POMGNT1:

THE GLYCAN STRUCTURES THAT REQUIRE IT AND THOSE WHICH ESCAPE IT

by

SALLY RIEWE BOYD

(Under the Direction of Lance Wells)

ABSTRACT

The enzyme POMGNT1 catalyzes the transfer of N-acetylglucosamine in a 2-linkage to O-linked mannose on polypeptides in the cis-Golgi apparatus to generate the core M1 glycan. Recently several proteins have been shown to contain O-Mannose glycans that exist as the simple monosaccharide, core M0, without extension. Given the established promiscuous nature of POMGNT1, we sought to explain this surprising finding of M0 structures on the cadherin family. Our work demonstrates that cadherins that carry non-elaborated M0 structures are resistant to POMGNT1 extension at the protein level, unlike alpha-dystroglycan expressed in a mutant background so as to contain M0 glycans. However, upon reduction, alkylation, and tryptic digestion, the M0-containing glycopeptides from both cadherin family proteins as well as alpha- dystroglycan can be modified by POMGNT1. Given that protein folding occurs in the endoplasmic reticulum of the secretory pathway and POMGNT1 is localized to the cis-Golgi, our experiments suggest that the M0 structures on folded cadherins, unlike alpha-dystroglycan, are not accessible for elaboration by POMGNT1 due to structural constraints. Loss of POMGNT1, responsible for the generation of the core M1 glycan, is seen in Muscle-Eye- Brain (MEB) Disease, one of the secondary dystroglycanopathies resulting from loss of

functional glycosylation of alpha-dystroglycan. However, our laboratory and others have determined the structure of the fully elaborated functional M3 glycan on alpha-dystroglycan and there is no apparent POMGNT1-dependent linkage required. Thus, we wanted to investigate the role of POMGNT1 in the generation of the functional M3 glycan that is defective in multiple forms of congenital muscular dystrophy. Several hypotheses were tested with regard to the role of POMGNT1 in functionally glycosylating alpha-dystroglycan. We found that only full-length, catalytically active POMGNT1 could rescue an MEB cell line, demonstrating that enzyme activity is required. Furthermore, while fully-elaborated M3 glycans are not observed on alpha- dystroglycan expressed in MEB cells, we were able to determine the presence of phosphoribitol- extended phosphotrisaccharide M3 glycans on alpha-dystroglycan. Thus, loss of active, full-length POMGNT1 inhibits late steps in the synthesis of the functional M3 glycan. Therefore, our research has shed light on how M0 structures escape POMGNT1 and how the functional M3 structures require POMGNT1 activity.

INDEX WORDS: POMGnT1/ O-glycosylation/ Mannose/ Cadherins/ alpha-Dystroglycan/ Muscle-Eye-Brain Disease

POMGNT1:

THE GLYCO-PROTEINS IT MODIFIES AND THOSE WHICH ESCAPE IT

by

SALLY BOYD

BS, University of Florida, 2006

MS, University of Florida, 2009

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

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA
2019

POMGNT1:

THE GLYCAN STRUCTURES THAT REQUIRE IT AND THOSE WHICH ESCAPE IT

by

SALLY RIEWE BOYD

Major Professor:	Lance Wells
Committee:	Michael Tiemeyer
	Carl Bergmann
	Daniel Mead

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
December 2019

ACKNOWLEDGEMENTS

I would like to acknowledge the many people who have helped me in my education and early career. This includes my parents who have coerced, supported, and enabled my education from an early age through college. My sister, Karen, for always being a little sister and attending my oral defense. One-year- old George for being so energetic and for providing edits to this dissertation in permanent marker. Jake for being a buddy, a roommate, and a support system. Bret for taking care of so many things that allow me to work. Your support and belief in my abilities is so important to me.

I'd like to thank the University of Georgia and the Integrated Life Sciences program for accepting me. I would like to thank Lance Wells and the Biochemistry department for welcoming me and giving me a chance. I hope I have lived up to your expectations. I would like to thank my committee members; Danny Mead, Michael Tiemeyer, and Carl Bergmann for offering their time, energy, and knowledge to enhance my educational experience. I would like to acknowledge the contribution members of the Wells' lab have made to my education especially Sneha and Trevor for being desk-mates and Osman for our scientific discussions.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

INTRODUCTION

O-glycosylation in mammals

O-linked glycosylation refers to the covalent linkage of a sugar to the hydroxyl oxygen of a serine or threonine residue. There are six sugars known to occupy the reducing terminus of O-glycans. They are fucose, xylose, glucose, N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), and mannose. O-linked fucose occurs on Epidermal Growth Factor-like (EGF) repeats and Thrombospondin Type 1 Repeats (TSR). Protein O-fucosyltransferase 1 (POFUT1) adds fucose only to folded EGF repeats (Wang, Y., Lee, G.F., et al. 1996). POFUT2 adds fucose to TSR (Luo, Y., Nita-Lazar, A., et al. 2006). The specificity of these glycosyltransferases is due to differences in their binding pockets which complement differences in the three-dimensional shape of their substrates (Holdener, B.C. and Haltiwanger, R.S. 2019, Li, Z., Han, K., et al. 2017, Valero-Gonzalez, J., Leonhard-Melief, C., et al. 2016). The O-fucose modifications function to stabilize structure and accelerate protein folding in the endoplasmic reticulum (ER) (Takeuchi, H., Yu, H., et al. 2017, Vasudevan, D., Takeuchi, H., et al. 2015). In the Golgi, O-fucose on EGF repeats can be further elaborated with GlcNAc, galactose, and sialic acid, by the enzymes Fringe, B4GALT1, and ST6GAL1, respectively (Moloney, D.J., Panin, V.M., et al. 2000). Mammals have three homologs of Fringe: Lunatic, Manic, and Radical. All three function to transfer GlcNAc to O-fucosylated EGF-repeats on Notch. Modification by Fringe modulates the binding of the Notch receptor with its Delta- like ligands 1, 3, and 4 and JAG ligands 1 and 2 (Delta and Serrate in *Drosophila*) (Luca, V.C.,

Jude, K.M., et al. 2015, Luca, V.C., Kim, B.C., et al. 2017, Panin, V.M., Papayannopoulos, V., et al. 1997). Lunatic and Manic Fringe enhance activation of Notch by DLL while inhibiting activation of Notch by JAG. Radical Fringe enhances signaling from both DLL and JAG ligands (Kakuda, S. and Haltiwanger, R.S. 2017, Rampal, R., Li, A.S.Y., et al. 2005).

In addition to O-fucose, EGF repeats are modified by O-glucose and O-GlcNAc (Harvey, B.M. and Haltiwanger, R.S. 2018). The O-glucose modification was first seen as a trisaccharide elongated with two xylose residues on the EGF repeats of blood coagulation factors VII and IX (Hase, S., Kawabata, S., et al. 1988). The O-GlcNAc modification was first reported on the extracellular domain of Notch receptors in 2008 (Matsuura, A., Ito, M., et al. 2008). This non-canonical O-GlcNAc on a protein destined for the cell surface is added by extracellular EGF-domain specific O-GlcNAc Transferase (EOGT) (Sakaidani, Y., Ichiyanagi, N., et al. 2012).

The O-GlcNAc modification is usually non-extended and occurs on nuclear and cytoplasmic proteins. Its discovery in the 1980s was the first report of glycosylation occurring inside a cell outside of the secretory pathway (Holt, G.D. and Hart, G.W. 1986). O-GlcNAc is added by O-GlcNActransferase (OGT) and removed by O-GlcNAcase (OGA) (Dong, D.L. and Hart, G.W. 1994, Haltiwanger, R.S., Holt, G.D., et al. 1990). The addition/removal cycling of O-GlcNAc functions as a nutrient sensor (Hart, G.W. 2019). There is extensive crosstalk between kinase/phosphatase phosphorylation cycling and OGT/OGA O-GlcNAc cycling with OGT existing in-complex with serine/threonine phosphatases (Wells, L., Kreppel, L.K., et al. 2004). Three decades of study has revealed O-GlcNAc cycling to regulate signaling, transcription, mitochondrial activity, and cytoskeletal functions (Hart, G.W. 2019). Recently, a similar system that utilizes O-linked mannose to modify nuclear and cytoplasmic proteins has been reported in yeast (Halim, A., Larsen, I.S., et al. 2015).

Another O-linked sugar, O-GalNAc, is added to nascent polypeptides in the Golgi by nearly 20 GalNAc transferases (GalNAc-Ts) (Bennett, E.P., Mandel, U., et al. 2012). The large number of homologous enzymes responsible for the initiation of O-GalNAc allows for great variation in differential regulation of mucin-type O-glycans in different cell types. The O-GalNAc, or Tn antigen, is extended by C1GalT1 to become the Core 1 T-antigen. Expression of functional C1GALT1 requires its private chaperone COSMC (Ju, T. and Cummings, R.D. 2002). Therefore, loss of COSMC results in loss of extended O-GalNAc glycans and concomitant expression of the Tn antigen (Wang, Y., Ju, T., et al. 2010).

O-mannosylation

O-mannosylation was originally discovered in yeast (Sentandreu, R. and Northcote, D.H. 1968). In mammals, O-mannosylation was discovered in 1979 in rat brain (Finne, J., Krusius, T., et al. 1979). Nearly one-third of O-glycans in the brain are O-mannose glycans (Chai, W., Yuen, C.T., et al. 1999). In mammalian O-mannosylation, nascent proteins are O-mannosylated in the endoplasmic reticulum (ER). The mannose donor is dolichol monophosphate-activated mannose (DPM) (Lommel, M. and Strahl, S. 2009). Defects in the DPM synthase complex (DPM1, DPM2, and DPM3) affect glycophosphatidylinositol anchor biosynthesis, N-glycan processing, and O- and C- mannosylation because DPM is the source of mannose for all known post-translational mannosylation. In the classical O-mannosylation pathway transferring of mannose from DPM to a serine or threonine of a nascent protein is done by the protein O-mannosyltransferases POMT1/POMT2 (Jurado, L.A., Coloma, A., et al. 1999, Manya, H., Chiba, A., et al. 2004).

It has recently been discovered that POMT1 and POMT2 are not necessary for the O-mannosylation of cadherins, plexins, and the IPT domain-containing protein HGFR and RON

(Larsen, I.S.B., Narimatsu, Y., et al. 2017b). The Ig-like/plexins/transcription factors domain (IPT domain) is found in the cell surface tyrosine kinase receptors *recepteur d'origine nantais* (RON) and hepatocyte growth factor receptor (HGFR) (Collesi, C., Santoro, M.M., et al. 1996). IPT domains are also found in intracellular transcription factors where they are involved in DNA binding (Bork, P., Doerks, T., et al. 1999). Plexins are cells surface proteins which function as receptors for semaphorins (Bussolino, F., Valdembri, D., et al. 2006). The plexin-binding semaphorins may be membrane-bound, GPI-linked, or viral-encdoded (Schmidt, E.F., Togashi, H., et al. 2010). As such, the plexin VESPR/plexinC1 acts as a binding site for virus-encoded semaphorins (Winberg, M.L., Noordermeer, J.N., et al. 1998). Plexins and HGFR are O-mannosylated in a POMT1/POMT2 independent manner, by an as-yet unknown mechanism (Larsen, I.S.B., Narimatsu, Y., et al. 2019).

The POMT1/2-independent O-mannosylation pathways lead to proteins with single, non-elongated O-mannose residues, referred to as the core M0 structure. In the classical POMT1/2-dependent O-mannosylation pathway, mannose is extended into a core M1, M2, or M3 structure (Praissman, J.L. and Wells, L. 2014). How the core M0 structure escapes elongation into the core M1 structure is not known. Work in Chapter 2 explores the possibility that protein folding prevents elongation of the O-mannose. Work in chapter two shows that POMGnT1 will modify the core M0 structure on tryptic peptides, but not on undigested recombinant cadherin proteins.

Core M1

Core M1 structures are formed when protein O-linked-mannose beta-1,2-N-acetylglucosaminyltransferase 1 (POMGnT1) catalyzes the transfer of GlcNAc from uridine

5'- diphosphate (UDP)-GlcNAc to the O-mannose in a β1-2 linkage. B3GAT1 then adds galactose in a β1,3-linkage to the GlcNAc of the disaccharide. The resulting trisaccharide may then be modified to several mature structures. It may be modified with sialic acid to become the classical tetrasaccharide, or with α1,3-linked fucose to form the Lewis x structure, or with sulfated glucuronic acid to form the HNK-1 epitope. The function of these core M1 structures is not known. Furthermore, our knowledge of the glycosylation frequency, patterns, and even identity of the proteins they modify is incomplete.

Proteins which have been identified as modified by POMGnT1 include CD24, RPTPζ, and alpha-dystroglycan. The most well-studied core M1-containing glycoprotein is alpha-dystroglycan. However, when comparing the O-glycans released from the brains of alpha-dystroglycan-lacking mice to their wild-type littermates, Stalnaker et al. found similar amounts of O-mannose initiated glycan structures present. It is therefore assumed that other mannose-initiated glycoproteins must be present in the brain (Stalnaker, S.H., Aoki, K., et al. 2011). The cell adhesion molecule CD24 has been shown to be modified by a diverse array of O-mannose linked glycans, including some containing 3-linked sialic acid, fucose, and the HNK-1 epitope (Bleckmann, C., Geyer, H., et al. 2009).

The products of the *ptprz1* gene, trans-membrane receptor type protein tyrosine phosphatase (RPTPζ) and its secreted splice variant, phosphacan, have also been shown to be O-mannosylated (Dwyer, C.A., Baker, E., et al. 2012). Alternative splicing of the *ptprz1* gene yields either the secreted phosphacan, the trans-membrane RPTPζ, or one of two shortened proteins; a short, secreted variant or a short receptor form. Phosphacan is a secreted proteoglycan found in the neural extracellular matrix (Maurel, P., Rauch, U., et al. 1994). RPTPζ is a trans-membrane receptor of the R5 family of receptor type protein

tyrosine phosphatases. The N-terminal domain contains a hydrophobic binding pocket with sequence homology to a carbonic anhydrase domain, but is not functional as an anhydrase as it lacks an active-site histidine residue (Peles, E., Nativ, M., et al. 1995). The O-mannose glycans on RPTPζ/phosphacan include the classical tetrasaccharide, the fucosylated tetrasaccharide, and sulfated glycans including a terminal O-mannose-linked HNK-1 epitope glycan (Dwyer, C.A., Baker, E., et al. 2012, Dwyer, C.A., Katoh, T., et al. 2015). These glycans were found to be differentially expressed in a cell-specific manner, with glial cells expressing different glycoforms than neurons (Dwyer, C.A., Katoh, T., et al. 2015).

Core M2

Core M1 glycans are converted to core M2 by the addition of a second GlcNAc in a β1,6- linkage by GnT-Vb. This structure has only been reported in the brain and prostate, making up less than five percent of the identified O-glycans in those tissues (Lange, T., Ullrich, S., et al. 2012, Stalnaker, S.H., Aoki, K., et al. 2011).

GnT-Vb null mice showed no adverse neurological phenotypic effects (Lee, J.K., Matthews, R.T., et al. 2012). However, integrin-dependent cell adhesion and migration were shown to be affected *in vitro* due to an increased level of tyrosine phosphorylation of β-catenin caused by core M2 glycan-based inhibition of RPTPβ activity (Abbott, K.L., Matthews, R.T., et al. 2008).

Core M3

The Core M3 O-mannose structure is formed when, while still in the ER, O-mannose is extended by Protein O-mannosyl N-acetylglucosamine transferase 2 (POMGnT2) through the addition of a β1,4-linked GlcNAc. The core M3 structure has only been observed on one protein, α-DG, at two specific sites: T371/T319 and T379/381. The rarity of this structure appears to be due to the primary amino acid substrate selectivity of POMGnT2 (Halmo, S.M., Singh, D., et al. 2017). The core M3 O-mannose glycan is unique, not just for its rarity, but because it may be extended through a complex pathway containing at least seven enzymes, all of which seem to be dedicated exclusively to performing one step of the synthesis pathway of this particular glycan (Sheikh, M.O., Halmo, S.M., et al. 2017).

After formation of the core m3 structure, B3GALNT2 forms a tri-saccharide with the addition of β-1,3-GalNAc to the GlcNAc. The trisaccharide is phosphorylated when POMK adds a phosphate to the 6-position of the mannose (Yoshida-Moriguchi, T., Willer, T., et al. 2013). Initially, this phosphate was believed to be the site from which the structure was further extended because this phospho-tri-saccharide structure is what remains on α–DG after treatment with aqueous hydrofluoric acid. However, recent work has shown that a ribitol-phosphate is added to the GalNAc (Gerin, I., Ury, B., et al. 2016, Praissman, J.L., Willer, T., et al. 2016). The enzymes forming the phospho-tri-saccharide all reside in the ER. The remaining modifications occur in the Golgi apparatus, as illustrated in Figure 1-1.

CDP-ribitol is synthesized from CTP and ribitol-phosphate by isoprenoid synthase domain-containing protein (ISPD) (Gerin, I., Ury, B., et al. 2016). Fukutin and fukutin-related protein (FKRP) transfer ribitol phosphate from CDP-ribitol to the developing glycan sequentially (Gerin, I., Ury, B., et al. 2016). Next, RXYLT1 (TMEM5) adds a priming xylose

(Manya, H., Yamaguchi, Y., et al. 2016). B4GAT1 adds GlcA to the β-linked xylose (Praissman, J.L., Live, D.H., et al. 2014). Finally, like-acetylglucosaminyltransferase (LARGE) acts as a bifunctional glycosyltransferase to add a repeating Xyl-GlcA disaccharide (Inamori, K., Yoshida-Moriguchi, T., et al. 2012). The repeating Xyl-GlcA is termed matriglycan and is the binding motif for LG-domain containing proteins in the extra-cellular matrix (Yoshida-Moriguchi, T. and Campbell, K.P. 2015).

Core M0

Core M0 glycans are a single, non-elongated mannose. The first report of a non-elongated O-mannose was found on the light chain of a recombinant IgG2 expressed in Chinese hamster ovary (CHO) cell (Martinez, T., Pace, D., et al. 2007). In 2013, Winterhalter et al. reported rabbit T-cadherin as a novel O-mannosylated protein and determined that the O-mannose residue was not extended (Winterhalter, P.R., Lommel, M., et al. 2013) . The same year, Vester-Christensen et al. identified cadherins and plexins as major O- mannosylated glycoproteins, identifying 37 cadherins and 6 plexins as being O-mannosylated in a POMGnT1/COSMC knockout cell line, termed SimpleCells (Vester-Christensen, M.B., Halim, A., et al. 2013). The SimpleCells, which are made in MDAMB231, CHO, and HEK293 cell backgrounds using zinc finger nucleases or CRISPER/Cas to knockout POMGnT1 and COSMC, lack elongated O-mannose or O-GalNAc structures. These knockouts simplify the process of mapping sites of O-mannose or O-GalNAc modification by mass spectrometry.

Further work to characterize the O-mannose glycans on cadherins and plexins in POMGnT1/COSMC/POMT1/POMT2 knockout cells revealed the surprising revelation that the POMT1 and POMT2 enzymes were not necessary for O-mannosylation of cadherins

(Larsen, I.S.B., Narimatsu, Y., et al. 2017b). This report also characterized the O-mannose glycans in wild-type cells, confirming that the O- mannose glycans are not extended but were in fact core M0 non-elongated O-mannose.

Later work identified the transmembrane and tetratricopeptide repeat containing (TMTC) proteins TMTC1-4 as being necessary for O-mannosylation of the cadherin superfamily. The TMTC proteins were identified by bioinformatic searches for sequence homology to POMT1 and POMT2 after it was discovered that POMT1 and POMT2 were dispensable for the O-mannosylation of cadherins (Larsen, I.S.B., Narimatsu, Y., et al. 2017a). TMTC1-4 are similar to POMT1/2 in that they have conserved DD motifs predicted to be their catalytic site in the first lumen-facing loop of their multi-pass domains (Petrou, V.I., Herrera, C.M., et al. 2016). TMTC proteins also contain C-terminal tetratricopeptide repeats (TPR). TPR domains are present on OGT and function in substrate recognition (Hart, G.W., Slawson, C., et al. 2011). Larsen et al. have postulated that the TPR domains on TMTC proteins may account for their specificity toward cadherins (Larsen, I.S.B., Narimatsu, Y., et al. 2019).

TMTC1 and TMTC2 are localized to the ER where they are involved in calcium homeostasis (Sunryd, J.C., Cheon, B., et al. 2014). TMTC proteins are also believed to be involved in the regulation of proteasomal protein degradation in the ER as a stress response. This was shown from an experiment testing HeLa cell sensitivity to the proteasome inhibitor Bortezomib, a drug that induces ER stress via protein overload. TMTC3 silencing rendered HeLa cells more sensitive to Bortezomib (Racape, M., Duong Van Huyen, J.P., et al. 2011). Interestingly, biallelic mutations in TMTC3 had been shown to cause cobblestone lissencephaly (Jerber, J., Zaki, M.S., et al. 2016). Cobblestone lissencephaly is a symptom of Muscle-Eye-Brain disease, and other CMDs, as described below. CMDs result from a loss of

functional O-mannosylation of alpha-dystroglycan (Endo, T. 2014, Live, D., Wells, L., et al. 2013). However, alpha-dystroglycan is O-mannosylated via the classical (POMT1/POMT2) O-mannosylation pathway (Martin, P.T. 2007, Manya, H., Chiba, A., et al. 2004).

The mammalian POMT1 and POMT2 are homologs of the protein mannosyltransferases PMT1-7 in yeast, which were known prior to the discovery of O-mannosylation in mammals (Loibl, M. and Strahl, S. 2013, Lommel, M. and Strahl, S. 2009, Sentandreu, R. and Northcote, D.H. 1968, Strahl-Bolsinger, S., Immervoll, T., et al. 1993). The classical O-mannosylation pathway requires coexpression of both POMT1 and POMT2 for enzyme activity (Manya, H., Chiba, A., et al. 2004). By contrast, all four TMTC proteins are not required for O-mannosylation of cahderins. Cadherin proteins expressed in TMTC1 and TMTC3 double knockout cells were O-mannosylated on the G strand, but not the B strand of the β-strand folds on cadherins (See Cadherins, below) (Larsen, I.S.B., Narimatsu, Y., et al. 2017a).

The IPT-domain containing proteins, such as plexins and HGFR, require neither POMT1/2 nor TMTC1-4 to be O-mannosylated (Larsen, I.S.B., Narimatsu, Y., et al. 2017a). These surprising findings that the core M0 mannose residues did not require POMT1/2 and were not modified by the promiscuous POMGnT1 brought forth the question of whether the hexose modifications were indeed mannose. Work in Chapter 2 of this dissertation shows mannose on E-cadherin and N-cadherin to be susceptible to an α-mannosidase enzyme. Additionally, Chapter 2 shows that when digested into tryptic peptides the O-mannosylated peptides can act as a substrate for POMGnT1. Taken together, the glycan's susceptibility to α-mannosidase, and ability to be modified by

POMGnT1 in vitro after treatment, provide evidence that these non-elongated hexose residues are in fact core M0 O-mannose structures.

Cadherins

Cadherins are cell-surface glycoproteins that function in calcium-dependent cell-cell adhesion. Cadherins function in cell sorting and morphogenesis including embryonic cell layer separation (Halbleib, J.M. and Nelson, W.J. 2006). Cadherins are a large family of proteins including the classical type I and type II cadherins. Other cadherin types include the desmosomal cadherins which form desmosome cell-cell junctions in tissues exposed to high mechanical stress (Kowalczyk, A.P. and Green, K.J. 2013). The protocadherins are expressed in the mammalian nervous system and believed to play a role in neural patterning (Schreiner, D. and Weiner, J.A. 2010). Flamingo cadherins are seven-pass transmembrane proteins containing nine EC repeats, EGF, laminin-G like and hormone receptor-like domains (Usui, T., Shima, Y., et al. 1999).

E-cadherin, also known as CDH1, and N-cadherin, also known as CDH2, are classical type I cadherins. They are single-pass transmembrane proteins. Their cytoplasmic domain is highly conserved within the subfamily and contains binding motifs for β-catenin (armadillo in *Drosophila*) (Brasch, J., Harrison, O.J., et al. 2012). The ectodomain of classical type I cadherins consists of five extracellular cadherin (EC) domains arranged in tandem. Each EC domain is approximately 110 residues and folds into a seven stranded β-barrel consisting of a two-layer β-sheet sandwich with the β-strands running antiparallel (See Figure 1-2). As a result, the N- and C-termini of the β-fold are located on opposite sides of the EC domain allowing the EC domains to line up in tandem (Harrison, O.J., Jin, X., et al.

2011). The EC domains are connected with linker regions stabilized by three calcium ions which provide rigidity and a crescent shape (Pokutta, S., Herrenknecht, K., et al. 1994).

The classical cadherins mediate cell-cell adhesion through adherens junctions. These junctions are initiated by *trans* dimerization in which a cadherin on one cell dimerizes with a cadherin on another cell. The formation of *trans* adhesive homodimers in Type I classical cadherins occurs through exchange of the N-terminal β-strands on their EC1 domains (Vendome, J., Felsovalyi, K., et al. 2014). That is, the beta-strand of EC1 of one cadherin swaps places with the beta-strand of EC1 on a cadherin on a different cell. A conserved tryptophan residue anchors the adhesive interface by docking in a complementary hydrophobic pocket on the opposing strand. This strand swap is preceded by a fast binding X-dimer intermediate. The X-diamer localizes around the linker region between EC1 and EC2. This interface functions as a kinetic intermediate in the formation of the strand-swapped diamer (Vendome, J., Felsovalyi, K., et al. 2014).

Trans binding lowers the entropic penalty for the formation of *cis* interactions and is followed by *cis* clustering at a lateral interface formed between the base of an EC1 domain and the apex of a parallel (*cis*) EC2 domain. These form a lattice that is the structural basis for the adherens junctions (Harrison, O.J., Jin, X., et al. 2011). After the *trans* dimerization of cadherins on two difference cells, *cis* clustering occurs among cadherins on the same cell. Interactions between cadherins on proximal cells results in cell-cell adhesion, while *cis* interactions among cadherins on the same cell stabilize the junction (Harrison, O.J., Jin, X., et al. 2011).

α-Dystroglycan

α-Dystroglycan (α-DG) is the most well-studied O-mannosylated protein. It was discovered in chicken brains, rodent NG108-15 neural hybridoma cells, and mouse 3T3 fibroblasts under the name cranin (Smalheiser, N. and B Schwartz, N. 1987). α-DG has been found to be widely expressed in mammalian tissues with higher prevalence in skeletal muscle, brain, and epidermal tissues (Ibraghimov-Beskrovnaya, O., Milatovich, A., et al. 1993). α-DG is part of the dystrophin-dystroglycan complex which functions in attaching cells to the extracellular matrix (Ervasti, J.M. and Campbell, K.P. 1993, Ibraghimov-Beskrovnaya, O., Ervasti, J.M., et al. 1992). In this complex the intracellular actin cytoskeleton is linked to dystrophin which attaches to dystroglycan (Ervasti, J.M. and Campbell, K.P. 1991). The loss of functional dystrophin results in Duchenne's muscular dystrophy, a progressive disease of muscle often diagnosed in childhood but not apparent at birth (Hoffman, E.P., Brown, R.H., Jr., et al. 1987). Dystrophin is attached to the beta subunit of dystroglycan. Dystroglycan is translated from the DAG1 gene as a single 895-residue precursor protein and proteolytically cleaved into an alpha and beta subunit (Holt, K.H., Crosbie, R.H., et al. 2000). The alpha subunit contains an extended, proline-rich, mucin-like domain between N- and C-globular domains. It undergoes significant processing in the secretory pathway including N- and O-linked glycosylation and cleavage of the N-globular domain by furin (Ervasti, J.M. and Campbell, K.P. 1991). The mature alpha-subunit, residues 313-653 of the precursor protein, is extracellular and binds non-covalently to the transmembrane beta subunit (Holt, K.H., Crosbie, R.H., et al. 2000).

The mucin domain of α-DG, residues 316-485, is heavily modified by both O-mannose and O-GalNAc-initiated glycans. Site-mapping of the O-glycans has been done in

skeletal muscle from rabbit and human, as well as recombinant mouse α-DG expressed in HEK 293T cells (Harrison, R., Hitchen, P.G., et al. 2012, Nilsson, J., Nilsson, J., et al. 2010, Stalnaker, S.H., Hashmi, S., et al. 2010). Combined, these works identified a total of 18 sites of O-linked glycosylation on α-DG. Chapter 2 of this dissertation identifies four novel sites found on recombinant rabbit α- DG expressed in HEK 293 SimpleCells.

α-DG is the only protein on which the core M3 structure has been detected, likely due to the selectivity of the POMGnT2 enzyme (Halmo, S.M., Singh, D., et al. 2017). The matriglycan extended from the core M3 O- mannose glycan (see Core M3, above) is the cell surface receptor for the LG domain-containing ligands of laminin, agrin, merosin, and perlecan in the extra-cellular matrix (Yoshida-Moriguchi, T., Yu, L., et al. 2010). The importance of proper glycosylation of alpha-dystroglycan is evident from the congenital muscular dystrophies which occur in its absence.

Congenital Muscular Dystrophies

Congenital muscular dystrophies (CMDs) are a heterogeneous group of neuromuscular disorders characterized by muscle hypotonia and atrophy, and in some cases, brain and eye abnormalities. Signs of disease appear at birth or during infancy (Bertini, E., D'Amico, A., et al. 2011). Approximately 2 out of every 100,000 infants are affected by CMD (Norwood, F.L., Harling, C., et al. 2009). Many CMDs are caused by defects in the mechanisms of glycosylation. CMDs caused by defects of glycosylation are known as congenital disorders of glycosylation (CDG) (Sparks, S.E., Quijano-Roy, S., et al. 1993).

A subset of the CMDs is the secondary dystroglycanopathies, which are caused by hypo-glycosylation of α-DG (Clement, E., Mercuri, E., et al. 2008, Yoshida-Moriguchi, T. and

Campbell, K.P. 2015). The secondary dystroglycanopathies include, in decreasing order of severity, Walker-Warburg syndrome, Fukuyama CMD, Muscle-Eye-brain disease, and Limb-girdle muscular dystrophy (Brown, S.C. and Winder, S.J. 2012, Godfrey, C., Clement, E., et al. 2007). Of these, only Limb-girdle muscular dystrophy does not include neurological symptoms. The secondary dystroglycanopathies are named as such because they result not from defects in the dystroglycan protein, but from defects in the glycosylation of dystroglycan (Wells, L. 2013).

Mutations in *Pomt1*/*Pomt2*, not surprisingly, cause some of the most severe disease phenotypes. In mice, knocking out *Pomt1* results in embryonic lethality between E7.5 and E9.5 while knockout of *Pomt2* is embryonic lethal by day E3.5 (Willer, T., Prados, B., et al. 2004). The difference between these two knockouts is believed to be due to partial rescue by maternal Pomt1 mRNA (Lommel, M., Winterhalter, P.R., et al. 2013). In humans, mutations in *Pomt1* or *Pomt2* have been associated with Walker-Warburg syndrome, the most severe of the dystroglycanopathies (Beltrán-Valero de Bernabé, D., Currier, S., et al. 2002, van Reeuwijk, J., Janssen, M., et al. 2005).

Patients with secondary dystroglycanopathies have been found to have deleterious mutations in the enzymes POMT1, POMT2, POMGnT1, POMGnT2, B3GALNT2, POMK, FKTN, FKRP, FKRP, RXYLT1 (TMEM5), B4GAT1, or LARGE1. Each of these enzymes has a described role in building the functional glycan with one glaring exception; POMGnT1. POMGnT1 has no known role in building the functional glycan. Nevertheless, cell lines and murine tissues from POMGnT1 knockouts show a loss of matriglycan and loss of laminin binding activity (Liu, J., Ball, S.L., et al. 2006). Chapter 3 of this dissertation examines how loss of POMGnT1 results in loss of functional glycosylation of alpha-dystroglycan.

POMGnT1

POMGnT1 catalyzes the transfer of GlcNAc from UDP-GlcNAc to O-linked mannose in a β1-2 linkage to form the core M1 glycan, as described above. POMGnT1 is a single-pass type II transmembrane protein with residues 1-37 in the cytoplasm, 38-58 a transmembrane signal-anchor, and 59-660 in the lumen of the Golgi apparatus. As shown in Figure 1-3, the lumenal domain consists of a stem domain of two stacked β-sheets and two α-helices consisting of residues 97-250 proximal to the transmembrane domain, a linker region (residues 250-299), and a catalytic domain divided into a Rossmann-like fold motif structured N-lobe (residues 300-497) and an open sheet α/β-structured C-lobe (residues 498-646) (Kuwabara, N., Manya, H., et al. 2016). The Rossmann-like fold motif is a five-stranded version of the six-stranded Rossmann fold motif which functions to bind nucleotide co-factors. The Rossman fold is composed of alternating beta strands and alpha helices with the initial beta-alpha-beta fold as the most conserved segment (Rao, S.T. and Rossmann, M.G. 1973).

Loss of POMGnT1 has been associated with two different diseases. The first is Muscle-Eye-Brain disease (MEB), a congenital muscular dystrophy with etiology that greatly overlaps with the secondary dystroglycanopathies. A second disease associated with mutations in POMGnT1 is retinitis pigmentosa 76 (RP76). RP76 is a much milder disease than MEB, with no extra-ocular involvement. Symptoms include night vision blindness and loss of midperipheral visual field progressing to further loss of vision with age. RP76 is a Mendelian-heritable autosomal recessive disorder affecting about 1 in every 5,000 people (Xu, M., Yamada, T., et al. 2016). Mutations L120R, E156K, I287S, and G502A in POMGnT1 have been identified as causal for RP76 (Xu, M., Yamada, T., et al. 2016). Three of these mild disease-causing mutations are found in the stem region of POMGnT1.

In contrast to RP76, the symptoms of MEB are severe. Patients have muscle weakness, with or without brain and eye abnormalities. Most patients have severe neurological symptoms, including hydrocephalus, cobblestone lissencephaly, severe intellectual disability, and gross motor defects (Haltia, M., Leivo, I., et al. 1997, Taniguchi, K., Kobayashi, K., et al. 2003). MEB is also autosomal recessive with patients having either homozygous or compound heterozygous mutations in POMGnT1 (Yoshida, A., Kobayashi, K., et al. 2001). Over 30 POMGnT1 mutations are associated with MEB (Manya, H., Sakai, K., et al. 2003, Yoshida, A., Kobayashi, K., et al. 2001). These mutations occur throughout the luminal-facing structure of the enzyme, from T176 to R605, with several occurring in the catalytic region of the enzyme: R311Q, R367H, W425W, D427H, R442C, C490Y, P493R, and S550N (Biancheri, R., Bertini, E., et al. 2006, Clement, E., Mercuri, E., et al. 2008, Diesen, C., Saarinen, A., et al. 2004, S Vervoort, V., R Holden, K., et al. 2004).

Recent advances in our understanding of O-mannosylation have left significant questions regarding the POMGnT1 enzyme. With the transferase enzymes required to build the entire matriglycan-containing core M3 structure now known, what is the role of POMGnT1 in functional glycosylation of α-DG? Given the discovery of the M0 structure, how are the O-mannose residues on cadherins escaping modification by POMGnT1? The purpose of the work presented in this dissertation is to add knowledge and understanding to how some glycan structures indirectly require POMGnT1 and how others escape modification by it.

References

Abbott KL, Matthews RT, Pierce M. 2008. Receptor Tyrosine Phosphatase β (RPTPβ) Activity and Signaling Are Attenuated by Glycosylation and Subsequent Cell Surface Galectin-1 Binding. Journal of Biological Chemistry, 283:33026-33035.

Beltrán-Valero de Bernabé D, Currier S, Steinbrecher A, Celli J, van Beusekom E, van der Zwaag B, Kayserili H, Merlini L, Chitayat D, Dobyns WB, et al. 2002. Mutations in the O-mannosyltransferase gene POMT1 give rise to the severe neuronal migration disorder Walker-Warburg syndrome. American journal of human genetics, 71:1033-1043.

Bennett EP, Mandel U, Clausen H, Gerken TA, Fritz TA, Tabak LA. 2012. Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Glycobiology, 22:736-756.

Bertini E, D'Amico A, Gualandi F, Petrini S. 2011. Congenital muscular dystrophies: a brief review. Seminars in pediatric neurology, 18:277-288.

Biancheri R, Bertini E, Falace A, Pedemonte M, Rossi A, D'Amico A, Scapolan S, Bergamino L, Petrini S, Cassandrini D, et al. 2006. POMGnT1 mutations in congenital muscular dystrophy: genotype-phenotype correlation and expanded clinical spectrum. Archives of neurology, 63:1491-1495.

Bleckmann C, Geyer H, Lieberoth A, Splittstoesser F, Liu Y, Feizi T, Schachner M, Kleene R, Reinhold V, Geyer R. 2009. O-glycosylation pattern of CD24 from mouse brain. Biological chemistry, 390:627-645.

Bork P, Doerks T, Springer TA, Snel B. 1999. Domains in plexins: links to integrins and transcription factors. Trends Biochem Sci, 24:261-263.

Brasch J, Harrison OJ, Honig B, Shapiro L. 2012. Thinking outside the cell: how cadherins drive adhesion. Trends in cell biology, 22:299-310.

Brown SC, Winder SJ. 2012. Dystroglycan and dystroglycanopathies: report of the 187th ENMC Workshop 11-13 November 2011, Naarden, The Netherlands. Neuromuscul Disord, 22:659-668.

Bussolino F, Valdembri D, Caccavari F, Serini G. 2006. Semaphoring vascular morphogenesis. Endothelium : journal of endothelial cell research, 13:81-91.

Chai W, Yuen CT, Kogelberg H, Carruthers RA, Margolis RU, Feizi T, Lawson AM. 1999. High prevalence of 2-mono- and 2,6-di-substituted manol-terminating sequences among O-glycans released from brain glycopeptides by reductive alkaline hydrolysis. European journal of biochemistry, 263:879-888.

Clement E, Mercuri E, Godfrey C, Smith J, Robb S, Kinali M, Straub V, Bushby K, Manzur A, Talim B, et al. 2008. Brain involvement in muscular dystrophies with defective dystroglycan glycosylation. Annals of neurology, 64:573-582.

Collesi C, Santoro MM, Gaudino G, Comoglio PM. 1996. A splicing variant of the RON transcript induces constitutive tyrosine kinase activity and an invasive phenotype. Mol Cell Biol, 16:5518-5526.

Diesen C, Saarinen A, Pihko H, Rosenlew C, Cormand B, Dobyns WB, Dieguez J, Valanne L, Joensuu T, Lehesjoki AE. 2004. POMGnT1 mutation and phenotypic spectrum in muscle-eye-brain disease. Journal of medical genetics, 41:e115.

Dong DL, Hart GW. 1994. Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol. The Journal of biological chemistry, 269:19321-19330.

Dwyer CA, Baker E, Hu H, Matthews RT. 2012. RPTPzeta/phosphacan is abnormally glycosylated in a model of muscle-eye-brain disease lacking functional POMGnT1. Neuroscience, 220:47-61.

Dwyer CA, Katoh T, Tiemeyer M, Matthews RT. 2015. Neurons and glia modify receptor protein-tyrosine phosphatase zeta (RPTPzeta)/phosphacan with cell-specific O-mannosyl glycans in the developing brain. The Journal of biological chemistry, 290:10256-10273.

Endo T. 2014. Glycobiology of α-dystroglycan and muscular dystrophy. The Journal of Biochemistry, 157:1-12.

Ervasti JM, Campbell KP. 1991. Membrane organization of the dystrophin-glycoprotein complex. Cell, 66:1121-1131.

Ervasti JM, Campbell KP. 1993. A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin. J Cell Biol, 122:809-823.

Finne J, Krusius T, Margolis RK, Margolis RU. 1979. Novel mannitol-containing oligosaccharides obtained by mild alkaline borohydride treatment of a chondroitin sulfate proteoglycan from brain. The Journal of biological chemistry, 254:10295-10300.

Gerin I, Ury B, Breloy I, Bouchet-Seraphin C, Bolsée J, Halbout M, Graff J, Vertommen D, Muccioli GG, Seta N, et al. 2016. ISPD produces CDP-ribitol used by FKTN and FKRP to transfer ribitol phosphate onto α-dystroglycan. Nat Commun, 7:11534-11534.

Godfrey C, Clement E, Mein R, Brockington M, Smith J, Talim B, Straub V, Robb S, Quinlivan R, Feng L, et al. 2007. Refining genotype phenotype correlations in muscular dystrophies with defective glycosylation of dystroglycan. Brain : a journal of neurology, 130:2725-2735.

Halbleib JM, Nelson WJ. 2006. Cadherins in development: cell adhesion, sorting, and tissue morphogenesis. Genes & development, 20:3199-3214.

Halim A, Larsen IS, Neubert P, Joshi HJ, Petersen BL, Vakhrushev SY, Strahl S, Clausen H. 2015. Discovery of a nucleocytoplasmic O-mannose glycoproteome in yeast. Proc Natl Acad Sci U S A, 112:15648-15653.

Halmo SM, Singh D, Patel S, Wang S, Edlin M, Boons GJ, Moremen KW, Live D, Wells L. 2017. Protein O-Linked Mannose beta-1,4-N-Acetylglucosaminyl-transferase 2 (POMGNT2) Is a Gatekeeper Enzyme for Functional Glycosylation of alpha-Dystroglycan. The Journal of biological chemistry, 292:2101-2109.

Haltia M, Leivo I, Somer H, Pihko H, Paetau A, Kivela T, Tarkkanen A, Tome F, Engvall E, Santavuori P. 1997. Muscle-eye-brain disease: a neuropathological study. Annals of neurology, 41:173-180.

Haltiwanger RS, Holt GD, Hart GW. 1990. Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-acetylglucosamine:peptide beta-N-acetylglucosaminyltransferase. The Journal of biological chemistry, 265:2563-2568.

Harrison OJ, Jin X, Hong S, Bahna F, Ahlsen G, Brasch J, Wu Y, Vendome J, Felsovalyi K, Hampton CM, et al. 2011. The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. Structure (London, England : 1993), 19:244-256.

Harrison R, Hitchen PG, Panico M, Morris HR, Mekhaiel D, Pleass RJ, Dell A, Hewitt JE, Haslam SM. 2012. Glycoproteomic characterization of recombinant mouse alpha-dystroglycan. Glycobiology, 22:662-675.

Hart GW. 2019. Nutrient regulation of signaling and transcription. The Journal of biological chemistry, 294:2211-2231.

Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O. 2011. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annual review of biochemistry, 80:825-858.

Harvey BM, Haltiwanger RS. 2018. Regulation of Notch Function by O-Glycosylation. Advances in experimental medicine and biology, 1066:59-78.

Hase S, Kawabata S, Nishimura H, Takeya H, Sueyoshi T, Miyata T, Iwanaga S, Takao T, Shimonishi Y, Ikenaka T. 1988. A new trisaccharide sugar chain linked to a serine residue in bovine blood coagulation factors VII and IX. Journal of biochemistry, 104:867-868.

Hoffman EP, Brown RH, Jr., Kunkel LM. 1987. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. Cell, 51:919-928.

Holdener BC, Haltiwanger RS. 2019. Protein O-fucosylation: structure and function. Current opinion in structural biology, 56:78-86.

Holt GD, Hart GW. 1986. The subcellular distribution of terminal N-acetylglucosamine moieties. Localization of a novel protein-saccharide linkage, O-linked GlcNAc. The Journal of biological chemistry, 261:8049-8057.

Holt KH, Crosbie RH, Venzke DP, Campbell KP. 2000. Biosynthesis of dystroglycan: processing of a precursor propeptide. FEBS letters, 468:79-83.

Ibraghimov-Beskrovnaya O, Ervasti JM, Leveille CJ, Slaughter CA, Sernett SW, Campbell KP. 1992. Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. Nature, 355:696-702.

Ibraghimov-Beskrovnaya O, Milatovich A, Ozcelik T, Yang B, Koepnick K, Francke U, Campbell KP. 1993. Human dystroglycan: skeletal muscle cDNA, genomic structure, origin of tissue specific isoforms and chromosomal localization. Hum Mol Genet, 2:1651-1657.

Inamori K, Yoshida-Moriguchi T, Hara Y, Anderson ME, Yu L, Campbell KP. 2012. Dystroglycan function requires xylosyl- and glucuronyltransferase activities of LARGE. Science (New York, N.Y.), 335:93-96.

Jerber J, Zaki MS, Al-Aama JY, Rosti RO, Ben-Omran T, Dikoglu E, Silhavy JL, Caglar C, Musaev D, Albrecht B, et al. 2016. Biallelic Mutations in TMTC3, Encoding a Transmembrane and TPR-Containing Protein, Lead to Cobblestone Lissencephaly. American journal of human genetics, 99:1181-1189.

Ju T, Cummings RD. 2002. A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase. Proc Natl Acad Sci U S A, 99:16613-16618.

Jurado LA, Coloma A, Cruces J. 1999. Identification of a human homolog of the Drosophila rotated abdomen gene (POMT1) encoding a putative protein O-mannosyl-transferase, and assignment to human chromosome 9q34.1. Genomics, 58:171-180.

Kakuda S, Haltiwanger RS. 2017. Deciphering the Fringe-Mediated Notch Code: Identification of Activating and Inhibiting Sites Allowing Discrimination between Ligands. Dev Cell, 40:193-201.

Kowalczyk AP, Green KJ. 2013. Structure, function, and regulation of desmosomes. Prog Mol Biol Transl Sci, 116:95-118.

Kuwabara N, Manya H, Yamada T, Tateno H, Kanagawa M, Kobayashi K, Akasaka-Manya K, Hirose Y, Mizuno M, Ikeguchi M, et al. 2016. Carbohydrate-binding domain of the POMGnT1 stem region modulates O-mannosylation sites of α-dystroglycan. Proc Natl Acad Sci U S A, 113:9280-9285.

Lange T, Ullrich S, Muller I, Nentwich MF, Stubke K, Feldhaus S, Knies C, Hellwinkel OJ, Vessella RL, Abramjuk C, et al. 2012. Human prostate cancer in a clinically relevant xenograft mouse model: identification of beta(1,6)-branched oligosaccharides as a marker of tumor progression. Clinical cancer research : an official journal of the American Association for Cancer Research, 18:1364-1373.

Larsen ISB, Narimatsu Y, Clausen H, Joshi HJ, Halim A. 2019. Multiple distinct O-Mannosylation pathways in eukaryotes. Current opinion in structural biology, 56:171-178.

Larsen ISB, Narimatsu Y, Joshi HJ, Siukstaite L, Harrison OJ, Brasch J, Goodman KM, Hansen L, Shapiro L, Honig B, et al. 2017a. Discovery of an O-mannosylation pathway selectively serving cadherins and protocadherins. Proceedings of the National Academy of Sciences, 114:11163-11168.

Larsen ISB, Narimatsu Y, Joshi HJ, Yang Z, Harrison OJ, Brasch J, Shapiro L, Honig B, Vakhrushev SY, Clausen H, et al. 2017b. Mammalian O-mannosylation of cadherins and plexins is independent of protein O-mannosyltransferases 1 and 2. The Journal of biological chemistry, 292:11586-11598.

Lee JK, Matthews RT, Lim JM, Swanier K, Wells L, Pierce JM. 2012. Developmental expression of the neuron-specific N-acetylglucosaminyltransferase Vb (GnT-Vb/IX) and identification of its in vivo glycan products in comparison with those of its paralog, GnT-V. The Journal of biological chemistry, 287:28526-28536.

Li Z, Han K, Pak JE, Satkunarajah M, Zhou D, Rini JM. 2017. Recognition of EGF-like domains by the Notch-modifying O-fucosyltransferase POFUT1. Nature chemical biology, 13:757-763.

Liu J, Ball SL, Yang Y, Mei P, Zhang L, Shi H, Kaminski HJ, Lemmon VP, Hu H. 2006. A genetic model for muscle-eye-brain disease in mice lacking protein O-mannose 1,2-N-acetylglucosaminyltransferase (POMGnT1). Mechanisms of development, 123:228-240.

Live D, Wells L, Boons G-J. 2013. Dissecting the molecular basis of the role of the O-mannosylation pathway in disease: α-dystroglycan and forms of muscular dystrophy. Chembiochem, 14:2392-2402.

Loibl M, Strahl S. 2013. Protein O-mannosylation: what we have learned from baker's yeast. Biochimica et biophysica acta, 1833:2438-2446.

Lommel M, Strahl S. 2009. Protein O-mannosylation: conserved from bacteria to humans. Glycobiology, 19:816-828.

Lommel M, Winterhalter PR, Willer T, Dahlhoff M, Schneider MR, Bartels MF, Renner-Muller I, Ruppert T, Wolf E, Strahl S. 2013. Protein O-mannosylation is crucial for E-cadherin-mediated cell adhesion. Proc Natl Acad Sci U S A, 110:21024-21029.

Luca VC, Jude KM, Pierce NW, Nachury MV, Fischer S, Garcia KC. 2015. Structural biology. Structural basis for Notch1 engagement of Delta-like 4. Science (New York, N.Y.), 347:847-853.

Luca VC, Kim BC, Ge C, Kakuda S, Wu D, Roein-Peikar M, Haltiwanger RS, Zhu C, Ha T, Garcia KC. 2017. Notch-Jagged complex structure implicates a catch bond in tuning ligand sensitivity. Science (New York, N.Y.), 355:1320-1324.

Luo Y, Nita-Lazar A, Haltiwanger RS. 2006. Two distinct pathways for O-fucosylation of epidermal growth factor-like or thrombospondin type 1 repeats. The Journal of biological chemistry, 281:9385-9392.

Manya H, Chiba A, Yoshida A, Wang X, Chiba Y, Jigami Y, Margolis RU, Endo T. 2004. Demonstration of mammalian protein O-mannosyltransferase activity: coexpression of POMT1 and POMT2 required for enzymatic activity. Proc Natl Acad Sci U S A, 101:500-505.

Manya H, Sakai K, Kobayashi K, Taniguchi K, Kawakita M, Toda T, Endo T. 2003. Loss-of-function of an N-acetylglucosaminyltransferase, POMGnT1, in muscle-eye-brain disease. Biochemical and biophysical research communications, 306:93-97.

Manya H, Yamaguchi Y, Kanagawa M, Kobayashi K, Tajiri M, Akasaka-Manya K, Kawakami H, Mizuno M, Wada Y, Toda T, et al. 2016. The Muscular Dystrophy Gene TMEM5 Encodes a Ribitol beta1,4-Xylosyltransferase Required for the Functional Glycosylation of Dystroglycan. The Journal of biological chemistry, 291:24618-24627.

Martin PT. 2007. Congenital muscular dystrophies involving the O-mannose pathway. Curr Mol Med, 7:417-425.

Martinez T, Pace D, Brady L, Gerhart M, Balland A. 2007. Characterization of a novel modification on IgG2 light chain. Evidence for the presence of O-linked mannosylation. Journal of chromatography. A, 1156:183-187.

Matsuura A, Ito M, Sakaidani Y, Kondo T, Murakami K, Furukawa K, Nadano D, Matsuda T, Okajima T. 2008. O-linked N-acetylglucosamine is present on the extracellular domain of notch receptors. The Journal of biological chemistry, 283:35486-35495.

Maurel P, Rauch U, Flad M, Margolis RK, Margolis RU. 1994. Phosphacan, a chondroitin sulfate proteoglycan of brain that interacts with neurons and neural cell-adhesion molecules, is an extracellular variant of a receptor-type protein tyrosine phosphatase. Proc Natl Acad Sci U S A, 91:2512-2516.

Moloney DJ, Panin VM, Johnston SH, Chen J, Shao L, Wilson R, Wang Y, Stanley P, Irvine KD, Haltiwanger RS, et al. 2000. Fringe is a glycosyltransferase that modifies Notch. Nature, 406:369-375.

Nilsson J, Nilsson J, Larson G, Grahn A. 2010. Characterization of site-specific O-glycan structures within the mucin-like domain of alpha-dystroglycan from human skeletal muscle. Glycobiology, 20:1160-1169.

Norwood FL, Harling C, Chinnery PF, Eagle M, Bushby K, Straub V. 2009. Prevalence of genetic muscle disease in Northern England: in-depth analysis of a muscle clinic population. Brain : a journal of neurology, 132:3175-3186.

Panin VM, Papayannopoulos V, Wilson R, Irvine KD. 1997. Fringe modulates Notch-ligand interactions. Nature, 387:908-912.

Peles E, Nativ M, Campbell PL, Sakurai T, Martinez R, Lev S, Clary DO, Schilling J, Barnea G, Plowman GD, et al. 1995. The carbonic anhydrase domain of receptor tyrosine phosphatase beta is a functional ligand for the axonal cell recognition molecule contactin. Cell, 82:251-260.

Petrou VI, Herrera CM, Schultz KM, Clarke OB, Vendome J, Tomasek D, Banerjee S, Rajashankar KR, Belcher Dufrisne M, Kloss B, et al. 2016. Structures of aminoarabinose transferase ArnT suggest a molecular basis for lipid A glycosylation. Science (New York, N.Y.), 351:608-612.

Pokutta S, Herrenknecht K, Kemler R, Engel J. 1994. Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding. European journal of biochemistry, 223:1019-1026.

Praissman JL, Live DH, Wang S, Ramiah A, Chinoy ZS, Boons GJ, Moremen KW, Wells L. 2014. B4GAT1 is the priming enzyme for the LARGE-dependent functional glycosylation of alpha-dystroglycan. eLife, 3.

Praissman JL, Wells L. 2014. Mammalian O-mannosylation pathway: glycan structures, enzymes, and protein substrates. Biochemistry, 53:3066-3078.

Praissman JL, Willer T, Sheikh MO, Toi A, Chitayat D, Lin YY, Lee H, Stalnaker SH, Wang S, Prabhakar PK, et al. 2016. The functional O-mannose glycan on alpha-dystroglycan contains a phospho-ribitol primed for matriglycan addition. eLife, 5.

Racape M, Duong Van Huyen JP, Danger R, Giral M, Bleicher F, Foucher Y, Pallier A, Pilet P, Tafelmeyer P, Ashton-Chess J, et al. 2011. The involvement of SMILE/TMTC3 in endoplasmic reticulum stress response. PloS one, 6:e19321.

Rampal R, Li ASY, Moloney DJ, Georgiou SA, Luther KB, Nita-Lazar A, Haltiwanger RS. 2005. Lunatic Fringe, Manic Fringe, and Radical Fringe Recognize Similar Specificity Determinants in O-Fucosylated Epidermal Growth Factor-like Repeats. Journal of Biological Chemistry, 280:42454-42463.

Rao ST, Rossmann MG. 1973. Comparison of super-secondary structures in proteins. Journal of Molecular Biology, 76:241-256.

S Vervoort V, R Holden K, C Ukadike K, S Collins J, Saul R, Anand K. 2004. POMGnT1 gene alterations in a family with neurological abnormalities. Annals of neurology, 56:143-148.

Sakaidani Y, Ichiyanagi N, Saito C, Nomura T, Ito M, Nishio Y, Nadano D, Matsuda T, Furukawa K, Okajima T. 2012. O-linked-N-acetylglucosamine modification of mammalian Notch receptors by an atypical O-GlcNAc transferase Eogt1. Biochemical and biophysical research communications, 419:14-19.

Schmidt EF, Togashi H, Strittmatter SM. 2010. Chapter 240 - Semaphorins and their Receptors in Vertebrates and Invertebrates. In: Bradshaw RA, Dennis EA editors. Handbook of Cell Signaling (Second Edition). San Diego: Academic Press. p. 1961-1966.

Schreiner D, Weiner JA. 2010. Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion. Proc Natl Acad Sci U S A, 107:14893-14898.

Sentandreu R, Northcote DH. 1968. The structure of a glycopeptide isolated from the yeast cell wall. The Biochemical journal, 109:419-432.

Sheikh MO, Halmo SM, Wells L. 2017. Recent advancements in understanding mammalian O-mannosylation. Glycobiology, 27:806-819.

Smalheiser N, B Schwartz N. 1987. Cranin: A Laminin-Binding Protein of Cell Membranes. Proc Natl Acad Sci U S A, 84:6457-6461.

Sparks SE, Quijano-Roy S, Harper A, Rutkowski A, Gordon E, Hoffman EP, Pegoraro E. 1993. Congenital Muscular Dystrophy Overview. GeneReviews.

Stalnaker SH, Aoki K, Lim JM, Porterfield M, Liu M, Satz JS, Buskirk S, Xiong Y, Zhang P, Campbell KP, et al. 2011. Glycomic analyses of mouse models of congenital muscular dystrophy. The Journal of biological chemistry, 286:21180-21190.

Stalnaker SH, Hashmi S, Lim JM, Aoki K, Porterfield M, Gutierrez-Sanchez G, Wheeler J, Ervasti JM, Bergmann C, Tiemeyer M, et al. 2010. Site mapping and characterization of O-glycan structures on alpha-dystroglycan isolated from rabbit skeletal muscle. The Journal of biological chemistry, 285:24882-24891.

Strahl-Bolsinger S, Immervoll T, Deutzmann R, Tanner W. 1993. PMT1, the gene for a key enzyme of protein O-glycosylation in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A, 90:8164-8168.

Sunryd JC, Cheon B, Graham JB, Giorda KM, Fissore RA, Hebert DN. 2014. TMTC1 and TMTC2 are novel endoplasmic reticulum tetratricopeptide repeat-containing adapter proteins involved in calcium homeostasis. The Journal of biological chemistry, 289:16085-16099.

Takeuchi H, Yu H, Hao H, Takeuchi M, Ito A, Li H, Haltiwanger RS. 2017. O-Glycosylation modulates the stability of epidermal growth factor-like repeats and thereby regulates Notch trafficking. Journal of Biological Chemistry.

Taniguchi K, Kobayashi K, Saito K, Yamanouchi H, Ohnuma A, Hayashi YK, Manya H, Jin DK, Lee M, Parano E, et al. 2003. Worldwide distribution and broader clinical spectrum of muscle-eye-brain disease. Hum Mol Genet, 12:527-534.

Usui T, Shima Y, Shimada Y, Hirano S, Burgess RW, Schwarz TL, Takeichi M, Uemura T. 1999. Flamingo, a seven-pass transmembrane cadherin, regulates planar cell polarity under the control of Frizzled. Cell, 98:585-595.

Valero-Gonzalez J, Leonhard-Melief C, Lira-Navarrete E, Jimenez-Oses G, Hernandez-Ruiz C, Pallares MC, Yruela I, Vasudevan D, Lostao A, Corzana F, et al. 2016. A proactive role of water molecules in acceptor recognition by protein O-fucosyltransferase 2. Nature chemical biology, 12:240-246.

van Reeuwijk J, Janssen M, van den Elzen C, Beltran-Valero de Bernabe D, Sabatelli P, Merlini L, Boon M, Scheffer H, Brockington M, Muntoni F, et al. 2005. POMT2 mutations cause alpha-dystroglycan hypoglycosylation and Walker-Warburg syndrome. Journal of medical genetics, 42:907-912.

Vasudevan D, Takeuchi H, Johar SS, Majerus E, Haltiwanger RS. 2015. Peters plus syndrome mutations disrupt a noncanonical ER quality-control mechanism. Current biology : CB, 25:286-295.

Vendome J, Felsovalyi K, Song H, Yang Z, Jin X, Brasch J, Harrison OJ, Ahlsen G, Bahna F, Kaczynska A, et al. 2014. Structural and energetic determinants of adhesive binding specificity in type I cadherins. Proc Natl Acad Sci U S A, 111:E4175-4184.

Vester-Christensen MB, Halim A, Joshi HJ, Steentoft C, Bennett EP, Levery SB, Vakhrushev SY, Clausen H. 2013. Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. Proceedings of the National Academy of Sciences, 110:21018-21023.

Wang Y, Ju T, Ding X, Xia B, Wang W, Xia L, He M, Cummings RD. 2010. Cosmc is an essential chaperone for correct protein O-glycosylation. Proc Natl Acad Sci U S A, 107:9228-9233.

Wang Y, Lee GF, Kelley RF, Spellman MW. 1996. Identification of a GDP-L-fucose:polypeptide fucosyltransferase and enzymatic addition of O-linked fucose to EGF domains. Glycobiology, 6:837-842.

Wells L. 2013. The o-mannosylation pathway: glycosyltransferases and proteins implicated in congenital muscular dystrophy. The Journal of biological chemistry, 288:6930-6935.

Wells L, Kreppel LK, Comer FI, Wadzinski BE, Hart GW. 2004. O-GlcNAc transferase is in a functional complex with protein phosphatase 1 catalytic subunits. The Journal of biological chemistry, 279:38466-38470.

Willer T, Prados B, Falcon-Perez JM, Renner-Muller I, Przemeck GK, Lommel M, Coloma A, Valero MC, de Angelis MH, Tanner W, et al. 2004. Targeted disruption of the Walker-Warburg syndrome gene Pomt1 in mouse results in embryonic lethality. Proc Natl Acad Sci U S A, 101:14126-14131.

Winberg ML, Noordermeer JN, Tamagnone L, Comoglio PM, Spriggs MK, Tessier-Lavigne M, Goodman CS. 1998. Plexin A is a neuronal semaphorin receptor that controls axon guidance. Cell, 95:903-916.

Winterhalter PR, Lommel M, Ruppert T, Strahl S. 2013. O-glycosylation of the non-canonical T-cadherin from rabbit skeletal muscle by single mannose residues. FEBS letters, 587:3715-3721.

Xu M, Yamada T, Sun Z, Eblimit A, Lopez I, Wang F, Manya H, Xu S, Zhao L, Li Y, et al. 2016. Mutations in POMGNT1 cause non-syndromic retinitis pigmentosa. Hum Mol Genet, 25:1479-1488.

Yoshida-Moriguchi T, Campbell KP. 2015. Matriglycan: a novel polysaccharide that links dystroglycan to the basement membrane. Glycobiology, 25:702-713.

Yoshida-Moriguchi T, Willer T, Anderson ME, Venzke D, Whyte T, Muntoni F, Lee H, Nelson SF, Yu L, Campbell KP. 2013. SGK196 is a glycosylation-specific O-mannose kinase required for dystroglycan function. Science (New York, N.Y.), 341:896-899.

Yoshida-Moriguchi T, Yu L, Stalnaker SH, Davis S, Kunz S, Madson M, Oldstone MB, Schachter H, Wells L, Campbell KP. 2010. O-mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. Science (New York, N.Y.), 327:88-92.

Yoshida A, Kobayashi K, Manya H, Taniguchi K, Kano H, Mizuno M, Inazu T, Mitsuhashi H, Takahashi S, Takeuchi M, et al. 2001. Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. Dev Cell, 1:717-724.

Figure 1-1.

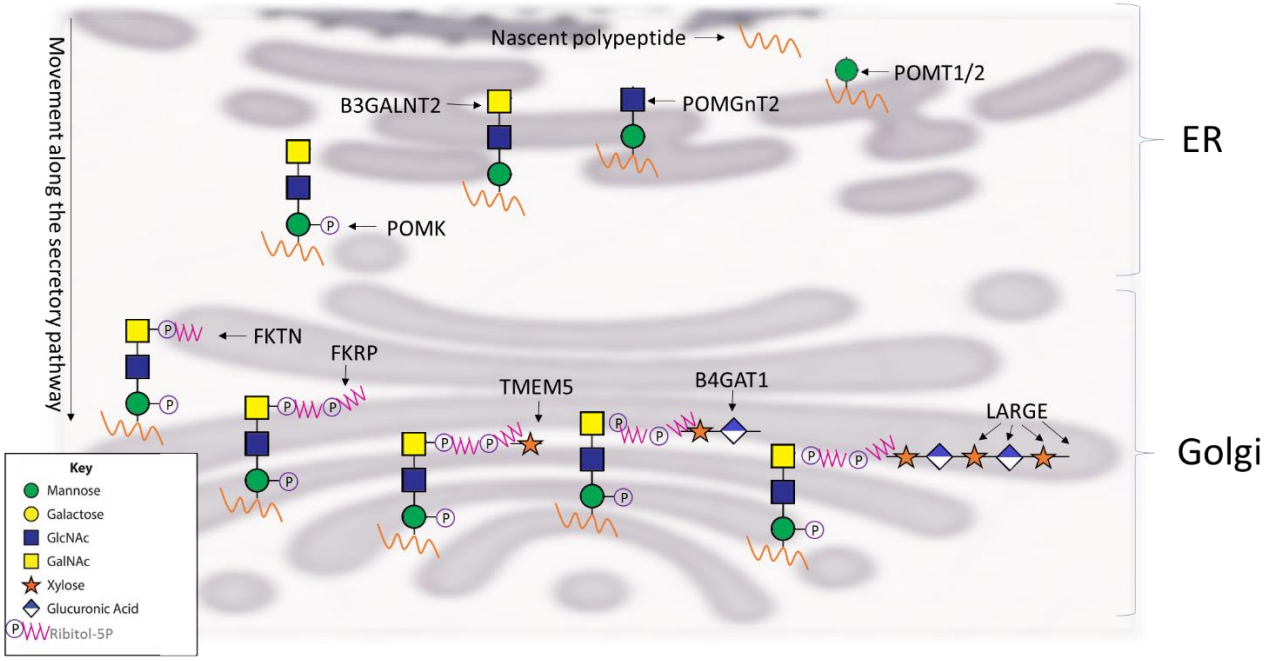

Movement along the secretory pathway
Nascent polypeptide
POMT1/2
B3GALNT2
POMGnT2
POMK
ER
FKTN
FKRP
TMEM5
B4GAT1
LARGE
Golgi
Key
Mannose
Galactose
GlcNAc
GalNAc
Xylose
Glucuronic Acid
Ribitol-5P

Figure 1-1. The process of building the matriglycan-containing core M3 O-mannose glycan on alpha-dystroglycan. The first four steps, catalyzed by enzymes POMT1/POMT2, POMGNT2, B3GALNT2, and POMK, occur in the ER. The latter steps, catalyzed by FKTN, FKRP, TMEM5, B4GAT1, and LARGE, occur in the Golgi.

Figure 1-2.

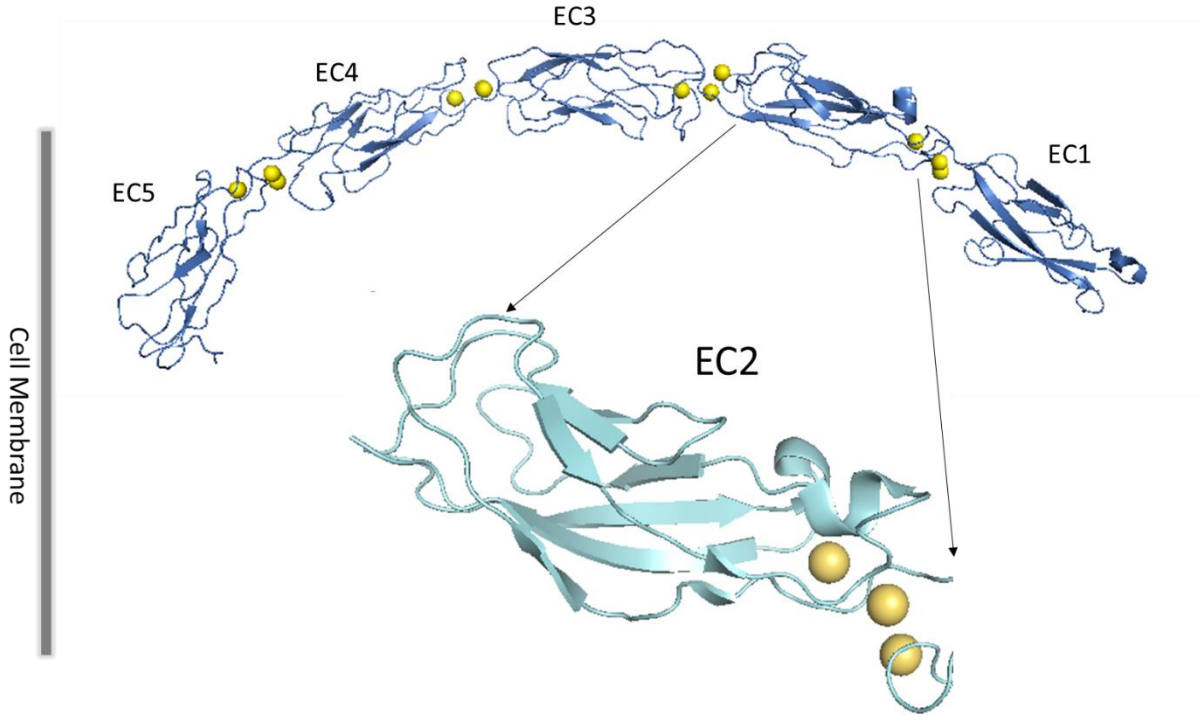


Figure 1-2. The ectodomain of E-cadherin. The ectodomain of E-cadherin, consisting of EC1-EC5, is shown in blue. Three calcium ions between each EC domain are shown in yellow. An enlarged image of EC2 shows the seven strands arranged antiparallel which form the β-barrel. pdb-ID: 1L3W (full ectodomain) and 1Q1P (EC2).

Figure 1-3.

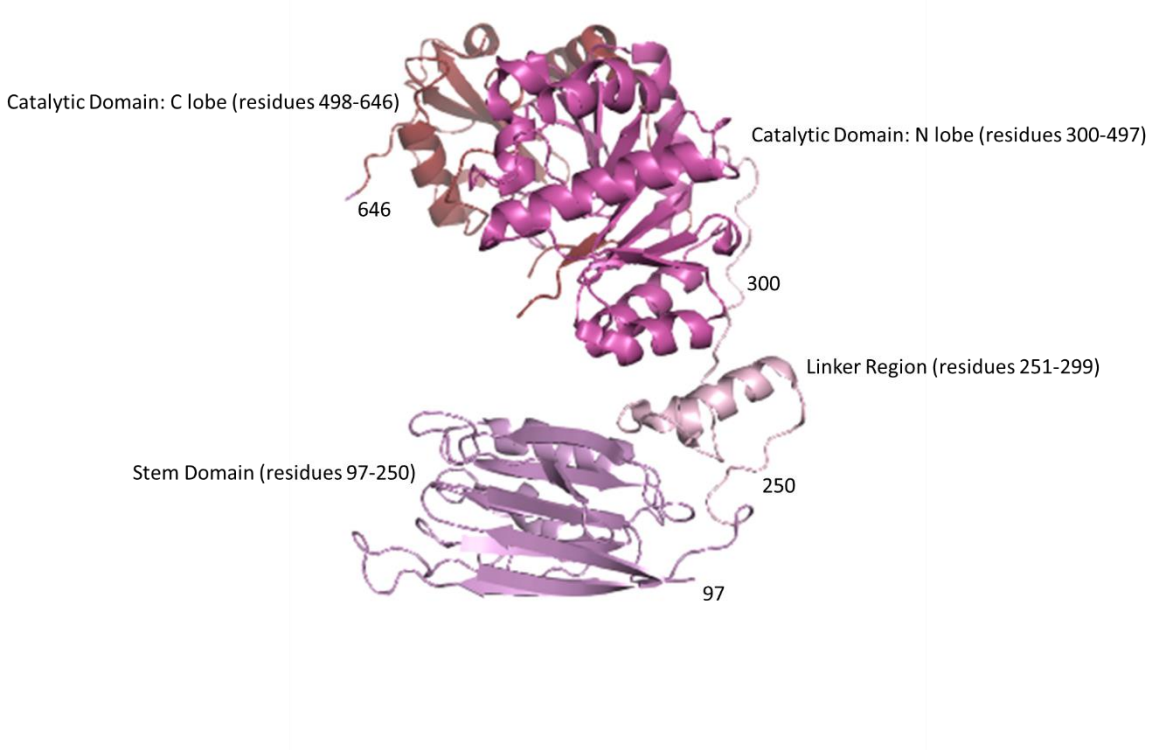

Catalytic Domain: C lobe (residues 498-646)
Catalytic Domain: N lobe (residues 300-497)
646
300
Linker Region (residues 251-299)
Stem Domain (residues 97-250)
250
97

Figure 1-3. The structure of POMGnT1. Domains are labelled with catalyctic domain C-lobe in ruby, N-lobe in raspberry, the linker region in light pink, and the membrane-proximal stem domain in violet. pdb ID: 5GGG.

CHAPTER 2

PROTEIN FOLDING PREVENTS ELONGATION OF CORE M0 STRUCTURES ON CADHERINS

Background

Mammalian O-mannosylation was first reported in rabbit and rat brain, but is now recognized as a major glycoform in a broad range of species (Chai, W., Yuen, C.T., et al. 1999, Finne, J., Krusius, T., et al. 1979). Until recently only a handful of proteins were known to be O-mannosylated: receptor protein tyrosine phosphatase beta, phosphacan, CD24, and alpha-dystroglycan (Abbott, K.L., Matthews, R.T., et al. 2008, Bleckmann, C., Geyer, H., et al. 2009, Dwyer, C.A., Baker, E., et al. 2012). Alpha-dystroglycan is the most extensively studied O-mannosylated glycoprotein due to the congenital muscular dystrophies associated with aberrant glycoforms (Sheikh, M.O., Halmo, S.M., et al. 2017, Wells, L. 2013, Yoshida-Moriguchi, T. and Campbell, K.P. 2015, Yoshida-Moriguchi, T., Yu, L., et al. 2010). The O-mannosylation pathway of alpha-dystroglycan begins in the endoplasmic reticulum (ER) where the protein O-mannosyltransferases POMT1/POMT2 transfer mannose from dolichol monophosphate-activated mannose to serine or threonine residues on a nascent polypeptide (Lommel, M. and Strahl, S. 2009). In the ER, the enzyme POMGnT2 may add a β-1,4-linked N-acetylglucosamine to create a core M3 structure. However, the product of POMGnT2 is only seen on two sites of the protein alpha-dystroglycan and the enzyme has been shown to exhibit significant primary amino acid specificity in its substrates (Halmo, S.M., Singh, D., et al. 2017). Modification by POMGnT2 in the ER leads to a unique functional core M3 structure

which is elongated by several enzymes and has been reviewed previously (Manya, H. and Endo, T. 2017, Sheikh, M.O., Halmo, S.M., et al. 2017).

After passing to the cis-Golgi, the sites of O-mannosylation which remain unmodified by POMGnT2 should be available to POMGnT1. POMGnT1 adds a β-1,2-linked N-acetylglucosamine to the O-mannose forming the core M1 structure. Unlike POMGnT2, POMGnT1 has not been shown to exhibit substrate specificity and is considered a promiscuous enzyme (Halmo, S.M., Singh, D., et al. 2017). Core M1 structures may then be extended by β-1,4-linked galactose and either a sialic acid, to form the classical tetrasaccharide, a fucose, to form the Lewis x structure, or a glucuronic acid which can then be sulfated to form the HNK-1 epitope. The function of these core M1 structures remains unknown. They are highly prevalent in the brain, making up nearly one-third of O-linked glycans, even in the absence of alpha-dystroglycan (Chai, W., Yuen, C.T., et al. 1999, Stalnaker, S.H., Aoki, K., et al. 2011).

Our understanding of the mammalian O-mannose glycoproteome has expanded considerably in recent years. O-linked glycosylation was observed on crystal structures of the ectodomains of E- and N-cadherins published by Harrison et al (Harrison, O.J., Jin, X., et al. 2011). The O-linked glycan was determined by two independent approaches to be a non-elongated O-mannose, the M0 structure. Weak lectin affinity chromatography of glycopeptides from SimpleCells revealed cadherins as major O- mannosylated glycoproteins (Vester-Christensen, M.B., Halim, A., et al. 2013). In total, 37 cadherins and 6 plexins were identified as being O-mannosylated. Additionally, Winterhalter et al. used an antibody developed against O-mannose to identify rabbit T-cadherin as a novel O-mannosylated protein and determined that the O-mannose residue was not extended by N-acetylglucosamine to form

any of the three core M structures (Winterhalter, P.R., Lommel, M., et al. 2013). O-mannosylation of cadherins is essential for E-cadherin mediated cell adhesion and its absence prevents preimplantation development of mouse embryos from the morula to the blastocyst stage (Lommel, M., Winterhalter, P.R., et al. 2013).

However, surprisingly, POMT1 and POMT2, the enzyme complex that adds O-mannose to alpha-dystroglycan, is expendable for O-mannosylation of cadherins (Larsen, I.S.B., Narimatsu, Y., et al. 2017b). Using bioinformatic searches for sequence homology, Larsen et al. identified the genes for the tetra-trico-peptide repeat-containing proteins (TMTC1-4) as candidates for being novel O-mannosyltransferases. Knockout of these genes resulted in loss of O-mannosylation on cadherin proteins in HEK 293 cells, demonstrating that cadherin O-mannosylation is dependent on TMTC1—4 (Larsen, I.S.B., Narimatsu, Y., et al. 2017a). The O-mannose sites identified on E-cadherin were mapped to the B and G strands of the seven-stranded beta barrel structures of the EC domains. Loss of TMTC1/3 resulted in a loss of O-mannose on the G strand, while loss of TMTC1/2/3/4 resulted in a loss of all O-mannose on recombinantly expressed E-cadherin.

The finding that the O-mannose glycans on cadherins and plexins are M0 structures that are not elongated was surprising given that POMGnT1 has been shown to lack peptide specificity for synthetic glycopeptides of sequences derived from alpha-dystroglycan (Halmo, S.M., Singh, D., et al. 2017). Given the promiscuity of POMGnT1, we sought to determine how the O-mannose residues on cadherins escaped modification. We hypothesized that protein folding in the ER rendered the sites of O-mannose inaccessible to POMGnT1 residing in the cis-Golgi.

Results

Expressed Proteins are modified by a non-elongated O-Mannose

Recombinant E-cadherin, N-cadherin, and alpha-dystroglycan were expressed in both HEK 293 Wt and POMGNT1-knockout cells (Figure 2-1A). As expected, alpha-dystroglycan expressed in HEK 293 Wt was functionally glycosylated, as determined by reactivity of the IIH6 antibody, while alpha-dystroglycan expressed in POMGnT1-knockout cells was not (Figure 2-1A).

E-cadherin and N-cadherin were initially characterized as being O-mannosylated in previous work via their interaction with concanavalin-A when expressed in cells lacking POMGNT1. Since concanavalin-A can interact with both mannose and glucose, we first sought to confirm the identity of the hexose. The E-cadherin, N-cadherin, and alpha-dystroglycan expressed in HEK Wt cells was analyzed using tandem mass spectrometry. As expected, we were able to detect hexose modifications with no evidence of extended structures (Figures 2-1B, C and Appendix B). Similarly, when we expressed alpha-dystroglycan in POMGNT1 knock-out cells, we also observed multiple sites of O-hexose modification without extension (Figure 2-1D) as compared to our and others' original studies that demonstrated that the majority of O-Man sites on alpha-dystroglycan are extended when expressed or purified from wild-type sources (Harrison, R., Hitchen, P.G., et al. 2012, Stalnaker, S.H., Hashmi, S., et al. 2010). The approximate location of each hexose is illustrated in Figure 2-1E. Figure 2-1F identifies the specific residues on E-cadherin and N-cadherin found to be modified by O-hexose. The location of the O-hexoses on the EC2-4 domains of each cadherin is mapped in Figure 2-1G.

To verify that the observed hexoses were indeed mannose, the resulting peptides were treated with alpha-mannosidase. The ability of the alpha-mannosidase to remove protein O-linked mannose was confirmed using a synthetic O-Mannose modified peptide. The alpha-mannosidase reduced the major peak detected by MALDI-TOF in the sample from 2373 m/z to 2211 m/z, a loss of 162 m/z which corresponds to the loss of mannose (Figure 2-2A). On the expressed proteins, alpha-mannosidase treatment reduced the number of hexose-containing peptides detected by >90% (Figures 2-2B, C, D and E) demonstrating that the detected hexoses are indeed non-elongated O-mannose M0 structures.

POMGnT1 elongates tryptic peptides of E-Cadherin and N-Cadherin

To determine whether lack of extension by POMGnT1 was due to tertiary structure of the folded proteins, the E-cadherin, N-cadherin, and alpha-dystroglycan constructs were expressed in HEK-POMGnT1 knockout cells and a portion of each protein was reduced, alkylated, and digested into tryptic peptides. We then utilized UDP-[^{3}H]GlcNAc to quantify the transfer of GlcNAc to the O-mannose by POMGnT1 on equal amounts of protein and derived peptide for each sample. E-cadherin and N-cadherin were both poor substrates at the protein level whereas alpha-dystroglycan protein was efficiently modified. By contrast, the tryptic peptides of E-cadherin, N-cadherin, and alpha-dystroglycan were all efficiently modified by POMGnT1 (Figure 2-3). Therefore, digesting the cadherin proteins into peptides converted them into substrates for POMGNT1.

Analysis by Mass Spectrometry confirms cadherin peptides as substrates for POMGnT1

To orthogonally confirm that tryptic peptides of E-cadherin and N-cadherin are substrates for POMGnT1, we performed the same reaction with only non-radiolabelled UDP-

GlcNAc and analyzed the resulting products by mass spectrometry. Results confirmed that O-mannose glycans on the peptides were extended by POMGnT1 to a Man-GlcNAc disaccharide. Performing the same reaction with intact proteins yielded a 78% decrease in the number of mannose glycans modified by POMGNT1 (Figure 2-4). Use of tandem mass spectrometry allowed us to confirm the O-mannose glycans on peptides of E-cadherin and N-cadherin were extended with GlcNAc.

Discussion

The classical cadherins mediate calcium-dependent cell-cell adhesion through adherens junctions. These junctions are initiated by *trans* dimerization in which a cadherin on one cell dimerizes with a cadherin on another cell. The formation of *trans* adhesive homodimers in Type I classical cadherins occurs through exchange of the N-terminal β-strands on their EC1 domains (Vendome, J., Felsovalyi, K., et al. 2014). That is, the beta-strand of EC1 of one cadherin swaps places with the beta-strand of EC1 on a cadherin on a different cell. A conserved tryptophan residue anchors the adhesive interface by docking in a complementary hydrophobic pocket on the opposing strand.

This strand swap is preceded by a fast binding X-dimer intermediate. The X-diamer localizes around the linker region between EC1 and EC2. This interface functions as a kinetic intermediate in the formation of the strand-swapped diamer (Vendome, J., Felsovalyi, K., et al. 2014). *Trans* binding lowers the entropic penalty for the formation of *cis* interactions and is followed by *cis* clustering at a lateral interface formed between the base of an EC1 domain and the apex of a parallel (*cis*) EC2 domain. These form a lattice that is the structural basis for the adherens junctions (Harrison, O.J., Jin, X., et al. 2011). Interactions between cadherins on

proximal cells results in cell-cell adhesion, while *cis* interactions among cadherins on the same cell stabilize the junction (Harrison, O.J., Jin, X., et al. 2011).

Since Harrison et al.'s initial observation that O-linked glycans are present on cadherins, we and others have identified the glycans as non-elongated O-mannose. (Harrison, O.J., Jin, X., et al. 2011, Vester-Christensen, M.B., Halim, A., et al. 2013, Winterhalter, P.R., Lommel, M., et al. 2013). Expressed recombinant E-cadherin and N-cadherin contain only non-elongated hexose residues in wildtype cells. Removal of these residues with α-mannosidase confirms they are α-linked O-mannose residues.

O-mannosylation of cadherins is essential for E-cadherin mediated cell adhesion and its absence prevents preimplantation development of mouse embryos from the morula to the blastocyst stage (Lommel, M., Winterhalter, P.R., et al. 2013). O-mannosylation of cadherins is dependent on TMTC1—4 (Larsen, I.S.B., Narimatsu, Y., et al. 2017a). Biallelic mutations in TMTC3 lead to cobblestone lissencephaly, a condition also associated with improper O-mannosylation of alpha-dystroglycan (Jerber, J., Zaki, M.S., et al. 2016).

The O-mannose sites identified on E-cadherin were mapped to the B and G strands of the seven-stranded beta barrel structures of the EC domains. Loss of TMTC1/3 resulted in a loss of O-mannose on the G strand, while loss of TMTC1/2/3/4 resulted in a loss of all O-mannose on recombinantly expressed E-cadherin.

The finding that O-mannose on cadherins and plexins are M0 structures that are not elongated was an apparent incongruity with the previously described promiscuous behavior of POMGNT1 (Halmo, S.M., Singh, D., et al. 2017). However, the lack of substrate specificity by POMGNT1 had only been demonstrated in synthetic peptides containing sequences from the mucin-like region of alpha-dystroglycan. It was possible that testing the specificity of

POMGNT1 only against synthetic peptides derived from a known substrate of the enzyme led to over extrapolation of its promiscuity. The work in this report demonstrated that POMGNT1 was, in fact, able to modify tryptic peptides derived from E-cadherin and N-cadherin proteins expressed in HEK 293 cells. We therefore confirmed the finding that POMGNT1 is a promiscuous enzyme for a variety of O-mannosylated peptides.

By contrast, the E-cadherin and N-cadherin proteins were not efficiently modified by POMGNT1 unless they were proteolytically digested. Alpha-dystroglycan expressed in HEK 293 cells was efficiently modified by POMGNT1 with or without prior proteolytic digestion. These findings are consistent with the findings that alpha-dystroglycan expressed in mammalian muscle and brain is heavily modified by POMGNT1 while cadherins are not.

We sought to determine how the O-mannose residues on cadherins escaped modification. We hypothesized that protein folding in the ER rendered the sites of O-mannose inaccessible to POMGnT1 residing in the cis-Golgi. Digestion of proteins allowed POMGNT1 to modify them efficiently. POMGnT1 modified a-DG in both protein and peptide format. Tryptic peptides of E-cadherin and N-cadherin are better POMGnT1 substrates than the undigested protein. To evaluate whether the failure of POMGnT1 to elongate the O-mannose residues was due to protein folding, we compared untreated protein against tryptic peptides of the digested proteins. For E-cadherin and N-cadherin, only the tryptic peptides were efficient POMGnT1 substrates whereas alpha-dystroglycan, expressed in a POMGNT1 KO, was effectively extended at both the protein and peptide level. Peptides from E-cadherin, N-cadherin, and alpha- dystroglycan were all found to be extended to the O-Man-GlcNAc disaccharide by POMGNT1.

The successful modification of peptides by POMGnT1 is in agreement with previous findings that POMGnT1 does not show substrate specificity with respect to amino acid sequence. This study expands the previous work showing POMGnT1 to be a promiscuous modifier of peptides derived from alpha-dystroglycan by showing that POMGnT1 modifies peptides from other proteins. This study also provides an explanation for the apparent incongruity of the promiscuous POMGnT1 failing to modify O-mannose residues on cadherin proteins. Folding of the nascent polypeptides occurs in the ER resulting in the Golgi-resident POMGnT1 only encountering folded proteins. Therefore, the folded structure allows O-mannose modified residues on cadherins to escape modification by POMGnT1.

Materials and Methods

Protein Expression

DNA constructs: N-Cadherin (amino acid residues 1-724; UniProt P19022) extracellular domains were synthesized by GeneArt (Thermo Fisher). They were placed in a pGEc2 expression construct provided by Kelley Moremen using the LR Clonase kit (Thermo Fisher) (Moremen, K.W., Ramiah, A., et al. 2018). The mouse ECAD construct was a kind gift from Laurence Shapiro (Harrison, O.J., Jin, X., et al. 2011).

Expression of recombinant proteins: HEK 293 Wt and HEK 293 POMGnT1 knockout cells were described previously (Larsen, I.S.B., Narimatsu, Y., et al. 2017b). Cells were grown in suspension with FreeStyle 293 Expression Medium (Thermo Fisher Scientific). One day prior to transfection cells were pelleted and resuspended to a concentration of 1.5×10^6 cells/mL in 30mL volume. Transfections were performed by combining 20ug DNA, 4mL FreeStyle media, and 60uL 1mg/mL PEI and allowing complexes to form for 10 minutes. The transfection

mixture was added drop-wise to the culture. 24 hours later, 600uL of 375mM valproic acid and 30mL additional FreeStyle media were added to transfected cells. Proteins were harvested six days post- transfection and purified via His-tag by binding to Ni-NTA His-Bind Resin (Millipore-Sigma). Proteins were eluted in 250mM imidazole and dialyzed into PBS utilizing D-tube Dialyzer MWCO 3.5kDa (Millipore-Sigma).

Radiolabel Transfer Assays: Protein substrate for the radiolabel transfer assay was prepared either by dialysis into 100mM MES pH 6.5 (untreated protein) or reduction with 5 mM DTT at 56°C for 60 minutes, alkylation with 5.5 mM iodoacetamide in the dark at room temperature for 45 minutes, and digestion with sequencing grade modified trypsin at 37°C overnight (tryptic peptides). The tryptic digests were cleaned by reverse-phase separation using C18 Sep-Pak macro-spin columns, dried, and resuspended in 100mM MES pH 6.5 (tryptic peptides).

The radiometric assays were carried out in reactions containing 100mM MES pH 6.5, 4mM MnCl2, 1mM UDP-GlcNAc mixed with 10nM 3H-labeled UDP-GlcNAc and 0.2 nmol recombinant protein or peptides. Reactions were incubated for 1 hour at 37°C, then quenched by addition of 5uL 5% TFA. Unbound ^{3}H-labeled UDP-GlcNAc was removed by reverse-phase separation using C18 Sep-Pak micro-spin columns (Nest Group) by loading and washing with 0.1% formic acid followed by 0.1% formic acid with 4% acetonitrile. Scintillation fluid (Thermo Fisher Scientific) was loaded onto columns and disintegrations per minute (DPM) from the entire column were counted using a liquid scintillation counter (Beckman Coulter) as described previously (Halmo, S.M., Singh, D., et al. 2017).

Data Analysis for Radiolabel Transfer Assays: The data presented represent the average of three independent experiments. DPM were used to determine the amount of [^{3}H]-GlcNAc incorporated onto the glycopeptides. This amount was corrected by subtracting

background from a no-substrate-containing control reaction (typically 100-200 DPM). The nmol of GlcNAc transferred per hour, per mg of POMGnT1 were calculated and reported as a percentage of the amount transferred by the trypsin-digested samples. A280 readings of trypsin-digested samples and untreated samples were used to correct for sample loss during preparation.

Mannosidase treatment of HEK Wt expressed proteins:

50ug of each purified protein was trypsin digested. Tryptic peptides were treated with PNGaseF overnight in 50mM sodium phosphate pH 7.7. Peptides were separated from released N-glycans by reverse-phase separation using C18 Sep- Pak microspin columns (Nest Group). Dried peptides were re-suspended in 100mM ammonium acetate pH 4.5. 10 ug of peptide (as measured by A280) was digested with mannosidase (Millipore- Sigma) for 4 hours at 37°C in 100mM ammonium acetate pH 4.5 with 2mM $ZnCl_2$. Another 10ug of peptide was used as a control and not treated with mannosidase. Peptides were de-salted by reverse-phase separation using C18 Sep-Pak microspin columns (Nest Group). Dried peptides were stored at -20°C until analysis by mass spectrometry.

Analysis by mass spectrometry

Peptides were reconstituted in 3 µL 5% solvent B (0.1% formic acid in 80% acetonitrile) and 57 µL solvent A (0.1% formic acid), separated on an Acclaim PepMap RSLC C18 column (75um x 15cm), and eluted into the nano-electrospray ion source of an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) with a 180-minute linear gradient consisting of 1-99% solvent B over 180 minutes at a flow rate of 200 nL/min. Spray voltage was set to 2.2 kV and the temperature of the heated capillary was set to 280°C. Full MS scans

were acquired from m/z 300 to 2000 at 60k resolution, and MS2 scans following high energy collision dissociation stepped at 25% +/- 10, were collected in the ion trap for the most intense ions in the Top-Speed mode with a 3 second cycle using Fusion instrument software (v2.0, Thermo Fisher Scientific).

The resulting spectra were analyzed using Byonic (v2.6.46, Protein Metrics Inc) with precursor mass tolerance of 20 ppm and HCD fragment mass tolerance of 20 ppm. Cleavage sites were C- terminal R and K, fully specific, allowing for up to 2 missed cleavages. Modifications allowed were carbamidomethyl (+57.021464) @C fixed, oxidation (+15.994915) @M, Hexose (+162.052824) @ S or T, HexNAc-Hex (+365.1322) @ S or T, and HexNAc-HexNAc-Hex (+568.2116) @ S or T. Identified spectra were confirmed manually by viewing with Xcalibur Qual Browser (2.2 SP1 build 48, Thermo Fisher Scientific).

References

Abbott KL, Matthews RT, Pierce M. 2008. Receptor Tyrosine Phosphatase β (RPTPβ) Activity and Signaling Are Attenuated by Glycosylation and Subsequent Cell Surface Galectin-1 Binding. *Journal of Biological Chemistry*, 283:33026-33035.

Bleckmann C, Geyer H, Lieberoth A, Splittstoesser F, Liu Y, Feizi T, Schachner M, Kleene R, Reinhold V, Geyer R. 2009. O-glycosylation pattern of CD24 from mouse brain. *Biological chemistry*, 390:627-645.

Chai W, Yuen CT, Kogelberg H, Carruthers RA, Margolis RU, Feizi T, Lawson AM. 1999. High prevalence of 2-mono- and 2,6-di-substituted manol-terminating sequences among O-glycans released from brain glycopeptides by reductive alkaline hydrolysis. *European journal of biochemistry*, 263:879-888.

Dwyer CA, Baker E, Hu H, Matthews RT. 2012. RPTPzeta/phosphacan is abnormally glycosylated in a model of muscle-eye-brain disease lacking functional POMGnT1. *Neuroscience*, 220:47-61.

Finne J, Krusius T, Margolis RK, Margolis RU. 1979. Novel mannitol-containing oligosaccharides obtained by mild alkaline borohydride treatment of a chondroitin sulfate proteoglycan from brain. *The Journal of biological chemistry*, 254:10295-10300.

Halmo SM, Singh D, Patel S, Wang S, Edlin M, Boons GJ, Moremen KW, Live D, Wells L. 2017. Protein O-Linked Mannose beta-1,4-N-Acetylglucosaminyl-transferase 2 (POMGNT2) Is a Gatekeeper Enzyme for Functional Glycosylation of alpha-Dystroglycan. *The Journal of biological chemistry*, 292:2101-2109.

Harrison OJ, Jin X, Hong S, Bahna F, Ahlsen G, Brasch J, Wu Y, Vendome J, Felsovalyi K, Hampton CM, *et al.* 2011. The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. *Structure (London, England : 1993)*, 19:244-256.

Harrison R, Hitchen PG, Panico M, Morris HR, Mekhaiel D, Pleass RJ, Dell A, Hewitt JE, Haslam SM. 2012. Glycoproteomic characterization of recombinant mouse alpha-dystroglycan. *Glycobiology*, 22:662-675.

Jerber J, Zaki MS, Al-Aama JY, Rosti RO, Ben-Omran T, Dikoglu E, Silhavy JL, Caglar C, Musaev D, Albrecht B, *et al.* 2016. Biallelic Mutations in TMTC3, Encoding a Transmembrane and TPR-Containing Protein, Lead to Cobblestone Lissencephaly. *American journal of human genetics*, 99:1181-1189.

Larsen ISB, Narimatsu Y, Joshi HJ, Siukstaite L, Harrison OJ, Brasch J, Goodman KM, Hansen L, Shapiro L, Honig B, *et al.* 2017a. Discovery of an O-mannosylation pathway selectively serving cadherins and protocadherins. *Proceedings of the National Academy of Sciences*, 114:11163-11168.

Larsen ISB, Narimatsu Y, Joshi HJ, Yang Z, Harrison OJ, Brasch J, Shapiro L, Honig B, Vakhrushev SY, Clausen H, *et al.* 2017b. Mammalian O-mannosylation of cadherins and plexins is independent of protein O-mannosyltransferases 1 and 2. *The Journal of biological chemistry*, 292:11586-11598.

Lommel M, Strahl S. 2009. Protein O-mannosylation: conserved from bacteria to humans. *Glycobiology*, 19:816-828.

Lommel M, Winterhalter PR, Willer T, Dahlhoff M, Schneider MR, Bartels MF, Renner-Muller I, Ruppert T, Wolf E, Strahl S. 2013. Protein O-mannosylation is crucial for E-cadherin-mediated cell adhesion. *Proc Natl Acad Sci U S A*, 110:21024-21029.

Manya H, Endo T. 2017. Glycosylation with ribitol-phosphate in mammals: New insights into the O-mannosyl glycan. *Biochimica et biophysica acta. General subjects*, 1861:2462-2472.

Moremen KW, Ramiah A, Stuart M, Steel J, Meng L, Forouhar F, Moniz HA, Gahlay G, Gao Z, Chapla D, *et al.* 2018. Expression system for structural and functional studies of human glycosylation enzymes. *Nature chemical biology*, 14:156-162.

Sheikh MO, Halmo SM, Wells L. 2017. Recent advancements in understanding mammalian O-mannosylation. *Glycobiology*, 27:806-819.

Stalnaker SH, Aoki K, Lim JM, Porterfield M, Liu M, Satz JS, Buskirk S, Xiong Y, Zhang P, Campbell KP, *et al.* 2011. Glycomic analyses of mouse models of congenital muscular dystrophy. *The Journal of biological chemistry*, 286:21180-21190.

Stalnaker SH, Hashmi S, Lim JM, Aoki K, Porterfield M, Gutierrez-Sanchez G, Wheeler J, Ervasti JM, Bergmann C, Tiemeyer M, *et al.* 2010. Site mapping and characterization of O-glycan structures on alpha-dystroglycan isolated from rabbit skeletal muscle. *The Journal of biological chemistry*, 285:24882-24891.

Vendome J, Felsovalyi K, Song H, Yang Z, Jin X, Brasch J, Harrison OJ, Ahlsen G, Bahna F, Kaczynska A, *et al.* 2014. Structural and energetic determinants of adhesive binding specificity in type I cadherins. *Proc Natl Acad Sci U S A*, 111:E4175-4184.

Vester-Christensen MB, Halim A, Joshi HJ, Steentoft C, Bennett EP, Levery SB, Vakhrushev SY, Clausen H. 2013. Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. *Proceedings of the National Academy of Sciences*, 110:21018-21023.

Wells L. 2013. The o-mannosylation pathway: glycosyltransferases and proteins implicated in congenital muscular dystrophy. *The Journal of biological chemistry*, 288:6930-6935.

Winterhalter PR, Lommel M, Ruppert T, Strahl S. 2013. O-glycosylation of the non-canonical T-cadherin from rabbit skeletal muscle by single mannose residues. *FEBS letters*, 587:3715-3721.

Yoshida-Moriguchi T, Campbell KP. 2015. Matriglycan: a novel polysaccharide that links dystroglycan to the basement membrane. *Glycobiology,* 25:702-713.

Yoshida-Moriguchi T, Yu L, Stalnaker SH, Davis S, Kunz S, Madson M, Oldstone MB, Schachter H, Wells L, Campbell KP. 2010. O-mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. *Science (New York, N.Y.)*, 327:88-92.

Figure 2-1A

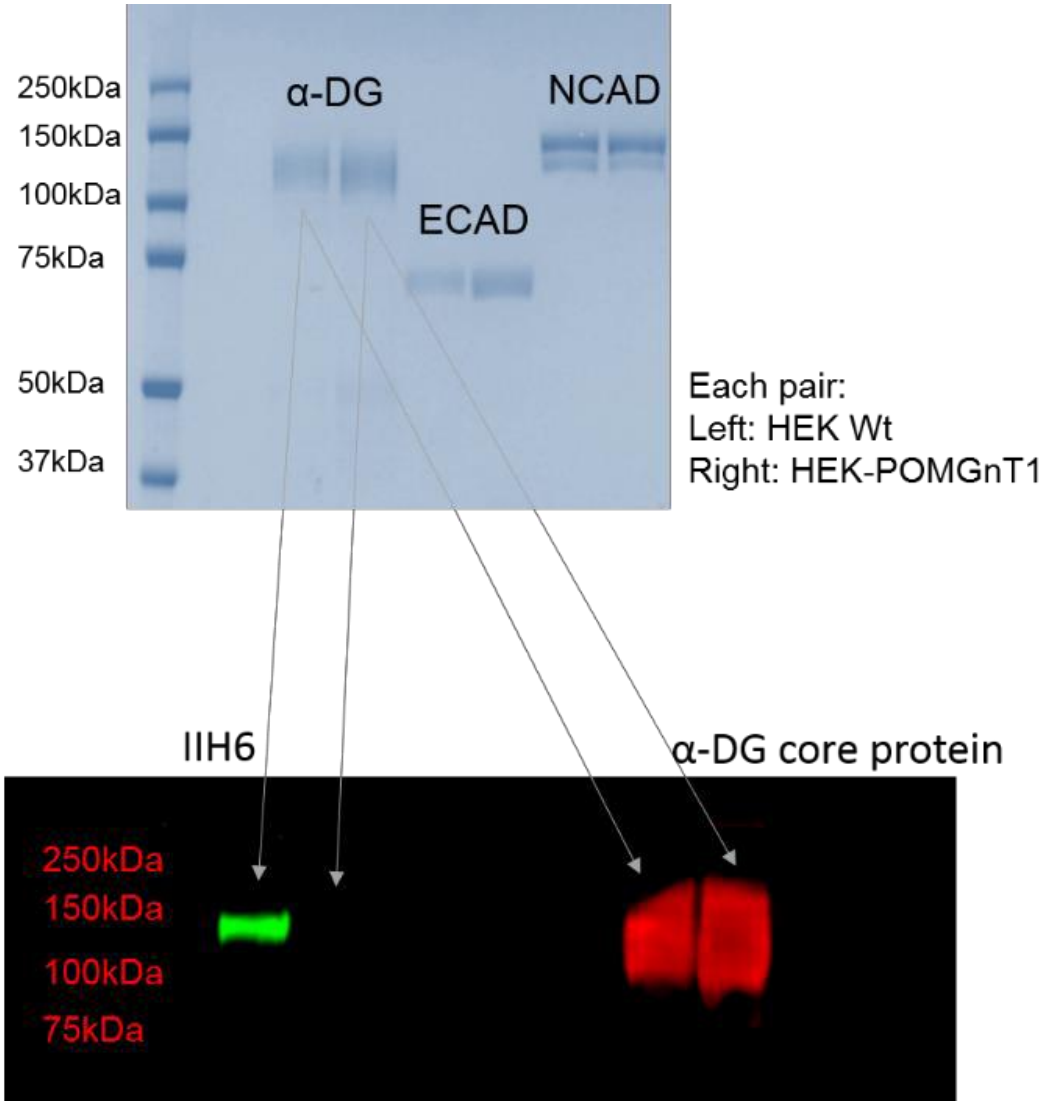

250kDa
150kDa
100kDa
75kDa
50kDa
37kDa
α-DG
NCAD
ECAD
Each pair:
Left: HEK Wt
Right: HEK-POMGnT1
IIH6
α-DG core protein
250kDa
150kDa
100kDa
75kDa

Figure 2-1B

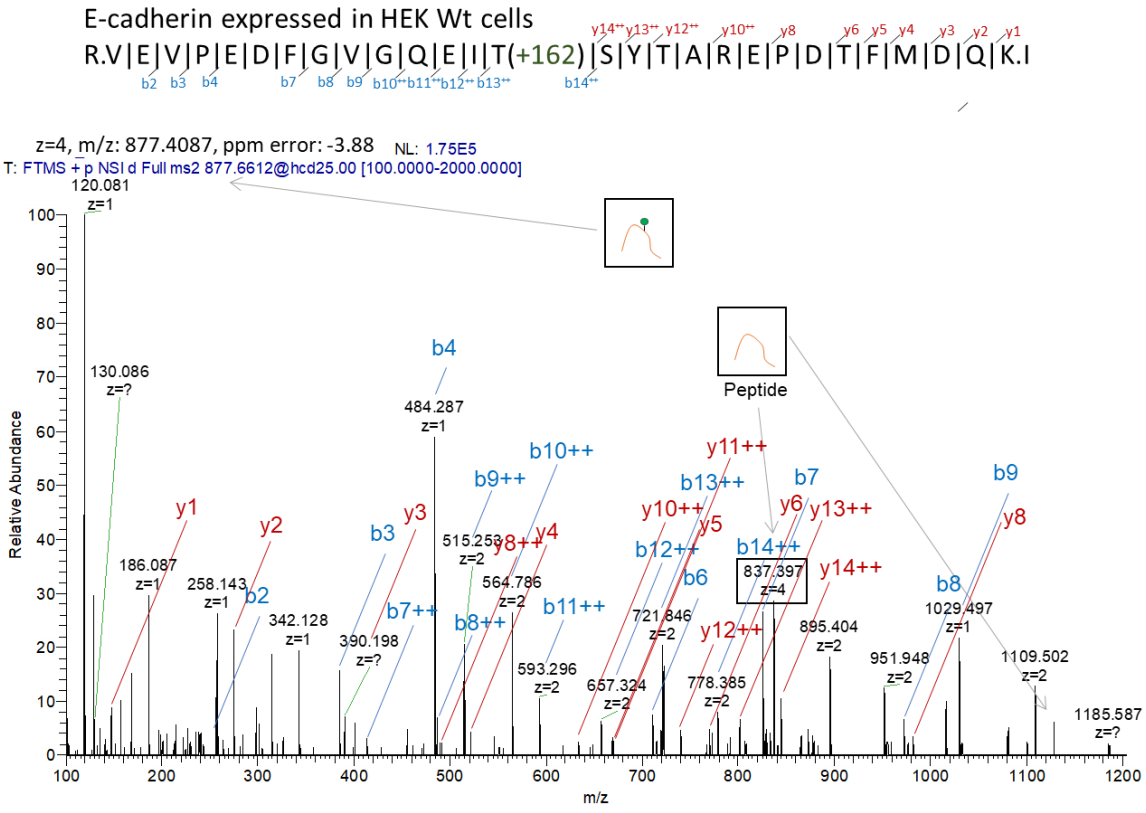

E-cadherin expressed in HEK Wt cells
R.V|E|V|P|E|D|F|G|V|G|Q|E|I|T(+162)|S|Y|T|A|R|E|P|D|T|F|M|D|Q|K.I
z=4, m/z: 877.4087, ppm error: -3.88 NL: 1.75E5
T: FTMS + p NSI d Full ms2 877.6612@hcd25.00 [100.0000-2000.0000]
Peptide
Relative Abundance
m/z
120.081 z=1
130.086 z=?
186.087 z=1
258.143 z=1
342.128 z=1
390.198 z=?
484.287 z=1
515.253 z=2
564.786 z=2
593.296 z=2
657.324 z=2
721.846 z=2
778.385 z=2
837.397 z=4
895.404 z=2
951.948 z=2
1029.497 z=1
1109.502 z=2
1185.587 z=?

Figure 2-1C

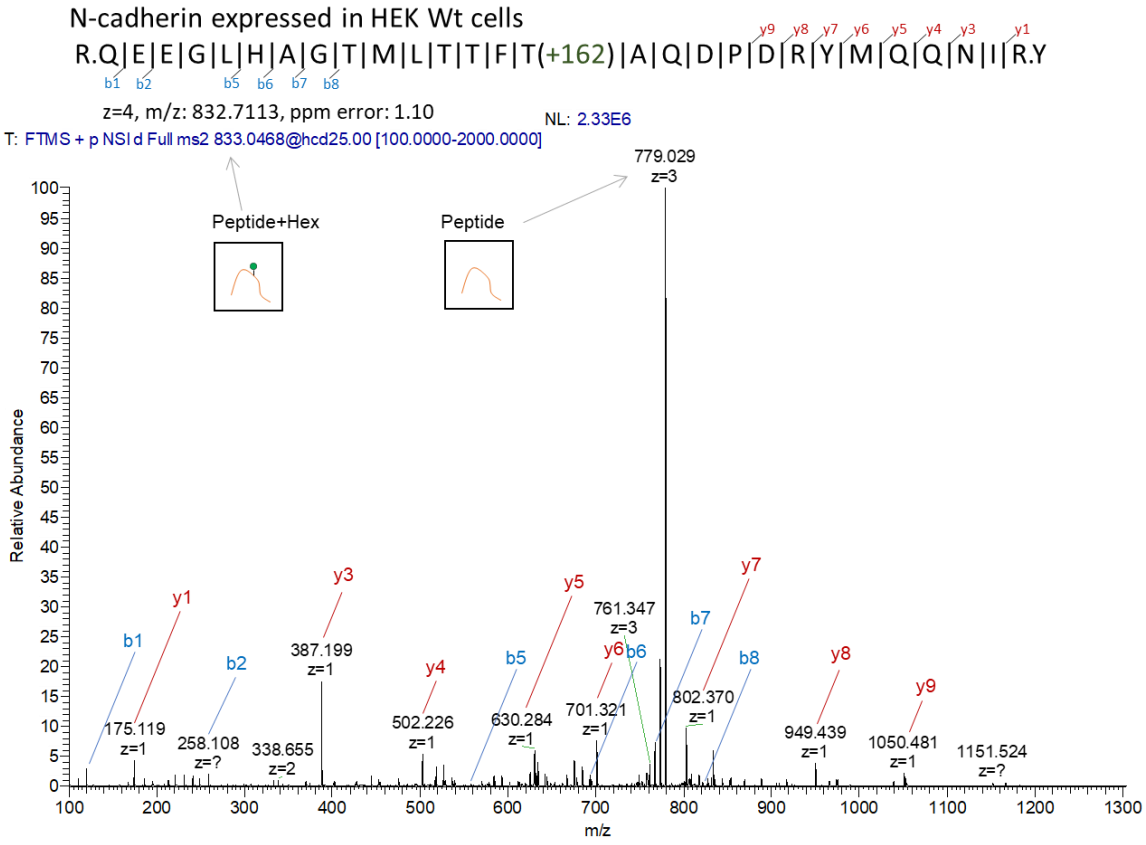

N-cadherin expressed in HEK Wt cells
R.Q|E|E|G|L|H|A|G|T|M|L|T|T|F|T(+162)|A|Q|D|P|D|R|Y|M|Q|Q|N|I|R.Y
b1 b2 b5 b6 b7 b8
y9 y8 y7 y6 y5 y4 y3 y1
z=4, m/z: 832.7113, ppm error: 1.10
NL: 2.33E6
T: FTMS + p NSI d Full ms2 833.0468@hcd25.00 [100.0000-2000.0000]
Peptide+Hex
Peptide
779.029
z=3
Relative Abundance
m/z
175.119
z=1
258.108
z=?
338.655
z=2
387.199
z=1
502.226
z=1
630.284
z=1
701.321
z=1
761.347
z=3
802.370
z=1
949.439
z=1
1050.481
z=1
1151.524
z=?

Figure 2-1D

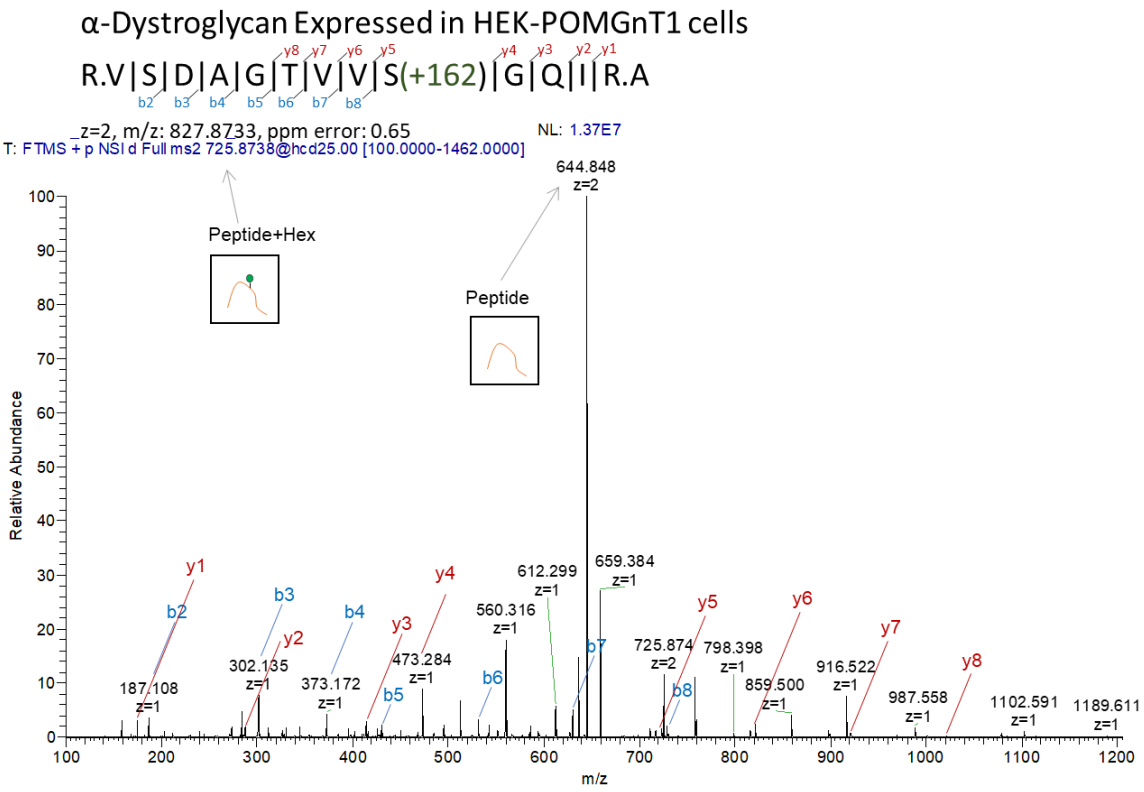

α-Dystroglycan Expressed in HEK-POMGnT1 cells
R.V|S|D|A|G|T|V|V|S(+162)|G|Q|I|R.A
b2 b3 b4 b5 b6 b7 b8
y8 y7 y6 y5 y4 y3 y2 y1
_z=2, m/z: 827.8733, ppm error: 0.65
NL: 1.37E7
T: FTMS + p NSI d Full ms2 725.8738@hcd25.00 [100.0000-1462.0000]
Peptide+Hex
Peptide
644.848 z=2
659.384 z=1
612.299 z=1
560.316 z=1
473.284 z=1
373.172 z=1
302.135 z=1
187.108 z=1
725.874 z=2
798.398 z=1
859.500 z=1
916.522 z=1
987.558 z=1
1102.591 z=1
1189.611 z=1
Relative Abundance
m/z

Figure 2-1E

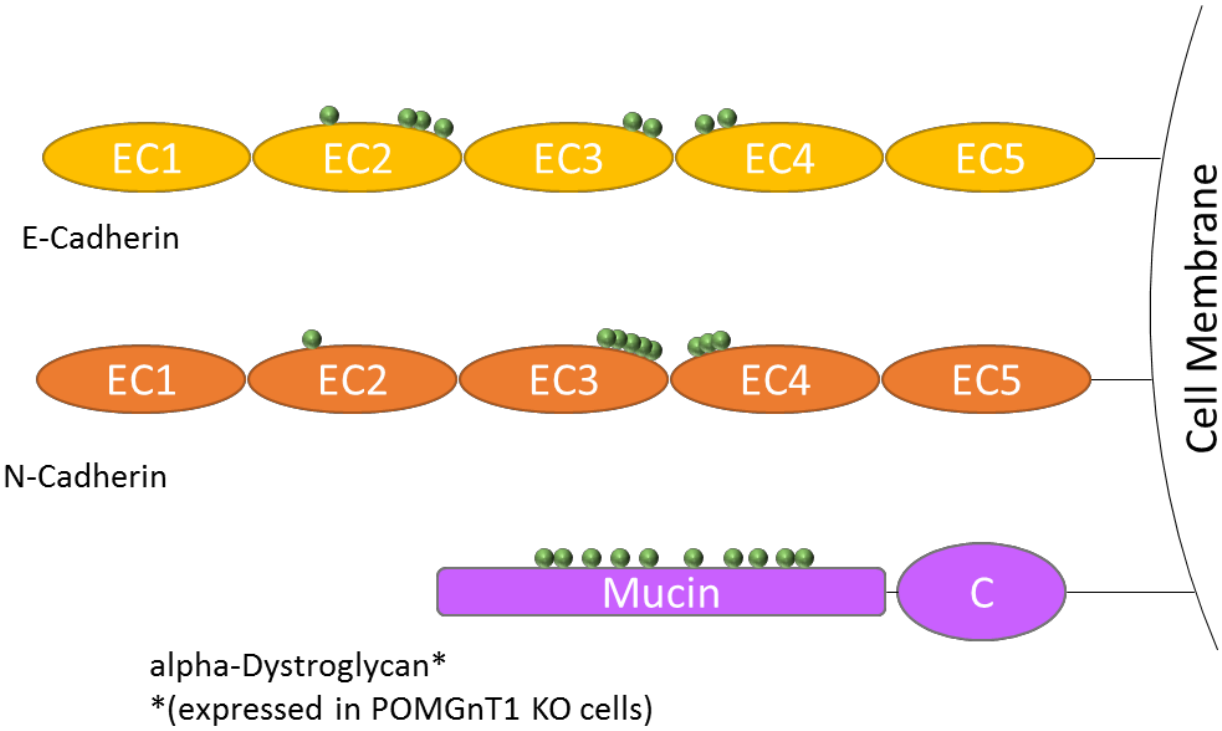

EC1
EC2
EC3
EC4
EC5
E-Cadherin
EC1
EC2
EC3
EC4
EC5
N-Cadherin
Mucin
C
alpha-Dystroglycan*
*(expressed in POMGnT1 KO cells)
Cell Membrane

Figure 2-1F

E-Cadherin

EC1:DWVIPPISCPENEKGEFPKNLVQIKSNRDKETKVFYSITGQGADKPPVGVFIIERETGWLKVTQPLDREAIAKYILYSHAVSSNGEAVEDPMEIVITVTDQNDNRPEF
EC2:TQPVFEGFVAEGAVPGTSVMKVSATDADDDVNTYNAAIAY**T**IVSQDPELPHKNMFTVNRDTGVISVLTSGLDRESYP**T**Y**T**LVVQAADLQGEGL**S**TTAKAVITVKDINDNAPVF
EC3:NPSTYQGQVPENEVNARIATLKVTDDDAPNTPAWKAVYTVVNDPDQQFVVVTDPTTNDGILKTAKGLDFEAKQQYILHVRVENEEPFEGSLVPSTA**T**V**T**VDVVDVNEAPIF
EC4:MPAERRVEVPEDFGVGQEI**T**SY**T**AREPDTFMDQKITYRIWRDTANWLEINPETGAIFTRAEMDREDAEHVKNSTYVALIIATDDGSPIATGTGTLLLVLLDVNDNAP
EC5:IPEPRNMQFCQRNPQPHIITILDPDLPPNTSPFTAELTHGASVNWTIEYNDAAQESLILQPRKDLEIGEYKIHLKLADNQNKDQVTTLDVHVCDCEGTVNNCMKA

N-Cadherin

EC1: DWVIPPINLPENSRGPFPQELVRIRSDRDKNLSLRYSVTGPGADQPPTGIFIINPISGQLSVTKPLDREQIARFHLRAHAVDINGNQVENPIDIVINVIDMNDNRPEF
EC2: LHQVWNGTVPEGSKPGTYVMTV**T**AIDADDPNALNGMLRYRIVSQAPSTPSPNMFTINNETGDIITVAAGLDREKVQQYTLIIQATDMEGNPTYGLSNTATAVITVTDVNDNPPEF
EC3: TAMTFYGEVPENRVDIIVANLTVTDKDQPHTPAWNAVYRISGGDPTGRFAIQTDPNSNDGLVTVVKPIDFETNRMFVLTVAAENQVPLAKGIQHPPQ**ST**A**T**V**S**V**T**VIDVNENPYF
EC4: APNPKIIRQEEGLHAG**T**ML**T**TF**T**AQDPDRYMQQNIRYTKLSDPANWLKIDPVNGQITTIAVLDRESPNVKNNIYNATFLASDNGIPPMSGTGTLQIYLLDINDNAP
EC5: QVLPQEAETCETPDPNSINITALDYDIDPNAGPFAFDLPLSPVTIKRNWTITRLNGDFAQLNLKIKFLEAGIYEVPIIITDSGNPPKSNISILRVKVCQCDSNGDCTDVDR

alpha-Dystroglycan

Mucin-like domain: ATPTPVTAIGPPTTAIQEPPSRIVPTPTSPAIAPPTETMAPPVRDPVPGKPTVTTRTRGAIIQ**T**P**T**LGPIQP**T**RVSDAG**T**VV**S**GQIRATVTIPGYVEP**T**
AVATPPTTTTKKPRVSTPKPA**T**PSTD**S**SAT**TT**RRPTKKPRTPRPVPRVTTKAPITRLETASPPTRIRTTTS

Figure 2-1G

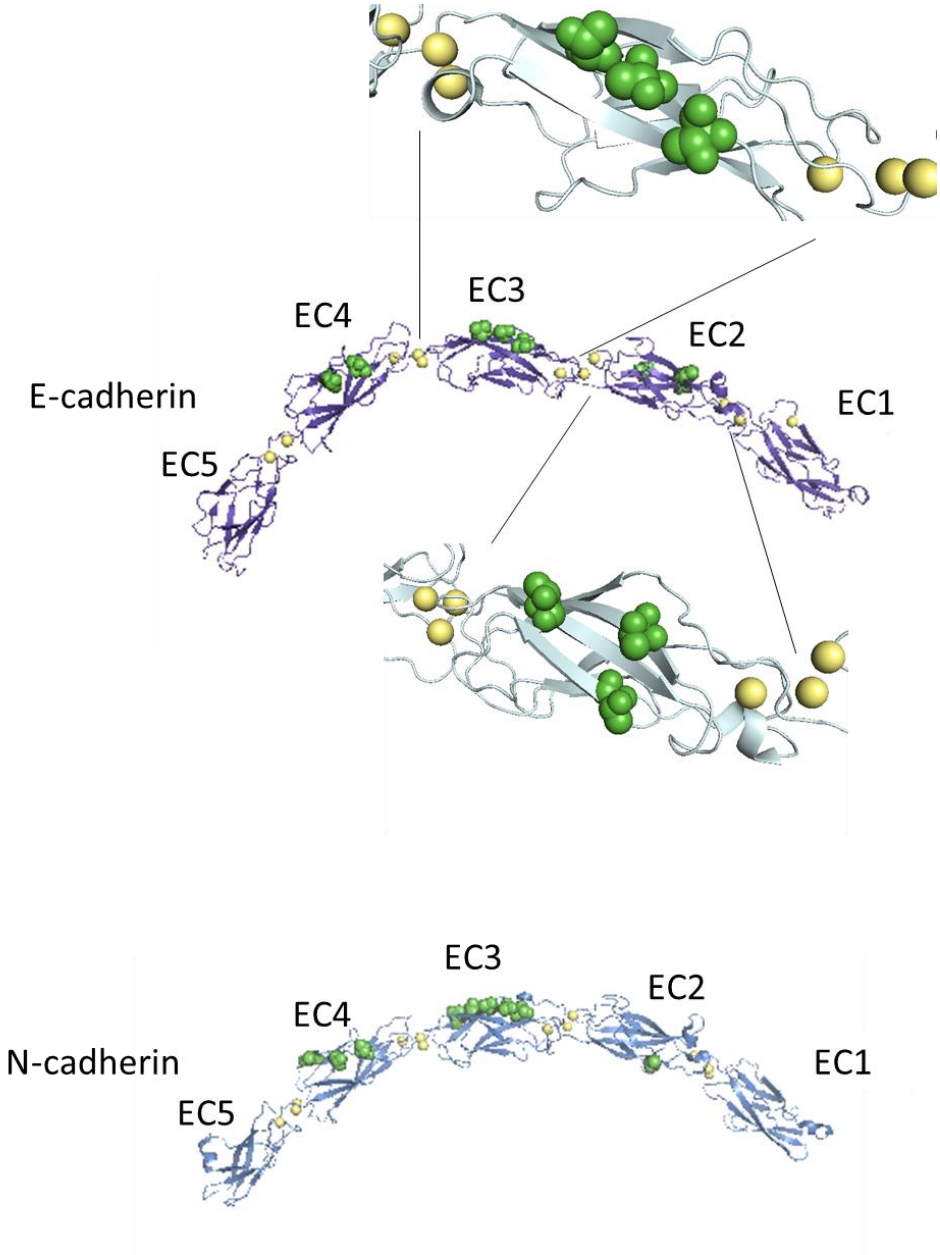


Figure 2-1. E-cadherin, N-cadherin, and alpha-dystroglycan proteins expressed in HEK Wt and HEK-POMGNT1 cells. (A) His-tag purified recombinant proteins expressed in HEK Wt (left side of each pair) and HEK-POMGNT1 KO (right). Western blot of α-DG expressed in HEK Wt and HEK-POMGNT1 KO. Both express α-DG core protein (red), but only the Wt-expressed protein expressed matriglycan to bind IIH6 antibody (green). (B) Annotated spectrum of HEK Wt- expressed E-cadherin modified by hexose at T508. (C) Annotated spectrum of HEK Wt-expressed N-cadherin modified by hexose at T520. (D) Annotated spectrum of HEK POMGNT1-expressed alpha-dystroglycan modified by hexose at S398. (E) Illustration of mapped sites of hexose modification. (F) Sites of hexose mapped to E-cadherin, N- cadherin, and alpha-dystroglycan. (G) Sites of O-mannose modification mapped onto E-cadherin (pdb-ID:3Q2V) and N-cadherin (pdb-ID: 3Q2W) ectodomains.

Figure 2-2A

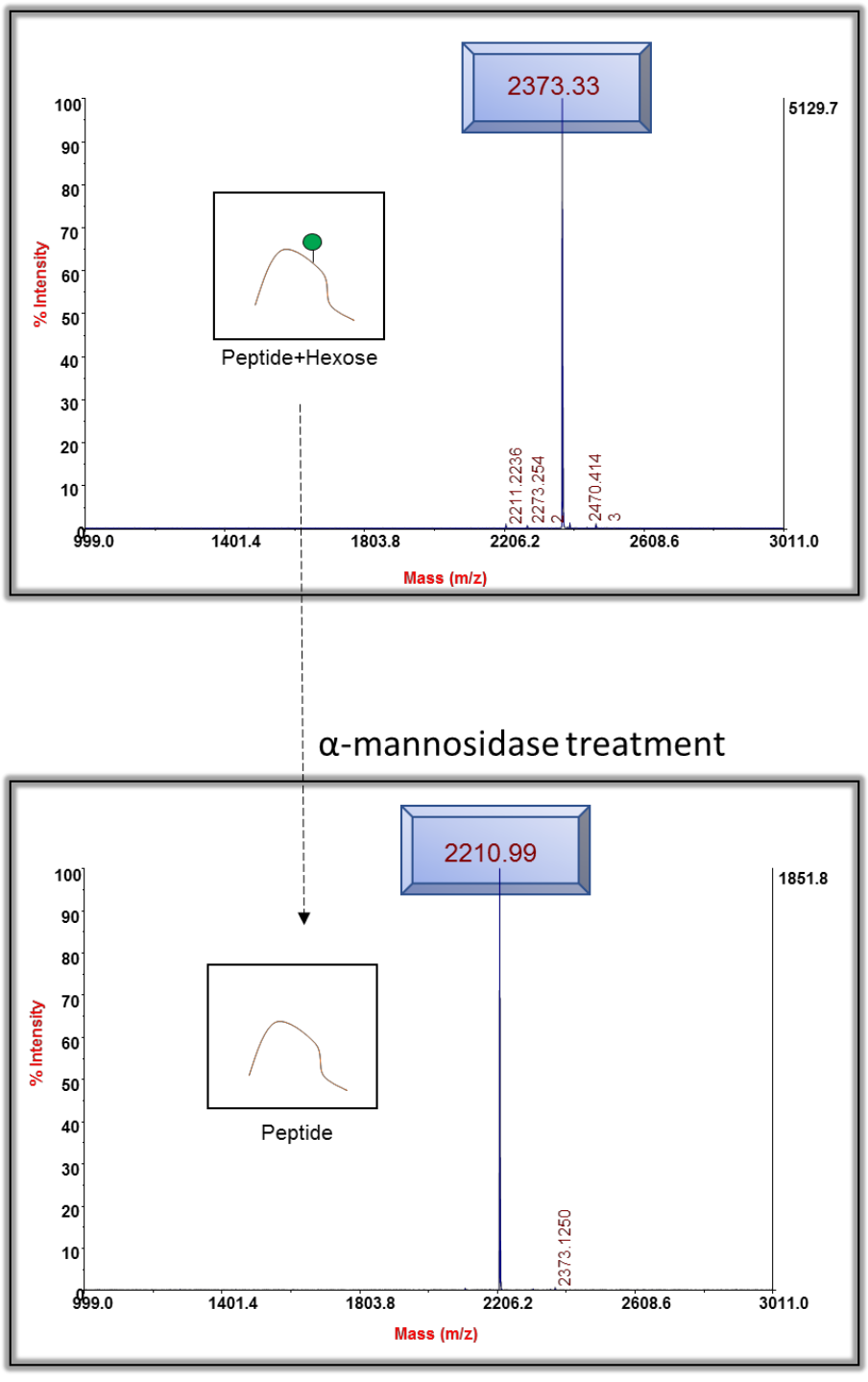

2373.33
5129.7
100
90
80
70
60
50
40
30
20
10
0
% Intensity
Peptide+Hexose
2211.2236
2273.254
2
2470.414
3
999.0
1401.4
1803.8
2206.2
2608.6
3011.0
Mass (m/z)
α-mannosidase treatment
2210.99
1851.8
Peptide
2373.1250
Mass (m/z)

Figure 2-2B

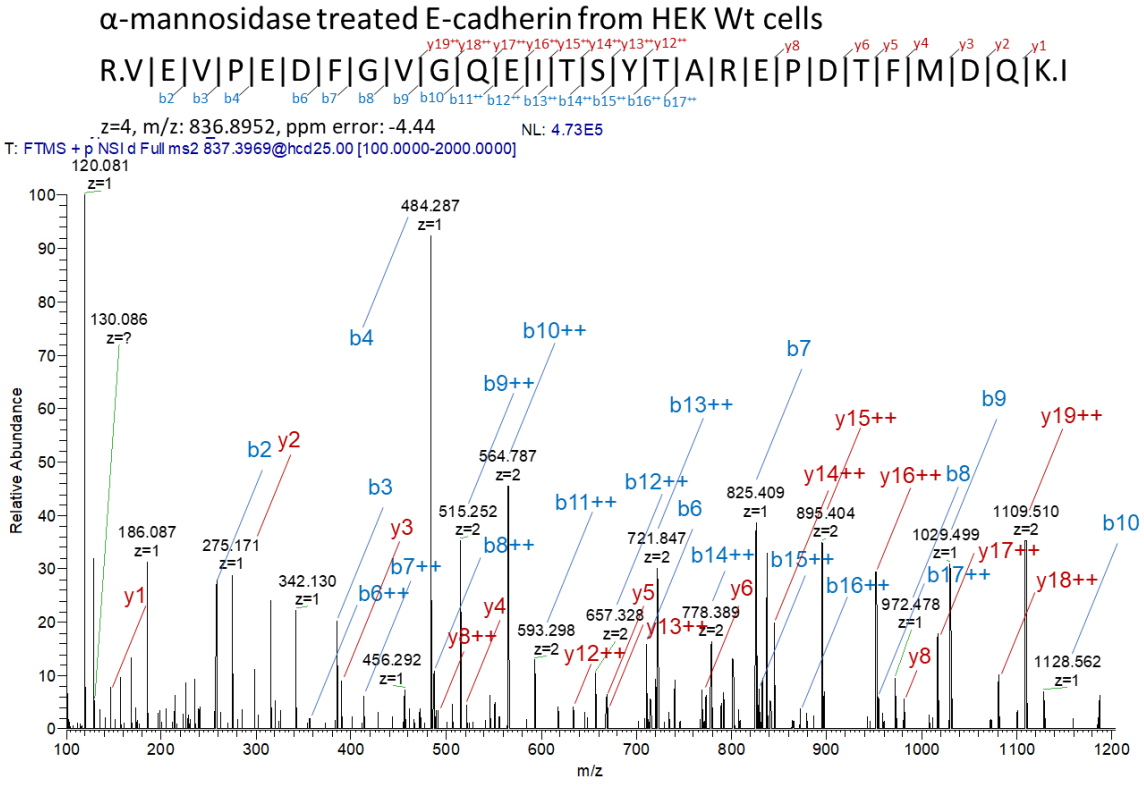

α-mannosidase treated E-cadherin from HEK Wt cells
R.V|E|V|P|E|D|F|G|V|G|Q|E|I|T|S|Y|T|A|R|E|P|D|T|F|M|D|Q|K.I
z=4, m/z: 836.8952, ppm error: -4.44
NL: 4.73E5
T: FTMS + p NSI d Full ms2 837.3969@hcd25.00 [100.0000-2000.0000]
Relative Abundance
m/z
120.081 z=1
484.287 z=1
130.086 z=?
564.787 z=2
515.252 z=2
825.409 z=1
895.404 z=2
1109.510 z=2
1029.499 z=1
186.087 z=1
275.171 z=1
342.130 z=1
721.847 z=2
657.328 z=2
778.389 z=2
972.478 z=1
593.298 z=2
456.292 z=1
1128.562 z=1
b4
b10++
b9++
b7
b13++
b9
y19++
y15++
b2
y2
b3
y3
b12++
b11++
b6
y14++
y16++
b8
b10
b8++
b14++
b15++
y17++
b7++
y1
y5
y6
b16++
b17++
y18++
b6++
y4
y13++
y8++
y12++
y8

Figure 2-2C

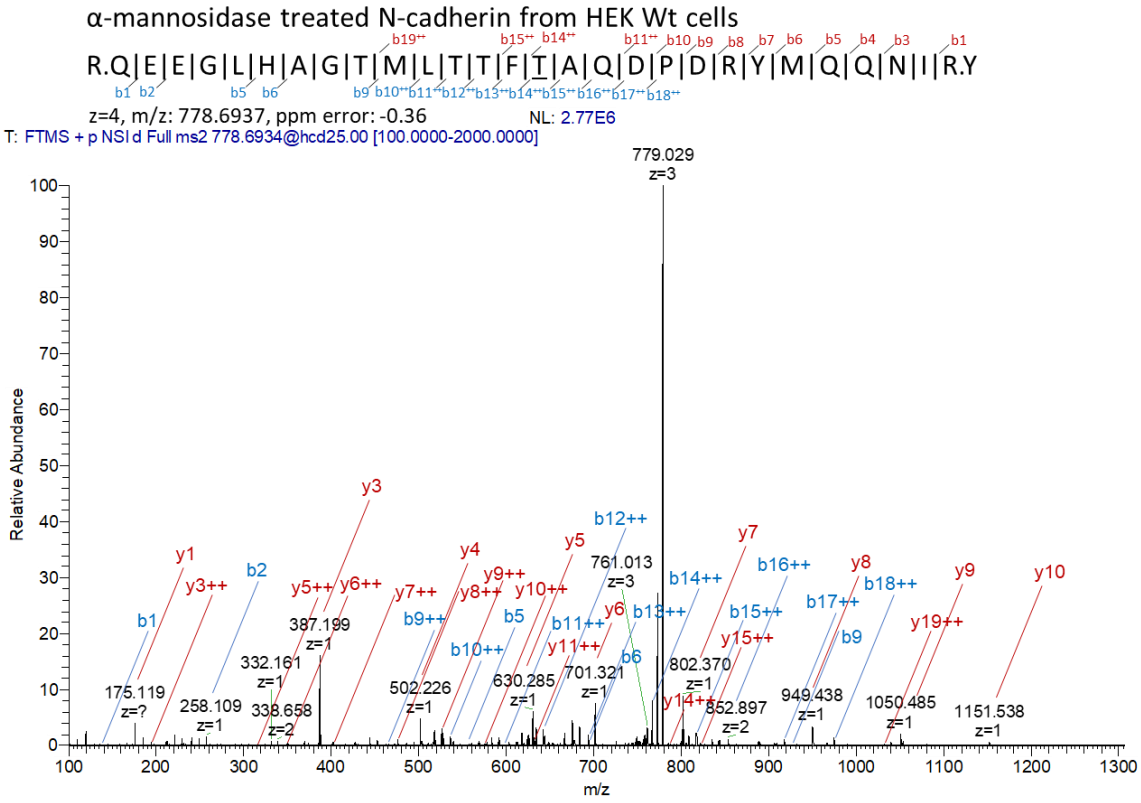

α-mannosidase treated N-cadherin from HEK Wt cells
R.Q|E|E|G|L|H|A|G|T|M|L|T|T|F|T|A|Q|D|P|D|R|Y|M|Q|Q|N|I|R.Y
z=4, m/z: 778.6937, ppm error: -0.36
NL: 2.77E6
T: FTMS + p NSI d Full ms2 778.6934@hcd25.00 [100.0000-2000.0000]
779.029
z=3
Relative Abundance
m/z

Figure 2-2D

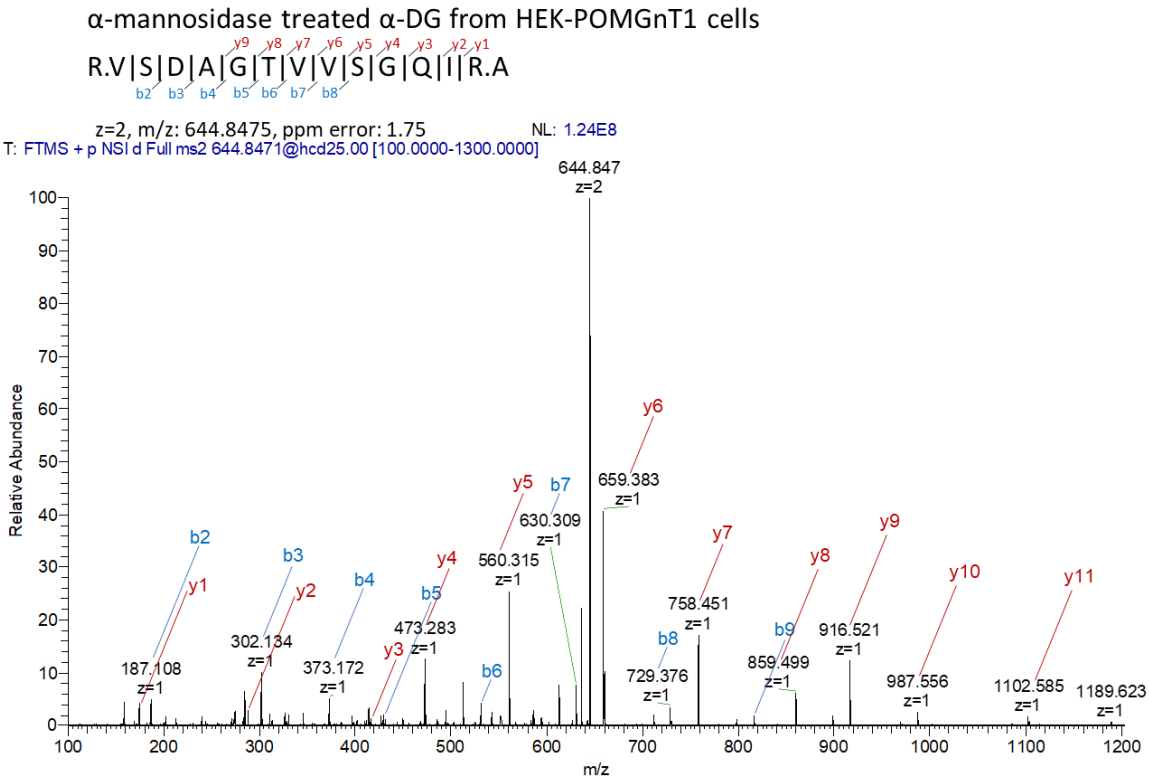

α-mannosidase treated α-DG from HEK-POMGnT1 cells
R.V|S|D|A|G|T|V|V|S|G|Q|I|R.A
y9 y8 y7 y6 y5 y4 y3 y2 y1
b2 b3 b4 b5 b6 b7 b8
z=2, m/z: 644.8475, ppm error: 1.75
NL: 1.24E8
T: FTMS + p NSI d Full ms2 644.8471@hcd25.00 [100.0000-1300.0000]
Relative Abundance
m/z
644.847 z=2
659.383 z=1
630.309 z=1
560.315 z=1
473.283 z=1
373.172 z=1
302.134 z=1
187.108 z=1
758.451 z=1
729.376 z=1
859.499 z=1
916.521 z=1
987.556 z=1
1102.585 z=1
1189.623 z=1

Figure 2-2E

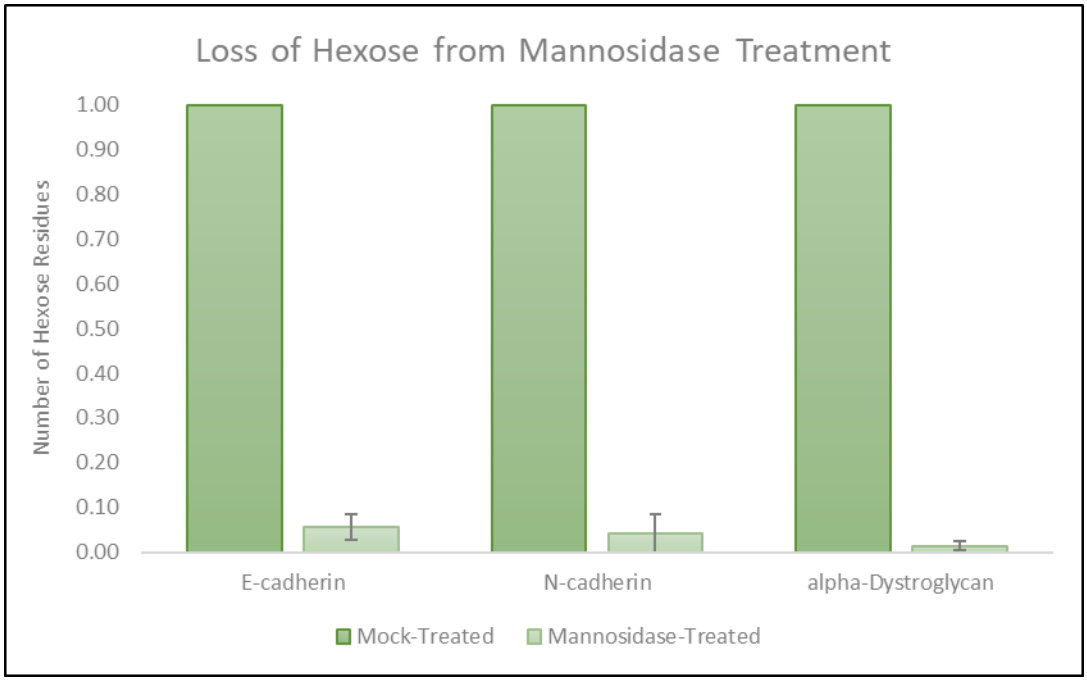

Loss of Hexose from Mannosidase Treatment
Number of Hexose Residues
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
E-cadherin
N-cadherin
alpha-Dystroglycan
Mock-Treated
Mannosidase-Treated

Figure 2-2. Hexose modification is lost during treatment with α-mannosidase. (A) Synthetic O- mannosylated peptide derived from residues 331-351 of alpha-dystroglycan (IQEPPSRIVPT(man)PTSPAIAPPT) before and after α-mannosidase treatment. (B) Annotated spectrum of HEK Wt-expressed E-cadherin without a hexose at T508. (C) Annotated spectrum of HEK Wt-expressed N-cadherin without a hexose. (D) Annotated spectrum of HEK POMGNT1-expressed alpha-dystroglycan without a hexose at S398. (E) Total number of hexoses detected after treatment with mannosidase, relative to the number of hexose detected in mock-treated samples.

Figure 2-3

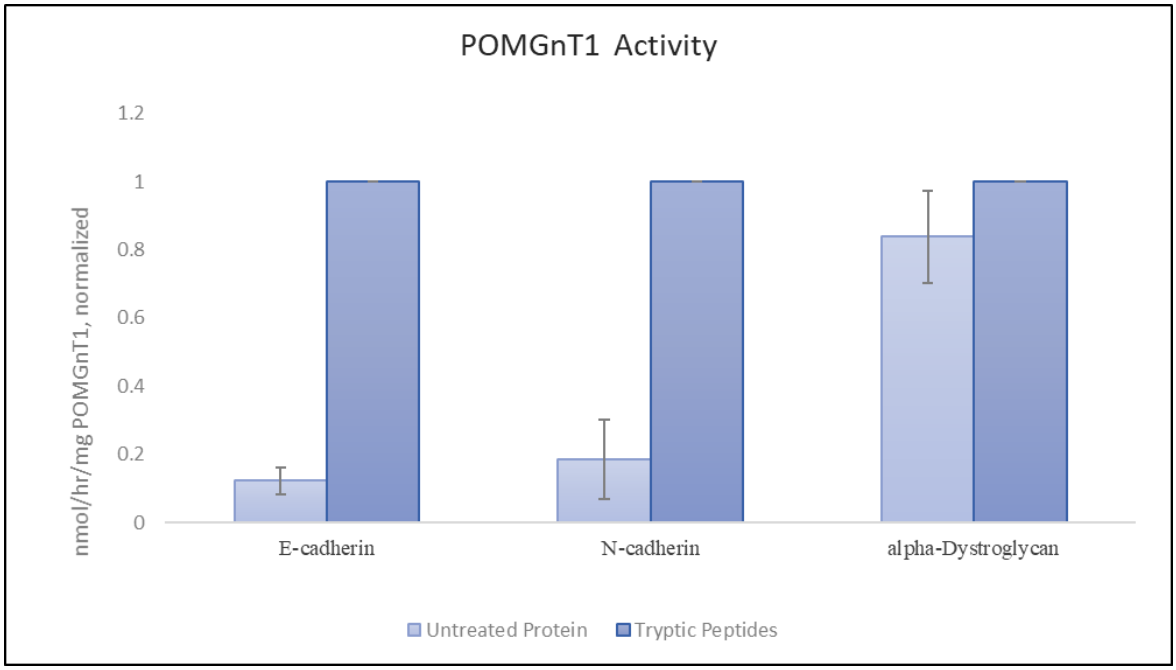

POMGnT1 Activity
nmol/hr/mg POMGnT1, normalized
1.2
1
0.8
0.6
0.4
0.2
0
E-cadherin
N-cadherin
alpha-Dystroglycan
Untreated Protein
Tryptic Peptides

Figure 2-3. Cadherins become POMGNT1 substrates after digestion into tryptic peptides. The rate of transfer of labeled GlcNAc in nmol of GlcNAc transferred/hr/mg of POMGnT1 as a percentage of the rate of transfer in the trypsin-digested samples. POMGnT1 efficiently transferred labelled GlcNAc to both untreated protein and tryptic peptides of alpha-dystroglycan. POMGnT1 was not able to efficiently transfer to E- cadherin or N-cadherin proteins but was able to modify the tryptic peptides of the proteins. The data represent the average of three independent experiments with the error bars reporting the standard deviation.

Figure 2-4A

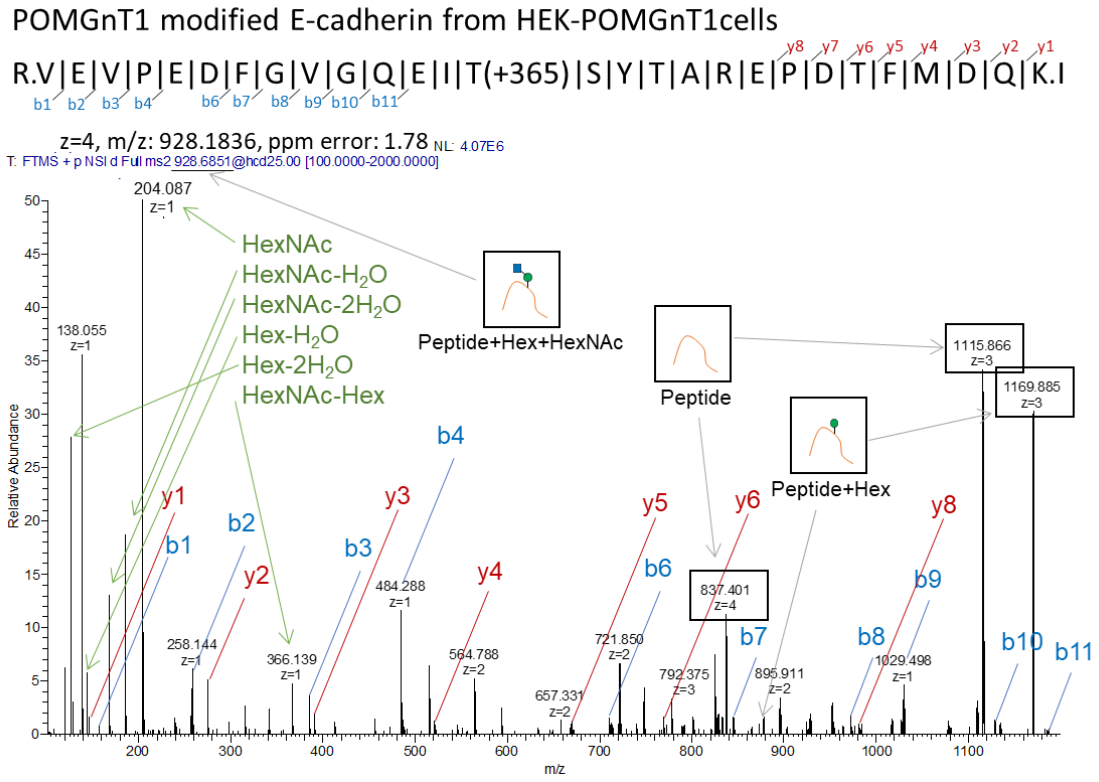

POMGnT1 modified E-cadherin from HEK-POMGnT1cells
R.V|E|V|P|E|D|F|G|V|G|Q|E|I|T(+365)|S|Y|T|A|R|E|P|D|T|F|M|D|Q|K.I
b1 b2 b3 b4 b6 b7 b8 b9 b10 b11
y8 y7 y6 y5 y4 y3 y2 y1
z=4, m/z: 928.1836, ppm error: 1.78 NL: 4.07E6
T: FTMS + p NSI d Full ms2 928.6851@hcd25.00 [100.0000-2000.0000]
204.087 z=1
138.055 z=1
HexNAc
HexNAc-H2O
HexNAc-2H2O
Hex-H2O
Hex-2H2O
HexNAc-Hex
Peptide+Hex+HexNAc
Peptide
Peptide+Hex
1115.866 z=3
1169.885 z=3
837.401 z=4
484.288 z=1
258.144 z=1
366.139 z=1
564.788 z=2
721.850 z=2
657.331 z=2
792.375 z=3
895.911 z=2
1029.498 z=1
Relative Abundance
m/z

Figure 2-4B

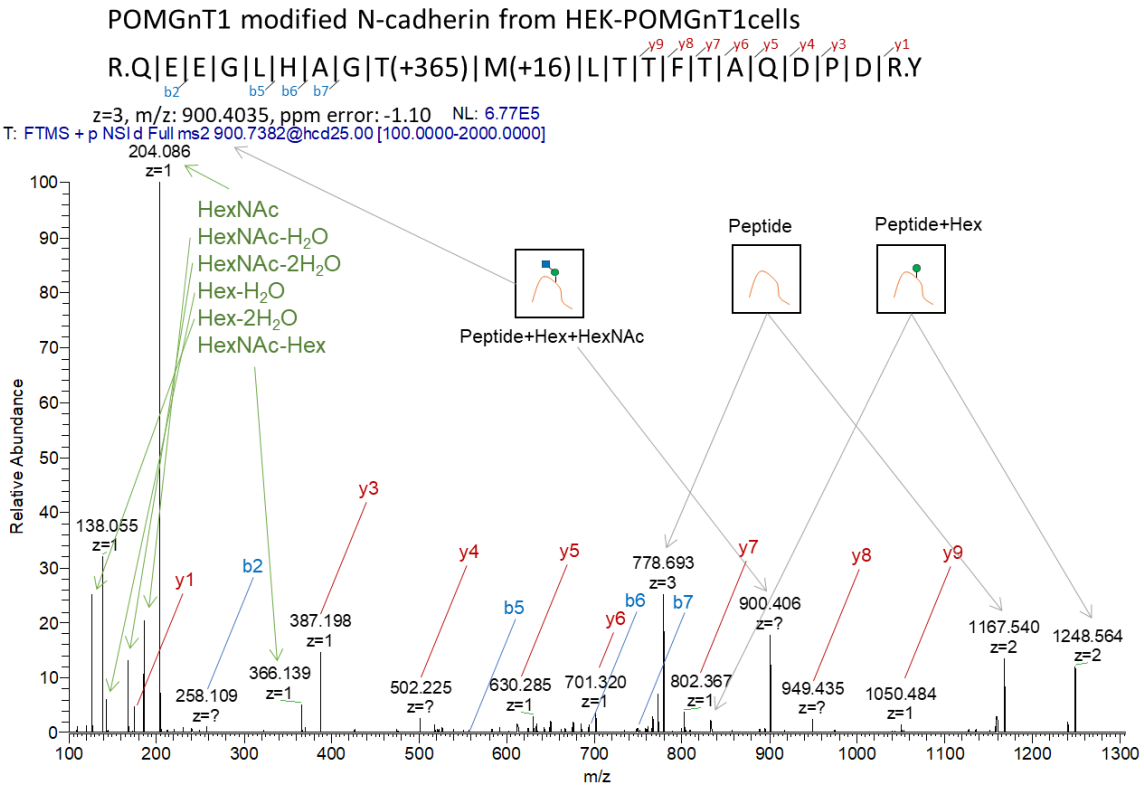

POMGnT1 modified N-cadherin from HEK-POMGnT1cells
R.Q|E|E|G|L|H|A|G|T(+365)|M(+16)|L|T|T|F|T|A|Q|D|P|D|R.Y
z=3, m/z: 900.4035, ppm error: -1.10 NL: 6.77E5
T: FTMS + p NSI d Full ms2 900.7382@hcd25.00 [100.0000-2000.0000]
HexNAc
HexNAc-H2O
HexNAc-2H2O
Hex-H2O
Hex-2H2O
HexNAc-Hex
Peptide+Hex+HexNAc
Peptide
Peptide+Hex
Relative Abundance
m/z
204.086 z=1
138.055 z=1
258.109 z=?
366.139 z=1
387.198 z=1
502.225 z=?
630.285 z=1
701.320 z=1
778.693 z=3
802.367 z=1
900.406 z=?
949.435 z=?
1050.484 z=1
1167.540 z=2
1248.564 z=2

Figure 2-4C

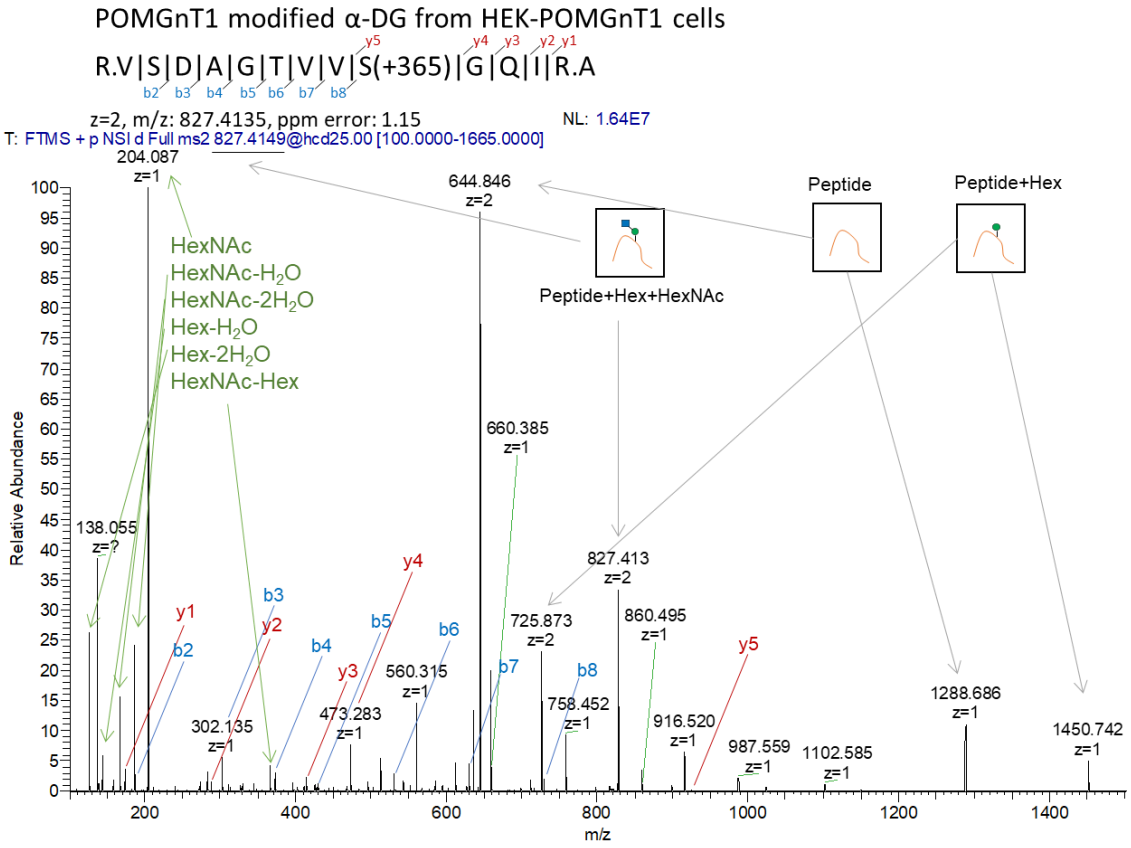

POMGnT1 modified α-DG from HEK-POMGnT1 cells
R.V|S|D|A|G|T|V|V|S(+365)|G|Q|I|R.A
z=2, m/z: 827.4135, ppm error: 1.15
NL: 1.64E7
T: FTMS + p NSI d Full ms2 827.4149@hcd25.00 [100.0000-1665.0000]
Peptide
Peptide+Hex
Peptide+Hex+HexNAc
HexNAc
HexNAc-H2O
HexNAc-2H2O
Hex-H2O
Hex-2H2O
HexNAc-Hex
Relative Abundance
m/z

Figure 2-4. A HexNAc-Hexose modification is detected on peptides after treatment with POMGNT1. (A) Annotated spectrum of POMGNT1-modfied E-cadherin peptide from protein expressed in HEK-POMGNT1 cells. (B) Annotated spectrum of POMGNT1-modfied N-cadherin peptide from protein expressed in HEK-POMGNT1 cells. (C) Annotated spectrum of POMGNT1-modfied alpha-dystroglycan peptide from protein expressed in HEK-POMGNT1 cells.

Table 2-1

Average number of Sites Modified by POMGNT1			
	Peptide	Protein	Percent Reduction
E-Cadherin	75±28	17±11	78%
N-Cadherin	210±111	47±31	78%
alpha-Dystroglycan	106±27	73±27	31%

Table 2-1. E-cadherin and N-cadherin samples treated with POMGNT1 as proteins had 78% fewer sites of POMGNT1 modification than samples treated with POMGNT1 as peptides. There was a 31% reduction in the number of POMGNT1 modified glycans in alpha-dystroglycan. The data represent the average of three independent experiments.

CHAPTER 3

LOSS OF POMGNT1 AND FUNCTIONAL GLYCOSYLATION OF ALPHA-DYSTROGLYCAN

Background

Muscle-Eye-Brain disease (MEB) is a rare congenital muscular dystrophy. MEB patients typically have muscle hypotonia and atrophy, eye conditions such as myopia and glaucoma, and brain abnormalities including hydrocephalus and cobblestone complex (Santavuori, P., Somer, H., et al. 1989). MEB is an autosomal recessive disorder caused by homozygous or compound heterozygous mutations in the *Pomgnt1* gene (Manya, H., Sakai, K., et al. 2003). The *Pomgnt1* gene encodes the enzyme protein O-mannose beta-1,2-N-acetylglucosaminyltransferase (POMGnT1).

MEB is part of a subset of congenital muscular dystrophies known as the secondary dystroglycanopathies. These disorders are characterized by muscle weakness, often with brain and eye abnormalities (Bertini, E., D'Amico, A., et al. 2011). The secondary dystroglycanopathies range from the mild Limb- Girdle muscular dystrophy, which lacks neurological involvement, to the Walker-Warburg Syndrome where severe symptoms including brain malformation are usually fatal within the first 2 years of life (Mercuri, E., Messina, S., et al. 2009).

The secondary dystroglycanopathies are named as such because they result not from defects in the dystroglycan peptide backbone, but from defects in the glycosylation of dystroglycan. The *Dag1* gene codes for the dystroglycan precursor which is proteolytically

cleaved into alpha- and beta-subunits (Barresi, R. and Campbell, K.P. 2006). The beta-subunit connects to dystrophin, spans the plasma membrane, and is non-covalently linked to the extracellular alpha-subunit (Holt, K.H., Crosbie, R.H., et al. 2000). The alpha-subunit (α-DG) has a mucin-like domain which is heavily modified by a great diversity of O-mannose and O-GalNAc initiated glycans (Dobson, C.M., Hempel, S.J., et al. 2013, Endo, T. 1999, Ervasti, J.M. and Campbell, K.P. 1993, Harrison, R., Hitchen, P.G., et al. 2012, Wells, L. 2013). α-DG is also the only protein known to be modified with a unique core M3 O-mannose initiated glycan which terminates in matriglycan, a repeating disaccharide of xylose and glucuronic acid (Ibraghimov-Beskrovnaya, O., Ervasti, J.M., et al. 1992, Yoshida-Moriguchi, T., Yu, L., et al. 2010). This glycan structure attaches via the matriglycan to LG- domain containing proteins in the extracellular matrix (Ervasti, J.M. and Campbell, K.P. 1991). Secondary dystroglycanopathies result from mutations which prevent assembly of this functional glycan.

Patients with secondary dystroglycanopathies have been found to have deleterious mutations in the enzymes POMT1, POMT2, POMGnT1, POMGnT2, B3GALNT2, POMK, FKTN, FKRP, RXYLT1 (TMEM5), B4GAT1, or LARGE1 (Meilleur, K.G., Zukosky, K., et al. 2014). Each of these enzymes has a described role in building the functional glycan with one glaring exception; POMGnT1. POMGnT1 has no known role in building the functional glycan. Nevertheless, cell lines and murine tissues from POMGnT1 knockouts show a loss of matriglycan as determined by the IIH6 antibody which recognizes the repeating disaccharide (Liu, J., Ball, S.L., et al. 2006).

One possible role for POMGnT1 in building the functional glycan is to recruit one of the necessary enzymes to the site of the nascent glycan. It has been suggested that the lectin domain (residues 97-250) of POMGnT1 interacts with fukutin to recruit fukutin to the site of

construction of the functional glycan. It has further been suggested that the catalytic activity of POMGnT1 is not necessary for the functional glycosylation of α-DG (Kuwabara, N., Manya, H., et al. 2016). The severe clinical symptoms of MEB patients with mutations in the catalytic domain of POMGnT1 seemed inconsistent with this proposal. We therefore sought to further determine what role POMGnT1 played in building the functional glycan.

To study the effects of loss of POMGnT1, we utilized both HEK293 POMGnT1 knockout cells and HEK 293 Simple Cells. Use of HEK293 POMGnT1 knockout cells allowed us to assess the ability of several variants of POMGnT1 to rescue expression of matriglycan. The HEK293 Simple Cells have both POMGnT1 and COSMC knocked out. COSMC is a chaperone essential for expression of active T-synthase, the enzyme that Gal-1,3 to O-linked GalNAc on mucin glycans. We carefully evaluated whether COSMC was necessary for matriglycan expression through a series of experiments. The HEK Simple Cells have only O-mannose or O-GalNAc where we would expect to see core M1 structures and mucin glycans. These simplified glycans allow us to use mass spectrometry to determine which sites have been modified on glycopeptides by allowing for the addition of either +162 for hexose or +204 for hexnac, rather than performing several complex searches to attempt to identify more complex glycan structures.

Results

Catalytically defective POMGnT1 enzyme is not able to rescue matriglycan expression

We sought to determine what domains of POMGnT1 were necessary for building the functional glycan. We utilized HEK-POMGnT1 to express variants of POMGnT1 and tested their reactivity to the IIH6 antibody which detects matriglycan in the functional glycan. As

expected, HEK Wt was IIH6 reactive while HEK-POMGnT1 was not (Figure 3-1, lanes 1 and 2). Full-length POMGnT1 with a FLAG tag was able to rescue IIH6 reactivity (Figure 3-1, lane 3). We tested *Pomgnt1* constructs containing a mutation in the DXD motif, the catalytic base, and a construct lacking the stem domain residues 97-250. We were not able to detect IIH6 reactivity from cells expressing any of the mutant POMGnT1 enzymes (Figure 3-1).

The core M3 structure in POMGnT1 KO cells contains a phospho-ribitol

We sought to determine where exactly the breakdown in assembly of the functional glycan occurred in the absence of POMGnT1. We utilized a soluble, truncated α-DG construct consisting of the N-terminal 340 amino acids of α-DG, previously described as αDG340. Furin

cleavage at residue 312, and tryptic cleavage after R337, yields a 28-mer which contains a site known to be modified by the functional glycan; T317/319 (Figure 3-2A).

We originally tried expressing αDG340 in HEK-POMGnT1 cells. However, having six sites which may or may not be occupied by O-mannose, the unknown mass of the partial core M3 glycan, or any combination of mucin glycans was too much complexity to allow for meaningful interpretation of the mass spectrometry data. This was partially solved by introducing T/S→A mutations at all but the T317/T319 sites. To further reduce unnecessary complexity, this aDG340 (TTTS→AAAA) construct was expressed in HEK Simple Cells in which both POMGnT1 and COSMC had been knocked out. This reduced the possible modifications at T317 and T319 to only O-mannose, O-GalNAc, or the partial core M3 structure.

We were able to identify scans containing the phospho-trisaccharide with one phospho- ribitol (Figure 3-2B). Therefore, it is possible for the product of fukutin to be made in a POMGnT1 knockout cell. However, the number of spectra identified to contain the M3 glycan assembled beyond the phospho-tri-saccharide was a small percentage (10%) of the number of M3 structures observed in total (Table 3-1). Separating the effects of over-expression of α-DG from the effects of loss of POMGnT1 will require additional controls. Determining the actual frequency of fukutin-modified α-DG in POMGnT1 knockout cells is further complicated by our inability to detect the fully-elaborated matriglycan structure due to its large mass.

Functional Glycosylation of aDG in POMGnT1 KO cells is not detectable

We considered that POMGnT1 mutants may be able to produce the functional glycan, but at a greatly reduced amount relative to wild type. To probe for the presence matriglycan on the surface of cells, we utilized immunofluorescent staining with the IIH6 antibody. Fluorescence was imaged on a Licor Odyssey Clx. For the positive control HAP1-Wt cells, both α-DG core protein and matriglycan were detected, as expected (Figure 3). For the negative control, HAP1-DAG1, neither a-DG nor matriglycan were detected, also as expected. The HAP1-POMGnT1 cells expressed a-DG, but not matriglycan (Figure 3).

The purpose of this experiment was to determine whether HAP1-POMGnT1 cells would express a very reduced amount of matriglycan, rather than none at all. The findings were that the amount of matriglycan detected on HAP1-POMGnT1 cells did not look different that the amount detected on the negative control. While a small, below the limit of detection, amount of matriglycan expression cannot be ruled out, this experiment gave no indication that matriglycan is expressed in any quantity on HAP1-POMGnT1 cells.

Lack of IIH6 reactivity in WGA-enriched COSMC cell lysates:

We considered the possibility that loss of COSMC could interfere with assembly of the core M3 glycan on α-DG. WGA-enriched lysates of HEK Wt and HEK-COSMC cells were tested for IIH6 reactivity. The HEK-COSMC cells showed markedly reduced IIH6 reactivity, indicating abnormally low production of matriglycan (Figure 3-4). However, these results were misleading, as determined by additional experiments described below.

HAP1-COSMC cells do express matriglycan

We sought to confirm loss of matriglycan in COSMC knockout cells using orthogonal methods. As an alternative method of detecting matriglycan, we took advantage of its known ability to act as a receptor for Lassa virus and employed a VSV-Lassa infectivity assay. Here we used Vesicular Stomatitis Indiana virus pseudo-typed with Lassa glycoprotein to probe for matriglycan on the cell surface. HAP1-COSMC cells were obtained for this assay as they were more adherent, allowed us to repeat the results in a different cell line, and allowed for use of our HAP1-DAG1 cell line an alpha-dystroglycan-lacking negative control.

By viewing green fluorescent protein expression it was determined that greater than 50% of the HAP1-Wt cells were infected, while very few of the HAP1-DAG1 controls were infected. (Figure 3-5A). The HAP1-COSMC cells, however, were somewhat in between, more similar to the HAP1-Wt cells and showing at least some susceptibility to infection. This indicated there was a receptor for the VSV-Lassa on the surface of the HAP1-COSMC cells that was not present on the HAP1-DAG1 cells. This failed to show that HAP1-COSMC cells lack matriglycan. However, since COSMC modifies a lot of proteins we considered the

possibility that its absence made available a new receptor allowing for endocytosis of VSV-Lassa.

To clarify the results of the VSV-Lassa infectivity assay, we probed for the presence of matriglycan on HAP1 cells utilizing an on-cell western in which the cells are fixed and stained in place, allows for staining of the cells surface with antibodies. This allows for the staining of natively expressed α-DG on the cell surface without the need for enriching via immunoprecipitation.

The HAP1 Wt cells showed both IIH6 reactivity and reactivity to an antibody to the core protein, as was expected. The HAP1-DAG1 knockout cells lacked reactivity to both the IIH6 antibody and the αDG core protein antibody. The HAP1-COSMC cells showed both IIH6 and α-DG protein antibody reactivity. This was surprising as based on the results of Figure 3-1, we expected the HAP1-COSMC cells to lack IIH6 reactivity.

Both COSMC KO cell lines express Tn antigen

With contradicting results, we considered the possibility that one of our cell lines may have been incorrectly described as a COSMC knockout. We utilize ligand binding assays to verify the phenotype of both HEK-COSMC and HAP1-COSMC cell lines. We probed with SNA lectin which recognizes a2,6-linked sialic acid present on the sialyl-Tn antigen. All cells, including the negative controls, bound to SNA (data not shown). We concluded that there was too much background sialic acid for this to be an effective test.

We next tried VVA which binds α- or β-linked terminal N-acetylgalactosamine such as the Tn antigen. In this case, while both HAP1-Wt and HEK Wt control cells stained poorly, the HAP1 and HEK COSMC knockout cells both bound the VVA lectin (Figure 3-6). This

indicates that both cell lines are expressing the Tn antigen, thus we expect both to be COSMC knockout cell lines.

Lack of a-DG protein in WGA-enriched COSMC cell lysates

With the knowledge that both HAP1-COSMC and HEK-COSMC cell lines displayed the Tn antigen, indicating that they were both phenotypically COSMC KOs, we wanted to know why the WGA-enriched HEK cell lysates showed decrease IIH6 reactivity.

We hypothesized that the WGA enrichment, which enriches for glycoproteins expressing GlcNAc residues, could be failing to enrich α-DG from the COSMC cell lysates. To test this, we evaluated a battery of eight different α-DG antibodies, finding one that worked; R & D Biosystems AF6868. When WGA-enriched lysates were probed with both IIH6 and this a-DG core protein antibody, we confirmed that less α-DG protein was present in the enriched HEK- COSMC cell lysates than HEK Wt, accounting for the discrepancy in their IIH6 reactivity (Figure 3-7).

aDG340 expressed in COSMC cells

To confirm our findings, we expressed the his-tagged aDG340 in HEK Wt and HEK-COSMC cells. This allowed for a His-tag enrichment process that was not biased with respect to glycosylation. When equal amounts of His-tagged α-DG340 protein were used in the western blot, the HEK-COSMC expressed protein was IIH6 reactive (Figure 3-8). Thus, COSMC knockout cells are not deficient in their ability to produce sufficient matriglycan to bind the IIH6 antibody.

Discussion

POMGnT1 knockouts fail to express matriglycan, as determined by IIH6 reactivity (Figure 1). We were able to rescue matriglycan expression in POMGnT1 knockout cells by expressing a POMGnT1-FLAG tagged construct. However, the rescue effect was not seen in mutants generated from the POMGnT1-FLAG construct. Mutations to disrupt the catalytic activity of POMGnT1 included mutating the catalytic base, D476, and a residue of the DXD motif, D395. Neither catalytic mutant rescued IIH6 reactivity. We also tested a construct lacking the stem (97-250) domain which also failed to rescue expression of matriglycan as measured by IIH6 reactivity. Our inability to rescue matriglycan expression using POMGnT1 with a mutation at the catalytic base is inconsistent with results obtained by Kuwabara et al. who saw rescue of the IIH6 epitope in a POMGnT1 W473/M477A mutant construct. One possible explanation for this inconsistency is that the W473/M477A mutant exhibited some POMGnT1 activity which was below the limit of detection (Kuwabara, N., Manya, H., et al. 2016). Other possibilities include non-specific antibody staining or sample-cross contamination.

We sought to determine at what point in the synthesis of the core M3 glycan was truncated in POMGnT1 knockout cells. Expressing the αDG340(TTTS→AAAA) construct in HEK Simple Cells, missing POMGnT1 and COSMC activity, reduced the possible modifications on the target peptide to only O-mannose, O-GalNAc, or the partial core M3 structure at T317 or T319. We were able to identify a phospho-ribitol, phospho-trisaccharide containing peptide.

While studying POMGnT1 using POMGnT1/COSMC knockout cells, we initially had trouble rescuing IIH6-reactivity with POMGnT1. That led us to ask whether COSMC was

necessary functional glycosylation of alpha-dystroglycan. We tested COSMC KO HEK and HAP1 cells and ultimately determined that knockout of COSMC does not prevent the expression of matriglycan. We confirmed that COSMC is not necessary for expressing matriglycan.

Insights gained from this work reveal that the role of POMGnT1 in the functional glycosylation of alpha-dystroglycan is more complex than is presently described in the literature. POMGnT1 knockout cells are capable of extending the M3 structure beyond the phospho-tri- saccharide, but with unknown frequency. The POMGnT1 knockouts require more than just the stem domain of POMGnT1 to rescue functional glycosylation of alpha-dystroglycan. Future work is needed to elucidate the entire role of POMGnT1 in functional glycosylation of alpha- dystroglycan and, thus, understand why loss of POMGnT1 results in the congenital muscular dystrophy of Muscle-Eye-Brain disease.

Materials and Methods

DNA constructs

aDG340TAA: As previously described (Praissman, J.L., Willer, T., et al. 2016) except used site-directed mutagenesis to mutate T322, T328, T329, and S336 to Alanine.

POMGnT1: The full-length POMGnT1 was provided by Kevin Campbell. The construct was put into pcDNA3.1 (Thermo Fisher Scientific) via restriction cloning with XhoI/XbaI. Site-directed mutagenesis was used to mutate D395 and D476 using primers TGGTTCTGGAAGAGAACCTGGACATTGCTGTG/ACAGCAATGTCCAGGTTCTCTTC CA GAACCAC and TCTGGGATTGGAACATGTGGATGCG/CGCATCCACATGTTCCAATCCCAGAG, respectively. Deletion of the stem domain residues 97-250 was done with Q5 mutagenesis (New England Biolabs) using primers GAGGCAGAGTGCCACTGG and CCGGGGACCACTGCCTCT designed using NEBaseChanger (New England Biolabs).

Cell Culture

HEK Wt, HEK-POMGnT1, and HEK Simple Cells (POMGnT1/COSMC) were provided by Henrick Clausen and described previously (Larsen, I.S.B., Narimatsu, Y., et al. 2017a, Larsen, I.S.B., Narimatsu, Y., et al. 2017b). Cells were grown in DMEM (Gibco) with 10% FBS (Gibco) at 37°C with 5% CO2.

Expression of aDG340(TTTS→AAAA) was done via transient transfection with PEI as described in Chapter 2. Expression of POMGnT1 mutants was done via stable transfection using linearize DNA, PEI, and G-418 (Gibco) to select for stable transfectants.

WGA Enrichment

HEK Wt and HEK COSMC cells were washed twice in cold PBS and lysed with 1% Trition X- 100 in PBS with 1X EDTA-free protease inhibitor cocktail (Millipore-Sigma). Cell debris was removed by centrifugation at 8,000 rcf for 10 minutes. Lysate was incubated with 20uL equilibrated WGA resin overnight. WGA resin with glycoproteins bound was washed 2X in PBS and eluted in western blot loading buffer.

Western Blot of POMGnT1 variants

Adherent cells were washed twice with cold PBS. PBS with 1% Trition X-100 was used to lyse cells and solubilize proteins. After centrifugation at 8,000 rcf for 10 minutes, pellet was discarded, and the protein concentration of the supernatant was measured by BCA assay. Equal amounts of protein from each supernatant were incubated with WGA resin (Vector Labs) overnight at 4C. WGA resin was washed three times in PBS and bound glycoproteins were eluted by incubating at 95C for 10 minutes in LDS sample buffer (Fisher). Samples were loaded onto 4-12% Bis-Tris Bolt SDS-PAGE Gel (Fisher) and run for 30 minutes at 200V. Proteins were transferred to 0.45um FL low-fluorescence PVDF membrane. Membranes were blocked in 0.5% cold-water fish skin gelatin and incubated with IIH6 antibody diluted 1:500 (DSHB) or anti-FLAG antibody 1:3000 (Sigma) overnight at 4C.

Mass Spec of glycopeptides

Ni-NTA purified samples were buffer-exchanged into PBS with D-tube midi dialyzers (Millipore-Sigma). Samples were reduced with 5mM dithiothreitol in 40mM NH_4CO_3 at 56°C

for 60 minutes, alkylated in 27.5 mM iodoacetamide for 45 minutes at room temperature (~22°C) in the dark and digested with 2 µg sequencing grade modified trypsin (Promega) at 37°C overnight. Peptides were recovered using reverse-phase separation on C18 spin column (MicroSpin, The Nest Group), eluted with 0.1% Formic Acid in 80% acetonitrile, and dried under a vacuum. Peptides were reconstituted in 3 µL 5% solvent B (0.1% formic acid in 80% acetonitrile) and 57 µL solvent A (0.1% formic acid), separated on an Acclaim PepMap RSLC C18 column (75um x 15cm), and eluted into the nano-electrospray ion source of an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) with a 180-minute linear gradient consisting of 1-99% solvent B over 180 minutes at a flow rate of 200 nL/min. Spray voltage was set to 2.2 kV and the temperature of the heated capillary was set to 280°C. Full MS scans were acquired from m/z 300 to 2000 at 60k resolution, and MS2 scans following high energy collision dissociation stepped at 25% +/- 10, were collected in the ion trap for the most intense ions in the Top-Speed mode with a 3 second cycle using Fusion instrument software (v2.0, Thermo Fisher Scientific).

References

Barresi R, Campbell KP. 2006. Dystroglycan: from biosynthesis to pathogenesis of human disease. *Journal of cell science*, 119:199-207.

Bertini E, D'Amico A, Gualandi F, Petrini S. 2011. Congenital muscular dystrophies: a brief review. *Seminars in pediatric neurology*, 18:277-288.

Dobson CM, Hempel SJ, Stalnaker SH, Stuart R, Wells L. 2013. O-Mannosylation and human disease. *Cellular and molecular life sciences : CMLS*, 70:2849-2857.

Endo T. 1999. O-mannosyl glycans in mammals. *Biochimica et biophysica acta*, 1473:237-246.

Ervasti JM, Campbell KP. 1991. Membrane organization of the dystrophin-glycoprotein complex. *Cell*, 66:1121-1131.

Ervasti JM, Campbell KP. 1993. A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin. *J Cell Biol*, 122:809-823.

Harrison R, Hitchen PG, Panico M, Morris HR, Mekhaiel D, Pleass RJ, Dell A, Hewitt JE, Haslam SM. 2012. Glycoproteomic characterization of recombinant mouse alpha-dystroglycan. *Glycobiology*, 22:662-675.

Holt KH, Crosbie RH, Venzke DP, Campbell KP. 2000. Biosynthesis of dystroglycan: processing of a precursor propeptide. *FEBS letters*, 468:79-83.

Ibraghimov-Beskrovnaya O, Ervasti JM, Leveille CJ, Slaughter CA, Sernett SW, Campbell KP. 1992. Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature*, 355:696-702.

Kuwabara N, Manya H, Yamada T, Tateno H, Kanagawa M, Kobayashi K, Akasaka-Manya K, Hirose Y, Mizuno M, Ikeguchi M, *et al.* 2016. Carbohydrate-binding domain of the POMGnT1 stem region modulates O-mannosylation sites of α-dystroglycan. *Proc Natl Acad Sci U S A*, 113:9280-9285.

Larsen ISB, Narimatsu Y, Joshi HJ, Siukstaite L, Harrison OJ, Brasch J, Goodman KM, Hansen L, Shapiro L, Honig B, *et al.* 2017a. Discovery of an O-mannosylation pathway selectively serving cadherins and protocadherins. *Proceedings of the National Academy of Sciences*, 114:11163-11168.

Larsen ISB, Narimatsu Y, Joshi HJ, Yang Z, Harrison OJ, Brasch J, Shapiro L, Honig B, Vakhrushev SY, Clausen H, *et al.* 2017b. Mammalian O-mannosylation of cadherins and plexins is independent of protein O-mannosyltransferases 1 and 2. *The Journal of biological chemistry*, 292:11586-11598.

Liu J, Ball SL, Yang Y, Mei P, Zhang L, Shi H, Kaminski HJ, Lemmon VP, Hu H. 2006. A genetic model for muscle-eye-brain disease in mice lacking protein O-mannose 1,2-N-acetylglucosaminyltransferase (POMGnT1). *Mechanisms of development*, 123:228-240.

Manya H, Sakai K, Kobayashi K, Taniguchi K, Kawakita M, Toda T, Endo T. 2003. Loss-of-function of an N-acetylglucosaminyltransferase, POMGnT1, in muscle-eye-brain disease. *Biochemical and biophysical research communications*, 306:93-97.

Meilleur KG, Zukosky K, Medne L, Fequiere P, Powell-Hamilton N, Winder TL, Alsaman A, El-Hattab AW, Dastgir J, Hu Y, *et al.* 2014. Clinical, pathologic, and mutational spectrum of dystroglycanopathy caused by LARGE mutations. *Journal of neuropathology and experimental neurology*, 73:425-441.

Mercuri E, Messina S, Bruno C, Mora M, Pegoraro E, Comi GP, D'Amico A, Aiello C, Biancheri R, Berardinelli A, *et al.* 2009. Congenital muscular dystrophies with defective glycosylation of dystroglycan: a population study. *Neurology*, 72:1802-1809.

Praissman JL, Willer T, Sheikh MO, Toi A, Chitayat D, Lin YY, Lee H, Stalnaker SH, Wang S, Prabhakar PK, *et al.* 2016. The functional O-mannose glycan on alpha-dystroglycan contains a phospho-ribitol primed for matriglycan addition. *eLife*, 5.

Santavuori P, Somer H, Sainio K, Rapola J, Kruus S, Nikitin T, Ketonen L, Leisti J. 1989. Muscle-eye-brain disease (MEB). *Brain & development*, 11:147-153.

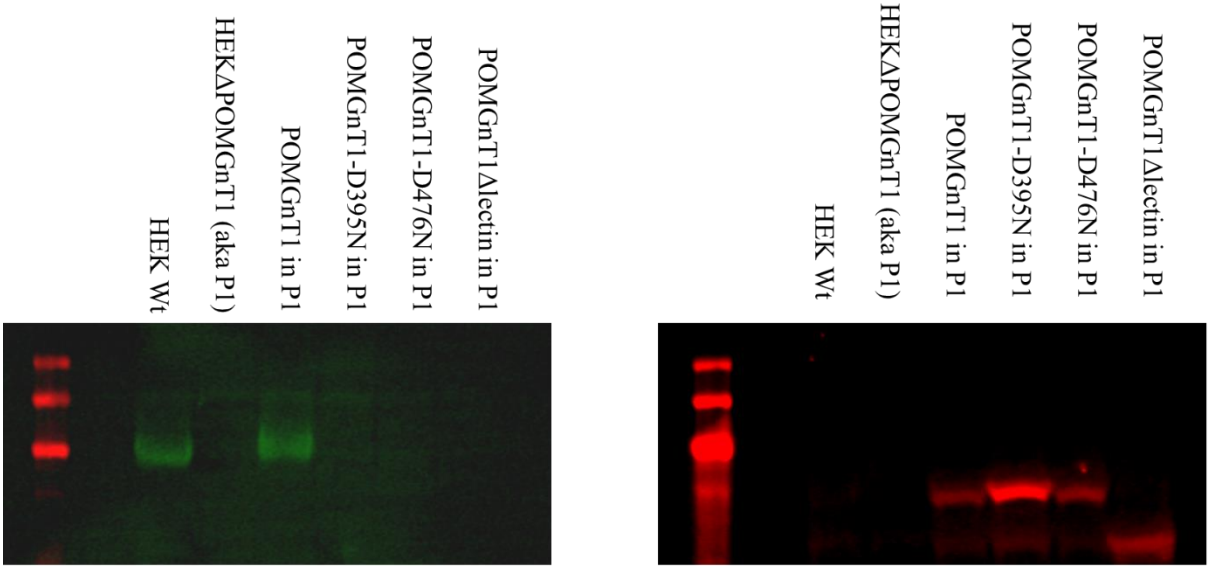
Wells L. 2013. The o-mannosylation pathway: glycosyltransferases and proteins implicated in congenital muscular dystrophy. *The Journal of biological chemistry*, 288:6930-6935.

Yoshida-Moriguchi T, Yu L, Stalnaker SH, Davis S, Kunz S, Madson M, Oldstone MB, Schachter H, Wells L, Campbell KP. 2010. O-mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. *Science (New York, N.Y.)*, 327:88-92.

Figure 3-1A


Full Length POMGnT1 (1-660)
lectin
linker
catalytic
DXD Mutant (D395N)
lectin
linker
Catalytic Base Mutant (D476N)
lectin
linker
Lectin Mutant (Δ97-250)
linker
catalytic

Figure 3-1B



Green: IIH6 (matriglycan)
(WGA-Enriched cell lysate)

Red: FLAG
(whole cell lysate)

Figure 3-1. Western Blot of POMGnT1 constructs. (A) IIH6 reactivity of HEK Wt and HEK POMGNT1 cells shows that HEK Wt are IIH6 positive while HEK POMGnT1 are IIH6 negative. Expressing a full-length POMGnT1-FLAG construct in HEK-POMGnT1 rescued IIH6 reactivity. However, POMGnT1-FLAG constructs in which the DXD motif (D395N), catalytic base (D476N), or stem domain (Δ97-250) had been mutated failed to rescue IIH6 reactivity. (B) Anti-FLAG reactivity was negative in the parental HEK Wt and HEK POMGnT1 cell lines.

FLAG expression was detected cells expressing the four POMGnT1-FLAG mutants. The FLAG antibody was detected at the expected 77kDa for all constructs except the (Δ97-250) stem domain deletion, as expected.

Figure 3-2A

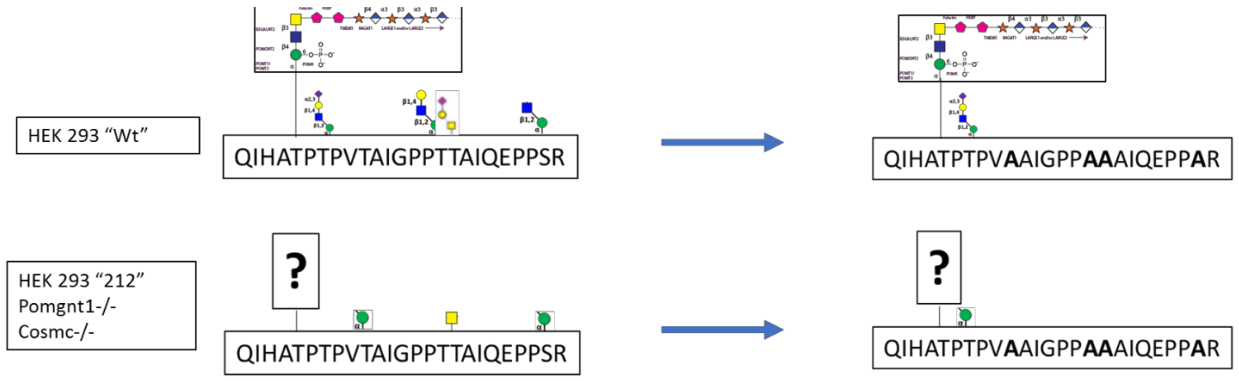

HEK 293 "Wt"
QIHATPTPVTAIGPPTTAIQEPPSR
QIHATPTPVAAIGPPAAAIQEPPAR
HEK 293 "212"
Pomgnt1-/-
Cosmc-/-
?
QIHATPTPVTAIGPPTTAIQEPPSR
?
QIHATPTPVAAIGPPAAAIQEPPAR

Figure 3-2B

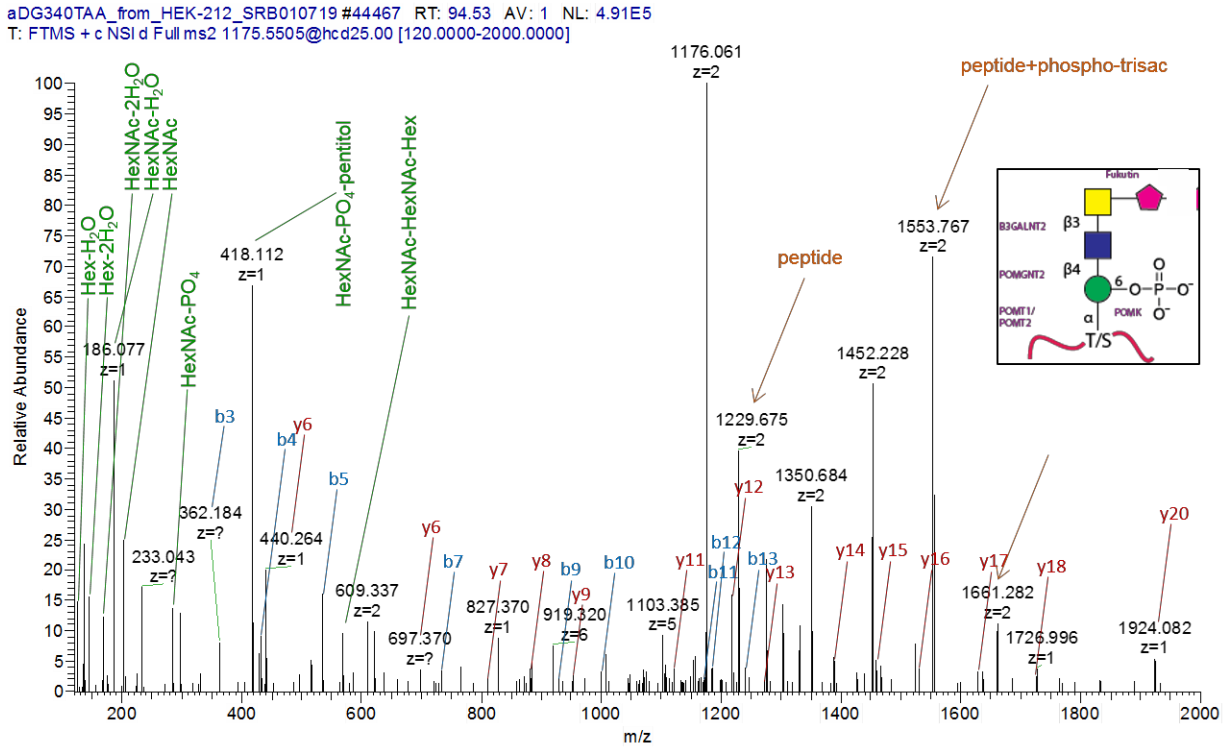

aDG340TAA_from_HEK-212_SRB010719 #44467 RT: 94.53 AV: 1 NL: 4.91E5
T: FTMS + c NSI d Full ms2 1175.5505@hcd25.00 [120.0000-2000.0000]
Relative Abundance
m/z
peptide+phospho-trisac
peptide
1176.061 z=2
1553.767 z=2
1452.228 z=2
1229.675 z=2
1350.684 z=2
418.112 z=1
186.077 z=1
362.184 z=?
440.264 z=1
233.043 z=?
609.337 z=2
697.370 z=?
827.370 z=1
919.320 z=6
1103.385 z=5
1661.282 z=2
1726.996 z=1
1924.082 z=1
HexNAc-2H2O
HexNAc-H2O
HexNAc
Hex-H2O
Hex-2H2O
HexNAc-PO4
HexNAc-PO4-pentitol
HexNAc-HexNAc-Hex
b3
b4
b5
b7
b9
b10
b11
b12
b13
y6
y7
y8
y9
y11
y12
y13
y14
y15
y16
y17
y18
y20
Fukutin
B3GALNT2
β3
POMGNT2
β4
6
POMT1/
POMT2
α
POMK
T/S

Figure 3-2. The glycan structure present at T317/T319 on α-DG expressed in cells lacking POMGnT1 can be extended past the phospho-tri-saccharide. (A) The aDG340 construct is a soluble, secreted construct of the first N-terminal 340 amino acids of alpha-dystroglycan. It contains the T317/319 site of core M3 O-mannosylation. The illustration depicts glycans expected to be present on both aDG340 and the aDG340(TTTS→AAAA) mutant when expressed in either wild type or SimpleCells. (B) Annotated spetrum of phospho-ribitol-phospho-tri-saccharide-modified peptide from α-DG expressed in HEK SimpleCells.

Table 3-1

Modification	Spectral Counts	Enzyme needed to obtain modification
Peptide (all mods)	769	
+204 (not part of M3)	572	
+162	8 (6%)	POMT1/2
+365	26 (19%)	POMGnT2
+568	48 (35%)	B3GALNT2
+648	42 (30%)	POMK
+862	12 (9%)	Fukutin
+1076	1*	FKRP
+1208	0	RXYLT1/ TMEM5
+1384	1*	B4GAT1
+1516	0	LARGE

*Same scan, z=8

Table 3-1. Over-expression of the α-DG340(TTTS→AAAA)-pGEc2 in HEK SimpleCells. The first column contains the mass of the modification allowed for in the search. The second column reports the number of spectral counts identified as having that modification. The third column lists the enzyme necessary to obtain the modification listed in the first column.

Figure 3-3

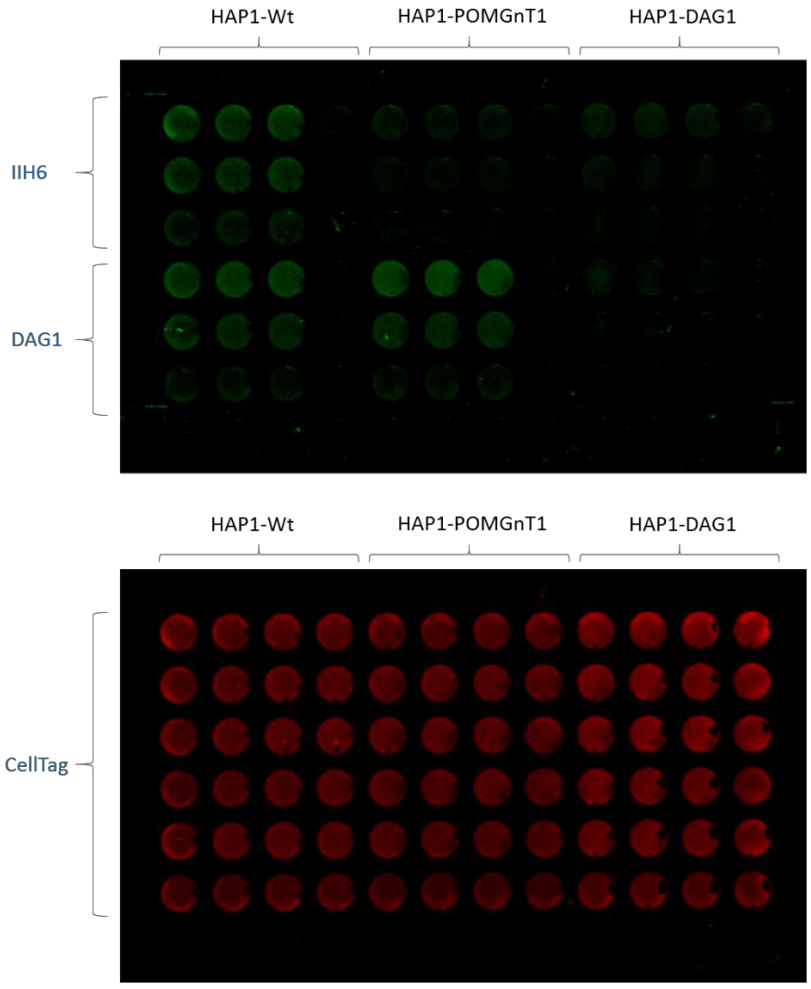

HAP1-Wt
HAP1-POMGnT1
HAP1-DAG1
IIH6
DAG1
HAP1-Wt
HAP1-POMGnT1
HAP1-DAG1
CellTag

Figure 3-3. HAP1-POMGnT1 cells express α-DG but not matriglycan. CellTag stain on bottom panel is a positive control indicating that all wells contain cells and cells were not washed away during the assay. The plate is divided into three columns, containing HAP1-Wt (left), HAP1- POMGNT1 (middle), and HAP-DAG1 (right). The first three wells of each cell type were treated with primary and secondary antibody. Fourth well was treated with secondary antibody only as a negative control.

The top of the plate was probed with IIH6 while the bottom with the core aDG protein antibody. As expected, the Wt cells were positive for both aDG protein and IIH6. Also, as expected the DAG1 knockout cells were negative for both. The POMGnT1 KO cells, center, were positive for aDG protein, but negative for IIH6 reactivity.

Figure 3-4

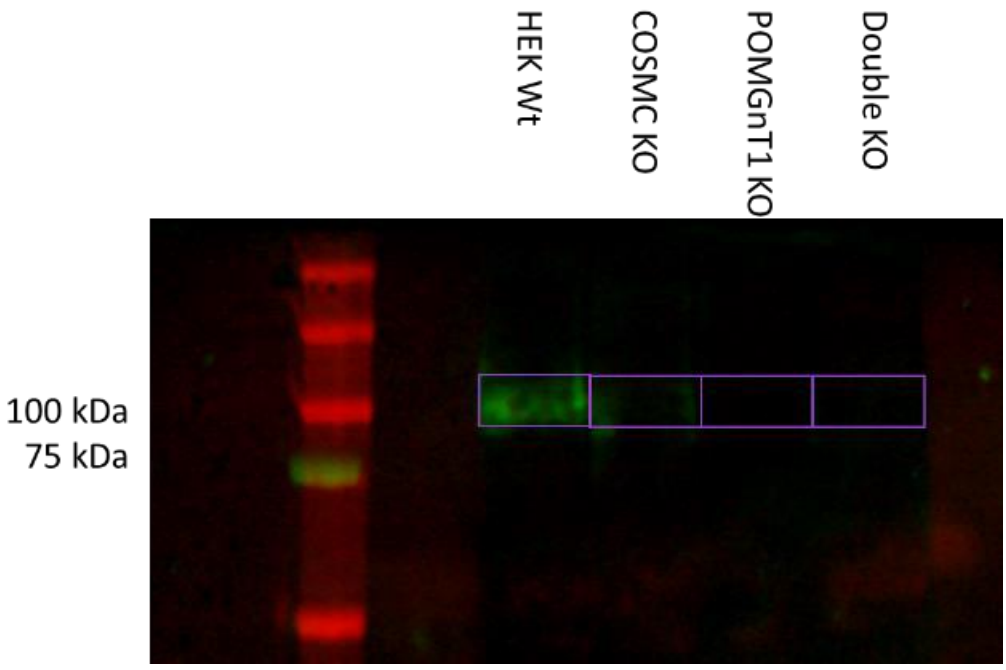


Figure 3-4. Western Blot probing IIH6 reactivity of WGA-enriched lysates from HEK Wt, COSMC, POMGnT1, and COSMC/POMGNT1 KO cells. Only the Wt cells are IIH6 reactive, indicating that the COSMC and POMGnT1 knockouts fail to express matriglycan.

Figure 3-5A

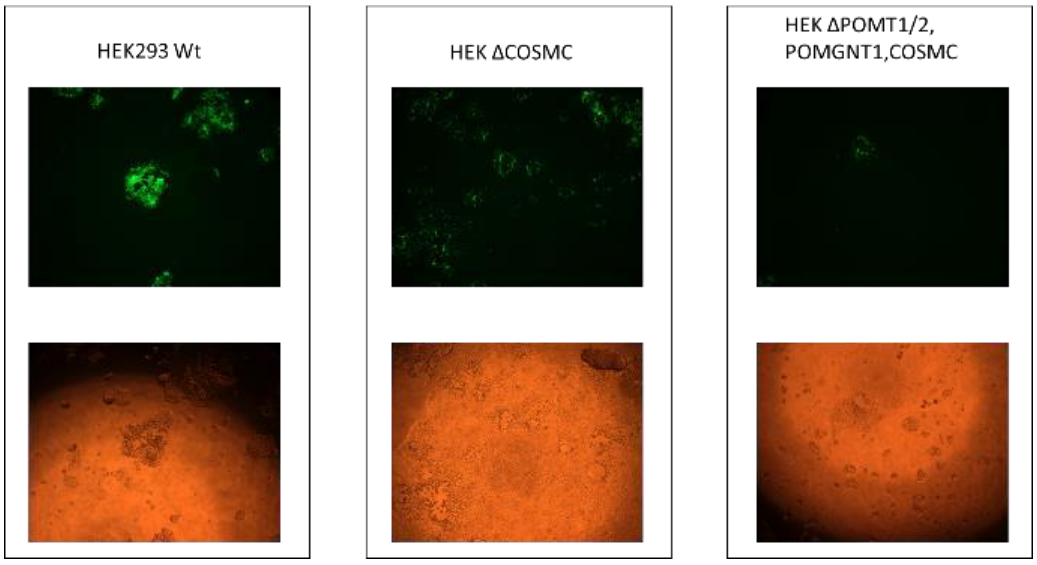


Figure 3-5B

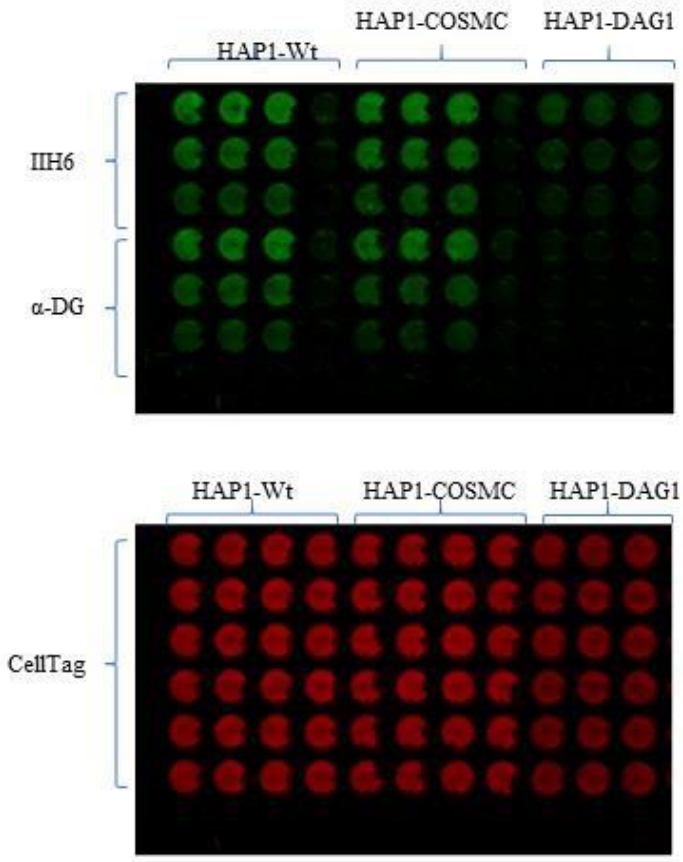

HAP1-Wt
HAP1-COSMC
HAP1-DAG1
IIH6
α-DG
HAP1-Wt
HAP1-COSMC
HAP1-DAG1
CellTag

Figure 3-5. HAP1-COSMC knockout cells express functionally glycosylated α-DG. (A) GFP+ HAP1-COSMC cells indicate that HAP1-COSMC was infected by VSV-LASV. VSV-LASV is able to infect the matriglycan-expressing HAP1-control cells but not the HAP1-DAG1 knockout cells which lack α-DG and therefore lack the matriglycan cell surface receptor for the Lassa virus glycoproteins expressed on the pseudo-typed VSV. (B) On-cell western of HAP1-Wt, HAP1-COSMC, and HAP1-DAG1. CellTag stain on bottom panel is a positive control indicating that all wells contain cells. The plate is divided into three columns for the HAP1-Wt (positive control, left four columns of wells), HAP1-COSMC (experimental, middle four columns of wells) and HAP1-DAG1 (negative control, four columns of wells, one of which was cut off in the image.) The top of the plate received IIH6 as the primary antibody while the bottom of the plate received the antibody to the α-DG core protein. As expected, Wt cells were positive for both and DAG-1 cells were negative for both. The HAP1-COSMC cells clearly showed the presence of both α-DG core protein and matriglycan.

Figure 3-6

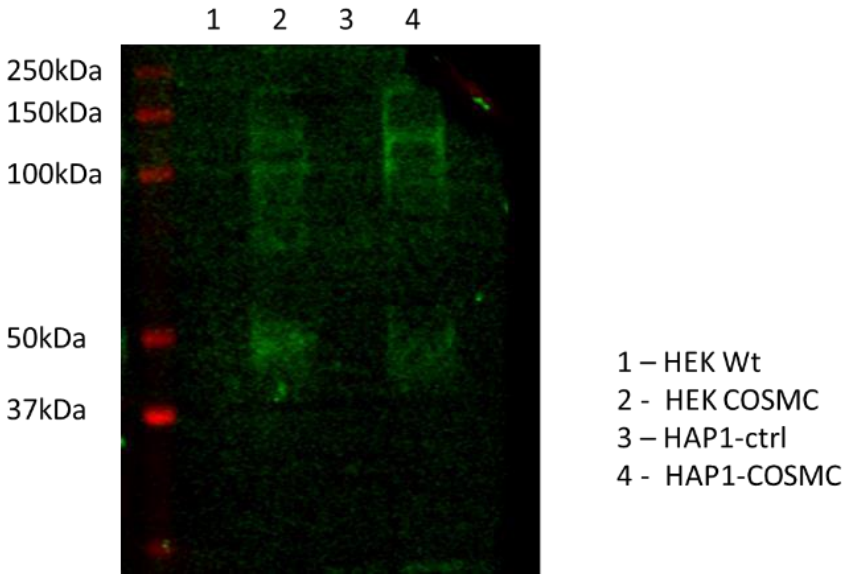
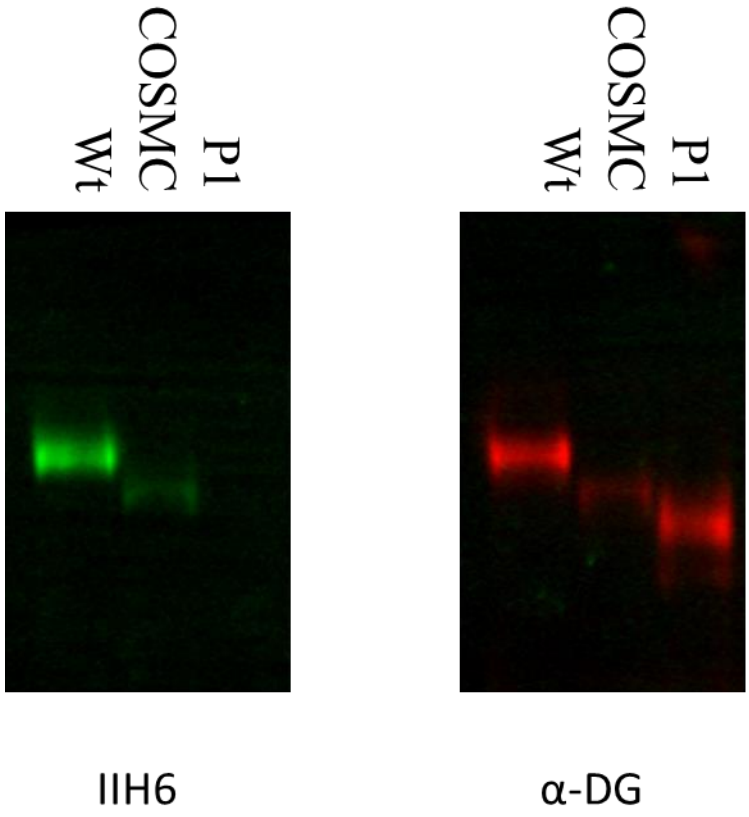

1
2
3
4
250kDa
150kDa
100kDa
50kDa
37kDa
1 – HEK Wt
2 - HEK COSMC
3 – HAP1-ctrl
4 - HAP1-COSMC

Figure 3-6. VVA lectin blot of HEK and HAP1 cells. Both HEK Wt and HAP1-Wt cells failed to stain, as expected. Both HEK-COSMC and HAP1-COSMC bound the VVA lectin, demonstrating that both COSMC knockout cell lines express Tn antigen.

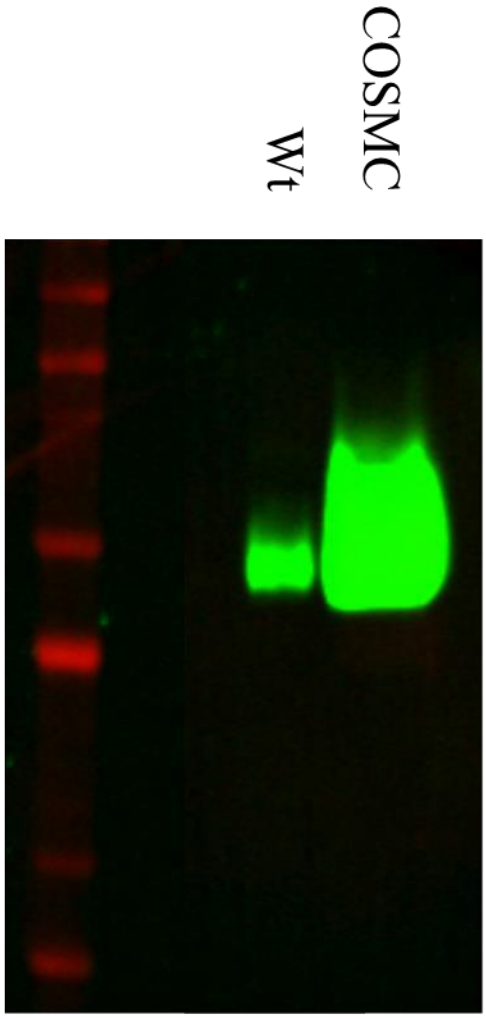
Figure 3-7



Wt – WGA-enriched HEK Wt cell lysate
COSMC – WGA-enriched HEK-COSMC cell lysate
P1 – WGA enriched HEK POMGnT1/COSMC/POMT1/POMT2 knockout

Figure 3-7. Western blots of WGA-enriched lysates from HEK cells. HEK Wt, in lane 1, expressed both α-DG and matriglycan. The *Pomgn1/Cosmc/Pomt1/Pomt2* knockout in lane 3 expressed α-DG but not matriglycan. The samples from HEK-COSMC cells in the middle lane on each blot contain very little α-DG and a similar amount of IIH6 reactivity. Sample loading was based on the amount of protein, as measured by BCA assay, before the WGA enrichment. The reduced amount of IIH6 reactivity seen in the left blot is a result of the reduced amount of α- DG from the WGA-enrichment process, not a reduction in the amount of matriglycan on α-DG.

Figure 3-8



Wt – aDG340 expressed in HEK Wt cells
COSMC – aDG340 expressed in HEK COSMC cells

Figure 3-8. Western blot of IIH6 reactivity of aDG340 from HEK Wt and HEK-COSMC cells. Both are IIH6 reactive, indicating that both are expressing matriglycan. The large difference in signal is due to different amounts of protein being loaded. This blot confirms that HEK-COSMC cells are capable of functionally glycosylating α-DG.

CHAPTER 4

CONCLUSION

O-linked mannose has been known to modify mammalian glycoproteins for over four decades. However, the prevalence and function of O-mannosylation and O-mannose glycans is still being elucidated. O-mannose glycans have been found necessary for proper development and function of cell adhesion factors. Loss of proper O-mannosylation has been identified as the cause of congenital disorders involving the brain, muscles, and eyes (Sparks, S.E., Quijano-Roy, S., et al. 1993).

Recently, our understanding of the number and type of O-mannosylated proteins has expanded dramatically. Cadherins and IPT domain-containing proteins have recently been identified as being modified by a novel core M0, non-elongated, O-mannose (Vester-Christensen, M.B., Halim, A., et al. 2013). The O- mannose modification has been shown to be crucial for E-cadherin-based cell adhesion (Lommel, M., Winterhalter, P.R., et al. 2013). The O-mannose on cadherins is different from the previously studied O-mannose glycans on alpha- dystroglycan in that it is not extended and is not dependent on the Protein O-mannosyl transferases 1 and 2, previously believe to be the only O-mannosyl transferases in the mammalian secretory pathway (Larsen, I.S.B., Narimatsu, Y., et al. 2019).

The enzyme POMGnT1 catalyzes the transfer of N-acetylglucosamine to O-linked mannose in the cis-Golgi apparatus. Although POMGnT1 has previously been shown to lack peptide specificity, it does not modify the O-mannose residues recently discovered on

cadherins. As protein folding occurs in the endoplasmic reticulum, the Golgi-residing POMGnT1 encounters folded proteins.

To study whether cadherins could be converted into POMGnT1 substrates, and investigate how they escape modification, we expressed recombinant E-cadherin and N-cadherin in HEK-POMGnT1 KO cells. Consistent with previous results, the cadherin proteins were modified by non-elongated hexoses presumed to be mannose. Expressing these same proteins in HEK Wt cells resulted in a similar pattern of hexose-modification indicating that POMGnT1 does not alter the expression of these O-linked hexoses, consistent with previous results (Larsen, I.S.B., Narimatsu, Y., et al. 2017b).

Given that these hexose modifications were unlike all previously known Core M1, M2, and M3 O-mannose glycans, and that they were still present after knockout of the only known mammalian O-mannosyl transferases (POMT1 and POMT2), there was a need to confirm the identity of the hexose. The initial method of identifying over 50 different cadherin and plexin proteins as being O-mannosylated utilized their weak affinity to concanavalin A (Con A) (Vester-Christensen, M.B., Halim, A., et al. 2013). Con A has affinity for both mannose and glycose (Mandal, D.K., Bhattacharyya, L., et al. 1994). To confirm that the hexoses were mannose, we utilized α-mannosidase. The hexose was found to be susceptible to removal by α- mannosidase when tryptic peptides of the proteins were treated. Thus, we confirmed that the hexose is an O-mannose glycan which is not modified by POMGnT1.

Previous work has shown that POMGnT1 lacks substrate selectivity with respect to amino acid sequence when presented with synthetic O-mannosylated peptides (Halmo, S.M., Singh, D., et al. 2017). However, the peptides used in that study had been derived exclusively

from sequences of alpha-dystroglycan, a protein well-established as a substrate of POMGnT1 (Biancheri, R., Bertini, E., et al. 2006, Halmo, S.M., Singh, D., et al. 2017, Kuwabara, N., Manya, H., et al. 2016, Stalnaker, S.H., Aoki, K., et al. 2011). A possible explanation for this was that there is a more complex motif for POMGnT1 modification, one present throughout the mucin-like domain of α-DG but absent on cadherins. A more straightforward explanation was

that steric hinderance prevented the sites of O-mannosylation from being available to POMGnT1. We hypothesized that protein folding, which occurs in the ER antecedent to entry into the Golgi, rendered the sites of O-mannosylation inaccessible to POMGnT1.

We tested this by digesting the recombinantly expressed proteins with the protease trypsin and comparing the ability of POMGnT1 to modify the tryptic peptides versus untreated protein. First we counted disintegrations per minute to quantify UDP-[3H]GlcNAc transferred by POMGnT1 to the different potential substrates. Second, we used only cold UDP-GlcNAc and assessed the number of spectral counts of POMGnT1-modifed peptides obtained from samples treated with POMGnT1 as either protein or peptides. In both experiments, E-cadherin and N- cadherin were inefficient POMGnT1 substrates until being converted to peptides. Tryptic peptides of E-cadherin and N-cadherin were efficiently modified by POMGnT1, as were both the untreated protein and tryptic peptides of α-dystroglycan. Thus, we conclude that secondary, tertiary, or quaternary structure, not the primary sequence of amino acids, prevents extension of the O-mannose on cadherins by POMGnT1.

Future study of the core M0 structure may take several directions. First, the initial work by Vester-Christensen et al. revealed a second class of proteins to be modified by non-elongated O-mannose; the IPT domain-containing proteins, including the plexins and

hepatocyte growth factor receptor (HGFR) [11]. Repeating the work done on cadherins with plexins would confirm that they too are modified by mannose which can be extended by POMGnT1 if unhindered by tertiary structure. We began this work with HGFR (see Appendix A) and demonstrated that the hexose modification on HGFR at T761 is somewhat sensitive to α-mannosidase (180 spectral counts of hexose-containing peptides vs. 60 counts after mannosidase treatment). Additionally, treatment of tryptic peptides of HGFR with POMGnT1 resulted in transfer of UDP-GlcNAc to the hexose (Figure A-2.). Taken together, these results confirm that the hexoses on HGFR are O-mannose.

An important future direction of study remaining is to determine the function of O-mannose residues on cadherins and IPT domain-containing proteins. Lommel et al. have demonstrated that O-mannosylation is crucial for E-cadherin-mediated cell adhesion, and that its loss results in embryonic lethality between the morula and blastocyst stages in mice (Lommel, M., Winterhalter, P.R., et al. 2013).

However, they demonstrated this by knocking out *Pomt2* and through use of the fungal protein O-mannosyltransferase inhibitor R3A-5a. We now know that *Pomt2* is disposable for O- mannosylation of E-cadherin (Larsen, I.S.B., Narimatsu, Y., et al. 2017a). R3A-5a was shown to be an inhibitor of mammalian O- mannosylation by its ability to inhibit synthesis of matriglycan on α-DG (Lommel, M., Winterhalter, P.R., et al. 2013). With our new knowledge of the alternative O-mannosylation pathway for cadherins, independent of the classical POMT1/2-dependent O-mannosylation of α-DG, these findings merit further study.

One possibility is that the morula-to-blastocyst embryonic lethality Lommel et al. reported in *Pomt2* KO mice was due to a loss of core M1, M2, or M3 structures on α-DG or other proteins from the classical O-mannosylation pathway. Several mouse models utilized in

the study of CMD can inform regarding this possibility. The absence of LARGE in the Large^{myd} mouse demonstrates that LARGE modification of α-DG or other proteins is not necessary for viability in mice (Grewal, P.K. and Hewitt, J.E. 2002). Liu et al. generated POMGnT1 null mice to model MEB (Liu, J., Ball, S.L., et al. 2006). The viability of this line discounts the necessity of core M1 and M2 structures for the morula to blastocyst transition in mice. Moore et al. deleted dystroglycan in mice, but specifically in the central nervous system using *Cre-loxP* methodology. The brain-selective expression of Cre recombinase wasn't detectable until day E13.5, well past the morula to blastocyst transition (Moore, S.A., Saito, F., et al. 2002). The work by Lommel et al. demonstrates that O-mannosylation by POMT1/2 is essential. However, it does not inform us of the function of O-mannosylation on cadherins or IPT domain containing proteins. With the recently acquired knowledge that cadherin O-mannosylation is dependent on the putative O-mannosyltransferases TMTC1-4, and IPT domain-containing proteins are O-mannosyalted by an as-yet undiscovered pathway, the O-mannose modification may function differently in the cadherins when compared to the IPT domain containing proteins.

For cadherins, Larsen et al. generated TMTC1-4 knockout HEK cells which lack O-mannosylation of cadherin proteins. These cells could be utilized to assay differences in cell migration, cell adhesion, and formation of adherins junctions through previously developed methods (Harrison, O.J., Jin, X., et al. 2011). Further, their responses to various external stimuli could be examined globally at the transcriptome level with RNAseq, and at the protein level to determine changes in expression or phosphorylation status of transcription factors such as β-catenin and downstream transcriptional activators.

Regarding the study of O-mannose in IPT domain-containing proteins, HGFR is an appealing starting place as its small number of sites of O-mannosylation make it a simple model. Early work in our laboratory indicated that HGFR had only one site of O-mannosylation, T761. T761 is located in the third IPT domain of HGFR, a location known to contain a high affinity binding site for HGF (Basilico, C., Arnesano, A., et al. 2008). Preliminary work with a secreted construct of HGFR, and a T761A mutant, determined that loss of O-mannosylation specifically at T761 did not affect the ability recombinant HGFR to bind its HFR ligand (Figures A-3 and A-4).

There are three possible explanations for the lack of difference in binding capacity of HGFR and HGFR-T761A. This first is that our hypothesis is wrong, and the O-mannose does not have a role in modulating ligand binding. The second is that we failed to remove all of the O-mannose residues from HGFR. More robust step-HCD MS/MS analysis of subsequent batches of HGFR expressed in HEK Wt cells received from Henrik Clausen's laboratory revealed three sites of O-mannosylation: S582, T676, and T761. It is possible the remaining sites of O-mannosylation could have compensated for the missing T761 in our mutant. Finally, a third possibility is that the O-mannose functions in poorly-understood mechanism for dimerization of the HGFR receptor which occurs after ligand binding to activate the downstream signaling cascade.

A possibility for future work would be to generate an HGFR construct with all three sites of O-mannosylation mutated (HGFR3Xmut) and express either it in an HGFR knockout hepatocyte cell line. The first experiment would be to confirm that HGFR and the mutant are both expressed at the cell surface, probably by the on-cell Western blot described in. Assuming both HGFR and HGFR3Xmut are expressed at the surface, I would then expose

both cell populations to the HGF ligand and asses their response using techniques similar to those described for the TMTC1-4 KO cells above.

Loss of POMGnT1 is seen in Muscle-Eye-Brain Disease, one of the secondary dystroglycanopathies resulting from loss of functional glycosylation of alpha-dystroglycan. POMGnT1 transfers GlcNAc in a 2-linkage to form the core M1 structure. Therefore, loss of the core M1 structure, and the core M2 structure built from it, are expected in POMGnT1 knockouts. Unexpectedly, loss of POMGnT1 also results in loss of functional glycosylation of alpha- dystroglycan. Functional glycosylation of alpha-dystroglycan refers to the assembly of a core M3 (POMGnT2-dependent) glycan structure that is extended with matriglycan, a repeating disaccharide of xylose and glucuronic acid. Matriglycan functions as a receptor for LG-domain containing proteins in the extracellular matrix.

There are several enzymes which, when defective, are known to be causal for secondary dystroglycanopathies. However, unlike POMGnT1, each of those enzymes has a known role in building the core M3, matriglycan-containing, functional glycan on alpha-dystroglycan. While POMGnT1 modifies other O-mannose glycans on alpha-dystroglycan, its role in building the matriglycan-containing structure is unknown.

We sought to better understand how POMGnT1 was necessary for expression of matriglycan-containing core M3 structure by asking what the core M3 structure looked like in the absence of POMGnT1. Western blots with the IIH6 antibody, which recognizes matriglycan, show a loss of matriglycan in POMGnT1 knockouts. However, beyond the loss of matriglycan, the composition of the core M3 in POMGnT1 knockouts is unknown.

The first four components of the functional glycan, the phospho-tri-saccharide, are assembled in the ER (Yoshida-Moriguchi, T., Willer, T., et al. 2013). POMGnT1 resides in

the cis-Golgi so it is expected that the functional glycan would be assembled up to or beyond the phospho-tri-saccharide. The next step in synthesis of the functional glycan addition of phospho-ribitol by the glycosyl transferase fukutin. Kuwabara et al. postulated that the function of POMGnT1 in assembly of the functional glycan is to recruit fukutin to the phospho-tri-saccharide (Kuwabara, N., Manya, H., et al. 2016). If so, we would expect to see α-DG expressed in POMGnT1 knockouts to consistently modified by the core M3 structure built up to the phospho-tri-saccharide, but not extended further.

We sought to determine what the core M3 structure looked like in POMGnT1 knockouts by expressing a simplified α-DG construct in POMGnT1 knockout cells. We determined that, in rare cases, it was possible for the glycan structure to be built beyond the phospho-tri-saccharide. We were, however, not able to discern a clear cut-off or point at which the assembly of the core M3 structure breaks down in POMGnT1 knockouts.

Evidence of support for the POMGnT1-fukutin hypothesis provided by Kuwabara et al. included a blot which showed recovery of IIH6 reactivity in POMGnT1 deficient cells transfected with a W473/M477A mutant of POMGnT1. This led to the conclusion that catalytic activity of POMGnT1 was not necessary for functional glycosylation of alpha-dystroglycan (Kuwabara, N., Manya, H., et al. 2016). This conclusion, however, is contradicted by the identification of seven different mutations in the catalytic domain of POMGnT1 identified in patients with MEB (Diesen, C., Saarinen, A., et al. 2004). Additionally, there are two examples of patients with mutations in the stem domain of POMGnT1 who have RP76, not MEB, and therefore no muscular or neurological disorder (Xu, M., Yamada, T., et al. 2016).

To better assess the necessity of various domains of POMGnT1 for recovery of functional glycosylation of alpha-dystroglycan we generated POMGnT1 constructs

containing the full-length Wt POMGnT1 protein, a mutation at the DXD motif (D395N), a mutation at the catalytic base (D474N), and deletion of the stem domain (Δ97-299). Only the full-length Wt POMGnT1 was able to rescue IIH6 reactivity of alpha-dystroglycan in POMGnT1 knockout cells. One possible explanation for these contradictory results is that the W473/M477A mutant utilized by Kuwabara et al. may have some residual activity.

Future work to test this explanation could be done in several ways. First, POMGnT1 activity assay on the secreted W473/M477A mutant could be performed overnight, rather than one hour. Secondly, the cell lysates showing IIH6 reactivity after cells were transfected with the full length W473/M477A mutant could be analyzed by mass spectrometry using glycoproteomic approaches to determine whether core M1 structures are present. The presence of core M1 structures would indicate that the W473/M477A mutant had some residual activity.

We identified spectra of the core M3 structure on α-DG modified by fukutin in POMGnT1 knockout cells. However, we were not able to determine the frequency with which the core M3 structure is extended past this step. This is due to our inability to detect peptides modified by the full matriglycan-containing structure by mass spectrometry due to its very large mass. One possible solution to this would be to utilize exo-deglycosylases to remove the terminal glucuronic acid and xylose from the core M3 structure. This would allow us to compare the frequency of various stages of the core M3 structure between our POMGnT1 knockout and Wt cells.

Support for the POMGnT1-fukutin hypothesis also comes from the previously reported molecular interaction between fukutin and POMGnT1 (Xiong, H., Kobayashi, K., et al. 2006). This interaction was demonstrated by co-immunoprecipitation of overexpressed FLAG- and V5-tagged fukutin and POMGnT1. An improved method of determining whether POMGnT1

interacts specifically with fukutin or several enzymes in the core M3 synthesis process is to repeat this pull-down of tagged POMGnT1, but analyze the sample by mass spectrometry to identify all proteins present.

However, this method is still not ideal in that any method utilized to lyse cells may disrupt the POMGnT1-protein interactions as well. Alternatively, a method of identifying proximal and interacting proteins in mammalian cells was recently described by Roux et al. They utilized a promiscuous biotin ligase fusion protein to biotin-label interacting proteins (Roux, K.J., Kim, D.I., et al. 2012). A construct of POMGnT1 with the BirA fusion protein linked to the c-terminal end could label the POMGnT1- interactome with biotin. These biotinylated proteins would then be isolated from cell lysates by affinity chromatography and identified by mass spectrometry.

The list of identified proteins would be compared to a list obtained from a negative control. An ideal negative control for this would be the first 60 amino acids of POMGnT1, the signal sequence and the transmembrane domain, attached to the BirA construct. This negative control would identify any proteins which do not directly interact with POMGnT1 but were labelled erroneously due to their proximity in the cis-Golgi. An interesting additional experimental construct to test would be the POMGnT1 protein without the stem domain. Kuwabara et al. showed that the stem domain of POMGnT1 recognizes β-linked GlcNAc. Thus, the list of biotin-labelled proteins interacting with full length POMGnT1 could be very different than the list of proteins interacting with POMGnT1Δstem. The list of proteins in the POMGnT1- interactome could also be used to begin identifying other proteins which are substrates for POMGnT1. This was the original intent of the work that led to the discovery of cadherins and plexins as major O-mannosylated proteins.

Determining the substrates of POMGnT1 would be a significant advancement. At present, only a handful of POMGnT1 substrates are known, the most well defined being α-DG. In the α-DG-lacking brain samples from mice analyzed by Stalnaker et al. the core M1 and core M2 O-mannose structures were not detectibly different than the structures obtained from the brains of wildtype. Thus, the O-mannose glycans in the brain, which comprise nearly a third of all O-glycans in the mammalian brain, modify undefined proteins. Furthermore, with the new understanding of how O-mannosylated cadherins escape modification by POMGnT1, we have a better understanding of what defines a POMGnT1 substrate.

The work presented in this dissertation advances knowledge of POMGnT1 substrates. Understanding how cadherins escape POMGnT1 modification adds to our understanding of the relationship between the recently discovered O-mannosylation pathway for cadherins and the classical O-mannosylation pathway for alpha-dystroglycan. A better understanding of how loss of POMGnT1 affects assembly of the core M3 structure aids in the search to better understand these severe congenital disorders.

References

Basilico C, Arnesano A, Galluzzo M, Comoglio PM, Michieli P. 2008. A high affinity hepatocyte growth factor-binding site in the immunoglobulin-like region of Met. The Journal of biological chemistry, 283:21267-21277.

Biancheri R, Bertini E, Falace A, Pedemonte M, Rossi A, D'Amico A, Scapolan S, Bergamino L, Petrini S, Cassandrini D, et al. 2006. POMGnT1 mutations in congenital muscular dystrophy: genotype-phenotype correlation and expanded clinical spectrum. Archives of neurology, 63:1491-1495.

Diesen C, Saarinen A, Pihko H, Rosenlew C, Cormand B, Dobyns WB, Dieguez J, Valanne L, Joensuu T, Lehesjoki AE. 2004. POMGnT1 mutation and phenotypic spectrum in muscle-eye-brain disease. Journal of medical genetics, 41:e115.

Grewal PK, Hewitt JE. 2002. Mutation of Large, which encodes a putative glycosyltransferase, in an animal model of muscular dystrophy. Biochimica et biophysica acta, 1573:216-224.

Halmo SM, Singh D, Patel S, Wang S, Edlin M, Boons GJ, Moremen KW, Live D, Wells L. 2017. Protein O-Linked Mannose beta-1,4-N-Acetylglucosaminyl-transferase 2 (POMGNT2) Is a Gatekeeper Enzyme for Functional Glycosylation of alpha-Dystroglycan. The Journal of biological chemistry, 292:2101-2109.

Harrison OJ, Jin X, Hong S, Bahna F, Ahlsen G, Brasch J, Wu Y, Vendome J, Felsovalyi K, Hampton CM, et al. 2011. The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. Structure (London, England : 1993), 19:244-256.

Kuwabara N, Manya H, Yamada T, Tateno H, Kanagawa M, Kobayashi K, Akasaka-Manya K, Hirose Y, Mizuno M, Ikeguchi M, et al. 2016. Carbohydrate-binding domain of the POMGnT1 stem region modulates O-mannosylation sites of α-dystroglycan. Proc Natl Acad Sci U S A, 113:9280-9285.

Larsen ISB, Narimatsu Y, Clausen H, Joshi HJ, Halim A. 2019. Multiple distinct O-Mannosylation pathways in eukaryotes. Current opinion in structural biology, 56:171-178.

Larsen ISB, Narimatsu Y, Joshi HJ, Siukstaite L, Harrison OJ, Brasch J, Goodman KM, Hansen L, Shapiro L, Honig B, et al. 2017a. Discovery of an O-mannosylation pathway selectively serving cadherins and protocadherins. Proceedings of the National Academy of Sciences, 114:11163-11168.

Larsen ISB, Narimatsu Y, Joshi HJ, Yang Z, Harrison OJ, Brasch J, Shapiro L, Honig B, Vakhrushev SY, Clausen H, et al. 2017b. Mammalian O-mannosylation of cadherins and plexins is independent of protein O-mannosyltransferases 1 and 2. The Journal of biological chemistry, 292:11586-11598.

Liu J, Ball SL, Yang Y, Mei P, Zhang L, Shi H, Kaminski HJ, Lemmon VP, Hu H. 2006. A genetic model for muscle-eye-brain disease in mice lacking protein O-mannose 1,2-N-acetylglucosaminyltransferase (POMGnT1). Mechanisms of development, 123:228-240.

Lommel M, Winterhalter PR, Willer T, Dahlhoff M, Schneider MR, Bartels MF, Renner-Müller I, Ruppert T, Wolf E, Strahl S. 2013. Protein O-mannosylation is crucial for E-cadherin–mediated cell adhesion. Proceedings of the National Academy of Sciences, 110:21024-21029.

Mandal DK, Bhattacharyya L, Koenig SH, Brown RD, 3rd, Oscarson S, Brewer CF. 1994. Studies of the binding specificity of concanavalin A. Nature of the extended binding site for asparagine-linked carbohydrates. Biochemistry, 33:1157-1162.

Moore SA, Saito F, Chen J, Michele DE, Henry MD, Messing A, Cohn RD, Ross-Barta SE, Westra S, Williamson RA, et al. 2002. Deletion of brain dystroglycan recapitulates aspects of congenital muscular dystrophy. Nature, 418:422-425.

Roux KJ, Kim DI, Raida M, Burke B. 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. The Journal of Cell Biology, 196:801-810.

Sparks SE, Quijano-Roy S, Harper A, Rutkowski A, Gordon E, Hoffman EP, Pegoraro E. 1993. Congenital Muscular Dystrophy Overview. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, Amemiya A editors. GeneReviews((R)). Seattle (WA): University of Washington, Seattle

University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved.

Stalnaker SH, Aoki K, Lim J-M, Porterfield M, Liu M, Satz JS, Buskirk S, Xiong Y, Zhang P, Campbell KP, et al. 2011. Glycomic Analyses of Mouse Models of Congenital Muscular Dystrophy. Journal of Biological Chemistry, 286:21180-21190.

Vester-Christensen MB, Halim A, Joshi HJ, Steentoft C, Bennett EP, Levery SB, Vakhrushev SY, Clausen H. 2013. Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. Proceedings of the National Academy of Sciences, 110:21018-21023.

Xiong H, Kobayashi K, Tachikawa M, Manya H, Takeda S, Chiyonobu T, Fujikake N, Wang F, Nishimoto A, Morris GE, et al. 2006. Molecular interaction between fukutin and POMGnT1 in the glycosylation pathway of alpha-dystroglycan. Biochemical and biophysical research communications, 350:935-941.

Xu M, Yamada T, Sun Z, Eblimit A, Lopez I, Wang F, Manya H, Xu S, Zhao L, Li Y, et al. 2016. Mutations in POMGNT1 cause non-syndromic retinitis pigmentosa. Hum Mol Genet, 25:1479-1488.

Yoshida-Moriguchi T, Willer T, Anderson ME, Venzke D, Whyte T, Muntoni F, Lee H, Nelson SF, Yu L, Campbell KP. 2013. SGK196 is a glycosylation-specific O-mannose kinase required for dystroglycan function. Science (New York, N.Y.), 341:896-899.

APPENDIX A

EXPERIMENTS PROBING THE FUNCTION OF M0 STRUCTURES ON HEPATOCYTE GROWTH FACTOR RECEPTOR

Figure A-1A

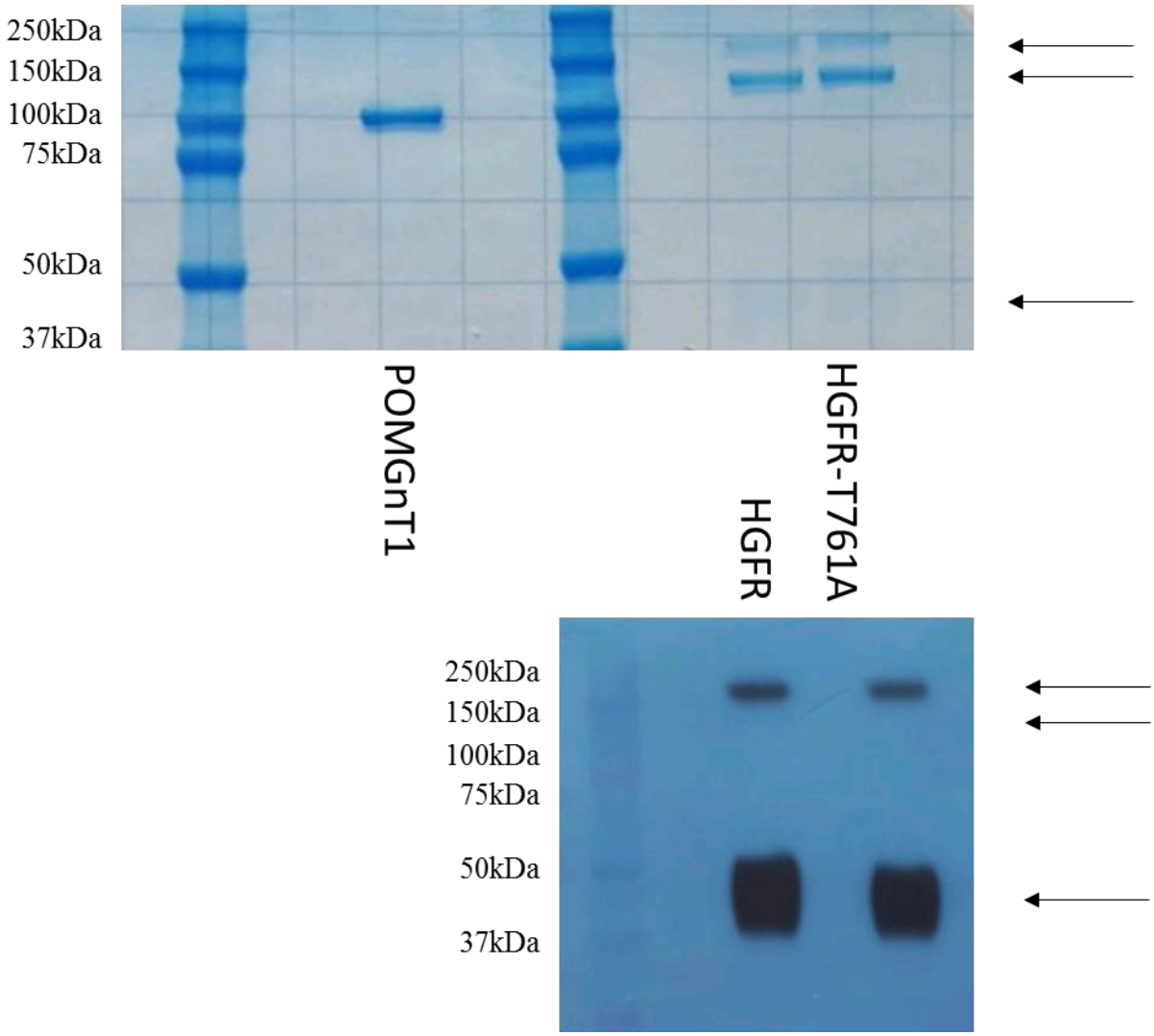

250kDa
150kDa
100kDa
75kDa
50kDa
37kDa
POMGnT1
HGFR
HGFR-T761A
250kDa
150kDa
100kDa
75kDa
50kDa
37kDa

Figure A-1B

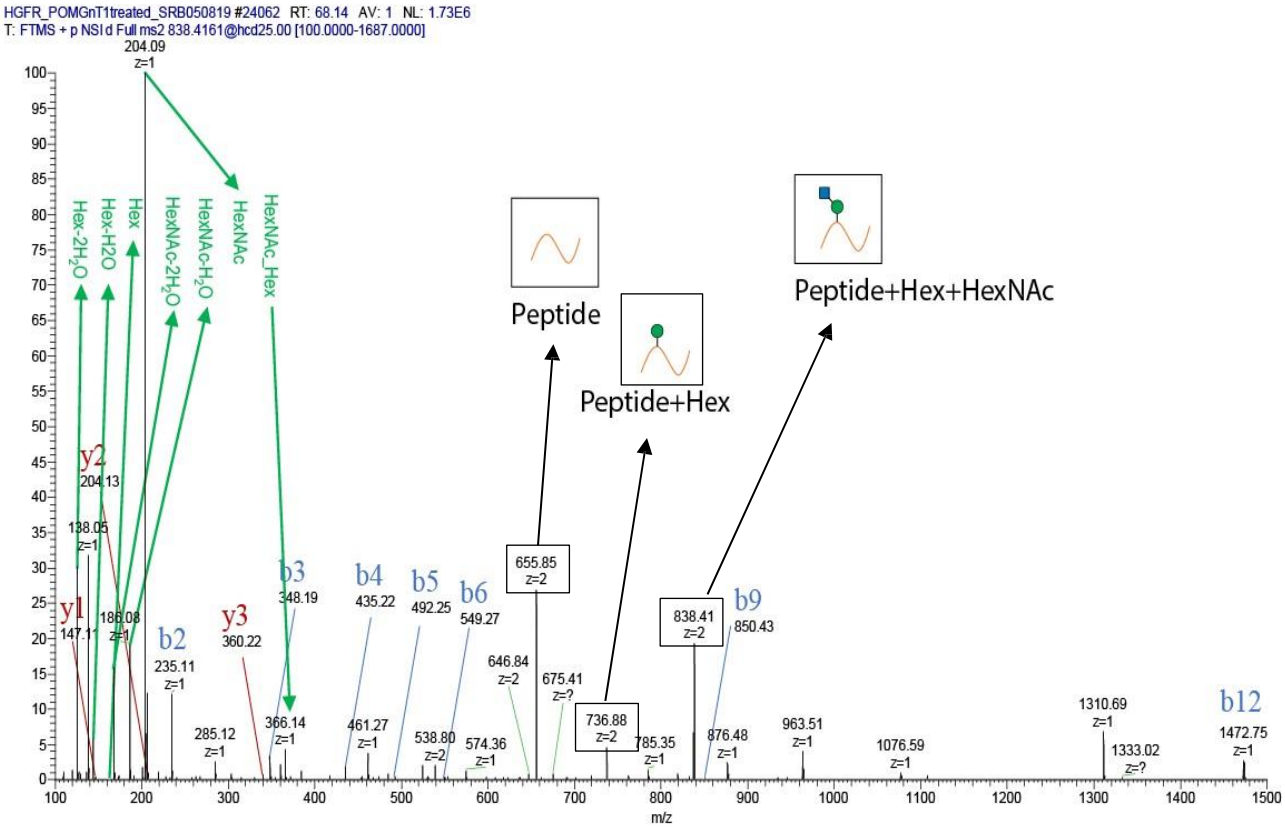

HGFR_POMGnT1treated_SRB050819 #24062 RT: 68.14 AV: 1 NL: 1.73E6
T: FTMS + p NSI d Full ms2 838.4161@hcd25.00 [100.0000-1687.0000]
204.09 z=1
Hex-2H2O
Hex-H2O
Hex
HexNAc-2H2O
HexNAc-H2O
HexNAc
HexNAc_Hex
Peptide
Peptide+Hex
Peptide+Hex+HexNAc
y2 204.13
138.05 z=1
y1 147.11
186.08 z=1
b2 235.11 z=1
y3 360.22
b3 348.19
b4 435.22
b5 492.25
b6 549.27
655.85 z=2
646.84 z=2
675.41 z=?
736.88 z=2
785.35 z=1
838.41 z=2
b9 850.43
876.48 z=1
963.51 z=1
1076.59 z=1
1310.69 z=1
1333.02 z=?
b12 1472.75 z=1
285.12 z=1
366.14 z=1
461.27 z=1
538.80 z=2
574.36 z=1
m/z

Figure A-1. Characterization of recombinantly expressed HGFR. (A) POMGnT1-pGEc2, HGFR-pGEc2, and HGFR-T761A mutant constructs expressed in HEK 293F cells. Both HGFR constructs appeared as three bands on the Coomassie gel. There was one band at, and one band just below, the expected 112kDa size. There was also a very faint, very wide band from ~37kDa to ~50kDa. A western blot probing with an antibody generated against the first 100 amino acid residues of the N-terminus of HGFR detected two bands, with the smaller wide band being stained disproportionally more by the antibody. (B) Annotated spectrum of an HGFR peptide modified by POMGnT1.

Figure A-2A

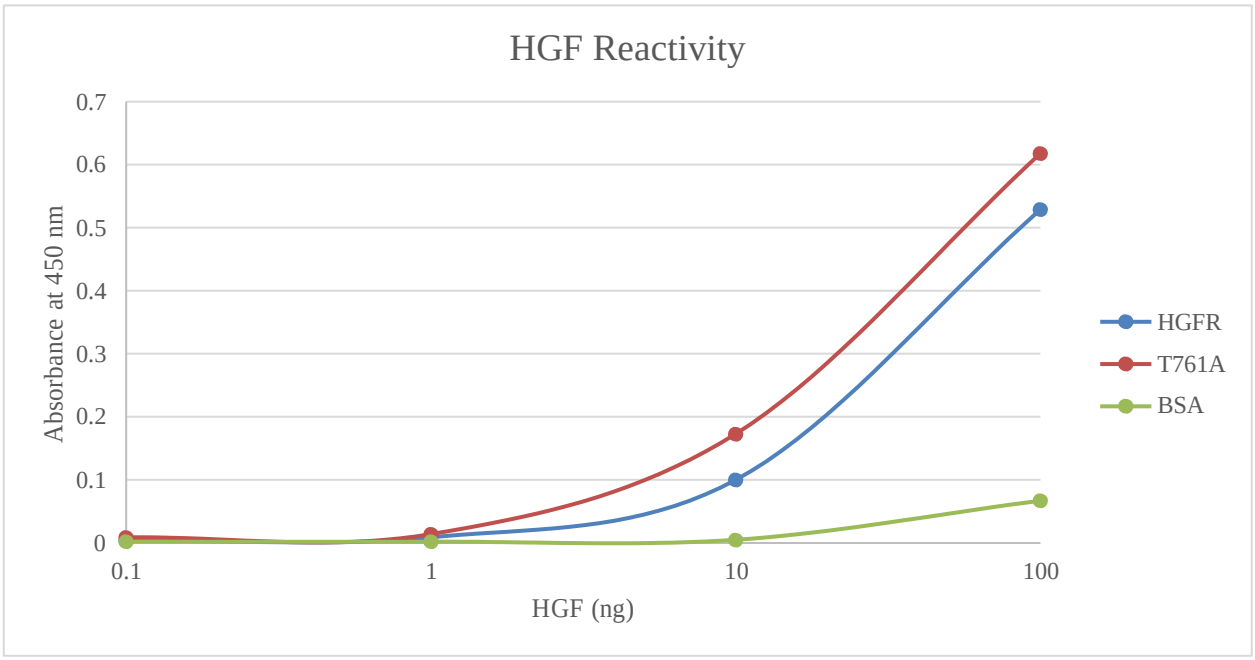

HGF Reactivity
Absorbance at 450 nm
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.1
1
10
100
HGF (ng)
HGFR
T761A
BSA

Figure A-2B

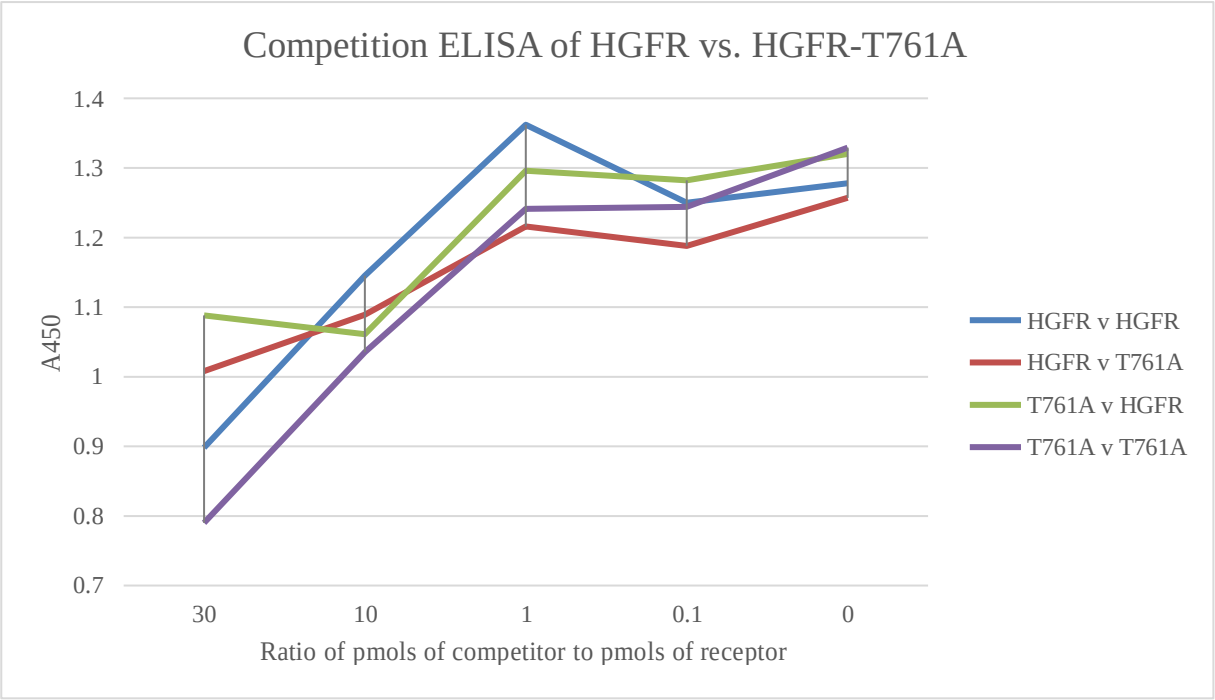

Competition ELISA of HGFR vs. HGFR-T761A
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
A450
30
10
1
0.1
0
Ratio of pmols of competitor to pmols of receptor
HGFR v HGFR
HGFR v T761A
T761A v HGFR
T761A v T761A

Figure A-2. Mutation of T761 to alanine does not affect the binding of HGFR to its ligand. (A) HGFR, HGFR-T761A, and BSA were coated onto a plate. After blocking wells were treated with an increasing amount of HGF ligand, then antibodies against the HGF ligand and a HRP- linked secondary antibody. The assay was developed with TMB substrate and quenched with 2N H_2SO_4. There was no significant difference between the amount of HGF bound to HGFR and the amount bound to HGFR-T761A. (B) The dynamic binding abilities of HGFR and HGFR- T761A were tested by incubating each with the HGF ligand and then testing the ability of HGFR or HGFR-T761A to bind the pre-bound HGF. There was no difference in the dynamic binding of HGFR or HGFR-T761A to the HGF ligand.

APPENDIX B

ANNOTATED SPECTRA OF SITES OF O-MANNOSYLATION ON E-CADHERIN, N-CADHERIN, AND ALPHA-DYSTROGLYCAN

Figure B-1A

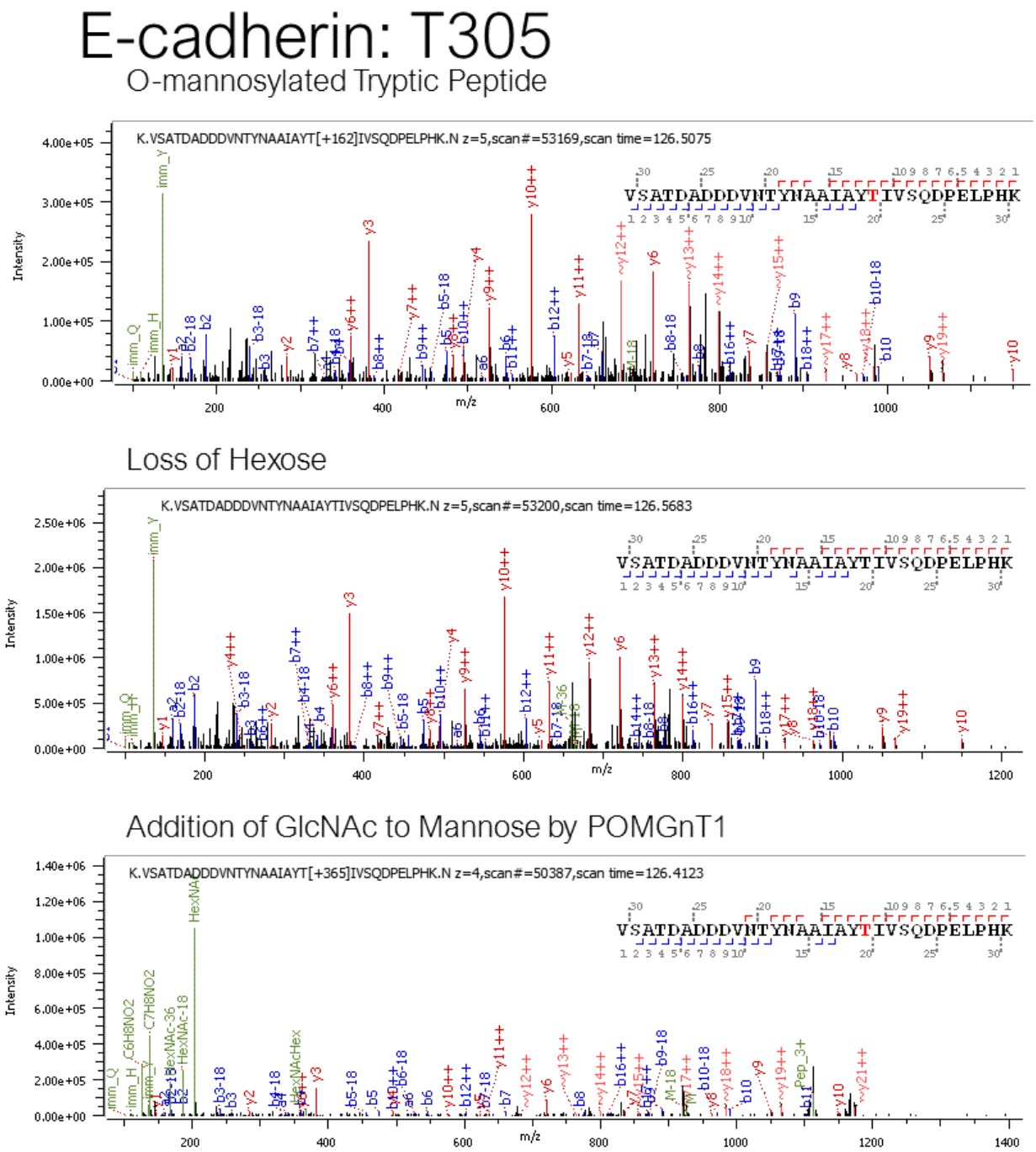

E-cadherin: T305
O-mannosylated Tryptic Peptide
K.VSATDADDDVNTYNAAIAYT[+162]IVSQDPELPHK.N z=5,scan#=53169,scan time=126.5075
VSATDADDDVNTYNAAIAYTIVSQDPELPHK
Intensity
m/z
Loss of Hexose
K.VSATDADDDVNTYNAAIAYTIVSQDPELPHK.N z=5,scan#=53200,scan time=126.5683
VSATDADDDVNTYNAAIAYTIVSQDPELPHK
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
K.VSATDADDDVNTYNAAIAYT[+365]IVSQDPELPHK.N z=4,scan#=50387,scan time=126.4123
VSATDADDDVNTYNAAIAYTIVSQDPELPHK
Intensity
m/z

Figure B-1B

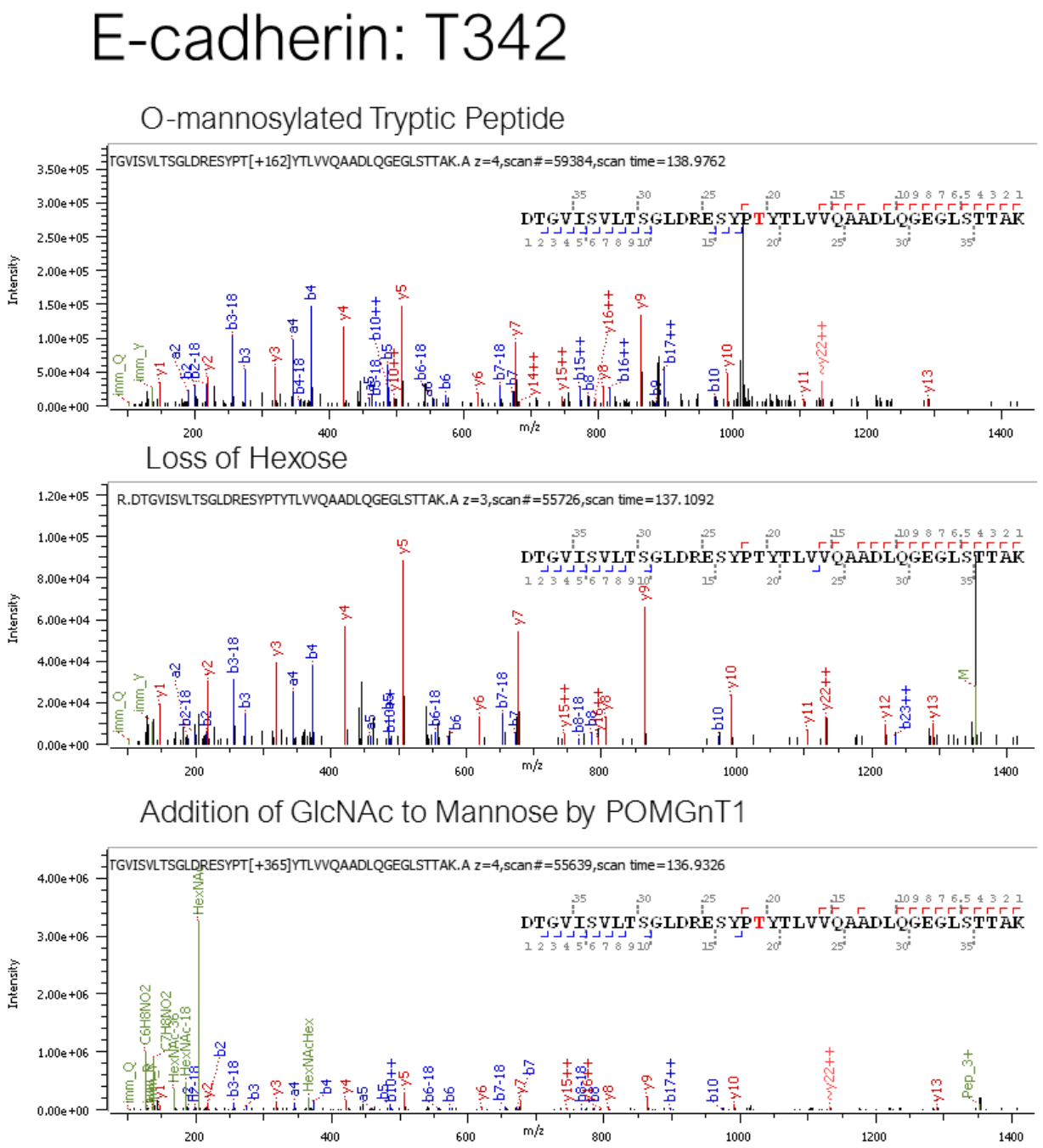

E-cadherin: T342
O-mannosylated Tryptic Peptide
TGVISVLTSGLDRESYPT[+162]YTLVVQAADLQGEGLSTTAK.A z=4,scan#=59384,scan time=138.9762
DTGVISVLTSGLDRESYPTYTLVVQAADLQGEGLSTTAK
Intensity
m/z
Loss of Hexose
R.DTGVISVLTSGLDRESYPTYTLVVQAADLQGEGLSTTAK.A z=3,scan#=55726,scan time=137.1092
DTGVISVLTSGLDRESYPTYTLVVQAADLQGEGLSTTAK
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
TGVISVLTSGLDRESYPT[+365]YTLVVQAADLQGEGLSTTAK.A z=4,scan#=55639,scan time=136.9326
DTGVISVLTSGLDRESYPTYTLVVQAADLQGEGLSTTAK
Intensity
m/z

Figure B-1C

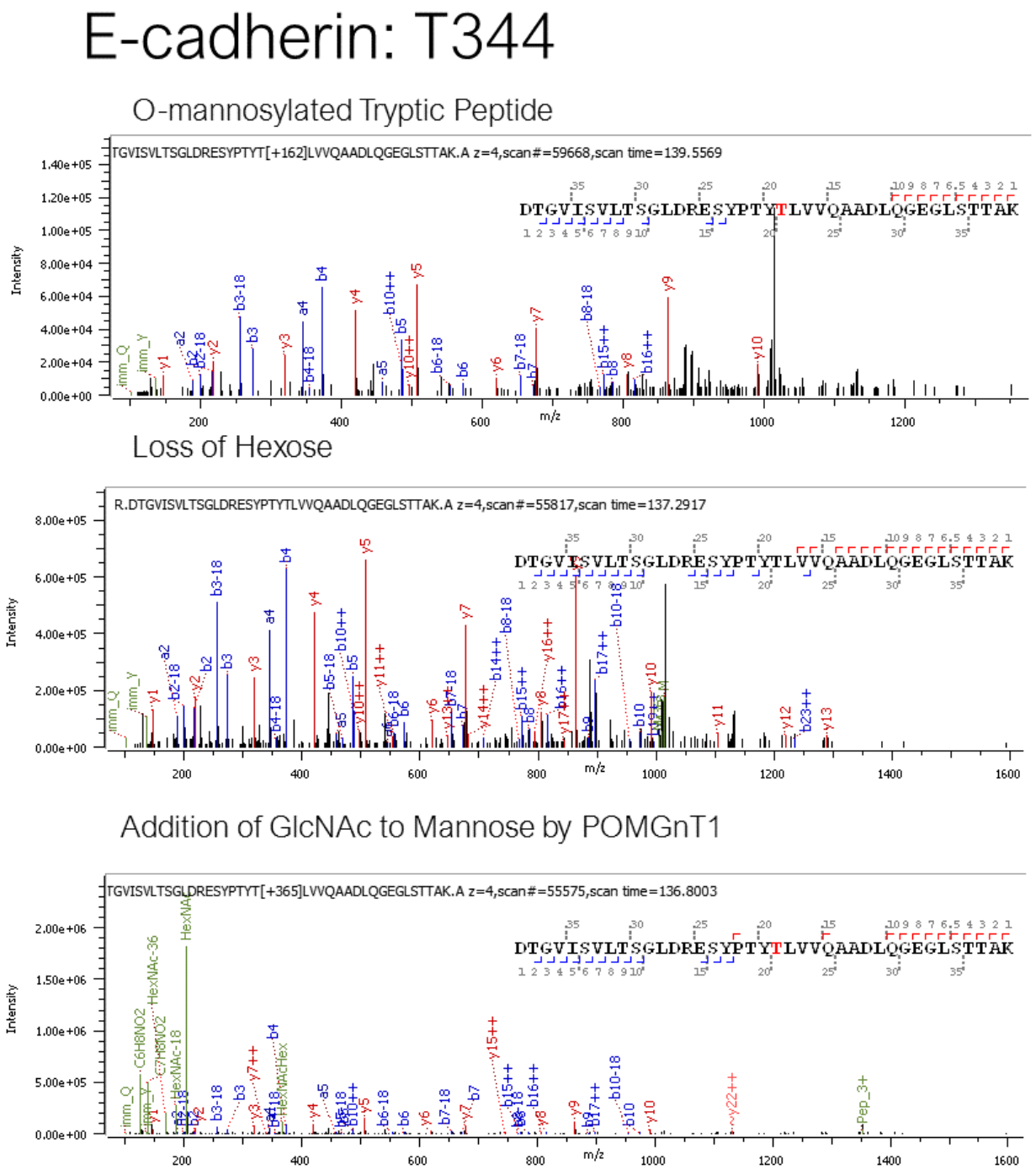

E-cadherin: T344
O-mannosylated Tryptic Peptide
TGVISVLTSGLDRESYPTYT[+162]LVVQAADLQGEGLSTTAK.A z=4,scan#=59668,scan time=139.5569
DTGVISVLTSGLDRESYPTYTLVVQAADLQGEGLSTTAK
Intensity
m/z
Loss of Hexose
R.DTGVISVLTSGLDRESYPTYTLVVQAADLQGEGLSTTAK.A z=4,scan#=55817,scan time=137.2917
DTGVISVLTSGLDRESYPTYTLVVQAADLQGEGLSTTAK
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
TGVISVLTSGLDRESYPTYT[+365]LVVQAADLQGEGLSTTAK.A z=4,scan#=55575,scan time=136.8003
DTGVISVLTSGLDRESYPTYTLVVQAADLQGEGLSTTAK
Intensity
m/z

Figure B-1D

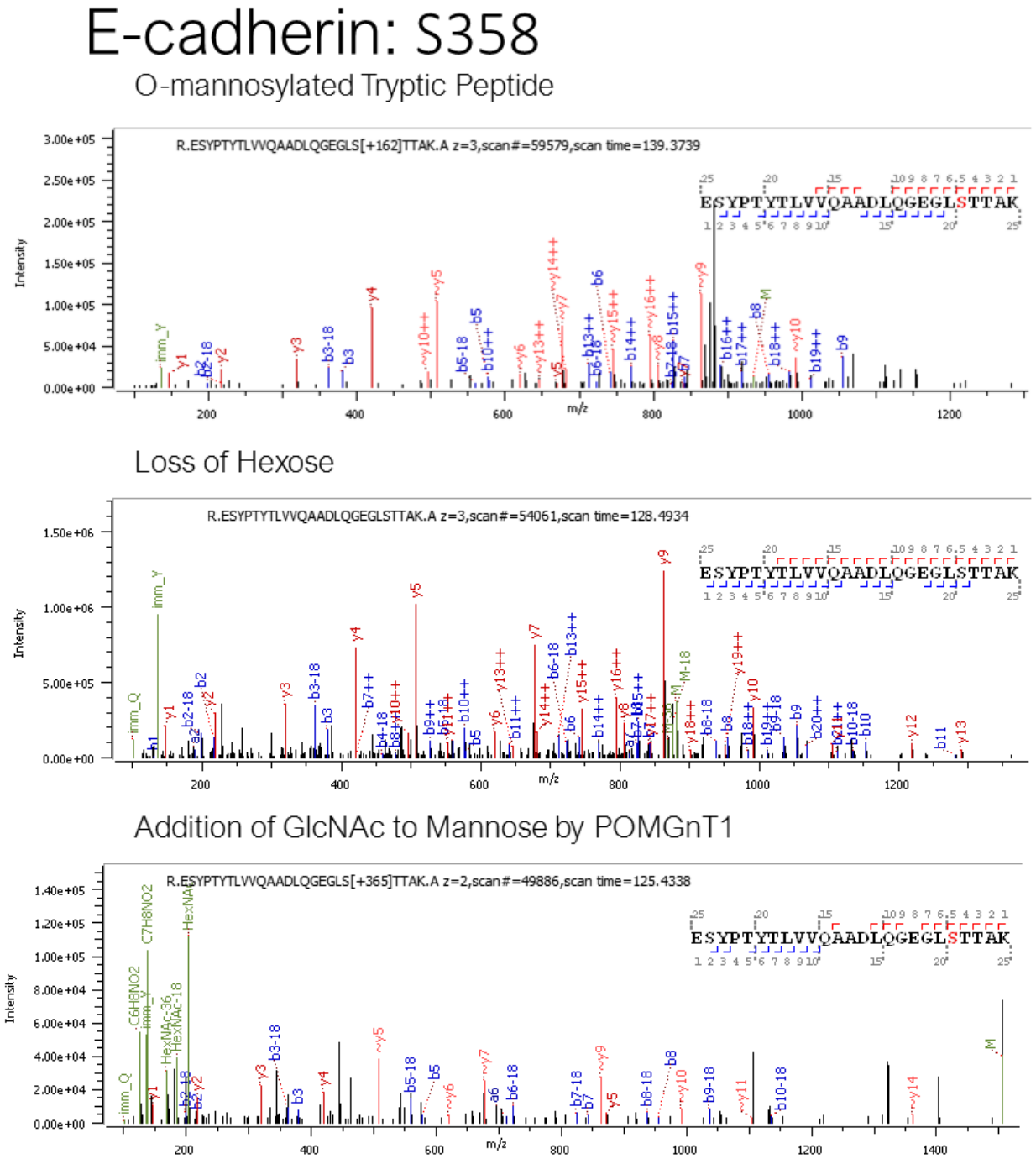

E-cadherin: S358
O-mannosylated Tryptic Peptide
R.ESYPTYTLVVQAADLQGEGLS[+162]TTAK.A z=3,scan#=59579,scan time=139.3739
ESYPTYTLVVQAADLQGEGLSTTAK
Intensity
m/z
Loss of Hexose
R.ESYPTYTLVVQAADLQGEGLSTTAK.A z=3,scan#=54061,scan time=128.4934
ESYPTYTLVVQAADLQGEGLSTTAK
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
R.ESYPTYTLVVQAADLQGEGLS[+365]TTAK.A z=2,scan#=49886,scan time=125.4338
ESYPTYTLVVQAADLQGEGLSTTAK
Intensity
m/z

Figure B-1E

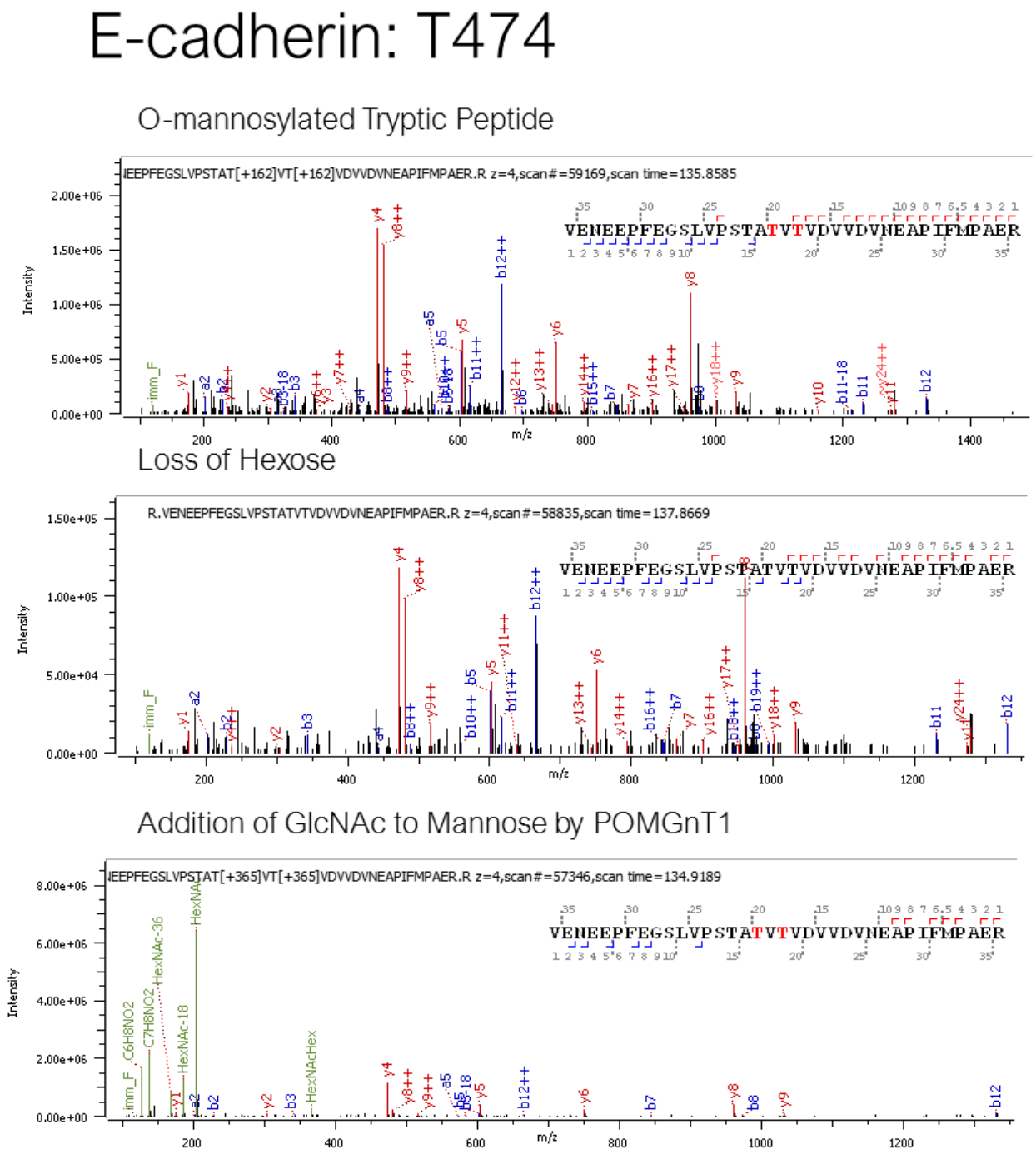

E-cadherin: T474
O-mannosylated Tryptic Peptide
IEEPFEGSLVPSTAT[+162]VT[+162]VDVVDVNEAPIFMPAER.R z=4,scan#=59169,scan time=135.8585
VENEEPFEGSLVPSTATVTVDVVDVNEAPIFMPAER
Intensity
m/z
Loss of Hexose
R.VENEEPFEGSLVPSTATVTVDVVDVNEAPIFMPAER.R z=4,scan#=58835,scan time=137.8669
VENEEPFEGSLVPSTATVTVDVVDVNEAPIFMPAER
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
IEEPFEGSLVPSTAT[+365]VT[+365]VDVVDVNEAPIFMPAER.R z=4,scan#=57346,scan time=134.9189
VENEEPFEGSLVPSTATVTVDVVDVNEAPIFMPAER
Intensity
m/z

Figure B-1F

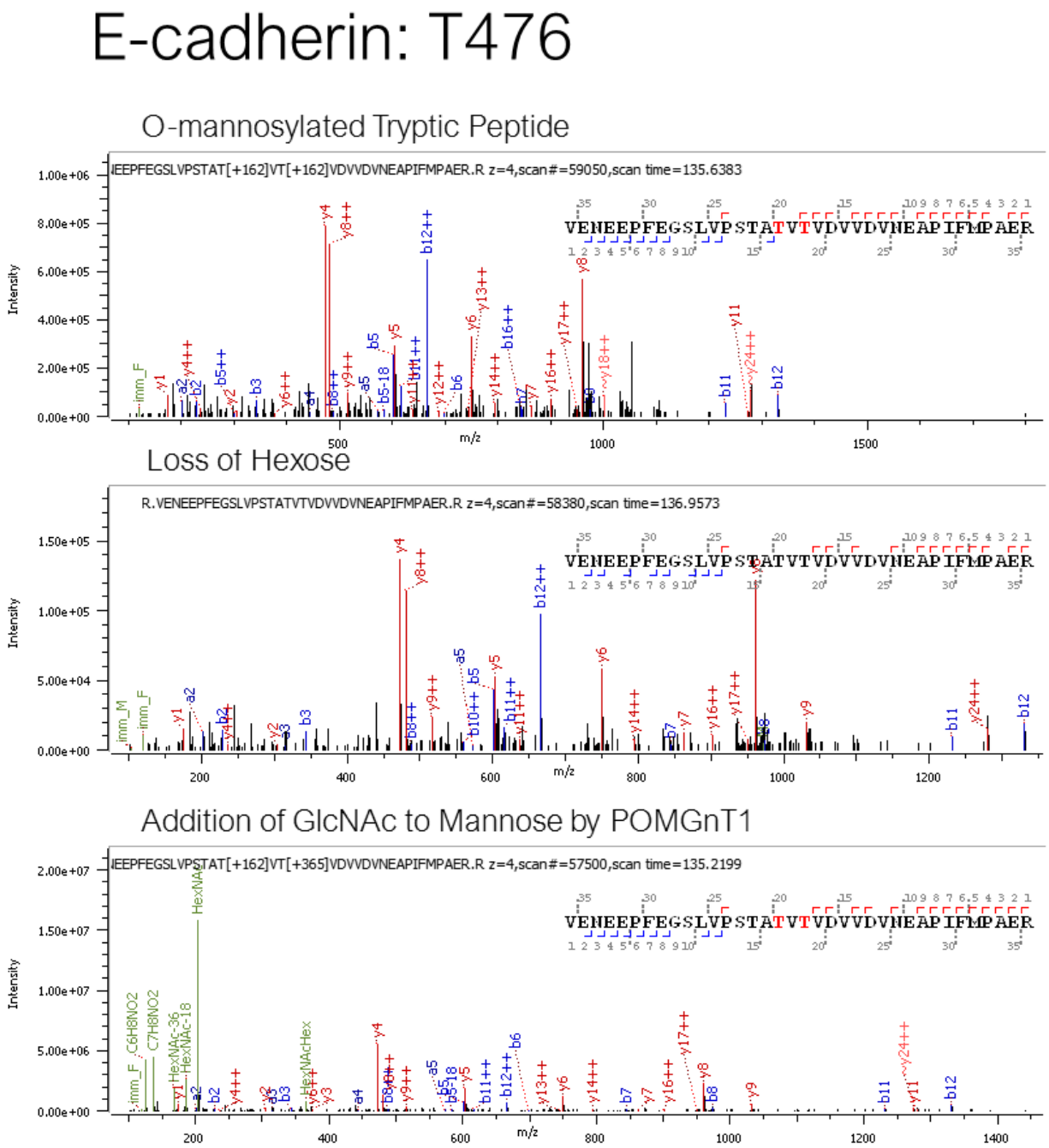

E-cadherin: T476
O-mannosylated Tryptic Peptide
IEEPFEGSLVPSTAT[+162]VT[+162]VDVVDVNEAPIFMPAER.R z=4,scan#=59050,scan time=135.6383
VENEEPFEGSLVPSTATVTVDVVDVNEAPIFMPAER
Intensity
m/z
Loss of Hexose
R.VENEEPFEGSLVPSTATVTVDVVDVNEAPIFMPAER.R z=4,scan#=58380,scan time=136.9573
VENEEPFEGSLVPSTATVTVDVVDVNEAPIFMPAER
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
IEEPFEGSLVPSTAT[+162]VT[+365]VDVVDVNEAPIFMPAER.R z=4,scan#=57500,scan time=135.2199
VENEEPFEGSLVPSTATVTVDVVDVNEAPIFMPAER
Intensity
m/z

Figure B-1G

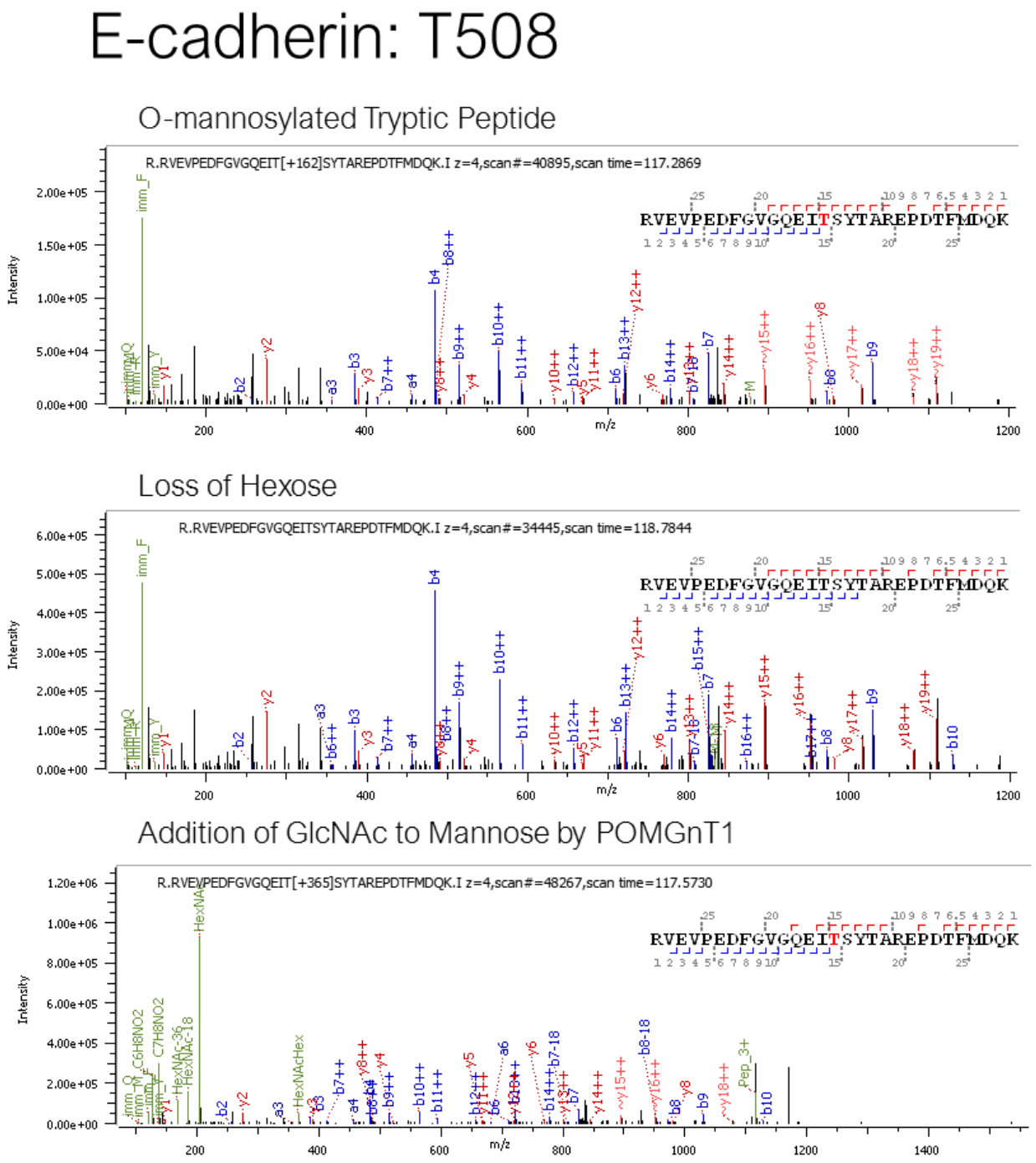

E-cadherin: T508
O-mannosylated Tryptic Peptide
R.RVEVPEDFGVGQEIT[+162]SYTAREPDTFMDQK.I z=4,scan#=40895,scan time=117.2869
RVEVPEDFGVGQEITSYTAREPDTFMDQK
Intensity
m/z
Loss of Hexose
R.RVEVPEDFGVGQEITSYTAREPDTFMDQK.I z=4,scan#=34445,scan time=118.7844
RVEVPEDFGVGQEITSYTAREPDTFMDQK
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
R.RVEVPEDFGVGQEIT[+365]SYTAREPDTFMDQK.I z=4,scan#=48267,scan time=117.5730
RVEVPEDFGVGQEITSYTAREPDTFMDQK
Intensity
m/z

Figure B-1H

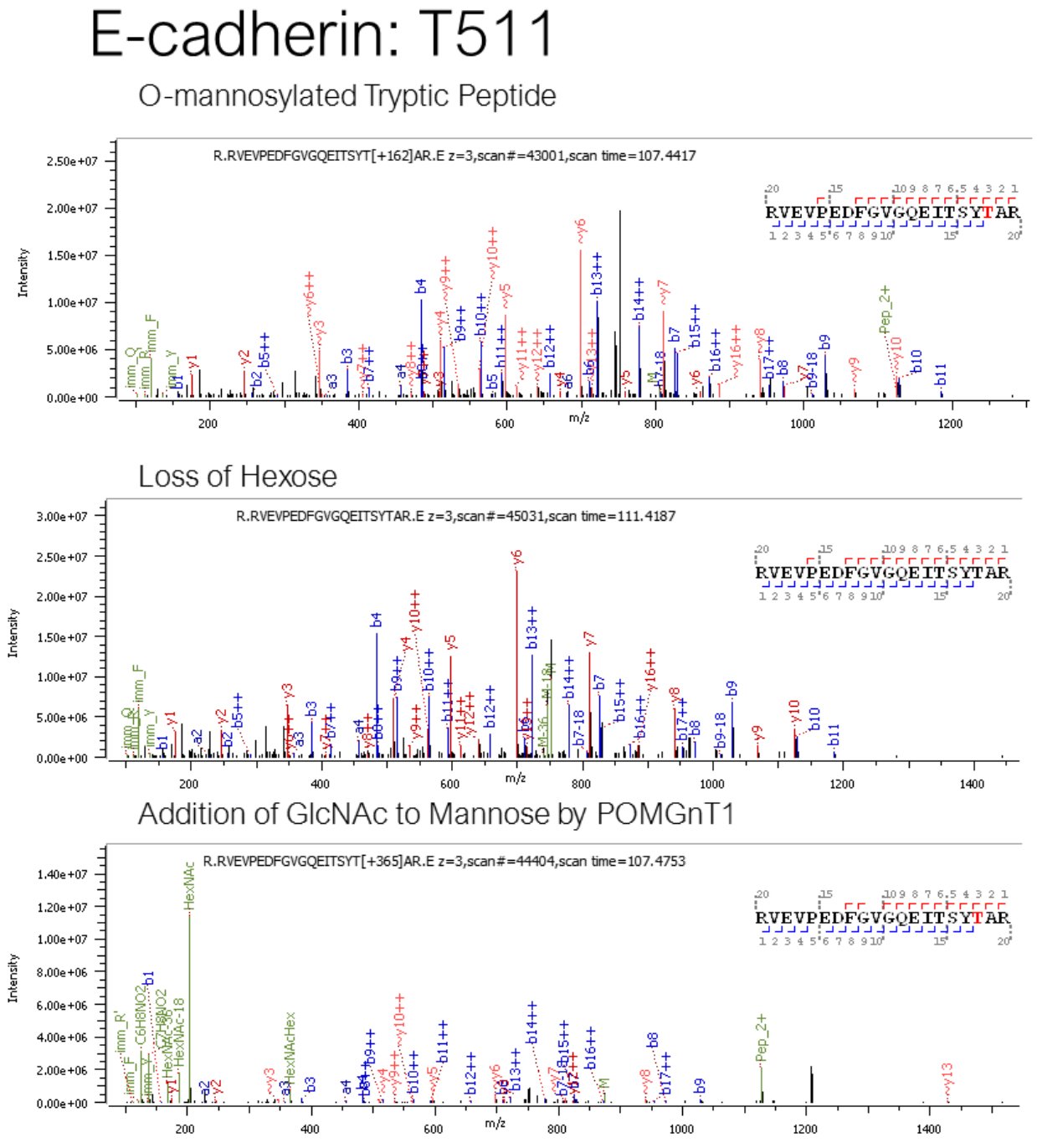
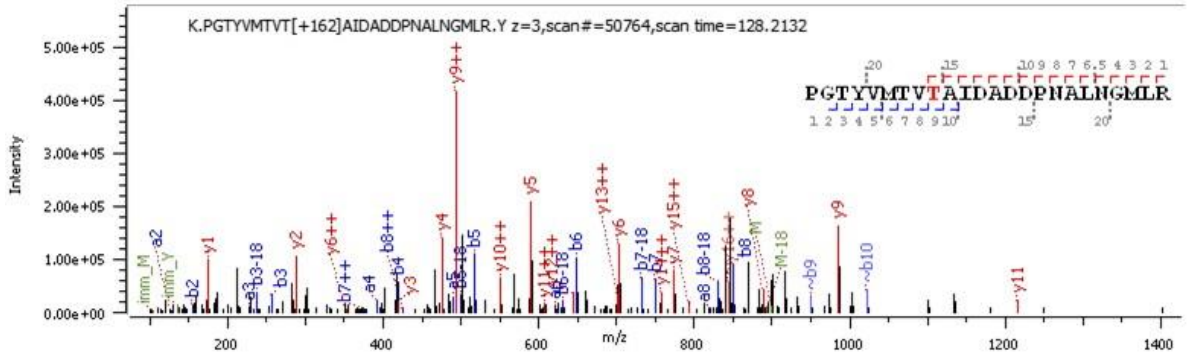

E-cadherin: T511
O-mannosylated Tryptic Peptide
R.RVEVPEDFGVGQEITSYT[+162]AR.E z=3,scan#=43001,scan time=107.4417
RVEVPEDFGVGQEITSYTAR
Intensity
m/z
Loss of Hexose
R.RVEVPEDFGVGQEITSYTAR.E z=3,scan#=45031,scan time=111.4187
RVEVPEDFGVGQEITSYTAR
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
R.RVEVPEDFGVGQEITSYT[+365]AR.E z=3,scan#=44404,scan time=107.4753
RVEVPEDFGVGQEITSYTAR
HexNAc
HexNAcHex
Intensity
m/z

Figure B-1. Tandem MS/MS spectra of glycopeptides from E-cadherin. The top panel shows hexose (+162) on peptides from HEK Wt-expressed E-cadherin. The middle panel shows the loss of O-mannose (lack of +162) after α-mannosidase treatment on the HEK Wt-expressed E-cadherin peptides. The bottom panel shows the addition of GlcNAc to O-mannose (+365) after treatment of peptides from HEK-POMGnT1 expressed E-cadherin with POMGnT1. (A) T305. (B) T342. (C) T344. (D) S358. (E) T474. (F) 476. (G) T506. (H) T511.

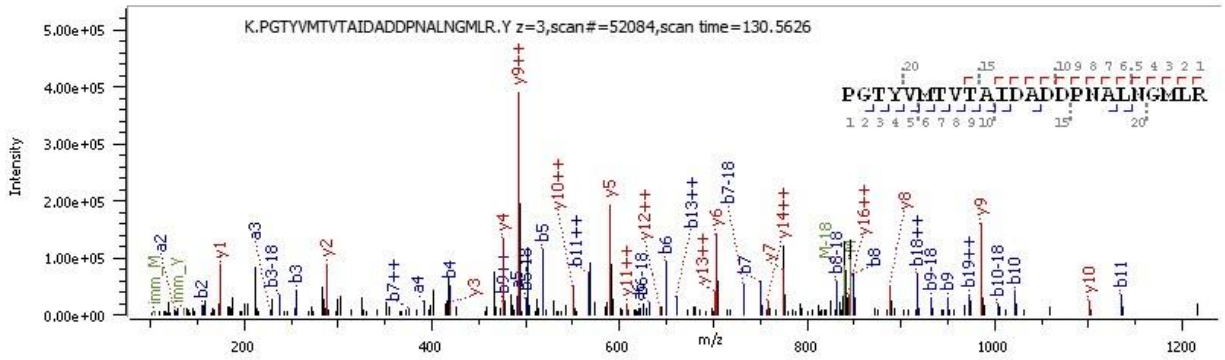
Figure B-2A

T290

O-mannosylated Tryptic Peptide



Loss of Hexose



Addition of GlcNAc to Mannose by POMGnT1

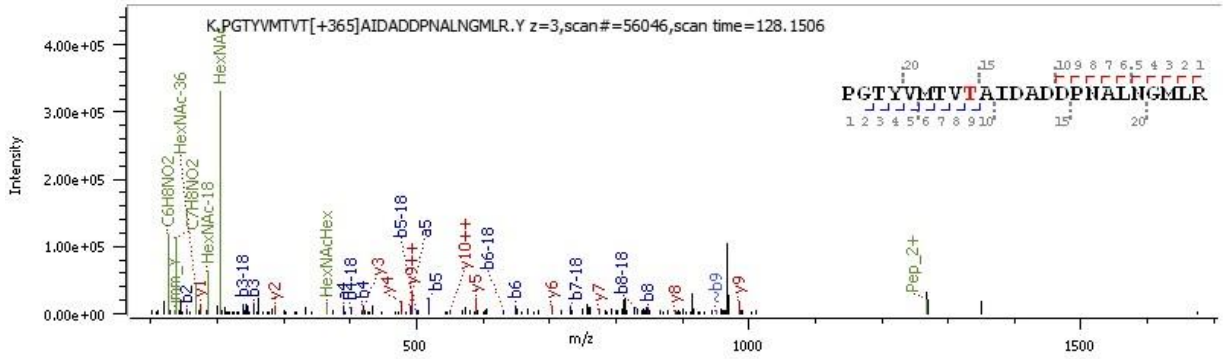


Figure B-2B

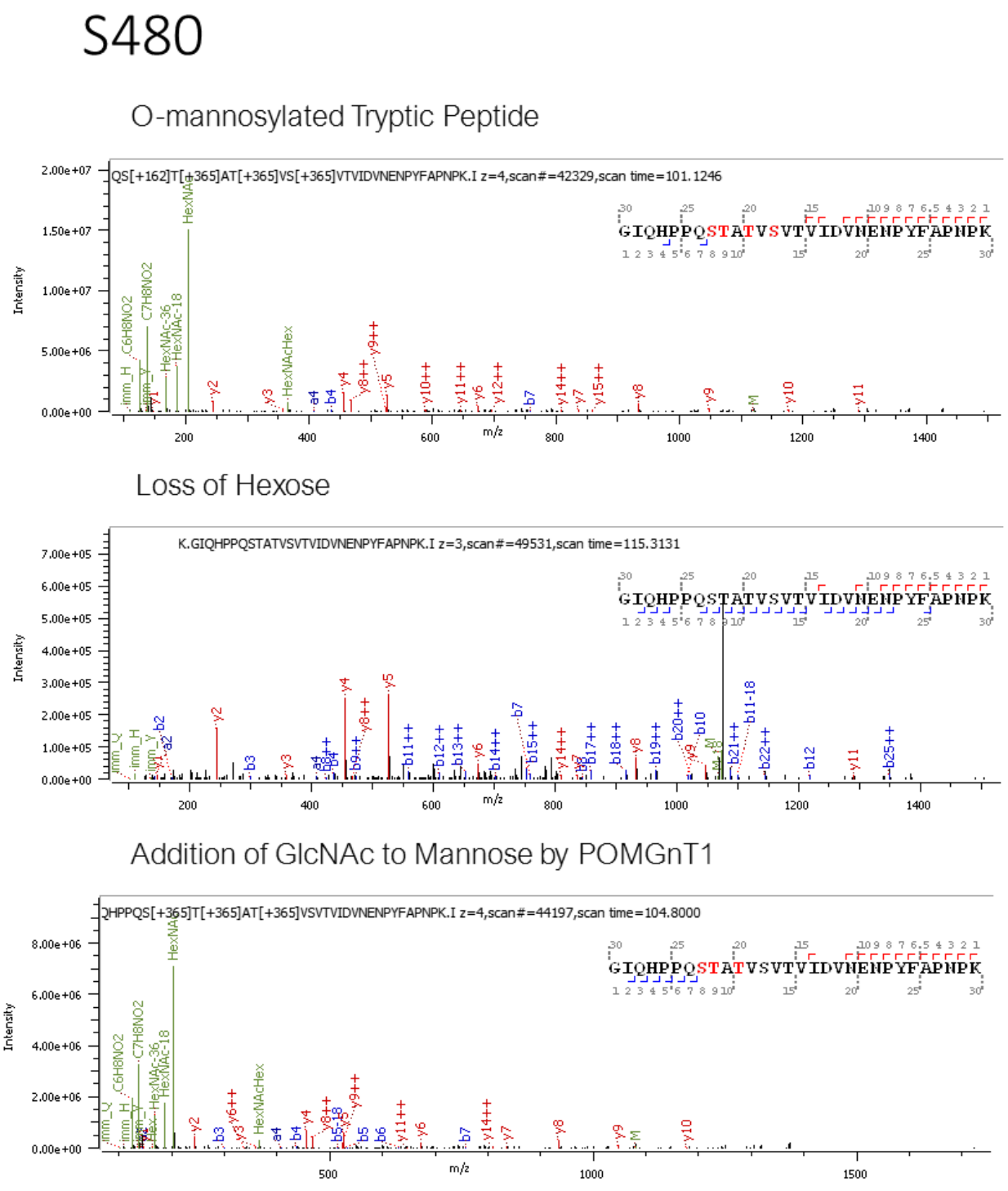
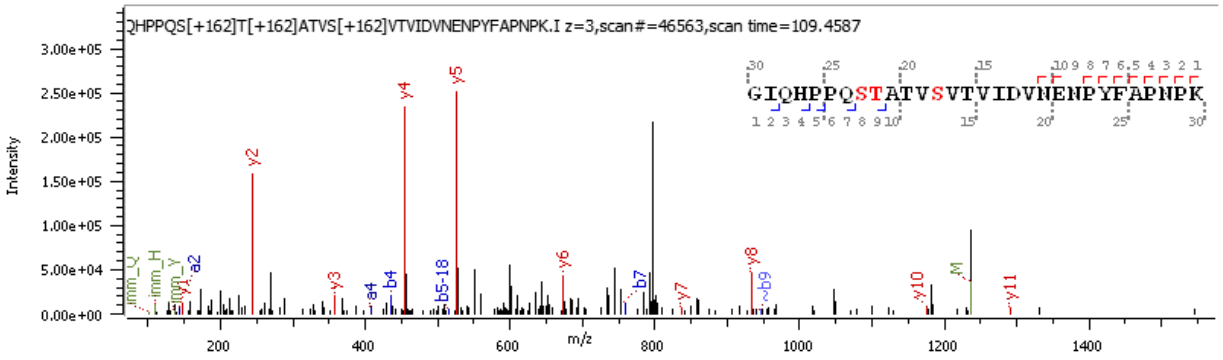

S480
O-mannosylated Tryptic Peptide
QS[+162]T[+365]AT[+365]VS[+365]VTVIDVNENPYFAPNPK.I z=4,scan#=42329,scan time=101.1246
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
Intensity
m/z
Loss of Hexose
K.GIQHPPQSTATVSVTVIDVNENPYFAPNPK.I z=3,scan#=49531,scan time=115.3131
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
QHPPQS[+365]T[+365]AT[+365]VSVTVIDVNENPYFAPNPK.I z=4,scan#=44197,scan time=104.8000
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
Intensity
m/z

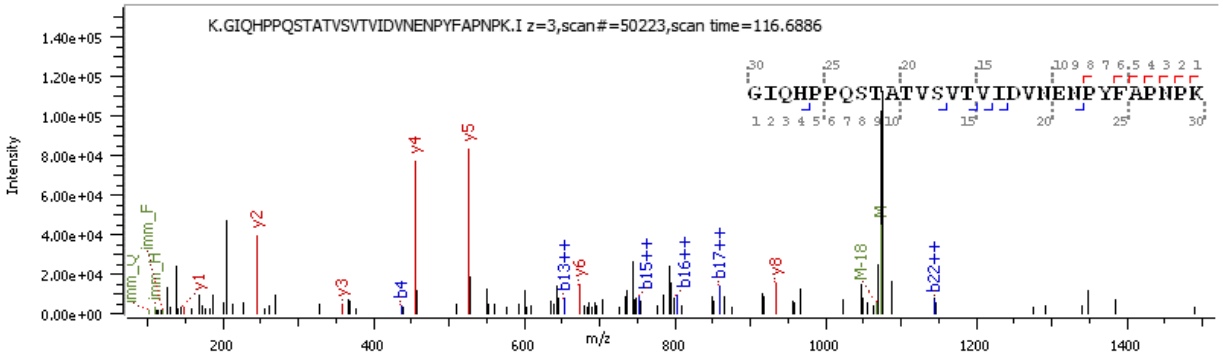
Figure B-2C

T481

O-mannosylated Tryptic Peptide



Loss of Hexose



Addition of GlcNAc to Mannose by POMGnT1

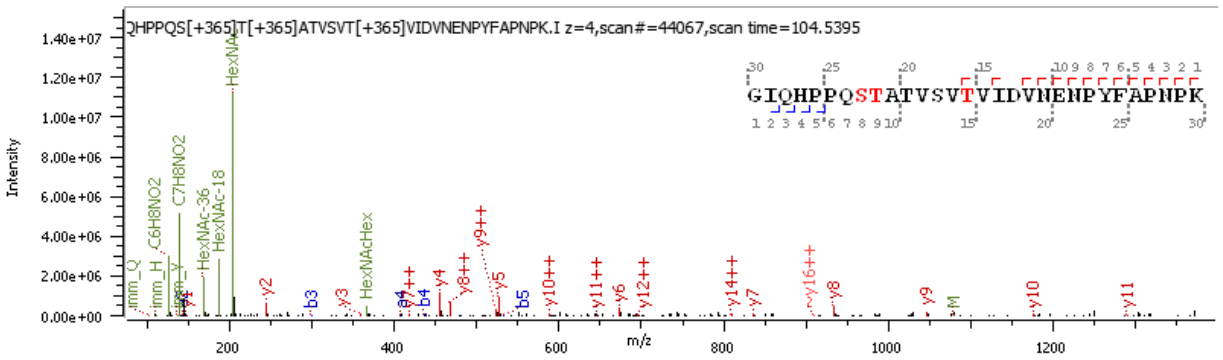


Figure B-2D

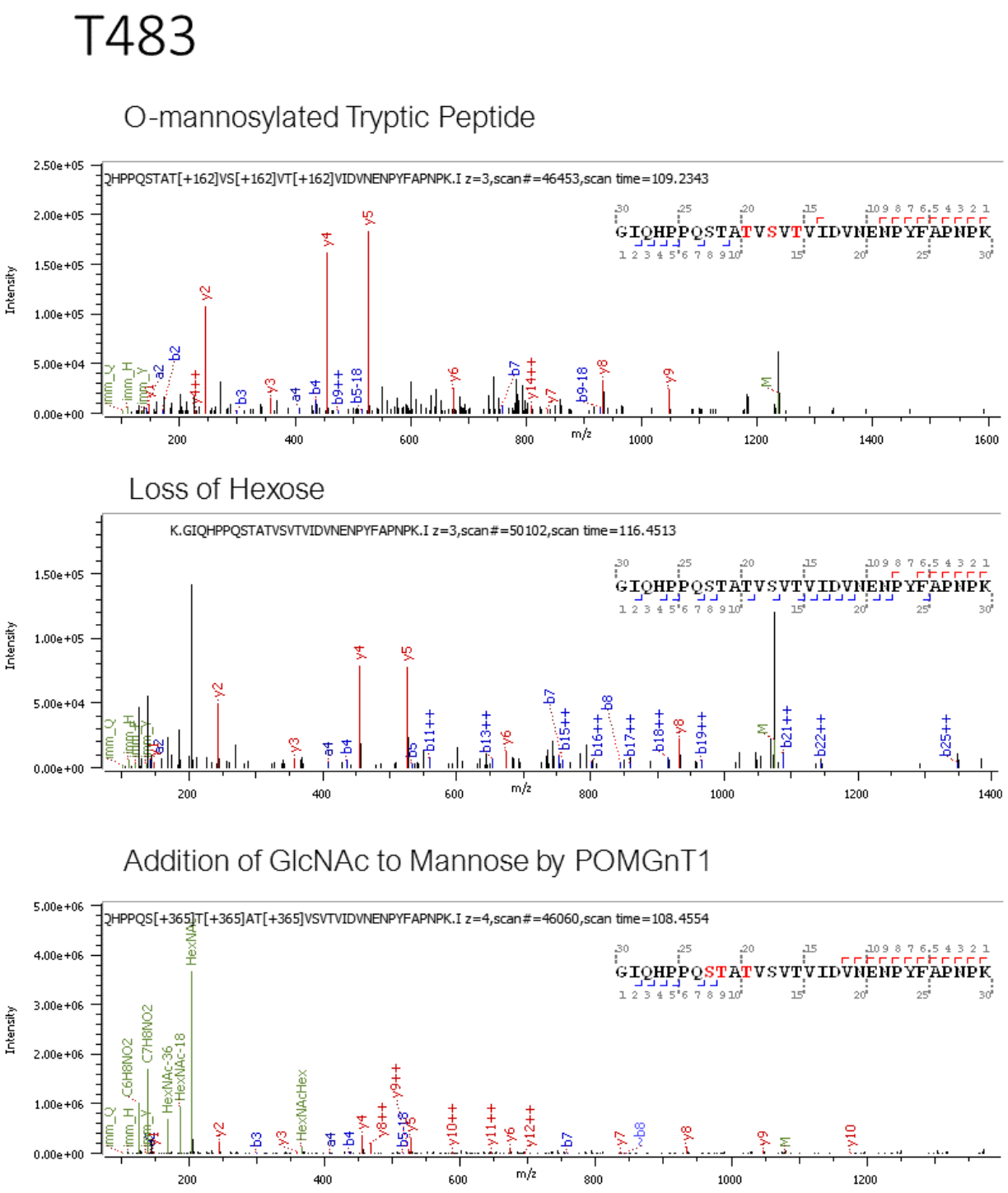

T483
O-mannosylated Tryptic Peptide
QHPPQSTAT[+162]VS[+162]VT[+162]VIDVNENPYFAPNPK.I z=3,scan#=46453,scan time=109.2343
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
Intensity
m/z
Loss of Hexose
K.GIQHPPQSTATVSVTVIDVNENPYFAPNPK.I z=3,scan#=50102,scan time=116.4513
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
QHPPQS[+365]T[+365]AT[+365]VSVTVIDVNENPYFAPNPK.I z=4,scan#=46060,scan time=108.4554
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
HexNAc
C6H8NO2
C7H8NO2
HexNAc-36
HexNAc-18
HexNAcHex
Intensity
m/z

Figure B-2E

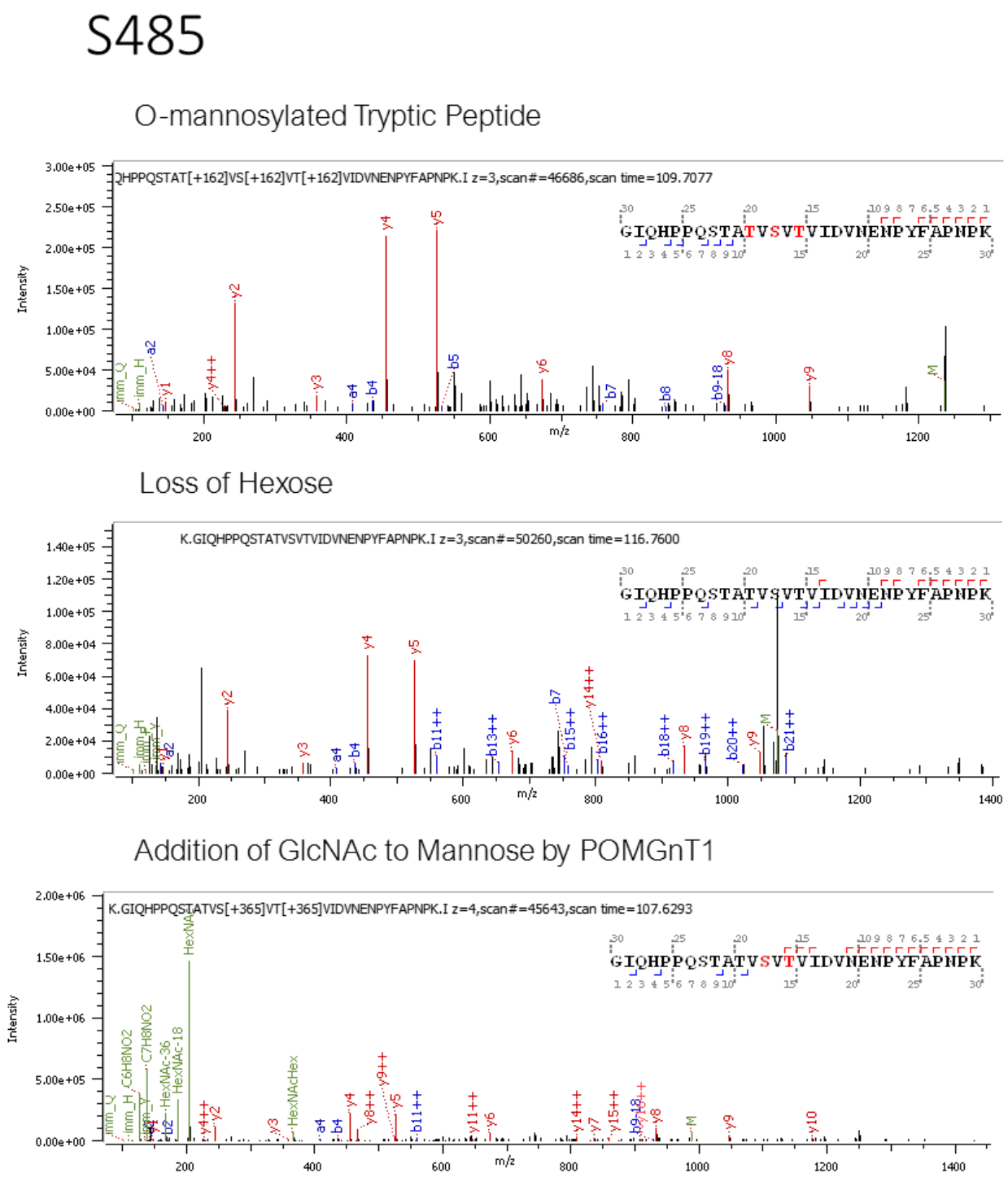

S485
O-mannosylated Tryptic Peptide
QHPPQSTAT[+162]VS[+162]VT[+162]VIDVNENPYFAPNPK.I z=3,scan#=46686,scan time=109.7077
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
Intensity
m/z
Loss of Hexose
K.GIQHPPQSTATVSVTVIDVNENPYFAPNPK.I z=3,scan#=50260,scan time=116.7600
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
K.GIQHPPQSTATVS[+365]VT[+365]VIDVNENPYFAPNPK.I z=4,scan#=45643,scan time=107.6293
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
HexNAc
C6H8NO2
C7H8NO2
HexNAc-36
HexNAc-18
HexNAcHex
Intensity
m/z

Figure B-2F

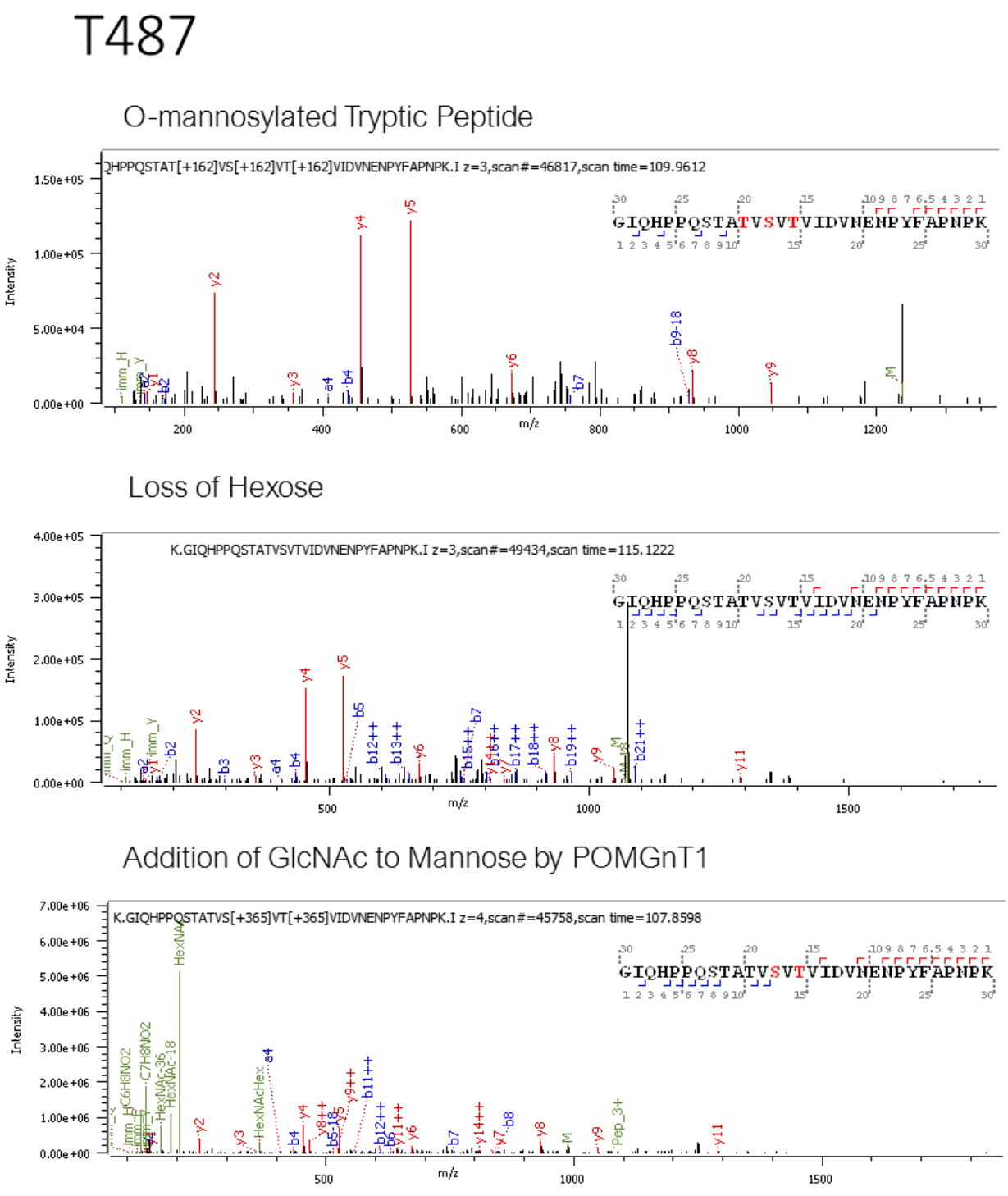

T487
O-mannosylated Tryptic Peptide
QHPPQSTAT[+162]VS[+162]VT[+162]VIDVNENPYFAPNPK.I z=3,scan#=46817,scan time=109.9612
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
Intensity
m/z
Loss of Hexose
K.GIQHPPQSTATVSVTVIDVNENPYFAPNPK.I z=3,scan#=49434,scan time=115.1222
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
K.GIQHPPQSTATVS[+365]VT[+365]VIDVNENPYFAPNPK.I z=4,scan#=45758,scan time=107.8598
GIQHPPQSTATVSVTVIDVNENPYFAPNPK
Intensity
m/z

Figure B-2G

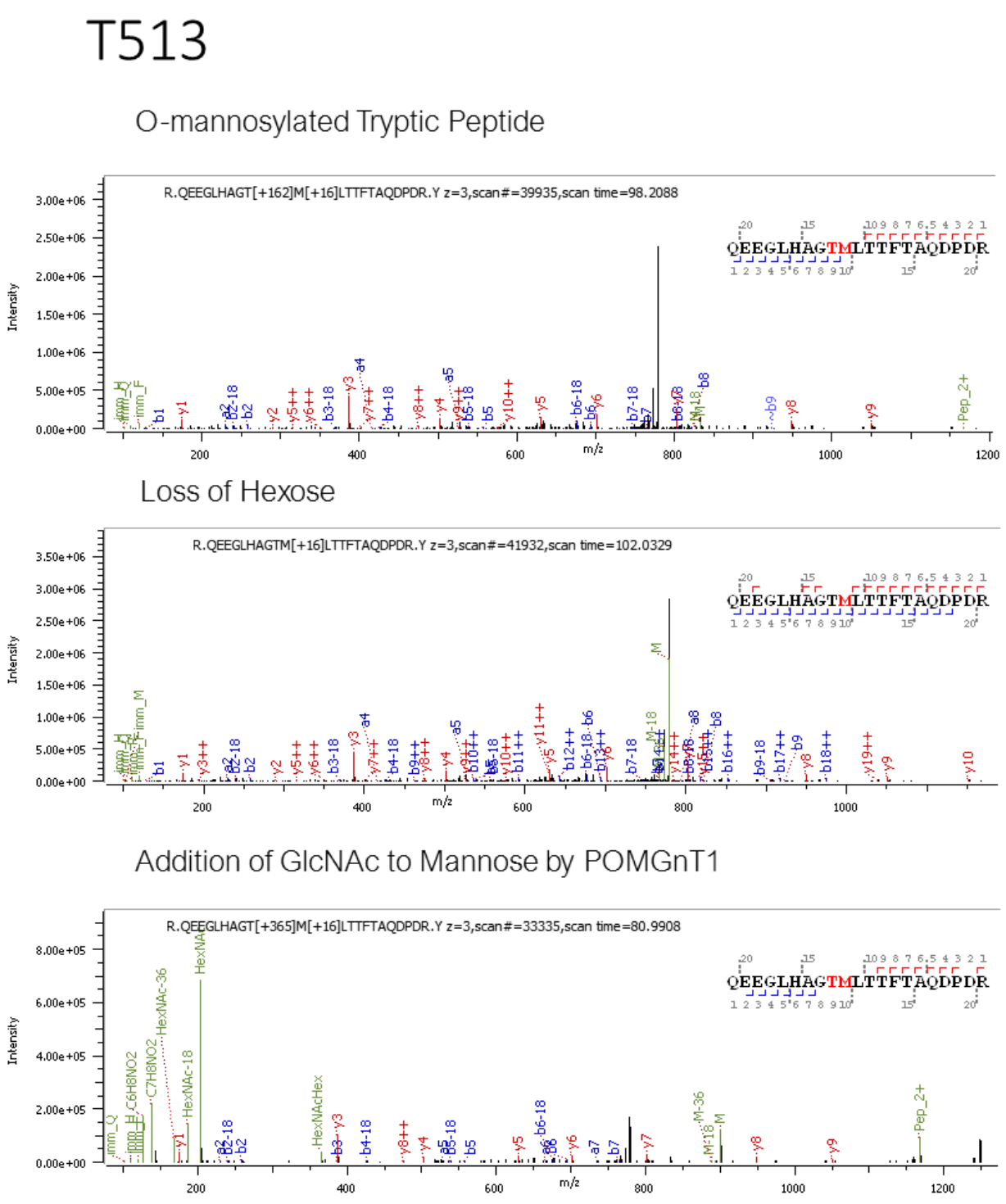

T513
O-mannosylated Tryptic Peptide
R.QEEGLHAGT[+162]M[+16]LTTFTAQDPDR.Y z=3,scan#=39935,scan time=98.2088
QEEGLHAGTMLTTFTAQDPDR
Intensity
m/z
Loss of Hexose
R.QEEGLHAGTM[+16]LTTFTAQDPDR.Y z=3,scan#=41932,scan time=102.0329
QEEGLHAGTMLTTFTAQDPDR
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
R.QEEGLHAGT[+365]M[+16]LTTFTAQDPDR.Y z=3,scan#=33335,scan time=80.9908
QEEGLHAGTMLTTFTAQDPDR
Intensity
m/z

Figure B-2H

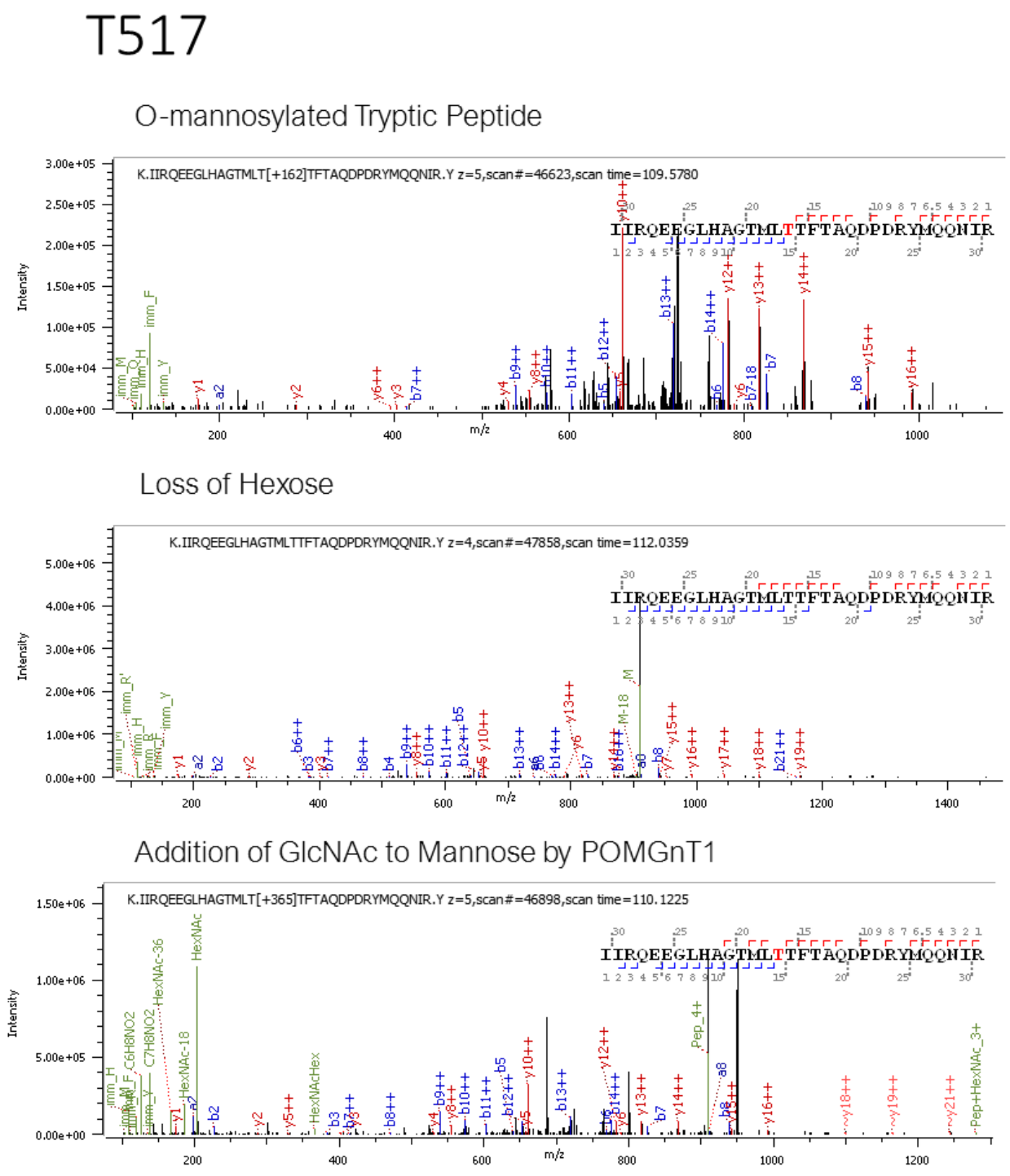

T517
O-mannosylated Tryptic Peptide
K.IIRQEEGLHAGTMLT[+162]TFTAQDPDRYMQQNIR.Y z=5,scan#=46623,scan time=109.5780
IIRQEEGLHAGTMLTTFTAQDPDRYMQQNIR
Intensity
m/z
Loss of Hexose
K.IIRQEEGLHAGTMLTTFTAQDPDRYMQQNIR.Y z=4,scan#=47858,scan time=112.0359
IIRQEEGLHAGTMLTTFTAQDPDRYMQQNIR
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
K.IIRQEEGLHAGTMLT[+365]TFTAQDPDRYMQQNIR.Y z=5,scan#=46898,scan time=110.1225
IIRQEEGLHAGTMLTTFTAQDPDRYMQQNIR
Intensity
m/z

Figure B-2I

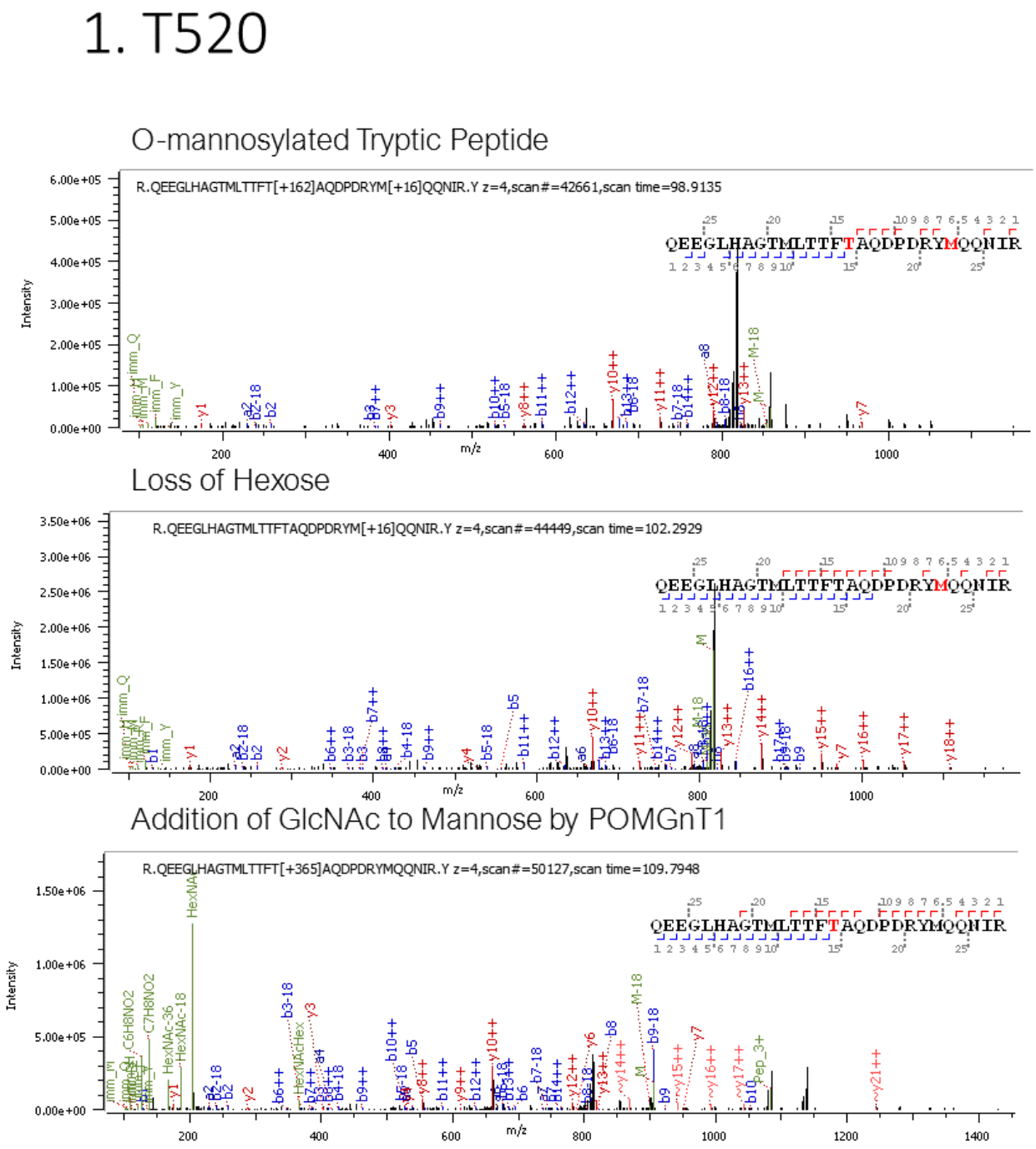

1. T520
O-mannosylated Tryptic Peptide
R.QEEGLHAGTMLTTFT[+162]AQDPDRYM[+16]QQNIR.Y z=4,scan#=42661,scan time=98.9135
QEEGLHAGTMLTTFTAQDPDRYMQQNIR
Intensity
m/z
Loss of Hexose
R.QEEGLHAGTMLTTFTAQDPDRYM[+16]QQNIR.Y z=4,scan#=44449,scan time=102.2929
QEEGLHAGTMLTTFTAQDPDRYMQQNIR
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
R.QEEGLHAGTMLTTFT[+365]AQDPDRYMQQNIR.Y z=4,scan#=50127,scan time=109.7948
QEEGLHAGTMLTTFTAQDPDRYMQQNIR
Intensity
m/z

Figure B-2. Tandem MS/MS spectra of glycopeptides from N-cadherin. The top panel shows hexose (+162) on peptides from HEK Wt-expressed N-cadherin. The middle panel shows the loss of O-mannose (lack of +162) after α-mannosidase treatment on the HEK Wt-expressed N-cadherin peptides. The bottom panel shows the addition of GlcNAc to O-mannose (+365) after treatment of peptides from HEK-POMGnT1 expressed N-cadherin with POMGnT1. (A) T290. (B) S480. (C) T481. (D) T483. (E) S485. (F) T487. (G) T513. (H) T517. (I) T520.

Figure B-3A

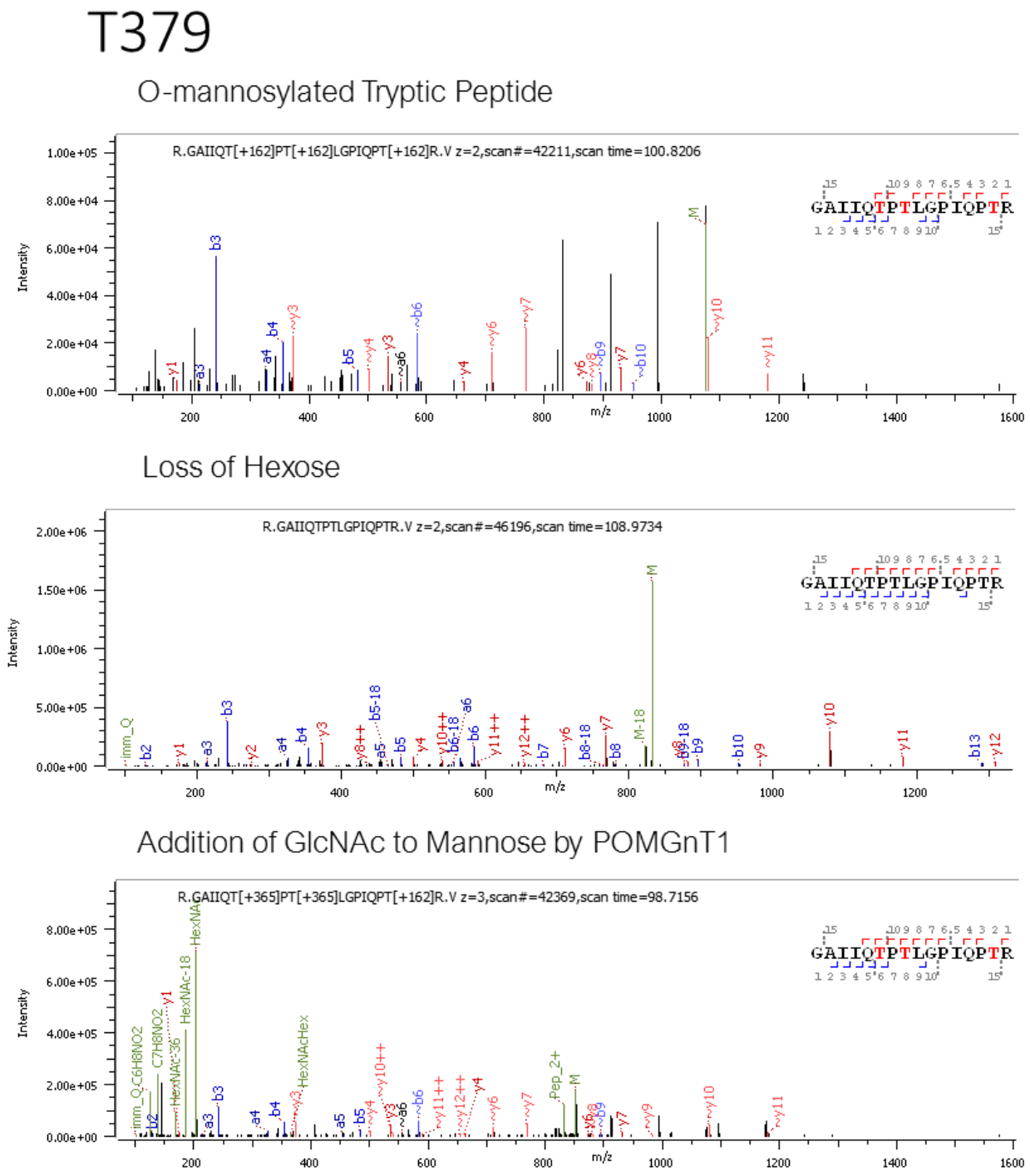
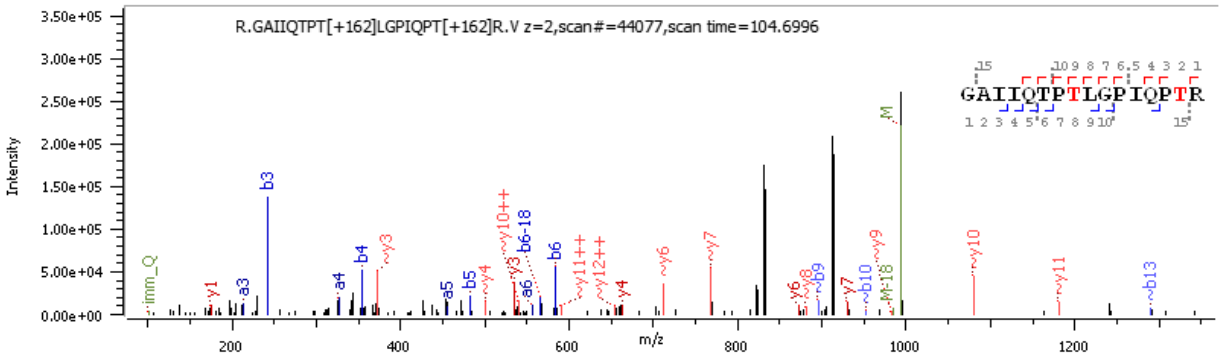

T379
O-mannosylated Tryptic Peptide
R.GAIIQT[+162]PT[+162]LGPIQPT[+162]R.V z=2,scan#=42211,scan time=100.8206
GAIIQTPTLGPIQPTR
Intensity
m/z
Loss of Hexose
R.GAIIQTPTLGPIQPTR.V z=2,scan#=46196,scan time=108.9734
GAIIQTPTLGPIQPTR
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
R.GAIIQT[+365]PT[+365]LGPIQPT[+162]R.V z=3,scan#=42369,scan time=98.7156
GAIIQTPTLGPIQPTR
Intensity
m/z

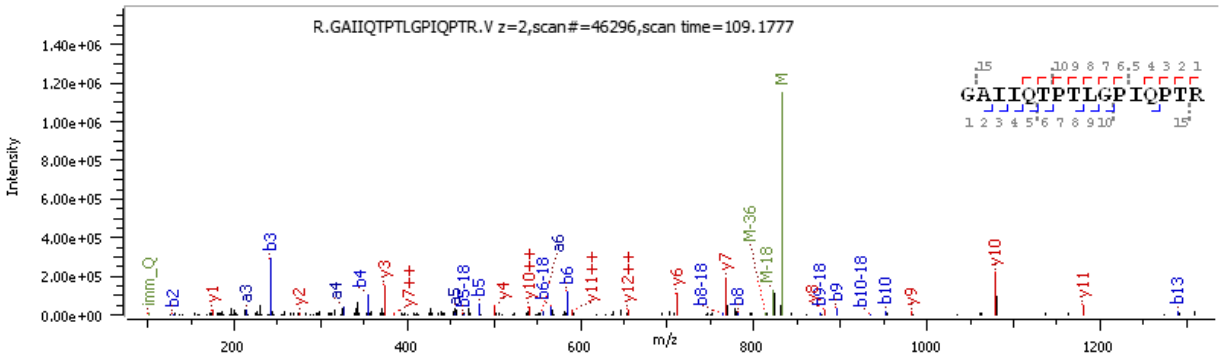
Figure B-3B

T381

O-mannosylated Tryptic Peptide



Loss of Hexose



Addition of GlcNAc to Mannose by POMGnT1

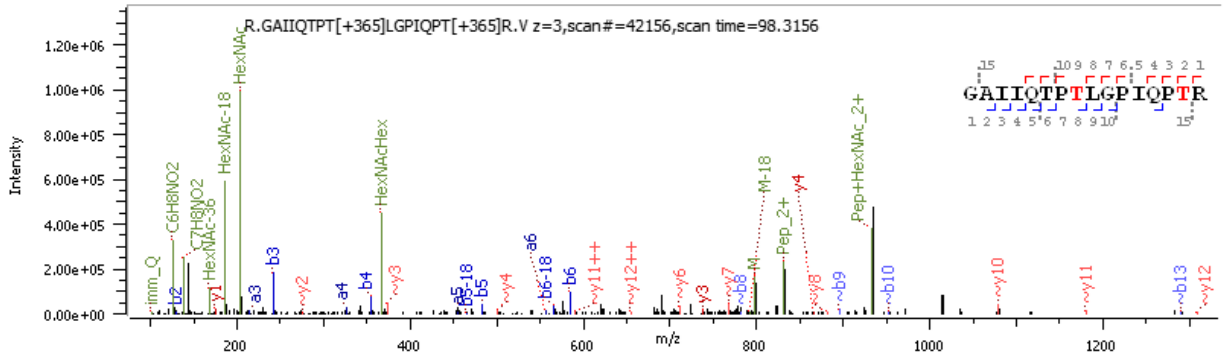


Figure B-3C

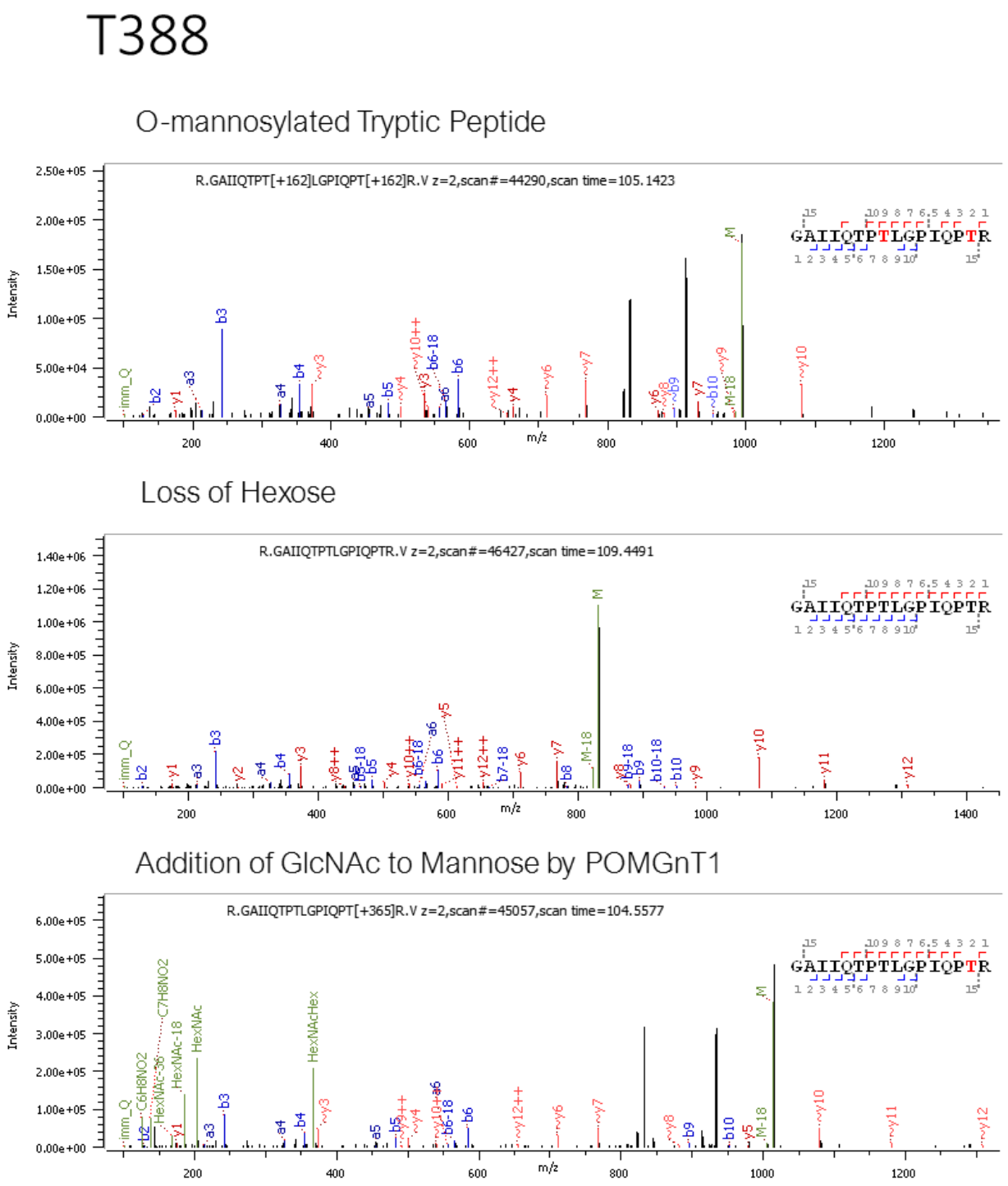
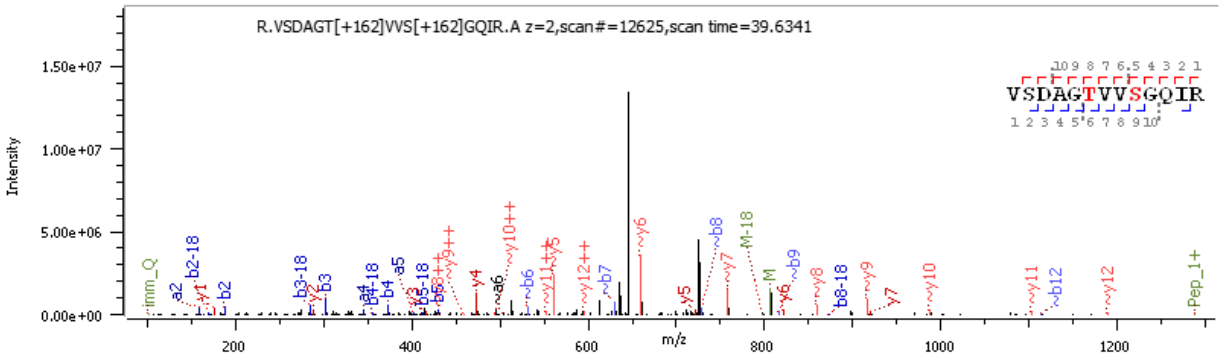

T388
O-mannosylated Tryptic Peptide
R.GAIIQTPT[+162]LGPIQPT[+162]R.V z=2,scan#=44290,scan time=105.1423
GAIIQTPTLGPIQPTR
Intensity
m/z
Loss of Hexose
R.GAIIQTPTLGPIQPTR.V z=2,scan#=46427,scan time=109.4491
GAIIQTPTLGPIQPTR
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
R.GAIIQTPTLGPIQPT[+365]R.V z=2,scan#=45057,scan time=104.5577
GAIIQTPTLGPIQPTR
Intensity
m/z

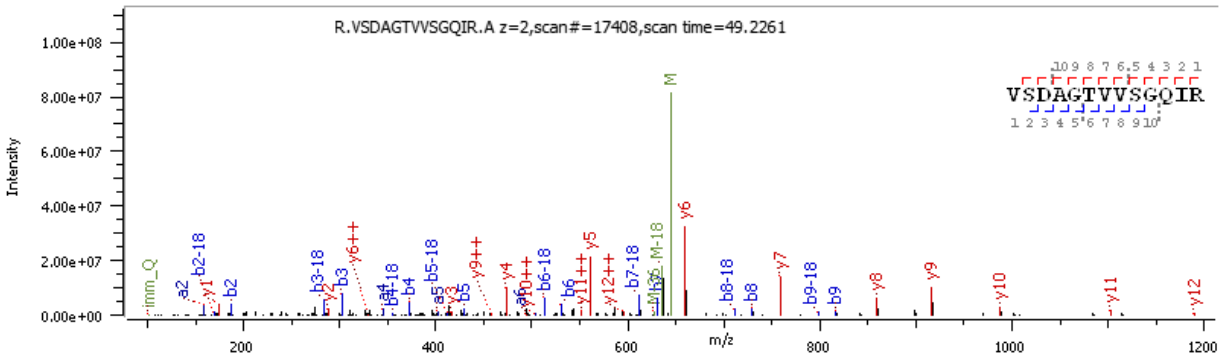
Figure B-3D

T395

O-mannosylated Tryptic Peptide



Loss of Hexose



Addition of GlcNAc to Mannose by POMGnT1

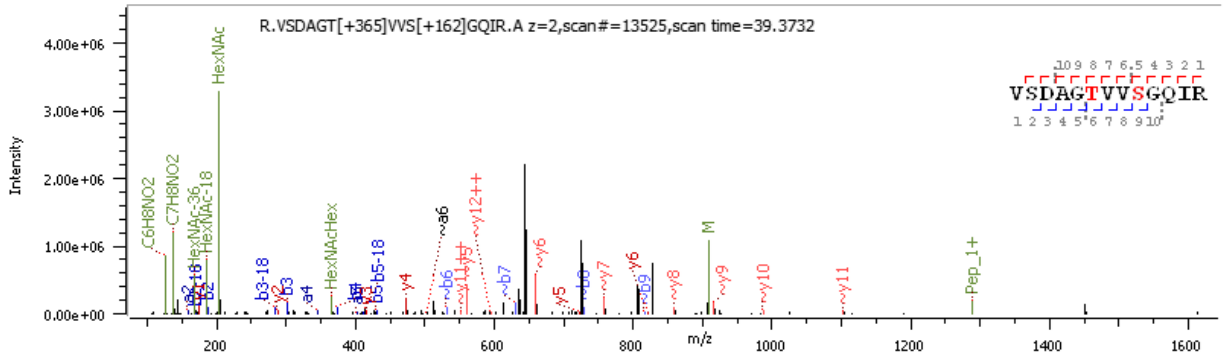


Figure B-3E

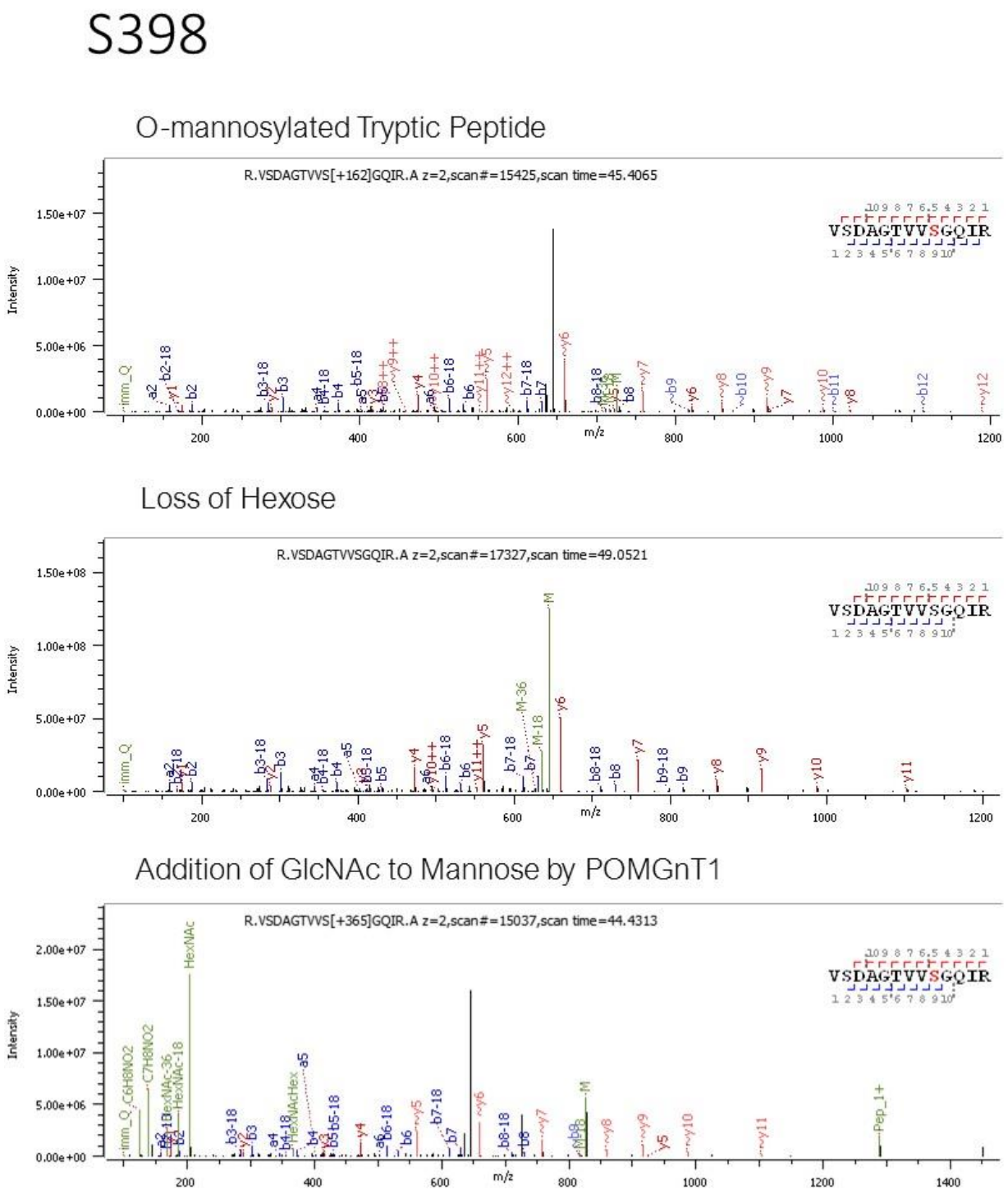
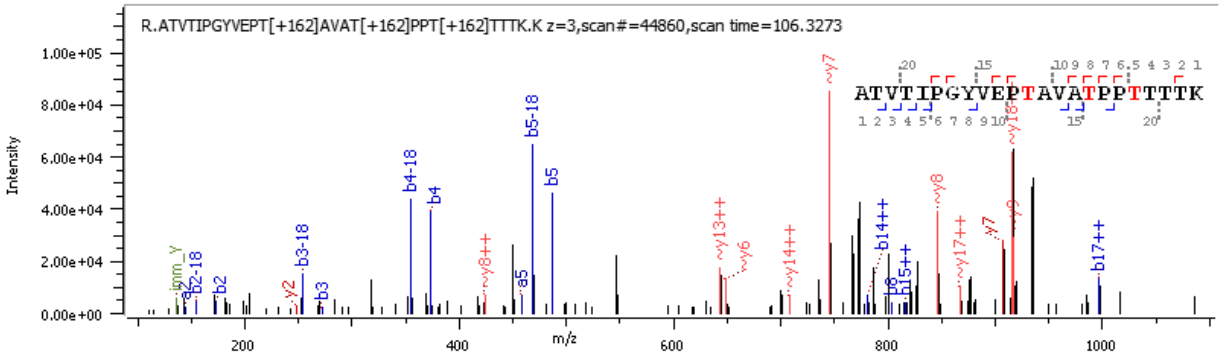

S398
O-mannosylated Tryptic Peptide
R.VSDAGTVVS[+162]GQIR.A z=2,scan#=15425,scan time=45.4065
VSDAGTVVSGQIR
Intensity
m/z
Loss of Hexose
R.VSDAGTVVSGQIR.A z=2,scan#=17327,scan time=49.0521
VSDAGTVVSGQIR
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
R.VSDAGTVVS[+365]GQIR.A z=2,scan#=15037,scan time=44.4313
VSDAGTVVSGQIR
HexNAc
Pep_1+
Intensity
m/z

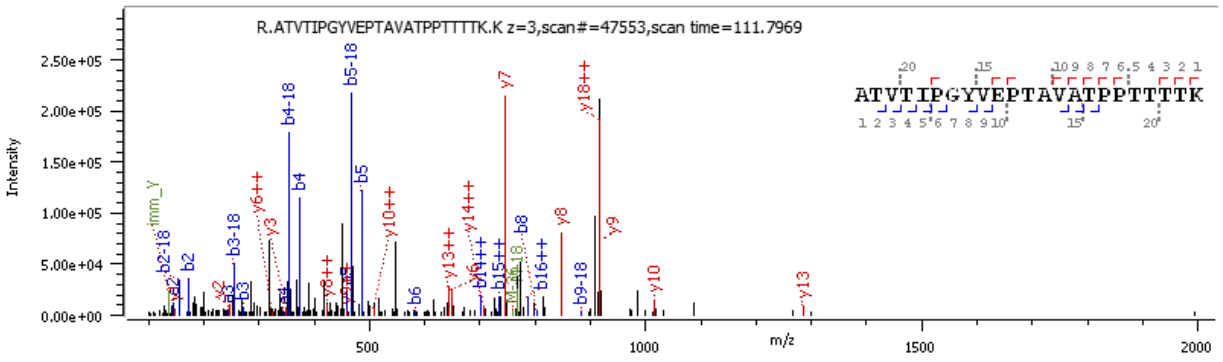
Figure B-3F

T414

O-mannosylated Tryptic Peptide



Loss of Hexose



Addition of GlcNAc to Mannose by POMGnT1

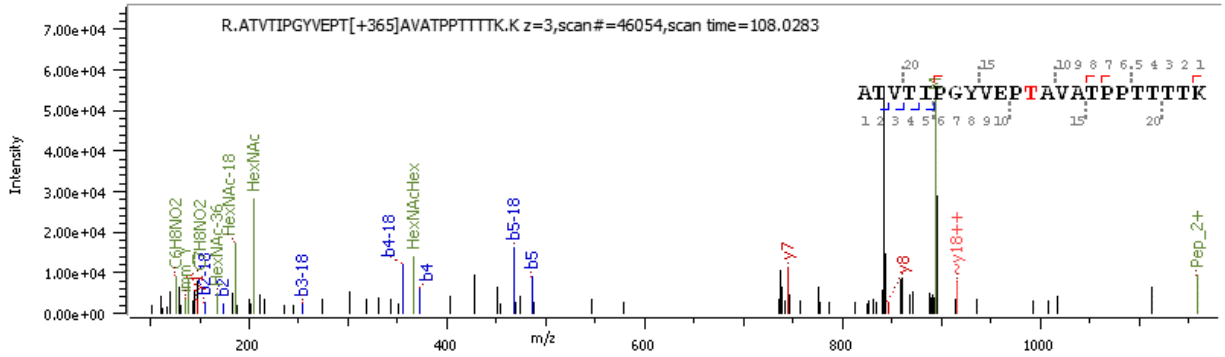


Figure B-3G

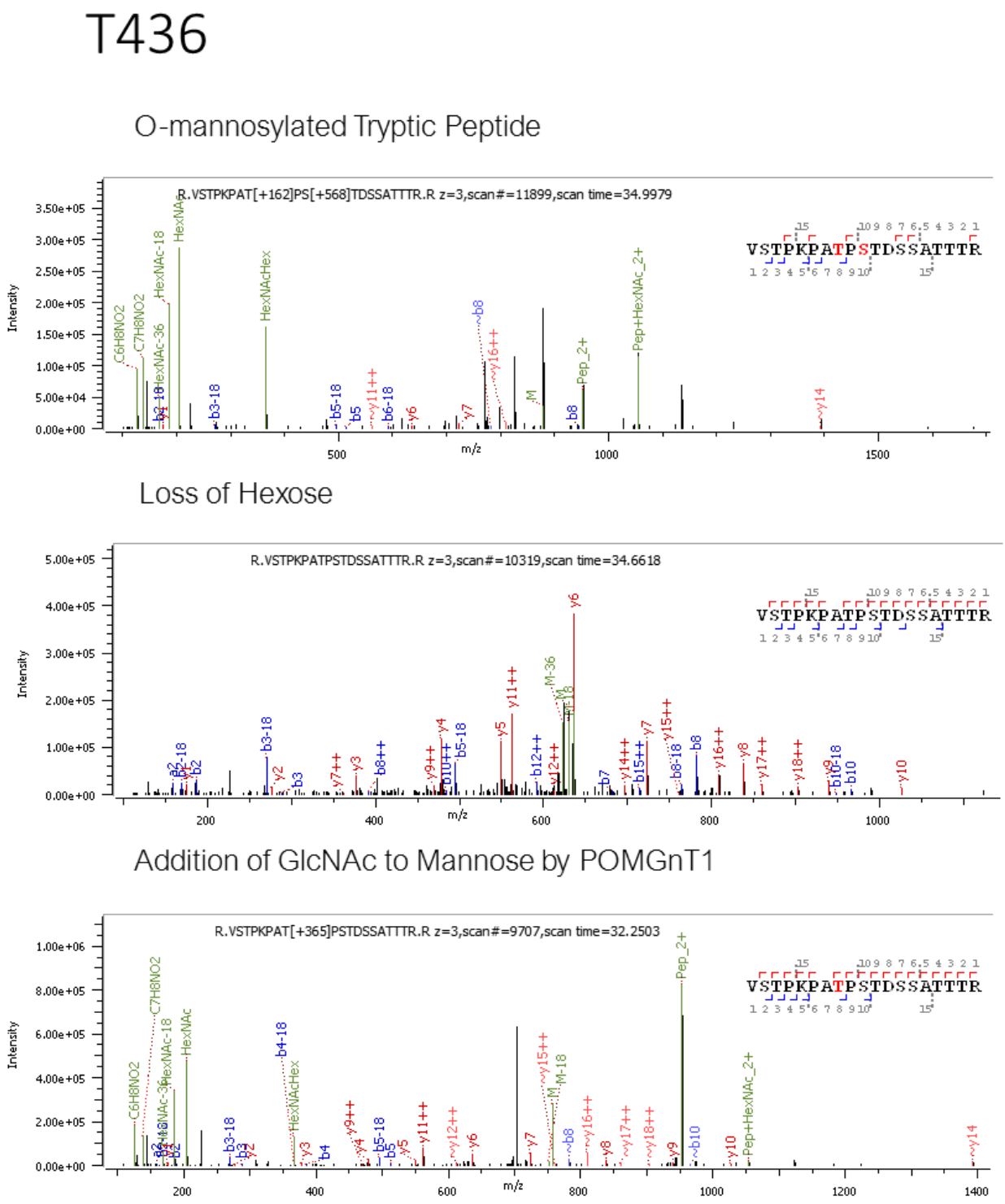
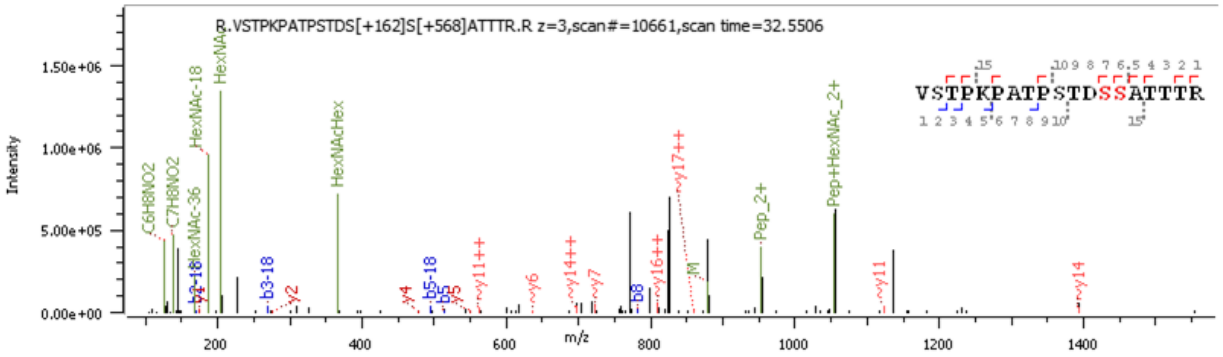

T436
O-mannosylated Tryptic Peptide
R.VSTPKPAT[+162]PS[+568]TDSSATTTR.R z=3,scan#=11899,scan time=34.9979
VSTPKPATPSTDSSATTTR
Intensity
m/z
Loss of Hexose
R.VSTPKPATPSTDSSATTTR.R z=3,scan#=10319,scan time=34.6618
VSTPKPATPSTDSSATTTR
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
R.VSTPKPAT[+365]PSTDSSATTTR.R z=3,scan#=9707,scan time=32.2503
VSTPKPATPSTDSSATTTR
Intensity
m/z

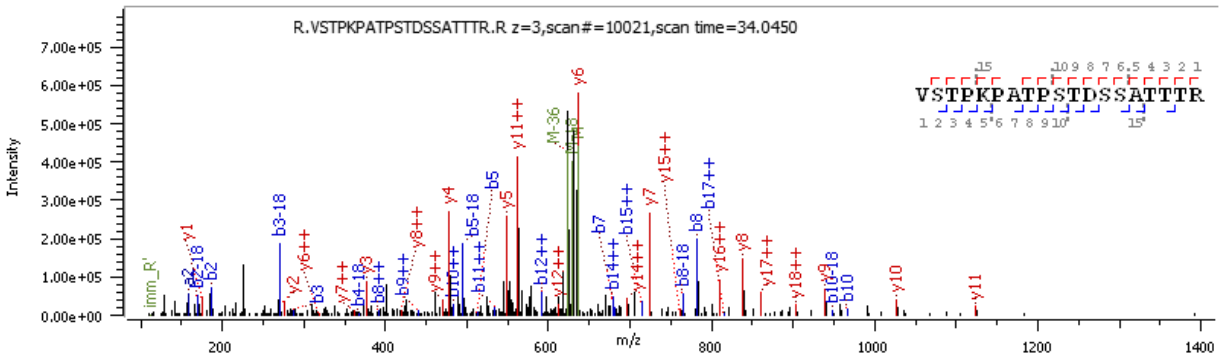
Figure B-3H

S441

O-mannosylated Tryptic Peptide



Loss of Hexose



Addition of GlcNAc to Mannose by POMGnT1

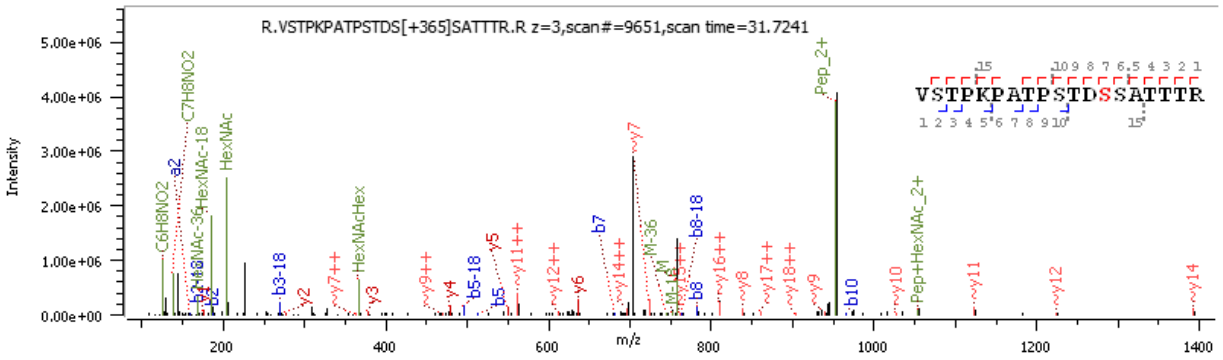
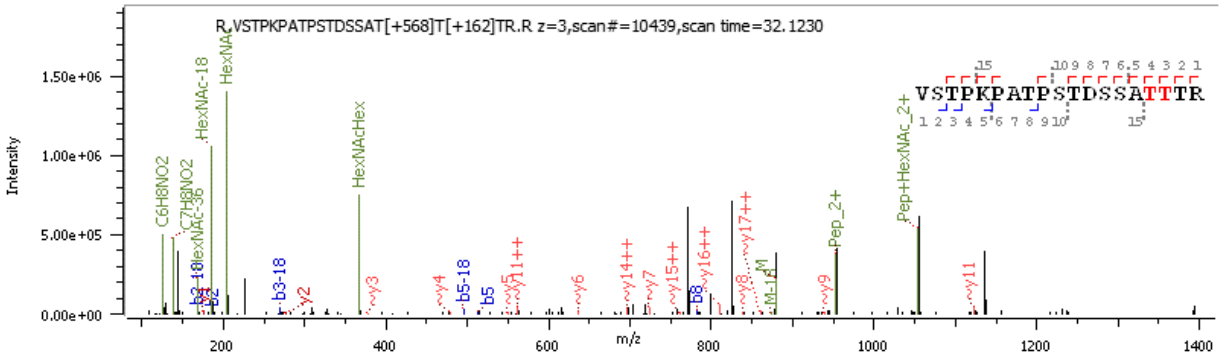


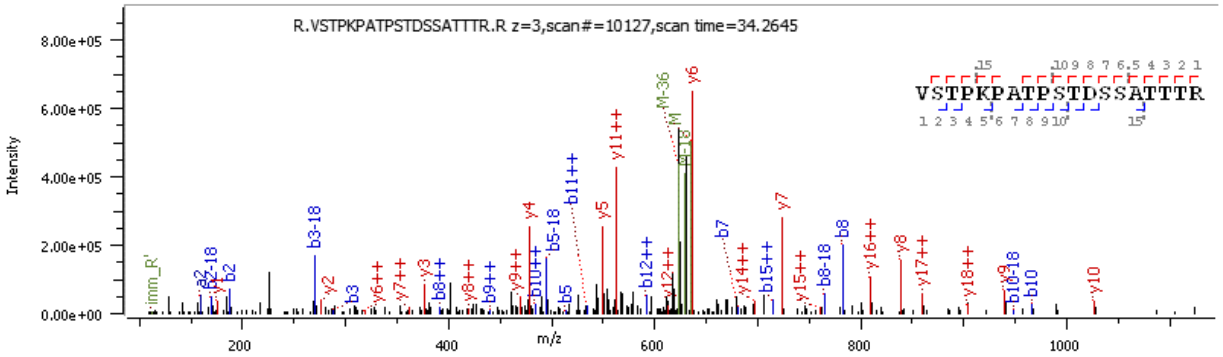
Figure B-3I

T445

O-mannosylated Tryptic Peptide



Loss of Hexose



Addition of GlcNAc to Mannose by POMGnT1

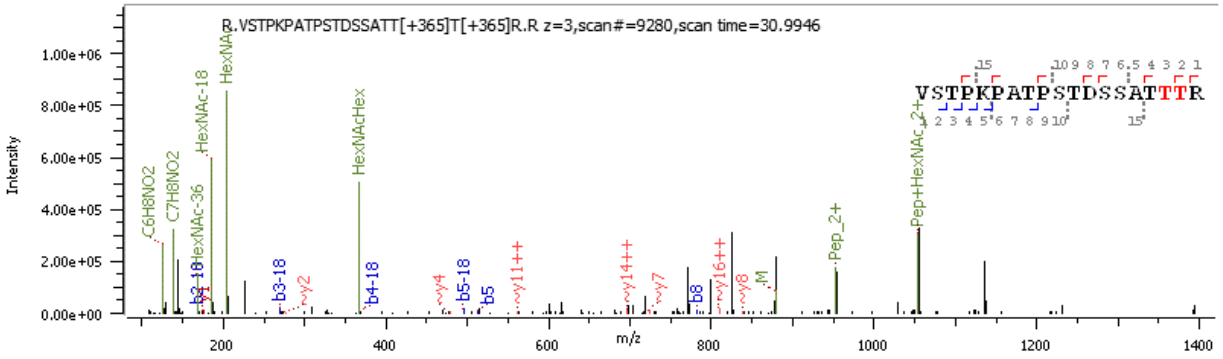


Figure B-3J

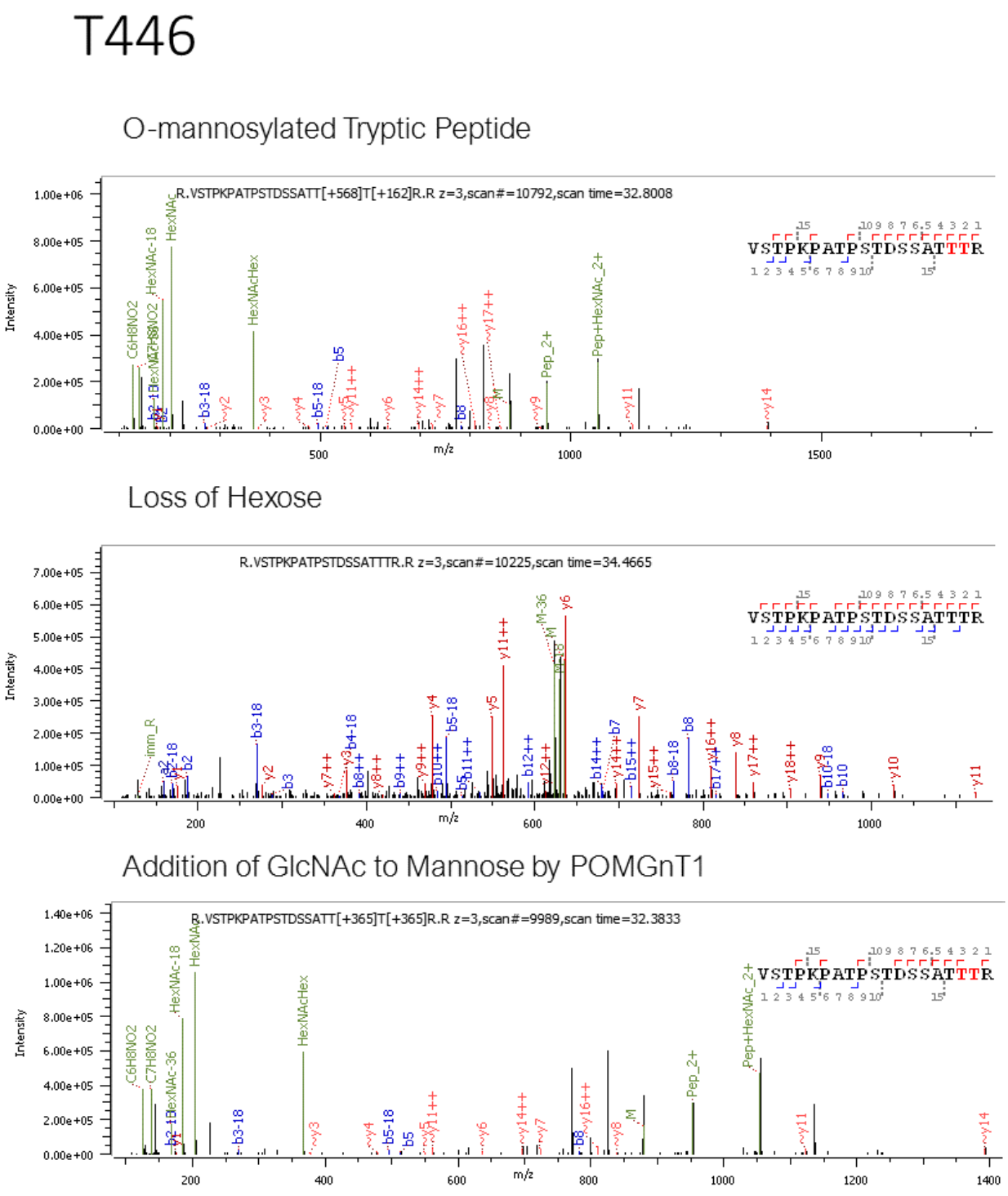

T446
O-mannosylated Tryptic Peptide
R.VSTPKPATPSTDSSATT[+568]T[+162]R.R z=3,scan#=10792,scan time=32.8008
VSTPKPATPSTDSSATTTR
Intensity
m/z
Loss of Hexose
R.VSTPKPATPSTDSSATTTR.R z=3,scan#=10225,scan time=34.4665
VSTPKPATPSTDSSATTTR
Intensity
m/z
Addition of GlcNAc to Mannose by POMGnT1
R.VSTPKPATPSTDSSATT[+365]T[+365]R.R z=3,scan#=9989,scan time=32.3833
VSTPKPATPSTDSSATTTR
Intensity
m/z

Figure B-3. Tandem MS/MS spectra of glycopeptides from alpha-dystroglycan. The top panel shows hexose (+162) on peptides from HEK-POMGnT1 expressed alpha-dystroglycan. The middle panel shows the loss of O-mannose (lack of +162) after α-mannosidase treatment on the HEK-POMGnT1 expressed alpha-dystroglycan peptides. The bottom panel shows the addition of GlcNAc to O-mannose (+365) after treatment of peptides from HEK-POMGnT1 expressed alpha-dystroglycan with POMGnT1. (A) T379. (B) T381. (C) T388. (D) T395. (E) S398. (F) T414. (G) T436. (H) T441. (I) S445. (J) T446