostasis. One of the only other published OGT interactomes[155] supports this, where significant changes in the OGT interactome were observed under cellular oxidative stress. Their observation of a greater number of changes may be due to the OGT interactome being more responsive under oxidative stress than the conditions we've studied here. The immunoprecipitation-based method used in that study may also contribute to greater depth of protein identification due to the ability to reduce pulldown of nonspecific interactors, which may enhance their identification of differential interactors.

It is important to mention here that OGT is expressed ubiquitously and to reiterate that it performs an immense variety of cellular functions. It is impossible from the data presented here to fully generalize about how the OGT interactome functions as a whole. What we've attempted to do here is provide an initial look into the dynamics of this interactome and suggestions for further study. Recommendations for future study are discussed in detail in Chapter 6, but briefly, identifying OGT interactomes in primary cells and with very targeted cellular stimulus will be beneficial in further elucidating how and when the OGT interactome is altered.

Table 4.1 TPR-TurboID interacting proteins under different glucose conditions in HeLa cells

Dark gray - identified only in TPR-TurboID, not present in eGFP-TurboID
Light gray - identified in both TPR-TurboID and eGFP-TurboID, significantly enriched in TPR-TurboID

Accession	Description	High over eGFP	4mM over eGFP	1mM over eGFP
NUP54_HUMAN	Nucleoporin p54			
RREB1_HUMAN	Ras-responsive element-binding protein 1			
CCAR1_HUMAN	Cell division cycle and apoptosis regulator protein 1			
UBN1_HUMAN	Ubinuclein-1			
ARI5B_HUMAN	AT-rich interactive domain-containing protein 5B			
ZN609_HUMAN	Zinc finger protein 609			
LRIF1_HUMAN	Ligand-dependent nuclear receptor-interacting factor 1			
TBL1X_HUMAN	F-box-like/WD repeat-containing protein TBL1X			
NUP62_HUMAN	Nuclear pore glycoprotein p62			
RBM33_HUMAN	RNA-binding protein 33			
NUP58_HUMAN	Nucleoporin p58/p45			
TAF6_HUMAN	Transcription initiation factor TFIID subunit 6			
MEF2D_HUMAN	Myocyte-specific enhancer factor 2D			
NCOA2_HUMAN	Nuclear receptor coactivator 2			
SIN3A_HUMAN	Paired amphipathic helix protein Sin3a			
TNR6A_HUMAN	Trinucleotide repeat-containing gene 6A protein			
ZC3HE_HUMAN	Zinc finger CCCH domain-containing protein 14			
BCOR_HUMAN	BCL-6 corepressor			
CABIN_HUMAN	Calcineurin-binding protein cabin-1			
SMRC1_HUMAN	SWI/SNF complex subunit SMARCC1			
TET2_HUMAN	Methylcytosine dioxygenase TET2			
KMT2C_HUMAN	Histone-lysine N-methyltransferase 2C			
CIC_HUMAN	Protein capicua homolog			
PRR12_HUMAN	Proline-rich protein 12			
PCF11_HUMAN	Pre-mRNA cleavage complex 2 protein Pcf11			
PHC3_HUMAN	Polyhomeotic-like protein 3			
RBM26_HUMAN	RNA-binding protein 26			
KMT2D_HUMAN	Histone-lysine N-methyltransferase 2D			
ZEP1_HUMAN	Zinc finger protein 40			
ZHX3_HUMAN	Zinc fingers and homeoboxes protein 3			
RGPD3_HUMAN	RanBP2-like and GRIP domain-containing protein 3			
ZHX2_HUMAN	Zinc fingers and homeoboxes protein 2			
UBN2_HUMAN	Ubinuclein-2			
JHD2C_HUMAN	Probable JmjC domain-containing histone demethylation protein 2C			
SMRC2_HUMAN	SWI/SNF complex subunit SMARCC2			
PRC2B_HUMAN	Protein PRRC2B			
BAP1_HUMAN	Ubiquitin carboxyl-terminal hydrolase BAP1			
P121C_HUMAN	Nuclear envelope pore membrane protein POM 121C			
RGPD5_HUMAN	RANBP2-like and GRIP domain-containing protein 5/6			
RESF1_HUMAN	Retroelement silencing factor 1			

ZHX1_HUMAN	Zinc fingers and homeoboxes protein 1			
ARI1B_HUMAN	AT-rich interactive domain-containing protein 1B			
ARI1A_HUMAN	AT-rich interactive domain-containing protein 1A			
HIRA_HUMAN	Protein HIRA			
SET1A_HUMAN	Histone-lysine N-methyltransferase SETD1A			
ZN318_HUMAN	Zinc finger protein 318			
ZFR_HUMAN	Zinc finger RNA-binding protein			
RPRD2_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 2			
LIN54_HUMAN	Protein lin-54 homolog			
QSER1_HUMAN	Glutamine and serine-rich protein 1			
TAF4_HUMAN	Transcription initiation factor TFIID subunit 4			
ACL6A_HUMAN	Actin-like protein 6A			
TRRAP_HUMAN	Transformation/transcription domain-associated protein			
ASXL2_HUMAN	Putative Polycomb group protein ASXL2			
OGA_HUMAN	Protein O-GlcNAcase			
I2BP2_HUMAN	Interferon regulatory factor 2-binding protein 2			
POGZ_HUMAN	Pogo transposable element with ZNF domain			
NCOA6_HUMAN	Nuclear receptor coactivator 6			
TRPS1_HUMAN	Zinc finger transcription factor Trps1			
SF3A1_HUMAN	Splicing factor 3A subunit 1			
TASO2_HUMAN	Protein TASOR 2			
EP300_HUMAN	Histone acetyltransferase p300			
KDM2B_HUMAN	Lysine-specific demethylase 2B			
CEP85_HUMAN	Centrosomal protein of 85 kDa			
AHDC1_HUMAN	AT-hook DNA-binding motif-containing protein 1			
BCL9_HUMAN	B-cell CLL/lymphoma 9 protein			
PF21A_HUMAN	PHD finger protein 21A			
ZFHX3_HUMAN	Zinc finger homeobox protein 3			
DOT1L_HUMAN	Histone-lysine N-methyltransferase, H3 lysine-79 specific			
CBP_HUMAN	CREB-binding protein			
ICE1_HUMAN	Little elongation complex subunit 1			
RGPD8_HUMAN	RANBP2-like and GRIP domain-containing protein 8			
NUP50_HUMAN	Nuclear pore complex protein Nup50			
NCOR2_HUMAN	Nuclear receptor corepressor 2			
PRC2A_HUMAN	Protein PRRC2A			
ANR17_HUMAN	Ankyrin repeat domain-containing protein 17			
ZC3H1_HUMAN	Zinc finger C3H1 domain-containing protein			
KDM1A_HUMAN	Lysine-specific histone demethylase 1A			
TAF5_HUMAN	Transcription initiation factor TFIID subunit 5			
YLPM1_HUMAN	YLP motif-containing protein 1			
RBM27_HUMAN	RNA-binding protein 27			
PR40A_HUMAN	Pre-mRNA-processing factor 40 homolog A			
GANP_HUMAN	Germinal-center associated nuclear protein			
EMSY_HUMAN	BRCA2-interacting transcriptional repressor EMSY			
ZN281_HUMAN	Zinc finger protein 281			
SMCA4_HUMAN	Transcription activator BRG1			
DIDO1_HUMAN	Death-inducer obliterator 1			
UBAP2_HUMAN	Ubiquitin-associated protein 2			
NU153_HUMAN	Nuclear pore complex protein Nup153			
KDM3B_HUMAN	Lysine-specific demethylase 3B			

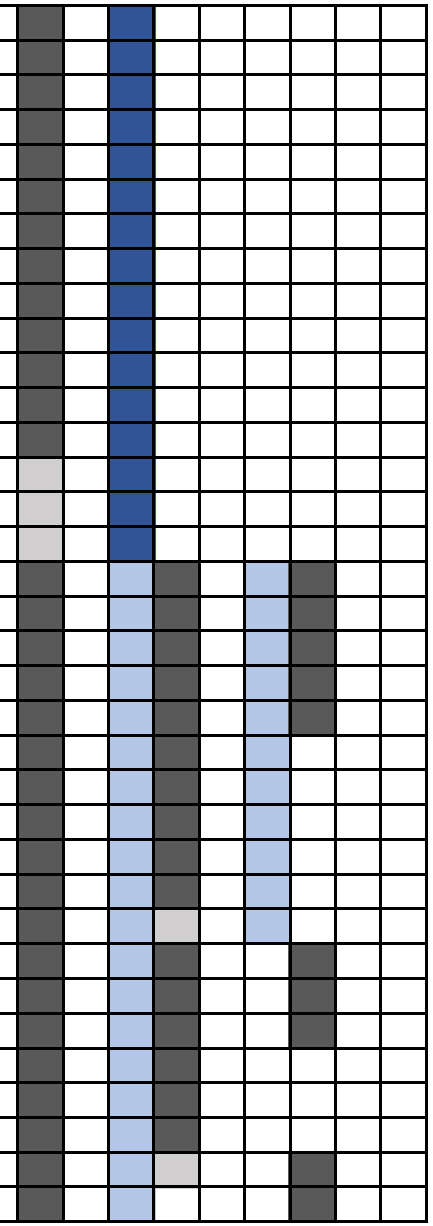
EP400_HUMAN	E1A-binding protein p400			
MINT_HUMAN	Msx2-interacting protein			
OGT1_HUMAN	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit			
ANKH1_HUMAN	Ankyrin repeat and KH domain-containing protein 1			
TBL1R_HUMAN	F-box-like/WD repeat-containing protein TBL1XR1			
PRC2C_HUMAN	Protein PRRC2C			
WNK1_HUMAN	Serine/threonine-protein kinase WNK1			
NU214_HUMAN	Nuclear pore complex protein Nup214			
RBM25_HUMAN	RNA-binding protein 25			
RAGP1_HUMAN	Ran GTPase-activating protein 1			
NUP88_HUMAN	Nuclear pore complex protein Nup88			
HCFC1_HUMAN	Host cell factor 1			
NCOR1_HUMAN	Nuclear receptor corepressor 1			
UBP2L_HUMAN	Ubiquitin-associated protein 2-like			
RBP2_HUMAN	E3 SUMO-protein ligase RanBP2			
KMT2A_HUMAN	Histone-lysine N-methyltransferase 2A			
BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF			
PHF12_HUMAN	PHD finger protein 12			
MTREX_HUMAN	Exosome RNA helicase MTR4			
CPSF1_HUMAN	Cleavage and polyadenylation specificity factor subunit 1			
TNR6B_HUMAN	Trinucleotide repeat-containing gene 6B protein			
ZN106_HUMAN	Zinc finger protein 106			
HNRPL_HUMAN	Heterogeneous nuclear ribonucleoprotein L			
ILF2_HUMAN	Interleukin enhancer-binding factor 2			
KNL1_HUMAN	Kinetochore scaffold 1			
TAB3_HUMAN	TGF-beta-activated kinase 1 and MAP3K7-binding protein 3			
NFRKB_HUMAN	Nuclear factor related to kappa-B-binding protein			
RGPD1_HUMAN	RANBP2-like and GRIP domain-containing protein 1			
SRCAP_HUMAN	Helicase SRCAP			
ATX1L_HUMAN	Ataxin-1-like			
KANL3_HUMAN	KAT8 regulatory NSL complex subunit 3			
SF3B2_HUMAN	Splicing factor 3B subunit 2			
LIMD1_HUMAN	LIM domain-containing protein 1			
SC24B_HUMAN	Protein transport protein Sec24B			
RAE1L_HUMAN	mRNA export factor			
WDR33_HUMAN	pre-mRNA 3' end processing protein WDR33			
GSE1_HUMAN	Genetic suppressor element 1			
BCL9L_HUMAN	B-cell CLL/lymphoma 9-like protein			
YTHD3_HUMAN	YTH domain-containing family protein 3			
NCOA3_HUMAN	Nuclear receptor coactivator 3			
YTHD2_HUMAN	YTH domain-containing family protein 2			
ZMYM4_HUMAN	Zinc finger MYM-type protein 4			
YTHD1_HUMAN	YTH domain-containing family protein 1			
DLDH_HUMAN	Dihydrolipoyl dehydrogenase, mitochondrial			
MGAP_HUMAN	MAX gene-associated protein			
ODO1_HUMAN	2-oxoglutarate dehydrogenase, mitochondrial			
PCCB_HUMAN	Propionyl-CoA carboxylase beta chain, mitochondrial			
NUP98_HUMAN	Nuclear pore complex protein Nup98-Nup96			
RBM10_HUMAN	RNA-binding protein 10			
BUB3_HUMAN	Mitotic checkpoint protein BUB3			

EMSA1_HUMAN	ELM2 and SANT domain-containing protein 1			
WDR5_HUMAN	WD repeat-containing protein 5			
CHD8_HUMAN	Chromodomain-helicase-DNA-binding protein 8			
ZN384_HUMAN	Zinc finger protein 384			
TRI33_HUMAN	E3 ubiquitin-protein ligase TRIM33			
YETS2_HUMAN	YEATS domain-containing protein 2			
NUFP2_HUMAN	Nuclear fragile X mental retardation-interacting protein 2			
ATX2L_HUMAN	Ataxin-2-like protein			

Table 4.2 TPR-TurboID interacting proteins under different serum conditions in HeLa cells

Dark gray - identified only in TPR-TurboID, not present in eGFP-TurboID
Light gray - identified in both TPR-TurboID and eGFP-TurboID, significantly enriched in TPR-TurboID
Dark blue - identified only in the indicated TPR-TurboID condition
Light blue - identified in both TPR-TurboID conditions, significantly enriched in indicated TPR-TurboID

		Serum			Starve			Pulse		
Accession	**Description**	Over eGFP	Over Starve	Over Pulse	Over eGFP	Over Serum	Over Pulse	Over eGFP	Over Serum	Over Starve
RCOR1_HUMAN	REST corepressor 1									
ARAF_HUMAN	Serine/threonine-protein kinase A-Raf									
INO80_HUMAN	Chromatin-remodeling ATPase INO80									
TLE3_HUMAN	Transducin-like enhancer protein 3									
CNOT1_HUMAN	CCR4-NOT transcription complex subunit 1									
NUP50_HUMAN	Nuclear pore complex protein Nup50									
BCL9_HUMAN	B-cell CLL/lymphoma 9 protein									
NCOA3_HUMAN	Nuclear receptor coactivator 3									
TIF1A_HUMAN	Transcription intermediary factor 1-alpha									
YETS2_HUMAN	YEATS domain-containing protein 2									
PRR12_HUMAN	Proline-rich protein 12									
RBM33_HUMAN	RNA-binding protein 33									
GGYF2_HUMAN	GRB10-interacting GYF protein 2									
PATL1_HUMAN	Protein PAT1 homolog 1									
MOV10_HUMAN	Helicase MOV-10									
ZN638_HUMAN	Zinc finger protein 638									
MYPN_HUMAN	Myopalladin									
ATN1_HUMAN	Atrophin-1									
CRTC3_HUMAN	CREB-regulated transcription coactivator 3									
YLPM1_HUMAN	YLP motif-containing protein 1									
SC24B_HUMAN	Protein transport protein Sec24B									

GANP_HUMAN	Germinal-center associated nuclear protein	
FUBP2_HUMAN	Far upstream element-binding protein 2	
SUMO1_HUMAN	Small ubiquitin-related modifier 1	
RPB1_HUMAN	DNA-directed RNA polymerase II subunit RPB1	
ZHX3_HUMAN	Zinc fingers and homeoboxes protein 3	
FOSL2_HUMAN	Fos-related antigen 2	
CUX1_HUMAN	Homeobox protein cut-like 1	
UBAP2_HUMAN	Ubiquitin-associated protein 2	
CD2B2_HUMAN	CD2 antigen cytoplasmic tail-binding protein 2	
MBD3_HUMAN	Methyl-CpG-binding domain protein 3	
BCL9L_HUMAN	B-cell CLL/lymphoma 9-like protein	
SBNO1_HUMAN	Protein strawberry notch homolog 1	
FBRS_HUMAN	Probable fibrosin-1	
RENT1_HUMAN	Regulator of nonsense transcripts 1	
AGFG1_HUMAN	Arf-GAP domain and FG repeat-containing protein 1	
SRRM2_HUMAN	Serine/arginine repetitive matrix protein 2	
KMT2D_HUMAN	Histone-lysine N-methyltransferase 2D	
WNK1_HUMAN	Serine/threonine-protein kinase WNK1	
TRI33_HUMAN	E3 ubiquitin-protein ligase TRIM33	
OGA_HUMAN	Protein O-GlcNAcase	
DIDO1_HUMAN	Death-inducer obliterator 1	
ANKH1_HUMAN	Ankyrin repeat and KH domain-containing protein 1	
P121C_HUMAN	Nuclear envelope pore membrane protein POM121C	
KDM3B_HUMAN	Lysine-specific demethylase 3B	
NFIC_HUMAN	Nuclear factor 1 C-type	
EPN4_HUMAN	Clathrin interactor 1	
ANR17_HUMAN	Ankyrin repeat domain-containing protein 17	
RBM26_HUMAN	RNA-binding protein 26	
NCOR2_HUMAN	Nuclear receptor corepressor 2	
CCAR2_HUMAN	Cell cycle and apoptosis regulator protein 2	
WDR5_HUMAN	WD repeat-containing protein 5	
TNR6B_HUMAN	Trinucleotide repeat-containing gene 6B protein	
CIC_HUMAN	Protein capicua homolog	
TFG_HUMAN	Protein TFG	
SIN3A_HUMAN	Paired amphipathic helix protein Sin3a	

PRC2C_HUMAN	Protein PRRC2C									
NUFP2_HUMAN	Nuclear fragile X mental retardation-interacting protein 2									
ATX2L_HUMAN	Ataxin-2-like protein									
ZN609_HUMAN	Zinc finger protein 609									
UBP2L_HUMAN	Ubiquitin-associated protein 2-like									
CHAP1_HUMAN	Chromosome alignment-maintaining phosphoprotein 1									
SC16A_HUMAN	Protein transport protein Sec16A									
PRPF3_HUMAN	U4/U6 small nuclear ribonucleoprotein Prp3									
RAB18_HUMAN	Ras-related protein Rab-18									
CHD7_HUMAN	Chromodomain-helicase-DNA-binding protein 7									
TMED5_HUMAN	Transmembrane emp24 domain-containing protein 5									
CLP1L_HUMAN	Cleft lip and palate transmembrane protein 1-like protein									
KMT2C_HUMAN	Histone-lysine N-methyltransferase 2C									
YTHD1_HUMAN	YTH domain-containing family protein 1									
ASXL2_HUMAN	Putative Polycomb group protein ASXL2									
BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF									
GFPT1_HUMAN	Glutamine--fructose-6-phosphate aminotransferase [isomerizing] 1									
CEBPB_HUMAN	CCAAT/enhancer-binding protein beta									
QSER1_HUMAN	Glutamine and serine-rich protein 1									
SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich									
HDAC3_HUMAN	Histone deacetylase 3									
ZC3H1_HUMAN	Zinc finger C3H1 domain-containing protein									
KDM1A_HUMAN	Lysine-specific histone demethylase 1A									
ZN217_HUMAN	Zinc finger protein 217									
NCOR1_HUMAN	Nuclear receptor corepressor 1									
NUP98_HUMAN	Nuclear pore complex protein Nup98-Nup96									
RAE1L_HUMAN	mRNA export factor									
KMT2A_HUMAN	Histone-lysine N-methyltransferase 2A									
GPS2_HUMAN	G protein pathway suppressor 2									
GSE1_HUMAN	Genetic suppressor element 1									
RBBP5_HUMAN	Retinoblastoma-binding protein 5									
SF01_HUMAN	Splicing factor 1									
ZMYM4_HUMAN	Zinc finger MYM-type protein 4									
TET2_HUMAN	Methylcytosine dioxygenase TET2									
ZC3HE_HUMAN	Zinc finger CCCH domain-containing protein 14									

EMSY_HUMAN	BRCA2-interacting transcriptional repressor EMSY										
DOT1L_HUMAN	Histone-lysine N-methyltransferase, H3 lysine-79 specific										
ZFR_HUMAN	Zinc finger RNA-binding protein										
RBM27_HUMAN	RNA-binding protein 27										
ASH2L_HUMAN	Set1/Ash2 histone methyltransferase complex subunit ASH2										
OGT1_HUMAN	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit										
RING2_HUMAN	E3 ubiquitin-protein ligase RING2										
HLAB_HUMAN	HLA class I histocompatibility antigen, B alpha chain										
RBM10_HUMAN	RNA-binding protein 10										
ZEP1_HUMAN	Zinc finger protein 40										
AHDC1_HUMAN	AT-hook DNA-binding motif-containing protein 1										
ARI1B_HUMAN	AT-rich interactive domain-containing protein 1B										
RPR1B_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 1B										
KNL1_HUMAN	Kinetochore scaffold 1										
SC23A_HUMAN	Protein transport protein Sec23A										
TBL1X_HUMAN	F-box-like/WD repeat-containing protein TBL1X										
CCNT1_HUMAN	Cyclin-T1										
ZN148_HUMAN	Zinc finger protein 148										
TAF6_HUMAN	Transcription initiation factor TFIID subunit 6										
MINT_HUMAN	Msx2-interacting protein										
CCAR1_HUMAN	Cell division cycle and apoptosis regulator protein 1										
HSP72_HUMAN	Heat shock-related 70 kDa protein 2										
PRC2B_HUMAN	Protein PRRC2B										
TNR6A_HUMAN	Trinucleotide repeat-containing gene 6A protein										
ILF2_HUMAN	Interleukin enhancer-binding factor 2										
PHF12_HUMAN	PHD finger protein 12										
HNRL1_HUMAN	Heterogeneous nuclear ribonucleoprotein U-like protein 1										
BAP1_HUMAN	Ubiquitin carboxyl-terminal hydrolase BAP1										
SP130_HUMAN	Histone deacetylase complex subunit SAP130										
YAP1_HUMAN	Transcriptional coactivator YAP1										
RU2B_HUMAN	U2 small nuclear ribonucleoprotein B''										
NFRKB_HUMAN	Nuclear factor related to kappa-B-binding protein										
BCOR_HUMAN	BCL-6 corepressor										
NUP58_HUMAN	Nucleoporin p58/p45										
LIN54_HUMAN	Protein lin-54 homolog										

ZC11A_HUMAN	Zinc finger CCCH domain-containing protein 11A									
UBC9_HUMAN	SUMO-conjugating enzyme UBC9									
CBX3_HUMAN	Chromobox protein homolog 3									
AKAP8_HUMAN	A-kinase anchor protein 8									
MTREX_HUMAN	Exosome RNA helicase MTR4									
PRP31_HUMAN	U4/U6 small nuclear ribonucleoprotein Prp31									
JUNB_HUMAN	Transcription factor jun-B									
RING1_HUMAN	E3 ubiquitin-protein ligase RING1									
HTF4_HUMAN	Transcription factor 12									
SET1A_HUMAN	Histone-lysine N-methyltransferase SETD1A									
BRD1_HUMAN	Bromodomain-containing protein 1									
POGZ_HUMAN	Pogo transposable element with ZNF domain									
YTHD2_HUMAN	YTH domain-containing family protein 2									
NUP88_HUMAN	Nuclear pore complex protein Nup88									
NUP62_HUMAN	Nuclear pore glycoprotein p62									
HNRPK_HUMAN	Heterogeneous nuclear ribonucleoprotein K									
REQU_HUMAN	Zinc finger protein ubi-d4									
PRP4_HUMAN	U4/U6 small nuclear ribonucleoprotein Prp4									
PABP4_HUMAN	Polyadenylate-binding protein 4									
HDAC1_HUMAN	Histone deacetylase 1									
RPRD2_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 2									
SF3A3_HUMAN	Splicing factor 3A subunit 3									
MGAP_HUMAN	MAX gene-associated protein									
SR140_HUMAN	U2 snRNP-associated SURP motif-containing protein									
SNW1_HUMAN	SNW domain-containing protein 1									
TBL1R_HUMAN	F-box-like/WD repeat-containing protein TBL1XR1									
NIPBL_HUMAN	Nipped-B-like protein									
IMB1_HUMAN	Importin subunit beta-1									
RBBP7_HUMAN	Histone-binding protein RBBP7									
EP400_HUMAN	E1A-binding protein p400									
ARI1A_HUMAN	AT-rich interactive domain-containing protein 1A									
PABP1_HUMAN	Polyadenylate-binding protein 1									
NU214_HUMAN	Nuclear pore complex protein Nup214									
RAGP1_HUMAN	Ran GTPase-activating protein 1									
NU153_HUMAN	Nuclear pore complex protein Nup153									

RBP2_HUMAN	E3 SUMO-protein ligase RanBP2										
HCFC1_HUMAN	Host cell factor 1										
HNRPM_HUMAN	Heterogeneous nuclear ribonucleoprotein M										
TurboID	TurboID TurboID with 2X FLAG and 3X GGGGS linker										
SF3A1_HUMAN	Splicing factor 3A subunit 1										
HDAC2_HUMAN	Histone deacetylase 2										
ZN281_HUMAN	Zinc finger protein 281										
RBM14_HUMAN	RNA-binding protein 14										
LARP4_HUMAN	La-related protein 4										
U5S1_HUMAN	116 kDa U5 small nuclear ribonucleoprotein component										
SNR40_HUMAN	U5 small nuclear ribonucleoprotein 40 kDa protein										
SMD2_HUMAN	Small nuclear ribonucleoprotein Sm D2										
U520_HUMAN	U5 small nuclear ribonucleoprotein 200 kDa helicase										
RSMB_HUMAN	Small nuclear ribonucleoprotein-associated proteins B and B'										
SF3B2_HUMAN	Splicing factor 3B subunit 2										
PRP8_HUMAN	Pre-mRNA-processing-splicing factor 8										
SYF1_HUMAN	Pre-mRNA-splicing factor SYF1										
CABIN_HUMAN	Calcineurin-binding protein cabin-1										
ARI5B_HUMAN	AT-rich interactive domain-containing protein 5B										
RALY_HUMAN	RNA-binding protein Raly										
PFKAL_HUMAN	ATP-dependent 6-phosphofructokinase, liver type										
CPSF2_HUMAN	Cleavage and polyadenylation specificity factor subunit 2										
JHD2C_HUMAN	Probable JmjC domain-containing histone demethylation protein 2C										
ATX2_HUMAN	Ataxin-2										
CARF_HUMAN	CDKN2A-interacting protein										
CHD8_HUMAN	Chromodomain-helicase-DNA-binding protein 8										
MCM3_HUMAN	DNA replication licensing factor MCM3										
CPSF1_HUMAN	Cleavage and polyadenylation specificity factor subunit 1										
P66A_HUMAN	Transcriptional repressor p66-alpha										
CPSF5_HUMAN	Cleavage and polyadenylation specificity factor subunit 5										
PF21A_HUMAN	PHD finger protein 21A										
CPSF7_HUMAN	Cleavage and polyadenylation specificity factor subunit 7										
SMG7_HUMAN	Protein SMG7										

Table 4.3: Comparison of TPR interactors between HeLa and SHSY5Y cells

Accession	Description	In HeLa?	In SHSY5Y?
ANKH1_HUMAN	Ankyrin repeat and KH domain-containing protein 1		
ANR17_HUMAN	Ankyrin repeat domain-containing protein 17		
ARI1A_HUMAN	AT-rich interactive domain-containing protein 1A		
ARI1B_HUMAN	AT-rich interactive domain-containing protein 1B		
BCL9_HUMAN	B-cell CLL/lymphoma 9 protein		
BCOR_HUMAN	BCL-6 corepressor		
BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF		
CCAR1_HUMAN	Cell division cycle and apoptosis regulator protein 1		
CIC_HUMAN	Protein capicua homolog		
DIDO1_HUMAN	Death-inducer obliterator 1		
EMSY_HUMAN	BRCA2-interacting transcriptional repressor EMSY		
GANP_HUMAN	Germinal-center associated nuclear protein		
HCFC1_HUMAN	Host cell factor 1		
HIRA_HUMAN	Protein HIRA		
JHD2C_HUMAN	Probable JmjC domain-containing histone demethylation protein 2C		
KDM1A_HUMAN	Lysine-specific histone demethylase 1A		
KDM3B_HUMAN	Lysine-specific demethylase 3B		
KMT2A_HUMAN	Histone-lysine N-methyltransferase 2A		
KMT2D_HUMAN	Histone-lysine N-methyltransferase 2D		
NCOA6_HUMAN	Nuclear receptor coactivator 6		
NCOR1_HUMAN	Nuclear receptor corepressor 1		
NCOR2_HUMAN	Nuclear receptor corepressor 2		
NU153_HUMAN	Nuclear pore complex protein Nup153		
NU214_HUMAN	Nuclear pore complex protein Nup214		
NUP54_HUMAN	Nucleoporin p54		
NUP58_HUMAN	Nucleoporin p58/p45		
NUP62_HUMAN	Nuclear pore glycoprotein p62		
NUP88_HUMAN	Nuclear pore complex protein Nup88		
OGT1_HUMAN	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit		
P121C_HUMAN	Nuclear envelope pore membrane protein POM 121C		
PF21A_HUMAN	PHD finger protein 21A		
PRC2A_HUMAN	Protein PRRC2A		
PRC2B_HUMAN	Protein PRRC2B		
PRC2C_HUMAN	Protein PRRC2C		
PRR12_HUMAN	Proline-rich protein 12		
QSER1_HUMAN	Glutamine and serine-rich protein 1		
RAGP1_HUMAN	Ran GTPase-activating protein 1		
RBM27_HUMAN	RNA-binding protein 27		

RBP2_HUMAN	E3 SUMO-protein ligase RanBP2		
RPRD2_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 2		
SET1A_HUMAN	Histone-lysine N-methyltransferase SETD1A		
SIN3A_HUMAN	Paired amphipathic helix protein Sin3a		
TBL1R_HUMAN	F-box-like/WD repeat-containing protein TBL1XR1		
TBL1X_HUMAN	F-box-like/WD repeat-containing protein TBL1X		
TET2_HUMAN	Methylcytosine dioxygenase TET2		
TNR6A_HUMAN	Trinucleotide repeat-containing gene 6A protein		
TRRAP_HUMAN	Transformation/transcription domain-associated protein		
UBAP2_HUMAN	Ubiquitin-associated protein 2		
UBN2_HUMAN	Ubinuclein-2		
UBP2L_HUMAN	Ubiquitin-associated protein 2-like		
WNK1_HUMAN	Serine/threonine-protein kinase WNK1		
YLPM1_HUMAN	YLP motif-containing protein 1		
ZC3H1_HUMAN	Zinc finger C3H1 domain-containing protein		
ZC3HE_HUMAN	Zinc finger CCCH domain-containing protein 14		
ZN281_HUMAN	Zinc finger protein 281		
ZN318_HUMAN	Zinc finger protein 318		
ZN609_HUMAN	Zinc finger protein 609		
ZHX3_HUMAN	Zinc fingers and homeoboxes protein 3		
RBM26_HUMAN	RNA-binding protein 26		
ZHX2_HUMAN	Zinc fingers and homeoboxes protein 2		
TRPS1_HUMAN	Zinc finger transcription factor Trps1		
OGA_HUMAN	Protein O-GlcNAcase		
KMT2C_HUMAN	Histone-lysine N-methyltransferase 2C		
CEP85_HUMAN	Centrosomal protein of 85 kDa		
ARI5B_HUMAN	AT-rich interactive domain-containing protein 5B		
RGPD8_HUMAN	RANBP2-like and GRIP domain-containing protein 8		
RGPD5_HUMAN	RANBP2-like and GRIP domain-containing protein 5/6		
ZFR_HUMAN	Zinc finger RNA-binding protein		
RGPD3_HUMAN	RanBP2-like and GRIP domain-containing protein 3		
BAP1_HUMAN	Ubiquitin carboxyl-terminal hydrolase BAP1		
ZEP1_HUMAN	Zinc finger protein 40		
PHC3_HUMAN	Polyhomeotic-like protein 3		
CABIN_HUMAN	Calcineurin-binding protein cabin-1		
MEF2D_HUMAN	Myocyte-specific enhancer factor 2D		
PCF11_HUMAN	Pre-mRNA cleavage complex 2 protein Pcf11		
TAF4_HUMAN	Transcription initiation factor TFIID subunit 4		
LIN54_HUMAN	Protein lin-54 homolog		
RESF1_HUMAN	Retroelement silencing factor 1		
TASO2_HUMAN	Protein TASOR 2		
DOT1L_HUMAN	Histone-lysine N-methyltransferase, H3 lysine-79 specific		
ASXL2_HUMAN	Putative Polycomb group protein ASXL2		
RBM33_HUMAN	RNA-binding protein 33		
SMRC1_HUMAN	SWI/SNF complex subunit SMARCC1		
ICE1_HUMAN	Little elongation complex subunit 1		

CBP_HUMAN	CREB-binding protein		
SMRC2_HUMAN	SWI/SNF complex subunit SMARCC2		
LRIF1_HUMAN	Ligand-dependent nuclear receptor-interacting factor 1		
KDM2B_HUMAN	Lysine-specific demethylase 2B		
NCOA2_HUMAN	Nuclear receptor coactivator 2		
UBN1_HUMAN	Ubinuclein-1		
ZHX1_HUMAN	Zinc fingers and homeoboxes protein 1		
ZFHX3_HUMAN	Zinc finger homeobox protein 3		
EP300_HUMAN	Histone acetyltransferase p300		
SF3A1_HUMAN	Splicing factor 3A subunit 1		
ACL6A_HUMAN	Actin-like protein 6A		
TAF6_HUMAN	Transcription initiation factor TFIID subunit 6		
POGZ_HUMAN	Pogo transposable element with ZNF domain		
AHDC1_HUMAN	AT-hook DNA-binding motif-containing protein 1		
I2BP2_HUMAN	Interferon regulatory factor 2-binding protein 2		
RREB1_HUMAN	Ras-responsive element-binding protein 1		
PHF12_HUMAN	PHD finger protein 12		
NUP50_HUMAN	Nuclear pore complex protein Nup50		
TNR6B_HUMAN	Trinucleotide repeat-containing gene 6B protein		
PR40A_HUMAN	Pre-mRNA-processing factor 40 homolog A		
MINT_HUMAN	Msx2-interacting protein		
EP400_HUMAN	E1A-binding protein p400		
ILF2_HUMAN	Interleukin enhancer-binding factor 2		
SMCA4_HUMAN	Transcription activator BRG1		
MTREX_HUMAN	Exosome RNA helicase MTR4		
TAF5_HUMAN	Transcription initiation factor TFIID subunit 5		
RBM25_HUMAN	RNA-binding protein 25		
CPSF1_HUMAN	Cleavage and polyadenylation specificity factor subunit 1		
HNRPL_HUMAN	Heterogeneous nuclear ribonucleoprotein L		
ZN106_HUMAN	Zinc finger protein 106		
KNL1_HUMAN	Kinetochore scaffold 1		
SC24B_HUMAN	Protein transport protein Sec24B		
P121A_HUMAN	Nuclear envelope pore membrane protein POM 121		
PABP4_HUMAN	Polyadenylate-binding protein 4		
SF01_HUMAN	Splicing factor 1		
NUP98_HUMAN	Nuclear pore complex protein Nup98-Nup96		
PHX2B_HUMAN	Paired mesoderm homeobox protein 2B		
CARM1_HUMAN	Histone-arginine methyltransferase CARM1		
SC16A_HUMAN	Protein transport protein Sec16A		
TNR6C_HUMAN	Trinucleotide repeat-containing gene 6C protein		
GSE1_HUMAN	Genetic suppressor element 1		
HDAC3_HUMAN	Histone deacetylase 3		
SC23B_HUMAN	Protein transport protein Sec23B		
SP130_HUMAN	Histone deacetylase complex subunit SAP130		
TLE1_HUMAN	Transducin-like enhancer protein 1		
CPSF7_HUMAN	Cleavage and polyadenylation specificity factor subunit 7		

ZMYM4_HUMAN	Zinc finger MYM-type protein 4		
RCOR3_HUMAN	REST corepressor 3		
ITF2_HUMAN	Transcription factor 4		
KANL3_HUMAN	KAT8 regulatory NSL complex subunit 3		
PSPC1_HUMAN	Paraspeckle component 1		
RING1_HUMAN	E3 ubiquitin-protein ligase RING1		
CD2B2_HUMAN	CD2 antigen cytoplasmic tail-binding protein 2		
ATN1_HUMAN	Atrophin-1		
ASH2L_HUMAN	Set1/Ash2 histone methyltransferase complex subunit ASH2		
PRP4_HUMAN	U4/U6 small nuclear ribonucleoprotein Prp4		
RCOR1_HUMAN	REST corepressor 1		
MYT1_HUMAN	Myelin transcription factor 1		
WDR5_HUMAN	WD repeat-containing protein 5		
PP1A_HUMAN	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit		
CBX3_HUMAN	Chromobox protein homolog 3		
MEX3A_HUMAN	RNA-binding protein MEX3A		
SYNP2_HUMAN	Synaptopodin-2		
DCAF7_HUMAN	DDB1- and CUL4-associated factor 7		
RBBP7_HUMAN	Histone-binding protein RBBP7		
RBBP5_HUMAN	Retinoblastoma-binding protein 5		
CSK21_HUMAN	Casein kinase II subunit alpha		
IMPA1_HUMAN	Inositol monophosphatase 1		
SSBP3_HUMAN	Single-stranded DNA-binding protein 3		
DDX6_HUMAN	Probable ATP-dependent RNA helicase DDX6		
MBD3_HUMAN	Methyl-CpG-binding domain protein 3		
LS14B_HUMAN	Protein LSM14 homolog B		
HNRPK_HUMAN	Heterogeneous nuclear ribonucleoprotein K		
RAE1L_HUMAN	mRNA export factor		
REQU_HUMAN	Zinc finger protein ubi-d4		
NHSL2_HUMAN	NHS-like protein 2		
ZN148_HUMAN	Zinc finger protein 148		
RUXF_HUMAN	Small nuclear ribonucleoprotein F		
YTHD2_HUMAN	YTH domain-containing family protein 2		
SPF45_HUMAN	Splicing factor 45		
RSMB_HUMAN	Small nuclear ribonucleoprotein-associated proteins B and B'		
F193A_HUMAN	Protein FAM193A		
IMB1_HUMAN	Importin subunit beta-1		
HTF4_HUMAN	Transcription factor 12		
PHX2A_HUMAN	Paired mesoderm homeobox protein 2A		
PABP1_HUMAN	Polyadenylate-binding protein 1		
HDAC1_HUMAN	Histone deacetylase 1		
MGAP_HUMAN	MAX gene-associated protein		
SR140_HUMAN	U2 snRNP-associated SURP motif-containing protein		
BAP18_HUMAN	Chromatin complexes subunit BAP18		
IMA7_HUMAN	Importin subunit alpha-7		
PCNA_HUMAN	Proliferating cell nuclear antigen		

MTA2_HUMAN	Metastasis-associated protein MTA2		
SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich		
DDX23_HUMAN	Probable ATP-dependent RNA helicase DDX23		
PRP8_HUMAN	Pre-mRNA-processing-splicing factor 8		
SNR40_HUMAN	U5 small nuclear ribonucleoprotein 40 kDa protein		
HDAC2_HUMAN	Histone deacetylase 2		
SF3B2_HUMAN	Splicing factor 3B subunit 2		
CHERP_HUMAN	Calcium homeostasis endoplasmic reticulum protein		
U5S1_HUMAN	116 kDa U5 small nuclear ribonucleoprotein component		
AGFG1_HUMAN	Arf-GAP domain and FG repeat-containing protein 1		
U520_HUMAN	U5 small nuclear ribonucleoprotein 200 kDa helicase		
RBM14_HUMAN	RNA-binding protein 14		
TCRG1_HUMAN	Transcription elongation regulator 1		
PRP19_HUMAN	Pre-mRNA-processing factor 19		
BUB3_HUMAN	Mitotic checkpoint protein BUB3		
TBG1_HUMAN	Tubulin gamma-1 chain		
CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4		
CPSF5_HUMAN	Cleavage and polyadenylation specificity factor subunit 5		

Table 4.4 TPR-TurboID interacting proteins under different serum conditions in SHSY5Y cells

Dark gray - identified only in TPR-TurboID, not present in eGFP-TurboID
Light gray - identified in both TPR-TurboID and eGFP-TurboID, significantly enriched in TPR-TurboID
Dark blue - identified only in the indicated TPR-TurboID condition
Light blue - identified in both TPR-TurboID conditions, significantly enriched in indicated TPR-TurboID

		Serum			Starve			Pulse		
Accession	**Description**	Over eGFP	Over Starve	Over Pulse	Over eGFP	Over Serum	Over Pulse	Over eGFP	Over Serum	Over Starve
HNRH3_HUMAN	Heterogeneous nuclear ribonucleoprotein H3									
IF4G3_HUMAN	Eukaryotic translation initiation factor 4 gamma 3									
UBP16_HUMAN	Ubiquitin carboxyl-terminal hydrolase 16									
PLRG1_HUMAN	Pleiotropic regulator 1									
SURF4_HUMAN	Surfeit locus protein 4									
KANL3_HUMAN	KAT8 regulatory NSL complex subunit 3									
EFNMT_HUMAN	eEF1A lysine and N-terminal methyltransferase									
ZFR_HUMAN	Zinc finger RNA-binding protein									
CARF_HUMAN	CDKN2A-interacting protein									
CDK9_HUMAN	Cyclin-dependent kinase 9									
PRPS2_HUMAN	Ribose-phosphate pyrophosphokinase 2									
SYFA_HUMAN	Phenylalanine--tRNA ligase alpha subunit									
IPO9_HUMAN	Importin-9									
HDAC3_HUMAN	Histone deacetylase 3									
MINT_HUMAN	Msx2-interacting protein									
OGT1_HUMAN	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit									
KDM1A_HUMAN	Lysine-specific histone demethylase 1A									
OGA_HUMAN	Protein O-GlcNAcase									
ILF2_HUMAN	Interleukin enhancer-binding factor 2									
SF01_HUMAN	Splicing factor 1									

ZMYM4_HUMAN	Zinc finger MYM-type protein 4									
ZC3HE_HUMAN	Zinc finger CCCH domain-containing protein 14									
NUP54_HUMAN	Nucleoporin p54									
CPSF7_HUMAN	Cleavage and polyadenylation specificity factor subunit 7									
ARI1B_HUMAN	AT-rich interactive domain-containing protein 1B									
DIDO1_HUMAN	Death-inducer obliterator 1									
RBM27_HUMAN	RNA-binding protein 27									
P121C_HUMAN	Nuclear envelope pore membrane protein POM 121C									
QSER1_HUMAN	Glutamine and serine-rich protein 1									
RBM26_HUMAN	RNA-binding protein 26									
CCAR1_HUMAN	Cell division cycle and apoptosis regulator protein 1									
WNK1_HUMAN	Serine/threonine-protein kinase WNK1									
ANR17_HUMAN	Ankyrin repeat domain-containing protein 17									
RAE1L_HUMAN	mRNA export factor									
TFG_HUMAN	Protein TFG									
WDR5_HUMAN	WD repeat-containing protein 5									
CNOT1_HUMAN	CCR4-NOT transcription complex subunit 1									
YLPM1_HUMAN	YLP motif-containing protein 1									
SF3A1_HUMAN	Splicing factor 3A subunit 1									
CARM1_HUMAN	Histone-arginine methyltransferase CARM1									
NUP58_HUMAN	Nucleoporin p58/p45									
SP130_HUMAN	Histone deacetylase complex subunit SAP130									
CUX1_HUMAN	Homeobox protein cut-like 1									
TLE3_HUMAN	Transducin-like enhancer protein 3									
CBX3_HUMAN	Chromobox protein homolog 3									
POGZ_HUMAN	Pogo transposable element with ZNF domain									
REQU_HUMAN	Zinc finger protein ubi-d4									
PTN23_HUMAN	Tyrosine-protein phosphatase non-receptor type 23									
YTHD2_HUMAN	YTH domain-containing family protein 2									
CDIPT_HUMAN	CDP-diacylglycerol--inositol 3-phosphatidyltransferase									
RT05_HUMAN	28S ribosomal protein S5, mitochondrial									
IMPA1_HUMAN	Inositol monophosphatase 1									
NU188_HUMAN	Nucleoporin NUP188									

GGYF2_HUMAN	GRB10-interacting GYF protein 2									
BCL9_HUMAN	B-cell CLL/lymphoma 9 protein									
FUBP2_HUMAN	Far upstream element-binding protein 2									
TBL1R_HUMAN	F-box-like/WD repeat-containing protein TBL1XR1									
NUP62_HUMAN	Nuclear pore glycoprotein p62									
ATX2L_HUMAN	Ataxin-2-like protein									
MBD3_HUMAN	Methyl-CpG-binding domain protein 3									
NUP98_HUMAN	Nuclear pore complex protein Nup98-Nup96									
PHX2A_HUMAN	Paired mesoderm homeobox protein 2A									
UBP2L_HUMAN	Ubiquitin-associated protein 2-like									
PHX2B_HUMAN	Paired mesoderm homeobox protein 2B									
SIN3A_HUMAN	Paired amphipathic helix protein Sin3a									
NU153_HUMAN	Nuclear pore complex protein Nup153									
KMT2A_HUMAN	Histone-lysine N-methyltransferase 2A									
RAGP1_HUMAN	Ran GTPase-activating protein 1									
HNRPK_HUMAN	Heterogeneous nuclear ribonucleoprotein K									
ZN609_HUMAN	Zinc finger protein 609									
ARI1A_HUMAN	AT-rich interactive domain-containing protein 1A									
CPSF5_HUMAN	Cleavage and polyadenylation specificity factor subunit 5									
U520_HUMAN	U5 small nuclear ribonucleoprotein 200 kDa helicase									
PABP1_HUMAN	Polyadenylate-binding protein 1									
NU214_HUMAN	Nuclear pore complex protein Nup214									
RBP2_HUMAN	E3 SUMO-protein ligase RanBP2									
HCFC1_HUMAN	Host cell factor 1									
RHGA_ASPAC	TurboID TurboID with 2X FLAG and 3X GGGGS linker									
CPSF1_HUMAN	Cleavage and polyadenylation specificity factor subunit 1									
PABP4_HUMAN	Polyadenylate-binding protein 4									
MTA1_HUMAN	Metastasis-associated protein MTA1									
IMB1_HUMAN	Importin subunit beta-1									
U5S1_HUMAN	116 kDa U5 small nuclear ribonucleoprotein component									
CPSF6_HUMAN	Cleavage and polyadenylation specificity factor subunit 6									
PRP8_HUMAN	Pre-mRNA-processing-splicing factor 8									
CCAR2_HUMAN	Cell cycle and apoptosis regulator protein 2									

RBM14_HUMAN	RNA-binding protein 14									
NUP88_HUMAN	Nuclear pore complex protein Nup88									
SRRM2_HUMAN	Serine/arginine repetitive matrix protein 2									
HNRPM_HUMAN	Heterogeneous nuclear ribonucleoprotein M									
PRC2C_HUMAN	Protein PRRC2C									
NFIB_HUMAN	Nuclear factor 1 B-type									
SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich									
SNR40_HUMAN	U5 small nuclear ribonucleoprotein 40 kDa protein									
TIF1B_HUMAN	Transcription intermediary factor 1-beta									
HNRL1_HUMAN	Heterogeneous nuclear ribonucleoprotein U-like protein 1									
CHAP1_HUMAN	Chromosome alignment-maintaining phosphoprotein 1									
BUB3_HUMAN	Mitotic checkpoint protein BUB3									
SC16A_HUMAN	Protein transport protein Sec16A									
ROA1_HUMAN	Heterogeneous nuclear ribonucleoprotein A1									
BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF									
EP400_HUMAN	E1A-binding protein p400									
RUVB2_HUMAN	RuvB-like 2									
HNRPC_HUMAN	Heterogeneous nuclear ribonucleoproteins C1/C2									
SMD3_HUMAN	Small nuclear ribonucleoprotein Sm D3									
SNF5_HUMAN	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily B member 1									
HDAC2_HUMAN	Histone deacetylase 2									
PRP19_HUMAN	Pre-mRNA-processing factor 19									
CHD8_HUMAN	Chromodomain-helicase-DNA-binding protein 8									
TAF9_HUMAN	Transcription initiation factor TFIID subunit 9									
LDHA_HUMAN	L-lactate dehydrogenase A chain									
RBBP4_HUMAN	Histone-binding protein RBBP4									
PRP4_HUMAN	U4/U6 small nuclear ribonucleoprotein Prp4									
CDC5L_HUMAN	Cell division cycle 5-like protein									
PCH2_HUMAN	Pachytene checkpoint protein 2 homolog									
RPRD2_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 2									
PF21A_HUMAN	PHD finger protein 21A									
SC24B_HUMAN	Protein transport protein Sec24B									
RING1_HUMAN	E3 ubiquitin-protein ligase RING1									

PRR12_HUMAN	Proline-rich protein 12									
BAP18_HUMAN	Chromatin complexes subunit BAP18									
SET1A_HUMAN	Histone-lysine N-methyltransferase SETD1A									
TCRG1_HUMAN	Transcription elongation regulator 1									
SNW1_HUMAN	SNW domain-containing protein 1									
RING2_HUMAN	E3 ubiquitin-protein ligase RING2									
ZMYM2_HUMAN	Zinc finger MYM-type protein 2									
AP2B_HUMAN	Transcription factor AP-2-beta									
TRI33_HUMAN	E3 ubiquitin-protein ligase TRIM33									
RBBP5_HUMAN	Retinoblastoma-binding protein 5									
ROA3_HUMAN	Heterogeneous nuclear ribonucleoprotein A3									
ZN281_HUMAN	Zinc finger protein 281									
KDM3B_HUMAN	Lysine-specific demethylase 3B									
RALY_HUMAN	RNA-binding protein Raly									
HDAC1_HUMAN	Histone deacetylase 1									
PRC2A_HUMAN	Protein PRRC2A									
SPF45_HUMAN	Splicing factor 45									
ROA2_HUMAN	Heterogeneous nuclear ribonucleoproteins A2/B1									
HTF4_HUMAN	Transcription factor 12									
RUVB1_HUMAN	RuvB-like 1									
KPCA_HUMAN	Protein kinase C alpha type									
MATR3_HUMAN	Matrin-3									
ZC11A_HUMAN	Zinc finger CCCH domain-containing protein 11A									
HNRPL_HUMAN	Heterogeneous nuclear ribonucleoprotein L									
NFIX_HUMAN	Nuclear factor 1 X-type									
NIPBL_HUMAN	Nipped-B-like protein									
PSPC1_HUMAN	Paraspeckle component 1									
ASH2L_HUMAN	Set1/Ash2 histone methyltransferase complex subunit ASH2									
TRRAP_HUMAN	Transformation/transcription domain-associated protein									
NCOR2_HUMAN	Nuclear receptor corepressor 2									
NCOR1_HUMAN	Nuclear receptor corepressor 1									
PP1A_HUMAN	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit									
ZFHX3_HUMAN	Zinc finger homeobox protein 3									

ANKH1_HUMAN	Ankyrin repeat and KH domain-containing protein 1									
H2A1A_HUMAN	Histone H2A type 1-A									
RCOR3_HUMAN	REST corepressor 3									
PUM1_HUMAN	Pumilio homolog 1									
NDUAA_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10, mitochondrial									
PATL1_HUMAN	Protein PAT1 homolog 1									
RAB2B_HUMAN	Ras-related protein Rab-2B									
RAB2A_HUMAN	Ras-related protein Rab-2A									
CFA20_HUMAN	Cilia- and flagella-associated protein 20									
NHSL2_HUMAN	NHS-like protein 2									
XPO1_HUMAN	Exportin-1									
SR140_HUMAN	U2 snRNP-associated SURP motif-containing protein									
SMC2_HUMAN	Structural maintenance of chromosomes protein 2									
PCNA_HUMAN	Proliferating cell nuclear antigen									
RL22_HUMAN	60S ribosomal protein L22									
GARS_HUMAN	Glycine--tRNA ligase									
RL27A_HUMAN	60S ribosomal protein L27a									
CAPZB_HUMAN	F-actin-capping protein subunit beta									
CTBP2_HUMAN	C-terminal-binding protein 2									
RAB14_HUMAN	Ras-related protein Rab-14									
PPIG_HUMAN	Peptidyl-prolyl cis-trans isomerase G									
EPN4_HUMAN	Clathrin interactor 1									
PCBP2_HUMAN	Poly(rC)-binding protein 2									
TAF3_HUMAN	Transcription initiation factor TFIID subunit 3									
DHCR7_HUMAN	7-dehydrocholesterol reductase									
RAB1A_HUMAN	Ras-related protein Rab-1A									

Table 4.5 TPR-TurboID interacting proteins under depolarization of SHSY5Y cells

Dark gray - identified only in TPR-TurboID, not present in eGFP-TurboID
Light gray - identified in both TPR-TurboID and eGFP-TurboID, significantly enriched in TPR-TurboID
Dark blue - identified only in the indicated TPR-TurboID condition
Light blue - identified in both TPR-TurboID conditions, significantly enriched in indicated TPR-TurboID

		Label		Depolarize	
Accession	**Description**	Over eGFP	Over Depolarize	Over eGFP	Over Label
IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 3				
CDS2_HUMAN	Phosphatidate cytidylyltransferase 2				
G3P_HUMAN	Glyceraldehyde-3-phosphate dehydrogenase				
DIM1_HUMAN	Probable dimethyladenosine transferase				
PCLO_HUMAN	Protein piccolo				
CDK9_HUMAN	Cyclin-dependent kinase 9				
RTCB_HUMAN	RNA-splicing ligase RtcB homolog				
CHD8_HUMAN	Chromodomain-helicase-DNA-binding protein 8				
XRN2_HUMAN	5'-3' exoribonuclease 2				
E2AK2_HUMAN	Interferon-induced, double-stranded RNA-activated protein kinase				
NXF1_HUMAN	Nuclear RNA export factor 1				
GANP_HUMAN	Germinal-center associated nuclear protein				
MBD3_HUMAN	Methyl-CpG-binding domain protein 3				
QSER1_HUMAN	Glutamine and serine-rich protein 1				
LS14B_HUMAN	Protein LSM14 homolog B				
MEX3A_HUMAN	RNA-binding protein MEX3A				
P121C_HUMAN	Nuclear envelope pore membrane protein POM 121C				
BCL9_HUMAN	B-cell CLL/lymphoma 9 protein				
KMT2D_HUMAN	Histone-lysine N-methyltransferase 2D				
SET1A_HUMAN	Histone-lysine N-methyltransferase SETD1A				
OGT1_HUMAN	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit				
HDAC3_HUMAN	Histone deacetylase 3				
ZC3H1_HUMAN	Zinc finger C3H1 domain-containing protein				
KDM1A_HUMAN	Lysine-specific histone demethylase 1A				
TBL1X_HUMAN	F-box-like/WD repeat-containing protein TBL1X				
NCOR1_HUMAN	Nuclear receptor corepressor 1				
CD2B2_HUMAN	CD2 antigen cytoplasmic tail-binding protein 2				
SC24B_HUMAN	Protein transport protein Sec24B				
ITF2_HUMAN	Transcription factor 4				
DDX6_HUMAN	Probable ATP-dependent RNA helicase DDX6				
YLPM1_HUMAN	YLP motif-containing protein 1				
NUP98_HUMAN	Nuclear pore complex protein Nup98-Nup96				

HIRA_HUMAN	Protein HIRA				
ATN1_HUMAN	Atrophin-1				
DCAF7_HUMAN	DDB1- and CUL4-associated factor 7				
WDR5_HUMAN	WD repeat-containing protein 5				
PP1A_HUMAN	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit				
CSK21_HUMAN	Casein kinase II subunit alpha				
MYT1_HUMAN	Myelin transcription factor 1				
TLE1_HUMAN	Transducin-like enhancer protein 1				
RING1_HUMAN	E3 ubiquitin-protein ligase RING1				
CBX3_HUMAN	Chromobox protein homolog 3				
GSE1_HUMAN	Genetic suppressor element 1				
RBBP5_HUMAN	Retinoblastoma-binding protein 5				
SC23B_HUMAN	Protein transport protein Sec23B				
SF01_HUMAN	Splicing factor 1				
JHD2C_HUMAN	Probable JmjC domain-containing histone demethylation protein 2C				
UBAP2_HUMAN	Ubiquitin-associated protein 2				
ZN318_HUMAN	Zinc finger protein 318				
ZMYM4_HUMAN	Zinc finger MYM-type protein 4				
ZC3HE_HUMAN	Zinc finger CCCH domain-containing protein 14				
BCOR_HUMAN	BCL-6 corepressor				
NUP54_HUMAN	Nucleoporin p54				
EMSY_HUMAN	BRCA2-interacting transcriptional repressor EMSY				
CARM1_HUMAN	Histone-arginine methyltransferase CARM1				
CCAR1_HUMAN	Cell division cycle and apoptosis regulator protein 1				
CPSF7_HUMAN	Cleavage and polyadenylation specificity factor subunit 7				
PSPC1_HUMAN	Paraspeckle component 1				
PF21A_HUMAN	PHD finger protein 21A				
CIC_HUMAN	Protein capicua homolog				
NUP58_HUMAN	Nucleoporin p58/p45				
SSBP3_HUMAN	Single-stranded DNA-binding protein 3				
SP130_HUMAN	Histone deacetylase complex subunit SAP130				
TNR6C_HUMAN	Trinucleotide repeat-containing gene 6C protein				
RCOR3_HUMAN	REST corepressor 3				
RBM27_HUMAN	RNA-binding protein 27				
KANL3_HUMAN	KAT8 regulatory NSL complex subunit 3				
ASH2L_HUMAN	Set1/Ash2 histone methyltransferase complex subunit ASH2				
RCOR1_HUMAN	REST corepressor 1				
PRR12_HUMAN	Proline-rich protein 12				
NCOR2_HUMAN	Nuclear receptor corepressor 2				
PABP4_HUMAN	Polyadenylate-binding protein 4				
PRC2B_HUMAN	Protein PRRC2B				
RPRD2_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 2				
KDM3B_HUMAN	Lysine-specific demethylase 3B				
ANKH1_HUMAN	Ankyrin repeat and KH domain-containing protein 1				
DIDO1_HUMAN	Death-inducer obliterator 1				
TBL1R_HUMAN	F-box-like/WD repeat-containing protein TBL1XR1				
ZN281_HUMAN	Zinc finger protein 281				
SC16A_HUMAN	Protein transport protein Sec16A				
PRP4_HUMAN	U4/U6 small nuclear ribonucleoprotein Prp4				
PHX2B_HUMAN	Paired mesoderm homeobox protein 2B				
UBN2_HUMAN	Ubinuclein-2				
TET2_HUMAN	Methylcytosine dioxygenase TET2				

IMPA1_HUMAN	Inositol monophosphatase 1				
RBBP7_HUMAN	Histone-binding protein RBBP7				
TNR6A_HUMAN	Trinucleotide repeat-containing gene 6A protein				
SYNP2_HUMAN	Synaptopodin-2				
NCOA6_HUMAN	Nuclear receptor coactivator 6				
P121A_HUMAN	Nuclear envelope pore membrane protein POM121				
SPF45_HUMAN	Splicing factor 45				
WNK1_HUMAN	Serine/threonine-protein kinase WNK1				
ZN148_HUMAN	Zinc finger protein 148				
TRRAP_HUMAN	Transformation/transcription domain-associated protein				
YTHD2_HUMAN	YTH domain-containing family protein 2				
NU153_HUMAN	Nuclear pore complex protein Nup153				
RAE1L_HUMAN	mRNA export factor				
NHSL2_HUMAN	NHS-like protein 2				
RAGP1_HUMAN	Ran GTPase-activating protein 1				
HNRPK_HUMAN	Heterogeneous nuclear ribonucleoprotein K				
KMT2A_HUMAN	Histone-lysine N-methyltransferase 2A				
RUXF_HUMAN	Small nuclear ribonucleoprotein F				
REQU_HUMAN	Zinc finger protein ubi-d4				
RSMB_HUMAN	Small nuclear ribonucleoprotein-associated proteins B and B'				
F193A_HUMAN	Protein FAM193A				
PHX2A_HUMAN	Paired mesoderm homeobox protein 2A				
SR140_HUMAN	U2 snRNP-associated SURP motif-containing protein				
ANR17_HUMAN	Ankyrin repeat domain-containing protein 17				
PRC2A_HUMAN	Protein PRRC2A				
BAP18_HUMAN	Chromatin complexes subunit BAP18				
PRC2C_HUMAN	Protein PRRC2C				
HCFC1_HUMAN	Host cell factor 1				
IMB1_HUMAN	Importin subunit beta-1				
MGAP_HUMAN	MAX gene-associated protein				
HTF4_HUMAN	Transcription factor 12				
NUP88_HUMAN	Nuclear pore complex protein Nup88				
HDAC1_HUMAN	Histone deacetylase 1				
ZN609_HUMAN	Zinc finger protein 609				
PABP1_HUMAN	Polyadenylate-binding protein 1				
UBP2L_HUMAN	Ubiquitin-associated protein 2-like				
ARI1B_HUMAN	AT-rich interactive domain-containing protein 1B				
PCNA_HUMAN	Proliferating cell nuclear antigen				
IMA7_HUMAN	Importin subunit alpha-7				
MTA2_HUMAN	Metastasis-associated protein MTA2				
SNR40_HUMAN	U5 small nuclear ribonucleoprotein 40 kDa protein				
SIN3A_HUMAN	Paired amphipathic helix protein Sin3a				
BUB3_HUMAN	Mitotic checkpoint protein BUB3				
ARI1A_HUMAN	AT-rich interactive domain-containing protein 1A				
CPSF5_HUMAN	Cleavage and polyadenylation specificity factor subunit 5				
U520_HUMAN	U5 small nuclear ribonucleoprotein 200 kDa helicase				
SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich				
NU214_HUMAN	Nuclear pore complex protein Nup214				
NUP62_HUMAN	Nuclear pore glycoprotein p62				
RBP2_HUMAN	E3 SUMO-protein ligase RanBP2				
AGFG1_HUMAN	Arf-GAP domain and FG repeat-containing protein 1				
RHGA_ASPAC	TurboID TurboID with 2X FLAG and 3X GGGGS linker				

BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF				
CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4				
U5S1_HUMAN	116 kDa U5 small nuclear ribonucleoprotein component				
PRP8_HUMAN	Pre-mRNA-processing-splicing factor 8				
HDAC2_HUMAN	Histone deacetylase 2				
RBM14_HUMAN	RNA-binding protein 14				
DDX23_HUMAN	Probable ATP-dependent RNA helicase DDX23				
TBG1_HUMAN	Tubulin gamma-1 chain				
TCRG1_HUMAN	Transcription elongation regulator 1				
SF3B2_HUMAN	Splicing factor 3B subunit 2				
CHERP_HUMAN	Calcium homeostasis endoplasmic reticulum protein				
PRP19_HUMAN	Pre-mRNA-processing factor 19				
RBM33_HUMAN	RNA-binding protein 33				
RPB1_HUMAN	DNA-directed RNA polymerase II subunit RPB1				
TLE3_HUMAN	Transducin-like enhancer protein 3				
SNW1_HUMAN	SNW domain-containing protein 1				
VEZF1_HUMAN	Vascular endothelial zinc finger 1				
AHDC1_HUMAN	AT-hook DNA-binding motif-containing protein 1				
POGZ_HUMAN	Pogo transposable element with ZNF domain				
YTHD3_HUMAN	YTH domain-containing family protein 3				
KMT2C_HUMAN	Histone-lysine N-methyltransferase 2C				
ZFR_HUMAN	Zinc finger RNA-binding protein				
MINT_HUMAN	Msx2-interacting protein				
ATX2_HUMAN	Ataxin-2				
YTHD1_HUMAN	YTH domain-containing family protein 1				
RPR1B_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 1B				
NUP50_HUMAN	Nuclear pore complex protein Nup50				
CTBP2_HUMAN	C-terminal-binding protein 2				
SCAF4_HUMAN	SR-related and CTD-associated factor 4				
RBM10_HUMAN	RNA-binding protein 10				
NIPBL_HUMAN	Nipped-B-like protein				
ATX2L_HUMAN	Ataxin-2-like protein				
TFG_HUMAN	Protein TFG				
CHAP1_HUMAN	Chromosome alignment-maintaining phosphoprotein 1				
NU188_HUMAN	Nucleoporin NUP188				
AKAP8_HUMAN	A-kinase anchor protein 8				
HNRPF_HUMAN	Heterogeneous nuclear ribonucleoprotein F				
SMD3_HUMAN	Small nuclear ribonucleoprotein Sm D3				
EP400_HUMAN	E1A-binding protein p400				
CYFP1_HUMAN	Cytoplasmic FMR1-interacting protein 1				
CDC5L_HUMAN	Cell division cycle 5-like protein				
HNRPM_HUMAN	Heterogeneous nuclear ribonucleoprotein M				
RBBP4_HUMAN	Histone-binding protein RBBP4				
CPSF1_HUMAN	Cleavage and polyadenylation specificity factor subunit 1				
TADBP_HUMAN	TAR DNA-binding protein 43				
NONO_HUMAN	Non-POU domain-containing octamer-binding protein				
NUFP2_HUMAN	Nuclear fragile X mental retardation-interacting protein 2				
OGA_HUMAN	Protein O-GlcNAcase				
KDM6A_HUMAN	Lysine-specific demethylase 6A				
MTREX_HUMAN	Exosome RNA helicase MTR4				
HELZ_HUMAN	Probable helicase with zinc finger domain				
KDM2B_HUMAN	Lysine-specific demethylase 2B				

SMG7_HUMAN	Protein SMG7				
CARF_HUMAN	CDKN2A-interacting protein				
RUXE_HUMAN	Small nuclear ribonucleoprotein E				
CBX1_HUMAN	Chromobox protein homolog 1				
BAP1_HUMAN	Ubiquitin carboxyl-terminal hydrolase BAP1				
DMAP1_HUMAN	DNA methyltransferase 1-associated protein 1				
NFIB_HUMAN	Nuclear factor 1 B-type				
GPS2_HUMAN	G protein pathway suppressor 2				
NCOA2_HUMAN	Nuclear receptor coactivator 2				
RBM26_HUMAN	RNA-binding protein 26				
PR40B_HUMAN	Pre-mRNA-processing factor 40 homolog B				
GGYF2_HUMAN	GRB10-interacting GYF protein 2				
I2BP2_HUMAN	Interferon regulatory factor 2-binding protein 2				
XRN1_HUMAN	5'-3' exoribonuclease 1				
DDX1_HUMAN	ATP-dependent RNA helicase DDX1				
RING2_HUMAN	E3 ubiquitin-protein ligase RING2				
HNRL1_HUMAN	Heterogeneous nuclear ribonucleoprotein U-like protein 1				
PTN23_HUMAN	Tyrosine-protein phosphatase non-receptor type 23				
TNR6B_HUMAN	Trinucleotide repeat-containing gene 6B protein				
S23IP_HUMAN	SEC23-interacting protein				
GARS_HUMAN	Glycine--tRNA ligase				
GCP3_HUMAN	Gamma-tubulin complex component 3				
MYEF2_HUMAN	Myelin expression factor 2				
SF3B4_HUMAN	Splicing factor 3B subunit 4				
PP1RA_HUMAN	Serine/threonine-protein phosphatase 1 regulatory subunit 10				
RBM12_HUMAN	RNA-binding protein 12				
POTEF_HUMAN	POTE ankyrin domain family member F				

CHAPTER 5

DEFINING THE INTERACTOME OF OGT VARIANTS CAUSAL FOR X-LINKED INTELLECTUAL DISABILITY

Introduction

In recent years, several mutations in the O-GlcNAc Transferase gene that are causal for X-Linked Intellectual Disability (XLID) have been identified. XLID is a set of neurodevelopmental disorders caused by alterations in genes on the X chromosome, with hallmark characteristics of intellectual disability (IQ less than 70) along with defects in adaptive functioning[30]. *OGT* mutations causing XLID were first described in the TPR domain of the protein[26,156], with several following in the catalytic domain[28,29]. Patients carrying these OGT mutations have varying degrees of severe intellectual disability, and a variety of additional phenotypes, but commonly have clinodactyly, hypogonadism, and dysmorphic facial features. One of the prevailing questions about these mutations is, considering that OGT is ubiquitously expressed[2] and necessary for embryonic development[4], what is the molecular mechanism of these mutations causing XLID?

OGT mutations in the catalytic domain seem to cause reduced catalytic activity of OGT. However, the 5 reported TPR mutations (**Fig. 5.1**) result in OGT variant enzymes that are catalytically comparable to the wild-type enzyme, are thermodynamically stable, and cause no noticeable effects on global O-GlcNAc levels[26,156]. Therefore, the mechanism of these mutations causing XLID is still unknown and of great interest. Since these mutations occur in the TPR domain, the domain of OGT hypothesized to be responsible for OGT substrate interaction[18,46], it has been suggested that the XLID mutations alter protein-protein interaction. In Chapters 3 and 4, we discussed the basal OGT TPR interactome, which had not previously been studied. Here we adapt the TurboID method to assess the OGT TPR interactome with OGT XLID mutations and

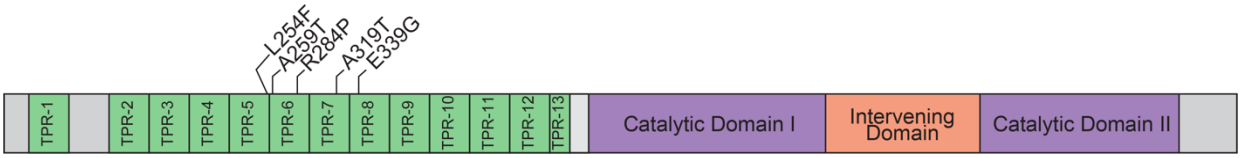


Figure 5.1: Diagram of the structure of OGT and location of the TPR XLID variants.

determine if protein-protein interaction is altered and may underlie the mechanism of these mutations causing XLID. Since XLID is primarily a neurological disorder, we examined OGT TPR interactomes in the neuroblastoma cell line, SHSY5Y, which has been extensively used to study intellectual disability[33–35]. Looking at both basal interactomes, and interactomes under membrane depolarization, we identified several differential interactors that represent potential mechanistic proteins underlying the OGT XLID phenotype.

Materials and Methods

Plasmid Constructs

OGT XLID mutation plasmids were generated with site-directed mutagenesis from the previously described TPR-TurboID plasmids.

Determination of OGT XLID TPR interactomes

SHSY5Y cell transfection, labeling, and identification of biotinylated proteins were performed as described previously. Method details are reiterated here for convenience.

Cell Culture, Transfection, and Collection

SHSY5Y cells were grown in DMEM/F12 50:50 with 10% FBS, 20mM Glutagro (Corning) and antibiotic-antimycotic. SHSY5Y cells were plated on 10cm dishes with 5.7e6 cells per plate for transfection. Transfection occurred 48 hours after plating with 31.3µg eGFP-Turbo or 61.2µg TPR-Turbo, wild-type (WT), or one of 5 XLID mutations (L254F, A259T, R284P, A319T, or E339G). 24 hours after transfection, cells were either labeled with 500µM biotin or labeled and depolarized with 50mM KCl and 500µM biotin. Cells were placed on ice to stop labeling, then plates scraped to collect cells. Cells were centrifuged to pellet, then resuspended in RIPA lysis

buffer (50mM Tris, 150mM NaCl, 1% IGEPAL-CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 1mM DTT, 1mM EDTA, protease and phosphatase inhibitors, and universal nuclease (ThermoFisher Scientific)) for 15 minutes on ice with vortexing for 30 seconds at 10 and 5 minutes. Three biological replicates were performed, each consisting of eGFP-TurboID, WT-TPR-TurboID, and 5 XLID mutant TPR-TurboIDs. All replicates were performed simultaneously, except 5 XLID samples which were repeated at a later time: A259T Label Replicate 3, R284P Label Replicate 3, L254F Depolarize Replicate 3, R284P Depolarize Replicate 2, and A319T Depolarize Replicate 3.

Isolation of biotinylated interactors

2mg of protein was added to 100uL neutravidin magnetic beads (Cytiva Life Sciences) and incubated overnight, rotating at 4C. Beads were washed 5X with RIPA buffer without inhibitors, then once with 1X TBS. Biotinylated proteins were eluted by adding 5% SDS in 50mM TEAB buffer and boiling at 95C for 10 minutes. Eluates were then digested on S-trap columns (Protifi) following the manufacturer's recommendation. For both serum starvation interactomes, a C18 peptide cleanup was performed following S-trap elution. Dried peptides were resuspended in 20uL sample buffer as described previously[134] for mass spectrometry.

Mass Spectrometry and Data Analysis

3.5uL of resuspended peptides was injected for each instrument run. LC-MS/MS was performed on an Orbitrap Lumos Tribrid mass spectrometer (ThermoFisher Scientific) equipped with an Ultimate 3000 RSLCnano HPLC system (Thermofisher Scientific). Repeated replicates were run on an Orbitrap Eclipse Tribrid mass spectrometer (ThermoFisher Scientific) equipped with an Ultimate 3000 RSLCnano HPLC system (Thermofisher Scientific). In all cases, gradient and instrument settings were identical to those previously described[134], excepting dynamic exclusion occurring for 20 seconds after 2 occurrences in 10 seconds.

Raw files were searched using ThermoFisher Scientific Proteome Discoverer (2.5.0.400) using a combined SequestHT and Mascot (2.6.2) search. The proteomic database consisted of all SwissProt annotated human protein sequences (obtained 03-2021) plus a list of common contaminants and the sequences for eGFP and TurboID. Tryptic cleavage was specific at Arg and Lys with two missed cleavages allowed. For both search engines, precursor mass tolerance was 10ppm and fragment mass tolerance was 0.6Da. Methylthio on Cys was a fixed modification, and variable modifications were allowed: Oxidation of Met, HexNAc on Ser/Thr, and biotin on Lys. For eGFP-TurboID, peptide and PSM FDR were set at 2%, and for TPR-TurboID, peptide and PSM FDR were set at 1%. Proteins were included in each final protein list were required to be the master protein of a protein group, have 2 or greater unique peptides, and a positive score for both Sequest and Mascot.

We utilized the proteomic data analysis software Simplifi (Protifi) for statistical analysis. Statistical significance among groups was determined using a combination of Wilcoxon Signed Rank Test and T-Test. Since Simplifi does not require data to be parametric we did not utilize NSAF of normal logarithmic transformation and instead performed statistical tests on raw spectral counts with 1 added to all spectral counts to avoid zero values. In order to be considered significant, a Wilcoxon p value of less than 0.05, a T-test p value of less than 0.05, and a 2-fold higher spectral count was required. This was required for both comparisons of experimental condition to negative control (TPR to eGFP) and across experimental conditions.

Western Blots and Antibodies

SDS-PAGE gels (4-15%) were run using the BioRad Mini-PROTEAN gel system and transferred onto Immobilin-FL PVDF membranes (Sigma-Aldritch) using the BioRad Trans-Blot SD Semi-Dry Transfer Cell system. Membrane blocking was in 1% cold water fish skin gelatin (Sigma-Aldrich). Primary antibody ratios were as follows: FLAG (1:2500 Sigma F3165), anti-biotin (1:1000, Jackson 200-002-211), anti-β-actin (1:2000, Cell Signaling 4970). Secondary antibodies

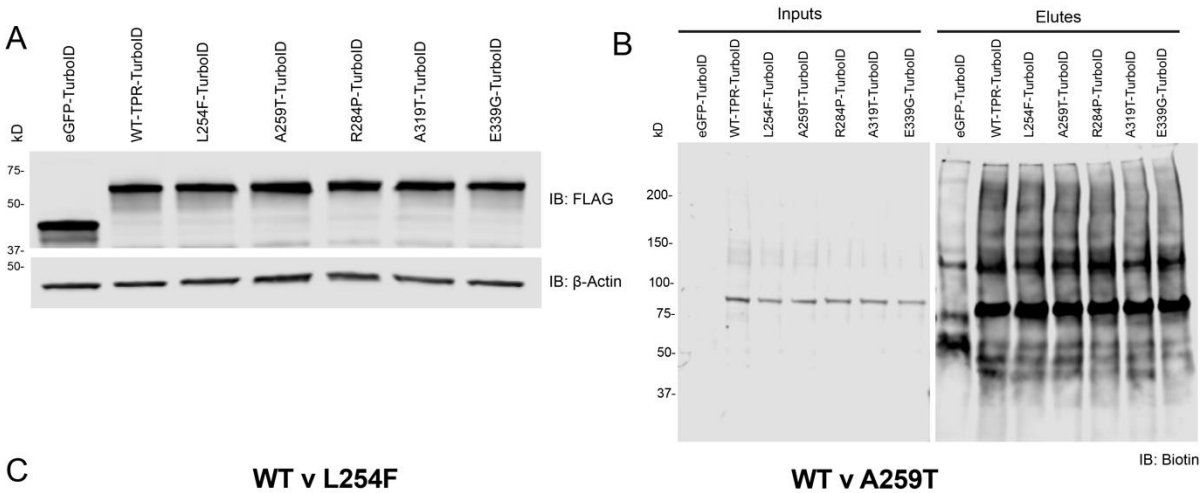
were LiCor IRDye: 680RD donkey-anti-mouse 680 (1:10000) and 800CW Goat anti-rabbit (1:20000). Three washes in TBST (0.1% tween) were performed after each antibody incubator. All Western blots were imaged on a Li-Cor Odyssey Clx system.

Results

OGT XLID Basal Interactomes

We first sought out to describe basal OGT TPR interactomes for OGT XLID TPR variant enzymes in SHSY5Y cells using the same TurboID system described in chapter 4. Expression of fusion proteins is shown in **Figure 5.2A**, and enrichment of biotinylated proteins is shown in **Figure 5.2B**. Final interactomes are shown in **Table 5.2**, and volcano plots of interacting proteins are shown in **Figure 5.2C**. When considering proteins that are candidates for mechanistic interactors underlying the OGT XLID phenotype, proteins that fail to interact with several XLID variants over wild-type or proteins that gain an interaction with several XLID variants over wild-type are of highest interest. As before, in order for a protein interactor to be considered significantly enriched, it was required to have t-test and Wilcoxon p-values of less than 0.05 and a spectral count fold change of at least two across TPR-TurboID conditions.

One interactor of interest immediately jumps out from these basal interactome comparisons. Tet2, a known and well-studied OGT interactor[20], interacts with wild-type TPR-Turbo but fails to interact with four of the five XLID variants. Tet2 is involved in the demethylation of DNA and therefore is an important transcriptional regulator. According to the OMIM catalogue, Tet2 has not previously been linked to neurodevelopmental disorders but is associated with immunodeficiency and myelodysplastic syndromes[107]. It is important to note that this interactor is only of medium confidence (meaning that it was identified with less than 10 spectral counts in all groups, but was significantly enriched in wild-type). However, it was unidentified entirely in three of the four variant interactomes (L254F, A259T, and R284P) and is enriched over A319T. Since Tet2 has already been shown to interact with OGT, validation of these differential interactions *in*


A
kD
eGFP-TurboID
WT-TPR-TurboID
L254F-TurboID
A259T-TurboID
R284P-TurboID
A319T-TurboID
E339G-TurboID
75-
50-
37-
50-
IB: FLAG
IB: β-Actin
B
Inputs
Elutes
kD
200-
150-
100-
75-
50-
37-
IB: Biotin

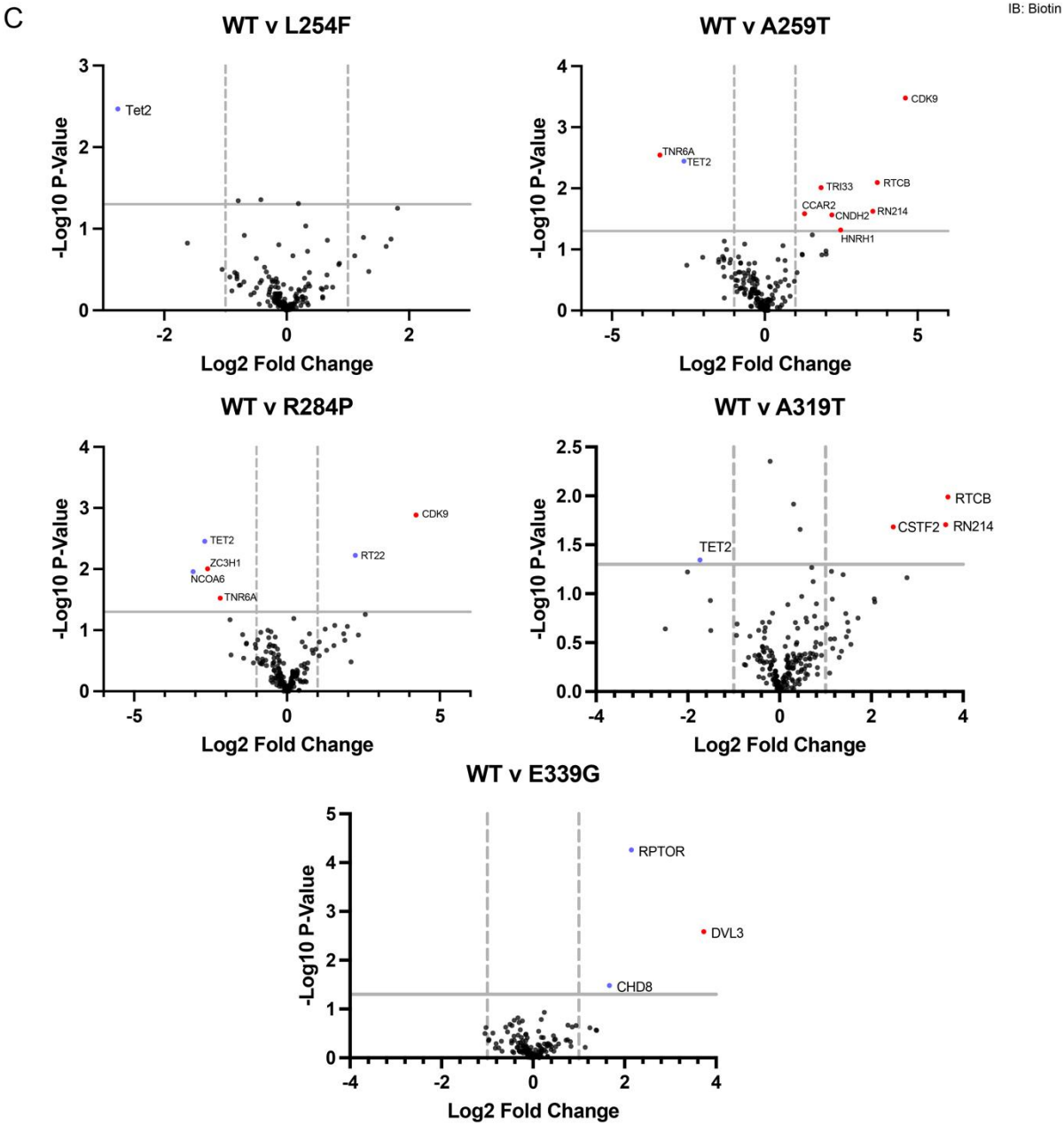

C
WT v L254F
Tet2
-Log10 P-Value
Log2 Fold Change
WT v A259T
CDK9
TNR6A
TET2
TRI33
RTCB
CCAR2
CNDH2
RN214
HNRH1
WT v R284P
CDK9
TET2
ZC3H1
NCOA6
RT22
TNR6A
WT v A319T
RTCB
CSTF2
RN214
TET2
WT v E339G
RPTOR
DVL3
CHD8

Figure 5.2: OGT XLID TPR variant basal interactomes in SHSY5Y cells. ***A,*** Expression of TurboID fusion proteins ***B,*** Biotin western blot showing enrichment of biotinylated proteins following neutravidin pulldown ***C,*** showing abundance and significance of proteins across glucose conditions. Vertical dashed lines indicate log2 fold change cutoff of 1 (fold change of 2) and horizontal solid line indicates T-test p-value cutoff of 0.05. Red points are "high confidence", meaning that at least one of the two conditions had higher than 10 total spectral counts. Purple points are "medium confidence", meaning that while statistical and fold metrics were achieved, both conditions had less than ten spectral counts.

vitro is high priority in determining whether the interaction differences seen here can be recapitulated.

Also of interest is the protein TNR6A, a protein involved in RNA-mediated gene-silencing. TNR6A interacts with the wild-type TPR-Turbo but fails to interact with A259T and has reduced interaction with R284P. TNR6A has previously been linked to an epilepsy disorder, so it may have relevant functional roles in the nervous system. While this interaction deficit is only observed in two of the five XLID variants, this result is still interesting to note because it is possible that there is heterogeneity among the molecular mechanisms that lead to *OGT* mutations causing XLID. While the simplest mechanism to investigate would be one that is unifying among all XLID variants, it is possible that heterogenous interaction differences could lead to a similar phenotype. Additionally, causality for all of the published variants has not been extensively established (except for L254F, for which a full family tree and X-inactivation studies have been performed[26]), so it is possible that mutations that appear to be mechanistic outliers may simply not be the casual mutation for the disorder observed. Therefore, an interaction difference consistently occurring in even two of the variants is still of interest.

Additionally, there are a few interactors that appear to gain interaction with several XLID OGT variants. RTCB, which is involved in the maturation of tRNAs, interacts uniquely with both A259T and A319T. Also, intriguingly, the protein CDK9 appears to newly interact with A259T and R284P. This transcription regulating kinase also interacts with the wild-type TPR-TurboID protein under depolarization, as shown in Chapter 4. Since we hypothesized there that the interaction of

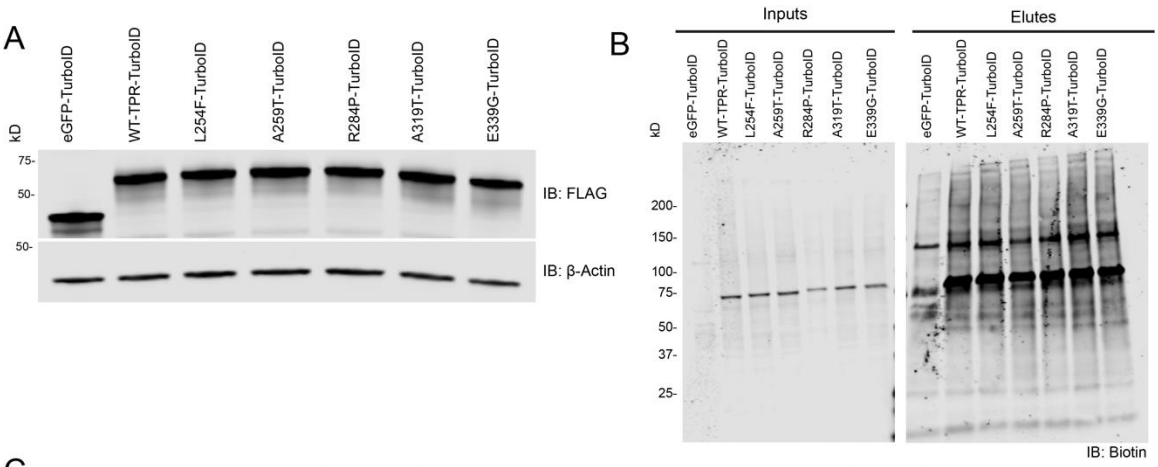
CDK9 and OGT happens under stress, it is interesting to note that this interaction also occurs under basal cellular conditions with expression of XLID variant TPR-TurboIDs. This could indicate that the cells carrying XLID variant TPR-TurboIDs are under additional stress as compared to wild-type, or that the sequence alteration of these XLID variants uniquely induces them to constitutively interact with CDK9.

OGT XLID Interactomes under depolarization

Since XLID is a neurodevelopmental disorder, we also wanted to observe how the XLID TPR-TurboID interactomes would change in neuronal cells experiencing an action potential, which can be simulated here by depolarizing SHSY5Y cells with KCl. Since we previously observed that changes in the OGT TPR interactome with depolarization of SHSY5Y cells largely involves non-neuron-specific proteins, this can also serve as an assessment of how the OGT TPR interactomes may change under stress. Expression of TPR-TurboID fusion proteins for depolarized cells is shown in **Figure 5.3A**, and enrichment of biotinylated proteins is shown in **Figure 5.3B**.

As shown in **Table 5.3** and **Figure 5.3C**, many more interaction changes were observed under depolarization than in the basal interactomes. This is unsurprising, since we reported in chapter 4 that OGT gains additional interactions under depolarization in SHSY5Y cells. First viewing these interaction differences in comparison to the basal interactomes reveals some interesting differential interactors. The protein ZC3H1, which is involved in the targeting of pre-mRNAs for degradation, has reduced interaction with R284P over wild-type in the basal interactome, but under depolarization has reduced interaction with A259T, R284P, and A319T. Additionally, the protein RN214 interacts exclusively with A259T and A319T in the basal interactomes, but additionally newly interacts with R284P under depolarization.

Additionally, there are several new interactors of interest in the depolarization condition. One such is the protein piccolo, which we reported in chapter 4 newly interacts with OGT under


A
kD
eGFP-TurboID
WT-TPR-TurboID
L254F-TurboID
A259T-TurboID
R284P-TurboID
A319T-TurboID
E339G-TurboID
75-
50-
50-
IB: FLAG
IB: β-Actin
B
Inputs
Elutes
kD
200-
150-
100-
75-
50-
37-
25-
IB: Biotin

C

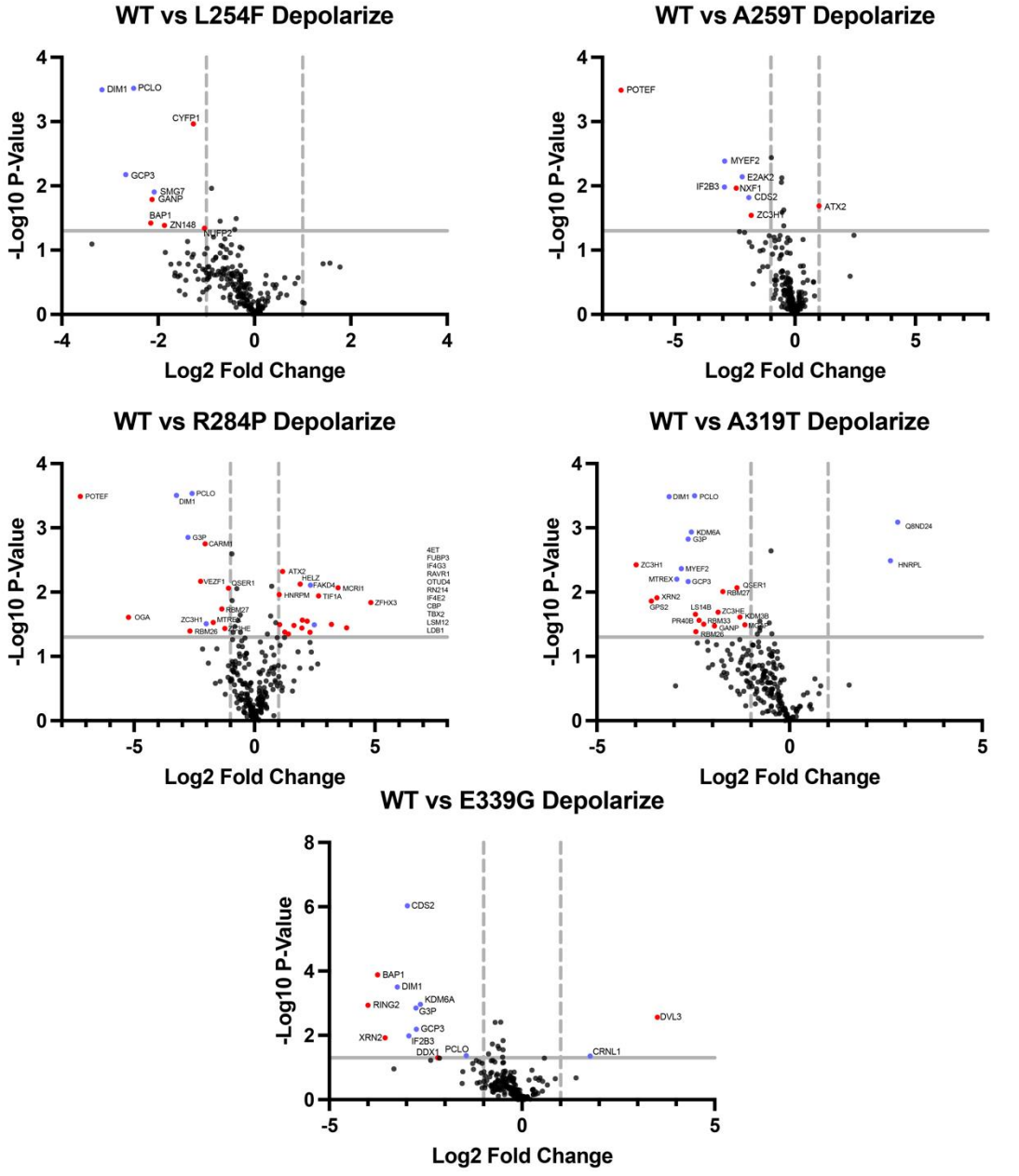

WT vs L254F Depolarize
WT vs A259T Depolarize
WT vs R284P Depolarize
WT vs A319T Depolarize
WT vs E339G Depolarize
-Log10 P-Value
Log2 Fold Change
DIM1
PCLO
CYFP1
GCP3
SMG7
GANP
BAP1
ZN148
MYEF2
POTEF
E2AK2
IF2B3
NXF1
CDS2
ZC3H1
ATX2
OGA
CARM1
VEZF1
QSER1
HELZ
FAKD4
MCRI1
HNRPM
TIF1A
ZFHX3
RBM27
MTREX
RBM26
ZC3HE
4ET
FUBP3
IF4G3
RAVR1
CTUD4
RN214
IF4E2
C9P
TBX2
LSM12
LDB1
Q8ND24
HNRPL
KDM6A
G3P
XRN2
GPS2
LS14B
PR40B
RBM33
GANP
KDM3B
MK13
DVL3
CRNL1
RING2
IF2B3
DDX1

Figure 5.3: OGT XLID TPR variant interactomes in depolarized SHSY5Y cells. ***A,*** Expression of TurboID fusion proteins ***B,*** Biotin western blot showing enrichment of biotinylated proteins following neutravidin pulldown ***C,*** showing abundance and significance of proteins across glucose conditions. Vertical dashed lines indicate log2 fold change cutoff of 1 (fold change of 2) and horizontal solid line indicates T-test p-value cutoff of 0.05. Red points are "high confidence", meaning that at least one of the two conditions had higher than 10 total spectral counts. Purple points are "medium confidence", meaning that while statistical and fold metrics were achieved, both conditions had less than ten spectral counts.

depolarization. Piccolo fails to interact with three of the five XLID variants (L254F, R284P, A319T) and has reduced interaction with E339G. Since this protein is involved in synaptic signaling, this is an interactor of high interest to possibly explain the neurological phenotype of XLID. As with Tet2, however, this interaction is of medium confidence due to spectral counts so the interaction differences must be validated with a separate method. A new interactor of interest, Dim1, which is involved in rRNA processing, confidently interacts with the wild-type TPR-TurboID but fails to interact with four of the five XLID variants (L254F, R284P, A319T, and E339G). While this protein has limited research available on its function and is not currently linked to any neurological disorders, that it completely fails to interact with a majority of the XLID variants makes it of high interest in pursuing the XLID mechanism. It is important to note that this interaction seems to only occur under depolarization (and with high confidence over its negative control – see **Table 4.5**) so specific targeted study of how this interaction occurs will be necessary to parse out how it may contribute to the XLID phenotype.

Several proteins interact with the wild-type TPR-TurboID but fail to interact or have reduced interaction with three of the XLID variants. These are GCP3, ZC3H1 (as previously mentioned), and G3P. Both GCP3 and G3P are ubiquitously expressed with globally vital functions (with overlap in cytoskeletal organization) and have not been previously linked to neurological disorders so the possible function of their differential interaction with XLID variants is not immediately clear.

Discussion

In this chapter, we set out to determine if OGT interactomes are altered with XLID variant enzymes, and to identify a set of candidate mechanistic proteins that might underlie the XLID phenotype. Here, no extremely strong candidates present themselves (e.g. a protein that is previously linked to intellectual disability and strongly interacts with the wild-type OGT and fails to interact with all of the XLID variant enzymes). However, we have identified a handful of interactors as possible mechanistic proteins and are worthy of additional research. Tet2, Piccolo, Dim1, and ZC3H1 are the strongest identified candidates and would benefit from future study.

Tet2, as previously mentioned, is known to interact with OGT and in fact assists in targeting OGT to chromatin and promoting its O-GlcNAc modification of histones[20]. Here, Tet2 interacts with OGT only under basal conditions and not under depolarization. Tet2 is one of several Tet proteins that catalyzes the conversion of methylated cytosine (5mC) to hydroxymethylcytosine (5hmC), which is the first step in DNA demethylation. While mutations in Tet2 have not specifically been linked to neurodevelopmental disorders, it is linked to a neurodegenerative disorder[157] and myelodysplastic syndrome, which results in poor differentiation of hematopoietic stem cells. Tet proteins play essential roles in the maintenance and early specification of embryonic stem cells[158,159]. Since intellectual disability disorders can be observed on a molecular level very early on in neural deveopment[160], it is possible that Tet2's role in early embryonic development could be affected by a lack of interaction with OGT and lead to XLID. However, research on Tet2 and OGT's interaction has shown that O-GlcNAc modification of Tet2 does not have an impact on its function[20]. Considering this, it is also possible that the reverse is true: that OGT failing to interact with Tet2 results in OGT not being targeted to chromatin correctly, resulting in altered O-GlcNAc modification of histones, leading to developmental deficits. OGT itself is essential for early embryonic development[4], and plays important roles in embryonic stem cell fate and differentiation, especially of neural

development[113,161–163]. The hypothesis that OGT fails to interact with Tet2 during early embryonic development, leading to altered OGT localization and modification of histone proteins is an intriguing possible mechanism for OGT XLID.

Piccolo is another interesting case of an OGT interactor that fails to interact with OGT variants under depolarization. Piccolo has been linked to the neurodevelopmental disorder pontocerebellar hypoplasia[164], but in fact has very limited research on its expression and function in embryonic and early neuronal cells. In mature neurons, piccolo forms a part of the presynaptic cytomatrix, and is an important organization component of the presynapse[165]. OGT has also been shown to localize to the presynapse[148], and piccolo is known to be extensively O-GlcNAc modified[152]. However, the function of this O-GlcNAc modification is not known. In the absence of information about what function, if any, piccolo serves in neuronal development, it is difficult to speculate on what the functional consequence of this interaction being interrupted may be. It is possible that any deficit in interaction with OGT is more a byproduct of the XLID phenotype, i.e. that abnormal neuronal development occurs due to some other altered interaction, and then piccolo also fails to interact with OGT in mature neurons. It will be interesting to confirm this interaction deficit and look toward future research in describing if piccolo is expressed and functional during neurodevelopment and what the functional role of its interaction with OGT is.

Another interactor of high interest is DIM1, which interacts with OGT only under depolarization. DIM1 has not been studied extensively in any system, except to be linked to poor survival in gastric carcinoma[166] and suggested to have an essential role in ribosome assembly by virtue of its function as an rRNA processing enzyme[167]. DIM1 is ubiquitously expressed[168], so it having a neuronal specific function is unlikely. Here we report that DIM1 interacts with OGT only under depolarization, and hypothesize that it may be a more generic stress response interaction rather than one specific to depolarization. However, one other published OGT interactome did not report Dim1 as an OGT interactor under oxidative stress[155]. This interactome was performed in

mouse embryonic fibroblast cells, so it is possible that this interaction is still somewhat cell and/or species specific. Without much published information on the function of Dim1, it is difficult to speculate on what the role of this interaction might be. However, with a function as essential and ubiquitous as ribosome assembly, if OGT plays some important role in this function, it is possible that an interruption of that interaction and function could have an important impact on early neurodevelopment. Validating this interaction and further studying the functional roles of Dim1 will be essential to identifying whether this interaction mechanistically underlies the OGT XLID phenotype.

Finally, the protein ZC3H1 has reduced interaction with three of the five XLID variant enzymes, although the degree of this effect varies between the basal and depolarized interactomes. This protein interactor, which is involved in the targeting of polyadenylated RNAs for exosomal degradation as part of the PAXT complex[169], jumps out as a candidate protein of interest due to previous research on its roles in embryonic stem cell development. A study performed in mouse embryonic stem cells found that ZC3H1 knockout cells exhibit reduced PAXT-mediated RNA decay and a reduced ability to differentiate[170]. This phenotype was linked to deregulation and reduced function of polycomb repressive complex 2 (PRC2) proteins. OGT itself is a polycomb group protein in *Drosophila*[112,113], and is essential for normal gene silencing mediated by polycomb group proteins. While the function OGT serves toward mammalian polycomb group proteins and PRC2 is still unclear[171], the potential link between this role and its interaction with ZC3H1 is intriguing. While ZC3H1 itself has not been linked to any developmental disorders, PRC2 proteins are linked to a variety of congenital disorders with similarities to our OGT XLID patients, including intellectual disability and dysmorphic facial features[172]. If OGT's interaction with or O-GlcNAc modification of ZC3H1 (ZC3H1 has been reported to be O-GlcNAc modified[173]) plays a role in its RNA-degradation function, it is possible that reduced OGT interaction could lead to a similar deficit in PRC2 function as previously observed, and thus lead

to the neurodevelopmental phenotype we see in XLID. Confirmation of the interaction deficit between ZC3H1 and XLID variant OGTs and additional research on OGT's impact on ZC3H1's function are necessary to confirm this link.

Recommendations for future study are detailed in chapter 6, but briefly, a few obvious avenues for future work present themselves. Validation of these differential interactions is essential, and specifically targeting the mechanism or downstream effects of differential interactors to see if they are altered with OGT XLID variants will help to reveal if these proteins may mechanistically underlie the OGT XLID phenotype. Additionally, repeating variant interactomes in a more immediately neurodevelopmental system, such as embryonic stem cells, neural precursor cells, or neural crest cells, may allow for accurate capture of when these interactions are most relevant to the neurodevelopmental process.

Table 5.1: TPR-TurboID interactors with TPR XLID mutations in SHSY5Y cells

Dark gray - identified only in TPR-TurboID, not present in eGFP-TurboID
Light gray - identified in both TPR-TurboID and eGFP-TurboID, significantly enriched in TPR-TurboID
Dark blue - identified only in the indicated TPR-TurboID condition
Light blue - identified in both TPR-TurboID conditions, significantly enriched in indicated TPR-TurboID

		WT						L254F		A259T		R284P		A319T		E339G	
Accession	Description	Over eGFP	Over L254F	Over A259T	Over R284P	Over A319T	Over E339G	Over eGFP	Over WT	Over eGFP	Over WT	Over eGFP	Over WT	Over eGFP	Over WT	Over eGFP	Over WT
TET2_HUMAN	Methylcytosine dioxygenase TET2																
TNR6A_HUMAN	Trinucleotide repeat-containing gene 6A protein																
NCOA6_HUMAN	Nuclear receptor coactivator 6																
ZC3H1_HUMAN	Zinc finger C3H1 domain-containing protein																
RN214_HUMAN	RING finger protein 214																
RTCB_HUMAN	RNA-splicing ligase RtcB homolog																
CCAR2_HUMAN	Cell cycle and apoptosis regulator protein 2																
TRI33_HUMAN	E3 ubiquitin-protein ligase TRIM33																
CND2_HUMAN	Condensin complex subunit 2																
HNRH1_HUMAN	Heterogeneous nuclear ribonucleoprotein H																
CDK9_HUMAN	Cyclin-dependent kinase 9																
RT22_HUMAN	28S ribosomal protein S22, mitochondrial																
CSTF2_HUMAN	Cleavage stimulation factor subunit 2																
RPTOR_HUMAN	Regulatory-associated protein of mTOR																
DVL3_HUMAN	Segment polarity protein dishevelled homolog DVL-3																
CHD8_HUMAN	Chromodomain-helicase-DNA-binding protein 8																
P121C_HUMAN	Nuclear envelope pore membrane protein POM 121C																
BCL9_HUMAN	B-cell CLL/lymphoma 9 protein																

SC16A_HUMAN	Protein transport protein Sec16A																		
SET1A_HUMAN	Histone-lysine N-methyltransferase SETD1A																		
OGT1_HUMAN	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit																		
KDM1A_HUMAN	Lysine-specific histone demethylase 1A																		
TBL1X_HUMAN	F-box-like/WD repeat-containing protein TBL1X																		
CD2B2_HUMAN	CD2 antigen cytoplasmic tail-binding protein 2																		
SC24B_HUMAN	Protein transport protein Sec24B																		
ITF2_HUMAN	Transcription factor 4																		
YLPM1_HUMAN	YLP motif-containing protein 1																		
NUP98_HUMAN	Nuclear pore complex protein Nup98-Nup96																		
WDR5_HUMAN	WD repeat-containing protein 5																		
RING1_HUMAN	E3 ubiquitin-protein ligase RING1																		
CBX3_HUMAN	Chromobox protein homolog 3																		
PABP4_HUMAN	Polyadenylate-binding protein 4																		
GSE1_HUMAN	Genetic suppressor element 1																		
RBBP5_HUMAN	Retinoblastoma-binding protein 5																		
SF01_HUMAN	Splicing factor 1																		
QSER1_HUMAN	Glutamine and serine-rich protein 1																		
PRC2B_HUMAN	Protein PRRC2B																		
RPRD2_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 2																		
KDM3B_HUMAN	Lysine-specific demethylase 3B																		
NUP54_HUMAN	Nucleoporin p54																		
ANKH1_HUMAN	Ankyrin repeat and KH domain-containing protein 1																		
CPSF7_HUMAN	Cleavage and polyadenylation specificity factor subunit 7																		
NUP58_HUMAN	Nucleoporin p58/p45																		
TBL1R_HUMAN	F-box-like/WD repeat-containing protein TBL1XR1																		
SP130_HUMAN	Histone deacetylase complex subunit SAP130																		
RCOR3_HUMAN	REST corepressor 3																		
RBM27_HUMAN	RNA-binding protein 27																		
ASH2L_HUMAN	Set1/Ash2 histone methyltransferase complex subunit ASH2																		
RCOR1_HUMAN	REST corepressor 1																		
PRR12_HUMAN	Proline-rich protein 12																		
ZN281_HUMAN	Zinc finger protein 281																		
PRP4_HUMAN	U4/U6 small nuclear ribonucleoprotein Prp4																		
DCAF7_HUMAN	DDB1- and CUL4-associated factor 7																		
PHX2B_HUMAN	Paired mesoderm homeobox protein 2B																		
ZMYM4_HUMAN	Zinc finger MYM-type protein 4																		
ZC3HE_HUMAN	Zinc finger CCCH domain-containing protein 14																		

CCAR1_HUMAN	Cell division cycle and apoptosis regulator protein 1																		
NCOR2_HUMAN	Nuclear receptor corepressor 2																		
HDAC3_HUMAN	Histone deacetylase 3																		
SC23B_HUMAN	Protein transport protein Sec23B																		
PF21A_HUMAN	PHD finger protein 21A																		
DIDO1_HUMAN	Death-inducer obliterator 1																		
TNR6C_HUMAN	Trinucleotide repeat-containing gene 6C protein																		
NCOR1_HUMAN	Nuclear receptor corepressor 1																		
CARM1_HUMAN	Histone-arginine methyltransferase CARM1																		
SSBP3_HUMAN	Single-stranded DNA-binding protein 3																		
MYT1_HUMAN	Myelin transcription factor 1																		
RBBP7_HUMAN	Histone-binding protein RBBP7																		
UBAP2_HUMAN	Ubiquitin-associated protein 2																		
EMSY_HUMAN	BRCA2-interacting transcriptional repressor EMSY																		
MBD3_HUMAN	Methyl-CpG-binding domain protein 3																		
ATN1_HUMAN	Atrophin-1																		
TLE1_HUMAN	Transducin-like enhancer protein 1																		
CIC_HUMAN	Protein capicua homolog																		
UBN2_HUMAN	Ubinuclein-2																		
PSPC1_HUMAN	Paraspeckle component 1																		
BCOR_HUMAN	BCL-6 corepressor																		
DDX6_HUMAN	Probable ATP-dependent RNA helicase DDX6																		
KANL3_HUMAN	KAT8 regulatory NSL complex subunit 3																		
KMT2D_HUMAN	Histone-lysine N-methyltransferase 2D																		
IMPA1_HUMAN	Inositol monophosphatase 1																		
MEX3A_HUMAN	RNA-binding protein MEX3A																		
PP1A_HUMAN	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit																		
CSK21_HUMAN	Casein kinase II subunit alpha																		
ZN318_HUMAN	Zinc finger protein 318																		
SYNP2_HUMAN	Synaptopodin-2																		
JHD2C_HUMAN	Probable JmjC domain-containing histone demethylation protein 2C																		
LS14B_HUMAN	Protein LSM14 homolog B																		
GANP_HUMAN	Germinal-center associated nuclear protein																		
P121A_HUMAN	Nuclear envelope pore membrane protein POM 121																		
HIRA_HUMAN	Protein HIRA																		
RAGP1_HUMAN	Ran GTPase-activating protein 1																		
NU153_HUMAN	Nuclear pore complex protein Nup153																		
RAE1L_HUMAN	mRNA export factor																		

KMT2A_HUMAN	Histone-lysine N-methyltransferase 2A																		
REQU_HUMAN	Zinc finger protein ubi-d4																		
HNRPK_HUMAN	Heterogeneous nuclear ribonucleoprotein K																		
YTHD2_HUMAN	YTH domain-containing family protein 2																		
RSMB_HUMAN	Small nuclear ribonucleoprotein-associated proteins B and B'																		
NHSL2_HUMAN	NHS-like protein 2																		
SPF45_HUMAN	Splicing factor 45																		
WNK1_HUMAN	Serine/threonine-protein kinase WNK1																		
TRRAP_HUMAN	Transformation/transcription domain-associated protein																		
RUXF_HUMAN	Small nuclear ribonucleoprotein F																		
F193A_HUMAN	Protein FAM193A																		
ZN148_HUMAN	Zinc finger protein 148																		
ZN609_HUMAN	Zinc finger protein 609																		
IMA7_HUMAN	Importin subunit alpha-7																		
ANR17_HUMAN	Ankyrin repeat domain-containing protein 17																		
PABP1_HUMAN	Polyadenylate-binding protein 1																		
PCNA_HUMAN	Proliferating cell nuclear antigen																		
PRC2A_HUMAN	Protein PRRC2A																		
HCFC1_HUMAN	Host cell factor 1																		
HDAC1_HUMAN	Histone deacetylase 1																		
UBP2L_HUMAN	Ubiquitin-associated protein 2-like																		
IMB1_HUMAN	Importin subunit beta-1																		
BAP18_HUMAN	Chromatin complexes subunit BAP18																		
ARI1B_HUMAN	AT-rich interactive domain-containing protein 1B																		
NUP88_HUMAN	Nuclear pore complex protein Nup88																		
PRC2C_HUMAN	Protein PRRC2C																		
MTA2_HUMAN	Metastasis-associated protein MTA2																		
PHX2A_HUMAN	Paired mesoderm homeobox protein 2A																		
SR140_HUMAN	U2 snRNP-associated SURP motif-containing protein																		
HTF4_HUMAN	Transcription factor 12																		
MGAP_HUMAN	MAX gene-associated protein																		
ARI1A_HUMAN	AT-rich interactive domain-containing protein 1A																		
U520_HUMAN	U5 small nuclear ribonucleoprotein 200 kDa helicase																		
NU214_HUMAN	Nuclear pore complex protein Nup214																		
NUP62_HUMAN	Nuclear pore glycoprotein p62																		
RBP2_HUMAN	E3 SUMO-protein ligase RanBP2																		
BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF																		
U5S1_HUMAN	116 kDa U5 small nuclear ribonucleoprotein component																		

PRP8_HUMAN	Pre-mRNA-processing-splicing factor 8																		
CHERP_HUMAN	Calcium homeostasis endoplasmic reticulum protein																		
HDAC2_HUMAN	Histone deacetylase 2																		
SNR40_HUMAN	U5 small nuclear ribonucleoprotein 40 kDa protein																		
RBM14_HUMAN	RNA-binding protein 14																		
SIN3A_HUMAN	Paired amphipathic helix protein Sin3a																		
DDX23_HUMAN	Probable ATP-dependent RNA helicase DDX23																		
SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich																		
AGFG1_HUMAN	Arf-GAP domain and FG repeat-containing protein 1																		
TurboID	TurboID TurboID with 2X FLAG and 3X GGGGS linker																		
TBG1_HUMAN	Tubulin gamma-1 chain																		
BUB3_HUMAN	Mitotic checkpoint protein BUB3																		
TCRG1_HUMAN	Transcription elongation regulator 1																		
CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4																		
CPSF5_HUMAN	Cleavage and polyadenylation specificity factor subunit 5																		
SF3B2_HUMAN	Splicing factor 3B subunit 2																		
PRP19_HUMAN	Pre-mRNA-processing factor 19																		
TLE3_HUMAN	Transducin-like enhancer protein 3																		
POGZ_HUMAN	Pogo transposable element with ZNF domain																		
RPR1B_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 1B																		
NUP50_HUMAN	Nuclear pore complex protein Nup50																		
MINT_HUMAN	Msx2-interacting protein																		
TFG_HUMAN	Protein TFG																		
NIPBL_HUMAN	Nipped-B-like protein																		
RBM33_HUMAN	RNA-binding protein 33																		
VEZF1_HUMAN	Vascular endothelial zinc finger 1																		
RPB1_HUMAN	DNA-directed RNA polymerase II subunit RPB1																		
ATX2L_HUMAN	Ataxin-2-like protein																		
RENT1_HUMAN	Regulator of nonsense transcripts 1																		
RBM10_HUMAN	RNA-binding protein 10																		
SMD3_HUMAN	Small nuclear ribonucleoprotein Sm D3																		
CHAP1_HUMAN	Chromosome alignment-maintaining phosphoprotein 1																		
AKAP8_HUMAN	A-kinase anchor protein 8																		
NCOA2_HUMAN	Nuclear receptor coactivator 2																		
NFIB_HUMAN	Nuclear factor 1 B-type																		
DOT1L_HUMAN	Histone-lysine N-methyltransferase, H3 lysine-79 specific																		
BICRL_HUMAN	BRD4-interacting chromatin-remodeling complex-associated protein-like																		
SMG7_HUMAN	Protein SMG7																		

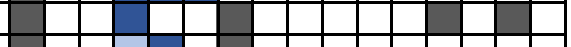
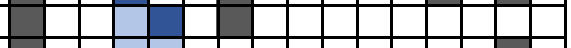
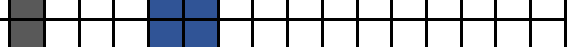
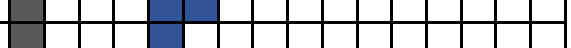
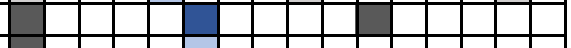
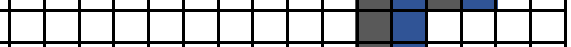
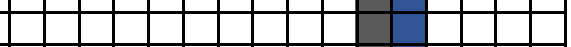
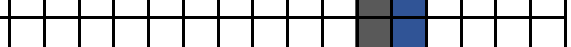
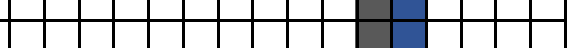
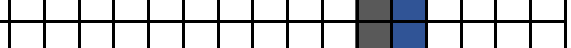
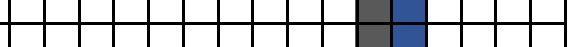
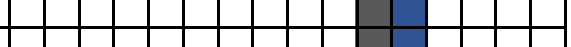
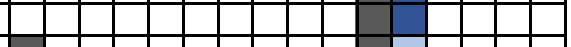
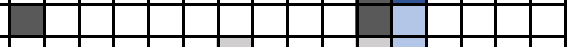
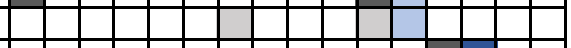
DYL2_HUMAN	Dynein light chain 2, cytoplasmic																		
RBM26_HUMAN	RNA-binding protein 26																		
JUNB_HUMAN	Transcription factor jun-B																		
THA11_HUMAN	THAP domain-containing protein 11																		
ZN207_HUMAN	BUB3-interacting and GLEBS motif-containing protein ZNF207																		
AF17_HUMAN	Protein AF-17																		
GPS2_HUMAN	G protein pathway suppressor 2																		
ROA1_HUMAN	Heterogeneous nuclear ribonucleoprotein A1																		
RP9_HUMAN	Retinitis pigmentosa 9 protein																		
YETS2_HUMAN	YEATS domain-containing protein 2																		
MYEF2_HUMAN	Myelin expression factor 2																		
BCORL_HUMAN	BCL-6 corepressor-like protein 1																		
CNOT1_HUMAN	CCR4-NOT transcription complex subunit 1																		
I2BP2_HUMAN	Interferon regulatory factor 2-binding protein 2																		
PUM1_HUMAN	Pumilio homolog 1																		
FUBP2_HUMAN	Far upstream element-binding protein 2																		
PTN23_HUMAN	Tyrosine-protein phosphatase non-receptor type 23																		
DC1L1_HUMAN	Cytoplasmic dynein 1 light intermediate chain 1																		
YTHD1_HUMAN	YTH domain-containing family protein 1																		
IF4G3_HUMAN	Eukaryotic translation initiation factor 4 gamma 3																		
GGYF2_HUMAN	GRB10-interacting GYF protein 2																		
RUVB1_HUMAN	RuvB-like 1																		
CARF_HUMAN	CDKN2A-interacting protein																		
SF3A1_HUMAN	Splicing factor 3A subunit 1																		
NUFP2_HUMAN	Nuclear fragile X mental retardation-interacting protein 2																		
LDB1_HUMAN	LIM domain-binding protein 1																		
YTHD3_HUMAN	YTH domain-containing family protein 3																		
FUBP1_HUMAN	Far upstream element-binding protein 1																		
ATX2_HUMAN	Ataxin-2																		
RL10A_HUMAN	60S ribosomal protein L10a																		
PRP16_HUMAN	Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16																		
HNRPM_HUMAN	Heterogeneous nuclear ribonucleoprotein M																		
TCPE_HUMAN	T-complex protein 1 subunit epsilon																		
S23IP_HUMAN	SEC23-interacting protein																		
WDR82_HUMAN	WD repeat-containing protein 82																		
DDX1_HUMAN	ATP-dependent RNA helicase DDX1																		
RUVB2_HUMAN	RuvB-like 2																		
AHDC1_HUMAN	AT-hook DNA-binding motif-containing protein 1																		

CEP85_HUMAN	Centrosomal protein of 85 kDa																	
DMAP1_HUMAN	DNA methyltransferase 1-associated protein 1																	
DNJA1_HUMAN	DnaJ homolog subfamily A member 1																	
ILF2_HUMAN	Interleukin enhancer-binding factor 2																	
EP400_HUMAN	E1A-binding protein p400																	
RBBP4_HUMAN	Histone-binding protein RBBP4																	
OGA_HUMAN	Protein O-GlcNAcase																	
SEC13_HUMAN	Protein SEC13 homolog																	
NAB1_HUMAN	NGFI-A-binding protein 1																	
TNR6B_HUMAN	Trinucleotide repeat-containing gene 6B protein																	
ISL1_HUMAN	Insulin gene enhancer protein ISL-1																	
SSBP4_HUMAN	Single-stranded DNA-binding protein 4																	
SF3B4_HUMAN	Splicing factor 3B subunit 4																	
PP1RA_HUMAN	Serine/threonine-protein phosphatase 1 regulatory subunit 10																	
PPIL1_HUMAN	Peptidyl-prolyl cis-trans isomerase-like 1																	
TOX_HUMAN	Thymocyte selection-associated high mobility group box protein TOX																	
MATR3_HUMAN	Matrin-3																	
ROA3_HUMAN	Heterogeneous nuclear ribonucleoprotein A3																	
NU133_HUMAN	Nuclear pore complex protein Nup133																	
CUX1_HUMAN	Homeobox protein cut-like 1																	
CBX1_HUMAN	Chromobox protein homolog 1																	
CCNT1_HUMAN	Cyclin-T1																	
LIN54_HUMAN	Protein lin-54 homolog																	
ZFR_HUMAN	Zinc finger RNA-binding protein																	
SBNO1_HUMAN	Protein strawberry notch homolog 1																	
SNW1_HUMAN	SNW domain-containing protein 1																	
GTF2I_HUMAN	General transcription factor II-I																	
HNRPC_HUMAN	Heterogeneous nuclear ribonucleoproteins C1/C2																	
RAB14_HUMAN	Ras-related protein Rab-14																	
PRPF3_HUMAN	U4/U6 small nuclear ribonucleoprotein Prp3																	
IF4G2_HUMAN	Eukaryotic translation initiation factor 4 gamma 2																	
EPN4_HUMAN	Clathrin interactor 1																	
CDC5L_HUMAN	Cell division cycle 5-like protein																	
SUGP2_HUMAN	SURP and G-patch domain-containing protein 2																	
SRRM2_HUMAN	Serine/arginine repetitive matrix protein 2																	
HMGX4_HUMAN	HMG domain-containing protein 4																	
RPB2_HUMAN	DNA-directed RNA polymerase II subunit RPB2																	

Table 5.2: TPR-TurboID interactors with TPR XLID mutations in SHSY5Y cells under depolarization

Dark gray - identified only in TPR-TurboID, not present in eGFP-TurboID
Light gray - identified in both TPR-TurboID and eGFP-TurboID, significantly enriched in TPR-TurboID
Dark blue - identified only in the indicated TPR-TurboID condition
Light blue - identified in both TPR-TurboID conditions, significantly enriched in indicated TPR-TurboID

		WT						L254F		A259T		R284P		A319T		E339G	
Accession	Description	Over eGFP	Over L254F	Over A259T	Over R284P	Over A319T	Over E339G	Over eGFP	Over WT	Over eGFP	Over WT	Over eGFP	Over WT	Over eGFP	Over WT	Over eGFP	Over WT
DIM1_HUMAN	Probable dimethyladenosine transferase																
PCLO_HUMAN	Protein piccolo																
GCP3_HUMAN	Gamma-tubulin complex component 3																
GANP_HUMAN	Germinal-center associated nuclear protein																
BAP1_HUMAN	Ubiquitin carboxyl-terminal hydrolase BAP1																
SMG7_HUMAN	Protein SMG7																
ZN148_HUMAN	Zinc finger protein 148																
NUFP2_HUMAN	Nuclear fragile X mental retardation-interacting protein 2																
CYFP1_HUMAN	Cytoplasmic FMR1-interacting protein 1																
POTEF_HUMAN	POTE ankyrin domain family member F																
MYEF2_HUMAN	Myelin expression factor 2																
IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 3																
ZC3H1_HUMAN	Zinc finger C3H1 domain-containing protein																
CDS2_HUMAN	Phosphatidate cytidylyltransferase 2																
NXF1_HUMAN	Nuclear RNA export factor 1																
E2AK2_HUMAN	Interferon-induced, double-stranded RNA-activated protein kinase																

G3P_HUMAN	Glyceraldehyde-3-phosphate dehydrogenase	
OGA_HUMAN	Protein O-GlcNAcase	
MTREX_HUMAN	Exosome RNA helicase MTR4	
RBM26_HUMAN	RNA-binding protein 26	
RBM27_HUMAN	RNA-binding protein 27	
ZC3HE_HUMAN	Zinc finger CCCH domain-containing protein 14	
QSER1_HUMAN	Glutamine and serine-rich protein 1	
VEZF1_HUMAN	Vascular endothelial zinc finger 1	
CARM1_HUMAN	Histone-arginine methyltransferase CARM1	
XRN2_HUMAN	5'-3' exoribonuclease 2	
KDM6A_HUMAN	Lysine-specific demethylase 6A	
GPS2_HUMAN	G protein pathway suppressor 2	
LS14B_HUMAN	Protein LSM14 homolog B	
PR40B_HUMAN	Pre-mRNA-processing factor 40 homolog B	
RBM33_HUMAN	RNA-binding protein 33	
KDM3B_HUMAN	Lysine-specific demethylase 3B	
MGAP_HUMAN	MAX gene-associated protein	
RING2_HUMAN	E3 ubiquitin-protein ligase RING2	
DDX1_HUMAN	ATP-dependent RNA helicase DDX1	
ATX2_HUMAN	Ataxin-2	
RN214_HUMAN	RING finger protein 214	
CBP_HUMAN	CREB-binding protein	
LSM12_HUMAN	Protein LSM12 homolog	
RAVR1_HUMAN	Ribonucleoprotein PTB-binding 1	
TBX2_HUMAN	T-box transcription factor TBX2	
4ET_HUMAN	Eukaryotic translation initiation factor 4E transporter	
FUBP3_HUMAN	Far upstream element-binding protein 3	
LDB1_HUMAN	LIM domain-binding protein 1	
FAKD4_HUMAN	FAST kinase domain-containing protein 4	
OTUD4_HUMAN	OUT domain-containing protein 4	
TIF1A_HUMAN	Transcription intermediary factor 1-alpha	
IF4E2_HUMAN	Eukaryotic translation initiation factor 4E type 2	
MCRI1_HUMAN	Mapk-regulated corepressor-interacting protein 1	
ZFHX3_HUMAN	Zinc finger homeobox protein 3	
HELZ_HUMAN	Probable helicase with zinc finger domain	
IF4G3_HUMAN	Eukaryotic translation initiation factor 4 gamma 3	
HNRPL_HUMAN	Heterogeneous nuclear ribonucleoprotein L	
CRNL1_HUMAN	Crooked neck-like protein 1	

DVL3_HUMAN	Segment polarity protein dishevelled homolog DVL-3																		
ATN1_HUMAN	Atrophin-1																		
SET1A_HUMAN	Histone-lysine N-methyltransferase SETD1A																		
UBAP2_HUMAN	Ubiquitin-associated protein 2																		
KMT2D_HUMAN	Histone-lysine N-methyltransferase 2D																		
P121C_HUMAN	Nuclear envelope pore membrane protein POM 121C																		
BCL9_HUMAN	B-cell CLL/lymphoma 9 protein																		
OGT1_HUMAN	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit																		
KDM1A_HUMAN	Lysine-specific histone demethylase 1A																		
NCOR1_HUMAN	Nuclear receptor corepressor 1																		
SC24B_HUMAN	Protein transport protein Sec24B																		
ITF2_HUMAN	Transcription factor 4																		
NUP98_HUMAN	Nuclear pore complex protein Nup98-Nup96																		
TLE1_HUMAN	Transducin-like enhancer protein 1																		
TLE3_HUMAN	Transducin-like enhancer protein 3																		
CBX3_HUMAN	Chromobox protein homolog 3																		
SNW1_HUMAN	SNW domain-containing protein 1																		
GSE1_HUMAN	Genetic suppressor element 1																		
RBBP5_HUMAN	Retinoblastoma-binding protein 5																		
SC23B_HUMAN	Protein transport protein Sec23B																		
NCOA2_HUMAN	Nuclear receptor coactivator 2																		
SF01_HUMAN	Splicing factor 1																		
ZMYM4_HUMAN	Zinc finger MYM-type protein 4																		
NUP54_HUMAN	Nucleoporin p54																		
POGZ_HUMAN	Pogo transposable element with ZNF domain																		
CPSF7_HUMAN	Cleavage and polyadenylation specificity factor subunit 7																		
SP130_HUMAN	Histone deacetylase complex subunit SAP130																		
WNK1_HUMAN	Serine/threonine-protein kinase WNK1																		
RPR1B_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 1B																		
RCOR3_HUMAN	REST corepressor 3																		
ASH2L_HUMAN	Set1/Ash2 histone methyltransferase complex subunit ASH2																		
NUP50_HUMAN	Nuclear pore complex protein Nup50																		
PRR12_HUMAN	Proline-rich protein 12																		
TRRAP_HUMAN	Transformation/transcription domain-associated protein																		
YTHD2_HUMAN	YTH domain-containing family protein 2																		
MEX3A_HUMAN	RNA-binding protein MEX3A																		
GGYF2_HUMAN	GRB10-interacting GYF protein 2																		
I2BP2_HUMAN	Interferon regulatory factor 2-binding protein 2																		

WDR5_HUMAN	WD repeat-containing protein 5																		
TBL1X_HUMAN	F-box-like/WD repeat-containing protein TBL1X																		
CSK21_HUMAN	Casein kinase II subunit alpha																		
DMAP1_HUMAN	DNA methyltransferase 1-associated protein 1																		
CIC_HUMAN	Protein capicua homolog																		
DDX6_HUMAN	Probable ATP-dependent RNA helicase DDX6																		
RING1_HUMAN	E3 ubiquitin-protein ligase RING1																		
PSPC1_HUMAN	Paraspeckle component 1																		
SPF45_HUMAN	Splicing factor 45																		
RCOR1_HUMAN	REST corepressor 1																		
S23IP_HUMAN	SEC23-interacting protein																		
PF21A_HUMAN	PHD finger protein 21A																		
NUP58_HUMAN	Nucleoporin p58/p45																		
TNR6B_HUMAN	Trinucleotide repeat-containing gene 6B protein																		
MYT1_HUMAN	Myelin transcription factor 1																		
CD2B2_HUMAN	CD2 antigen cytoplasmic tail-binding protein 2																		
HDAC3_HUMAN	Histone deacetylase 3																		
SSBP3_HUMAN	Single-stranded DNA-binding protein 3																		
HIRA_HUMAN	Protein HIRA																		
YLPM1_HUMAN	YLP motif-containing protein 1																		
CCAR1_HUMAN	Cell division cycle and apoptosis regulator protein 1																		
KANL3_HUMAN	KAT8 regulatory NSL complex subunit 3																		
TNR6C_HUMAN	Trinucleotide repeat-containing gene 6C protein																		
CHD8_HUMAN	Chromodomain-helicase-DNA-binding protein 8																		
YTHD3_HUMAN	YTH domain-containing family protein 3																		
RPB1_HUMAN	DNA-directed RNA polymerase II subunit RPB1																		
YTHD1_HUMAN	YTH domain-containing family protein 1																		
PTN23_HUMAN	Tyrosine-protein phosphatase non-receptor type 23																		
HNRL1_HUMAN	Heterogeneous nuclear ribonucleoprotein U-like protein 1																		
NFIB_HUMAN	Nuclear factor 1 B-type																		
CARF_HUMAN	CDKN2A-interacting protein																		
RTCB_HUMAN	RNA-splicing ligase RtcB homolog																		
ZFR_HUMAN	Zinc finger RNA-binding protein																		
RBM12_HUMAN	RNA-binding protein 12																		
PP1RA_HUMAN	Serine/threonine-protein phosphatase 1 regulatory subunit 10																		
KDM2B_HUMAN	Lysine-specific demethylase 2B																		
RUXE_HUMAN	Small nuclear ribonucleoprotein E																		
CBX1_HUMAN	Chromobox protein homolog 1																		

MBD3_HUMAN	Methyl-CpG-binding domain protein 3																		
DCAF7_HUMAN	DDB1- and CUL4-associated factor 7																		
PP1A_HUMAN	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit																		
EMSY_HUMAN	BRCA2-interacting transcriptional repressor EMSY																		
SF3B4_HUMAN	Splicing factor 3B subunit 4																		
MINT_HUMAN	Msx2-interacting protein																		
ZN318_HUMAN	Zinc finger protein 318																		
XRN1_HUMAN	5'-3' exoribonuclease 1																		
BCOR_HUMAN	BCL-6 corepressor																		
JHD2C_HUMAN	Probable JmjC domain-containing histone demethylation protein 2C																		
AHDC1_HUMAN	AT-hook DNA-binding motif-containing protein 1																		
CDK9_HUMAN	Cyclin-dependent kinase 9																		
KMT2C_HUMAN	Histone-lysine N-methyltransferase 2C																		
NCOR2_HUMAN	Nuclear receptor corepressor 2																		
SR140_HUMAN	U2 snRNP-associated SURP motif-containing protein																		
ANR17_HUMAN	Ankyrin repeat domain-containing protein 17																		
PRC2A_HUMAN	Protein PRRC2A																		
NU153_HUMAN	Nuclear pore complex protein Nup153																		
RAE1L_HUMAN	mRNA export factor																		
PABP4_HUMAN	Polyadenylate-binding protein 4																		
NHSL2_HUMAN	NHS-like protein 2																		
PRC2B_HUMAN	Protein PRRC2B																		
RPRD2_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 2																		
ANKH1_HUMAN	Ankyrin repeat and KH domain-containing protein 1																		
BAP18_HUMAN	Chromatin complexes subunit BAP18																		
ATX2L_HUMAN	Ataxin-2-like protein																		
SNR40_HUMAN	U5 small nuclear ribonucleoprotein 40 kDa protein																		
SIN3A_HUMAN	Paired amphipathic helix protein Sin3a																		
DIDO1_HUMAN	Death-inducer obliterator 1																		
TBL1R_HUMAN	F-box-like/WD repeat-containing protein TBL1XR1																		
PRC2C_HUMAN	Protein PRRC2C																		
ZN281_HUMAN	Zinc finger protein 281																		
NIPBL_HUMAN	Nipped-B-like protein																		
PHX2A_HUMAN	Paired mesoderm homeobox protein 2A																		
CHAP1_HUMAN	Chromosome alignment-maintaining phosphoprotein 1																		
TFG_HUMAN	Protein TFG																		
SCAF4_HUMAN	SR-related and CTD-associated factor 4																		
CTBP2_HUMAN	C-terminal-binding protein 2																		

RBM10_HUMAN	RNA-binding protein 10																		
SC16A_HUMAN	Protein transport protein Sec16A																		
RAGP1_HUMAN	Ran GTPase-activating protein 1																		
HCFC1_HUMAN	Host cell factor 1																		
KMT2A_HUMAN	Histone-lysine N-methyltransferase 2A																		
IMB1_HUMAN	Importin subunit beta-1																		
EP400_HUMAN	E1A-binding protein p400																		
NUP88_HUMAN	Nuclear pore complex protein Nup88																		
AKAP8_HUMAN	A-kinase anchor protein 8																		
HTF4_HUMAN	Transcription factor 12																		
NU188_HUMAN	Nucleoporin NUP188																		
HNRPK_HUMAN	Heterogeneous nuclear ribonucleoprotein K																		
SMD3_HUMAN	Small nuclear ribonucleoprotein Sm D3																		
PHX2B_HUMAN	Paired mesoderm homeobox protein 2B																		
HNRPF_HUMAN	Heterogeneous nuclear ribonucleoprotein F																		
PRP4_HUMAN	U4/U6 small nuclear ribonucleoprotein Prp4																		
HNRPM_HUMAN	Heterogeneous nuclear ribonucleoprotein M																		
ARI1A_HUMAN	AT-rich interactive domain-containing protein 1A																		
ZN609_HUMAN	Zinc finger protein 609																		
U520_HUMAN	U5 small nuclear ribonucleoprotein 200 kDa helicase																		
PABP1_HUMAN	Polyadenylate-binding protein 1																		
SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich																		
NU214_HUMAN	Nuclear pore complex protein Nup214																		
NUP62_HUMAN	Nuclear pore glycoprotein p62																		
RBP2_HUMAN	E3 SUMO-protein ligase RanBP2																		
TurboID	TurboID TurboID with 2X FLAG and 3X GGGGS linker																		
BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF																		
UBP2L_HUMAN	Ubiquitin-associated protein 2-like																		
U5S1_HUMAN	116 kDa U5 small nuclear ribonucleoprotein component																		
PRP8_HUMAN	Pre-mRNA-processing-splicing factor 8																		
ARI1B_HUMAN	AT-rich interactive domain-containing protein 1B																		
HDAC2_HUMAN	Histone deacetylase 2																		
REQU_HUMAN	Zinc finger protein ubi-d4																		
RBM14_HUMAN	RNA-binding protein 14																		
DDX23_HUMAN	Probable ATP-dependent RNA helicase DDX23																		
RUXF_HUMAN	Small nuclear ribonucleoprotein F																		
AGFG1_HUMAN	Arf-GAP domain and FG repeat-containing protein 1																		
HDAC1_HUMAN	Histone deacetylase 1																		

RBBP4_HUMAN	Histone-binding protein RBBP4																		
CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4																		
CPSF5_HUMAN	Cleavage and polyadenylation specificity factor subunit 5																		
TADBP_HUMAN	TAR DNA-binding protein 43																		
GARS_HUMAN	Glycine--tRNA ligase																		
BUB3_HUMAN	Mitotic checkpoint protein BUB3																		
CPSF1_HUMAN	Cleavage and polyadenylation specificity factor subunit 1																		
NONO_HUMAN	Non-POU domain-containing octamer-binding protein																		
CDC5L_HUMAN	Cell division cycle 5-like protein																		
DOT1L_HUMAN	Histone-lysine N-methyltransferase, H3 lysine-79 specific																		
CNOT1_HUMAN	CCR4-NOT transcription complex subunit 1																		
CHERP_HUMAN	Calcium homeostasis endoplasmic reticulum protein																		
FUBP2_HUMAN	Far upstream element-binding protein 2																		
CCAR2_HUMAN	Cell cycle and apoptosis regulator protein 2																		
CDIPT_HUMAN	CDP-diacylglycerol--inositol 3-phosphatidyltransferase																		
RBBP7_HUMAN	Histone-binding protein RBBP7																		
RPB2_HUMAN	DNA-directed RNA polymerase II subunit RPB2																		
MED23_HUMAN	Mediator of RNA polymerase II transcription subunit 23																		
ROA3_HUMAN	Heterogeneous nuclear ribonucleoprotein A3																		
CSTF2_HUMAN	Cleavage stimulation factor subunit 2																		
FUBP1_HUMAN	Far upstream element-binding protein 1																		
CND2_HUMAN	Condensin complex subunit 2																		
FKB15_HUMAN	FK506-binding protein 15																		
WDR82_HUMAN	WD repeat-containing protein 82																		
TRI33_HUMAN	E3 ubiquitin-protein ligase TRIM33																		
SEC13_HUMAN	Protein SEC13 homolog																		
NAB1_HUMAN	NGFI-A-binding protein 1																		
UBN2_HUMAN	Ubinuclein-2																		
BCORL_HUMAN	BCL-6 corepressor-like protein 1																		
AKT1_HUMAN	RAC-alpha serine/threonine-protein kinase																		
HMGX4_HUMAN	HMG domain-containing protein 4																		
F193A_HUMAN	Protein FAM193A																		
N4BP2_HUMAN	NEDD4-binding protein 2																		
CPSF6_HUMAN	Cleavage and polyadenylation specificity factor subunit 6																		
SF3A1_HUMAN	Splicing factor 3A subunit 1																		
PUM1_HUMAN	Pumilio homolog 1																		
RLA2_HUMAN	60S acidic ribosomal protein P2																		
ENOA_HUMAN	Alpha-enolase																		

HNRPC_HUMAN	Heterogeneous nuclear ribonucleoproteins C1/C2																	
CEP85_HUMAN	Centrosomal protein of 85 kDa																	
TBX3_HUMAN	T-box transcription factor TBX3																	
MATR3_HUMAN	Matrin-3																	
EP300_HUMAN	Histone acetyltransferase p300																	
RBM20_HUMAN	RNA-binding protein 20																	
PUM2_HUMAN	Pumilio homolog 2																	
PROX1_HUMAN	Prospero homeobox protein 1																	
YETS2_HUMAN	YEATS domain-containing protein 2																	
LARP4_HUMAN	La-related protein 4																	
PATL1_HUMAN	Protein PAT1 homolog 1																	
RUVB1_HUMAN	RuvB-like 1																	
ROA1_HUMAN	Heterogeneous nuclear ribonucleoprotein A1																	
ILF2_HUMAN	Interleukin enhancer-binding factor 2																	
RALY_HUMAN	RNA-binding protein Raly																	
RUVB2_HUMAN	RuvB-like 2																	
ZN638_HUMAN	Zinc finger protein 638																	
ZCHC3_HUMAN	Zinc finger CCHC domain-containing protein 3																	
PRPF3_HUMAN	U4/U6 small nuclear ribonucleoprotein Prp3																	
RENT1_HUMAN	Regulator of nonsense transcripts 1																	
ADNP_HUMAN	Activity-dependent neuroprotector homeobox protein																	
SUGP2_HUMAN	SURP and G-patch domain-containing protein 2																	
RSMB_HUMAN	Small nuclear ribonucleoprotein-associated proteins B and B'																	
EPN4_HUMAN	Clathrin interactor 1																	
SRRM2_HUMAN	Serine/arginine repetitive matrix protein 2																	
AF17_HUMAN	Protein AF-17																	
BRCC3_HUMAN	Lys-63-specific deubiquitinase BRCC36																	

CHAPTER 6

CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions

This dissertation has described several O-GlcNAc Transferase TPR interactomes. The scope of this work consists of two different cancer cell lines (HeLa and SHSY5Y) and perturbations including different glucose levels, serum starvation, and depolarization. In addition, we have described interactomes with OGT enzymes carrying mutations causal for X-Linked Intellectual Disability in SHSY5Y cells. In summarizing such a large and complex dataset, for both the conclusion and future directions statements, the discussion will be focused on either wild-type OGT or the XLID variants.

The role of the TPR domain in OGT substrate selection

One of the most compelling results from the work presented here comes from our use of just the TPR domain of OGT to identify interactors. We have identified many proteins that are known to be O-GlcNAc modified and/or are known OGT interactors that interact exclusively with the TPR domain of OGT without the presence of the catalytic domain. That these substrate proteins interact with just the TPR domain is an excellent reinforcer of the hypothesis that the TPR domain of OGT is responsible for substrate selection. In the future, it will be interesting to repeat studies like these with the full-length OGT and just the catalytic domain of OGT to see how many of these TPR interactors still interact and if any novel catalytic domain interacting proteins can be identified.

The "basal" OGT interactome

Given the variety of cellular states we have explored here, it is possible to begin to describe an absolute "basal" OGT TPR interactome. This interactome would consist of the

Table 6.1 Constitutive OGT TPR interactors across all studied conditions

Accession	Description
ANR17_HUMAN	Ankyrin repeat domain-containing protein 17
CCAR1_HUMAN	Cell division cycle and apoptosis regulator protein 1
DIDO1_HUMAN	Death-inducer obliterator 1
RBP2_HUMAN	E3 SUMO-protein ligase RanBP2
TBL1R_HUMAN	F-box-like/WD repeat-containing protein TBL1XR1
QSER1_HUMAN	Glutamine and serine-rich protein 1
SET1A_HUMAN	Histone-lysine N-methyltransferase SETD1A
HCFC1_HUMAN	Host cell factor 1
KDM3B_HUMAN	Lysine-specific demethylase 3B
KDM1A_HUMAN	Lysine-specific histone demethylase 1A
NU153_HUMAN	Nuclear pore complex protein Nup153
NU214_HUMAN	Nuclear pore complex protein Nup214
NUP88_HUMAN	Nuclear pore complex protein Nup88
NUP62_HUMAN	Nuclear pore glycoprotein p62
NCOR1_HUMAN	Nuclear receptor corepressor 1
NCOR2_HUMAN	Nuclear receptor corepressor 2
NUP58_HUMAN	Nucleoporin p58/p45
SIN3A_HUMAN	Paired amphipathic helix protein Sin3a
PRR12_HUMAN	Proline-rich protein 12
PRC2C_HUMAN	Protein PRRC2C
RAGP1_HUMAN	Ran GTPase-activating protein 1
RPRD2_HUMAN	Regulation of nuclear pre-mRNA domain-containing protein 2
RBM27_HUMAN	RNA-binding protein 27
WNK1_HUMAN	Serine/threonine-protein kinase WNK1
UBP2L_HUMAN	Ubiquitin-associated protein 2-like
OGT1_HUMAN	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit
YLPM1_HUMAN	YLP motif-containing protein 1
ZC3HE_HUMAN	Zinc finger CCCH domain-containing protein 14
ZN281_HUMAN	Zinc finger protein 281
ZN609_HUMAN	Zinc finger protein 609

proteins that constitutively interact with OGT – e.g. the proteins that are present in all of our wild-type interactomes, regardless of cell type or state. These proteins may perform some universal essential function that is mediated by their interaction with OGT. It is important to note that this interactome will likely be pared down further by future studies since OGT is ubiquitously expressed[2] and involved in many more cellular functions than we explored here.

The "basal" OGT TPR interactome from the conditions we studied consists of 31 proteins (**Table 6.1**). Several of these proteins are well-known and previously studied OGT interactors, especially HCF-1[19] and Sin3a[5]. Nuclear pore proteins, which are well established as OGT interactors[120,174], also dominate the list. It is interesting to note that many of these interactions which appear to be highly stable and ubiquitous were some of the few previously captured OGT interactors. Several of the proteins in this list are also nucleic acid binding proteins (DIDO1, RPRD2, RBM27, ZC3HE, ZN281, ZN609) and chromatin remodeling proteins (SET1A, KDM3B, KDM1A, NOCR1, NCOR2, SIN3A), which further cements one of OGT's primary cellular roles as a high-level regulator of other regulatory proteins. When assessing the role of OGT protein interaction in the cell, these constitutive interactors should be considered high priority because their functional interaction with OGT may apply universally. Additionally, future interactome studies will aid in condensing this list and confirming the importance of the proteins identified here.

The diversity of the OGT interactome

It has been thoroughly established that thousands of nuclear and cytosolic proteins are O-GlcNAc modified[22]. When also considering the possibility of OGT partner proteins that interact with OGT but may not be functionally O-GlcNAc modified, the diversity of OGT protein interactors becomes immensely vast. Despite this, our list of OGT protein interactors per interactome is comparatively small – we identify approximately 100-200 proteins per interactome. It may seem that given the sheer number of O-GlcNAc modified proteins that we may not be achieving sufficient depth. However, our comparison of OGT TPR interactomes in two cell lines, HeLa and

SHSY5Y, suggests that all OGT interactors may not be fully represented within a single cell type. Rather, the vast abundance of O-GlcNAc modified proteins (which are by definition OGT interactors, and likely TPR interactors[5,19,20], although not certainly[54]) may be achieved by significantly differing interactomes across cell types. The HeLa and SHSY5Y OGT TPR interactomes only have around 30% of interactors in common. When considering the diversity of cell types in which OGT is expressed, it is easy to see how over a thousand protein interactors globally could be achieved.

The degree to which the OGT interactome varies with different cellular states has been a significant question in the field for some time. It stands to reason that with only one mammalian OGT[4] that protein interaction with effector substrates specific to that cellular state would be a major driver of OGT's diverse cellular functions. For most of the cellular conditions we've studied, however, we don't observe large changes in interactome like we may have expected. This could be for several reasons. We have only examined three cellular states here: differential glucose concentrations, serum starvation, and depolarization. It is possible that the OGT interactome is more dynamic under different conditions. Indeed, one other OGT interactome describes significant changes in the interactome under oxidative stress[155], and previous links to OGT as a stress response protein[7] support cellular stress response as a major OGT function. Additionally, the cell lines that we utilized in this study are both cancer cell lines. OGT plays a role in cancer progression and prognosis[23], although this role depends on cancer type, so caution must be taken when attempting to generalize about typical cellular behavior from these cell lines.

In order to understand why the OGT interactome is less dynamic than anticipated, the way OGT affects cellular function may need to be considered differently. O-GlcNAc is often compared to protein phosphorylation, and the two modifications do share many similarities, but protein phosphorylation often serves as more of a cellular "switch", i.e. a modification that has a rapid, singular effect on a protein with an immediate effector function. We see from our interactomes

that OGT doesn't tend to interact with those proteins (e.g. proteins involved in signaling cascades). Instead, we see a strong enrichment toward interaction with proteins that are high-level cellular regulators: chromatin remodelers, transcription factors, nuclear pore proteins. This raises the possibility that OGT doesn't need to alter its interaction with many proteins in order to respond to changes in cellular status. In Chapter 2, we characterized OGT as a "regulator of the regulators" and the scope of its TPR interactomes supports this idea. OGT may be involved in more subtle, top-level regulation that occurs across cellular regulators, thus enabling broader cellular changes with a more subtle shift. These shifts many be more difficult to parse out than, for example, rapid phosphorylation of a signaling protein, so future research focused on downstream effectors rather than the direct outcome of OGT interaction may be useful.

Another intriguing possibility to address this issue comes from a study which shows that OGT constitutively interacts with the protein neurofilament-H, but only after phosphorylated p38 interacts with OGT is neurofilament-H O-GlcNAc modified[54]. It is possible that this sort of phenomenon is common with OGT – that the interactome itself is more static, but O-GlcNAc modification of interacting proteins could be induced by additional stimulus.

The largest difference in OGT interactome that we observed was under differential serum conditions in HeLa cells, where brief stimulation with serum seems to cause a large decrease in proteins interacting with OGT. The rest of the interactome changes are more modest, with just a handful of interactors being altered, or in the case of differential glucose conditions in HeLa cells, no interactors being altered. This points toward the essentiality of not thinking of the OGT interactome as a standard set of predictable proteins and functions, but rather that its interacting proteins and even roles may be quite divergent across cell types. As discussed below, additional research into how and when the OGT interactome is altered in additional cell types and conditions will be essential to continue to parse out just how diverse OGT's function and functional effectors are.

The OGT interactome in disease states

One major driver for this research is the knowledge of the diseases that OGT and O-GlcNAc have been connected to, without a mechanistic link being established. Diseases like Alzheimer's Disease[130], type 2 diabetes[175], cancer[23], and X-Linked Intellectual Disability[38] are all correlated to OGT and O-GlcNAc, but a mechanistic understanding of their roles is essential in order to inform potential therapeutics. One essential step in establishing those mechanisms is in understanding what proteins OGT interacts with. Simply establishing basal protein interactors in cancer cells, as we've done here, may allow for future comparison with the basal OGT interactome in non-cancer cells and help to identify OGT interactors that may be drivers of the cancer phenotype. It is also possible that our glucose studies in HeLa cells could provide insight into diabetes; however, as discussed previously, HeLa cells being highly adapted to grow in high glucose limits our confidence in the utility of this data for this purpose.

A primary driver for our interactome studies has been the role of OGT in X-Linked Intellectual Disability, since OGT TPR variants causal for XLID are biochemically identical to the wild-type enzyme. Given that these mutations are localized to the TPR domain of *OGT*, it stands to reason that the mutations could alter protein-protein interaction in a way that leads to this neurodevelopmental phenotype. When assessing our interactomes for protein candidates for future study, we focused on proteins that 1) differentially interacted with a majority of XLID variants and 2) were previously linked to neurodevelopmental disorders, especially intellectual disability disorders. While it is tempting to focus on proteins that are expressed uniquely in mature neurons, it is more likely that the defect occurs during early neurodevelopment, since OGT XLID patients present not only with intellectual disability but also with dysmorphic facial features and other developmental defects. Therefore, as we screened our differential interactors to identify mechanistic candidates, we have largely focused on proteins with possible neurodevelopmental

Table 6.2 Interaction of proteins of interest to XLID across conditions

Dark Gray - identified only in TPR-TurboID, not present in eGFP-TurboID
Light gray - identified in both TPR-TurboID and eGFP-TurboID, significantly enriched in TPR-TurboID

Condition	Tet2	Piccolo	Dim1	ZC3H1
HeLa Basal/High Glucose				
HeLa 4mM Glucose				
HeLa 1mM Glucose				
HeLa Serum				
HeLa Serum Starve				
HeLa Serum Pulse				
SHSY5Y Basal				
SHSY5Y Serum				
SHSY5Y Serum Starve				
SHSY5Y Serum Pulse				
SHSY5Y Depolarize				

roles. Those proteins are Tet2, Piccolo, Dim1, and ZC3H1, and suggestions for further research into these candidate proteins is discussed below.

For these proteins of high interest, we also examined their interaction with the OGT TPR across conditions studied throughout this dissertation (**Table 6.2**). Tet2 interacts with OGT under all conditions in HeLa cells but is only identified in our SHSY5Y basal interactome. ZC3H1 is similarly constitutive, interacting under all HeLa conditions and the basal and depolarized SHSY5Y interactome. Piccolo only interacts in SHSY5Y cells under depolarization, which is unsurprising given its neuron-specific expression and synaptic signaling function. Dim1 only interacts in our SHSY5Y depolarized interactome, which may point toward a highly cell- and function-specific interaction. It is important to note here that none of these four interactors appear in our SHSY5Y serum interactomes in any condition. For Tet2, this may be due to a low rate of overall interaction in SHSY5Y cells as reflected by low spectral counts where it is identified. For Piccolo and Dim1, it is possible that they both truly only interact in SHSY5Y cells under depolarization since the stress of serum starvation doesn't induce the interaction. That ZC3H1 and Tet2 don't appear in the SHSY5Y serum interactome despite interacting under basal conditions points to the inherent variability of a proteomic dataset and why it is so essential to validate novel interactions with other methods, as discussed below.

Future Directions

Recommendations for future study of the wild-type OGT interactome

In continuing study of the wild-type OGT interactome, both additional global study and more targeted study building off of the interactomes presented here will be useful.

Continuing to explore the breadth of OGT interactors across different cell types and states is essential to further understand the diversity of the OGT interactome. Although more technically challenging, assessing interactomes in primary cells would be highly beneficial toward understanding what proteins OGT interacts with under basal conditions. Pancreatic β-cells and

mature cortical neurons are of high interest due to OGT's high expression and multiple disease connections across those cell types. Examining additional cellular states will be beneficial as well. Focusing more narrowly on the effects of serum stimulation by adding specific extracellular signaling molecules such as ERK or insulin to see the downstream effects on the OGT interactome will be useful in determining what role OGT plays in those specific pathways. Expanding on the depolarization studies here, performing similar experiments in cells with fully functioning synapses may further elucidate what neuron-specific proteins OGT interacts with and how those function during synaptic signaling.

In addition, although the TurboID method is useful for identifying protein interactors on a short time scale and including transient interactors, it is not without its limitations. One challenge of this method is that nonspecific biotin labeling of highly abundant and "sticky" proteins inherently occurs, which results in suppression of lower abundance interactors. We have attempted to remove these nonspecific interactors by the inclusion of a negative control TurboID fusion protein, but this doesn't change that many spectral IDs in the mass spectrometer come from highly abundant proteins that are biotin labeled, such as myosins. Cross-referencing the interactomes described here with another interaction approach would be useful in validating these interactors and may also allow for more depth. One such method is an immunoprecipitation-based method recently used to describe an OGT interactome in mouse embryonic fibroblasts[155]. The basal interactome identified in this reference differs greatly from what we have identified here, although this may be due in part to the different cell types utilized. Cross-referencing this interactome method with TurboID may allow for greater depth of interactor identification.

Finally, alterations to the method used here may contribute to higher confidence in interactor identification. In our statistical analysis, we have been very careful to thoroughly vet differential interactors to ensure that we reduce the incidence of false positives as much as possible. This is necessary in large part due to the inherent variability in a multi-stage sample

prep whose quantification relies on spectral counting. Altering our method to use pooled sample prep with methods such as SILAC[176] or TMT[177] may help to reduce this variability and allow for more confidence in determination of differential interactors.

For targeted study of OGT interactors, the first necessary next step is to validate several novel interactors with an orthogonal technique. This is not necessarily trivial. Since many of these interactions are likely to be transient, it may be difficult to capture the interaction with a traditional method like immunoprecipitation. Alternative methods like BiFC[89] or FRET[88] may be useful for validating interactions. From there, determining the effect that OGT interaction has on interactors of interest is of high priority. The approaches utilized here will vary depending on the function of the interactor of interest, but considering OGT's interaction with proteins involved in well-studied systems like transcriptional regulation, many applicable methods are available for this purpose. A possible workflow for determining the effect of OGT on a protein interactor would be biochemically mapping the site of the interaction, then expressing an OGT interaction-deficient variant of the interactor and assessing how this affects its downstream function, localization, and the overall cellular phenotype.

One interactor of interest for future study due to its identification as a differential OGT interactor across several conditions is CDK9. While a bulk of the OGT interactome seems to be fairly stable, CDK9 may represent an OGT interactor that is more labile and condition dependent. CDK9 is a kinase that is an important component of each step in RNA Pol II transcription, phosphorylating the Pol II CTD during pre-initiation, transcription, and transcription termination. OGT itself is part of the pre-initiation complex (PIC) and is essential for the assembly of the PIC[6]. In fact, crosstalk of mutually exclusive O-GlcNAc modification or phosphorylation of the RNA Pol II CTD has been shown to be essential in transcription initiation[59]. What hasn't been previously studied is OGT directly interacting with transcriptional regulatory kinases such as CDK9, despite reports that CDK9 is O-GlcNAc modified[178]. Determining the functional outcome of this interaction

may provide additional insight into both proteins' roles in transcriptional regulation. It is possible that O-GlcNAc modification of CDK9 has some effect on its function, but also that CDK9 could phosphorylate OGT, resulting in a change in its transcription-regulating function. Determining the sites and abundance of modification on either protein, and how this affects their activity toward the CTD of RNA Pol II would be an interesting first step in discovering the regulatory effects of this interaction.

Recommendations for future study of OGT TPR XLID

We have identified several candidate proteins that may underlie the mechanism of OGT XLID. Additional validation is necessary to both confirm these proteins as differential interactors and to further elucidate the underlying mechanism. First, these interactors must be independently validated to interact with the wild-type OGT and to fail to interact with XLID variants. Immunoprecipitation or BiFC may be an excellent method to accomplish this, although this may be a challenge for some of the interacting proteins as several of them are large (especially piccolo which is 560kD), which may make isolation and/or expression challenging. Additionally, three of the four protein candidates here only interacted with OGT in SHSY5Y cells under depolarization, so utilizing this or another stress condition may be essential in order to recapitulate the interaction. However, the first "gold standard" for validating these interactions as differential will be confirming their loss of interaction with XLID variants biochemically.

From there, and if this can't be achieved, looking at downstream aspects of these proteins' functions provide clues into whether or not their function is altered with OGT XLID variants. We fortunately already have in hand human embryonic stem cells carrying each OGT XLID mutation[27], so study of the downstream effector functions of each of these proteins of interest will be achievable in this homozygous system. Additionally, these cells can be differentiated down neural lineages to assess various functional aspects during neurodevelopment.

For Tet2, the clear next step is to look at methylated cytosine in ES cells carrying the XLID mutations versus wild-type cells. Global levels of 5-methylcytosine and 5-hydroxymethylcytosine (5-hmC) (the product of Tet protein reactions) are easily assessable with immune detection[179,180]. If global 5-hmC is significantly altered in XLID variant-carrying cells, this would point strongly toward a Tet protein dysfunction. Considering the possibility that a defect in OGT/Tet2 interaction causes an issue with OGT chromatin targeting, this could be assessed by looking at global histone O-GlcNAc modification, since this is the known outcome of OGT and Tet2 protein interaction[20]. Looking at OGT chromatin occupancy on a global scale, as discussed below, may also be useful.

In assessing a possible role for piccolo, first determining where in neurodevelopment it is expressed is essential. If it is expressed exclusively in more mature neurons, differentiating ES cells carrying XLID variants to neurons and assessing the organization of their synapses as compared to wild-type may hint toward a dysfunction with piccolo, although many proteins contribute to synaptic organization. Previous piccolo knockouts in mice have had a limited observable phenotype, although synaptic vesicle clustering may be affected[181]. Additionally, a piccolo knockdown study revealed that neurons lacking piccolo exhibit poor F-actin polymerization[182]. Looking for these cellular phenotypes in XLID neurons may corroborate a connection to the piccolo protein. However, this does not necessarily directly connect the OGT mutation to piccolo's function. If piccolo can be confirmed to have reduced interaction with OGT XLID variants, mapping the exact site of the interaction biochemically and looking for altered piccolo O-GlcNAc modification may allow for more targeted look at the contribution of OGT to piccolo's function. Practically, this would look like replacing piccolo with an OGT-interaction-deficient variant of piccolo in stem cells and seeing if this recapitulates the phenotype observed in the cells carrying the XLID OGT. This sort of reciprocal assay may be very useful in parsing out what outcomes are passenger vs driver of the XLID phenotype.

Dim1, due to its global function and limited previous research, is more challenging to determine a specific assay to assess its downstream function. One previous paper reported that Dim1 knockout HEK293T cells exhibit decreased 40S assembly and decreased global protein synthesis[167]. If Dim1 can be confirmed to have reduced interaction with OGT XLID variants biochemically, assessing the polysome and global protein translation in XLID ES cells and derived neural lineage cells, especially neural precursor or neural crest, may provide additional evidence toward Dim1 as a mechanistic protein underlying XLID. However, protein translation is a very complex process with a myriad of essential proteins involved, so this evidence alone would not be sufficient to solely implicate Dim1. As with piccolo, determining the exact region of OGT interaction and/or where Dim1 is O-GlcNAc modified will be essential to fully determine if Dim1 is responsible for the XLID phenotype. If this is the case, replacing Dim1 with an OGT-interaction-deficient Dim1 in wild-type cells should recapitulate the phenotype observed in OGT XLID mutation-carrying cells.

A downstream phenotype for ZC3H1 dysfunction is easier to identify due to previous work on its function in embryonic stem cell development[170]. In this manuscript, the authors reported that ZC3H1 knockout ES cells exhibited increased expression of PAXT complex target RNAs and increased expression of polycomb repressive complex 2 transcripts. These molecular phenotypes will be simple to examine in XLID ES cells and derived neural lineage cells to determine if an abnormality in ZC3H1 function is observed.

It is important to note that while we are focusing on single protein candidates here, it is very possible that the mechanism underlying OGT XLID is not so simple as one single failed protein interaction. For this reason, continuing to study the XLID phenotype on a global scale may still provide useful insights into what aspect of OGT function is dysregulated. For example, we have already observed global transcriptome changes in XLID variant ES cells[26,27]. This points to a possible high-level dysregulation of transcription which could be caused by a variety of

interactions with OGT. Assessing chromatin status may be useful in determining where this defect originates. Using a technique like ATAC-Seq[183] to determine if chromatin is uniquely accessible or inaccessible on a global scale in XLID ES cells may reveal a larger-scale chromatin regulation issue. Additionally, using a method like ChIP-Seq to determine where OGT is bound to chromatin in wild-type versus XLID cells will be useful to determine if OGT localization contributes to transcriptional changes.

Finally, although the SHSY5Y cell line we chose to use here has been used extensively to study intellectual disability[33–35], it is possible that the XLID relevant interactions simply cannot be accurately captured from this cell line. Therefore, repeating these interactome studies in embryonic stem cells and during early neural differentiation may be an essential step in correctly parsing out what OGT interactions are altered. Performing these interactomes both with TurboID (made possible by stem-cell specific transfection reagents, although still challenging due to rapid silencing of transgenes in ES cells) and an immunoprecipitation method will be ideal for this. It will also be important to assess interactomes in both wild-type cells (expressing each XLID variant) and in the XLID mutation-carrying cells (expressing each corresponding XLID variant, or simply immunoprecipitation the endogenous OGT and its interactors), because of the significant transcriptomic changes we observe in the XLID cells. Simply assessing interaction in the XLID cells may result in missed identifications due to a relevant protein having altered expression. Conversely, only utilizing the wild-type cells may have the effect of reducing detection of interactions that are altered with the XLID variant enzymes due to their interaction with the endogenous, wild-type OGT. The most thorough method to assess the question of OGT XLID interactomes during early neurodevelopment will result in six interactome sets: Immunoprecipitation of OGT variants from XLID ES cells, TurboID XLID variants expressed in wild-type ES cells, TurboID XLID variants expressed in each corresponding XLID ES cell, and each of the previous repeated in neural precursor or neural crest cells.

In conclusion, through a series of descriptive studies we have provided necessary insight to the O-GlcNAc field on what proteins OGT interacts with, under what conditions, and how those interactions might play a role in OGT-related disease. While significant additional work is necessary to put these interactions into a molecular, functional context, we hope that these interactomes being made available will allow for many jumping-off-points to spur additional research and understanding of this important glycosyltransferase's function on a cellular level.

REFERENCES

1. Haltiwanger, R. S., Blomberg, M. A. & Hart, G. W. Glycosylation of nuclear and cytoplasmic proteins. *J. Biol. Chem.* **1992,** *267,* 9005–9013.
2. Kreppel, L. K., Blomberg, M. A. & Hart, G. W. Dynamic Glycosylation of Nuclear and Cytosolic Proteins. *J. Biol. Chem.* **1997,** *272,* 9308–9315.
3. Lubas, W. A., Frank, D. W., Hanover, J. A. & Krause, M. O -Linked GlcNAc Transferase Is a Conserved Nucleocytoplasmic Protein Containing Tetratricopeptide Repeats. *J. Biol. Chem.* **1997,** *272,* 9316–9324.
4. Shafi, R., Iyer, S. P. N., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W. & Marth, J. D. The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc. Natl. Acad. Sci.* **2000,** *97,* 5735–5739.
5. Yang, X., Zhang, F. & Kudlow, J. E. Recruitment of O -GlcNAc Transferase to Promoters by Corepressor mSin3A : Coupling Protein O -GlcNAcylation to Transcriptional Repression. *Cell* **2002,** *110,* 69–80.
6. Kelly, W. G., Dahmus, M. E. & Hart, G. W. RNA polymerase II is a glycoprotein. *J. Biol. Chem.* **1993,** *268,* 10416–10424.
7. Martinez, M. R., Dias, T. B., Natov, P. S. & Zachara, N. E. Stress-Induced O-GlcNAcylation, an Adaptive Process of Injured Cells. *Biochem. Soc. Trans.* **2017,** *45,* 237–249.
8. Vaidyanathan, K. & Wells, L. Multiple tissue-specific roles for the O-GlcNAc post-translational modification in the induction of and complications arising from type II diabetes. *J. Biol. Chem.* **2014,** *289,* 34466–34471.

9. Varki, A. Biological roles of glycans. *Glycobiology* **2017,** *27,* 3–49.

10. Hart, G. W. Dynamic O-Linked Glycosylation of Nuclear and Cytoskeletal Proteins. *Annu. Rev. Biochem.* **1997,** *66,* 315–335.

11. Leney, A. C., El Atmioui, D., Wu, W., Ovaa, H. & Heck, A. J. R. Elucidating crosstalk mechanisms between phosphorylation and O-GlcNAcylation. *Proc. Natl. Acad. Sci.* **2017,** 201620529. doi:10.1073/pnas.1620529114

12. Wang, Z., Gucek, M. & Hart, G. W. Cross-talk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated O-GlcNAc. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105,* 13793–8.

13. Wang, Z., Udeshi, N. D., Slawson, C., Compton, P. D., Cheung, W. D., Shabanowitz, J., Hunt, D. F. & Hart, G. W. Extensive Crosstalk Between O-GlcNAcylation and Phosphorylation Regulates Cytokinesis. *Sci Signal* **2010,** *3,* 1–22.

14. Tarrant, M. K., Rho, H., Xie, Z., Jiang, Y. L., Gross, C., Culhane, J. C., Yan, G., Qian, J., Ichikawa, Y., Matsuoka, T., Zachara, N., Etzkorn, F. A., Hart, G. W., Jeong, J. S., Zhu, H. & Cole, P. A. Regulation of CK2 by Phosphorylation and O-GlcNAcylation Revealed by Semisynthesis. *Nat. Chem. Biol.* **2012,** *8,* 262–269.

15. Oruganty, K. & Kannan, N. Design principles underpinning the regulatory diversity of protein kinases. *Philos. Trans. R. Soc. B Biol. Sci.* **2012,** *367,* 2529–2539.

16. Lazarus, M. B., Nam, Y., Jiang, J., Sliz, P. & Walker, S. Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature2* **2011,** *469,* 564–567.

17. D'Andrea, L. D. & Regan, L. TPR proteins: The versatile helix. *Trends Biochem. Sci.* **2003,** *28,* 655–662.

18. Kreppel, L. K. & Hart, G. W. Regulation of a Cytosolic and Nuclear O-GlcNAc Transferase. *J. Biol. Chem.* **1999,** *274,* 32015–32022.

19. Capotosti, F., Guernier, S., Lammers, F., Waridel, P., Cai, Y., Jin, J., Conaway, J. W.,

Conaway, R. C. & Herr, W. O-GlcNAc transferase catalyzes site-specific proteolysis of HCF-1. *Cell* **2011,** *144,* 376–388.

20. Chen, Q., Chen, Y., Bian, C., Fujiki, R. & Xiaochun, Y. Tet2 promotes histone O-GlcNAcylation during gene transcription. *Nature* **2013,** *493,* 561–564.
21. Ito, R., Katsura, S., Shimada, H., Tsuchiya, H., Hada, M., Okumura, T., Sugawara, A. & Yokoyama, A. TET3-OGT interaction increases the stability and the presence of OGT in chromatin. *Genes to Cells* **2013,** *19,* 52–65.
22. Wulff-Fuentes, E., Berendt, R. R., Massman, L., Danner, L., Malard, F., Vora, J., Kahsay, R. & Olivier-Van Stichelen, S. The human O-GlcNAcome database and meta-analysis. *Sci. Data* **2021,** *8,* 1–11.
23. Hanover, J. A., Chen, W. & Bond, M. R. O- GlcNAc in cancer : An Oncometabolism-fueled vicious cycle. *J. Bioenerg. Biomembr.* **2018,** *50,* 155–173.
24. Yuzwa, S. A., Shan, X., MacAuley, M. S., Clark, T., Skorobogatko, Y., Vosseller, K. & Vocadlo, D. J. Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation. *Nat. Chem. Biol.* **2012,** *8,* 393–399.
25. Jensen, R. V., Zachara, N. E., Nielsen, P. H., Kimose, H. H., Kristiansen, S. B. & Bøtker, H. E. Impact of O-GlcNAc on cardioprotection by remote ischaemic preconditioning in non-diabetic and diabetic patients. *Cardiovasc. Res.* **2013,** *97,* 369–378.
26. Vaidyanathan, K., Niranjan, T., Selvan, N., Teo, C. F., May, M., Patel, S., Weatherly, B., Skinner, C., Opitz, J., Carey, J., Viskochil, D., Gecz, J., Shaw, M., Peng, Y., Alexov, E., Wang, T., Schwartz, C. & Wells, L. Identification and Characterization of a Missense Mutation in the O-GlcNAc Transferase Gene that Segregates with X-Linked Intellectual Disability. *J. Biol. Chem.* **2017,** *292,* jbc.M116.771030.
27. Selvan, N., George, S., Serajee, F. J., Shaw, M., Hobson, L., Kalscheuer, V., Prasad, N., Levy, S. E., Taylor, J., Aftimos, S., Schwartz, C. E., Huq, A. M., Gecz, J. & Wells, L. O-

GlcNAc transferase missense mutations linked to X-linked intellectual disability deregulate genes involved in cell fate determination and signaling. *J. Biol. Chem.* **2018,** *293,* 10810–10824.

28. Pravata, V. M., Muha, V., 1+, G., Ferenbach, A. T., Kakade, P., Vandadi, V., Wilmes, A. C., Borodkin, V. S., Joss, S., Stavridis, M. P. & Van Aalten, D. M. F. Catalytic deficiency of O-GlcNAc transferase leads to X-linked intellectual disability. **2019,** *116,* 14961–14970.
29. Pravata, V. M., Gundogdu, M., Bartual, S. G., Ferenbach, A. T. & Wojcik, M. H. A missense mutation in the catalytic domain of O -GlcNAc transferase links perturbations in protein O -GlcNAcylation to X-linked intellectual disability. **2019,** 1–11. doi:10.1002/1873-3468.13640
30. Raymond, F. L. X linked mental retardation: A clinical guide. *J. Med. Genet.* **2006,** *43,* 193–200.
31. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **2012,** *196,* 801–810.
32. Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., Feldman, J. L., Perrimon, N. & Ting, A. Y. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **2018,** *36,* 880–898.
33. Hurst, S. E., Liktor-Busa, E., Moutal, A., Parker, S., Rice, S., Szelinger, S., Senner, G., Hammer, M. F., Johnstone, L., Ramsey, K., Narayanan, V., Perez-Miller, S., Khanna, M., Dahlin, H., Lewis, K., Craig, D., Wang, E. H., Khanna, R. & Nelson, M. A. A novel variant in TAF1 affects gene expression and is associated with X-linked TAF1 intellectual disability syndrome. *Neuronal Signal.* **2018,** *2,* 1–17.
34. Kurosaki, T., Imamachi, N., Pröschel, C., Mitsutomi, S., Nagao, R., Akimitsu, N. & Maquat, L. E. Loss of the fragile X syndrome protein FMRP results in misregulation of

nonsense-mediated mRNA decay. *Nat. Cell Biol.* **2021,** *23,* 40–48.

35. Loi, M., Trazzi, S., Fuchs, C., Galvani, G., Medici, G., Gennaccaro, L., Tassinari, M. & Ciani, E. Increased DNA Damage and Apoptosis in CDKL5-Deficient Neurons. *Mol. Neurobiol.* **2020,** *57,* 2244–2262.
36. Torres, C. R. & Hart, G. W. Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *J. Biol. Chem.* **1984,** *259,* 3308–3317.
37. Hardville, S. & Hart, G. Nutrient Regulation of Signaling, Transcription, and Cell Physiology by O-GlcNAcylation. *Cell Metab.* **2014,** *20,* 208–213.
38. Pravata, V. M., Omelková, M., Stavridis, M. P., Desbiens, C. M., Stephen, H. M., Lefeber, D. J., Gecz, J., Gundogdu, M., Õunap, K., Joss, S., Schwartz, C. E., Wells, L. & van Aalten, D. M. F. An intellectual disability syndrome with single-nucleotide variants in O-GlcNAc transferase. *Eur. J. Hum. Genet.* **2020,** *28,* 706–714.
39. Hart, G. W. Nutrient regulation of signaling and transcription. *J. Biol. Chem.* **2019,** *294,* 2211–2231.
40. Lubas, W. a & Hanover, J. a. Functional Expression of O -linked GlcNAc Transferase. *J. Biol. Chem.* **2000,** *275,* 10983–10988.
41. Trapannone, R., Mariappa, D., Ferenbach, A. T. & van Aalten, D. M. F. Nucleocytoplasmic human O-GlcNAc transferase is sufficient for O-GlcNAcylation of mitochondrial proteins. *Biochem. J.* **2016,** *473,* 1693–1702.
42. Sacoman, J. L., Dagda, R. Y., Burnham-Marusich, A. R., Dagda, R. K. & Berninsone, P. M. Mitochondrial O-GlcNAc Transferase (mOGT) regulates mitochondrial structure, function, and survival in HeLa cells. *J. Biol. Chem.* **2017,** *292,* 4499–4518.
43. Love, D. C., Kochran, J., Cathey, R. L., Shin, S.-H. & Hanover, J. A. Mitochondrial and nucleocytoplasmic targeting of O-linked GlcNAc transferase. *J. Cell Sci.* **2003,** *116,* 647–

654.

44. Liu, L., Li, L., Ma, C., Shi, Y., Liu, C., Xiao, Z., Zhang, Y., Tian, F., Gao, Y., Zhang, J., Ying, W., WAng, P. G. & Zhang, L. O-GlcNAcylation of Thr12/Ser56 in short form O-GlcNAc transferase (sOGT) regulates its substrate selectivity. *J. Biol. Chem.* **2019,** *294,* 16620–16633.

45. Haltiwanger, R. . S. ., Holt, G. . D. . & Hart, G. . W. . Enzymatic addition of O-GlcNAc to Nuclear and Cytoplasmic Proteins. *J. Biol. Chem.* **1990,** *265,* 2563–2568.

46. Jinek, M., Rehwinkel, J., Lazarus, B. D., Izaurralde, E., Hanover, J. A. & Conti, E. The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha. *Nat Struct Mol Biol* **2004,** *11,* 1001–1007.

47. Lazarus, M. B., Jiang, J., Gloster, T. M., Zandberg, W. F., Whitworth, G. E., Vocaldo, D. J. & Walker, S. Structural snapshots of the reaction coordinate for O-GlcNAc transferase. *Nat. Chem. Biol.* **2012,** *8,* 966–968.

48. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A., Lynn, A. J., Snedecor, J. O., Guan, S., Medzihradszky, K. F., Maltby, D. A., Schoepfer, R. & Burlingame, A. L. O-linked N-acetylglucosamine proteomics of postsynaptic density preparation using lectin weak affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* **2006,** *5,* 923–934.

49. Perrine, C., Ju, T., Cummings, R. D. & Gerken, T. A. Systematic determination of the peptide acceptor preferences for the human UDP-Gal:glycoprotein-α-GalNAc β3 galactosyltranferase (T-synthase). *Glycobiology* **2009,** *19,* 321–328.

50. Pathak, S., Alonso, J., Schimpl, M., Rafie, K., Blair, D. E., Borodkin, V. S., Schüttelkopf, A. W., Albarbarawi, O. & van Aalten, D. M. F. The active site of O-GlcNAc transferase imposes constraints on substrate sequence. *Nat. Struct. Mol. Biol.* **2015,** *22,* 744–750.

51. Joiner, C. M., Levine, Z. G., Aonbangkhen, C., Woo, C. M. & Walker, S. Aspartate

residues far from the active site drive O-GlcNAc transferase substrate selection. *J. Am. Chem. Soc.* **2019,** jacs.9b06061. doi:10.1021/jacs.9b06061

52. Britto-borges, T. & Barton, G. J. A study of the structural properties of sites modified by the O -linked 6-N- acetylglucosamine transferase. *PLoS One* **2017,** *12,* 1–14.

53. Yang, X., Ongusaha, P. P., Miles, P. D., Havstad, J. C., Zhang, F., So, W. V., Kudlow, J. E., Michell, R. H., Olefsky, J. M., Field, S. J. & Evans, R. M. Phosphoinositide signalling links O-GlcNAc transferase to insulin resistance. *Nature* **2008,** *451,* 964–969.

54. Cheung, W. D. & Hart, G. W. AMP-activated protein kinase and p38 MAPK activate O-GlcNAcylation of neuronal proteins during glucose deprivation. *J. Biol. Chem.* **2008,** *283,* 13009–13020.

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

56. Levine, Z., Fan, C., Melicher, M. S., Orman, M., Benjamin, T. & Walker, S. O-GlcNAc transferase recognizes protein substrates using an asparagine ladder in the TPR superhelix. *J. Am. Chem. Soc.* **2018,** jacs.7b13546. doi:10.1021/jacs.7b13546

57. Martinez-Fleites, C., Macauley, M. S., He, Y., Shen, D. L., Vocadlo, D. J. & Davies, G. J. Structure of an O-GlcNAc transferase homolog provides insight into intracellular glycosylation. *Nat. Struct. Mol. Biol.* **2008,** *15,* 764–765.

58. Clarke, A. J., Hurtado-Guerrero, R., Pathak, S., Schüttelkopf, A. W., Borodkin, V., Shepherd, S. M., Ibrahim, A. F. M. & Van Aalten, D. M. F. Structural insights into mechanism and specificity of O-GlcNAc transferase. *EMBO J.* **2008,** *27,* 2780–2788.

59. Comer, F. I. & Hart, G. W. Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II. *Biochemistry* **2001,** *40,* 7845–7852.

60. Cheung, W. D., Sakabe, K., Housley, M. P., Dias, W. B. & Hart, G. W. O-linked β-N-

acetylglucosaminyltransferase substrate specificity is regulated by myosin phosphatase targeting and other interacting proteins. *J. Biol. Chem.* **2008,** *283,* 33935–33941.

61. Iyer, S. P. N. & Hart, G. W. Roles of the Tetratricopeptide Repeat Domain in O-GlcNAc Transferase Targeting and Protein Substrate Specificity. *J. Biol. Chem.* **2003,** *278,* 24608–24616.

62. Hrit, J., Goodrich, L., Li, C., Wang, B.-A., Nie, J., Cui, X., Martin, E. A., Simental, E., Fernandez, J., Liu, M. Y., Nery, J. R., Castanon, R., Kohli, R. M., Tretyakova, N., He, C., Ecker, J. R., Goll, M. & Panning, B. OGT binds a conserved C-terminal domain of TET1 to regulate TET1 activity and function in development. *Elife* **2018,** *7,* 1–23.

63. Whisenhunt, T. R., Yang, X., Bowe, D. B., Paterson, A. J., Van Tine, B. A. & Kudlow, J. E. Disrupting the enzyme complex regulating O-GlcNAcylation blocks signaling and development. *Glycobiology* **2006,** *16,* 551–563.

64. Riu, I. H., Shin, I. S. & Do, S. Il. Sp1 modulates ncOGT activity to alter target recognition and enhanced thermotolerance in E. coli. *Biochem. Biophys. Res. Commun.* **2008,** *372,* 203–209.

65. Seo, H. G., Kim, H. B., Kang, M. J., Ryum, J. H., Yi, E. C. & Cho, J. W. Identification of the nuclear localisation signal of O-GlcNAc transferase and its nuclear import regulation. *Sci. Rep.* **2016,** *6,* 1–12.

66. Chen, Q. & Yu, X. OGT restrains the expansion of DNA damage signaling. *Nucleic Acids Res.* **2016,** *44,* 9266–9278.

67. Li, Z., Li, X., Nai, S., Geng, Q., Liao, J., Xu, X. & Li, J. Checkpoint kinase 1-induced phosphorylation of O-linked β-N-acetylglucosamine transferase regulates the intermediate filament network during cytokinesis. *J. Biol. Chem.* **2017,** *292,* jbc.M117.811646.

68. Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V. & Skrzypek, E.

PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Res.* **2015,** *43,* D512–D520.

69. Bullen, J. W., Balsbaugh, J. L., Chanda, D., Shabanowitz, J., Hunt, D. F., Neumann, D. & Hart, G. W. Cross-talk between two essential nutrient-sensitive enzymes O-GlcNAc Transferase (OGT) and Amp-Activated Protein Kinase (AMPK). *J. Biol. Chem.* **2014,** *289,* 10592–10606.

70. Gao, J., Yang, Y., Qiu, R., Zhang, K., Teng, X., Liu, R. & Wang, Y. Proteomic analysis of the OGT interactome: novel links to epithelial-mesenchymal transition and metastasis of cervical cancer. *Carcinogenesis* **2018,** 1–13. doi:10.1093/carcin/bgy097

71. Deng, R. P., He, X., Guo, S. J., Liu, W. F., Tao, Y. & Tao, S. C. Global identification of O-GlcNAc transferase (OGT) interactors by a human proteome microarray and the construction of an OGT interactome. *Proteomics* **2014,** *14,* 1020–1030.

72. Stephen, H. M., Praissman, J. L. & Wells, L. Generation of an Interactome for the Tetratricopeptide Repeat Domain of O-GlcNAc Transferase Indicates a Role for the Enzyme in Intellectual Disability. *J. Proteome Res.* **2020,** doi:10.1021

73. Gao, Y., Wells, L., Comer, F. I., Parker, G. J. & Hart, G. W. Dynamic O-glycosylation of nuclear and cytosolic proteins: Cloning and characterization of a neutral, cytosolic β-N-acetylglucosaminidase from human brain. *J. Biol. Chem.* **2001,** *276,* 9838–9845.

74. Toleman, C., Paterson, A. J., Whisenhunt, T. R. & Kudlow, J. E. Characterization of the histone acetyltransferase (HAT) domain of a bifunctional protein with activable O-GlcNAcase and HAT activities. *J. Biol. Chem.* **2004,** *279,* 53665–53673.

75. Butkinaree, C., Cheung, W. D., Park, S., Park, K., Barber, M. & Hart, G. W. Characterization of B-N-acetylglucosaminidase cleavage by caspase-3 during apoptosis. *J. Biol. Chem.* **2008,** *283,* 23557–23566.

76. Rao, F. V, Schüttelkopf, A. W., Dorfmueller, H. C., Ferenbach, A. T., Navratilova, I. & van

Aalten, D. M. F. Structure of a bacterial putative acetyltransferase defines the fold of the human O-GlcNAcase C-terminal domain. *Open Biol.* **2013,** *3,* 130021.

77. He, Y., Roth, C., Turkenburg, J. P. & Davies, G. J. Three-dimensional structure of a Streptomyces sviceus GNAT acetyltransferase with similarity to the C-terminal domain of the human GH84 O-GlcNAcase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2014,** *70,* 186–195.

78. Wells, L., Gao, Y., Mahoney, J. A., Vosseller, K., Chen, C., Rosen, A., Hart, G. W., Wells, L., Gao, Y., Mahoney, J. A., Vosseller, K., Chen, C., Rosen, A. & Hart, G. W. Dynamic O-Glycosylation of Nuclear and Cytosolic Proteins : Further Characterization of the Nucleocytoplasmic B-N-Acetylglucosaminidase, O-GlcNAcase. *J. Biol. Chem.* **2002,** *277,* 1755–1761.

79. Zeidan, Q., Wang, Z., Maio, A. De & Hart, G. W. O-GlcNAc Cycling Enzymes Associate with the Translational Machinery and Modify Core Ribosomal Proteins. *Mol. Biol. Cell* **2010,** *21,* 1922–1936.

80. Comtesse, N., Maldener, E. & Meese, E. Identification of a nuclear variant of MGEA5, a cytoplasmic hyaluronidase and a β-A-acetylglucosaminidase. *Biochem. Biophys. Res. Commun.* **2001,** *283,* 634–640.

81. Shen, D. L., Gloster, T. M., Yuzwa, S. A. & Vocadlo, D. J. Insights into O-linked N-acetylglucosamine (O-GlcNAc) processing and dynamics through kinetic analysis of O-GlcNAc transferase and O-GlcNAcase activity on protein substrates. *J. Biol. Chem.* **2012,** *287,* 15395–15408.

82. Li, B., Li, H., Lu, L. & Jiang, J. Structures of human O-GlcNAcase and its complexes reveal a new substrate recognition mode. *Nat. Struct. Mol. Biol.* **2017,** *24,* 362–369.

83. Schimpl, M., Schüttelkopf, A. W., Borodkin, V. S. & Van Aalten, D. M. F. Human OGA binds substrates in a conserved peptide recognition groove. *Biochem. J.* **2010,** *432,* 1–7.

84. Elsen, N. L., Patel, S. B., Ford, R. E., Hall, D. L., Hess, F., Kandula, H., Kornienko, M., Reid, J., Selnick, H., Shipman, J. M., Sharma, S., Lumb, K. J., Soisson, S. M. & Klein, D. J. Insights into activity and inhibition from the crystal structure of human O-GlcNAcase. *Nat. Chem. Biol.* **2017,** *13,* 613–615.

85. Roth, C., Chan, S., Offen, W. A., Hemsworth, G. R., Willems, L. I., King, D. T., Varghese, V., Britton, R., Vocadlo, D. J. & Davies, G. J. Structural and functional insight into human O-GlcNAcase. *Nat. Chem. Biol.* **2017,** *13,* 610–612.

86. Groves, J. A., Maduka, A. O., O'Meally, R. N., Cole, R. N. & Zachara, N. E. Fatty acid synthase inhibits the O-GlcNAcase during oxidative stress. *J. Biol. Chem.* **2017,** *292,* 6493–6511.

87. Groussaud, D., Khair, M., Tollenaere, A. I., Waast, L., Kuo, M. S., Mangeney, M., Martella, C., Fardini, Y., Coste, S., Souidi, M., Benit, L., Pique, C. & Issad, T. Hijacking of the O-GlcNAcZYME complex by the HTLV-1 Tax oncoprotein facilitates viral transcription. *PLoS Pathog.* **2017,** *13,* 1–23.

88. Brzostowski, J. A., Meckel, T., Hong, J., Chen, A. & Jin, T. Imaging Protein-Protein Interactions by Forster Resonance Energy Transfer (FRET) Microscopy in Live Cells. *Curr. Protoc. Protein Sci.* **2009,** *56,* 19.5.1-19.5.12.

89. Kerppola, T. K. Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu. Rev. Biophys.* **2008,** *37,* 465–487.

90. Holt, G. D. & Hart, G. W. The subcellular distribution of terminal N-acetylglucosamine moieties. Localization of a novel protein-saccharidie linkage, O-linked GlcNAc. *J. Biol. Chem.* **1986,** *261,* 8049–8057.

91. Walgren, J. L. E., Vincent, T. S., Schey, K. L. & Buse, M. G. High glucose and insulin promote O-GlcNAc modification of proteins, including α-tubulin. *Am. J. Physiol. - Endocrinol. Metab.* **2003,** *284,* 424–434.

92. Wang, Z., Pandey, A. & Hart, G. W. Dynamic interplay between O-linked N-acetylglucosaminylation and glycogen synthase kinase-3-dependent phosphorylation. *Mol. Cell. Proteomics* **2007,** *6,* 1365–1379.

93. Samavarchi-Tehrani, P., Samson, R. & Gingras, A. C. Proximity dependent biotinylation: Key enzymes and adaptation to proteomics approaches. *Mol. Cell. Proteomics* **2020,** *19,* 757–773.

94. Elzi, D. J., Song, M., Hakala, K., Weintraub, S. T. & Shiio, Y. Proteomic Analysis of the EWS-Fli - 1 Interactome Reveals the Role of the Lysosome in EWS-Fli - 1 Turnover. *J. Proteome Res.* **2015,**

95. Couzens, A., Knight, J., Kean, M., Teo, G., Weiss, A., Dunham, W., Lin, Z., Bagshaw, R., Sicheri, F., Pawson, T., Wrana, J., Choi, H. & Gingras, A. Protein Interaction Metwork of the Mammalian Hippo Pathway Reveals Mechanisms of Kinase-Phosphatase Interactions. *Sci. Signal.* **2013,** *6,* 1–13.

96. Ramirez, D. H., Aonbangkhen, C., Wu, H.-Y., Naftaly, J. A., Tang, S., O'Meara, T. R. & Woo, C. M. Engineering a proximity-directed O-GlcNAc transferase for selective protein O-GlcNAcylation in cells. *ACS Chem. Biol.* **2020,** *15,* 1059–1066.

97. Zhu, Y. & Hart, G. W. Targeting the O-GlcNAc Transferase to Specific Proteins Using RNA Aptamers. *FASEB J.* **2020,** *31,* 1–1.

98. Yu, S. H., Zhao, P., Sun, T., Gao, Z., Moremen, K. W., Boons, G. J., Wells, L. & Steet, R. Selective exo-enzymatic labeling detects increased cell surface sialoglycoprotein expression upon megakaryocytic differentiation. *J. Biol. Chem.* **2016,** *291,* 3982–3989.

99. Weatherly, D. B., Atwood, J. A., Minning, T. A., Cavola, C., Tarleton, R. L. & Orlando, R. A Heuristic Method for Assigning a False-discovery Rate for Protein Identifications from Mascot Database Search Results. *Mol. Cell. Proteomics* **2005,** *4,* 762–772.

100. Zybailov, B., Mosley, A. L., Sardiu, M. E., Coleman, M. K., Florens, L. & Washburn, M. P.

Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J. Proteome Res.* **2006,** *5,* 2339–2347.

101. UniProt Consortium, T. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019,** *47,* D506–D515.

102. Schreiber, E., Matthias, P., Muller, M. M. & Schaffner, W. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. *Nucleic Acids Res.* **1989,** *17,* 6419.

103. Gene Ontology Consortium, T., Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M. & Rubin, Gerald M. Sherlock, G. The Gene Ontology Consortium, Michael Ashburner1, Catherine A. Ball3, Judith A. Blake4, David Botstein3, Heather Butler1, J. Michael Cherry3, Allan P. Davis4, Kara Dolinski3, Selina S. Dwight3, Janan T. Eppig4, Midori A. Harris3, David P. Hill4, Laurie Is. *Nat. Genet.* **2000,** *25,* 25–29.

104. Gene Ontology Consortium, T. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **2019,** *47,* D330–D338.

105. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS One* **2011,** *6,*

106. Kuznetsova, I., Lugmayr, A., Siira, S. J., Rackham, O. & Filipovska, A. CirGO: An alternative circular way of visualising gene ontology terms. *BMC Bioinformatics* **2019,** *20,* 1–7.

107. Online Mendelian In heritance in Man, OMIM. *McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD)* **https://omim.org/ (2020)**

108. Slawson, C., Lakshmanan, T., Knapp, S. & Hart, G. W. A Mitotic GlcNAcylation/Phosphorylation Signaling Complex Alters the Posttranslational State of

the Cytoskeletal Protein Vimentin. *Mol. Biol. Cell* **2008,** *19,* 4130–4140.

109. Lee, J. S. & Zhang, Z. O-linked N-acetylglucosamine transferase (OGT) interacts with the histone chaperone HIRA complex and regulates nucleosome assembly and cellular senescence. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113,* E3213–E3220.

110. Dey, A., Seshasayee, D., Noubade, R., French, D. M., Chaurushiya, M. S., Kirkpatrick, D. S., Pham, V. C., Lill, J. R., Bakalarski, E., Wu, J., Phu, L., Katavolos, P., Saunders, L. M., Modrusan, Z., Seshagiri, S., Dong, K., Lin, Z., Martin, F., Levine, R. L. & Dixit, V. M. Loss of the tumor suppressor BAP1 causes myeloid transformation. *Science (80-.).* **2012,** *337,* 1541–1546.

111. Alfaro, J. F., Gong, C., Monroe, M. E., Aldrich, J. T., Clauss, T. R. W., Purvine, S. O., Wang, Z., Camp, D. G., Shabanowitz, J., Stanley, P., Hart, G. W., Hunt, D. F., Yang, F. & Smith, R. D. Tandem mass spectrometry identifies many mouse brain O -GlcNAcylated proteins including EGF domain-specific O -GlcNAc transferase targets. *Pnas* **2012,** *109,* 7280–7285.

112. Gambetta, M. C., Oktaba, K. & Muller, J. Essential Role of the Glycosyltransferase Sxc/Ogt in Polycomb Repression. *Science (80-.).* **2009,** *325,* 93–96.

113. Sinclair, D. A. R., Syrzycka, M., Macauley, M. S., Rastgardani, T., Komljenovic, I., Vocadlo, D. J., Brock, H. W. & Honda, B. M. Drosophila O-GlcNAc transferase (OGT) is encoded by the Polycomb group (PcG) gene, super sex combs (sxc). *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106,* 13427–13432.

114. Wu, D., Zhao, L., Feng, Z., Yu, C., Ding, J., Wang, L., Wang, F., Liu, D., Zhu, H., Xing, F., Conaway, J. W., Conaway, R. C., Cai, Y. & Jin, J. O-Linked N-acetylglucosamine transferase 1 regulates global histone H4 acetylation via stabilization of the nonspecific lethal protein NSL3. *J. Biol. Chem.* **2017,** *292,* 10014–10025.

115. Sakabe, K., Wang, Z. & Hart, G. W. β-N-acetylglucosamine (O-GlcNAc) is part of the

histone code. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107,* 19915–19920.

116. Brownlee, P. M., Meisenberg, C. & Downs, J. A. The SWI/SNF chromatin remodelling complex: Its role in maintaining genome stability and preventing tumourigenesis. *DNA Repair (Amst).* **2015,** *32,* 127–133.

117. Huang, L., Jolly, L. A., Willis-Owen, S., Gardner, A., Kumar, R., Douglas, E., Shoubridge, C., Wieczorek, D., Tzschach, A., Cohen, M., Hackett, A., Field, M., Froyen, G., Hu, H., Haas, S. A., Ropers, H. H., Kalscheuer, V. M., Corbett, M. A. & Gecz, J. A noncoding, regulatory mutation implicates HCFC1 in nonsyndromic intellectual disability. *Am. J. Hum. Genet.* **2012,** *91,* 694–702.

118. Kaasik, K., Kivimae, S., Allen, J. J., Chalkley, R. J., Huang, Y., Baer, K., Kissel, H., Burlingame, A. L., Shokat, K. M., Ptacek, L. J. & Fu, Y.-H. Glucose Sensor O-GlcNAcylation Coordinates with Phosphorylation to Regulate Circadian Clock. *Cell Metab.* **2013,** *17,* 291–302.

119. Ranuncolo, S. M., Ghosh, S., Hanover, J. A., Hart, G. W. & Lewis, B. A. Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription in vitro and in vivo. *J. Biol. Chem.* **2012,** *287,* 23549–23561.

120. Eustice, M., Bond, M. R. & Hanover, J. A. O-GlcNAc cycling and the regulation of nucleocytoplasmic dynamics. *Biochem. Soc. Trans.* **2017,** *45,* 427–436.

121. Wang, Q., Fang, P., He, R., Li, M., Yu, H., Zhou, L., Yi, Y., Wang, F., Rong, Y., Zhang, Y., Chen, A., Peng, N., Lin, Y., Lu, M., Zhu, Y., Peng, G., Rao, L. & Liu, S. O-GlcNAc transferase promotes influenza A virus-induced cytokine storm by targeting interferon regulatory factor-5. *Sci. Adv.* **2020,** *6,* 1–13.

122. Baldini, S. F., Steenackers, A., Olivier-Van Stichelen, S., Mir, A. M., Mortuaire, M., Lefebvre, T. & Guinez, C. Glucokinase expression is regulated by glucose through O-GlcNAc glycosylation. *Biochem. Biophys. Res. Commun.* **2016,** *478,* 942–948.

123. Yi, W., Clark, P. M., Mason, D. E., Keenan, M. C., Hill, C., William, A., Iii, G., Peters, E. C., Driggers, E. M. & Hsieh-wilson, L. C. PFK1 Glycosylation Is a Key Regulator of Cancer Cell Growth and Central Metabolic Pathways Wen. *Science (80-.).* **2013,** *337,* 975–980.

124. Olivier-Van Stichelen, S., Dehennaut, V., Buzy, A., Zachayus, J. L., Guinez, C., Mir, A. M., El Yazidi-Belkoura, I., Copin, M. C., Boureme, D., Loyaux, D., Ferrara, P. & Lefebvre, T. O-GlcNAcylation stabilizes β-catenin through direct competition with phosphorylation at threonine 41. *FASEB J.* **2014,** *28,* 3325–3328.

125. Akan, I., Stichelen, S. O.-V., Bond, M. R. & Hanover, J. A. Nutrient-driven O-GlcNAc in proteostasis and neurodegeneration. *J. Neurochem.* **2017,** *12,* 3218–3221.

126. Danino, Y. M., Even, D., Ideses, D. & Juven-Gershon, T. The core promoter: At the heart of gene expression. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2015,** *1849,* 1116–1131.

127. Selvan, N., George, S., Serajee, F. J., Shaw, M., Hobson, L., Kalscheuer, V., Prasad, N., Levy, S. E., Taylor, J., Aftimos, S., Schwartz, C. E., Huq, A. H. M. M., Gecz, J. & Wells, L. Characterization of missense mutations in the tetratricopeptide region of O-GlcNAc transferase causal for X-linked intellectual disability. *J. Biol. Chem.* **2018,**

128. Leturcq, M., Mortuaire, M., Hardivillé, S., Schulz, C., Lefebvre, T. & Vercoutter-Edouart, A.-S. O-GlcNAc transferase associates with the MCM2-7 complex and its silencing destabilizes MCM-MCM interactions. *Cell. Mol. Life Sci.* **2018,** 16–18. doi:10.1007/s00018-018-2874-0

129. Choi, H., Kim, C., Song, H., Cha, M.-Y., Cho, H. J., Son, S. M., Kim, H. J. & Mook-Jung, I. Amyloid β-induced elevation of O-GlcNAcylated c-Fos promotes neuronal cell death. *Aging Cell* **2018,** e12872. doi:10.1111/acel.12872

130. Ma, X., Li, H., He, Y. & Hao, J. The emerging link between O-GlcNAcylation and

neurological disorders. *Cell. Mol. Life Sci.* **2017,** *74,* 3667–3686.

131. Peng, C., Zhu, Y., Zhang, W., Liao, Q., Chen, Y., Zhao, X., Guo, Q., Shen, P., Zhen, B., Qian, X., Yang, D., Zhang, J.-S., Xiao, D., Qin, W. & Pei, H. Regulation of the Hippo-YAP Pathway by Glucose Sensor O-GlcNAcylation. *Mol. Cell* **2017,** *68,* 591-604.e5.

132. Sekine, H., Okazaki, K., Kato, K., Shima, H., Katsuoka, F., Tsujita, T., Suzuki, N., Kobayashi, A., Yamamoto, M., Motohashi, H., Science, F., Sciences, M., Motohashi, H. & Yamamoto, M. O -GlcNAcylation Signal Mediates Proteasome Inhibitor Resistance in Cancer Cells by Stabilizing NRF1. *Mol. Cell. Biol.* **2018,** *38,*

133. Yang, Y., Yin, X., Yang, H. & Xu, Y. Histone demethylase LSD2 Acts as an E3Ubiquitin ligase and inhibits cancer cell growth through promoting proteasomal degradation of OGT. *Mol. Cell* **2015,** *58,* 47–59.

134. Stephen, H. M., Praissman, J. L. & Wells, L. Generation of an Interactome for the Tetratricopeptide Repeat Domain of O-GlcNAc Transferase Indicates a Role for the Enzyme in Intellectual Disability. *J. Proteome Res.* **2021,** *20,* 1229–1242.

135. Marshall, S., Nadeau, O. & Yamasaki, K. Dynamic actions of glucose and glucosamine on hexosamine biosynthesis in isolated adipocytes: Differential effects on glucosamine 6-phosphate, UDP-N-acetylglucosamine, and ATP levels. *J. Biol. Chem.* **2004,** *279,* 35313–35319.

136. Kang, J. G., Park, S. Y., Ji, S., Jang, I., Park, S., Kim, H. S., Kim, S. M., Yook, J. I., Park, Y. Il, Roth, J. & Cho, J. W. O-GlcNAc protein modification in cancer cells increases in response to glucose deprivation through glycogen degradation. *J. Biol. Chem.* **2009,** *284,* 34777–34784.

137. Rodríguez-Enríquez, S., Marín-Hernández, A., Gallardo-Pérez, J. C. & Moreno-Sánchez, R. Kinetics of transport and phosphorylation of glucose in cancer cells. *J. Cell. Physiol.* **2009,** *221,* 552–559.

138. German, M. S. Glucose sensing in pancreatic islet beta cells: The key role of glucokinase and the glycolytic intermediates. *Proc. Natl. Acad. Sci. U. S. A.* **1993,** *90,* 1781–1785.

139. Sharabi, K., Tavares, C. D. J., Rines, A. K. & Puigserver, P. Molecular Pathophysiology of Hepatic Glucose Production. *Mol. Aspects Med.* **2015,** *46,* 21–33.

140. Güemes, A. & Georgiou, P. Review of the role of the nervous system in glucose homoeostasis and future perspectives towards the management of diabetes. *Bioelectron. Med.* **2018,** *4,* 1–18.

141. Brunner, D., Frank, J., Appl, H., Schöffl, H., Pfaller, W. & Gstraunthaler, G. Serum-free cell culture: the serum-free media interactive online database. *ALTEX* **2010,** *27,* 53–62.

142. Olivier-Van Stichelen, S., Drougat, L., Dehennaut, V., El Yazidi-Belkoura, I., Guinez, C., Mir, A. M., Michalski, J. C., Vercoutter-Edouart, A. S. & Lefebvre, T. Serum-stimulated cell cycle entry promotes ncOGT synthesis required for cyclin D expression. *Oncogenesis* **2012,** *1,* e36-6.

143. MacDonald, B. T., Tamai, K. & He, X. Wnt/β-Catenin Signaling: Components, Mechanisms, and Diseases. *Dev. Cell* **2009,** *17,* 9–26.

144. Mahé, D., Mähl, P., Gattoni, R., Fischer, N., Mattei, M. G., Stévenin, J. & Fuchs, J. P. Cloning of human 2H9 heterogeneous nuclear ribonucleoproteins: Relation with splicing and early heat shock-induced splicing arrest. *J. Biol. Chem.* **1997,** *272,* 1827–1836.

145. Gradi, A., Imataka, H., Svitkin, Y. V., Rom, E., Raught, B., Morino, S. & Sonenberg, N. A Novel Functional Human Eukaryotic Translation Initiation Factor 4G. *Mol. Cell. Biol.* **1998,** *18,* 334–342.

146. Zhang, X., Yan, C., Hang, J., Finci, L. I., Lei, J. & Shi, Y. An Atomic Structure of the Human Spliceosome. *Cell* **2017,** *169,* 918-929.e14.

147. Cole, R. N. & Hart, G. W. Cytosolic O-glycosylation is abundant in nerve terminals. *J. Neurochem.* **2001,** *79,* 1080–1089.

148. Lagerlöf, O., Hart, G. W. & Huganir, R. L. O-GlcNAc transferase regulates excitatory synapse maturity. *Proc. Natl. Acad. Sci.* **2017,** *114,* 1684–1689.

149. Taylor, E. W., Wang, K., Nelson, A. R., Bredemann, T. M., Fraser, K. B., Clinton, S. M., Puckett, R., Marchase, R. B., Chatham, J. C. & McMahon, L. L. O-GlcNAcylation of AMPA Receptor GluA2 Is Associated with a Novel Form of Long-Term Depression at Hippocampal Synapses. *J. Neurosci.* **2014,** *34,* 10–21.

150. Jahn, K., Wieltsch, C., Blumer, N., Mehlich, M., Pathak, H., Khan, A. Q., Hildebrandt, H. & Frieling, H. A cell culture model for investigation of synapse influenceability: epigenetics, expression and function of gene targets important for synapse formation and preservation in SH-SY5Y neuroblastoma cells differentiated by retinoic acid. *J. Neural Transm.* **2017,** *124,* 1341–1367.

151. Seo, M., Kim, Y., Lee, Y. Il, Kim, S. Y., Ahn, Y. M., Kang, U. G., Roh, M. S., Kim, Y. S. & Juhnn, Y. S. Membrane depolarization stimulates the proliferation of SH-SY5Y human neuroblastoma cells by increasing retinoblastoma protein (RB) phosphorylation through the activation of cyclin-dependent kinase 2 (Cdk2). *Neurosci. Lett.* **2006,** *404,* 87–92.

152. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A., Lynn, A. J., Snedecor, J. O., Guan, S., Medzihradszky, K. F., Maltby, D. A., Schoepfer, R. & Burlingame, A. L. O -Linked N -Acetylglucosamine Proteomics of Postsynaptic Density Preparations Using Lectin Weak Affinity Chromatography and Mass Spectrometry. *Mol. Cell. Proteomics* **2006,** *5,* 923–934.

153. Pizzinga, M., Harvey, R. F., Garland, G. D., Mordue, R., Dezi, V., Ramakrishna, M., Sfakianos, A., Monti, M., Mulroney, T. E., Poyry, T. & Willis, A. E. The cell stress response: extreme times call for post-transcriptional measures. *Wiley Interdiscip. Rev. RNA* **2020,** *11,* 1–22.

154. Ma, J. & Hart, G. W. O-GlcNAc profiling: from proteins to proteomes. *Clin. Proteomics*

2014, *11,* 8.

155. Martinez, M., Renuse, S., Kreimer, S., O'Meally, R., Natov, P., Madugundu, A. K., Nirujogi, R. S., Tahir, R., Cole, R., Pandey, A. & Zachara, N. E. Quantitative proteomics reveals that the OGT interactome is remodeled in response to oxidative stress. *Mol. Cell. Proteomics* **2021,** *20,* 100069.

156. Selvan, N., George, S., Serajee, F. J., Shaw, M., Hobson, L., Kalscheuer, V., Prasad, N., Levy, S. E., Taylor, J., Aftimos, S., Schwartz, C. E., Huq, A. M., Gecz, J. & Wells, L. O-GlcNAc transferase missense mutations linked to X-linked intellectual disability deregulate genes involved in cell rate determination and signaling. *J. Biol. Chem.* **2018,** doi:10.1074/jbc.RA118.002583

157. Cochran, J. N., Geier, E. G., Bonham, L. W., Newberry, J. S., Amaral, M. D., Thompson, M. L., Lasseigne, B. N., Karydas, A. M., Roberson, E. D., Cooper, G. M., Rabinovici, G. D., Miller, B. L., Myers, R. M. & Yokoyama, J. S. Non-coding and Loss-of-Function Coding Variants in TET2 are Associated with Multiple Neurodegenerative Diseases. *Am. J. Hum. Genet.* **2020,** *106,* 632–645.

158. Ito, S., Alessio, A. C. D., Taranova, O. V, Hong, K., Lawrence, C. & Zhang, Y. Role of Tet proteins in 5mC to 5hmC conversion, ES cell self-renewal, and ICM specification. *Nature* **2010,** *466,* 1129–1133.

159. Mulholland, C. B., Traube, F. R., Ugur, E., Parsa, E., Eckl, E. M., Schönung, M., Modic, M., Bartoschek, M. D., Stolz, P., Ryan, J., Carell, T., Leonhardt, H. & Bultmann, S. Distinct and stage-specific contributions of TET1 and TET2 to stepwise cytosine oxidation in the transition from naive to primed pluripotency. *Sci. Rep.* **2020,** *10,* 1–12.

160. Bustos, F., Segarra-Fas, A., Chaugule, V. K., Brandenburg, L., Branigan, E., Toth, R., Macartney, T., Knebel, A., Hay, R. T., Walden, H. & Findlay, G. M. RNF12 X-Linked Intellectual Disability Mutations Disrupt E3 Ligase Activity and Neural Differentiation. *Cell*

Rep. **2018,** *23,* 1599–1611.

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

162. Mariappa, D., Zheng, X., Schimpl, M., Raimi, O., Ferenbach, A. T., Müller, H.-A. J. & van Aalten, D. M. F. Dual functionality of O-GlcNAc transferase is required for Drosophila development. *Open Biol.* **2015,** *5,* 150234.

163. Constable, S., Lim, J.-M., Vaidyanathan, K. & Wells, L. O-GlcNAc transferase regulates transcriptional activity of human Oct4. *Glycobiology* **2017,** *27,* 1–11.

164. Ahmed, M. Y., Chioza, B. A., Rajab, A., Schmitz-Abe, K., Al-Khayat, A., Al-Turki, S., Baple, E. L., Patton, M. A., Al-Memar, A. Y., Hurles, M. E., Partlow, J. N., Hill, R. S., Evrony, G. D., Servattalab, S., Markianos, K., Walsh, C. A., Crosby, A. H. & Mochida, G. H. Loss of PCLO function underlies pontocerebellar hypoplasia type III. *Neurology* **2015,** *84,* 1745–1750.

165. Gundelfinger, E. D., Reissner, C. & Garner, C. C. Role of Bassoon and Piccolo in Assembly and Molecular Organization of the Active Zone. *Front. Synaptic Neurosci.* **2016,** *7,*

166. Liu, G., Peng, X., Cai, Y., Cheng, A., Zha, L. & Wang, Z. DIMT1 overexpression correlates with progression and prognosis in gastric carcinoma. *Hum. Pathol.* **2017,** *70,* 35–42.

167. Shen, H., Stoute, J. & Liu, K. F. Structural and catalytic roles of the human 18S rRNA methyltransferases DIMT1 in ribosome assembly and translation. *J. Biol. Chem.* **2020,** *295,* 12058–12070.

168. Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A.,

Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I. M., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A. K., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., Edqvist, P. H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., Von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., Von Heijne, G., Nielsen, J. & Pontén, F. Tissue-based map of the human proteome. *Science (80-.).* **2015,** *347,*

169. Ogami, K., Richard, P., Chen, Y., Hoque, M., Li, W., Moresco, J. J., Yates, J. R., Tian, B. & Manley, J. L. An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. *Genes Dev.* **2017,** *31,* 1257–1271.

170. Garland, W., Comet, I., Wu, M., Radzisheuskaya, A., Rib, L., Vitting-Seerup, K., Lloret-Llinares, M., Sandelin, A., Helin, K. & Jensen, T. H. A Functional Link between Nuclear RNA Decay and Transcriptional Control Mediated by the Polycomb Repressive Complex 2. *Cell Rep.* **2019,** *29,* 1800-1811.e6.

171. Decourcelle, A., Loison, I., Baldini, S., Leprince, D. & Dehennaut, V. Evidence of a compensatory regulation of colonic O-GlcNAc transferase and O-GlcNAcase expression in response to disruption of O-GlcNAc homeostasis. *Biochem. Biophys. Res. Commun.* **2019,** doi:10.1016/j.bbrc.2019.10.090

172. Deevy, O. & Bracken, A. P. PRC2 functions in development and congenital disorders. *Dev.* **2019,** *146,*

173. Huo, B., Zhang, W., Zhao, X., Dong, H., Yu, Y., Wang, J., Qian, X. & Qin, W. A triarylphosphine-trimethylpiperidine reagent for the one-step derivatization and enrichment of protein post-translational modifications and identification by mass spectrometry. *Chem. Commun.* **2018,** *54,* 13790–13793.

174. Li, B. & Kohler, J. J. Glycosylation of the nuclear pore. *Traffic* **2014,** *15,* 347–361.

175. Vaidyanathan, K., Durning, S. & Wells, L. Functional O-GlcNAc modifications: Implications in molecular regulation and pathophysiology. *Crit. Rev. Biochem. Mol. Biol.* **2014,** *49,* 140–163.

176. Chen, X., Wei, S., Ji, Y., Guo, X. & Yang, F. Quantitative proteomics using SILAC: Principles, applications, and developments. *Proteomics* **2015,** *15,* 3175–3192.

177. Zecha, J., Satpathy, S., Kanashova, T., Avanessian, S. C., Kane, M. H., Clauser, K. R., Mertins, P., Carr, S. A. & Kuster, B. TMT labeling for the masses: A robust and cost-efficient, in-solution labeling approach. *Mol. Cell. Proteomics* **2019,** *18,* 1468–1478.

178. Zhang, W., Liu, T., Dong, H., Bai, H., Tian, F., Shi, Z., Chen, M., Wang, J., Qin, W. & Qian, X. Synthesis of a Highly Azide-Reactive and Thermosensitive Biofunctional Reagent for Efficient Enrichment and Large-Scale Identification of O-GlcNAc Proteins by Mass Spectrometry. *Anal. Chem.* **2017,** *89,* 5810–5817.

179. Tellez-Plaza, M., Tang, W. Y., Shang, Y., Umans, J. G., Francesconi, K. A., Goessler, W., Ledesma, M., Leon, M., Laclaustra, M., Pollak, J., Guallar, E., Cole, S. A., Fallin, M. D. & Navas-Acien, A. Association of global DNA methylation and global DNA hydroxymethylation with metals and other exposures in human blood DNA samples. *Environ. Health Perspect.* **2014,** *122,* 946–954.

180. Coppieters, N., Dieriks, B. V., Lill, C., Faull, R. L. M., Curtis, M. A. & Dragunow, M. Global changes in DNA methylation and hydroxymethylation in Alzheimer's disease human brain. *Neurobiol. Aging* **2014,** *35,* 1334–1344.

181. Mukherjee, K., Yang, X., Gerber, S. H., Kwon, H. B., Ho, A., Castillo, P. E., Liu, X. & Südhof, T. C. Piccolo and bassoon maintain synaptic vesicle clustering without directly participating in vesicle exocytosis. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107,* 6504–6509.

182. Waites, C. L., Leal-Ortiz, S. A., Andlauer, T. F. M., Sigrist, S. J. & Garner, C. C. Piccolo regulates the dynamic assembly of presynaptic F-actin. *J. Neurosci.* **2011,** *31,* 14250–

14263.

183. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **2015,** *2015,* 21.29.1-21.29.9.