Role of O-GlcNAc in the Vertebrate Secretory Pathway

by

Nathan James Cox

Department of Pharmacology and Cancer Biology
Duke University

Date:_______________________

Approved:

___________________________
Michael Boyce, Supervisor

___________________________
Tso-Pang Yao

___________________________
Timothy Arthur James Haystead

___________________________
Jeffrey C. Rathmell

___________________________
Huanghe Yang

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Pharmacology and Cancer Biology in the Graduate School of Duke University

2018
ABSTRACT

Role of O-GlcNAc in the Vertebrate Secretory Pathway

by

Nathan James Cox

Department of Pharmacology and Cancer Biology
Duke University

Date:_______________________

Approved:

________________________________________
Michael Boyce, Supervisor

________________________________________
Tso-Pang Yao

________________________________________
Timothy Arthur James Haystead

________________________________________
Jeffrey C. Rathmell

________________________________________
Huanghe Yang

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Pharmacology and Cancer Biology in the Graduate School of Duke University

2018
Abstract

O-linked β-N-acetylglucosamine (O-GlcNAc) exerts myriad effects on protein localization, activation, inhibition, stability, conformational changes, or degradation. However, the biochemical effects of O-GlcNAc on the vast majority of substrates is unknown. Recently, we and others have shown that several coat protein complex II (COPII) components including SEC23A, SEC24C, and SEC31A are O-GlcNAcylated. The COPII coat complex consists of protein coated carriers that mediate secretory trafficking from the endoplasmic reticulum. To determine the effects of O-GlcNAc on COPII we used a combination of chemical, biochemical, cellular and genetic approaches to demonstrate that site-specific O-GlcNAcylation of COPII proteins mediates their protein-protein interactions and modulates cargo secretion. We demonstrate that individual O-GlcNAcylation sites of SEC23A are required for its function in human cells and vertebrate development, because mutation of these sites impairs SEC23A-dependent in vivo collagen trafficking and skeletogenesis in a zebrafish model of cranio-lenticulo-sutural dysplasia (CLSD).

Next, we developed a proteomic workflow to address the challenges of identifying and quantifying novel changes in substrate O-GlcNAcylation in response to a stimulus. Current methods of O-GlcNAcome enrichment suffer from issues with specificity, reproducibility, time-resolution, or require specialized hardware. We
developed a novel, unbiased glycoproteomics workflow to survey global changes in O-GlcNAc in response to stimuli. Our approach utilizes both stable isotope labeling with amino acids in cell culture (SILAC) for quantitation and metabolic labeling of O-GlcNAc for enrichment. Using our glycoproteomics workflow we examined the effects of brefeldin A (BFA), a fungal metabolite that disrupts vesicle trafficking, and cytokine deprivation on a pro-B cell line. We identified changes in the O-GlcNAcylation of Coatomer subunit gamma-1 (COPG) a coat protein complex I (COPI) component in response to BFA. Interestingly, COPI mediates traffic from the Golgi to the ER, as well as within the Golgi, and is the specific target of BFA. O-GlcNAcylation of COPI components may have effects similar to O-GlcNAc on COPII, possibly altering membrane binding or the trafficking of specific cargo.

Finally, we identified a candidate O-GlcNAc-mediated binding part of SEC23A using a combination chemical biology tools and mass spectrometry (MS). We identified ankycorbin, a vertebrate specific protein with no known function, as the candidate SEC23A O-GlcNAc-mediated binding partner. However, our attempts to validate this interaction were inconclusive.

Overall, this work examines the role of O-GlcNAc in the vertebrate secretory pathway. We demonstrate the effects of O-GlcNAc on SEC23A in the COPII pathway and identify a potentially novel method of COPI protein trafficking regulation via the O-GlcNAcylation of COPG.
# Contents

Abstract ........................................................................................................................................ iv

List of Figures .................................................................................................................................. ix

Acknowledgements ........................................................................................................................ xi

List of Abbreviations ...................................................................................................................... xii

1. Introduction ................................................................................................................................. 1
   1.1 O-GlcNAc ............................................................................................................................. 1
   1.2 O-GlcNAc Transferase and O-GlcNAcase ........................................................................... 3
   1.3 Evolutionary Conservation of O-GlcNAc ............................................................................ 6
   1.4 O-GlcNAc as a Signaling Modification ................................................................................. 8
   1.5 COPII ..................................................................................................................................... 10
   1.6 SAR1 .................................................................................................................................... 13
   1.7 SEC23 .................................................................................................................................. 15
   1.8 SEC24 .................................................................................................................................. 18
   1.9 SEC13 and SEC31 .............................................................................................................. 22
   1.10 Regulation of COPII via PTMs ......................................................................................... 23

2. Dynamic Glycosylation Governs the Vertebrate COPII Protein Trafficking Pathway ........ 26
   2.1 Introduction .......................................................................................................................... 26
   2.2 Materials and Methods ......................................................................................................... 29
   2.3 Results ................................................................................................................................ 47
     2.3.1 Novel O-GlcNAc Sites Identified on SEC23A, SEC24C, and SEC31A ....................... 47
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.2 Normal O-GlcNAc Cycling is Required for COPII-dependent Secretion</td>
<td>50</td>
</tr>
<tr>
<td>2.3.3 Dynamic O-GlcNAcylation Mediates Interactions Among COPII Pathway Proteins</td>
<td>54</td>
</tr>
<tr>
<td>2.3.4 Site-specific O-GlcNAcylation of SEC23A is Required for Proper Collagen Trafficking</td>
<td>59</td>
</tr>
<tr>
<td>2.3.5 Site-specific O-GlcNAcylation of Sec23A Regulates in vivo Collagen Trafficking</td>
<td>65</td>
</tr>
<tr>
<td>2.4 Discussion</td>
<td>69</td>
</tr>
<tr>
<td>2.5 Chapter Supplemental Material</td>
<td>75</td>
</tr>
<tr>
<td>3. Identifying Novel O-GlcNAcylated proteins</td>
<td>83</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>83</td>
</tr>
<tr>
<td>3.2 Materials and Methods</td>
<td>88</td>
</tr>
<tr>
<td>3.3 Results</td>
<td>95</td>
</tr>
<tr>
<td>3.3.1 FL5.12 (XL4.1) Support Both SILAC and Metabolic Labeling Approaches</td>
<td>95</td>
</tr>
<tr>
<td>3.3.2 Glycoproteomic Profile After BFA Treatment</td>
<td>98</td>
</tr>
<tr>
<td>3.3.3 Glycoproteomic Profile After IL-3 Starvation</td>
<td>108</td>
</tr>
<tr>
<td>3.4 Discussion</td>
<td>111</td>
</tr>
<tr>
<td>4. Identification and Characterization of a Candidate O-GlcNAc-mediated Binding Partner of Sec23A</td>
<td>117</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>117</td>
</tr>
<tr>
<td>4.2 Materials and Methods</td>
<td>121</td>
</tr>
<tr>
<td>4.3 Results</td>
<td>124</td>
</tr>
<tr>
<td>4.3.1 Identifying Candidate SEC23A O-GlcNAc-mediated Binding Partners</td>
<td>124</td>
</tr>
</tbody>
</table>
4.3.2 Functional Impact of Ankycorbin on COPII Pathway Function ................. 127
4.2.3 Testing a Potential SEC23A-ankycorbin O-GlcNAc-mediated Interaction in 293T Cells .................................................................................................................. 130
4.2.4 Testing a SEC23A-ankycorbin O-GlcNAc-mediated Interaction in SW1353.. 133
4.4 Discussion ......................................................................................................... 135
5. Conclusions ......................................................................................................... 139
Appendix A ............................................................................................................. 145
  Chapter 2 Mass Spec Information and Data ......................................................... 145
  Chapter 3 Mass Spec Information and Data ......................................................... 145
References .............................................................................................................. 146
Biography ............................................................................................................. 172
List of Figures

Figure 1: O-GlcNAc, OGT, and OGA..............................................................2

Figure 2: Coat protein complex II (COPII) transport carrier formation..................13

Figure 3: Novel O-GlcNAc Sites Identified on SEC23A, SEC24C, and SEC31A........49

Figure 4: Normal O-GlcNAc Cycling is Required for COPII-dependent Secretion ......53

Figure 5: Dynamic O-GlcNAcylation Mediates Interactions Among COPII Pathway Proteins...........................................................................................................58

Figure 6: SEC23A\textsuperscript{+} SW1353 cells accumulate collagen when stimulated with ascorbate.62

Figure 7: Site-specific O-GlcNAcylation of SEC23A is Required for Proper Collagen Trafficking.............................................................................................................63

Figure 8: Quantitation of collagen retention in SEC23A\textsuperscript{+} SW1353 stably reconstituted with S184A SEC23A-myc-6xHis .................................................................................................64

Figure 9: SEC23A S184A and 4A mutants co-IPs with SEC24 paralogs.........................65

Figure 10: Site-specific O-GlcNAcylation of Sec23A Regulates in vivo Collagen Trafficking.............................................................................................................68

Figure 11: O-GlcNAcylation site mutants of TFG do not disrupt GlcNDAz crosslinking77

Figure 12: BFA washout in mutant SW1353 cells....................................................79

Figure 13: SEC23B\textsuperscript{+} SW1353 cells display no collagen retention when untreated........82

Figure 14: Schematic for BEMAD .............................................................................85

Figure 15: XL4.1 rapidly incorporate GalNAz..........................................................97

Figure 16: XL4.1 resist apoptosis induced by Brefeldin A ........................................99

Figure 17: Log\textsubscript{2} of vehicle/BFA ratio in the cytoplasm and nucleus...............101

Figure 18: CLTC is not enriched after biotin-click and NeutrAvidin pulldown..........102
Figure 19: Log₂ of BFA/vehicle ratio in the cytoplasm and nucleus ........................................ 104

Figure 20: COPG enrichment via biotin-click decreases after treatment with BFA in the nuclear fraction ........................................................................................................................................ 106

Figure 21: O-GlcNAc detected on endogenous COPG ........................................................................ 108

Figure 22: Dramatic change in SEC24C after IL-3 stimulation .......................................................... 111

Figure 23: Schematic of GlcNDAz crosslinking .................................................................................. 119

Figure 24: SEC23A crosslinking is retained after epitope tagging ...................................................... 125

Figure 25: Ankycorbin identified as a candidate SEC23A O-GlcNAc-mediated binding partner ........................................................................................................................................ 127

Figure 26: Functional impact of ankycorbin on secretion ................................................................. 129

Figure 27: SEC23A-ankycorbin interaction not detected by IP/WB in 293T cells ...................... 131

Figure 28: Ankycorbin is not required for SEC23A crosslinking in 293T cells ................................. 132

Figure 29: SEC23A crosslinking in SW1353 cells .................................................................................. 134
Acknowledgements

I would like to thank Michael Boyce for providing an environment for me to grow as a scientist, and for introducing me to the world of glycobiology. Boyce Lab clan, thank you for providing entertainment, conversation, food, and scientific advice. COPquistadors (Brett, Brittany, Peter, and Thomas), thanks for keeping our lab bay filled to the brim and providing scientific and experimental support. My undergraduates, Peter and Thomas, you both provided great opportunities for teaching and learning.

I would like to thank Tim Haystead, Jeff Rathmell, Tso-Pang Yao, and Huanghe Yang for agreeing to be on my committee and providing annual feedback. Michael Hanna and Anjon Audhya, thank you for providing reagents and insight into the world of COPII, SEC23, and TFG. Gokhan Unlu and Ela Knapik, thank you for providing the zebrafish expertise and beautiful in vivo data. Erik Soderblom, thank you for spending copious amounts of time answering all my mass spec related questions. I would also like to thank Lee Graves for giving me my first taste of research as an undergraduate and inspiring my interest in proteomics.

I would also like to thank Fredric Bard, Frances Brodsky, Larry Fisher, Josh Gamse, David Ginsburg, Rami Khoriaty, Jennifer Kohler, Chris Nicchitta, Randy Schekman, and Benjamin Swarts for providing reagents and advice.
List of Abbreviations

ANDD - Anderson’s disease

AR - ankyrin repeat

ARF1 - ADP-ribosylation factor 1

ATG9 - autophagy-related protein 9

BEMAD - β-elimination/Michael addition

BFA - brefeldin A

CaMKII - Ca\(^{2+}\)/calmodulin-dependent protein kinase II

CK1δ - casein kinase 1δ

CK2 - casein kinase 2

CLSD - cranio-lenticulo-sutural dysplasia

CLTC - clathrin heavy chain 1

CMRD - chylomicron retention disease

COL1A1 - collagen type I alpha 1 chain

COL1A2 - collagen type I alpha 2

COPG - coatomer subunit gamma-1

COPI - coat protein complex I

COPII - coat protein complex II

CuAAC - copper(I)-catalyzed alkyne-azide cycloaddition

DAT - dopamine transporter
ENO1 - alpha-enolase
ER - endoplasmic reticulum
ERES - endoplasmic reticulum exit sites
FL5.12 - murine pro-B cell line
GA - gibberellin
GABA - γ-aminobutyric acid
GALE - UDP-galactose 4-epimerase
GalNAz - N-azidoacetylgalactosamine
GalT - galactosyltransferase
GAP - GTPase activating protein
GAT1 - γ-aminobutyric acid transporter 1
GBF1 - Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1
GEF - guanine-nucleotide exchange factor

ghost
GlcNDAz - diazirine-functionalized GlcNAc analog
GT41 - CAZy glycosyl transferase 41 family
H2B - histone 2B
HAT - histone acetyl transferase

Haunted
HBP - hexosamine biosynthetic pathway
HCD - high-energy collisional dissociation

HCF-1 - Host cell factor-1

hpf - hours post-fertilization

IL-3 - interleukin 3

IP - immunoprecipitation

LWAC - lectin weak affinity chromatography

mOGT - mitochondrial OGT

MS - mass spectrometry

N6 - parental FL5.12 cell line

ncOGT - nucleocytoplasmic OGT

NET - noradrenaline transporter

NORPEG - novel retinal pigment epithelial cell gene

NPC - nuclear pore complex

NRF2 - nuclear factor erythroid 2-related factor 2

PDI – protein disulfide-isomerase

OGA - O-GlcNAcase

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

OGT - O-GlcNAc transferase

OI - osteogenesis imperfecta

P62 - nuclear pore glycoprotein P62
PFK1 - phosphofructokinase 1
PKM2 - pyruvate kinase M2
PP1 - protein phosphatase 1
PP6 - protein phosphatase 6
PTM - post-translational modification
RAI14 - retinoic acid induced gene 14
RBC - red blood cell
SEC - SECRET AGENT
SERT - serotonin transporter
sgRNA - single-guide RNA
SHV – AAVS1 “safe harbor” virus
SILAC - stable isotope labeling with amino acids in cell culture
SLC6 - solute carrier 6
SNP - single nucleotide polymorphism
sOGT - short OGT isoform
SPY - SPINDLY
ssHRP - secreted soluble horseradish peroxidase
SW1353 - human chondrosarcoma cell line
TET - ten eleven translocation
TFG - Trk-fused gene
TPR - tetratricopeptide repeats

tsVSVG-eGFP - temperature-sensitive mutant of the vesicular stomatitis virus glycoprotein

UAP1 - UDP-N-acetylhexosamine pyrophosphorylase

UBA1 - ubiquitin-like modifier-activating enzyme 1

UDP - uridine diphosphate

UPR - unfolded protein response

WGA - wheat germ agglutinin

XL4.1 - FL5.12 transduced with Bcl-X\textsubscript{L}

xvi
1. Introduction

1.1 O-GlcNAc

Post-translational modifications (PTM) are widely used within the cell to increase the complexity and functionality of its proteome. PTMs often consist of the covalent linkage of a small moiety to side chains of amino acids, and can exert myriad effects on protein localization, activity and functionality, stability etc. They range in size and complexity, from single phosphoryl groups to small proteins.

O-linked β-N-acetylglucosamine (O-GlcNAc), a form of intracellular glycosylation, was discovered in the early 1980s (1). O-GlcNAcylation consists of the dynamic addition and removal of a single N-acetylglucosamine monosaccharide covalently attached to the hydroxyl side-chains of serine and threonine residues (Figure 1A). This separates O-GlcNAc from the more familiar O- and N-linked glycans of the secretory pathway, which are large polysaccharides added to proteins in the endoplasmic reticulum (ER) and Golgi. Also, much of its regulation and cycling more closely resembles other signaling PTMs, such as phosphorylation, than it does other forms of glycosylation.

O-GlcNAc was discovered by Hart and coworkers using bovine milk galactosyltransferase, in conjunction with uridine diphosphate (UDP)-[^3H] galactose, to radiolabel terminal N-acetylglucosamine on the surface of lymphocytes. However, the
authors found that the bulk of the labeled GlcNAc residues were intracellular (1).

Shortly after, it was determined that the majority of the O-GlcNAc modified proteins reside in the nuclear and cytoplasmic compartments, and that the cycling of O-GlcNAc was rapid and variable on specific substrates within these compartments, similar to other intracellular PTMs (2,3).

Figure 1: O-GlcNAc, OGT, and OGA

(A) O-GlcNAc consists of the addition of a single β-N-acetylglucosamine reversibly attached to the hydroxyl side-chains of serines and threonines of intracellular proteins. O-GlcNAc is dynamically regulated by OGT and OGA. (B) The three isoforms of OGT. Nucleocytoplasmic OGT (ncOGT) contains twelve TPR repeats (TPR); mitochondrial OGT (mOGT) contains nine TPRs and a mitochondrial targeting sequence (MTS); short OGT (sOGT) contains two TPRs. The C-terminal
split catalytic glycosyltransferase domain (CD) is conserved in all three isoforms. (C) The two isoforms of OGA, long and short (OGA-L and OGA-S). The glycoside hydrolase CD is conserved in both. OGA-L also contains a C-terminal domain that resembles a histone acetyl transferase (HAT).

1.2 O-GlcNAc Transferase and O-GlcNAcase

In mammals, O-GlcNAc is under the control of two enzymes, O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA) (Figure 1B). OGT is the sole enzyme responsible for intracellular O-GlcNAc modification. While there is currently no known consensus sequence identified for OGT substrate proteins, a large portion share a proline-valine-serine/threonine motif (4). In vivo, it is thought that OGT gains specificity via peptide sequences or structural elementals flanking glycosylation sites, and through the aid of proteins that target it both towards substrates and subcellular locations. (5,6).

The human OGT gene is located on the X chromosome near the Xist locus (Xq13.1), which is a region associated with Parkinson’s disease (7,8). There are three human OGT isoforms generated through alternative splicing, which all contain an identical C-terminal catalytic domain, a linker region, and N-terminal tetratricopeptide repeats (TPRs) (9,10). The predominant and longest form of OGT is nucleocytoplasmic OGT (ncOGT) and contains 12 TPRs; an intermediate form, dubbed mitochondrial OGT (mOGT) contains 9 TPRs and a mitochondrial targeting sequence at its N-terminus; the shortest OGT isoform (sOGT) contains only 2 TPRs (9,10). TPRs are a structural motif found throughout nature and consist of a single repeat containing two anti-parallel
alpha-helices (11). When arranged serially, an amphipathic channel is formed, which aids in the binding of target proteins (11). In the case of OGT, these varying length TPRs are thought to modulate substrate specificity and enhance diversity by altering interactions with targeting proteins (5).

OGT activity is highly dependent on the concentration of its nucleotide-sugar donor substrate, UDP-GlcNAc (5,6,12). UDP-GlcNAc is the terminal product of the hexosamine biosynthetic pathway (HBP), which functions as a broad nutrient sensor, integrating flux through glucose, amino acid, fatty acid, and nucleotide metabolic pathways (13). Utilizing UDP-GlcNAc, OGT, a member of the CAZy glycosyl transferase 41 family (GT41) and an inverting glycosyltransferase, converts the α-linked donor to a β-linked PTM on its substrates (14-16). Structural analysis of human OGT points to an ordered bi-bi mechanism, in which UDP-GlcNAc binds in the catalytic site first, followed by the polypeptide acceptor (17,18). It has also been shown that OGT can function as a co-protease for Host cell factor-1 (HCF-1), a transcriptional co-regulator of human cell-cycle progression, which requires proteolytic cleavage for maturation (19,20). This process requires UDP-GlcNAc bound in the active site of OGT, which functions as a co-substrate for cleavage, and is an unprecedented reaction that takes place in the active site for glycosylation. The cleavage occurs between cysteine and glutamate within one of the six proteolytic repeats in HCF-1 and results in the formation
of a pyroglutamate product. Interestingly, mutating the glutamate at the cleavage site to a serine prevents cleavage but results in the serine being O-GlcNAcylated (19).

Currently, HCF-1 is the only known proteolytic target for OGT.

Analogous to OGT, OGA is the sole enzyme known to remove O-GlcNAc from intracellular substrates (Figure 1C). The human OGA gene is on chromosome 10q24.1, which is also associated with Alzheimer’s disease (21-23). Interestingly, a single nucleotide polymorphism (SNP) within OGA is also a known diabetes susceptibility factor in Mexican Americans (24). OGA belongs to CAZy glycoside hydrolase family 84 (GH84) and is a retaining enzyme, which retains the anomeric configuration of the O-linked β-N-acetylglucosamine after hydrolysis, resulting in a β-GlcNAc product (15,25,26). Unlike the vast majority of retaining glycoside hydrolases, which perform catalysis through the formation of a covalent glycosyl-enzyme intermediate, OGA uses an alternative mechanism in which catalysis occurs through an oxazoline intermediate (15).

Like OGT, the function and substrate specificity of OGA are thought to be driven via protein-protein interactions with various interacting partners, often forming in complexes with each other (27). There are two major splice isoforms of OGA within mammals, long and short. The long OGA isoform, found throughout the cytoplasm and nucleus, contains a catalytic domain and a domain that resembles histone acetyl
transf erases (HATs) (28,29). The OGA HAT domain is lacking key residues thought to be required for an active acetyltransferase, and most studies have failed to detect HAT activity by OGA (30-32), leaving its function an open question. The short isoform of OGA contains only the catalytic domain and has been shown to interact with lipid droplets, possibly controlling lipid droplet assembly and mobilization by remodeling the droplet surfaces via local proteasome activation (28).

1.3 Evolutionary Conservation of O-GlcNAc

O-GlcNAc is broadly conserved across animals, plants, and some fungi. OGT-like proteins from the protozoans Giardia lamblia and Cryptosporidium parvum have shown to have in vitro glycosyltransferase activity, and OGT-like sequences have been found in a large number of Gram-negative and Gram-positive eubacteria, but not in Archaea (33). Alignments of the OGT-like catalytic domain sequences suggest a high level of conservation within the functional amino acids, implying that OGT activity may be conserved across these predicted proteins (33). OGA has also been described in prokaryotes, and GH84 enzymes are encoded in many bacterial genomes (34). Many of these enzymes are distinct from eukaryotic OGA and apart from a conserved catalytic domain, contain additional domains or secretion signals, leaving their physiological function unclear and unstudied (34). However, in vitro studies using several bacterial OGAs have demonstrated effective removal of O-GlcNAc from O-GlcNAcylated
eukaryotic proteins (35,36). It has also been shown that *Trichoplax adhaerens*, the most extant basal animal, has functional OGT, OGA, and nucleocytoplasmic O-GlcNAcylation. Interestingly, OGT from *Trichoplax adhaerens* was able to rescue *Drosophila* OGT mutants that would normally die as pharate adults, demonstrating a conservation of activity and function (37).

O-GlcNAc cycling regulates many cellular processes, including transcription, cell metabolism, cell cycle progression, and cell death (38,39). Proper O-GlcNAc cycling is essential in mammals, as knockout of the OGT gene in mice is lethal at the one cell stage, whereas OGA knockout mice die as neonates with development delays (7,40). Due to the requirement of OGT and OGA in mammalian systems, other model organisms have been useful in the study of genetic alterations of O-GlcNAc cycling. Both *ogt* and *oga* knockouts in *C. elegans* alter macronutrient storage and dauer formation, exhibiting signaling changes with parallels to that of mammalian insulin resistance (41,42). A similar phenomenon is observed in *Drosophila melanogaster*, where Ogt function is required for viability, and the phenotypes associated with knockdown of Ogt and Oga and perturbed O-GlcNAc cycling mimic pancreatic β-cell dysfunction and glucose toxicity related to insulin resistance in mammals (43,44). In *Arabidopsis thaliana* there are two OGT homologs, SECRET AGENT (SEC) and SPINDLY (SPY) (33). SPY functions as a gibberellin (GA) signaling repressor, where loss of function results in a dwarf
phenotype, SEC mutants have a mild phenotype, though it has been shown to O-GlcNAcylate DELLA proteins, which mediate cross-talk between GA and other signaling pathways (33,45-49). There is likely some overlapping function between SEC and SPY, but a SEC/SPY double mutant is embryonic-lethal (33,46). Interestingly, in Arabidopsis, an OGA homolog has not yet been identified.

1.4 O-GlcNAc as a Signaling Modification

As mentioned earlier, OGA and OGT gain specificity through interacting partner proteins, either through changes in localization or by altering the binding affinity of substrates. Some examples of binding partners include TET2, Milton, and PP1 (50-52). Ten eleven translocation (TET) enzymes are regulators of gene transcription and it has been demonstrated that TET2 forms a complex with OGT, which may be responsible for chromatin and gene expression alterations. This TET2-OGT complex targets OGT to chromatin and regulates histone 2B (H2B) O-GlcNAcylation at S112 (51). O-GlcNAcylation at this residue has also been shown to promote K120 monoubiquitination by functioning as an anchor for Bre1A, a histone H2B ubiquitin ligase (53).

Milton, a mitochondrial adaptor protein, has also been shown to form a stable complex with OGT both in vitro and in vivo (52,54-56). This OGT-Milton complex is thought to mediate transport of mitochondria throughout the cell. Milton is required for
targeting OGT to the mitochondrial surface and O-GlcNAcylation of Milton on key residues is required to arrest mitochondrial mobility (52).

Serine/threonine protein phosphatase 1 (PP1) is responsible for the removal of phosphoryl groups from serines and threonines of thousands of intracellular proteins. Interestingly, PP1 is in stable and active complexes with OGT (50). During mitosis, a transient complex consisting of mitotic kinase Aurora B, PP1, OGT, and OGA alter the posttranslational status of target proteins, such as vimentin (57). These complexes provide evidence for the highly dynamic interplay between phosphorylation and O-GlcNAc within cells.

Improper cycling of O-GlcNAc has been implicated in myriad human diseases, including cancer, diabetes and cardiovascular dysfunction. For example, phosphofructokinase 1 (PFK1), an important glycolytic enzyme, becomes O-GlcNAcylated at S529 under hypoxic conditions, a feature common to many tumors (58,59). O-GlcNAc modification of PFK1 at S529 inhibits its activity and redirects metabolic flux to the oxidative pentose phosphate pathway, providing a selective growth advantage to cancer cells through increased availability of biosynthetic precursors and NADPH to maintain redox homeostasis (59). Blocking O-GlcNAcylation at S529 reduced cancer cell proliferation in vitro and reduces tumor mass in a mouse
xenograft model, demonstrating a novel mechanism for the regulation of metabolic flux, and possibly highlighting a target for future therapeutics (59).

Diabetes is accompanied by many complications, one of which is hyperglycemia, which increases flux through the HBP, leading to increases in UDP-GlcNAc (60). This boosts global O-GlcNAc levels and can alter proper cycling of O-GlcNAc on native substrates (60,61). Ca²⁺/calmodulin-dependent protein kinase II (CaMKII) is one substrate which becomes improperly regulated under hyperglycemic conditions. Under normoglycemic conditions, CaMKII has important regulatory functions in the heart and brain, but chronic activation can be pathological. Under hyperglycemic conditions, S279 is O-GlcNAcylated, which autonomously activates CaMKII in the heart, contributing to cardiac mechanical dysfunction and arrhythmias (61).

1.5 COPII

One-third of the eukaryotic proteome passes through the secretory pathway, and the journey of these secreted proteins begins in the endoplasmic reticulum (ER) (62-65). The ER is a site of protein folding, co- and posttranslational modification, and quality control. However, most proteins entering the ER are meant to function at another location. The coat protein complex II (COPII) consists of protein coated carriers that transport cargo early in the secretory pathway from the ER, a process essential for life (66-68).
The COPII pathway and core components were initially discovered and characterized in studies of yeast that eventually resulted in a 2013 Nobel Prize in Physiology (69). Initially, Schekman and Novick examined temperature-sensitive yeast mutants for secretory enzyme accumulation. Using transmission electron microscopy, they noticed that some of the mutants were full of vesicles and accumulated internal membranes. They hypothesized that these mutant cells would have different physicochemical properties. They developed a screening technique that took advantage of these mutants increased buoyant density, a result of retained secretory pathway products. This technique ushered in the discovery and characterization of the first genes responsible for ER to Golgi transport.

The formation of a COPII carrier begins when the small GTPase SAR1 is activated by its specific guanine-nucleotide exchange factor (GEF), SEC12, an ER-resident transmembrane protein (Figure 2) (67,68). Upon activation, SAR1-GTP undergoes a conformational change, inserting its N-terminal amphipathic α-helix into the ER membrane (67,68). SAR1-GTP then recruits the heterodimer made of SEC23/SEC24, also known as the inner COPII coat (67,68). SEC23A is proposed to function as a GTPase activating protein (GAP) for SAR1 (68,70,71), and SEC24 recruits cargo, either directly or through adaptor proteins (67,68,72,73). Next, the outer COPII coat, consisting of SEC13/SEC31 heterotetramers, a dimer of SEC13/31 dimers (67,68).
SEC31 potentiates the GAP activity of SEC23 and drives carrier formation by polymerizing into a polyhedral cage structure and SEC13 contributes to the rigidity of the COPII coat (68,74). Next, the mature COPII carrier splits from the ER. COPII carrier scission from the ER membrane is not well understood. However, there are a few contributing factors, which include enhanced SAR1 GTPase activity on membranes with elevated curvature, such as the neck of a budding carrier, and SEC31 enhancing the GAP activity of SEC23 (74,75).

There are five COPII components required for minimally reconstituted in vitro carrier formation: SAR1, SEC23, SEC24, SEC13, and SEC31 (76). These components are largely conserved across eukaryotes and gene duplication events have created multiple paralogs in many metazoans. Vertebrates have two SAR1 paralogs, two SEC23 paralogs, four SEC24 paralogs, one SEC13 paralog, and two SEC31A paralogs (77,78). Interestingly, Arabidopsis thaliana has more COPII components than any other known eukaryote, with five SAR1, two SEC13, seven SEC23, three SEC24, and two SEC31 isoforms (79,80).
COPII consist of protein coated carriers that transport cargo early in the secretory pathway from the ER. Briefly, COPII carrier formation begins when the cytosolic GTPase SAR1 binds GTP and inserts into the ER membrane. This process is mediated by SEC12, an ER-membrane anchored guanine nucleotide exchange factor. Next, SAR1-GTP recruits the inner coat consisting of SEC23/SEC24 heterodimers, which mediates cargo selection and concentration via protein-protein interactions with SEC24 or cargo adaptor proteins. The inner coat recruits the outer coat of SEC13/SEC31 heterotetramers, which promote polyhedral cage polymerization, membrane curvature and eventual scission.

1.6 SAR1

SAR1 is a smallGTPase and is a part of the COPII inner coat. It is required for ER-derived transport carriers and is conserved from yeast to mammals (65,81,82). The two SAR1 paralogs found in vertebrates, SAR1A and SAR1B, are 90% identical in humans, but they affect COPII transport carrier formation differently in vivo (78,83,84). SAR1 functions as a curvature-sensing component, preferentially binding to, and hydrolyzing
GTP faster on membranes with elevated curvature, which may explain a process of COPII transport carrier fission (75). As membrane curvature increases SAR1 membrane binding is elevated, which further deforms the local membrane, promoting more SAR1 binding and GTP hydrolysis (75). These events may sufficiently destabilize the bilayer to facilitate membrane fission and COPII carrier release (75).

In humans, mutations in SAR1B lead to chylomicron retention disease (CMRD), also known as Anderson’s disease (ANDD), which is characterized by malabsorption of lipids from the diet and accumulation of chylomicrons in enterocytes (78,83,85). Interestingly, CMRD patients express increased levels of SAR1A, but these are unable to compensate for the lack of SAR1B (84). Similar observations were made in a zebrafish model of CMRD where Sar1b was depleted (86). The resulting Sar1b-deficient zebrafish had dietary lipid accumulation in enterocytes, reminiscent of CMRD (86). However, they also had deficiencies in chondrocyte capacity to produce and maintain type II collagen (86). Interestingly, no gross morphological deficit in Sar1a-deficient zebrafish was observed, but the combined depletion of Sar1a and Sar1b resulted in a more severe Sar1b phenotype (86). Furthermore, Sar1a overexpression was only able to partially rescue Sar1b deficiency, suggesting some paralog overlap in function (86). To date no mutations in SAR1A have been associated with CMRD or any other human syndrome, and it is unknown what phenotypes could be associated with SAR1A variants (78).
1.7 SEC23

SEC23 functions as the GAP for SAR1, is a part of the COPII inner coat, and is essential for proper COPII function (67,78,87). Due to a gene duplication event, vertebrates have two paralogs of SEC23, SEC23A and SEC23B (77,78). In humans these two paralogs have 96% similarity and are 85% identical. High conservation is also seen across vertebrates, as human and zebrafish SEC23As are 98% similar and 91% identical. Interestingly, the differences in these proteins are localized to a highly divergent region of ~18 amino acids that is near the SEC24 interface, which could alter SEC24 paralog binding affinity, affecting cargo selection. Although these two paralogs are highly similar, they have distinct but largely overlapping functions, as exhibited by the different disease phenotypes in humans and animal models.

Mutations in SEC23A cause cranio-lenticulo-sutural dysplasia (CLSD), which is characterized by facial dysmorphisms, skeletal defects, late-closing fontanels, and cataracts, caused by the mistrafficking of collagen (88-90). The first characterized CLSD mutation, F382L SEC23A, is located near the region of SEC23A that binds and recruits SEC31 (88). Failure to recruit SEC31, thereby preventing membrane fission, leads to a reduction of packaging cargo proteins \textit{in vitro} and is accompanied by swelling of the ER with trapped cargo \textit{in vivo} (88). M702V SEC23A, was more recently characterized and does not suffer from a general ER export defect (89). M702V SEC23A enhances the GAP
activity of SEC23A against SAR1, causing premature dissociation of the COPPII transport carrier from the membrane (89). This premature fission resulted in retention of procollagen within the ER but did not disrupt the other cargo molecules tested. Both reported missense mutations are hypomorphic and no null mutations in SEC23A have been reported in humans, most likely due to their lethality. Interestingly, while expression of both paralogs of SEC23 is ubiquitous, calvarial osteoblasts, which are primarily responsible for ossification of the cranial suture, express very little SEC23B (91). This suggests that tissues affected in CLSD patients may lack enough total SEC23 to compensate for the SEC23A loss-of-function mutation (91).

Zebrafish crusher mutants have a nonsense mutation (L402X) in the sec23a gene and develop normally until the onset of craniofacial chondrogenesis (92). crusher mutants exhibit short body length, small and malformed head skeleton, and absence of cartilaginous ear capsules, a phenotype reminiscent of CLSD in humans (92). Similar to CLSD, this phenotype can be attributed to a specific cargo, collagen. In crusher mutant chondrocytes, collagen is not detected in the extracellular space, but retained within the cell in large vesicular structures (92).

Mice deficient in SEC23A die during mid-embryogenesis (E12.5), after neural tube opening (E11.5), a defect caused by collagen secretion defects in fibroblasts, which compromises the tensile strength of the skin (93). Interestingly, loss of one SEC23
paralog increases the amount of the remaining paralog in cells. However, this increased SEC23B expression in SEC23A-deficient mice is not sufficient to compensate (93). Also, reducing levels of the remaining paralog in SEC23A-deficient mice leads to earlier lethality of the embryos, indicating partial overlap of function in these paralogs in vivo (93).

Mutations of SEC23B in humans cause congenital dyserythropoietic anemia type II (CDAII), which is characterized by mild to moderate anemia, bi- or multinucleated erythroblasts, aberrant glycosylation of erythrocyte band 3, and red blood cell (RBC) lysis in some acidified normal sera (78,94,95). Knockdown of Sec23b in zebrafish leads to aberrant erythrocyte development, but other CDAII characteristics were not observed (95). Sec23b knockdown in zebrafish also causes craniofacial defects similar to that of crusher zebrafish (92). SEC23B-deficient mice die as neonates and have severe developmental defects in the pancreas and other exocrine glands, a phenotype distinct from patients with CDAII (96). Mice with the SEC23B deficiency restricted to the hematopoietic compartment do not develop anemia or any CDAII characteristics (97). However, levels of SEC23A in murine erythrocytes during differentiation remain high when compared to humans, where SEC23A is undetectable (98). These higher levels of SEC23A support the idea of paralog compensation and may explain the lack of common CDAII characteristics in hematopoietic compartment SEC23B-deficient mice (98). When
SEC23B deficiency is restricted to the pancreas in mice they die as neonates, which sufficiently explains the global SEC23B-deficient mice lethality (99,100). However, SEC23A is not required for murine pancreatic development (100). These phenotypes highlight the functional differences in these two highly similar proteins, as well as the differences in orthologs between species.

*Drosophila melanogaster* has only one Sec23 homolog which is encoded by the *haunted (hau)* gene (101). Embryos with a mutant *hau* produce a discontinuous or thin larval cuticle and die at the first instar larval stage within the egg case. Also, the ER of *hau* larval tracheal cells are dilated, likely a result of mistrafficked or retained cargo (101). Similarly, *Caenorhabditis elegans* has only one sec-23 gene, which is required for proper cuticle secretion, a collagenous exoskeleton that surrounds postembryonic worms (102,103). Loss of SEC-23 in *C. elegans* also causes defects in oogenesis, morphogenesis, and eventually death (102,103). In both flies and worms, SEC23 is necessary for proper organismal function, as well as maintaining the extracellular matrix.

**1.8 SEC24**

SEC24, a COPII inner coat protein that binds SEC23, is responsible for cargo recognition and binding (67,68,78). SEC24 proteins are characterized by clusters of highly conserved residues within the C-terminal domain and a divergent N-terminal
domain (104,105). In yeast, Sec24p is the only COPII component in which multiple orthologs have been identified (106). However, only Sec24p is essential (107,108). Interestingly, the other Sec24p homologues identified (Iss1p, Lst1p) interact directly with Sec23p and can form competent transport carriers in place of Sec24p, suggesting possible specialized function for cargo recognition and packaging (72,73,107,108).

Mammals express four paralogs of SEC24: A, B, C, and D. These can be grouped in pairs, SEC24A/SEC24B and SEC24C/SEC24D (68,104,105). These pairs retain high sequence similarity and preferentially bind similar ER export signals, as demonstrated in yeast and in vitro using model proteins (72,73). Interestingly, SEC24 paralog specific cargo trafficking requirements are observed in many diseases and animal models of COPII function (109-115). In mice, SEC24A is non-essential, as complete deficiency is compatible with normal survival and development (109). Interestingly, mice deficient in SEC24A exhibit markedly reduced plasma cholesterol due to a specific block in PCSK9 trafficking, a regulator of surface LDL receptor (109). This contrasts heavily with SEC24B mutant mice, who exhibit craniorachischisis, a severe birth defect that results in a neural tube closure deficit (111). This mutation disrupts the proper SEC24B-dependent trafficking of VANGL2, a core planar cell polarity signaling component (110,111).

Mice deficient in either SEC24C or SEC24D die during embryogenesis, but the embryonic lethality of SEC24C-deficient mice occurs significantly later than that of
SEC24D-deficient mice (112,114). Interestingly, mice haploinsufficient for either SEC24C or SEC24D exhibit no apparent abnormalities (112,114). Zebrafish bulldog mutants all have varied length truncations in the sec24d gene, but all resulting phenotypes are largely identical (115). This suggests that the C-terminus, common between all the mutants, is critical for proper protein function (115). bulldog mutants largely resemble the previously described crusher mutants, exhibiting craniofacial defects and defects in the collagen secretion of chondrocytes (115). Sec24c depletion in zebrafish does not affect craniofacial cartilage morphogenesis. However, Sec24c depletion in bulldog mutants led to earlier delays in craniofacial development, suggesting that these paralogs can compensate for one another at early stages of chondrogenesis (115). Drosophila melanogaster has only one Sec24 paralog which is encoded by the ghost (gho) gene and mutants exhibit a cuticle secretion defect reminiscent of haunted mutants (101). C. elegans has two SEC24 isoforms, SEC-24.1 and SEC24.2, which exhibit two distinct but similar phenotypes when depleted (103,104). Germline maintenance requires COPII-mediated trafficking. Depletion of SEC-24.1 causes compartmentalization to fail, resulting in multinucleation. However, depletion of SEC-24.2 causes oocytes to fail to bud away from the syncytium efficiently, but still causes sterility in adult animals (103).

In humans, mutations in SEC24B have been identified in stillborn fetuses with neural tube defects. These mutations inhibit protein stability and the recruitment of
VANGL2, similar to SEC24B-deficient mice (110). Mutations in SEC24D cause a form of osteogenesis imperfecta (OI), which is characterized by reduced bone mass, increased bone fragility, bone deformity, and growth deficiency (113). Up to 90% of individuals with a classical OI phenotype have mutations in the genes for collagen type I alpha 1 chain (COL1A1) and collagen type I alpha 2 (COL1A2). However, mutations in SEC24D cause similar clinical observations by disrupting efficient proper procollagen export from the ER (113).

In addition to PCSK9, VANGL2, and collagen some other examples of SEC24 paralog specific cargoes include Ebola virus matrix protein VP40, and solute carrier 6 (SLC6) family member neurotransmitters (116). SLC6 family members include serotonin transporter (SERT), γ-aminobutyric acid (GABA) transporter 1 (GAT1), dopamine transporter (DAT), and noradrenaline transporter (NET) (117,118). VP40 is important for virion formation and viral egress from cells. SEC24C interacts with amino acids 303 to 307 on VP40, which are required for its accumulation at the plasma membrane (116). SEC24C is also responsible for the specific trafficking of SERT, but not the other SLC6 family neurotransmitters, which are dependent on SEC24D for their specific trafficking (117,118). On all SLC6 family members there is a SEC24-binding motif, dubbed the RI/RL/RV motif. SERT has differences from GAT1, NET, and DAT in this binding motif, which may contribute to the specific SEC24 paralog requirements of these proteins (118).
1.9 SEC13 and SEC31

SEC13 and SEC31 form a heterotetramer consisting of two SEC13/SEC31 dimers and make up the outer coat of COPII transport carriers (68). Most eukaryotes have only one SEC13 paralog. However, vertebrates have two paralogs of SEC31, A and B. Both SEC13 and SEC31 are essential for proper anterograde traffic from the ER (119). However, yeast deficient in SEC13 can form COPII transport carriers that only select specific cargo proteins for ER export (120). Intriguingly, SEC13 has an important function outside of COPII and is also a part of the nuclear pore complex (NPC), which mediates nucleocytoplasmic traffic (121-123). This NPC functionality of Sec13 has been shown to play an essential role in the transcription of developmentally induced genes in Drosophila melanogaster (123).

In mice SEC13 knockout is lethal, but a mouse model with low levels of SEC13 resulted in specific immunological defects (124). These hypo-SEC13 mice had no bulk deficit in secretion, but the levels of certain immune factors including interferon-γ, MHC I, and MHC II were decreased (124). Because SEC13 functions as both a COPII and NPC component, determining what role SEC13 plays in some phenotypes can be difficult. The first model to distinguish these functions was in zebrafish, where the COPII function of Sec13 was determined to be responsible for the organogenesis of digestive organs and craniofacial cartilage, and the NPC function of Sec13 was required for retinal
development (121). Mutations in SEC13 have not been identified in humans, possibly due to its important role in both COPII and the NPC.

Knockdown of Sec31a in zebrafish largely resembles knockdown of Sec13, which is likely due to the requirement of a competent COPII outer coat, rather than individual components (121). Also, the specific functions and paralog differences of SEC31A and SEC31B have yet to be determined. However, this diversity in components most likely allows for further cargo discrimination or other specialized functions.

**1.10 Regulation of COPII via PTMs**

Although the fundamental steps of transport carrier assembly are relatively well understood, little is known about how vertebrate cells modulate COPII activity in response to developmental changes, metabolic demands, fluctuating signals, or stress. Recently, PTMs have emerged as a general regulator mechanism of the COPII pathway. Phosphorylation, ubiquitination and O-GlcNAcylation of COPII components have been shown to affect transport carrier biogenesis, protein localization, and cargo selectivity. A few examples will be highlighted here.

Phosphorylation governs several aspects of COPII trafficking. A Golgi-associated kinase casein kinase 1δ (CK1δ) phosphorylates SEC23 and is required for COPII transport carrier uncoating and fusion (125,126). Interestingly, this phosphorylated COPII coat cannot rebind the ER to initiate a new round of COPII biogenesis. Protein
phosphatase 6 (PP6) is responsible for dephosphorylating the COPII coat allowing recycling of the COPII coat (126). These events ensure orderly carrier formation, fusion, and directionality.

Recently, COPII transport carriers have been shown to become a membrane source for autophagosomes (127,128). Autophagy is an evolutionarily conserved response to starvation and stress, where superfluous or damage proteins or organelles are collected in autophagosomes and eventually degraded and recycled (127,128). Phosphorylation of three amino acids on the surface distal to the membrane of SEC24 enhances its interaction with autophagy-related protein 9 (ATG9), a component of the autophagy machinery required for autophagosome initiation (129).

Ubiquitination of Sec31 is responsible for the transport of larger cargoes (130). Normal COPII transport carriers range in size from 60 - 80 nm in diameter but must increase in size to accommodate 300 – 600 nm procollagen fibers or chylomicrons (130-133). The ubiquitin ligase CUL3-KLHL12 catalyzes the monoubiquitination of SEC31, which drives the assembly of large COPII coats required for collagen export (130). Interestingly, none of the 65 lysine residues was essential for ubiquitination and CUL3 targeted alternative lysines if the primary target of modification was blocked (130). How this monoubiquitination affects COPII coat size or structure is not known. SEC31 can also be phosphorylated by casein kinase 2 (CK2), a modification which reduces the
association of SEC31 with ER membranes and SEC23, decreasing ER-to-Golgi trafficking (134).

SEC24 is phosphorylated during mitosis, which prevents membrane binding (135). This could possibly contribute to the unknown mechanism by which COPII trafficking is suspended during this phase of the cell cycle. Interestingly, during interphase, SEC24 is modified by O-GlcNAc, a modification lost upon entering mitosis (135). Recently, we and others have shown that multiple COPII components including SEC23, SEC24, and SEC31 are modified by O-GlcNAc (87, 135-139). However, little is known about the functional effects of this modification on the COPII pathway.
2. Dynamic Glycosylation Governs the Vertebrate COPII Protein Trafficking Pathway

2.1 Introduction

One-third of eukaryotic proteins pass through the secretory pathway for targeting to specific locations, including the endoplasmic reticulum (ER), Golgi, plasma membrane and extracellular space (62-64). As a result, properly regulated protein secretion is required for cellular and organismal physiology. In particular, the coat protein complex II (COPII), which mediates anterograde trafficking from the ER, is a key component of the early secretory pathway (67,140-144). The formation of a COPII transport carrier begins when the cytosolic GTPase SAR1 binds GTP and inserts into the ER membrane, a process facilitated by the ER-anchored guanine nucleotide exchange factor SEC12 (67,140,141). Active SAR1-GTP recruits SEC23/SEC24 heterodimers, which effect cargo loading into the nascent carriers via protein-protein interactions among SEC24 and client or adaptor proteins (67,71,140,141,145). Then, heterotetramers of SEC13/SEC31 assemble over the SAR1/SEC23/SEC24 pre-budding complex, forming the outer layer of a polyhedral cage that promotes membrane curvature and scission, resulting in a mature COPII transport carrier (67,140,141).

The COPII pathway is required for protein sorting and cell viability in a wide range of organisms (67,140,141), and genetic defects in COPII components cause a variety of human diseases, including cranio-lenticulo-sutural dysplasia (CLSD), hematological disorders, a subtype of osteogenesis imperfecta (OI), and multiple

Interestingly, the molecular etiology of these diseases can often be attributed to aberrant trafficking of particular COPII client cargoes. For example, mutations in SEC23A (which cause CLSD) and SEC24D (which cause a subtype of OI) both disrupt collagen trafficking, leading to chondrocyte dysfunction, impaired skeletogenesis, craniofacial disease and bone deformities (88,90-92,113,115,149,152-154). Despite this broad pathophysiological importance, major aspects of COPII trafficking remain poorly understood. For instance, while the fundamental steps of COPII assembly are relatively well characterized, little is known about how vertebrate cells modulate this activity in response to developmental cues, fluctuating signals, metabolic demands, or stress (67,140,141). The COPII cargo load changes dramatically in both normal (e.g., stimulated B cells or differentiating pancreatic β islets) and pathological (e.g., nutrient, redox, or proteostasis stress) contexts (155-161), but the mechanisms through which the COPII machinery adjusts to these changes are largely unclear.

Post-translational modifications (PTMs) represent one likely general mode of COPII pathway regulation. For example, phosphorylation (125,126,134,162-165) and ubiquitination (130,166,167) of individual COPII proteins govern particular aspects of vesicular trafficking. Recently, we (136) and others (135,137,138,168) have also shown that multiple COPII components – including SEC23, SEC24 and SEC31 – are modified by O-linked β-N-acetylglucosamine (O-GlcNAc) in mammalian cells. O-GlcNAc is a major
intracellular PTM, reversibly decorating serine and threonine side-chains of thousands of nuclear, cytoplasmic and mitochondrial proteins, and is broadly conserved across animals, plants and perhaps some fungi (38,169-171). In animals, O-GlcNAc is added by O-GlcNAc transferase (OGT) and removed by O-GlcNAcase (OGA), both ubiquitous nuclear/cytoplasmic enzymes (38,169-171). O-GlcNAc cycling controls myriad processes, including cell metabolism, cell cycle progression and cell death (38,169-171), and is essential, as genetic ablation of OGT or OGA is lethal in mice (172,173). Because several core COPII proteins are O-GlcNAcylated (135-138,168), we reasoned that this intracellular glycosylation might regulate protein trafficking. Consistent with this notion, SEC24 O-GlcNAcylation is high during interphase but low in mitosis, when COPII-dependent trafficking is suspended through an unknown mechanism (135), suggesting that O-GlcNAc may dynamically control COPII function during cell cycle progression.

However, the mechanistic and functional effects of O-GlcNAc on the COPII pathway have not been explored.

Here, we address the prevalence, biochemical effects, and physiological importance of O-GlcNAcylation in the COPII pathway. We show that O-GlcNAc is widespread in the COPII system, and that site-specific glycosylation of key COPII proteins impacts protein secretion. Moreover, we use a chemical biology approach to demonstrate that dynamic O-GlcNAcylation mediates the protein-protein interactions of COPII components, identifying a new potential biochemical mechanism for cargo
trafficking regulation by this PTM. Using a genome-engineered human cell model, we show that individual glycosylation sites of SEC23A, an essential COPII component, are required to traffic endogenous collagen. Finally, we demonstrate that mutation of specific O-GlcNAc sites impairs in vivo SEC23A-dependent collagen trafficking in a zebrafish model of CLSD. Together, our results establish O-GlcNAc as a conserved and prevalent regulatory modification in the vertebrate COPII pathway.

2.2 Materials and Methods

Chemical synthesis – 5SGlcNAc was synthesized as described (174) and was a gift of Dr. Benjamin M. Swarts (Central Michigan University). Thiamet-G and Ac3GlcNDAz-1-P(Ac-SATE)2 (“GlcNDAz”) were synthesized as described (175,176) by the Duke Small Molecule Synthesis Facility. All other chemicals were purchased from Sigma-Aldrich unless otherwise indicated.

Western blotting – Samples were resolved on Tris-glycine SDS-PAGE gels and electroblotted onto PVDF membrane (88518, ThermoFisher) using standard methods (177). Membranes were blocked with Tris-buffered saline with 0.1% Tween (TBST) with 5% bovine serum albumin (BSA). All antibody dilutions were prepared in TBST with 5% BSA. Membranes were incubated with primary antibodies overnight at 4 °C, washed three times in TBST, incubated with secondary antibodies for 1 hour at room temperature, washed three times in TBST, and developed via enhanced chemiluminescence (ECL) according to the manufacturer’s instructions (WesternBright

**Cell culture** – 293T, HeLa, HeLa/UAP1(F383G), 293T/UAP1(F383G), COS7(tsVSVG-eGFP) and SW1353 (including all engineered derivative lines) were cultured in Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum, 100 units/mL penicillin, and 100 µg/mL streptomycin in 5% CO₂ at 37 °C. Transduced cell lines were also supplemented with puromycin during regular passaging at the following
concentrations: 1.5 µg/mL for HeLa/UAP1(F383G); 2 µg/mL for 293T/UAP1(F383G) and COS7(tsVSVG-eGFP).

**Mammalian expression vectors** – pSin-EF2-UAP1(F383G) was a gift from Dr. Jennifer Kohler, UT Southwestern (175). pmTurquoise2-Golgi was a gift from Dr. Dorus Gadella (Addgene plasmid # 36205). mCherry-ER-3 was a gift from Dr. Michael Davidson (Addgene plasmid # 55041). The pMSCV tsVSVG-eGFP was generated by cutting tsVSVG-eGFP from the parent vector (pEGFP-VSVG, Addgene, plasmid #11912) with BglII and NotI and ligating it into pMSCV via standard methods (177). pcDNA4-SEC23A-myc-6xHis, pcDNA4-SEC31A-myc-6xHis, and pcDNA3-myc-6xHis-SEC24C, p3xFLAG-SEC24C, pcDNA4-TFG-myc-6xHis and p3xFLAG-TFG were generated by amplifying the open reading frames from their respective cDNA and ligating into pcDNA4/myc-6xHis, pcDNA3-myc-6xHis, or p3xFLAG-CMV-10 using standard methods (177). Primers for site-directed mutagenesis were designed using QuikChange Primer Design (http://www.genomics.agilent.com/primerDesignProgram.jsp).

Mutagenesis reactions were performed using Phusion polymerase (M0530, New England Biolabs) essentially according to the manufacturer’s instructions, but with an extra initial 30-second 98 °C denaturation step prior to the addition of polymerase. Mutagenesis reactions were digested with DpnI and transformed into One Shot TOP10 chemically competent E. coli (C404010, ThermoFisher). The following primers and their reverse-complements were used for site-directed mutagenesis:
Stable cell line creation – UAP1(F383G) and tsVSVG-eGFP retro- and lentiviruses were generated in 293T-17 cells using calcium phosphate transfection and standard methods as described (136). Medium containing virus was filtered through pre-wetted 0.45 μm PVDF syringe filters and 1 mL of viral supernatant was added directly to 10 cm plates of 50% confluent cells. HeLa/UAP1(F383G) and 293T/UAP1(F383G) were selected using 1.5 μg/mL and 2 μg/mL puromycin, respectively (175). COS7(tsVSVG-
eGFP) were selected using 2 µg/mL puromycin until cells exhibited green ER fluorescence.

**Immunoprecipitation and tandem affinity purification** – Cells transfected with myc-6xHis- or 3xFLAG-tagged proteins were lysed in immunoprecipitation (IP) lysis buffer (1% Triton X-100, 0.1% SDS, 150 mM NaCl, 1 mM EDTA, 20 mM Tris-HCl pH 7.4) supplemented with protease inhibitor cocktail (P8340, Sigma), plus 50 µM UDP, and 5 µM PUGNAc (17151, Cayman Chemicals) to inhibit endogenous OGT and hexosaminidases, respectively. Lysates were probe-sonicated, cleared by centrifugation and quantified by BCA protein assay (23225, ThermoFisher) according to the manufacturer’s instructions. IPs were performed on 2-10 mg total protein (for WB) or 50-100 mg total protein (for MS analysis). Cleared lysates were adjusted to final total protein concentration of ~2-5 mg/mL using IP lysis buffer supplemented with protease inhibitor, UDP and PUGNAc, and 2 µg of mouse monoclonal anti-c-myc (9E10) (sc-40, Santa Cruz Biotech) per 1 mg of total protein were added and rotated overnight at 4 ºC. 50 µL washed protein A/G UltraLink Resin (53133, ThermoFisher) were added to the lysate and rotated at room temperature for 1 hour. Beads were washed three times with 1 mL of IP lysis buffer and then eluted. For IPs, elution was performed with IP lysis buffer plus 2X SDS-PAGE loading buffer (5X SDS-PAGE loading buffer: 250 mM Tris pH 6.8, 10% SDS, 30% glycerol, 5% β-mercaptoethanol, 0.02% bromophenol blue) and heating at 95 ºC for 5 minutes. Eluents were then analyzed by SDS-PAGE and WB. For
tandem affinity purification, beads were eluted twice in 500 µL using Ni-NTA wash buffer (8 M urea, 300 mM NaCl, 1% Triton X-100, 5 mM imidazole) with rotation. The two 500 µL elutions were pooled, and 50 µL washed 6xHisPur Ni-NTA resin (88223, ThermoFisher) were added to the eluate and rotated for 2 hours at room temperature. The Ni-NTA resin was washed three times with 1 mL of Ni-NTA wash buffer. Final elution from the Ni-NTA was performed using 8 M urea with 250 mM imidazole.

**Digitonin fractionation** – Cells were seeded such that they were ~80% confluent in a 10 cm plate on the day of harvest. Cells were washed with 10 mL cold PBS and incubated on ice for 10 minutes. PBS was removed from the plate and 1 mL permeabilization buffer (110 mM KOAc, 25 mM HEPES pH 7.2, 2.5 mM Mg(OAc)₂, 1 mM EGTA, 0.015% digitonin, 1 mM DTT, protease inhibitor cocktail) was added, taking care to not dislodge the cells, and incubated on ice for 5 minutes. Soluble material was collected from the plate into a new chilled 1.5 mL centrifuge tube and retained as the cytoplasmic fraction. The remaining cells were washed gently with 5 mL of wash buffer (110 mM KOAc, 25 mM HEPES pH 7.2, 2.5 mM Mg(OAc)₂, 1 mM EGTA, 0.004% digitonin, 1 mM DTT). After washing, the cells were incubated on ice in IP lysis buffer supplemented with protease inhibitor cocktail for 5 minutes and then scraped and moved into a chilled 1.5 mL centrifuge tube to obtain the membrane fraction. Both fractions were probe-sonicated and cleared by centrifugation. Equal volumes from each
fraction were methanol/chloroform-precipitated to concentrate. Fractions were analyzed by SDS-PAGE and WB.

**Densitometry** – Western blot band densitometry readings were performed on scanned blot images using the gel analyzer tool within the Fiji software package (178). Raw densitometry values were used to calculate cytoplasmic:membrane ratios for each protein within each experimental replicate.

**tsVSVG-eGFP imaging** – Two days prior to imaging, 200,000 COS7(tsVSVG-eGFP) cells were plated onto 35 mm glass-bottom dishes (P35GCol-1.5-14-C, MatTek) in phenol red-free Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum, 10 mM HEPES pH 7.4, 100 units/mL penicillin, and 100 µg/mL streptomycin in 5% CO₂. Sixteen hours prior to imaging, cells were shifted to 40 ºC and 5% CO₂; eight hours prior to imaging cells were treated with either vehicle (DMSO) or 25 µM Thiamet-G. Cells were then shifted to the permissive temperature (32 ºC and 5% CO₂) and images were collected every 10 seconds for 20 minutes. Imaging was performed on a Zeiss LSM 780 inverted confocal microscope, using 40X/1.4 NA Oil-Plan Apochromat DIC, (UV) VIS-IR (420762-9900), Zen 2011 software (Duke Light Microscopy Core Facility).

**Transfections** – Cells plated at ~50% confluence (or ~20% confluence, for GlcNDAz labeling experiments – see below) were transfected the following day. 750 µL pre-warmed OPTI-MEM (11058021, ThermoFisher) was placed 1.5 mL tubes with 30 µL TransIT-293 transfection reagent (Mirus), vortexed briefly, and incubated for 15 minutes.
at room temperature. After incubation, 10 µg of plasmid DNA were added to a 1.5 mL tube, vortexed briefly and incubated for 15 minutes at room temperature. After the final incubation, the transfection mixture was added dropwise to the cells. For GlcNDAz crosslinking experiments (see below), GlcNDAz was added 24 hours after transfection. For other experiments, cells were harvested 24 to 48 hours post-transfection.

**HRP secretion assay** – 293T cells were plated on 10 cm dishes, cultured until 80% confluent, and transfected with 20 µg ssHRP construct using TransIT-LT1 transfection reagent (Mirus) as above. Twenty-four hours post-transfection (and four hours before collecting supernatant), cells were pre-treated with 5SGlcNAc or vehicle only. After four hours, medium was aspirated and cells were gently washed with PBS. Fresh medium containing either 5SGlcNAc or vehicle was re-added. At each time point, 50 µL of medium was removed from the plate and stored at 4 ºC until analysis. To quantify ssHRP secretion, 10 µL of medium from each time-point was added to a clear-bottom white 96-well plate with 50 µL of pre-mixed ECL solution (WesternBright ECL, Advansta), and luminescence was measured on a Spectramax M5e plate reader (Molecular Devices) with an integration time of 500 milliseconds.

**Cell viability assay** – 293T cells were transfected with the ssHRP plasmid as above. After 24 hours, cells were trypsinized and 25,000 cells in 100 µL DMEM were seeded into clear-bottom 96-well plates and treated as described above. After an additional 24 hours, the MTS assay was performed according to manufacturer’s protocol.
(Promega, CellTiter 96 AQueous Proliferation Assay kit, G3581). Reported values are an average of technical triplicate reads minus the absorbance from control wells containing medium only.

**GlcNDAz crosslinking** – Cells stably expressing UAP1(F383G) were plated at ~50% confluence without puromycin. If transfecting first, cells were plated at ~20% confluence. The next day, 100 µM GlcNDAz (175) or vehicle was added to the cells. Medium was changed and fresh GlcNDAz was added every 24 hours. After 48-72 hours of GlcNDAz treatment, plates were placed on ice, lids were removed, and medium was replaced with ice-cold PBS. Cells were irradiated with a 365 nm UV lamp (Blak-Ray XX-20BLB UV Bench Lamp, 95-0045-04) for 20 minutes to induce crosslinking. Cells were collected by scraping into PBS and centrifugation, lysed in IP lysis buffer, and analyzed by BCA assay, SDS-PAGE and WB.

**Mass spectrometry (MS) O-GlcNAc site mapping** – Eight hours prior to harvest, 293T-17 cells transfected with myc-6xHis-tagged COPII components were treated with 25 µM Thiamet-G and 4 mM glucosamine. Tandem affinity-purified SEC23A, SEC24C or SEC31A was separated by SDS-PAGE, visualized by colloidal Coomassie blue, excised, and subjected to a standardized in-gel trypsin digestion (http://www.genome.duke.edu/cores/proteomics/sample-preparation/documents/In-gelDigestionProtocolrevised.pdf). Extracted peptides were lyophilized to dryness and resuspended in 20 µL of 0.2% formic acid/2% acetonitrile. Each sample was subjected to
chromatographic separation on a Waters NanoAquity UPLC equipped with a 1.7 µm BEH130 C₁₈ 75 µm I.D. X 250 mm reversed-phase column. The mobile phase consisted of (A) 0.1% formic acid in water and (B) 0.1% formic acid in acetonitrile. Following a 2 µL injection, peptides were trapped for 3 minutes on a 5 µm Symmetry C₁₈ 180 µm I.D. X 20 mm column at 5 µL/minute in 99.9% A. The analytical column was then switched in-line and a linear elution gradient of 5% B to 40% B was performed over 90 minutes at 400 nl/minute. The analytical column was connected to a fused silica PicoTip emitter (New Objective, Cambridge, MA) with a 10 µm tip orifice and coupled to a QExactive Plus mass spectrometer (ThermoFisher) through an electrospray interface operating in a data-dependent mode of acquisition. The instrument was set to acquire a precursor MS scan from m/z 375-1675 with MS/MS spectra acquired for the ten most abundant precursor ions. For all experiments, high-energy collisional dissociation (HCD) energy settings were 27 V, and a 120-second dynamic exclusion was employed for previously fragmented precursor ions.

Raw LC-MS/MS data files were processed in Proteome Discoverer (ThermoFisher) and then submitted to independent Mascot searches (Matrix Science) against a SwissProt database (human taxonomy) containing both forward and reverse entries of each protein (20,322 forward entries). Search tolerances were five parts per million for precursor ions and 0.02 Da for product ions using semi-trypsin specificity with up to two missed cleavages. Carbamidomethylation (+57.0214 Da on Cys) was set
as a fixed modification, whereas oxidation (+15.9949 Da on Met), deamidation (+0.98 Da on Asn/Gln), O-GlcNAc (+203.0793 Da on Ser/Thr), and phosphorylation (+79.9663 Da on Ser/Thr/Tyr) were considered dynamic mass modifications. All searched spectra were imported into Scaffold (v4.7, Proteome Software) and scoring thresholds were set to achieve a peptide false discovery rate below 1% using the PeptideProphet algorithm. Localized O-GlcNAcylated residues were manually verified based on annotated mass spectra generated within Mascot. In the SEC23A sample, peptides assigned by Mascot to SEC23B were assumed to be the cognate peptides in SEC23A, since the SEC23A was purified to homogeneity.

**Generation of SEC23A<sup>Δ</sup> SW1353 cells** – LentiCas9 virus was obtained from the Duke Functional Genomics Facility. SW1353 cells were plated in 6-well plates to be ~50% confluent at time of infection. Prior to infection, medium was replaced with fresh medium containing 4 µg/mL polybrene and 50 µL of LentiCas9 virus was added drop-wise to the cells. Plates were then centrifuged at 700 g for 30 minutes and incubated at 37 °C and 5% CO<sub>2</sub> overnight. The following morning, medium was replaced and cells were allowed to recover for 48 hours. After recovery, cells were selected using 3 µg/mL blasticidin and passaged until a control, uninfected plate, contained no living cells. These cells were dubbed SW1353/Cas9.
A single-guide RNA (sgRNA) sequence spanning the start codon of the human SEC23A locus was designed and validated via the Surveyor assay (179) by the Duke Functional Genomics Facility:

5' - TTCCAAATAGGGTCATTG - 3'

An sgRNA targeting AAVS1, commonly used as a control and deemed the “safe harbor” virus (SHV) locus (180), was used a control. Lentiviruses encoding the SEC23A sgRNA and the AAVS1 sgRNA were obtained from the Duke Functional Genomics Facility.

SW1353/Cas9 cells were plated in 6-well plates to be ~50% confluent at time of infection. Prior to infection, medium was replaced with fresh media containing 4 µg/mL polybrene and 50 µL of SEC23A or AAVS1 sgRNA lentivirus was added drop-wise to the cells. Plates were then centrifuged at 700 g for 30 minutes and incubated at 37 ºC and 5% CO₂ overnight. The following morning, medium was replaced and cells were allowed to recover for 48 hours. After recovery, infected cells were selected using 0.25 µg/mL puromycin and passaged until a control, uninfected plate, contained no living cells. Reduction of SEC23A expression was verified on a mixed population level via WB.

Next, using serial dilution, single cells were plated in 96-well plates for both SEC23A and AAVS1 sgRNA-containing populations, and clones were allowed to recover and expand. Cells were transferred to larger plates under selective pressure of 3 µg/mL blasticidin and 0.25 µg/mL puromycin. WB blot was used to verify loss of SEC23A
expression. One SEC23A\(^*\) clone (SW23A-3) and one AAVS1 safe harbor control clone (SHV-23) were used in subsequent reconstitution experiments.

**Generation of stable SEC23A-reconstituted SW23A-3** – SW23A-3 were plated to be ~30% confluent in 6-well plates the following day. Then, medium was replaced prior to transfections. 150 µL pre-warmed OPTI-MEM was placed 1.5 mL tubes with 9 µL TransIT-LT1 transfection reagent (Mirus), vortexed briefly, and incubated for 15 minutes at room temperature. After incubation, 3 µg of plasmid DNA for either pcDNA4-myc-6xHis (empty vector), pcDNA4-GFP-myc-6xHis, pcDNA4-SEC23A-myc-6xHis (wild type and serine-to-alanine point mutants) were added to a 1.5 mL tube, vortexed briefly and incubated for 15 minutes at room temperature. After the final incubation, the transfection mixture was added dropwise to the cells. Cells recovered at 37 °C and 5% CO\(_2\) for 48 hours and then were selected using 200 µg/mL Zeocin. Cells were passaged under selective pressure until a mock-transfected plate no longer contained living cells, the GFP-transfected control plate showed a majority positive cells under fluorescent illumination, and other resistant colonies were expanding. Cell populations were also examined for stable transfectants using myc immunofluorescence.

**SEC23B knockdown via siRNA in SW1353** – Glass coverslips were dipped in 100% ethanol, flame-sterilized, and placed into 6-well plates with media. Cells were plated to be ~15% the following day. The next morning, two reaction tubes per condition were prepared, both containing 125 µL pre-warmed OPTI-MEM. 100 pmol of a 1:1
mixture of SEC23B siRNA 1 and 2 was added to the first reaction tube (reaction 1) and incubated for 2 minutes. The same was done for the control SEC23B scrambled siRNA 1 and 2. Next, 3 µL per 100 pmol of Lipofectamine RNAiMAX Transfection Reagent (Thermo Fisher) was added to the other tube containing 125 µL pre-warmed OPTI-MEM (reaction 2). Then, reaction 1 was added to reaction 2, flicked to mix, and incubated at room temperature for 20 minutes. During the incubation, fresh medium was added to the cells. After 20 minutes, the combined mixture was added drop-wise to the cells.

After 72 hours of knockdown, cells were used in IF experiments.

siRNAs sequences targeting SEC23B have been described (181) and were ordered from Sigma-Aldrich as duplexes. Scrambled siRNA were generated from parent sequences using Invivogen’s siRNA Wizard scramble tool (http://www.invivogen.com/sirnawizard/scrambled.php). Scrambled sequences were checked using NCBI BLAST to verify no significant homology in the human genome.

SEC23B siRNA 1
5’-CACGUUACAUACACACGGA[dT][dT]

SEC23B siRNA 2
5’-CACUAUGAGAUGCUUGCUA[dT][dT]

SEC23B scrambled siRNA 1
5’-GAACCCGCACGUAACUAU[dT][dT]

SEC23B scrambled siRNA 2
Collagen immunofluorescence trafficking assay – After SEC23B knockdown via siRNA, cells were treated with 50 µg/mL ascorbate for 7 hours. Then, cells were washed 3 times with PBS and fixed with 4% paraformaldehyde in PBS for 20 minutes. Following fixation, cells were washed 3 times with PBS and permeabilized using 0.1% Triton X-100 in PBS for 10 minutes. After permeabilization, cells were washed 3 times using PBS and then blocked using 0.1% BSA in PBS for 30 minutes. Following blocking, glass coverslips were transferred to microcentrifuge caps inside a humidified 6-well plate. Primary antibody dilutions in 0.1% BSA were added to the coverslips and incubated overnight at 4 ºC. The next day, coverslips were transferred back to their original 6-well plate for washing and were washed 3 times with PBS. Coverslips were again transferred to microcentrifuge caps inside a humidified 6-well plate and incubated at room temperature for 1 hour with appropriate Alexa Fluor secondary antibodies in 0.1% BSA. Coverslips were washed 3 times in PBS and transferred to glass slides with ProLong™ Diamond Antifade Mountant with DAPI (Thermo Fisher). Slides were cured at room temperature in a dark cabinet for at least 24 hours before imaging. The following primary antibodies were used: anti-myc (9E10, BioLegend, 626801, 1:500) and anti-collagen I antiserum (LF-68, a gift of Dr. Larry Fisher, National Institute of Dental and Craniofacial Research, 1:2000). The following secondary antibodies were used: goat anti-rabbit Alexa Fluor 488 (Thermo-Fisher, A-11008, 1:2000)
and goat anti-mouse Alexa Fluor 594 (Thermo-Fisher, A-11005, 1:2000). At least 200 cells from each cell line were counted per biological replicate. Imaging was performed on a Zeiss LSM 780 inverted confocal microscope, using 40X/1.4 NA Oil-Plan Apochromat DIC, (UV) VIS-IR (420762-9900), Zen 2011 software (Duke Light Microscopy Core Facility).

**General immunofluorescence staining** – Cells were plated on glass cover-slips and treated as indicated. After treatment, cells were fixed, permeabilized, stained, and mounted as above. The following primary antibodies were used: Protein disulfide isomerase (PDI) (C81H6, Cell Signaling Technology, 3501, 1:100), SEC16 (A300-648A, Bethyl, 1:500), human procollagen type I C-peptide (QED Bioscience, 42024, 1:500). The following secondary antibodies were used: goat anti-rabbit Alexa Fluor 488 (Thermo-Fisher, A-11008, 1:2000), goat anti-mouse Alex Fluor 488 (Thermo-Fisher, A-11001; 1:2000), goat anti-mouse Alexa Fluor 594 (Thermo-Fisher, A-11005, 1:2000), goat anti-rabbit 594 (Thermo-Fisher, A-11012).

**Construction of Gateway vectors for Tol2kit system** – The zebrafish sec23a coding sequence was amplified from pCS2+ zSEC23a vector with primers containing attB1(forward) and attB2 (reverse) recombinase engineering sites. The fragment was purified and cloned into pDONR221 vector using Gateway BP clonase II enzyme mix. After confirming successful cloning by sequencing, this plasmid was used as the middle entry vector for the Multistep Gateway Recombineering system, in combination with
p5E_1.7kbCol2a1a promoter (182), p3E_v2a-EGFP (a gift of Dr. Josh Gamse, Vanderbilt University), pDestTol2pA2 and incubated with Gateway LR clonase II enzyme mix overnight at room temperature. After transformation and colony screening, a destination vector was obtained that allows expression of the zebrafish sec23a coding sequence fused with a self-cleaving viral 2a peptide-tagged EGFP (v2a-EGFP) under the tissue-specific Col2a1a promoter. Four candidate SEC23A O-GlcNAc sites were sequentially mutated using Q5 site-directed mutagenesis kit (New England Biolabs) with the following primers:

115-116F: TCCACAGTTTgcagctATTGAATATGTTGTC
115-116R: AGTAGCTCAGCAGGCTGG
137F: TGTGGTGGACgCGTGCATGGA
137R: TACAGGAAGTTCAGGGGCATTTG
168F: GGGCCTCATTgCATTTGGACGG
168R: ACTAGGCAGTAGGGGGC.

**Zebrafish microinjection** – Zebrafish embryos at the one-cell stage were collected from crusher<sup>m29</sup>/AB heterozygous crosses, injected with a combination of 50 pg medaka transposase mRNA and 10 pg of sec23a destination vector DNA, and grown for 3 days in embryo medium.

**Zebrafish cryosectioning and immunohistochemistry** – 80 hours post-fertilization (hpf), embryos were fixed with 4% paraformaldehyde at 4 °C overnight and
transferred to 30% sucrose in PBS. Embryos were embedded in Cryomatrix (ThermoFisher) and frozen at -80 ºC for 15 minutes. Then, 14 µm-thick sections were cut with a cryostat (Leica CM1900) and collected onto Fisherbrand Superfrost Plus slides. Slides were dried and rehydrated in PBS before staining. Following antigen retrieval (20 µg/mL proteinase K at room temperature for 5 minutes), embryos were permeabilized with 0.5% Triton X-100 in PBS for 10 minutes at room temperature. Samples were blocked (2% goat serum and 2 mg/mL BSA) for 30 minutes at room temperature, followed by overnight primary antibody incubation at 4 ºC, using 1:250 collagen-II antibody (Rockland) and 1:250 GFP antibody (Vanderbilt University Antibody and Protein Resource). Samples were incubated with secondary antibodies (rabbit AlexaFluor-555 and chicken AlexaFluor-488, LifeTechnologies) for one hour, and 1:4000 DAPI for 15 minutes at room temperature. Slides were mounted in ProlongGold (ThermoFisher).

**Zebrfish imaging and quantification** – Slides were imaged with Zeiss AxioImager.Z1. Percent collagen area in cytosol was calculated by the following formula: \( \text{Percent Collagen} = \left( \frac{\text{Collagen-positive intracellular area}}{\text{cytosol area - nucleus area marked by DAPI}} \right) \times 100 \). ImageJ (183) was used for intracellular area measurements.
2.3 Results

2.3.1 Novel O-GlcNAc Sites Identified on SEC23A, SEC24C, and SEC31A

Previous studies have indicated that COPII proteins are O-GlcNAcylated (135-138,168) but the biochemical and functional implications of these observations remain unclear. Indeed, O-GlcNAc sites have not been mapped systematically on any COPII protein, hindering subsequent biochemical and phenotypic experiments. As a first step towards understanding the role of O-GlcNAc cycling in anterograde trafficking, we analyzed the O-GlcNAcylation of three representative human COPII proteins, SEC23A, SEC24C and SEC31A. We expressed and purified each protein to homogeneity from human cells and used mass spectrometry (MS) to map O-GlcNAcylated residues. Using high-stringency criteria, we detected at least twenty-six, eleven and ten O-GlcNAc modifications on SEC23A, SEC24C and SEC31A, respectively, and manual inspection of the MS data permitted unambiguous assignment of many O-GlcNAc moieties to specific residues (Figure 3 and Appendix A). In most cases, the odds of modification were calculated by dividing the number of observed O-GlcNAc moieties by the number of modifiable residues on each peptide. In the case of SEC23A T168, O-GlcNAc was unambiguously assigned to the site upon manual inspection of MS/MS fragmentation data. Interestingly, the glycosylation sites were not evenly distributed across each protein, but rather were clustered in specific regions in all three cases. For example, five of the unambiguously assigned SEC23A O-GlcNAc sites (T137, T168, S184, S226, and
S241) lie near its site of interaction with SEC24, as determined by prior structural studies (184), whereas the others are distal to this interface (Figure 3B). In the case of SEC24C, the three definitively assigned glycosylation sites (S773, T775 and T776) reside in the relatively well-conserved β-sandwich domain. Finally, homology modeling based on structures of the yeast SEC31 ortholog suggests that all four assigned SEC31A glycosylation sites lie in the α-solenoid domain, which is thought to mediate protein-protein interactions and form a flexible hinge, permitting expansion of COPII carriers to accommodate a range of cargo sizes (185-188). Therefore, SEC31A glycosylation may modulate protein-protein interactions and/or transport carrier size in the mammalian COPII outer coat. Taken together, these results indicate that site-specific glycosylation is prevalent on COPII proteins and may impact their biochemical functions.
Figure 3: Novel O-GlcNAc Sites Identified on SEC23A, SEC24C, and SEC31A
(A) Schematic depicting twelve, three, and four O-GlcNAc sites identified on SEC23A, SEC24C, and SEC31A, respectively. Selected domains of each protein are also illustrated. Each modification shown here was unambiguously assigned to a single residue via manual inspection of MS data. (B) The unambiguously assigned O-GlcNAc sites identified in this report are depicted in red on a model of human SEC23A (silver, PDB: 5VNO) and SEC24C (blue, PDB: 3E3H). The SEC31 interaction interface of SEC23A74 is indicated in green. (C) Representative MS/MS fragmentation spectrum of a SEC23 glycopeptide. Accurate mass measurements of the intact QMSLSLLPPDALVGLITFGR peptide indicated the presence of two O-GlcNAc modifications. HCD fragmentation revealed a y-ion series through y10 (ALVGLITFGR fragment) as well as a fragmentation series y4 (TFGR fragment) plus 203.0794 Da, which is highlighted by y4+203 and expanded in Figure 1D. (D) Fragmentation series with an intact O-GlcNAc modification (magnified from the spectrum in Figure 1C). This fragment ion series localizes one O-GlcNAc modification (denoted +O) unambiguously to Thr168, with either Ser154 or Ser156 as the additional site. All O-GlcNAc sites depicted in Figures 1A and 1B were unambiguously assigned using similar analysis, or because the number of O-GlcNAc moieties detected equaled the number of Ser/Thr residues on a given tryptic peptide.

2.3.2 Normal O-GlcNAc Cycling is Required for COPII-dependent Secretion

Next, we asked whether dynamic O-GlcNAcylaion regulates the COPII pathway. We used immunoprecipitation (IP) and Western blotting (WB), in combination with specific small molecule inhibitors of OGT (5SGlcNAc) and OGA (Thiamet-G) (174,176), to demonstrate that O-GlcNAc cycles on COPII components, including SEC24C and Trk-fused gene (TFG), a protein required for COPII function in vivo in metazoans (Figure 4A) (146,189-191). These results suggest that dynamic glycosylation may regulate protein secretion. To test this hypothesis, we performed three functional assays to assess the requirement for O-GlcNAc cycling in COPII-dependent trafficking. First, we used a secreted soluble horseradish peroxidase (ssHRP) reporter that traffics
through the secretory pathway in a COPII-dependent manner (192). ssHRP can be
detected in culture supernatants via standard chemiluminescence, providing a
quantitative measure of protein secretion (192). We found that inhibition of OGT
significantly diminished the COPII-dependent secretion of ssHRP without reducing
intracellular ssHRP expression or cell viability (Figures 4B and data not shown). Second,
we expressed an enhanced green fluorescent protein (eGFP)-tagged temperature-
sensitive mutant of the vesicular stomatitis virus glycoprotein (tsVSVG-eGFP) in
mammalian cells (193,194). At the non-permissive temperature, tsVSVG-eGFP is
retained in the ER, as indicated by reticular green fluorescence and colocalization with
genetically encoded organelle markers (Figures 4C and data not shown) (193,194). Upon
shifting to the permissive temperature, tsVSVG-eGFP transits to the Golgi in a strictly
COPII-dependent manner, displaying characteristic juxtanuclear punctate fluorescence,
thereby providing a live-cell assay for pathway function (Figures 4C) (193,194). We
found that inhibition of OGA markedly delayed COPII-dependent tsVSVG-eGFP
trafficking (Figures 4C). Third, we asked whether O-GlcNAcylation affects the
subcellular distribution of COPII proteins. All components of the inner and outer COPII
coats exist in both solublecytoplasmic and membrane-associated pools, and reversible
membrane association is required for normal COPII function (67,140,141). To determine
whether O-GlcNAc regulates the membrane association of COPII components, we used
digitonin treatment to permeabilize the plasma membranes of human cells, permitting
the separation of soluble cytoplasmic proteins and membrane-associated proteins (195),
and then performed WBs for COPII components on these subcellular fractions.
Interestingly, we found that inhibition of OGA reduced the ratio of membrane-
associated to cytoplasmic SEC23A and SEC31A, indicating that O-GlcNAc cycling
regulates the subcellular distribution of essential COPII components (Figure 4D).
Together, these results indicate that dynamic O-GlcNAcylation is required for normal
COPII-dependent secretion.
Figure 4: Normal O-GlcNAc Cycling is Required for COPII-dependent Secretion

(A) O-GlcNAc dynamically modifies COPII components. 293T-17 cells were transfected with either myc-6xHis-SEC24C or TFG-myc-6xHis and treated with vehicle (DMSO), 25 µM Thiamet-G, or 25 µM 5S-GlcNAc eight hours prior to harvest. Tandem affinity-purified SEC24C and TFG were analyzed by WB. Both proteins demonstrated O-GlcNAc cycling as indicated by an increase in signal when treated with Thiamet-G, and a decrease in signal with 5S-GlcNAc, compared to vehicle. (B)
293T cells were treated with vehicle (DMSO) or 50 µM 5SGlcNAc and transfected with ssHRP. The amount of ssHRP in the medium supernatant was measured after 20 hours via luminescence assay. Inhibition of OGT via 5SGlcNAc significantly diminished secretion. n = 6, ** p = 0.006, Student’s t-test. Error bars are standard error of the mean. (C) COS7 cells stably expressing tsVSVG-eGFP were treated with vehicle (DMSO) or 25 µM Thiamet-G for 8 hours at the non-permissive temperature (40 ºC) and then shifted to the permissive temperature (32 ºC). Cells were imaged every 10 seconds for 20 minutes, monitoring the movement of the tsVSVG-eGFP from ER (reticular ER fluorescence) to the Golgi (juxtanuclear punctate fluorescence). (D) SW1353 cells were incubated with DMSO or 50 µM Thiamet-G for 24 hours. Cytoplasm (C) and endomembrane (M) fractions were prepared by digitonin fractionation, and SEC23A and SEC31A were analyzed by WB. Tubulin and TRAPα serve as loading controls for cytoplasmic and integral ER transmembrane proteins, respectively. One representative experiment is shown.

2.3.3 Dynamic O-GlcNAcylation Mediates Interactions Among COPII Pathway Proteins

Like other PTMs, O-GlcNAc can exert diverse biochemical effects on its substrates, including altering protein conformation or subcellular localization, or mediating or disrupting protein-protein interactions (38,169,172,173,196,197). Elegant biochemical and structural studies have revealed that numerous precise protein-protein interactions are required for COPII coat formation, cargo loading, membrane curvature, and scission (67,140,141). In addition, our MS site-mapping data demonstrate that O-GlcNAcylation occurs on known or predicted protein-protein interaction domains of COPII components (Figure 3). Therefore, we hypothesized that O-GlcNAc may mediate the protein-protein interactions of COPII components. However, physiologically important O-GlcNAc-mediated protein-protein interactions can be sub-stoichiometric, transient and low-affinity, presenting a significant challenge for their characterization.
To address this challenge, we harnessed a chemical biology method to covalently capture O-GlcNAc-mediated protein-protein interactions in live human cells (175). In this strategy, cells are metabolically labeled with a GlcNAc analog bearing a diazirine photocrosslinking moiety, termed “GlcNDAz” (175). GlcNDAz is accepted by the GlcNAc salvage pathway, converted to UDP-GlcNDAz, and used by OGT to decorate its native substrates (175). Briefly UV-treating GlcNDAz-labeled cells triggers the covalent crosslinking of O-GlcNDAz moieties to any binding partner proteins within ~2-4 Å of the sugar (175). Thanks to this short radius, GlcNDAz crosslinking occurs exclusively at sites where the glycan contributes to the interaction interface, without crosslinking to distant or nonspecific proteins (175). These covalent crosslinks can be conveniently visualized through the GlcNDAz-dependent appearance of high molecular weight species on a WB for the protein of interest. Therefore, GlcNDAz is a powerful chemical tool for identifying direct, glycosylation-mediated interactions between endogenous proteins in live cells (175).

Using GlcNDAz, we found that endogenous SEC23A, SEC24B, SEC24C, SEC31A and TFG all engage in O-GlcNAc-mediated protein-protein interactions in multiple human cell lines, detected as massive crosslinks on a WB (Figure 5A). Importantly, GlcNDAz crosslinking of SEC23A was enhanced when deglycosylation was blocked by brief Thiamet-G treatment (Figure 5B), indicating that these interactions are dynamic and
O-GlcNAc-specific. We concluded that O-GlcNAc mediates regulated interactions of COPII proteins.

To test the hypothesis that O-GlcNAc might mediate some of the well-characterized protein-protein interactions in the COPII coat (67,140,141), we examined TFG, SEC23A and SEC24C as model COPII glycoproteins. TFG homo-oligomerizes in vitro and in cells, and this property is required for its subcellular localization and function in COPII cargo trafficking (146,189-191). To determine whether O-GlcNAc participates in homotypic TFG-TFG interactions inside cells, we created constructs of human TFG with different epitope tags. IP/WBs with these constructs revealed that covalent, GlcNDAz-dependent crosslinked complexes of TFG contain both tagged versions of TFG simultaneously (Figure 5C). These results demonstrate that the GlcNDAz-dependent complexes are homo-oligomers, suggesting that O-GlcNAc may participate in TFG-TFG interactions in live cells.

Next, we examined the potential involvement of O-GlcNAc in the heterodimerization of SEC23 and SEC24, an essential step in the formation of the COPII pre-budding complex (67,140,141). We expressed FLAG-SEC24C in human cells and observed GlcNDAz-dependent crosslinking similar to that of endogenous SEC24C (Figure 3A, D). We IPed SEC24C and performed WBs for candidate interactors, including SEC23A (Figure 3D), SEC13, SEC31, SAR1 and the scaffolding protein SEC16 (not shown). In all cases, we failed to detect any candidate interactors in the crosslinked
SEC24C complexes (Figure 5D and not shown). Notably, endogenous SEC23A again crosslinked in this context, though not to SEC24C (Figure 3D). Similarly, IPed GlcNDAz crosslinks of SEC23A did not contain SEC24, SEC13, SEC31 or SAR1 (not shown). We concluded that SEC23A and SEC24C engage in specific but distinct O-GlcNAc-mediated protein-protein interactions, which may influence COPII trafficking. All together, these results demonstrate that dynamic O-GlcNAcylation mediates both known and novel interactions among COPII pathway proteins.
Figure 5: Dynamic O-GlcNAcylation Mediates Interactions Among COPII Pathway Proteins

(A) HeLa/UAP1 or 293T/UAP1 cells were treated with 100 µM GlcNDAz or vehicle (DMSO) and UV light (or not), and lysates were analyzed by WB. Circles indicate uncrosslinked proteins running at their predicted molecular weights. Arrows indicate GlcNDAz- and UV-dependent crosslinked species. (B) HeLa/UAP1 cells were treated with 100 µM GlcNDAz, 50 µM Thiamet-G, vehicle (DMSO), and UV light as indicated for 24 hours and lysates were analyzed by WB. Circle indicates uncrosslinked SEC23A running at its predicted molecular weight. Arrows indicate GlcNDAz- and UV-dependent crosslinked species. (C) Left: 293T/UAP1 cells were transfected with GFP (control), 3xFLAG-TFG only, TFG-myc-6xHis only, or co-
transfected with both FLAG and myc-6xHis constructs (“myc & FLAG”). Then, cells were treated with DMSO (vehicle) only or 100 µM GlcNDAz and were exposed to UV. Lysates were analyzed by WB, confirming that both myc- and FLAG-tagged TFG crosslinks similarly to endogenous TFG in a GlcNDAz-dependent manner. Covalently crosslinked bands of TFG (~150 kDa) are shown. Right: 293T/UAP1 cells were transfected with GFP (control), 3xFLAG-TFG only, TFG-myc-6xHis only, or co-transfected with both FLAG and myc-6xHis constructs (“myc & FLAG”). Then, cells were treated with 100 µM GlcNDAz and UV, and lysates were analyzed by IP and WB as indicated. Covalently crosslinked bands of TFG (~150 kDa) are shown. (D) 293T/UAP1 cells were transfected with 3xFLAG-SEC24C or GFP (control) and treated with 100 µM GlcNDAz or vehicle. Lysates were analyzed by anti-FLAG IP and WB. Crosslinked complexes of 3xFLAG-SEC24C were successfully purified (lane 8), but did not contain SEC23A. Notably, endogenous uncrosslinked SEC23A co-IPs with SEC24C (lanes 4 and 8) but endogenous crosslinked SEC23A (lane 5) does not, demonstrating that SEC24C and SEC23A each crosslink, but not to each other. Circles indicate uncrosslinked proteins running at their predicted molecular weights. Arrows indicate GlcNDAz- and UV-dependent crosslinked species.

2.3.4 Site-specific O-GlcNAcylation of SEC23A is Required for Proper Collagen Trafficking

We next focused on SEC23A, because SEC23 is an essential COPII component (67,71,140,141,198), and because human SEC23A mutations cause CLSD (88,90,91). To further elucidate the biochemical effects of SEC23A O-GlcNAcylation, we created unglycosylatable point mutants (Ser/Thr → Ala) in sixteen of the candidate sites we detected by MS, focusing on residues that are highly conserved across evolution (Figure 3A): S97, S102, S115, S116, T137, T168, S184, S226, T241, S508, S516, S627, S629, S639, S640, and S641. We expressed wild type or point-mutant SEC23A constructs in human cells and performed GlcNDAz crosslinking (Figure 7A). We found that the S115A, S116A, T137A, T168A, and S184A mutants reduced or abrogated GlcNDAz-mediated
crosslinking, whereas the others did not (Figure 7A). Interestingly, three of the residues required for SEC23A crosslinking, T137, T168 and S184, lie near its SEC24-binding interface (Figure 3A & B). However, as noted, SEC24 is not the GlcNDAz-mediated crosslinked partner of SEC23 (Figure 5). These observations suggest that the site-specific O-GlcNAcylation of SEC23A near the SEC24 binding interface mediates an interaction distinct from the SEC23/SEC24 pre-budding complex.

We next investigated the functional significance of site-specific SEC23A glycosylation in a human cell system. Chondrocytes create and maintain cartilage, largely through the regulated secretion of collagen (199), and SEC23A is required for this process in vivo (88,90,91). To assess the function of SEC23A with a physiologically relevant cell system and model cargo, we used CRISPR/Cas9 methods to delete the endogenous SEC23A gene in SW1353 human chondrosarcoma cells (Figure 6), which secrete endogenous collagen (200,201). Complete loss of SEC23A caused intense collagen accumulation and distended ER morphology (Figure 6), which resemble the reported previously COPII cargo defects of CLSD patient cells (88,90,91). These observations confirm that SEC23A+ SW1353 cells are an appropriate system for testing the function of unglycosylatable SEC23A mutants in trafficking endogenous collagen. Next, we stably re-expressed empty vector, wild type SEC23A, or unglycosylatable single point-mutant SEC23A alleles. To create an appropriately sensitive assay for SEC23A function, we knocked down SEC23B (data not shown) preventing compensation from this paralog, as
described previously (181). We then assessed SEC23A function in these reconstituted cell lines by tracking collagen secretion via immunofluorescence (IF) microscopy. Interestingly, SEC23A−/− SW1353 cells reconstituted with the S184A unglycosylatable point mutant exhibited a significantly higher proportion of retained collagen and characteristic dilated ER morphology, as compared to cells reconstituted with wild type SEC23A, despite partial rescue (Figures 7B & 8). Importantly, however, wild type and S184A mutant SEC23A co-IPed endogenous SEC24 proteins with similar efficiencies (Figure 9), indicating that the trafficking defect we observe in the S184A mutant (Figures 7B & 8) is not due to gross protein misfolding or lost SEC24 binding. Together, these results demonstrate that the S184 glycosylation site of SEC23A is required for its collagen trafficking function in human chondrosarcoma cells, possibly due to post-translational regulation at the SEC23A/SEC24 interface.
Figure 6: SEC23A\textsuperscript{−/−} SW1353 cells accumulate collagen when stimulated with ascorbate

(A) Single-cell clones derived from SW1353 cells stably expressing Cas9 and an sgRNA targeting a “safe harbor” locus (Control) or the SEC23A locus (SEC23A\textsuperscript{−/−}) were analyzed by WB. (B) SW1353 clones deleted for the “safe harbor” control locus (left, Control) or SEC23A (right, SEC23A\textsuperscript{−/−}) were treated with 50 µg/ml sodium ascorbate for 7 hours to stimulate collagen translation, as described (130,167). Then, cells were fixed and stained for endogenous collagen (LF68 anti-collagen, green) and DNA (DAPI, blue). Consistent with prior reports in other SEC23A loss-of-function systems (88,90,91), SEC23A\textsuperscript{−/−} cells exhibit dramatic intracellular collagen accumulation and distended ER phenotype.
Figure 7: Site-specific O-GlcNAcylation of SEC23A is Required for Proper Collagen Trafficking
(A) 293T/UAP1 cells were transfected with plasmids encoding GFP (control), wild type SEC23A-myc-6xHis, or the indicated SEC23A-myc-6xHis point mutants. Cells were treated with vehicle (DMSO) or 100 µM GlcNDAz and UV-irradiated. Lysates were analyzed by WB. The S115A, S116A, T137A, T168A and S184A mutants all show a marked reduction in crosslinking. Circles indicate uncrosslinked SEC23A running at its predicted molecular weight. Arrows indicate GlcNDAz- and UV-dependent crosslinked species. White rectangles indicate where irrelevant lanes have been cropped out of a single blot (bottom panels). Duplicate wild type SEC23A crosslinked samples were run on multiple blots due to space constraints. Therefore, apparent variability in wild type crosslinking efficiency is due to blot-to-blot variation in the enhanced chemiluminescence detection. (B) SEC23A−/− SW1353 cells stably reconstituted with empty vector (left), wild type SEC23A-myc-6xHis (middle) or S184A SEC23A-myc-6xHis (right) were transfected with anti-SEC23B siRNA (or scrambled SEC23B siRNA, not shown) and treated with 50 µg/ml sodium ascorbate for 7 hours to induce collagen translation (130,167). Then, cells were fixed and stained for SEC23A (anti-myc, red), endogenous collagen (LF68 anti-collagen, green) and DNA (DAPI, blue). Cells exhibiting both SEC23A expression and aberrant ER collagen staining (white arrows) were quantified as a fraction of the total number of SEC23A-positive cells. Representative images from eight biological replicates are shown at two magnifications (top and bottom panels).

Figure 8: Quantitation of collagen retention in SEC23A−/− SW1353 stably reconstituted with S184A SEC23A-myc-6xHis

Quantification of IF data presented in Figure 7B. Cells expressing S184A SEC23A were more than twice as likely to exhibit retained collagen, compared to those expressing wild type SEC23A. n = 8, *** p < 0.001, Student’s t-test. Error bars are standard error of the mean.
Figure 9: SEC23A S184A and 4A mutants co-IPs with SEC24 paralogs

Epitope-tagged human S184A (A) or 4A (B) mutant SEC23A was stably expressed in SEC23A−/− SW1353 cells, IPed and analyzed by WB. Both mutants co-IP endogenous SEC24 paralogs with efficiency similar to that of wild type SEC23A. Tubulin (arrow) is a loading control. Asterisk, IgG heavy chain from IPing antibody.

2.3.5 Site-specific O-GlcNAcylation of Sec23A Regulates in vivo Collagen Trafficking

In developing vertebrates, collagen trafficking and skeletogenesis are stringently dependent on SEC23A, because SEC23A partial loss-of-function mutations cause CLSD in humans and similar phenotypes in model organisms (88-92,152,202,203). Our biochemical and cellular data (Figure 3-5, 7-8) suggest that site-specific O-GlcNAcylation might regulate SEC23A function in developing tissues. To test this hypothesis, we took advantage of a vertebrate model of SEC23A dysfunction (92,152). We have shown
previously that mutations in zebrafish sec23a underlie the crusher phenotype, exhibiting collagen mis-trafficking, chondrocyte failure and skeletal dysmorphology (92,152) that closely resemble CLSD (88,90,91). Therefore, crusher fish provide a valuable vertebrate genetic model for studying SEC23A function in vivo (92,152). Importantly, the human and zebrafish SEC23A proteins are 91% identical, and all but one (S629) of the sixteen putative O-GlcNAc sites that we examined (Figure 7A) are conserved between these orthologs.

We tested the functional importance of SEC23A O-GlcNAcylation sites in vivo using the crusher model. First, we confirmed that wild type zebrafish SEC23A crosslinks in our GlcNDAz assay similarly to its human ortholog (Figure 10A). Then, we designed a compound-mutant SEC23A allele, in which four residues required for GlcNDAz crosslinking and conserved from humans to fish (S115, S116, T137, T168) were mutated to alanine ("4A mutant"). We created expression constructs of human and zebrafish SEC23A 4A mutant and confirmed that they are defective in GlcNDAz-mediated crosslinking (Figure 10A). These results indicate that the O-GlcNAc-mediated interactions of SEC23A are biochemically conserved across vertebrates.

Next, we developed an in vivo assay of SEC23A function in skeletogenesis (Figure 10B). We expressed wild type or 4A mutant SEC23A in Sec23a-deficient (crusher) zebrafish chondrocytes and tracked COPII-dependent collagen transport by immunohistochemistry (Figure 10B). We measured the cytosolic area occupied by
collagen as an established marker of its intracellular retention (92,115,152). We found that 5% of the cytosolic area of wild type chondrocytes was occupied by collagen at 80 hours post-fertilization (hpf), which corresponds to normal collagen traffic within secretory pathway compartments (Figure 10C, D). In contrast, 50-95% of the cytosolic area of sec23a-deficient chondrocytes was filled with aberrantly retained collagen (Figure 10C, D), consistent with our prior reports (92,152). Interestingly, while the expression of wild type SEC23A in crusher chondrocytes restored collagen secretion to wild type levels in a cell-autonomous fashion, the unglycosylatable SEC23A 4A mutant only partially restored collagen secretion, exhibiting significantly higher collagen retention than wild type (Figure 10C, D). Together, these results suggest that site-specific O-GlcNAcylation of SEC23A may regulate the in vivo COPII-dependent transport of SEC23A-dependent cargoes, including collagen.
Figure 10: Site-specific O-GlcNAcylation of Sec23A Regulates *in vivo* Collagen Trafficking

(A) 293T/UAP1 cells were transfected with either GFP (control), human wild type SEC23A-myc-6xHis, human 4A SEC23A-myc-6xHis, zebrafish wild type Sec23A-myc-6xHis, or zebrafish 4A Sec23A-myc-6xHis. Circle indicates uncrosslinked SEC23A running at its predicted molecular weight. Arrows indicate GlcNDAz- and UV-dependent crosslinked species. The 4A mutations reduce crosslinking in both human and zebrafish SEC23A. (B) Experimental strategy for mosaic expression of Sec23A in zebrafish chondrocytes. Sec23A wild type or 4A mutant was expressed under the
zebrafish Col2a1 promoter. Embryos injected with Sec23A constructs were grown until 80 hours post-fertilization (hpf) and cryosectioned for immunohistochemistry. (C) Cells expressing Sec23A constructs are marked by v2a-EGFP. Wild type cells secrete collagen to the extracellular space (arrows, top panel), whereas, crushe (sec23a loss-of-function) mutant chondrocytes accumulate intracellular collagen (arrowhead, middle panel). Overexpression of wild type Sec23A in crushe animals leads to clearance of intracellular collagen, whereas expression of the Sec23A 4A mutant fails to do so (arrowhead, lower panel). Double-headed arrows point to normal levels of collagen in the secretory compartment. (D) Quantification of zebrafish rescue experiments. Percent of cell area occupied by collagen shows that wild type Sec23A expression in crushe chondrocytes restores collagen secretion to normal levels, whereas the Sec23A 4A mutant does not, despite partial rescue. crushe (cru) chondrocytes accumulate the highest level of intracellular collagen. Investigator was blinded to wild type or 4A mutant genotype while performing quantification, per standard practice (92,115,152). p < 0.05, Student’s t-test. Error bars are standard error of the mean.

2.4 Discussion

The COPII trafficking pathway is essential for cell and tissue physiology in vertebrates and is dysregulated in several human diseases. While the core COPII machinery is relatively well understood, little is known about how trafficking is dynamically regulated in response to physiological or pathological cues. Our results indicate that O-GlcNAcylation of COPII proteins may be one important mode of pathway regulation in vertebrates.

We have shown that O-GlcNAc is a prevalent modification of COPII proteins, including SEC23, SEC24, SEC31 and TFG (Figure 3, 4A, 5A). Like phosphorylation, O-GlcNAcylation can exert a wide range of biochemical effects on its substrates, and it will be important to delineate the effects of COPII protein glycosylation at each individual modification site in order to build an integrated model for how O-GlcNAc affects
protein trafficking from the ER. As a first step towards this goal, our GlcNDAz results suggest that O-GlcNAc moieties lie at or very close to the interaction interface between TFG monomers, indicating a potential biochemical function for TFG glycosylation (Figure 5A, C). The N-terminal half of TFG, comprising its PB1 and coiled-coil domains, mediates homo-oligomerization, and this property is required for its trafficking function in vivo (146,189-191). While unglycosylated TFG can homo-oligomerize in vitro, our results suggest that O-GlcNAcylation may facilitate or regulate this process in live cells (146,191). Similarly, we show that both SEC23A and SEC24C also engage in O-GlcNAc-mediated protein-protein interactions, but not with each other, or with several other known binding partners in the COPII system (Figure 5A, B, D). These data suggest that novel O-GlcNAc-mediated protein-protein interactions may govern SEC23 and/or SEC24 function in the COPII system. It will be important to determine in future work how changes in site-specific glycosylation of TFG, SEC23 or SEC24 directly affect their biochemical activities in vivo.

At the cellular level, our results indicate that O-GlcNAc cycling in general (Figure 4), and specific SEC23A glycosylation sites in particular (Figures 7B & 8), are required for COPII-dependent protein secretion. Because native PTMs, including O-GlcNAc, are not essential for in vitro COPII vesicle assembly from minimal components (204), it is likely that O-GlcNAcylation and other PTMs are instead required to modulate the activity of the pathway in vivo. The precise molecular events affected by O-GlcNAc
and the full complement of substrates and modification sites most critical for this regulation remain incompletely understood. Interestingly, however, OGA inhibition both delayed the trafficking of tsVSVG-eGFP (Figure 4C) and decreased the pools of membrane-associated SEC23A and SEC31A (Figure 4D). These results suggest that unfettered O-GlcNAc cycling may be required for the efficient recruitment or recycling of COPII proteins to the ER or Golgi membranes. Other PTMs govern COPII in this way. For example, the Golgi-localized kinase CK1δ and the phosphatase PP6 act reciprocally on SEC23 and SEC24 to ensure orderly COPII carrier budding, fusion and directionality (125,126), and a recent study demonstrated that phosphorylation by CK2 inhibits the membrane association of SEC31 (134).

Notably, O-GlcNAc and O-phosphate exhibit a complex interplay in cells, frequently competing for nearby or identical residues on a given protein and exerting antagonistic effects (169). Indeed, a prior report suggests that the cell cycle-dependent phosphorylation and glycosylation of SEC24 are reciprocal and may affect its membrane association (135). Our data are consistent with this proposed model, but additional cell biological studies with unglycosylatable point mutants of COPII proteins will be required to elucidate the responsible modification sites and potential O-GlcNAc/O-phosphate cross-talk. In the case of SEC23A, none of the glycosylation sites that we examined (Figure 7A) is reported to be phosphorylated, though additional sites of both modifications may yet await discovery. In contrast, phosphorylation of S773 and T776 of
SEC24C has been reported (205), suggesting that the interplay between O-GlcNAcylation and phosphorylation at these sites could tune COPII activity. Future studies will address the potential cross-talk among PTMs of COPII proteins.

The COPII system of vertebrates is significantly more complex than that of lower eukaryotes. For example, vertebrates possess two paralogs of SEC23 (A and B) and four paralogs of SEC24 (A through D), whereas budding yeast has one essential ortholog of each. The functional significance of vertebrates’ multiple COPII paralogs remains largely unclear. All four SEC24 paralogs exhibit analogous biochemical functions in early COPII coat formation and cargo capture, and yet mouse knockout studies and differences in cargo binding clearly indicate distinct, if overlapping, physiological functions among these proteins (109,111,112,114,115). Studies of SEC23A and B revealed analogous results, despite their high level of conservation and similar biochemical functions (93,96,97,100). Currently, it is not well understood how these COPII components achieve distinct biological roles in vivo, though tissue-specific expression patterns of paralogs and cargoes likely contribute (202,203). Another possibility is that PTMs differentially regulate COPII paralogs. Consistent with this notion, we observed distinct GlcNDAz crosslinks of endogenous SEC24B and C (Figure 5A), suggesting that O-GlcNAcylation may mediate different protein-protein interactions among paralogs. In addition, one glycosylation site that we mapped on human SEC24C (T775) is conserved as a serine or threonine in all four human SEC24 paralogs, whereas the other two localized SEC24C O-
GlcNAc sites (Figures 3A & 3B) and several candidate sites are conserved in only one or none of the other three SEC24 paralogs. Therefore, glycosylation of multiple SEC24 paralogs at the T775-cognate site may impact all COPII traffic, whereas glycosylation of SEC24C at other sites may provide paralog-specific regulation to tune the trafficking of specific cargoes or cell types. In contrast, of the twelve unambiguously localized O-GlcNAc sites we identified on human SEC23A (Figures 3A-B), all except S516 are conserved in SEC23B (with a serine in SEC23B corresponding to SEC23A T508). These results suggest that glycosylation could affect SEC23 paralogs more uniformly than SEC24 paralogs, though this hypothesis remains to be tested.

Finally, our results demonstrate that several evolutionarily conserved glycosylation sites on SEC23A are required for O-GlcNAc-mediated protein-protein interactions in cells, and to support collagen secretion in SW1353 cells and skeletogenesis in developing zebrafish (Figures 7-8 & 10). One of these residues, S184, lies near the SEC23/SEC24 interface, potentially providing a site for a regulatory O-GlcNAc-mediated protein-protein interaction with a protein other than SEC24 (Figures 5D, 7A). In contrast, the residues altered in the SEC23A 4A mutant do not lie exclusively at the SEC23/SEC24 interface, and O-GlcNAcylation may play a biochemically distinct role at those sites. Although our cell-and animal-based experiments (Figures 4, 7, 8, 10) cannot stringently rule out the possibility that the S184A or 4A mutation induces a deleterious conformational change in SEC23A, both mutants co-IP with SEC24 proteins
with efficiencies similar to that of wild type SEC23A, arguing against dramatic conformational change or misfolding (Figure 9). Similarly, neither SEC23A deletion nor OGT or OGA inhibition disrupted ER exit sites (ERES), as judged by IF staining of the endogenous, canonical ERES marker SEC16 (data not shown). This result suggests that the functional effects that we observe upon OGT or OGA inhibition (Figure 4) or ablation of SEC23A glycosylation sites (Figures 7 & 8) may be due to regulation of SEC23A and/or other COPII proteins downstream of ERES formation. Therefore, we propose that SEC23A O-GlcNAcylation at these sites is required to regulate its interaction with binding partners beyond SEC24. We expect that future GlcNDAz crosslinking and MS proteomic experiments to identify the O-GlcNAc-dependent binding partner(s) of SEC23A will facilitate the elucidation of the molecular mechanisms through which site-specific O-GlcNAcylation of SEC23A influences COPII trafficking.

At the organismal level, our results suggest that dynamic O-GlcNAcylation of SEC23A may govern collagen trafficking during vertebrate development (Figures 7-8 & 10). Consistent with this idea, recent work identified a connection between whole-body O-GlcNAc levels and chondrogenic differentiation in mice (206), implicating O-GlcNAcylation in skeletogenesis. We have also shown that O-GlcNAc modifies SEC24C and SEC31A (Figure 3), both of which may participate in efficient in vivo collagen secretion as well (130,167,207). Therefore, O-GlcNAcylation may be a conserved mode of
COPII regulation in chondrocytes during vertebrate development. Testing this hypothesis will be an important goal of future studies.

2.5 Chapter Supplemental Material

This section contains brief results and discussion related to this chapter that were not included in the corresponding publication.

As previously described, O-GlcNAc cycles on TFG, suggesting dynamic glycosylation may regulate protein secretion (87). We also determined that TFG GlcNDAz-dependent crosslinks are homo-oligomers, suggesting that O-GlcNAc may participate in TFG-TFG interactions in live cells (87). A collaborator shared O-GlcNAc MS site-mapping data of TFG with us (personal communication, Anjon Audhya), indicating five novel O-GlcNAc-modified residues on TFG: T296, S297, S305, T330, and S376. To determine whether any of these played a role in O-GlcNAc-mediated protein-protein interactions, we created unglycosylatable point mutants (Ser/Thr → Ala). We expressed wild type and point-mutant TFG constructs in human cells and performed GlcNDAz crosslinking (Figure 11). We found that none of the single point mutations abrogated GlcNDAz-dependent crosslinking. We also tested GlcNDAz crosslinking in a S330A/T376A double mutant, which also did not reduce GlcNDAz-dependent crosslinking of TFG (Figure 11). Next, to determine the region of TFG important for GlcNDAz-dependent crosslinking, we examined crosslinking with a truncated form of TFG, which consists of amino acids 194 to 400 fused with GFP (208). This truncated TFG
contains the PQ-rich region of TFG which is required for its localization to the ER, can homodimerize, and facilitates the polymerization of TFG octamers (208). We determined that the C-terminal region of TFG was involved in GlcNDaz crosslinking (data not shown). Interestingly, all the O-GlcNAc residues we previously mutated were in this region. Although we saw no decrease in GlcNDaz-dependent crosslinking of TFG with any of these mutants, assessing the crosslinking of more TFG glycosylation-deficient compound mutants may give us insight into the sites required for O-GlcNAc-mediated interactions. Also, currently unidentified O-GlcNAc-modified residues within the C-terminal region could promote this interaction. Finally, because TFG homooligomerizes, it is possible that our TFG crosslinks contain both wild-type endogenous TFG and our epitope-tagged O-GlcNAc point mutants. A TFGΔ background may be required to fully assess which residues are important for the GlcNDaz-dependent crosslinking of TFG. Further studies with other TFG truncation mutants or compound Ser/Thr → Ala mutants will need to be done to determine which O-GlcNAc residues are required for TFG’s GlcNDaz-dependent crosslinking.
Figure 11: O-GlcNAcylation site mutants of TFG do not disrupt GlcNDAz crosslinking

Left: 293T cells were transfected with 3xFLAG-TFG O-GlcNAc point mutants. Then, cells were treated with DMSO or 100 μM GlcNDAz, and were exposed to UV. Lysates were analyzed by WB. None of the Ser/Thr → Ala mutants disrupted the GlcNDAz-dependent crosslinking of TFG. Crosslinked 3xFLAG-TFG indicated with an arrow, uncrosslinked with a circle. Right: 293T cells were transfected with 3xFLAG-TFG O-GlcNAc point mutants or a compound mutant. Then, cells were treated with DMSO or 100 μM GlcNDAz, and were exposed to UV. Lysates were analyzed by WB. Crosslinked 3xFLAG-TFG indicated with an arrow, uncrosslinked with a circle. None of the point or compound mutants of TFG disrupted GlcNDAz-dependent crosslinking.

As previously described, we evaluated the functional significance of site-specific SEC23A glycosylation in a human cell system (87). Using a genome-engineered human chondrosarcoma cell line, we evaluated endogenous collagen secretion and retention in cells expressing only SEC23A point mutants (87). In addition, we optimized another commonly used assay of COPII function. In the brefeldin A (BFA) washout assay, cells are treated with BFA, a fungal metabolite that disrupts vesicle trafficking by stabilizing an abortive intermediate of the COPI ARF complex (126,208-210). This disrupts COPI
traffic and subsequently causes the Golgi to collapse into the ER, thereby disrupting COPII traffic (209-211). The effects of BFA are reversible by BFA washout and subsequent recovery of the Golgi is COPII-dependent (126,208).

Using this assay, we first assessed whether the SEC23A−/− SW1353 cells reconstituted with wild type SEC23A responded to BFA treatment and washout with ER collapse and recovery, as reported previously (126,208). Cells were also transfected with an expression construct of alpha-mannosidase 2 (ManII), a resident Golgi protein, fused with GFP (ManII-GFP). After 30 minutes of BFA treatment, cells demonstrated diffuse GFP signal, indicating that the Golgi had collapsed into the ER (Figure 12A). After a washout and recovery period of 90 minutes, GFP distribution in the BFA treated cells was similar to untreated cells. (Figure 12A). Next, we used this BFA washout assay to assess the effect of our various SEC23A mutants on Golgi recovery. We detected no difference in SEC23A+/− SW1353 reconstituted with wild type, S184A, or 4A SEC23A in Golgi recovery time (Figure 12B and data not shown). In the collagen secretion assay we had observed no difference with any of the point or compound mutants without siRNA knockdown of SEC23B. This suggests that the BFA washout and recovery assay may show a more pronounced ER and collagen phenotype if sensitized via siRNA knockdown of SEC23B.
Figure 12: BFA washout in mutant SW1353 cells

(A) SEC23A−/− SW1353 cells stably reconstituted with wild type SEC23A-myc-6xHis were transfected with ManII-GFP and treated with BFA for 30 minutes (or untreated). After 30 minutes, BFA was removed and the Golgi was allowed to recover. ManII-GFP was diffusely localized after 0 minutes of recovery, indicating collapse of the Golgi into the ER. After a 90-minute recovery the ManII-GFP looked similar to the untreated control, displaying bright juxtanuclear fluorescence reminiscent of the Golgi. (B) SEC23A−/− SW1353 cells stably reconstituted with wild type or 4A SEC23A-myc-6xHis were transfected with ManII-GFP and treated with BFA for 30 minutes (or
untreated). No difference was observed between WT and 4A reconstituted cells after 60 minutes of recovery.

Also, the SEC23A 4A mutant showed no phenotype in any human chondrosarcoma cell system. However, the SEC23A 4A mutant demonstrated a significant increase in intracellular collagen retention when compared to wild type zebrafish chondrocytes or crusher zebrafish chondrocytes rescued with wild type SEC23A. Some of these observed differences could be attributed to differences in organism, cell culture model versus \textit{in vivo} model, or type of collagen imaged (collagen I in the SW1353 cells, collagen II in the zebrafish model). Future studies will focus on determining the molecular mechanism of site-specific O-GlcNAcylation of SEC23A on collagen secretion and COPII trafficking, which will allow further characterization in either model system.

We also created SEC23B\textsubscript{−/−} SW1353 cells using CRISPR/Cas9 methods. We verified SEC23B deletion via WB (Figure 13A). We assessed the ER morphology and collagen retention of SEC23B\textsubscript{−/−} SW1353 cells (Figure 13B). Compared to untreated SEC23A\textsubscript{−/−} SW1353 cells (Figure 6B and data not shown), there was no difference in ER morphology or collagen accumulation in the SEC23B\textsubscript{−/−} SW1353 cells. SEC23A\textsubscript{−/−} SW1353 cells exhibited a small but distinct population with enhanced collagen retention and expanded ER morphology (Figure 13B). In the future, we will treat SEC23B\textsubscript{−/−} SW1353 cells with ascorbate to stimulate collagen translation and secretion to assess whether the increased
secretory burden phenocopies ascorbate-treated SEC23A+ SW1353 cells. As previous reports indicate no skeletal abnormalities in SEC23B-deficient mice or humans with SEC23B mutations, and SEC23B-deficient MEFs exhibit normal collagen I staining, our observations were not surprising (94-100). Mutations in SEC23A cause CLSD, which consists of facial dysmorphisms, skeletal defects, late-closing fontanels, and cataracts, which are caused by improper collagen trafficking in the lens, whereas mutations in SEC23B cause CDAII, a form of anemia (88-90,94,95). Interestingly, human SEC23A and SEC23B are 85% identical and 96% similar, but, as seen in human disease, animal models, and our SW1353 knockout cells, their functions are not completely overlapping. Differences in tissue specific expression and the highly divergent ~18 amino acid region in SEC23 may explain some of the unique characteristics associated with each paralog.

Here we have demonstrated that TFG, a regulator of COPII function, engages in O-GlcNAc-mediated protein-protein interactions. Mutating specific O-GlcNAc sites within the C-terminus and truncating the protein, removing the first 193 amino acids of the N-terminus, did not disrupt GlcNDAz-dependent crosslinking. Further truncations, compound mutants, or novel O-GlcNAc point mutants will need to be made to determine the specific residues important for mediating these interactions. We also determined SEC23B is not required for collagen I secretion in a human chondrosarcoma cell line, whereas SEC23A is indispensable.
Figure 13: SEC23B−/− SW1353 cells display no collagen retention when untreated

(A) SW1353 cells stably expressing Cas9 and an sgRNA targeting a “safe harbor” locus (control) or the SEC23B locus (SEC23B sgRNA 1 & SEC23B sgRNA 2). The second guide (SEC23B sgRNA 2) is deficient in SEC23B. The first guide (SEC23B sgRNA 1) has no SEC23B knockdown when compared to a guide targeting the “safe harbor” locus (control). (B) Both SW1353 transfected with guides targeting SEC23B were examined for intracellular collagen staining (green) or expanded ER morphology (PDI, red). SW1343 SEC23B−/− (SEC23B sgRNA 2) cells do not exhibit any collagen or distended ER morphology, as observed with SW1353 SEC23A−/−.
3. Identifying Novel O-GlcNAcylated proteins

3.1 Introduction

O-GlcNAc regulates a broad range of physiological processes and modifies thousands of intracellular proteins (59,61,169,171,212,213). However, identifying changes in the O-GlcNAcome in response to a specific stimulus in an unbiased and quantitative manner remains difficult. In general, studying O-GlcNAc substrates remains challenging due to the lack of consensus sequences, sub-stoichiometric levels of modification, and the difficulty of performing MS on O-linked carbohydrates, which are labile during standard collision-induced dissociation fragmentation (214-217). Also, like phosphorylation, only a small fraction of O-GlcNAc substrates alter their glycosylation status in response to any given stimulus or signal. To address these challenges, we developed a novel, unbiased glycoproteomics workflow to survey global changes in O-GlcNAc in response to stimuli. This approach utilizes both stable isotope labeling with amino acids in cell culture (SILAC) and metabolic labeling for O-GlcNAc enrichment and profiling.

SILAC employs in vivo incorporation of a label into proteins for MS-based quantitative proteomics, relying on the metabolic inclusion of either ‘heavy’ or ‘light’ amino acids into proteins (218-220). This method utilizes amino acids with substituted stable isotopic nuclei (e.g. $^{13}$C,$^{15}$N). These ‘heavy’ labeled amino acids are stably incorporated into the cell’s proteome and after ~7 cell divisions the percent of proteins
labeled is >99% (218-220). Because there is no biochemical difference between the labeled ‘heavy’ amino acid and the natural ‘light’ amino acid isotopes, the ‘heavy’ and ‘light’ cells behave the same. However, the difference in amino acid weight is easily distinguished via MS, providing a quantitative approach to examine differences within these populations.

We aimed to couple SILAC with an O-GlcNAc enrichment method to profile global stimulus-induced changes in O-GlcNAc. Commonly used methods for profiling and enriching the O-GlcNAc come utilize O-GlcNAc antibodies, lectins, chemical derivatization methods, and chemoenzymatic and metabolic labeling (221,222).

Antibody-based enrichment methods suffer from the lack of availability of high-affinity O-GlcNAc antibodies (138,223). This problem stems from the difficulty of making O-GlcNAc antibodies, as the O-GlcNAc moiety is not highly immunogenic and carbohydrate-protein interactions are generally weak, both of which complicate antibody maturation and development (223). These problems can lead to false-negatives in experiments relying on current commercial anti-O-GlcNAc antibodies and necessitate verification via independent techniques (138,223).

Lectin-based enrichments also suffer from lack of specificity. Wheat germ agglutinin (WGA) is lectin that recognizes both terminal GlcNAc and sialic acid residues (224). WGA has been used with some success for purifying O-GlcNAcylated peptides using a method known as lectin weak affinity chromatography (LWAC) (225). In LWAC,
WGA conjugated to a substrate is packed into a lengthy column (~3 meters), which is coupled to a low flow rate HPLC. Compared to unmodified peptides, O-GlcNAc peptides are retained by the column and recovered in later fractions (225). However, the technique requires specialized hardware and relies on WGA, which lacks high affinity and specificity.

β-elimination/Michael addition (BEMAD) is a chemical derivatization-based method of O-GlcNAc enrichment (Figure 14). BEMAD uses a mild alkaline β-elimination to create dehydro-amino acids at sites of O-GlcNAc, followed by Michael addition with DTT (226).

![Chemical diagram](image)

**Figure 14: Schematic for BEMAD**

The β-elimination-Michael addition (BEMAD) workflow creates an MS-stable adduct in place of O-GlcNAc. First, modified hydroxyl residues on proteins of interest are β-eliminated using mild base to yield dehydro-amino acids at modification sites. This is followed by Michael addition at the resulting α,β-unsaturated carbonyl, often using dithiothreitol.

This approach circumvents the lability of the glycosidic bond during MS fragmentation, creating an adduct that is stable under collision-induced dissociation conditions. This methodology can be coupled with other enrichment strategies or be
used with other derivatizations, such as biotin pentyamine, placing an affinity handle at the site of O-GlcNAc modification after β-elimination (226). However, this technique is not without its drawbacks, such as labeling other O-linked PTMs, including phosphorylation, which may add confusion to determining the specific modification at a given site.

Chemoenzymatic labeling uses traditional galactosyltransferase (GalT) labeling supported by advanced chemical derivatization techniques, such as bioorthogonal chemistry. This method employs an engineered GalT with an enlarged binding pocket, GalT(Y289L), which relaxes substrate specificity, allowing the transfer of GalNAc or GalNAc analogs with N-acetyl substitutions, such as ketones or azides, onto O-GlcNAc acceptors (227). Commonly used synthetic nucleotide-sugar donors include UDP-keto-galactose and UDP-N-azidoacetylgalactosamine (UDP-GalNAz), which are then reacted with aminooxy- or alkyne-functionalized probes, respectively (59,228,229). Probes often feature Flag epitopes, PEG, biotin, fluorophores, etc. for downstream detection and/or purification. The variety in commercially available aminooxy- or alkyne-functionalized probes lends to the power of the GalT assay. However, because the GalT assay will indiscriminately label all O-GlcNAc moieties within a sample, whether relatively new or stable, it is not well suited for discriminating between novel, signal dependent changes and older, housekeeping glycans.
Metabolic labeling enrichment strategies aim to circumvent the issues with GaIT labeling, allowing time resolution labeling of signal-specific changes. In the following work we primarily use a synthetic precursor, peracetylated-GalNAz (AcGalNAz), which transits the endogenous GalNAc salvage pathway and is converted into UDP-GalNAz by cultured mammalian cells (136). Next, it is converted from UDP-GalNAz to UDP-GlcNAz by endogenous UDP-galactose 4-epimerase (GALE). Importantly, OGT accepts UDP-GlcNAz and decorates endogenous substrates with O-GlcNAz, whose azide moiety provides a bioorthogonal reactive handle (136). The azidosugar can be “clicked” to an alkyne-functionalized probe using copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC), also known as click chemistry (230). We demonstrated enrichment of GalNAz-modified proteins using two alkyne-functionalized probes, alkyne-biotin and alkyne-agarose, and there are many more available. This metabolic labeling approach has been previously used in our lab to profile nucleo-cytoplasmic O-GlcNAc substrates, to identify mitochondrial glycoproteins involved in organelle function and apoptosis, and to identify substrates that regulate NRF2 ubiquitination (136,230-232).

We developed a novel glycoproteomic workflow utilizing SILAC and metabolic labeling in tandem to identify global changes in O-GlcNAcylation in response to stimuli. Here we examine the effects of brefeldin A (BFA), a fungal metabolite that disrupts
vesicle trafficking, and cytokine deprivation on a pro-B cell line using our
glycoproteomic workflow.

3.2 Materials and Methods

**Cell culture** – Ramos were cultured in Roswell Park Memorial Institute medium (RPMI) containing 10% fetal bovine serum, 100 units/mL penicillin, and 100 µg/mL streptomycin in 5% CO2 at 37 °C. FL5.12 (XL4.1 and N6) were cultured in RPMI containing 10% fetal bovine serum, 100 units/mL penicillin, 100 µg/mL streptomycin, 55 µm β-mercaptoethanol, 2 mM L-glutamine, and 10 mM HEPES in 5% CO2 at 37 °C. FL5.12 media was also supplemented with 500 pg/mL recombinant mouse IL-3.

**Cellular Fractionation** – This protocol was adapted from a Lamond Lab cytoplasmic and nuclear fractionation protocol. Cells were pelleted and washed once with cold PBS. Cold PBS was removed and cells were resuspended in 5 mL of ice-cold Buffer A (1.5 mM MgCl₂, 10 mM KCl, 10 mM HEPES, pH 7.9) supplemented with protease inhibitors, PUGNAc, and UDP. Cells were broken open using a pre-chilled Dounce homogenizer and ~30 strokes with a tight pestle. Cell integrity was monitored using a hemocytometer. After Douncing, cells were centrifuged at 228 g for 5 min at 4º C for 5 minutes. The supernatant from this fraction was retained as crude cytoplasmic and the pellet as crude nuclear. To obtain a cleaner nuclear pellet, the crude nuclear pellet was resuspended in 3mL of buffer S1 (0.35 M sucrose, 0.5 mM MgCl₂) supplemented with protease inhibitors, PUGNAc, and UDP and layered over a cushion of buffer S3.
(0.88 M sucrose, 0.5 MgCl₂), and centrifuged at 2800 g for 10 minutes at 4°C. This results in a cleaner nuclear pellet. To obtain a cleaner cytoplasmic fraction the crude cytoplasmic fraction was centrifuged at >400,000 g for 1 hr at 4 °C. The supernatant was retained as the cytoplasmic fraction. The clean nuclear pellet was lysed in Click buffer (1% Triton X-100, 1% SDS, 150mM NaCl, 20 mM Tris pH 7.4) and the cytoplasmic fraction was brought to a final concentration of Click buffer.

**GalNAz Labeling and Affinity Purification** – Cells were treated with 100 µM GalNAz up to 24 hours prior to harvesting. After harvesting, cells were lysed in Click buffer supplemented with protease inhibitors, PUGNAc, and UDP. Also, metal chelators and reducing agents were avoided because of their interference with further downstream steps. After lysis or subcellular fractionation, reagents were assembled in an ice bucket. The following reaction components were added, in order, to the listed final concentration: protein sample, 5 mM sodium ascorbate, 25 µM alkyne-biotin, 100 µM TBTA, 1 mM CuSO₄. Reactions were mixed and rotated gently at room temperature for an hour. The reaction was quenched by adding EDTA to a final concentration of 10 mM. For immediate analysis, SDS-PAGE sample buffer was added directly to reactions and loaded on a gel (without boiling). For further processing and affinity purification, samples were precipitated using methanol. Reactions were mixed with ice cold methanol (10:1, methanol:lysate). After mixing, samples were placed on dry ice or stored in the -80 °C for 10 minutes to increase precipitation. Then, samples were spun at 17,000
g to pellet. Supernatant was removed and the pellet was resuspended in methanol and placed on ice. This process was repeated a total of four times. After the final precipitation, the protein pellet was dissolved in 4 M guanidine in PBS. A small fraction of this sample was saved for use as inputs and positive controls in later steps.

Biotinylated proteins were captured from the samples by incubating overnight at 4 °C with gentle rotation with NeutrAvidin beads (ThermoFisher, Pierce High Capacity NeutrAvidin Agarose, 29204). The following day, beads were washed three times with the following buffers, in order: 4 M guanidine in PBS, 5 M NaCl in H2O, 6 M urea in PBS, and 1% SDS in PBS. Proteins were eluted by boiling in 2X SDS-PAGE sample buffer. The reserved input samples were exchanged into another buffer either via column (BioRad, Bio-Spin 6, 7326221) or methanol/chloroform precipitation, because guanidine and SDS are incompatible.

**On-bead Click** – FL5.12 (XL4.1) cells were seeded for optimal density for treatment and harvest the at the time of GalNAz treatment. In all experimental conditions, 6 hours prior to harvesting cells were treated with 100 µM GalNAz. For Brefeldin A (BFA) treated samples, cells were treated 4 hours prior to harvesting. In interleukin 3 (IL-3) treated samples, cells were starved of IL-3 for 48 hours prior to harvesting and re-supplemented with IL-3 1 hour before harvesting. Heavy and Light-treated cells were pooled 1:1 and fractionated followed the Cellular Fractionation protocol. Protein amounts were quantified, and 2 mg total protein was used in both the
nuclear and cytoplasmic fraction in the following steps. Alkyne-agarose beads (Click Chemistry Tools, Alkyne Agarose, 1032-2) were equilibrated with Click buffer by washing three times. 2 mg total protein for both nuclear and cytoplasmic fractions were precleared with 150 µL bead volume of equilibrated alkyne-agarose beads with gentle rotation for 2 hours at room temperature. After preclearing, supernatant was retained, and beads used for preclearing were discarded. The retained cytoplasmic and nuclear fractions were combined with 50 µL of equilibrated alkyne-agarose beads and the following were added, in order, to the listed final concentration: 5mM sodium ascorbate, 100 µM TBTA, 1mM CuSO₄. The reaction was rotated at room temperature for 2 hours and EDTA was added to a final concentration of 10 mM to quench the reaction. Beads were then washed sequentially with three 1 mL washes of the following: 1% SDS, 20 mM Tris pH 7.4; 1% SDS, 10 mM DTT, 20 mM Tris pH 7.4; 1X PBS; 8 M urea; 1X PBS; 6 M guanidine hydrochloride; 1X PBS; 5 M NaCl; 1X PBS; 10X PBS; 1X PBS; 20% isopropanol; 20% acetonitrile; 50 mM ammonium bicarbonate. After the final ammonium bicarbonate wash, beads were stored at 4 ºC in 100 µL 50 mM ammonium bicarbonate until they were submitted for on-bead trypsin digestion, LC-MS/MS analysis, and quantification.

**Beta-elimination** – Alkyne-agarose beads coupled with proteins after click reaction were incubated with 5 mM NaOH, 25 mM NaOH or water at 37 ºC overnight.
Samples were resolved via SDS-PAGE and either analyzed via Western blot or silver stain (ThermoFisher, SilverQuest Silver Staining Kit, LC6070)

**Endogenous O-GlcNAc IP** – Cells were washed with cold PBS and lysed in IP lysis buffer without SDS (1% Triton X-100, 150 mM NaCl, 1 mM EDTA, 20 mM Tris-HCl pH 7.4) supplemented with protease inhibitors, PUGNAc, and UDP. Lysates were probe-sonicated, cleared by centrifugation and quantified by BCA protein assay according to the manufacturer’s instructions. IPs were performed on 5-10 mg total protein. Cleared lysates were adjusted to a final total protein concentration of ~2-5 mg/mL using IP lysis buffer supplemented with protease inhibitor, UDP and PUGNAc. For every 1 mg of protein lysate used 1 µg of each of the following antibodies was added: RL2 (SantaCruz Biotechnology, sc-59624), 1F5 (ThermoFisher, MA1-040), 9D1 (ThermoFisher, MA1-039), 18B10 (ThermoFisher, MA1-038) and rotated overnight at 4 ºC. The following day, 50 µL equilibrated protein A/G UltraLink Resin (53133, ThermoFisher) were added to the lysate and rotated at room temperature for 1 hour. Beads were washed three times with 1 ml of IP lysis buffer and then eluted in 2X SDS-PAGE sample buffer. Eluents were analyzed via WB.

**Coomassie** – After SDS-PAGE, gels were incubated in fixation buffer (30% methanol, 10% acetic acid) for at least 1 hour. After fixation, gels were washed with water for 15 minutes with rocking. Water wash was repeated 3 more times. Next, the colloidal “blue silver” solution (10% phosphoric acid, 10% ammonium sulfate, 0.12%
Coomassie Blue G-250, 20% methanol) was added to the gel for at least 3 hours, but overnight staining was preferred. After staining, gel was destained with water washes, changing as necessary.

**Methanol/Chloroform Precipitation** – 200 µL of lysate was added to a 1.5 mL microcentrifuge tube. Next, 600 µL of methanol was added and vortexed briefly. Then 150 µL of chloroform was added to the tube and vortexed briefly. Next, 500 µL of water was added and vortexed briefly. After the addition of water, the solution becomes cloudy. Samples were centrifuged at 17,000 g, then the upper aqueous layer was carefully removed, leaving the interface intact. Leaving excess aqueous layer above the interface in lieu of disturbing the interface was preferred. Next, 450 µL of methanol was added and samples were centrifuged at 17,000 g to pellet the protein. The supernatant was removed and discarded. Tubes were left open and covered with a lint-free tissue and the pellet was allowed to air dry. Protein was resuspended in buffers compatible with downstream protocols.

**SILAC Media Cell Culture** – L-lysine- and L-arginine-deficient RPMI 1640 medium for SILAC (ThermoFisher, A33823) was supplemented with 10% dialyzed and heat inactivated FBS (Corning, 35-071-CV), 1% PenStrep, and amino acids. Heavy medium was supplemented with 12.5 mg $^{13}$C$_6$$^{15}$N$_4$-arginine, 12.5 mg $^{13}$C$_6$$^{15}$N$_2$-lysine, and 5mg proline. Light medium was supplemented with 12.5 mg arginine, 12.5 mg lysine,
and 5 mg proline. Proline supplementation prevents conversion of arginine to proline (233).

**Conditioning Heavy and Light FL5.12 (XL4.1)** – Cells were passaged for at least 7 doublings in either Heavy or Light SILAC RPMI for >99% isotope incorporation. Heavy isotope incorporation was verified via MS at the Duke Proteomics Facility.

**Clathrin IP** – Cells were washed with cold PBS and then lysed in 1 mL clathrin IP buffer (1mM EDTA, 150 mM NaCl, 50 mM Tris, pH 7.05) supplemented with 0.5% NP-40 and protease inhibitors with rotation at 4 ºC for 1 hour. Samples were centrifuged at 500 g at 4 ºC for 20 minutes and 30 µL of supernatant was reserved as input. The remaining supernatant was added to a microcentrifuge tube with protein A/G UltraLink Resin and clathrin antibodies (CHC17: X-22 (mouse monoclonal antibody) or CHC22: mAB-26 (mouse monoclonal antibody), gifts of Dr. Frances Brodsky (University College London) and rotated at 4 ºC for 2 hours. Next, beads were washed three times with 1 mL of clathrin IP wash buffer (PBS with 0.05% NP-40), then once with clathrin IP buffer. Samples were eluted with 100 µL 2X SDS-PAGE sample buffer and heating at 95 ºC for 5 minutes. Eluents were analyzed via WB.

*Methods not explicitly described within this chapter can be found in previous chapters.*
3.3 Results

3.3.1 FL5.12 (XL4.1) Support Both SILAC and Metabolic Labeling Approaches

Many cell types are amenable to both AcGalNaz treatment and SILAC conditioning. However, we first verified that the cell type we chose, FL5.12, would efficiently incorporate the azidosugar. We chose FL5.12, a murine interleukin 3 (IL-3)-dependent pro-B cell line, because it is a model system for the cytokine-induced nutrient uptake and proliferation of B lymphocytes (234-236). Previous work demonstrated that lymphocyte activation induces dramatic changes in global O-GlcNAcylation, suggesting that O-GlcNAc may govern proliferative signaling (3). Although the O-GlcNAcylation of substrates within the nuclear and cytoplasmic compartments were observed changing with radiolabeled substrates resolved via SDS-PAGE, the specific identify of each protein was not obtained (3).

We used two varieties of FL5.12 in the following studies, N6, a parental cell line, and XL4.1, a line transduced with Bcl-XL (236,237). XL4.1 resists apoptosis under a wide variety of cell stresses and survives IL-3 deprivation, making them a tractable model for studying changes in O-GlcNAc in response to growth factors, nutrient availability, cell cycle arrest, ER stress, etc. Also, during B cell activation the COPII client burden is greatly enhanced by proliferation and immunoglobulin secretion (157,161). Importantly, FL5.12 cells are dependent on IL-3 for growth and proliferation, making them an ideal candidate for studying specific changes in O-GlcNAc in response to a specific growth
factor. We hypothesized that changing the secretory pathway burden in these cells would alter O-GlcNAcylation of interesting secretory pathway components, which we could identify and quantify \textit{a priori}.

We wanted a cell type that would rapidly incorporate the azidosugar into the cells’ endogenous glycans, allowing GalNAz labeling in a short pulse before stimuli, preventing the accumulation of O-GlcNAz on housekeeping proteins and reducing background within the assay. Under standard growth conditions XL4.1 incorporate the azidosugar into their glycans detectable via WB at 2 hours after alkyne-biotin click (Figure 15A). Next, to determine whether our enrichment strategy would efficiently capture azidoglycoproteins we treated cells with GalNAz and captured the azidoglycan-labeled proteins on alkyne-agarose via click reaction, washed stringently, released the proteins using β-elimination, and analyzed the eluents via WB and silver stain (Figure 15B). The β-elimination conditions used, although commonly used elsewhere, hydrolyzed not only the glycosidic bond, but the peptide backbone (222). This is demonstrated by the increased intensity in the GalNAz-treated and β-eliminated lanes, but lack of discrete bands in the silver stained gel (Figure 15B) and complete lack of signal via WB for control glycoproteins (e.g. P62, TFG, etc.) (data not shown). To further test our enrichment strategy, we performed an on-bead trypsin digestion after azidoglycoprotein capture and examined the captured substrates. Compared to a vehicle-treated control, lysates from GalNAz-treated cells showed a marked enrichment
in common O-GlcNAc substrates, including HCF-1 and nuclear pore complex proteins (data not shown).

Figure 15: XL4.1 rapidly incorporate GalNAz.

(A) XL4.1 were treated with DMSO or GalNAz for the indicated time. Lysates were clicked with alkyne-biotin and analyzed via WB. Biotin signal indicates incorporation of GalNAz into endogenous glycans. (B) XL4.1 were incubated with GalNAz for 4 hours, lysed and clicked onto alkyne-agarose beads. Samples were then β-eliminated with water (-), 5mM NaOH (+), or 25mM NaOH (++) at 37 ºC overnight. Eluents were resolved via SDS-PAGE and visualized by silver stain (above) or WB (not shown). We observed no discrete protein bands after β-elimination, most likely due to hydrolysis of both the glycosidic bond and the peptide backbone.

We first tested whether XL4.1 would incorporate heavy arginine ($^{13}$C$_6$^{15}$N$_4$-arginine) and heavy lysine ($^{13}$C$_6$^{15}$N$_2$-lysine) into their proteomes. XL4.1 were passaged for 7 doublings in both heavy and light SILAC media. After 7 doublings, heavy cells
were analyzed via MS and had incorporated the heavy amino acids into >99% of their proteome (personal communication, Erik Soderblom).

3.3.2 Glycoproteomic Profile After BFA Treatment

Next, we used our glycoproteomic workflow to identify changes in O-GlcNAcylated proteins induced by BFA, a fungal metabolite that disrupts COPI and COPII traffic. BFA also induces the unfolded protein response (UPR) (238,239). Previous work has indicated that the hexosamine biosynthetic pathway (HBP), which generates the nucleotide sugar donor for O-GlcNAc, UDP-GlcNAc, is transcriptionally activated in response to the UPR (240). We hypothesized that BFA treatment would trigger changes in O-GlcNAc signaling involved with secretory pathway trafficking.

We first determined a dose of BFA that would induce the UPR and inhibit the secretory pathway, without causing a large loss in cell viability. However, determining the dose of BFA using the XL4.1 is difficult because they are resistant to cell stress and apoptosis (236,237). While this attribute makes XL4.1 useful for profiling changes in response to a wide variety of stimuli, it makes determining a sub-lethal dose of BFA (or any stressor) difficult. Using the N6 parental FL5.12 cell line, we determined that 500 ng/mL of BFA was tolerated at both 4 and 24 hours, without a large reduction in cell viability, as determined by CellTiterGlo (Figure 16) and MTS (data not shown). However, higher doses (≥1µg/mL BFA) induced cell death after 24 hours.
Figure 16: XL4.1 resist apoptosis induced by Brefeldin A

N6 and XL4.1 cells were treated with increasing doses of BFA to induce cell death. Cells health was assessed at the indicated timepoints (4 or 24 hr) via CellTiterGlo. XL4.1 resist death induced by BFA at all doses, whereas N6 show a marked decrease in cell viability starting at 1 μg/mL BFA.

To begin our SILAC glycoproteomics workflow, we treated heavy and light SILAC-labeled XL4.1 cells with 100 μM GalNAz for 2 hours, followed by 500 ng/mL BFA (or vehicle) for another 4 hours. Next, we harvested the cells by pooling the heavy and light (BFA vs DMSO treated cells in the first biological replicate) one to one. Pooling the cells prior to any processing prevents differences due to sample handling and is tolerated because each cell population (treatment condition) is distinctly labeled via SILAC. Next, we fractionated cells into nuclear and cytoplasmic fractions with Douncing and sucrose gradient centrifugation. This step removes lumenal secretory pathway
glycans, which could contain O- or N-linked glycans containing either GalNAz or GlcNAz, and would interfere with O-GlcNAz capture. Also, previous reports have demonstrated that some proteins change cellular compartment in a PTM-dependent fashion, a change we would not detect without subcellular fractionation. One example of this is the transcription factor NeuroD1, which under low glucose conditions is mainly in the cytosol. However, under high glucose conditions, NeuroD1 becomes O-GlcNAcylated and translocates to the nucleus (241).

Next, using click chemistry, we covalently coupled the GalNAz labeled proteins to the alkyne-agarose substrate, and washed stringently. After the final wash, the beads were stored in 50 mM ammonium bicarbonate until submission to the Duke Proteomics facility for on-bead trypsin digestion, LC-MS/MS analysis, and quantification.

We analyzed the SILAC results for BFA-dependent changes in the O-GlcNAcome. In the first experiment, the heavy cells were treated with BFA (light with DMSO), and we identified 1253 and 793 proteins in the nuclear and cytoplasmic fractions, respectively (Figure 17 and Appendix A). Because the experimental workflow and data processing are time-consuming, we attempted to validate putative hits from this initial dataset as we simultaneously prepared a biological replicate. Clathrin heavy chain 1 (CLTC) decreased with a 2.7-fold-change in nuclear fraction after BFA treatment. However, further experiments demonstrated that clathrin heavy chain O-GlcNAc levels were undetectable and did not change after Thiamet-G treatment, as seen with both
endogenous IP and biotin enrichment after metabolic labeling (Figure 18 and data not shown).

Figure 17: Log$_2$ of vehicle/BFA ratio in the cytoplasm and nucleus.

Scatter plots representing the log$_2$ of vehicle treated intensity over BFA treated intensity from the first biological replicate of the BFA glycoproteomics workflow. We identified 793 proteins in the cytoplasmic fraction (top, blue) and 1253 proteins in the nuclear fraction (bottom, red).
XL4.1 cells were treated with either vehicle and GalNAz for 24 hours as indicated. Cells were also treated with either Thiamet-G, as indicated, to enhance global O-GlcNAcylation. Lysates were clicked with alkyne-biotin and enriched via NeutrAvidin. Eluents were analyzed by WB. No specific enrichment of clathrin was detected with either clathrin heavy chain antibody (CHC17 and CHC22). However, SEC24C, a positive control was enriched in a GalNAz-dependent manner.

In a biological replicate, heavy cells were treated with DMSO (light with BFA), and we identified 1704 and 1263 proteins in the nuclear and cytoplasmic fractions, respectively (Figure 19 and Appendix A). We calculated the heavy/light intensity ratio for every protein in each biological replicate. Overall, BFA barely altered the vast majority of captured proteins in either fraction (~96% and ~80%), that is they were
equally enriched in both fractions. This observation was expected, as with many
signaling events and PTMs only a small subset of proteins change after a stimulus.

Next, we applied stringent filters to a combined data set to identify candidate
BFA-dependent changes in O-GlcNAcylated proteins. First, we compared fold-changes
across biological replicates and retained only protein IDs with fold changes consistent
across biological samples (e.g. COPG went down in response to BFA). Next, we
examined proteins in the nuclear fraction with a fold-change greater than 2 or lower
than -2. A similar filter was placed on the cytoplasmic fraction, except we examined
proteins with a fold-change greater than 1.5 or less than -1.5. This less stringent filter
was placed on the cytoplasmic fraction because it had fewer total protein IDs and less
intense fold-changes overall. With these filtering conditions we identified 80 and 17
unique proteins of interest in the nuclear and cytoplasmic fractions, respectively.

Within this filtered set of proteins, we chose to pursue the following: 14-3-3,
alpha-enolase (ENO1), coatamer subunit gamma-1 (COPG), heat shock protein HSP 90-
alpha (HS90A), importin-5 (IPO5), E3 ubiquitin-protein ligase NEDD4, pyruvate kinase
M2 (PKM2), paired amphipathic helix protein Sin3a, SUN domain-containing protein 2
(SUN2), and ubiquitin-like modifier-activating enzyme 1 (UBA1). We selected these
proteins based on previous reports of O-GlcNAcylation, availability of antibodies,
and/or biological relevance. ENO1, HSP90, IPO5, NEDD4, PKM2, SIN3A, and UBA1 are
previously reported as O-GlcNAcylated, but the effect of O-GlcNAc on these substrates has yet to be elucidated (242-247).

Figure 19: Log2 of BFA/vehicle ratio in the cytoplasm and nucleus

Scatter plots representing the log2 of BFA treated intensity over vehicle treated intensity from the second biological replicate of the BFA glycoproteomics workflow. We identified 1263 proteins in the cytoplasmic fraction (top, blue) and 1704 proteins in the nuclear fraction (bottom, red).
Next, to examine our candidate hits, we used O-GlcNAc IPs followed by WB. Validating targets using the endogenous glycan is preferred, but this can be difficult due to the nature of O-GlcNAc antibodies that are available. All O-GlcNAc antibodies are prone to false-negative results because they recognize targets based partly on specific sequence contexts, relying on residues flanking the O-GlcNAc which vary widely on modified proteins (140,222,223). Also, the limited antigenic properties of O-GlcNAc prevent the development of high-affinity antibodies (222).

To circumvent some of these shortcomings we pooled four commercially available mouse monoclonal anti-O-GlcNAc antibodies (clones RL2, 1F5, 9D1, and 18B10) to make a cocktail for IP (138,248). We also included another cell line in our validation, Ramos, a human B lymphocyte cell line. We hypothesized that any changes in XL4.1 would be conserved across other B lymphocyte-derived cell lines. However, we were unable to validate any of the hits using this approach. With every candidate protein, we observed no specific IP of the protein, indicating either poor O-GlcNAc enrichment or non-specific binding of the protein to the beads (data not shown).

Next, we attempted to validate targets using biotin-click followed by streptavidin enrichment. We treated cells with GalNAz and fractionated into nuclear and cytoplasmic fractions. However, rather than covalently attaching the azidoglycoproteins to an alkyne-agarose substrate, we clicked on alkyne-biotin. Proteins were enriched via NeutrAvidin beads, stringently washed and analyzed by WB. Using
this approach, both COPG and UBA1 exhibited changes consistent with the BFA glycoproteomics workflow data and we chose to focus future validation experiments on these proteins (Figure 20 and data not shown).

<table>
<thead>
<tr>
<th>Input</th>
<th>Pulldown</th>
</tr>
</thead>
<tbody>
<tr>
<td>- - +</td>
<td>- - +</td>
</tr>
<tr>
<td>- + -</td>
<td>- + -</td>
</tr>
</tbody>
</table>

**Figure 20:** COPG enrichment via biotin-click decreases after treatment with BFA in the nuclear fraction.

XL4.1 were treated as in the BFA proteomics workflow. After fractionation, lysates were alkyne-biotin clicked and enriched via NeutrAvidin. Eluents were analyzed via WB. COPG decreases in response to BFA in the nuclear fraction. Biotin serves as an enrichment and pulldown control, P62 functions as a nuclear compartment and glycoprotein pulldown control.
COPG is a coatomer protein for COPI, which mediates traffic from the Golgi to the ER and between Golgi stacks (68). It is conserved from yeast to mammals and is essential in yeast (68,249). UBA1 catalyzes the first step in ubiquitination and participates in DNA damage response in cells (250). Neither are reported as O-GlcNAcylated, but because BFA is a secretory pathway poison that functions via COPI inhibition O-GlcNAc may function as a regulatory modification on COPG.

The other proteins examined showed no enrichment with GalNAz and alkyne-biotin (data not shown). However, we observed some inconsistencies in UBA1 and COPG. In one replicate UBA1 was not enriched from XL4.1 and was undetectable via WB, and in another replicate, we observed an increase in COPG enrichment in the nuclear fraction, a change opposite to our BFA proteomics result.

Next, we performed endogenous IPs of COPG and UBA1 and probed via WB. Both COPG and UBA1 are readily IPed from XL4.1 and Ramos lysates, but only the COPG IP was robust enough for further downstream analysis (Figure 21A). COPG O-GlcNAc signal is also greatly enhanced with Thiamet-G treatment (Figure 21B). Current efforts focus on validating changes in O-GlcNAc on COPG after BFA treatment, but the basal level of O-GlcNAcylation as detected by WB is low, which may make detecting BFA-dependent decreases difficult.
Figure 21: O-GlcNAc detected on endogenous COPG

(A) IP from Ramos lysates for endogenous COPG and UBA1. Beads only and irrelevant antibodies were included negative as controls (COPG– mouse monoclonal; UBA1 – rabbit polyclonal). Eluents were analyzed by WB. COPG endogenous IP was more efficient than UBA1. (B) Ramos and XL4.1 cells were treated with either DMSO or 50 μM Thiamet-G for 8 hours, and IPed for COPG. COPG O-GlcNAcylation was detected in both Ramos and XL4.1 with 18B10. However, only COPG IPed from Ramos was detected by RL2. Uncropped RL2 included as control for Thiamet-G treatment.

3.3.3 Glycoproteomic Profile After IL-3 Starvation

Next, we used our glycoproteomic workflow to identify changes in the O-GlcNAcome after specific growth factor deprivation and re-feeding. We withdrew IL-3 from SILAC-labeled XL4.1 for 48 hours to induce quiescence. Then, we pulse-labeled the cells with 100 μM azidosugar 4 hours prior to harvesting, followed by treatment with IL-3 one hour prior to harvesting to induce growth and proliferation. Then, we profiled IL-3-dependent changes in O-GlcNAc substrates via our glycoproteomic workflow. We
performed this experiment twice, where IL-3 re-fed cells were heavy-labeled in the first replicate and light-labeled in the second. In the first replicate we identified 1498 and 855 proteins in the nuclear and cytoplasmic fractions, respectively (Appendix A). In the second replicate we identified 1155 and 345 proteins in the nuclear and cytoplasmic fractions, respectively (Appendix A).

We initially applied stringent filtering criteria, including proteins from the nuclear and cytoplasmic fractions with a fold change greater than 2 or less than -2 and identifying proteins with consistent fold-changes across biological replicates. However, using these criteria there were no proteins common between the samples. Next, we applied a less stringent filter. We retained proteins identified by at least three unique peptides from each fraction. We then compared the fold-change of each individual peptide and assessed whether the ratio of peptides changing was consistent across biological replicates. Using this less stringent approach, we identified no candidate proteins. Unfortunately, the only proteins consistent across biological replicates were contaminants (e.g. keratin). This was due to the poor-quality data obtained by the Proteomics Facility from the second biological replicate, where independent injections from the same sample showed significant differences in the liquid chromatography (LC) traces and MS results.

Because of this problem, we pursued hits from the first biological replicate. Interestingly, SEC23A was depleted ~70-fold from the IL-3-stimulated sample, without
any detectable loss in total SEC23A level in whole-cell lysate, as judged by WB (Figure 22). However, our efforts to validate this change in SEC23A O-GlcNAcylation using an MS-independent method failed to consistently detect O-GlcNAc on SEC23A. Our attempts included metabolic labeling and chemoenzymatic labeling (Click-IT O-GlcNAc Enzymatic Labeling System) coupled with NeutrAvidin enrichment, SEC23A endogenous IPs, and O-GlcNAc IPs followed by WB. With both metabolic labeling and chemoenzymatic labeling we detected no consistent enrichment of SEC23A using NeutrAvidin pulldown. SEC23A endogenous IPs are specific but inefficient, preventing detection via O-GlcNAc antibodies, and O-GlcNAc IPs yield no detectable SEC23A via WB (data not shown).

Interestingly, our IP/WB validation attempts of SEC23A revealed an IL-3-dependent increase in O-GlcNAcylation of SEC24C (Figure 22). SEC24C is a COPII inner coat protein, which binds SEC23A, and is also O-GlcNAcylated (68,87,135). We also frequently use SEC24C as a positive control for biotin-click and NeutrAvidin enrichment. In concert with our data from Chapter 2, this suggests that O-GlcNAcylation of COPII may exert a wide variety of biochemical functions on COPII pathway activity.
Figure 22: Dramatic change in SEC24C after IL-3 stimulation

XL4.1 were starved of IL-3 for 48 hours and then treated with 100 μM GalNAz 6 hours prior to harvest. One hour prior to harvesting, cells were either stimulated with IL-3 of PBS, as indicated. Eight hours prior to harvesting cells were also treated with Thiamet-G (TG) and glucosamine or DMSO and PBS (vehicle controls), as indicated. An O-GlcNAc cocktail IP was performed on the lysates and eluents were analyzed by WB. SEC24C demonstrated a dramatic change in enrichment after IL-3 stimulation. SEC23A has no change in total or IPed amounts after IL-3 stimulation. Tubulin is a negative control and P62 functions as a glycoprotein positive control.

3.4 Discussion

O-GlcNAc is a highly dynamic PTM that modifies thousands of nuclear, cytoplasmic, and mitochondrial proteins. Although the number of O-GlcNAc modified substrates identified continues to increase, the specific function of O-GlcNAc on many of these proteins remains elusive. Assessing changes in the O-GlcNAcome after a stimulus is a major challenge within the field. There are myriad approaches for O-GlcNAcome
enrichment, but each suffers from its own shortcomings. Our results indicate that a time-resolved approach to purifying and quantifying the O-GlcNAcome in an unbiased manner after a stimulus is feasible and useful. Using our quantitative glycoproteomics approach we identified changes in both previously reported and novel O-GlcNAcylated substrates after defined stimuli.

Treating cells with BFA and profiling the O-GlcNAcome with our glycoproteomics workflow, we identified changes in 97 nucleocytoplasmic proteins, some of which are previously reported glycoproteins. Further MS-independent validation demonstrated that both COPG and UBA1 O-GlcNAc-based enrichment decreased with BFA treatment (Figure 20 and data not shown). We are currently focusing on validating these observations using endogenous proteins and glycans.

BFA disrupts COPI traffic by stabilizing an abortive intermediate of the ARF complex, which in turn disrupts both COPI and COPII traffic. COPI transport carriers mediate transport of cargo from the Golgi to the ER and between the Golgi stacks. Like COPII, COPI has a GEF (Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1, GBF1), which activates the small GTPase ADP-ribosylation factor 1 (ARF1). ARF1 then undergoes a conformational change, inserts an N-terminal amphipathic α-helix into the Golgi membrane, and recruits the stable heteroheptameric coat. When the heteroheptameric coat, which contains COPG, is recruited it undergoes a major conformational change that leads to oligomerization of
the coat complex (251). This process of coat polymerization is catalyzed by COPG binding to p24 proteins, which are a group of ~24kDa transmembrane proteins responsible for cargo selection and coat recruitment found in the Golgi (68,252). COPG was the only COPI component whose O-GlcNAcylation changed with BFA treatment in our SILAC datasets. Many other COPI components were identified, including coatmer subunit alpha (COPA), coatmer subunit beta-1 and 2 (COPB & COPB2), coatmer subunit delta (COPD), coatmer subunit gamma-2 (COPG2), and coatmer subunit zeta-1 (COPZ). Because we did not include controls without GalNAz in the BFA SILAC experiment, their presence could be due to a GalNAz-independent mechanism, such as non-specific binding to the alkyne-agarose substrate. Interestingly, O-GlcNAc modification of COPG has never been reported, nor has the O-GlcNAcylation of any other COPI components. As we have seen with COPII components, O-GlcNAc can alter membrane binding and residues modified by O-GlcNAc are important for proper COPII function (Figures 5, 7, 10) (87). O-GlcNAc may have a comparable role on COPI components, exerting similar effects on COPI function. Future studies will examine the functional effects of O-GlcNAc on COPG and examine the O-GlcNAc status of other COPI components.

Many previously reported O-GlcNAcylated COPII components were also identified in our glycoproteomics workflow, including SEC23A/B, SEC24A, and SEC31A, though their O-GlcNAcylation did not change with BFA treatment.
Interestingly, we also identified SEC13 and SAR1A/B, suggesting they may be modified by O-GlcNAc. Because we did not include a sample not treated with GalNAz, it is possible their presence in the samples is due to non-specific interactions. However, this suggests that COPII O-GlcNAcylation may extend beyond the proteins we site-mapped in Chapter 2.

Unfortunately, our data profiling IL-3-dependent signaling changes did not contain any consistent changes across biological replicates. This lack of consistent changes was likely due to undiagnosed technical problems with the LC/MS procedures in the Proteomics Facility (personal communication, Erik Soderblom). However, we chose to proceed with the initial biological replicate, because it contained consistent technical replicates. We observed a ~70-fold decrease in SEC23A enrichment without any change in total protein, as determined by WB (Figure 22). Due to the increased secretory burden, rapid expansion, and proliferation associated with IL-3 stimulation, we hypothesize that COPII function would need to rapidly increase in response to this cytokine. This would suggest that the SEC23A O-GlcNAcylation could be an inhibitory modification, an observation that is consistent with the altered membrane association of SEC23A after Thiamet-G treatment (Chapter 2) (87). Unfortunately, ascertaining the O-GlcNAc status of SEC23A via non-MS based methods has been technically challenging, as it is not readily recognized by any of the available O-GlcNAc antibodies.
As with all MS methods, candidates should be validated via MS-independent methods. Here, we employed a variety of MS-independent methods to validate candidates. Relying solely on our glycoproteomics workflow, we cannot distinguish between a protein with decreased stability but unchanged O-GlcNAcylation post-stimulus versus a stable protein with drastically reduced O-GlcNAcylation. Both of these would be enriched less than a control, as enrichment is based solely on the incorporation of GalNAz into the cells’ endogenous glycans. Similarly, subcellular fractionation gives us insight into whether protein stability or O-GlcNAcylation is varied between compartments, or after a stimulus. For example, a stable glycoprotein may shuttle between the nuclear and cytoplasmic compartments after a stimulus, which would appear as no change without fractionation. Using these MS-independent methods we can determine the O-GlcNAc status of a specific protein and also determine if total protein levels are changing after treatment.

Finally, our glycoproteomics workflow is cell-type independent, as it relies on the cells’ endogenous glycosylation machinery for the metabolic labeling of O-GlcNAc. It also utilizes SILAC, a widely used and versatile quantitative proteomics strategy. Using our workflow, we have identified both novel O-GlcNAcylated substrates and stimulus-dependent changes in O-GlcNAc on specific substrates in an unbiased, quantitative manner. Future studies will focus on determining what effects O-GlcNAc has on these specific substrates. The workflow defined here will allow the
characterization of O-GlcNAc signaling across a wide variety of biological contexts and is not restricted to the systems we have defined here.
4. Identification and Characterization of a Candidate O-GlcNAc-mediated Binding Partner of Sec23A

4.1 Introduction

Like many other PTMs, O-GlcNAc exerts myriad effects on protein localization, activation, inhibition, stability, conformational changes, or degradation (38,170,171,213,253). However, the biochemical effects of O-GlcNAc on the vast majority of specific substrates remain unknown.

One way that O-GlcNAc can alter protein function is by promoting protein-protein interactions (213). O-GlcNAc-mediated protein-protein interactions are challenging to study because they are often sub-stoichiometric, transient, and/or low-affinity (175,213). To address this difficulty, the Kohler lab recently described a strategy for covalently capturing these interactions. In this method, cells are metabolically labeled with GlcNDAz, a synthetic GlcNAc precursor functionalized with a diazirine, a commonly used photocrosslinking moiety (175). However, due to the large N-acyl diazirine substituent of the synthetic precursor, cells require a mutant UDP-N-acetylhexosamine pyrophosphorylase (UAP1) with an expanded binding pocket, UAP1(F383G), in order efficiently convert the synthetic precursor sugar to UDP-GlcNDAz (175). The Kohler lab rationally designed this UAP1 mutant by examining the crystal structure and deciding to mutate the bulky phenylalanine, which is near the N-acetyl group of bound UDP-GlcNAc, to a glycine. When a cell-permeable precursor of GlcNDAz is added to the culture medium of cells expressing UAP1(F383G), it is
converted into UDP-GlcNAz (Figure 23) (175). OGT readily accepts UDP-GlcNAz and places O-GlcNAz it on endogenous substrates. Upon brief irradiation with long-wave ultraviolet light (~365 nm), diazirines eliminate molecular nitrogen, forming a highly reactive carbene, which results in the covalent crosslinking of O-GlcNAz to any macromolecule within ~2-4 Å of the sugar (175).
Figure 23: Schematic of GlcNDAz crosslinking

A protected form of GlcNDAz is fed to cells expressing mutant UAP1(F383G). The cell converts the protected precursor to UDP-GlcNDAz which is added to endogenous substrates by OGT. After brief irradiation with long-wave ultraviolet light are high reactive carbene is formed, which covalently captures any macromolecules within ~2-4 Å of the sugar.

As described previously, several COPII proteins crosslink in a GlcNDAz-specific manner, indicating that O-GlcNAc mediates interactions of COPII proteins (87). We identified specific O-GlcNAc sites on SEC23A important for proper collagen trafficking in both a human chondrosarcoma cell line and a sec23a loss-of-function zebrafish model of CLSD (87). Interestingly, these key residues are not only required for proper collagen trafficking, but also required for GlcNDAz-dependent crosslinking (87). This suggests
that O-GlcNAcylation of SEC23A may mediate interactions with specific proteins involved in regulating COPII activity. To further characterize potential functionally relevant O-GlcNAc-mediated protein-protein interactions of SEC23A, we sought to identify the relevant binding partner(s). After GlcNDAz-treating cells expressing SEC23A-myc-6xHis, we tandem affinity purified SEC23A crosslinks, and analyzed them via MS. We identified ankycorbin, a vertebrate-specific protein with no known function, as a candidate novel O-GlcNAc-mediated binding partner of SEC23A.

Ankycorbin, also known as NORPEG (novel retinal pigment epithelial cell gene), and RAI14 (retinoic acid induced gene 14) was independently characterized by two labs in the early 2000s (254,255). Ankycorbin contains six ankyrin repeats (AR) in the N-terminal region and a coiled-coil domain in the C-terminal region (255). Ankycorbin interacts with the cytoskeleton but does not bind F-actin directly (255). It also exhibits cell density-dependent localization in retinal pigment epithelial cells, demonstrating a predominantly nuclear localization in nonconfluent cells, but a cytoplasmic localization reminiscent of the cytoskeleton in confluent cultures (256). It currently has no known function, but is predominantly expressed in the retina, placenta, and testes in humans, and has been shown to regulate spermatid polarity and transport during spermiogenesis in rats (257,258). In crusher (i.e., sec23a loss-of-function) zebrafish chondrocytes, ankycorbin mRNA levels are upregulated ~2 fold over WT chondrocytes, as determined by RNA-Seq (personal communication, Ela Knapik). Recently, ankycorbin has been
implicated as a potential biomarker in lung adenocarcinoma, as it was upregulated in A549 cells and 31 of 71 lung adenocarcinoma patient samples evaluated (259).

Interestingly, higher expression correlated with better prognosis. Also, in gastric cancer cells lines, siRNA knockdown of ankycorbin decreased cell proliferation (260).

Currently, there is no described connection between the COPII or any other secretory pathway and ankycorbin.

Here we describe our efforts to purify, identify, and validate the O-GlcNAc-mediated binding partner of SEC23A, including ankycorbin.

### 4.2 Materials and Methods

**Mammalian Expression Vectors** – Details of most of the vectors used within this chapter can be found in the previous chapters. The UAP1(F383G)-myc-6xHis was generated by PCR amplifying UAP1(F383G) from pSin-EF2-UAP1(F383G) (175) with the following primers: 5’- TAATACAAGCTTATGAACATTAATGACCTCAAACTCA -3’ and 5’ – TATCTAGAAATACCATTCCCCACCAGCTCATGAAC -3’ and digesting the product with HindIII and XbaI and ligating into pcDNA4/myc-6xHis using standard methods.

**Tandem Affinity Crosslink Purification** – Cells transfected with SEC23A-myc-6xHis were lysed in immunoprecipitation (IP) lysis buffer (1% Triton X-100, 0.1% SDS, 150 mM NaCl, 1 mM EDTA, 20 mM Tris-HCl pH 7.4) supplemented with protease
inhibitor cocktail (P8340, Sigma). Lysates were probe-sonicated, cleared by centrifugation and quantified by BCA protein assay (23225, ThermoFisher) according to the manufacturer’s instructions. IPs were performed on 50 mg total protein (for MS analysis). Cleared lysates were adjusted to final total protein concentration of ~2-5 mg/mL using IP lysis buffer supplemented with protease inhibitor and 2 µg of mouse monoclonal anti-c-myc (9E10) (sc-40, Santa Cruz Biotech) per 1 mg of total protein was added and rotated overnight at 4 ºC. 50 µL washed protein A/G UltraLink Resin (53133, ThermoFisher) was added to the lysate and the mixture rotated at room temperature for 1 hour. Beads were washed three times with 1 mL of IP lysis buffer and then eluted twice in 500 µL of Ni-NTA wash buffer (8 M urea, 300 mM NaCl, 1% Triton X-100, 5 mM imidazole) with rotation. The two 500 µL elutions were pooled, and 50 µl washed 6xHisPur Ni-NTA resin (88223, ThermoFisher) was added to the eluate and rotated for 2 hours at room temperature. The Ni-NTA resin was washed three times with 1 mL of Ni-NTA wash buffer. Final elution from the Ni-NTA was performed using 8 M urea with 250 mM imidazole.

**Generation of Ankycorbin⁻/⁻ Cells** – This protocol was performed as described in Chapter 2, with the following changes. Both 293T and SW1353 cells were used. 293T cells were treated with 3 µg/mL blasticidin after LentiCas9 infection and recovery, SW1353 were treated as previously described. Two sgRNA sequences
targeting exon 3 and 4 of human RAI14 (ankycorbin) locus were designed and validated via the Surveyor assay (179) by the Duke Functional Genomics Facility:

\[
RAI14 \text{ sgRNA 1} \\
5' - GAATGGAGATGCAGGAGAAGGTGG - 3' \\
\]

\[
RAI14 \text{ sgRNA 2} \\
5' - TGCTGCTGCAAAAGGACACGTGG - 3' \\
\]

After sgRNA lentivirus infection, 293T cells were selected using 0.5 µg/mL puromycin, and SW1353 were treated as previously described.

AAVS1 infected 293T cells were dubbed Ank-ShV (control), RAI14 sgRNA 1 and 2 infected cells were dubbed Ank-2 and Ank-3 (sgRNA 1 and 2), respectively.

AAVS1 infected SW1353 cells were dubbed SWAnk-ShV (control), RAI14 sgRNA 1 and 2 infected cells were dubbed SWAnk-2 and SWAnk-3 (sgRNA 1 and 2), respectively.

**Stable Expression of UAP1 in SW1353 Ankycorbin**—SWAnk-ShV, SWAnk-2, and SWAnk-3 cells were plated at ~10% confluence and transfected with UAP1(F383G)-myc-6xHis using TransIT-LT1 (Mirus), as previously described. 48 hours after transfection, cells were selected using 200 µg/µL Zeocin (InvivoGen, ant-zn-1p) and passaged until a control, untransfected plate contained no living cells.

*Methods not explicitly described within this chapter can be found in previous chapters.*
4.3 Results

4.3.1 Identifying Candidate SEC23A O-GlcNAc-mediated Binding Partners

We previously used GlcNDAz to determine that SEC23A, SEC24B, SEC24C, SEC31A, and TFG all engage in O-GlcNAc-mediated protein-protein interactions in human cells (87). To further investigate these interactions, we focused on SEC23A because it has a robust crosslinking pattern that is maintained after epitope-tagging the C-terminus (Figures 5, 7 & 24), which is required for large-scale enrichment. SEC23A also has relevant disease and animal models which will allow further functional characterization of an O-GlcNAc-mediated binding partner (88,90,92). We initially probed immunopurified SEC23A crosslinks for candidate COPII binding proteins, including SEC24, SEC13, SEC31, and SAR1, but none of these proteins was detected in the crosslinks (data not shown).
Figure 24: SEC23A crosslinking is retained after epitope tagging

293T cells were transiently transfected with SEC23A-myc-6xHis and treated with either DMSO or GlcNDAz and were exposed to UV. Lysates were analyzed by tandem affinity purification and WB. Input samples are denoted with an N, and tandem purified sample with a P. SEC23A-myc-6xHis crosslinks in a GlcNDAz-dependent manner and the crosslinks remain after epitope tagging.

To identify candidate O-GlcNAc-mediated binding partners, we purified SEC23A GlcNDAz-crosslinks from human cells and used MS to determine the content of the SEC23A crosslinks. To generate the crosslinks, we transfected human 293T cells stably expressing UAP1(F383G) with SEC23A-myc-6xHis. Then, we treated cells with GlcNDAz for ~72 hours, followed by brief UV irradiation. Next, we tandem affinity purified SEC23A-myc-6xHis crosslinks from the lysates. Initial attempts to identify the partner via MS yielded low-quality data, which included many contaminants (e.g. keratin, myosin), but SEC23A was identified within the higher molecular weight crosslinks (data not shown).
We repeated the purification on a larger scale, increasing the amount lysate for
crosslink purification from 15 mg to 50 mg. In this larger sample we identified SEC23A
in the crosslinks specifically in the GlcNDAz-treated sample, but not in a DMSO-treated
negative control sample. To determine the candidate binding partner, we applied
various filtering criteria. We rejected as candidates common contaminating proteins,
proteins common to both GlcNDAz and DMSO samples, and proteins whose size would
exclude them from a predicted crosslink. SEC23A has a molecular weight of 85 kDa and
we would expect an interactor to be ~100 - 175 kDa based on the molecular weight of the
GlcNDAz-dependent crosslink. We also excluded proteins with spectral counts much
lower or higher than SEC23A. Using our filtering criteria, we reduced our binding
partner candidate list to a single protein, ankycorbin. Ankycorbin has a predicted
molecular weight of 110 kDa, was identified with the same number of spectral counts as
SEC23A, and is not commonly identified in the CRAPome, a repository of commonly
identified MS contaminating proteins (Figure 25A,B) (261). Using the same sample
submitted for MS analysis, we detected ankycorbin in tandem affinity-purified SEC23A
crosslinks via WB (Figure 25C).
Figure 25: Ankycorbin identified as a candidate SEC23A O-GlcNAc-mediated binding partner

(A) 29T cells were transfected with SEC23A-myc-6xHis and treated with either DMSO or GlcNDAz. SEC23A crosslinks were tandem affinity purified. Eluents were analyzed via SDS-PAGE and the gel was visualized using Coomassie stain. Gel slices obtained from the indicated red and blue boxes were sent to Duke Proteomics Facility for analysis. (B) Ankycorbin was the only candidate protein identified after applying our filtering criteria. (C) Analyzing a retained portion of the sample sent for MS via WB indicated ankycorbin was detected only in the GlcNDAz treated samples. Blot has been cropped to only show crosslinks.

4.3.2 Functional Impact of Ankycorbin on COPII Pathway Function

We tested whether ankycorbin has a functional role in COPII vesicle trafficking. Although ankycorbin has no reported role in vesicle trafficking, it has previously been shown to localize to the ER (257,258). We hypothesized that an O-GlcNAc-mediated
SEC23A-ankycorbin interaction may regulate COPII vesicle trafficking. To test this, we employed commonly used COPII functional assays. First, we overexpressed ankycorbin in human 293T cells and found that it significantly enhanced COPII-dependent ssHRP secretion (Figure 26A) (192). However, a reciprocal experiment revealed no change in ssHRP secretion upon ankycorbin deletion (Figure 26B) (see below for deletion).

Next, we examined whether ankycorbin knockout affected ER morphology or collagen secretion. Using the ankycorbin⁻/⁻ SW1353 cells described below, we observed no detectable differences in collagen retention or ER morphology between SW1353 cells expressing the “safe harbor” sgRNA (control) or two independent sgRNAs targeting ankycorbin (sgRNA 1 and sgRNA 2) (Figure 26C, also compare to Figure 6B).
Figure 26: Functional impact of ankycorbin on secretion

(A) 293T cells were transfected with ssHRP and either GFP (control) or ankycorbin. The amount of ssHRP in the medium was measured after 20 hours via luminescence assay. Ankycorbin significantly enhanced ssHRP secretion. Ankycorbin overexpression significantly enhanced ssHRP secretion. n=3, Student’s t-test. Error bars are standard error of the mean. Ankycorbin was deleted from 293T cells (as described below). Ankycorbin deletion had no effect on COPII-dependent ssHRP secretion. n=2. Error bars are standard error of the mean.  

(C) No difference was detected in collagen retention (QED, green) or ER morphology (PDI, red) after ascorbate treatment in SW1353 control cells or SW1353 expressing Ankycorbin sgRNA 1 or 2.
4.2.3 Testing a Potential SEC23A-ankycorbin O-GlcNAc-mediated Interaction in 293T Cells

First, to validate the putative SEC23A-ankycorbin O-GlcNAc-mediated interaction, we performed IPs of ankycorbin from lysates of cells treated with Thiamet-G or 5SGlcNAc. We hypothesized that if the interaction was O-GlcNAc-mediated, the amount of SEC23A that co-IPs with ankycorbin would change with inhibition of OGT or OGA. Under both standard IP (1% Triton X-100, 150 mM NaCl, 1 mM EDTA, 20 mM TRIS, pH 7.4) and mild IP conditions (0.5% NP-40, 150 mM NaCl, 1 mM EDTA, 20 mM TRIS, pH 7.4), SEC23A did not co-IPed with endogenous ankycorbin (Figure 27A). We also performed the reciprocal experiment with overexpressed SEC23A-myc-6xHis myc IP but ankycorbin co-IP was not detected (data not shown).

Because attempts to validate the interaction using the endogenous glycan failed, we employed GlcNDAz in an effort to capture the endogenous SEC23A-ankycorbin interaction. We chose to IP ankycorbin because endogenous the SEC23A IP was less efficient (data not shown). We performed an endogenous ankycorbin IP coupled with GlcNDAz and either Thiamet-G or 5SGlcNAc treatment, but we were unable to detect any SEC23A-ankycorbin crosslinks in an ankycorbin IP (Figure 27B).

Because the SEC23A crosslinking is robust and consistent, we hypothesized that reduction of ankycorbin via shRNA would also reduce endogenous SEC23A crosslinking. However, our shRNA did not reduce ankycorbin levels and therefore
could not be used to assess whether it was SEC23A’s candidate binding partner (data not shown).

Figure 27: SEC23A-ankycorbin interaction not detected by IP/WB in 293T cells

(A) 293T cells were treated with DMSO, Thiamet-G or 5S-GlcNAc as indicated and lysates were analyzed by ankycorbin IP/WB. SEC23A did not co-IP with ankycorbin under any treatment conditions. RL2 was included as an indicator of O-GlcNAc levels after treatment. These representative IP/WB were washed with the more stringent conditions described. (B) 293T cells were treated with GlcNDAz, as indicated, and UV irradiated to covalently capture the O-GlcNAc-mediated interaction. Cells were also treated with either DMSO, Thiamet-G or 5S-GlcNAc to modulate O-GlcNAc levels. Lysates were analyzed by ankycorbin IP/WB. Endogenous SEC23A crosslinking was detected in lanes 2-4. Neither uncrosslinked nor crosslinked SEC23A co-IPed with ankycorbin.
Next, to test the requirement for ankycorbin in SEC23A crosslinking, we created ankycorbin\textsuperscript{-} cells using CRISPR/Cas9 methods. Using two independent sgRNAs targeting ankycorbin, we created clonal 293T cell lines with varying levels of ankycorbin reduction (Figure 28A). Using our ankycorbin\textsuperscript{-} 293T cells in concert with GlcNDAz crosslinking, we determined that SEC23A crosslinking was not reduced with ankycorbin knockout (Figure 28B).

Figure 28: Ankycorbin is not required for SEC23A crosslinking in 293T cells

(A) Single-cell clones derived from 293T cells stably expressing Cas9 and an sgRNA targeting a “safe harbor” locus (control) or the Ankycorbin locus (sgRNA 1 and sgRNA 2) were analyzed by WB. Clones with the lowest ankycorbin levels were used for experiments requiring ankycorbin\textsuperscript{-} cells. (B) Ankycorbin\textsuperscript{-} (sgRNA 1 and sgRNA 2) and control cells (control and parental) were transfected with UAP1(F383G) and treated with DMSO or GlcNDAz, and UV irradiated. Lysates were analyzed by WB. There was no difference in SEC23A crosslinking in ankycorbin\textsuperscript{-} cells when compared to 293T parental (no Cas9 or sgRNA) or control cells. SEC23A blot is cropped to only show crosslinks.
4.2.4 Testing a SEC23A-ankycorbin O-GlcNAc-mediated Interaction in SW1353

Because of the altered levels of ankycorbin mRNA in *crusher* zebrafish chondrocytes, the overlap of residues important for both SEC23A O-GlcNAc-mediated interactions and collagen trafficking in chondrocytes, and the essential function of SEC23A in chondrocytes, we tested for SEC23A-ankycorbin GlcNDAz crosslinking in a human chondrosarcoma cell line, SW1353, as a relevant model system for SEC23A function (87,88,90,92).

Using a CRISPR/Cas9 system, we deleted ankycorbin from SW1353 cells and verified reduction via WB (Figure 29A). We did not derive a clonal population due to the efficient elimination of ankycorbin expression in mixed cell populations (Figure 29A).

Next, we transiently transfected ankycorbin<sup>−</sup>-SW1353 with UAP1(F383G) to allow incorporation of GlcNDAz into the cells’ endogenous glycans. However, we did not observe a consistent reduction in SEC23A crosslinking in this system (Figure 29A). One drawback of using SW1353 is their low transfection efficiency, which generated inconsistent SEC23A crosslinking results when using UAP1(F383G) transient transfection. To solve the problem of transient transfection, we created ankycorbin<sup>−</sup>-SW1353 stably expressing UAP1(F383G), which allowed for more robust and consistent crosslinking, as seen with the control protein nucleoporin 62 (P62). However, we again did not consistently observe a reduction in SEC23A crosslinking in ankycorbin<sup>−</sup> cells (Figure 29B).
Figure 29: SEC23A crosslinking in SW1353 cells

(A) SW1353 cells stably expressing Cas9 and an sgRNA targeting a “safe harbor” locus (control) or the Ankycorbin locus (sgRNA 1 and sgRNA 2) were transfected with UAP1(F383G), treated with GlcNDAz (or DMSO), and UV irradiated. Lysates were analyzed by WB. SEC23A crosslinking is reduced with sgRNA 2, but not with control or sgRNA 1. However, this result has yet to be replicated. P62 is a crosslinking control. (B) Control, sgRNA 1, and sgRNA2 SW1353 cells stably expressing UAP1(F383G) were treated with DMSO or GlcNDAz, and UV irradiated. SEC23A crosslinking (denoted by *crosslinked) appears reduced with sgRNA 1 and 2. Representative blots shown for both (A) and (B) as SEC23A crosslinking is inconsistent in SW1353 cells.
4.4 Discussion

SEC23A is required for proper COPII pathway function in chondrocytes and is dysregulated in human disease (87,88,90,92). While the basics of COPII coat interactions and carrier assembly are well understood, relatively little is known about its regulation in response to growth and proliferative signaling, metabolic demands, or stress. Our results indicate that SEC23A O-GlcNAc-mediated protein-protein interactions may regulate SEC23A.

We identified a candidate SEC23A O-GlcNAc-mediated interacting partner, ankycorbin, via MS and detected the presence of ankycorbin in SEC23A GlcNDAz-crosslinks after tandem affinity purification (Figure 25). However, our attempts to confirm the interaction using the endogenous proteins and glycans were unsuccessful (Figure 27A). This may be due to challenges with verifying endogenous glycoprotein interactions, such as the detection limits of WB, and typically weak carbohydrate-protein interactions. Also, the ratio of crosslinked to uncrosslinked SEC23A suggests an interaction with low stoichiometry. These challenges made characterizing the interaction with this approach difficult (Figure 27B).

Because we did not detect a native O-GlcNAc-mediated interaction between endogenous SEC23A and ankycorbin, we next employed GlcNDAz-crosslinking coupled with CRISPR/Cas9 deletion of ankycorbin. In 293T cells, this method suggested that ankycorbin was not required for SEC23A-crosslinking (Figure 28B). Since GlcNDAz
crosslinking is mediated by a PTM, which may be dynamic and sub-stoichiometric, it is possible that only a small fraction of SEC23A and/or ankycorbin is required for this interaction. We verified the reduction of ankycorbin via WB, but because the ratio of uncrosslinked to crosslinked SEC23A observed via WB is high, a small amount of remaining ankycorbin could hypothetically fully support SEC23A-crosslinking. Also, we observed similar results in SW1353 cells, but our GlcNDAz-crosslinking was never as robust as in 293T (Figure 29). Another challenge associated with ankycorbin is the availability of reagents. Ankycorbin is a vertebrate-specific protein with no current known function and few reports in the literature. This modest background may explain the lack of quality commercial antibodies and reagents for the study of ankycorbin.

We have previously demonstrated that point mutations of S115, S116, T137, T168, and S184 in SEC23A disrupt GlcNDAz-dependent crosslinking (87). However, as seen in Chapter 2, Figure 3B (and not shown), the sites are distributed across the 3D structure of the protein. S184 is located directly at the interface with SEC24, whereas both T137 and T168 are close, but not directly at the interface. S115 and S116 are distal to the interface. In all cases mutating these serine and threonine residues to an unglycosylatable alanine reduces or abrogates GlcNDAz-mediated crosslinking. One possible explanation for this distribution of residues disrupting GlcNDAz crosslinking is that most sites are indirectly involved with regulating the O-GlcNAc-mediated interaction, but O-GlcNDAz modification at only a single key residue is important for directly interacting with and
covalently capturing the binding partner. PTM modification at any given serine or threonine important for GlcNDAz-dependent crosslinking may exert a conformational change, induce binding of important cofactor proteins, or enhance O-GlcNAcylation at distal sites. In this case, these sites would not directly participate in the interaction but would still be required for crosslinking.

Interestingly, ankycorbin contains six ankyrin repeats (255). AR are a 33-residue motif that contains two alpha-helices arranged in an anti-parallel fashion (262). AR are a common protein-protein interaction domain and do not rely on specific sequence identity or motifs for binding (262). They instead rely on discontinuous patches of residues dispersed in both the target protein and the AR containing protein (262). If ankycorbin is the binding partner, this may explain the decrease in SEC23A crosslinking observed across the various, spatially distant, mutants.

Unfortunately, our efforts to confirm ankycorbin as an O-GlcNAc-mediated binding partner for SEC23A remain inconclusive. It is possible that ankycorbin was erroneously identified in our MS data set. However, because of the stringent filtering criteria and inclusion of a DMSO-treated negative control, this is less likely. Before further experimentation is done to attempt to validate ankycorbin as the binding partner, the SEC23A GlcNDAz crosslink purification and MS analysis will be repeated. If we identify ankycorbin in the purified crosslinks again, we will reevaluate our approaches to characterizing this interaction. Generating our own ankycorbin and/or
SEC23A antibodies for IP, WB, or IF would be costly but justified if ankycorbin is the binding partner. However, if ankycorbin is not identified in a repeated SEC23A crosslink purification and analysis, we will refocus our efforts on any newly identified candidates. Ankycorbin may be the O-GlcNAc-mediated binding partner of SEC23A, but further experimentation is needed to test this hypothesis.
5. Conclusions

The results presented here contribute to our knowledge of O-GlcNAc and the effects it has on modulating COPII secretory pathway function. We also developed a glycoproteomics workflow to identify and study O-GlcNAc-dependent signaling changes in a wide variety of biological contexts and identified BFA-dependent changes in a COPI coatomer protein, COPG.

We demonstrated that evolutionarily conserved O-GlcNAcylation sites on SEC23A are important for collagen secretion in SW1353 cells and skeletogenesis in developing zebrafish (87). One of these modified sites, S184, lies near the SEC23/SEC24 interface (Figure 3B), potentially providing a site for a regulatory O-GlcNAc-modification. The SEC23/SEC24 interaction is required for COPII function and mutations at this interface could disrupt the inner coat complex, but we saw no change in SEC24 co-IP with the S184A mutant (Figure 9A). In contrast, the residues in the SEC23A 4A mutant are not localized exclusively at the SEC23/SEC24 interface, which may allow O-GlcNAc to play a distinct role at those sites (Figure 3B and not shown). Similar to the S184A mutant, the SEC23A 4A mutant was still competent to bind to SEC24 paralogs, as observed by IP/WB (Figure 9B). These sites are also important for SEC23A O-GlcNAc-mediated protein-protein interactions, as mutating any one of these residues to an unglycosylatable alanine disrupts the GlcNDaz-mediated crosslinking (Figure 7A). However, these sites are not all located in the same area on the SEC23A
crystal structure. There are a few possible explanations for this observation. One is that only one site is directly involved in the O-GlcNAc-mediated protein-protein interaction and the others play supporting roles, possibly by inducing conformational changes, recruiting cofactor proteins, or enhancing O-GlcNAcylation (or other PTMs) at distant sites. We also attempted to identify, validate, and characterize a candidate SEC23A O-GlcNAc-mediated binding partner, ankycorbin, but our results were inconclusive. Although we haven’t confidently identified the binding partner, we hypothesize that it may function as an inhibitor of COPII function. Three of the sites required for GlcNDAz crosslinking cluster near the SEC23/SEC24 interface, where modification by O-GlcNAc may recruit this binding partner and inhibit inner coat formation. Consistent with this hypothesis, our functional data suggest that elevated levels of O-GlcNAc after Thiamet-G treatment disrupt normal COPII pathway function and alter the membrane binding of COPII components (Figure 4C, D).

We used the GlcNDAz crosslinking results to inform our decision about which mutants to test in vivo, but it is possible that the GlcNDAz crosslinking and collagen trafficking phenotypes represent distinct functional roles of O-GlcNAc on SEC23A. As seen with SEC13, which functions as both an essential COPII component and a part of the nuclear pore complex, and SEC24, which regulates and is specifically required for autophagy, COPII components can moonlight in other pathways (121-124,127-129). SEC23 may also play a role in autophagy. In yeast, temperature-sensitive sec23 mutants
are deficient in the trafficking of an ER SNARE protein, Ufe1, which has a role in autophagosome biogenesis (263). Also, phosphorylation of SEC23A by the autophagy-activating kinase ULK1 causes ER exit sites to aggregate and reduces the transport of COPII cargo proteins (264). Identifying SEC23A’s binding partner(s) will allow us to discover what role, if any, O-GlcNAc-mediated protein-protein interactions play in COPII pathway function and/or other processes. We can use our mutant reconstituted SEC23A−/−SW1353 cells to determine which residues are required for the SEC23A-binding partner interaction, and knockdown/knockout in human cells or zebrafish to determine what function the binding partner has on the COPII pathway.

In addition to the sites required for SEC23A GlcNDAz crosslinking, we identified nine other unambiguously localized O-GlcNAc sites (Figure 3A). We also identified three and four unambiguously localized sites on SEC24C and SEC31A, respectively (Figure 3A). Currently there is no known role for O-GlcNAc on SEC24C or SEC31A but using COPII component- and paralog-specific cargoes in combination with unglycosylatable mutants we can determine whether O-GlcNAc at specific sites is required for the trafficking of these specific cargoes. For example, Ebola virus matrix protein VP40 and the serotonin transporter are SEC24C specific cargoes and collagen trafficking requires SEC31A (116,118,153). It is possible O-GlcNAcylation of these proteins plays no role in canonical COPII pathway function. Recently, phosphorylation of SEC24C has been shown to increase binding with the C-terminus of ATG9 to increase
autophagosome number during nutrient deprivation (127-129). This supports the idea that PTMs on these proteins can enhance moonlighting functions, as well as participate in COPII regulation.

SEC24 is phosphorylated during mitosis, which prevents membrane binding and may contribute to the unknown mechanism by which COPII trafficking is suspended during this phase of the cell cycle. This modification is replaced by O-GlcNAcylation during interphase (135). Extensive cross talk between O-GlcNAcylation and phosphorylation has been reported, frequently competing for nearby or identical residues on a given protein and exerting myriad effects (169). Interestingly, two of the three O-GlcNAc sites localized on SEC24C are reported as phosphorylated, suggesting the interplay of these PTMs could tune COPII activity (205). None of the sites on SEC23A have been reported to be phosphorylated, but modifications at these sites may yet be discovered. Using phosphomimetic mutants we could assess whether phosphorylation at any of these sites on SEC23A (or other COPII components) has any effect on COPII pathway function. Unfortunately, an in vivo O-GlcNAc mimetic is unavailable, so the reciprocal experiment is not feasible.

We also demonstrated that TFG is O-GlcNAcylated via WB and MS (Figure 4A and data not shown), that it crosslinks in a GlcNDAz-dependent manner (Figure 5A), and that the C-terminus is required for these interactions (data not shown). However, the individual and compound unglycosylatable mutants in the C-terminus that we
tested do not disrupt the GlcNDAz-dependent crosslinking of TFG (Figure 11). Because TFG engages in homotypic interactions, endogenous TFG could be interacting with our glycosylation deficient mutants in these experiments, supporting the GlcNDAz crosslinking (208). Generating TFG−/− cells may be necessary to determine which residues are important for O-GlcNAc-mediated interactions and to determine their function.

We developed a global quantitative glycoproteomics workflow to identify changes in O-GlcNAc in response to a stimulus. Using this technique, we identified a decrease in COPG, a COPI coatomer protein, in response to BFA treatment. Using GalNAz coupled with biotin-click and affinity purification, we demonstrated that these changes were reproduced using an MS-independent method (Figure 20). We were also able to detect the native glycan on endogenous COPG using an IP/WB approach (Figure 21). Currently, there are no published reports of O-GlcNAcylation of any COPI components. However, phosphorylation of COPI subunits, COPB and COPG, has been previously reported as a possible method of coatomer assembly, membrane recruitment, or altering the specificity of coatomer-organelle interactions (265). O-GlcNAc on COPG may function similarly to phosphorylation, which will be assessed in future studies.

As demonstrated by our data, O-GlcNAc modifies several COPII components and changes in O-GlcNAc levels alter normal COPII pathway functions. O-GlcNAc also modifies COPG, a component of the heteroheptameric COPI coat complex. Post-translational modification of both COPI and COPII components has been shown to alter
membrane binding, cargo selectivity, and vesicle biogenesis (87,130,135,164,264,265).

Biosynthesis of UDP-GlcNAc, the nucleotide sugar donor for OGT, integrates glucose, nitrogen, fatty acid, and nucleic acid metabolic pathways, linking nutrient sensing to signaling (38,266). Also, O-GlcNAc and OGT levels are elevated during autophagy, suggesting that O-GlcNAcylation could provide nutrient-sensitive modulation of COPII pathway activity and may support non-canonical functions of COPII components (267-270). Overall, this suggests O-GlcNAcylation may be a broadly conserved regulatory modification of the vertebrate secretory pathway.
Appendix A

Chapter 2 Mass Spec Information and Data

Complete mass spec data for Chapter 2 can be found online with the associated publication.


Chapter 3 Mass Spec Information and Data

Data from both BFA-treated and IL-3-stimulated SILAC data sets was processed using Rosetta Elucidator and is available as supplemental data.
References


8. Olivier-Van Stichelen S, Abramowitz LK, Hanover JA. X marks the spot: does it matter that O-GlcNAc transferase is an X-linked gene? Biochem Biophys Res Commun 2014;453:201-7


38. Hanover JA, Krause MW, Love DC. The hexosamine signaling pathway: O-GlcNAc cycling in feast or famine. Biochimica et biophysica acta 2010;1800:80-95


50. Wells L, Kreppel LK, Comer FI, Wadzinski BE, Hart GW. O-GlcNAc transferase is in a functional complex with protein phosphatase 1 catalytic subunits. The Journal of Biological Chemistry 2004;279:38466-70


54. Iyer SP, Akimoto Y, Hart GW. Identification and cloning of a novel family of coiled-coil domain proteins that interact with O-GlcNAc transferase. The Journal of biological chemistry 2003;278:5399-409

55. Iyer SP, Hart GW. Roles of the tetratricopeptide repeat domain in O-GlcNAc transferase targeting and protein substrate specificity. The Journal of biological chemistry 2003;278:24608-16

56. Brickley K, Pozo K, Stephenson FA. N-acetylglucosamine transferase is an integral component of a kinesin-directed mitochondrial trafficking complex. Biochimica et biophysica acta 2011;1813:269-81


Tang BL, Wang Y, Ong YS, Hong W. COPII and exit from the endoplasmic reticulum. Biochimica et biophysica acta 2005;1744:293-303


Yoshihisa T, Barlowe C, Schekman R. Requirement for a GTPase-activating protein in vesicle budding from the endoplasmic reticulum. Science 1993;259:1466-8


Miller E, Antonny B, Hamamoto S, Schekman R. Cargo selection into COPII vesicles is driven by the Sec24p subunit. The EMBO Journal 2002;21:6105-13


76. Salama NR, Yeung T, Schekman RW. The Sec13p complex and reconstitution of vesicle budding from the ER with purified cytosolic proteins. EMBO J 1993;12:4073-82


81. Kuge O. Sar1 promotes vesicle budding from the endoplasmic reticulum but not Golgi compartments. The Journal of cell biology 1994;125:51-65


117. Farhan H, Reiterer V, Korkhov VM, Schmid JA, Freissmuth M, Sitte HH. Concentrative export from the endoplasmic reticulum of the gamma-
aminobutyric acid transporter 1 requires binding to SEC24D. The Journal of biological chemistry 2007;282:7679-89


119. Salama NR, Chuang JS, Schekman RW. Sec31 encodes an essential component of the COPII coat required for transport vesicle budding from the endoplasmic reticulum. Mol Biol Cell 1997;8:205-17


122. Enninga J, Levay A, Fontoura BMA. Sec13 shuttles between the nucleus and the cytoplasm and stably interacts with Nup96 at the nuclear pore complex. Molecular and Cellular Biology 2003;23:7271-84


171. Hart GW. Three Decades of Research on O-GlcNAcylation - A Major Nutrient Sensor That Regulates Signaling, Transcription and Cellular Metabolism. Front Endocrinol (Lausanne) 2014;5:183


195. Reid DW, Chen Q, Tay AS, Shenolikar S, Nicchitta CV. The unfolded protein response triggers selective mRNA release from the endoplasmic reticulum. Cell 2014;158:1362-74


214. Myers SA, Daou S, Affar el B, Burlingame A. Electron transfer dissociation (ETD): the mass spectrometric breakthrough essential for O-GlcNAc protein site


221. Cox NJ, Meister TR, Boyce M. Chemical Biology of Glycoproteins. rsc 2017:94-149


232. Chen P-H, Chi J-T, Boyce M. KEAP1 has a sweet spot: A new connection between intracellular glycosylation and redox stress signaling in cancer cells. Molecular & Cellular Oncology 2017;4


236. Rathmell JC, Vander Heiden MG, Harris MH, Frauwirth KA, Thompson CB. In the absence of extrinsic signals, nutrient utilization by lymphocytes is insufficient to maintain either cell size or viability. Molecular Cell 2000;6:683-92


257. Qian X, Mruk DD, Cheng Y-h, Cheng YC. RAI14 (retinoic acid induced protein 14) is an F-actin regulator. Spermatogenesis 2014;3

258. Qian X, Mruk DD, Cheng YC. Rai14 (Retinoic Acid Induced Protein 14) Is Involved in Regulating F-Actin Dynamics at the Ectoplasmic Specialization in the Rat Testis*. PloS one 2013;8


262. Li J, Mahajan A, Tsai M-D. Ankyrin Repeat: A Unique Motif Mediating Protein–Protein Interactions*. Biochemistry 2006;45:15168-78


Biography

Nathan James Cox
Born January 27, 1988 in Greenville, NC

EDUCATION
PhD in Pharmacology
Duke University, Durham, North Carolina, May 2018

BS in Chemistry (focus in Biochemistry)
Minor in Japanese
University of North Carolina at Chapel Hill, May 2010

PUBLICATIONS


Cox NJ, Meister TR, Boyce M. Chemical Biology of Glycoproteins. rsc 2017:94-149


AWARDS

Pharmacological Sciences Training Program (PSTP)
Duke University, 2012 – 2014