INSIGHTS INTO O-MANNOSYLATION REVEALED BY GLYCOPROTEOMICS

by

STEPHANIE LAUREN HAMMOND

(Under the Direction of Lance Wells)

ABSTRACT

O-mannosylation accounts for approximately 30% of the reported *O*-linked glycan structures detected from mouse brain. However, *O*-mannosylation has rarely been observed on other mammalian proteins beside alpha-dystroglycan (α-DG). Previous studies have shown α-DG to be modified by both *O*-GalNAc and rare *O*-mannose initiated glycan structures but no sites of attachment have been mapped. Mutations in genes encoding known and putative glycosyltransferases involved in the *O*-mannosylation of α-DG have been implicated in several forms of Congenital Muscular Dystrophy (CMD). Thus this project sought to provide insights into global *O*-mannosylation and α-DG modification. Using tandem mass spectrometry approaches we were able to identify 21 sites of modification on 91 glycopeptides derived from α-DG. Additionally, through our glycomic analysis we were able to characterize 34 *O*-glycan structures, 9 *O*-mannose and 25 *O*-GalNAc initiated structures, from mouse models of CMD. This analysis of mouse brains allowed us to determine that POMGnT1 is essential for the extension of classical and branched *O*-mannose initiated glycan structures, that loss of LARGE has no impact on these observed structures or on *O*-GalNAc structures, and that additional *O*-mannose modified proteins besides α-DG exist in mouse brains. All the work outlined in this dissertation is based on the development and application of glycomic and glycoproteomic workflows.

INDEX WORDS: post-translational modifications, *O*-mannosylation, α-DG, glycomics, pseudo neutral loss MS^3, congenital muscular dystrophy

INSIGHTS INTO O-MANNOSYLATION REVEALED BY GLYCOPROTEOMICS

by

Stephanie Lauren Hammond

B.S. Lander University, 2005

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

2010

INSIGHTS INTO O-MANNOSYLATION REVEALED BY GLYCOPROTEOMICS

by

STEPHANIE LAUREN HAMMOND

Major Professor:	Lance Wells
Committee:	I. Jonathan Amster Ron Orlando

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2010

DEDICATION

To everyone who has instilled in me that I can accomplish anything I set my mind to and encouraged me to reach for the stars, thank you.

I love you.

ACKNOWLEDGEMENTS

I would like to express the most sincere gratitude and appreciation for the opportunities and knowledge that I have received here at the Complex Carbohydrate Research Center at the University of Georgia. I would like to especially express my gratitude to Lance Wells for his assistance and guidance throughout my studies, and providing me with the necessary resources. I would also like to thank all of my lab members past and present, Enas Gad El Karim, Jae-Min Lim, Chin Fen Teo, Olga Stulick, Peng (Linda) Zhao, Meng Fang, Edith Hayden, Sandii Brimble, Krithika Vaidyanathan, Sean Durning, Seongha Park, and Ryan Stuart, for your helpful advice and many discussions.

To my parents, thank you for EVERYTHING. Thank you for providing me with all I ever wanted and being my biggest supporters. I cannot express how much I appreciate the sacrifices you made for me over the years. I am blessed to be your daughter.

To my husband, thank you for encouragement and unconditional love.

To my friends, old and new, thank you for keeping me grounded and reminding me that the stressful times will pass.

Thank you everyone.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Post-translational modifications-

Post-translational modifications (PTMs) involve processing of a protein that changes its properties either by proteolytic cleavage or by the addition of a modifying group to one or more amino acid residues within the protein. Glycosylation is considered to be one of the most common post-translational modifications, and it is postulated that glycoproteins make up more than half of all of the proteins found with in the eukaryotic cell[1]. To date, 13 different monosaccharides and 8 different amino acids have been reported to be involved in numerous protein-carbohydrate linkages[2].

There are two main types of glycosylation, which are defined based upon the nature of the glycosidic bond attaching the monosaccharide(s) to the protein and are commonly referred to as N-linked and O-linked glycosylation. In the case of N-linked oligosaccharides, they are usually attached via a *N*-acetylglucosamine (GlcNAc) to an asparagine following the consensus sequence N-X-S/T (X cannot be proline). *O*-linked glycans however are attached to the hydroxy group of either serine (Ser) or threonine (Thr) through a variety of reducing terminal sugar residues such as *N*-acetylgalactosamine (GalNAc), fucose (Fuc), glucose, GlcNAc, xylose, galactose (Gal), and mannose (Man)[2].

O-mannosylation-

Identified in the late 1960's, *O*-mannosyl glycans were first observed linked to the Ser and Thr residues in baker's yeast[3]. In yeast and fungi, *O*-mannose that is attached to the protein backbone is typically extended by α-linked mannoses yielding a linear glycan structure[4]. Although *O*-mannosylation is prevalent in fungi, this modification is uncommon to mammalian proteins. To date, *O*-mannosyl glycans have only been identified from a limited number of sources, such as chondroitin sulfate proteoglycans from brain tissue[4], α-dystroglycan (α-DG) isolated from muscle, brain, and neuronal

tissues[5-7], and more recently was identified on a receptor protein tyrosine phosphatase β[8]. The *O*-mannosyl glycan structures found on α-DG are vastly different from those observed on fungal proteins. The most well characterized *O*-mannosyl glycan found on α-DG is a tetrasaccharide with the structure NeuAcα2-3Galβ1-4GlcNAcβ1-2Manα1Ser/Thr[5].

Dystroglycan-

Dystroglycan (DG) was first identified in 1987 as a novel laminin-binding protein (then named cranin), and was later cloned and characterized by Kevin Campbell's laboratory in the early 90's[9, 10]. Originally isolated from skeletal muscle, DG is a central component of the dystrophin glycoprotein complex (DGC)(Figure 1-1)[10, 11]. Although the exact function of the DGC remains undetermined, studies have shown the DGC to aid in providing structural stability to the sarcolemma during muscle contraction[12]. Within the DGC, DG has been shown to interact with proteins of the extracellular matrix (ECM), serving to connect the cytoskeleton of the muscle cell to the basement membrane[10, 13].

Encoded by a single gene, *DAG1,* DG is proteolytically cleaved into two non-convalently associated proteins α-dystroglycan (α-DG) and β-dystroglycan (β-DG)[10]. α-DG is an extracellular protein having a dumbbell shape which consists of two globular N- and C-terminal domains that contain 3 potential *N*-glycosylation sites separated by a extensively (heterogeneously) *O*-glycosylated mucin domain (Figure 1-2)[14, 15]. The *O*-linked glycan structures found on α-DG are initiated by both *O*-GalNAc and *O*-mannose. The *O*-GalNAc linked glycans are of the core 1 type, with structures Galβ1-3GalNAc, while the *O*-mannose linked glycans have a reported structure of NeuAcα2-3Galβ1-4GlcNAcβ1-2Man[5-7]. Preliminary studies focusing on the glycosylation of α-DG, have shown it to be extensively glycosylated based upon separation by SDS-PAGE[10]. While the predicted molecular weight of α-DG is ~70kDa, when detected by immunoblot α-DG has a molecular mass between 120 and156kDa, depending upon tissue source[13]. Following removal of *N*-linked glycans, a molecular weight shift of ~4kDa was observed, indicating that the variability of molecular weight arises from the extensive amount of *O*-glycosylation within the mucin domain[11].

Complete understanding of the enzymatic pathways involved in the post-translational processing of α-DG yielding fully, functionally active α-DG remains unclear. Following enzymatic removal of *N*-glycans, the ligand binding activity of α-DG has been shown to be undisturbed. However, upon full chemical deglycosylation of α-DG a complete loss of ligand binding activity was observed[16]. Thus research has shifted towards investigation of *O*-linked glycans on α-DG, of specific interest are the *O*-mannosyl glycan structures. Previous studies suggested that an *O*-mannosyl tetrasaccharide was required for the interaction between α-DG and extracellular ligands such as laminin. However, more recent studies show that upon degradation of this glycan structure with a combination of glycosidases that laminin binding is enhanced, indicating that there are other glycan structures present on α-DG that are required for laminin binding[17]. Supporting this finding, Yoshida-Moriguchi *et. al* recently identified a novel phosphorylated *O*-mannosyl glycan structure on α-DG which has been shown to be required for laminin binding activity[18].

Mutations in genes encoding (putative) glycosyltransferases involved in the post-translational processing of α-DG have been shown to have an effect on *O*-glycosylation pathways. Hypoglycosylation of α-DG results not only in a loss of laminin binding, but has also been associated with cancer metastasis and various forms of congenital muscular dystrophy (CMD)[19]. Mutations in the genes encoding protein *O*-mannosyltransferase 1 and 2 (POMT1 and POMT2), which catalyze the addition of *O*-mannose onto the protein backbone of α-DG have been shown to result in the severest form of CMD, Walker-Warburg Syndrome (WWS)[20, 21]. Muscle-eye-brain disease (MEB), results from a mutation in the gene protein *O*-mannosyl β-1,2-*N*-acetylglucosaminyltransferase 1 (POMGnT1), which extends *O*-mannose residues with a β-1,2 linked GlcNAc[22]. Mutations in other genes encoding putative glycosyltransferases LARGE, fukutin, and fukutin-related protein (FKRP) have also shown to result various forms of CMD[23-25]. Although the effect that these glycosyltransferases have on the post-translational processing of α-DG remain unknown, mutations in the genes encoding these putative glycosyltransferases has been shown to result in α-DG with a lower molecular mass and reduced laminin binding activity[26, 27].

Given the importance of *O*-mannosylation regarding the proper function of α-DG, we implemented the use of both proteomic and glycomic methodologies to map sites of post-translational modification and characterize *O*-glycans released from α-DG[28]. Additional work was also completed to characterize the structure and prevalence of *O*-glycans released from mouse brain models of CMD.

Glycoproteomic methodologies-

Glycosylation is one of the most common post-translational modifications, occurring on approximately 50% of all proteins[29]. Applications of analytical methodologies in glycoproteomics such as mass spectrometry (MS) have proven to be a valuable tool for the analysis of glycoproteins. These analyses have yielded information such as the structure of attached glycans on a glycoprotein or aided in identifying sites of glycosylation[30-32]. A typical protocol for the analysis of glycoproteins involves enzymatic digestion of the sample mixture, usually with trypsin or a combination of enzymes, followed by analysis by liquid chromatography/tandem mass spectrometry[33].

Identification of glycopeptides from a large peptide population have proven to be challenging due to the low abundance of glycosylated peptides relative to the total peptide population. To aid in mapping sites of glycosylation methods have been developed that incorporate chemical tags onto the modified glycopeptides at the site of attachment of the modifying glycan structure. These methods have also been applied to facilitate enrichment of glycopeptides[34]. Methods such as β-elimination followed by Michael addition with DTT (BEMAD) have also been developed, enabling sites of *O*-glycosylation to be tagged. Although these methods enable the site of modification to be identified, all information regarding the attached glycan structure is lost following enzymatic or chemical cleavage[35].

MS methods have been developed to assist in the analysis of simultaneously identifying glycopeptides and glycan structure and mapping modification sites[36]. Usually, MS/MS data obtained from the fragmentation of glycopeptides by collision induced dissociation (CID) yields glycosidic cleavage of the glycan structure and relatively few b & y type ions of the peptide[37]. Therefore consecutive losses equal to the masses of monosaccharides composing the glycan structure are able to be observed. If a peptide bears extensive glycosylation it is difficult to acquire information about the oligosaccharide and

peptide structure in a single spectra[38]. By taking advantage of the capabilities of a linear ion trap mass spectrometer, we developed and applied a pseudo-neutral loss MS^3 method to characterize the modifying glycan structure through a series of neutral losses, and identify the peptide sequence and the exact site of modification on α-DG. In order to map sites of modification, from each full MS scan (300-2000*m/z*), the top 5 most intense peaks were selected to undergo MS/MS fragmentation. Data dependent neutral loss scan was triggered if a neutral loss was observed equal to the singly or doubly charged mass of Hexose, HexNAc, or Neu5Ac.

An alternative method for the dissociation of peptide bonds is electron transfer dissociation (ETD). ETD results from ion-ion reactions of multiply protonated peptides with singly charged anions, and induces fragmentation along the peptide backbone[39]. The resulting fragmentation pattern from ETD has proven to be beneficial to studying PTMs, as an almost complete series of c- and z- type fragment ions are produced and, liable PTMs such as glycosylation or phosphorylation are retained, instead of being cleaved from the peptide backbone as with CID.

Additionally, studies focusing on glycomics have been rapidly expanding, and have revealed that mutations in glycosyltransferases affecting the synthesis of oligosaccharides has given rise to neural and developmental defects, connective tissue abnormalities, and muscular dystrophy[19, 40]. Glycome analysis poses more challenges than genome or proteome analysis because glycan synthesis is not template driven, and glycan structures are typically composed of several different monosaccharide units that are branched and attached to each other in a variety of linkages, which only increases the difficultly of analysis. Within the field of glycomics there are two main categories that are the focus for this type of research: glycome profiling and structure elucidation. Glycome profiling is typically used to compare and detect differences in glycan structures released from different disease states. Structural analysis of glycans is challenging due to their low abundance relative to the complex biological sample being analyzed.[41]

To characterize the structures of various *O*-glycans found to modify α-DG and mouse brain proteins from various models of CMD, *O*-glycans were chemically released from the protein backbone via reductive β-elimination. Following their release, *O*-glycans were permethylated to facilitate their

analysis by LC-MS/MS[42-44]. In addition to obtaining structural information about the released *O*-glycans, the prevalence of each glycan was determined using an unbiased approach called total ion mapping (TIM). Through the developed TIM method, MS/MS spectra are obtained in overlapping windows across a specified mass range ensuring that minor glycan structures are able to be detected[45]. From our analysis we identified 4 *O*-mannose initiated and 17 *O*-GalNAc initiated structures on α-DG isolated from rabbit skeletal muscle[28]. Additionally, similar techniques were applied for the analysis of *O*-glycans released from the mouse brain proteins of various models of CMD. Overall we identified 9 *O*-mannose and 25 *O*-GlaNAc initiated glycan structures from these model systems.

REFERENCES

1. Apweiler, R.; Hermjakob, H.; Sharon, N., On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* **1999,** 1473, (1), 4-8.

2. Spiro, R. G., Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **2002,** 12, (4), 43R-56R.

3. Sentandreu, R.; Northcote, D. H., The structure of a glycopeptide isolated from the yeast cell wall. *Biochem J* **1968,** 109, (3), 419-32.

4. Lommel, M.; Strahl, S., Protein O-mannosylation: conserved from bacteria to humans. *Glycobiology* **2009,** 19, (8), 816-28.

5. Chiba, A.; Matsumura, K.; Yamada, H.; Inazu, T.; Shimizu, T.; Kusunoki, S.; Kanazawa, I.; Kobata, A.; Endo, T., Structures of sialylated O-linked oligosaccharides of bovine peripheral nerve alpha-dystroglycan. The role of a novel O-mannosyl-type oligosaccharide in the binding of alpha-dystroglycan with laminin. *J Biol Chem* **1997,** 272, (4), 2156-62.

6. Smalheiser, N. R.; Haslam, S. M.; Sutton-Smith, M.; Morris, H. R.; Dell, A., Structural analysis of sequences O-linked to mannose reveals a novel Lewis X structure in cranin (dystroglycan) purified from sheep brain. *J Biol Chem* **1998,** 273, (37), 23698-703.

7. Sasaki, T.; Yamada, H.; Matsumura, K.; Shimizu, T.; Kobata, A.; Endo, T., Detection of O-mannosyl glycans in rabbit skeletal muscle alpha-dystroglycan. *Biochim Biophys Acta* **1998,** 1425, (3), 599-606.

8. Abbott, K. L.; Matthews, R. T.; Pierce, M., Receptor tyrosine phosphatase beta (RPTPbeta) activity and signaling are attenuated by glycosylation and subsequent cell surface galectin-1 binding. *J Biol Chem* **2008,** 283, (48), 33026-35.

9. Smalheiser, N. R.; Schwartz, N. B., Cranin: a laminin-binding protein of cell membranes. *Proc Natl Acad Sci U S A* **1987,** 84, (18), 6457-61.

10. Ibraghimov-Beskrovnaya, O.; Ervasti, J. M.; Leveille, C. J.; Slaughter, C. A.; Sernett, S. W.; Campbell, K. P., Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature* **1992,** 355, (6362), 696-702.

11. Ervasti, J. M.; Campbell, K. P., Membrane organization of the dystrophin-glycoprotein complex. *Cell* **1991,** 66, (6), 1121-31.

12. Petrof, B. J.; Shrager, J. B.; Stedman, H. H.; Kelly, A. M.; Sweeney, H. L., Dystrophin protects the sarcolemma from stresses developed during muscle contraction. *Proc Natl Acad Sci U S A* **1993,** 90, (8), 3710-4.

13. Durbeej, M.; Henry, M. D.; Campbell, K. P., Dystroglycan in development and disease. *Curr Opin Cell Biol* **1998,** 10, (5), 594-601.

14. Brancaccio, A.; Schulthess, T.; Gesemann, M.; Engel, J., The N-terminal region of alpha-dystroglycan is an autonomous globular domain. *Eur J Biochem* **1997,** 246, (1), 166-72.

15. Brancaccio, A.; Schulthess, T.; Gesemann, M.; Engel, J., Electron microscopic evidence for a mucin-like region in chick muscle alpha-dystroglycan. *FEBS Lett* **1995,** 368, (1), 139-42.

16. Ervasti, J. M.; Campbell, K. P., A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin. *J Cell Biol* **1993,** 122, (4), 809-23.

17. Combs, A. C.; Ervasti, J. M., Enhanced laminin binding by alpha-dystroglycan after enzymatic deglycosylation. *Biochem J* **2005,** 390, (Pt 1), 303-9.

18. Yoshida-Moriguchi, T.; Yu, L.; Stalnaker, S. H.; Davis, S.; Kunz, S.; Madson, M.; Oldstone, M. B.; Schachter, H.; Wells, L.; Campbell, K. P., O-mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. *Science* 327, (5961), 88-92.

19. Barresi, R.; Campbell, K. P., Dystroglycan: from biosynthesis to pathogenesis of human disease. *J Cell Sci* **2006,** 119, (Pt 2), 199-207.

20. Jurado, L. A.; Coloma, A.; Cruces, J., 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* **1999,** 58, (2), 171-80.

21. Manya, H.; Chiba, A.; Yoshida, A.; Wang, X.; Chiba, Y.; Jigami, Y.; Margolis, R. U.; Endo, T., Demonstration of mammalian protein O-mannosyltransferase activity: coexpression of POMT1 and POMT2 required for enzymatic activity. *Proc Natl Acad Sci U S A* **2004,** 101, (2), 500-5.

22. Yoshida, A.; Kobayashi, K.; Manya, H.; Taniguchi, K.; Kano, H.; Mizuno, M.; Inazu, T.; Mitsuhashi, H.; Takahashi, S.; Takeuchi, M.; Herrmann, R.; Straub, V.; Talim, B.; Voit, T.; Topaloglu, H.; Toda, T.; Endo, T., Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. *Dev Cell* **2001,** 1, (5), 717-24.

23. Longman, C.; Brockington, M.; Torelli, S.; Jimenez-Mallebrera, C.; Kennedy, C.; Khalil, N.; Feng, L.; Saran, R. K.; Voit, T.; Merlini, L.; Sewry, C. A.; Brown, S. C.; Muntoni, F., Mutations in the human LARGE gene cause MDC1D, a novel form of congenital muscular dystrophy with severe mental retardation and abnormal glycosylation of alpha-dystroglycan. *Hum Mol Genet* **2003,** 12, (21), 2853-61.

24. Kobayashi, K.; Nakahori, Y.; Miyake, M.; Matsumura, K.; Kondo-Iida, E.; Nomura, Y.; Segawa, M.; Yoshioka, M.; Saito, K.; Osawa, M.; Hamano, K.; Sakakihara, Y.; Nonaka, I.; Nakagome, Y.; Kanazawa, I.; Nakamura, Y.; Tokunaga, K.; Toda, T., An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy. *Nature* **1998,** 394, (6691), 388-92.

25. Brockington, M.; Yuva, Y.; Prandini, P.; Brown, S. C.; Torelli, S.; Benson, M. A.; Herrmann, R.; Anderson, L. V.; Bashir, R.; Burgunder, J. M.; Fallet, S.; Romero, N.; Fardeau, M.; Straub, V.; Storey, G.; Pollitt, C.; Richard, I.; Sewry, C. A.; Bushby, K.; Voit, T.; Blake, D. J.; Muntoni, F., Mutations in the fukutin-related protein gene (FKRP) identify limb girdle muscular dystrophy 2I as a milder allelic variant of congenital muscular dystrophy MDC1C. *Hum Mol Genet* **2001,** 10, (25), 2851-9.

26. Michele, D. E.; Barresi, R.; Kanagawa, M.; Saito, F.; Cohn, R. D.; Satz, J. S.; Dollar, J.; Nishino, I.; Kelley, R. I.; Somer, H.; Straub, V.; Mathews, K. D.; Moore, S. A.; Campbell, K. P., Post-translational disruption of dystroglycan-ligand interactions in congenital muscular dystrophies. *Nature* **2002,** 418, (6896), 417-22.

27. Mercuri, E.; Brockington, M.; Straub, V.; Quijano-Roy, S.; Yuva, Y.; Herrmann, R.; Brown, S. C.; Torelli, S.; Dubowitz, V.; Blake, D. J.; Romero, N. B.; Estournet, B.; Sewry, C. A.; Guicheney, P.; Voit, T.; Muntoni, F., Phenotypic spectrum associated with mutations in the fukutin-related protein gene. *Ann Neurol* **2003,** 53, (4), 537-42.

28. Stalnaker, S. H.; Hashmi, S.; Lim, J. M.; Aoki, K.; Porterfield, M.; Gutierrez-Sanchez, G.; Wheeler, J.; Ervasti, J. M.; Bergmann, C.; Tiemeyer, M.; Wells, L., Site-mapping and characterization of O-glycan structures on alpha-dystroglycan isolated from rabbit skeletal muscle. *J Biol Chem*.

29. Van den Steen, P.; Rudd, P. M.; Dwek, R. A.; Opdenakker, G., Concepts and principles of O-linked glycosylation. *Crit Rev Biochem Mol Biol* **1998,** 33, (3), 151-208.

30. Novotny, M. V.; Mechref, Y., New hyphenated methodologies in high-sensitivity glycoprotein analysis. *J Sep Sci* **2005,** 28, (15), 1956-68.

31. Zaia, J., Mass spectrometry of oligosaccharides. *Mass Spectrom Rev* **2004,** 23, (3), 161-227.

32. Mechref, Y.; Novotny, M. V.; Krishnan, C., Structural characterization of oligosaccharides using MALDI-TOF/TOF tandem mass spectrometry. *Anal Chem* **2003,** 75, (18), 4895-903.

33. Mechref, Y.; Novotny, M. V., Structural investigations of glycoconjugates at high sensitivity. *Chem Rev* **2002,** 102, (2), 321-69.

34. Witze, E. S.; Old, W. M.; Resing, K. A.; Ahn, N. G., Mapping protein post-translational modifications with mass spectrometry. *Nat Methods* **2007,** 4, (10), 798-806.

35. Wells, L.; Vosseller, K.; Cole, R. N.; Cronshaw, J. M.; Matunis, M. J.; Hart, G. W., Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. *Mol Cell Proteomics* **2002,** 1, (10), 791-804.

36. Dalpathado, D. S.; Desaire, H., Glycopeptide analysis by mass spectrometry. *Analyst* **2008,** 133, (6), 731-8.

37. Wuhrer, M.; Catalina, M. I.; Deelder, A. M.; Hokke, C. H., Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J Chromatogr B Analyt Technol Biomed Life Sci* **2007,** 849, (1-2), 115-28.

38. Wada, Y.; Tajiri, M.; Yoshida, S., Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal Chem* **2004,** 76, (22), 6560-5.

39. Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* **2004,** 101, (26), 9528-33.

40. Freeze, H. H., Human disorders in N-glycosylation and animal models. *Biochim Biophys Acta* **2002,** 1573, (3), 388-93.

41. Vanderschaeghe, D.; Festjens, N.; Delanghe, J.; Callewaert, N., Glycome profiling using modern glycomics technology: technical aspects and applications. *Biol Chem* 391, (2-3), 149-61.

42. Ciucanu, I.; Kerek, F., A Simple and Rapid Method for the Permethylation of Carbohydrates. *Carbohydrate Research* **1984,** 131, (2), 209-217.

43. Dell, A.; Reason, A. J., Carbohydrate analysis. *Curr Opin Biotechnol* **1993,** 4, (1), 52-6.

44. Aoki, K.; Porterfield, M.; Lee, S. S.; Dong, B.; Nguyen, K.; McGlamry, K. H.; Tiemeyer, M., The diversity of O-linked glycans expressed during Drosophila melanogaster development reflects stage- and tissue-specific requirements for cell signaling. *J Biol Chem* **2008**.

45. Aoki, K.; Perlman, M.; Lim, J. M.; Cantu, R.; Wells, L.; Tiemeyer, M., Dynamic developmental elaboration of N-linked glycan complexity in the Drosophila melanogaster embryo. *J Biol Chem* **2007,** 282, (12), 9127-42.

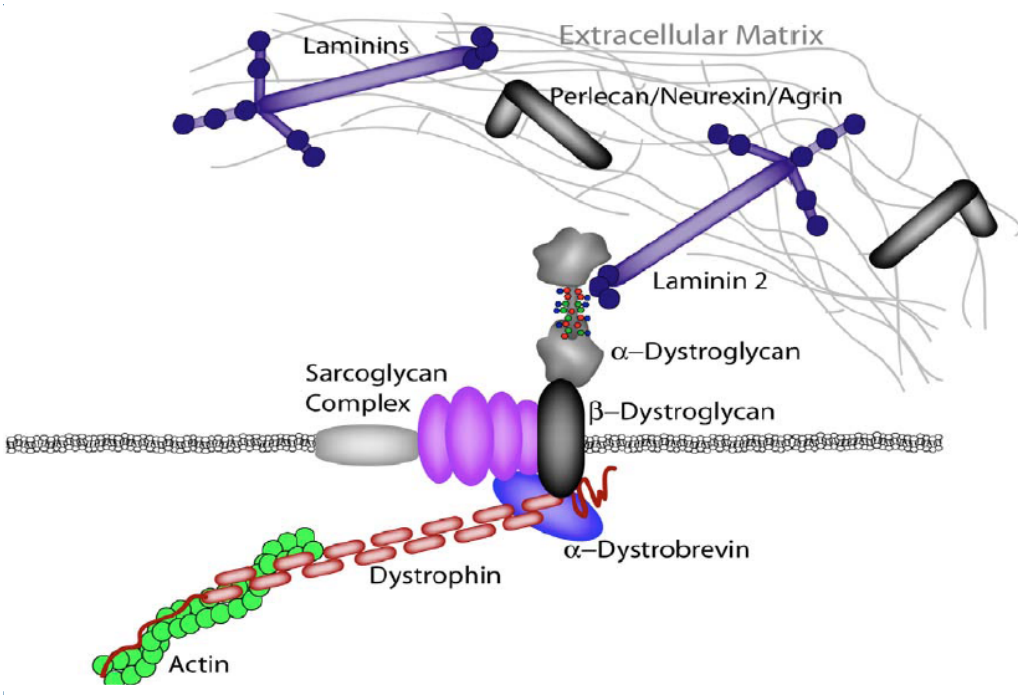


Figure 1-1. Alpha-dystroglycan is the central component of the dystrophin-glycoprotein complex (DGC).
Interaction between the α-subunit of dystroglycan (α-DG) and laminin-2 are dependent upon the addition of *O*-mannose initiated glycan structures through posttranslational modification. A genetic mutation in genes encoding analogous, putative glycosyltransferases has been shown to result in several forms of congenital muscular dystrophy.

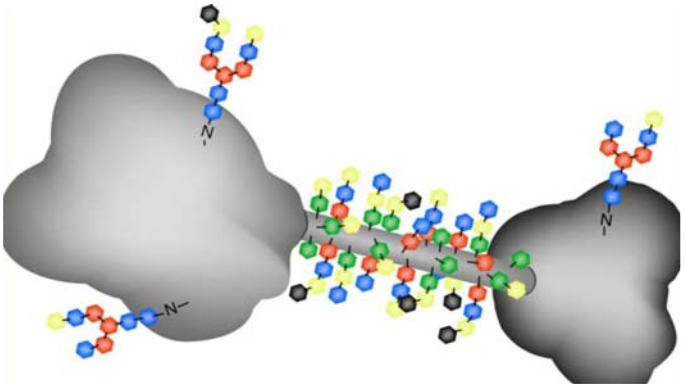


Figure 1-2. Proposed structure of α-DG.
α-DG is proposed to have a dumbbell shape, which is comprised of two globular domains at each end with 3 potential sites of *N*-linked glycosylation and a central stalk region that is heterogeneously glycosylated by *O*-linked glycans that are initiated either Man or GalNAc.

CHAPTER 2

SITE-MAPPING AND CHARACTERIZATION OF *O*-GLYCAN STRUCTURES ON ALPHA-DYSTROGLYCAN ISOLATED FROM RABBIT SKELETAL MUSCLE

Hammond S, Hashmi S, Lim JM, Aoki K, Porterfield M, Gutierrez-Sanchez G, Wheeler J, Ervasti JM, Bergmann C, Tiemeyer M, Wells L. Accepted by *Journal of Biological Chemistry*.
Reprinted here with permission of publisher, 07/12/2010

ABSTRACT

The main extracellular matrix binding component of the dystrophin-glycoprotein complex (DGC), alpha-dystroglycan (α-DG), which was originally isolated from rabbit skeletal muscle, is an extensively *O*-glycosylated protein. Previous studies have shown α-DG to be modified by both *O*-GalNAc and *O*-mannose initiated glycan structures. *O*-mannosylation while accounting for up to 30% of the reported *O*-linked structures in certain tissues, has been rarely observed on mammalian proteins. Mutations in multiple genes encoding defined or putative glycosyltransferases involved in *O*-mannosylation are causal for various forms of Congenital Muscular Dystrophy. Here we explore the glycosylation of purified rabbit skeletal muscle α-DG in detail. Using tandem mass spectrometry approaches, we were able identify 4 *O*-mannose initiated and 17 *O*-GalNAc initiated structures on α-DG isolated from rabbit skeletal muscle. Additionally, we demonstrate the use of tandem mass spectrometry-based workflows to directly analyze glycopeptides generated from the purified protein. By combining glycomics and tandem mass spectrometry analysis of 91 glycopeptides from α-DG, we are able to assign 21 different residues as being modified by *O*-glycosylation with differing degrees of microheterogeneity; 9 sites of *O*-mannosylation and 14 sites of *O*-GalNAcylation are observed with only two sites definitively exhibiting occupancy by either type of glycan. The distribution of identified sites of *O*-mannosylation suggests a limited role for local primary sequence in dictating sites of attachment.

INTRODUCTION

Defects in protein glycosylation related to human disease were first reported in the 1980s and since then upwards of 40 various types of congenital disorders of glycosylation (CDG) have been reported[1]. The term CDG was first used to describe alterations of the N-glycosylation pathway, and was later expanded to include the *O*-glycosylation pathways[1-3]. The importance and complexity of *O*-linked glycosylation has only recently begun to be appreciated[1,3,4]. In particular, mutations in genes encoding (putative) glycosyltransferases which catalyze the addition and extension of *O*-linked mannose-initiated glycans have garnered increased attention in the last decade given that they are causative for several forms of Congenital Muscular Dystrophy (CMD)[5,6].

The most common form of *O*-glycosylation on secretory proteins are the mucin-like *O*-GalNAc structures that are initiated by ppGalNAcTs in the endoplasmic reticulum-Golgi intermediate compartment and/or early cis-Golgi[7]. Additionally, other *O*-linked structures are initiated with alternative monosaccharides, such as *O*-mannose, *O*-glucose, *O*-fucose, *O*-xylose, and *O*-GlcNAc on Ser/Thr residues and the *O*-galactose modification of hydroxylysine residues in collagen domains[4]. The diversity of *O*-mannosylated proteins in mammals, while quite abundant in some tissues (~30% of *O*-glycans released from mouse brains[8]), has not been well characterized. The only clearly identified mammalian protein modified by *O*-mannosylation is alpha-dystroglycan (α-DG)[9].

α-DG is a subunit of dystroglycan and was originally isolated from rabbit skeletal muscle as the extracellular matrix-binding component of the dystrophin-glycoprotein complex (DGC,[10]). The binding to extracellular components such as laminin is dependent upon the addition of *O*-linked oligosaccharides. Numerous studies have shown that proper posttranslational processing of α-DG through the addition of *O*-mannose structures is crucial for proper muscle and brain development[5,6,9]. Of particular interest, several distinct forms of congenital muscular dystrophy have been linked to defects in glycosyltransferases involved in the *O*-mannosylation of α-DG[5,6,9]. Defects in glycosyltransferases involved in *O*-mannose attachment and extension including POMT1/2 and POMGnT1, as well as the putative glycosyltransferase LARGE, are present in various forms of CMD including Walker-Warburg syndrome and Muscle-Eye-Brain disease[5,6,9]. Furthermore, ablation of these gene products in mouse model systems recapitulates much of the pathophysiology of the corresponding human diseases[5,6,9].

Given the importance of *O*-mannosylation to the function of α-DG, we undertook glycomics and glycoproteomic site mapping of α-DG isolated from rabbit skeletal muscle. Since α-DG contains both *O*-Man and *O*-GalNAc initiated structures the use of tagging strategies following β-elimination (such as BEMAD[11]) cannot distinguish glycan type at individual sites. Therefore, we developed and employed methodology for the direct assignment of glycopeptides when coupled with glycomic analysis. *O*-glycan analysis was performed on released permethylated glycans using MS^n tandem mass spectrometry to define the structural diversity of *O*-Man and *O*-GalNAc initiated glycans present on the purified α-DG.

We then performed direct analysis of the peptides/glycopeptides of α-DG, following tryptic digestion with and without glycosidase treatment, via tandem mass spectrometry. Taking advantage of the ability of an ion trap instrument to perform pseudo-neutral loss triggered MS^3 analysis, we were able to assign specific *O*-glycan structures to peptides and in many cases to the exact sites of addition on α-DG. This study, which is the first to map endogenously added *O*-mannose sites from purified functional α-DG, facilitates our understanding of *O*-mannosylation in general. With respect to α-DG, the study highlights the interplay between the *O*-Man and *O*-GalNAc classes of *O*-glycosylation, which will further the development of future studies designed to unravel structure/function relationships for this important glycoprotein as it relates to the pathophysiology of Congenital Muscular Dystrophy.

EXPERIMENTAL PROCEDURES

Protein purification

α-DG was extracted and purified exactly as previously described[12].

Silver Staining and Western Blotting of Gels

SDS-PAGE was performed on a 4-20% Tris-HCl precast gel purchased from Bio-Rad Laboratories. Silver staining was conducted using an adapted protocol from Shevenchko and colleagues[13]. Western blots were performed on semi-dry transferred PVDF membranes using the VIA4 and IIH6 monoclonal antibodies followed by ECL detection as previously described[12].

Glycosidase treatment

The enzyme treated α-DG sample was prepared by combining 1.5μg of α-DG in 15 μl of water, 4μl of 5x incubation buffer (Prozyme), and 1.5 μl of 100 mM DTT and 6 μl of enzyme mixture (GLYCOPRO deglycosylation kit (Prozyme) combined with the PRO-LINK extender kit (Prozyme) containing the following enzymes: N-glycosidase (PNGaseF), sialidase A (*A. ureafaciens*), *O*-glycosidase, β(1-4)galactosidase, chondroitinase, and β-*N*-acetylglucosaminidase). The mock digested α-DG was prepared similarly but lacked enzyme. Both the mock digested and glycosidase treated fractions of α-DG were incubated overnight at 37°C.

Immobilization of laminin-1 on the sensor surface and Surface Plasmon Resonance (SPR)

Murine laminin-1 was not stable at acidic pHs for coupling to the CM5 chips using the standard amine coupling chemistry. In preparation for binding to the streptavidin chip (SA), laminin-1 was first treated with AEBSF as described previously by Colognato *et al.*[14]. Briefly, a 100mM solution of the serine protease inhibitor 4-(2-Aminoethyl)-benzenesulfonyl fluoride hydrochloride (AEBSF-HCl) containing laminin-1 was incubated overnight on ice. Free AEBSF was removed using a Microcon-10 (Amicon) microconcentrator. AEBSF-treated laminin-1 was then biotinylated at a molar ratio of 40:1 with NHSLC-biotin (Pierce, USA) according to the manufacturer's instructions. The reaction product was dialyzed free of unreacted NHSLC-biotin and the reaction product confirmed by HABA assay (data not shown). Approximately 609 RU of biotinylated laminin-1 was bound on the SA chip during a 70 μL injection (5 μL/min) of 100 μg/mL biotinylated laminin-1 in PBS. Samples of glycosidase treated and untreated α-DG were tested at a flow rate of 10 μL/min over the immobilized laminin-1 for 1 min, followed by a 2-min delayed wash to allow the dissociation phase to be recorded. Binding analysis of α-DG to laminin-1 was performed using a Biacore 3000 (Pharmacia Biosensor AB, Uppsala, Sweden). Binding causes a change in the surface Plasmon resonance (SPR), which was detected optically and measured in resonance units (RU). Sensograms were collected as the difference in binding to the laminin-1 versus a blank reference channel. The sensor chip surface was regenerated using 20 μL of glycine–HCl solution (pH 2.5) after each round of binding. The BIAevaluation software 3.0 (BIAcore) was used to analyze binding data.

Release of O-linked glycans

Purified α-DG, approximately 22μg, was transferred to a glass tube and stored at -80°C prior to drying on a lyophilizer. To remove any residual detergent that might have been present from the purification process, the dried protein powder was resuspended in acetone and centrifuged. The acetone supernatant was decanted from the protein pellet and the pellet and any remaining acetone was removed under a stream of nitrogen gas with mild warming (45°C). The dried sample was resuspended in 440μl of Milli-Q water, and a 200μl aliquot was taken for release of *O*-linked glycans. The aliquot was re-

lyophilized and subjected to reductive β-elimination (1M $NaBH_4$ in 50mM NaOH, 18 hours at 45°C). The reaction was neutralized by adding 10% acetic acid drop-wise while vortexing. The completely neutralized sample was desalted by loading onto a small column of AG-50-X8 (1 ml bed volume). Released oligosaccharides were eluted from the column with 3 volumes of 5% of acetic acid, collected, and evaporated to dryness using a Speed Vac. Borate was removed as an azeotrope with methanol and acetic acid by resuspending the dried sample in 9:1 methanol/acetic acid and then drying under a nitrogen stream at 37ºC four times.

Permethylation and analysis of released O-linked glycans

To aid in analysis of *O*-linked glycan structures, the released oligosaccharide mixture was permethylated according to the method of Ciucanu and Kerek[15]. Permethylated glycans were analyzed as described previously[16]. Briefly, following permethylation, glycans were dissolved in 1mM NaOH in 50% methanol. Using a nanoelectrospray source, the *O*-glycan mixture was directly infused into a linear ion trap mass spectrometer (LTQ, Thermo Fisher) at a flow rate of 0.4μl/min. As described previously, total ion mapping (TIM) was used to detect and quantify the prevalence of individual glycans[17]. The consortium for functional glycomics suggested nomenclature was used for all representations of glycan structures in the figures and tables with undefined hexose or HexNAc species displayed in gray.

Protein Digestion

α-DG purified from rabbit skeletal muscle was digested using either sequence grade trypsin (Promega) alone or in combination with endoproteinases Lys-c (Sigma). The samples were diluted to 40mM ammonium bicarbonate and reduced with 100mM DTT for 1 hour at 56ºC, carboxyamidomethylated with 55mM iodoacetamide in the dark for 45 minutes, and then protease digested overnight at 37ºC. In the case of Lys-C the digest was carried out in 6M Urea and the reduction temperature was held at 37ºC followed by dilution to 1M Urea with 40mM ammonium bicarbonate and then overnight trypsin digest. After digestion, the reaction was quenched with 1% trifluoroacetic acid (TFA) making final concentration ~0.1% TFA. The resulting peptides were dried down using a Speed Vac and stored at -20ºC until ready to analyze.

β-elimination followed by Michael addition of dithiothreitol (BEMAD)

The application of BEMAD to tryptic peptides was as previously described [11].

Nano-LC-MS^3

α-DG glycopeptides were analyzed on a linear ion trap mass spectrometer (LTQ; ThermoFisher) using a MS^3 data dependent neutral loss method. The glycopeptides were resuspended in 0.5 μl of solvent B (0.1% formic acid/80% acetonitrile) and 19.5 μl of solvent A (0.1% formic acid), filtered using a 0.2μm spin filter at 12,000 rpm, and loaded on a 75 mm x 8.5 cm C18 reverse phase column/emitter (packed in-house, YMC GEL ODS-AQ120ÅS-5) using a nitrogen pressure bomb. Peptides were eluted over a 160 minute linear gradient increasing from 5% to 100% B over 90 minutes at a flow rate of 250 nl/min. Each full MS scan from 300-2000*m/z* yielded 5 MS/MS scans of the top 5 most intense peaks with a dynamic exclusion of 2 for 30 seconds. Data dependent MS^3 scans were triggered if a neutral loss was observed equal to the singly or doubly charged mass of Hexose, HexNAc, Fucose, or Neu5Ac (sialic acid) within the top 3 peaks from the MS/MS scan.

Data Analysis

The acquired data was searched against a non-redundant rabbit database (generated March 26, 2004) obtained from the National Center for Biotechnology Information (NCBI) using the TurboSequest algorithm (Bio-Works, Thermo Fisher). To aid in identification of glycopeptides, we allowed for a mass increase of 162.1, 203.1, and 365.2 daltons on both threonines and serines looking for the addition of Man, GalNAc, and Hex-HexNAc respectively. Additionally, peptides that were subjected to BEMAD were searched looking for a mass increase of 136.2 daltons on both serines and threonines as previously described. Output files that failed to a yield a final score (Sf) and probability score (P) above .45 and 30 respectively were not considered further. All remaining spectra were manually evaluated for the presence of glycopeptides and sites of modification and were further validated by TurboSequest searches against the rabbit α-DG FASTA sequence combined with the Turbosequest common contaminants database.

RESULTS

Characterization of purified α-DG from rabbit skeletal muscle

α-DG was purified from rabbit skeletal muscle as previously described and validated for purity via silver staining (Fig. 2-1a) and for functional glycosylation using the IIH6 and $VIA4_1$ antibodies (Fig. 2-1b) that have previously been demonstrated to bind functional and glycosylated α-DG, respectively[12]. Purity of the sample was also determined via trypsin digestion followed by LC-MS/MS. Based on the full-length sequence of α-DG, coverage at 1% false-discovery rate (FDR) was only 13%. Decorin and Calsequestrin were also identified in the sample but contributed less than 5% of the total spectral counts assigned to proteins and thus represent minor co-purifying/contaminating proteins. In order to increase coverage, α-DG was subjected to glycosidase treatment (with N-glycosidase, sialidase A [*A. ureafaciens*], *O*-glycosidase, β(1-4)galactosidase, and β-N-acetylglucosaminidase that increased the mobility of the protein upon electrophoresis (Fig. 2-1a). Further, mature α-DG is known to be processed by proteases that cleave off the N-terminus[18] and we were unable to detect any peptides corresponding to this region (Fig. 1C, attempts to determine the N-terminus by automated Edman degradation sequencing were unsuccessful suggesting that the N-terminus is blocked, data not shown). LC-MS/MS analysis following glycosidase and trypsin/endoproteinase-LysC treatment increased overall coverage to 65% (Fig. 2-1c) when one takes into account the proposed cleavage site for the mature protein by Kanagawa and colleagues and the glycopeptides we observed (Supplemental Table 2-S1). Furthermore, SPR experiments were used to confirm that the purified α-DG could bind to laminin in agreement with the method of purification (laminin affinity column) and antibody binding (Fig. 2-1d). As previously observed, using different methodologies[12], treatment of α-DG with sialidase and galactosidase did not have a detrimental effect on laminin binding. Thus, this characterization of the starting material made us confident in moving forward with further analysis of functionally active, glycosylated α-DG.

O-glycans released from α-DG

O-linked glycans were released from α-DG purified from rabbit skeletal muscle by β-elimination, permethylated, and analyzed by NSI-MS/MS. The generated full scans allowed for detection of released

O-linked glycans (Fig. 2-2a) and structure of the *O*-glycans observed in the full MS was assigned based on MS/MS fragmentation. In order to detect glycans in an unbiased manner, the sample was subjected to total ion mapping as previously described (Fig. 2-2b,[17]). Total ion mapping generates MS/MS fragmentation profiles in small overlapping m/z ranges, allowing the detection of fragments that predict the presence of glycans across the full range of detected m/z values. Detected glycans were further confirmed by MS^n fragmentation as needed to define the structure (data not shown). In Figure 2-2c and 2-2d, we present two such MS/MS profiles (from a total of over 700) to display the identification of an *O*-GalNAc (disialyated T antigen) and an *O*-Man (the classical *O*-Man tetrasaccharide, Siaα2-3Galβ1-4GlcNAcβ1-2Man) initiated structure. Table 2-1 includes a list of all of the glycan structures observed from rabbit skeletal muscle α-DG. While there are more total *O*-GalNAc initiated structures observed, *O*-Man-initiated structures represent ~50% of the structures by prevalence.

Assignment of Glycopeptides and Sites of Attachment

Having established the range of structures observed on α-DG, we set out to assign these structures to the polypeptide backbone. Purified α-DG was digested using sequence grade trypsin alone or in combination the endoproteinase Lys-c to increase protein coverage and/or glycosidase treatment to improve digestion, yielding a mixture of peptides and glycopeptides. The resulting mixtures were then analyzed via LC-MS^3 using a linear ion trap mass spectrometer. By taking advantage of the capabilities of the linear ion trap mass spectrometer, we were able to apply MS^3 fragmentation to glycopeptides that generated neutral losses of glycans in MS/MS. In order to identify the glycopeptides, a full MS scan was acquired from 300-2000*m/z* (Fig. 2-3a and 2-4a). From the acquired full scan, MS/MS fragmentation spectra were generated for the top 5 peaks (Fig. 2-3b and 2-4b). Upon fragmentation, if a predetermined neutral loss corresponding to a glycan was observed, a data dependent MS^3 scan was triggered on the neutral-loss peptide that yielded further fragmentation data for the glycopeptide (Fig. 2-3b, 2-3c, 2-4b, and 2-4c).

Through application of this pseudo-neutral loss triggered MS^3 method, we were able to observe, in many cases, sequential monosaccharide losses, defining the glycan structure from its distal end to its

glycosidic attachment to Ser/Thr. The observed losses of glycans (Hexose, HexNAc, and Neu5Ac) species were then fitted to the existing confirmed structures on α-DG that had been determined through reductive β-elimination, permethylation, and MS^n analysis (Table 2-1). The modified peptide was able to be determined upon calculating the neutral loss of glycans and the generation of b- and y-ions in MS/MS and/or MS^3. The peptide sequence was able to be determined by comparing a list of the generated peptide $(M+H)^+$ values against a theoretical list of generated peptides for the α-DG proteins sequence using the MS-digest application from the Prospector web-site created by the University of California at San Francisco. We also used the BEMAD method to aid in mapping sites modified by *O*-linked glycans[11]. While this method proved to be beneficial by indicating the modified residue in a limited set of cases (Supplemental Table 2-S1), it is not capable of distinguishing between *O*-GalNAc or *O*-mannose initiated glycan structure. Thus, in order to make more confident assignments of the glycan structure responsible for modification at a particular Ser/Thr residue, we examined the b- and y- ions that were generated from MS/MS and MS^3 fragmentation of the peptide backbone. By comparing the theoretical b- and y- ions of *O*-GalNAc or *O*-Man containing fragments with those that were observed in the two spectra, we were able to determine in many cases the exact residue modified. For example, Figure 2-3c and 2-4c show glycopeptides from α-DG that were modified by the addition of an *O*-GalNAc initiated glycan structure, disialyated T-antigen, and an *O*-mannose initiated glycan structure, Siaa2-3Galb1-4GlcNAcb1-2Man, at S475 and S485, respectively. Upon fragmentation, b- and y- ions still modified by Hexose or HexNAc allow unequivocal assignment of the structures to specific residues. Similar strategies were applied for all *O*-Man and *O*-GalNAc initiated structures and the results are summarized in Table 2-2 and supplemental Table 2-S1.

DISCUSSION

O-linked glycans containing mannose were first isolated from an enriched mixture of brain chondroitin sulfate proteoglycans, with a core structure suggested to be Galβ1-4GlcNAcβ-1-2Man-Ser/Thr nearly 30 years ago[19]. However, sites of modification have not previously been mapped from native sources for the most well characterized *O*-mannosylated protein, α-DG. Given the importance of

O-glycosylation for proper function of α-DG, we sought here to map defined glycan structures to sites of attachment on the polypeptide from endogenously glycosylated α-DG isolated from rabbit skeletal muscle.

α-DG was purified from rabbit skeletal muscle as previously described and shown to be highly enriched and functionally glycosylated (Fig. 2-1,[12]). In order to map glycan structures to specific sites, we first released and permethylated the glycans from the glycoprotein so that we could get detailed fragmentation defining the full set of glycans present on α-DG (Fig. 2-2). This allowed us to determine that there were at least 21 different *O*-linked glycans present on α-DG purified from rabbit skeletal muscle. Four of these structures were initiated by *O*-Man with the classical, previously described[20], Siaα2-3Galβ1-4GlcNAcβ1-2Man tetrasaccharide structure being the most prevalent (Table 2-1). Only one branched *O*-Man structure was observed in rabbit skeletal muscle-derived α-DG at less than 0.1% prevalence (Table 2-1), consistent with the proposal that the brain-specific enzyme, GnT-Vb (GnT-IX), is responsible for *O*-Man branching[21].

With the glycans on α-DG defined, direct glycopeptide analysis following enzymatic digestion of the protein was performed via pseudo-neutral loss-triggered MS^3 analysis (Figs. 2-3 and 2-4). This procedure relies on the neutral loss of a glycan mass to trigger further fragmentation of the glycopeptide. Given the lability of the glycosidic linkage, most glycopeptides generate dominant neutral loss peaks associated with glycan fragmentation upon collision-induced dissociation (Fig. 2-3 and 2-4). Further fragmentation of the species that has undergone a neutral loss provides further glycan losses, as well as peptide b- and y-ions to assist in the assignment of the peptide and the site(s) of glycosylation. To facilitate improved digestion and better coverage for the resulting peptides, endoproteinase Lys-C under denaturing conditions followed by trypsin digestion was used. Furthermore, we found that partial deglycosylation greatly facilitated digestion and glycopeptides assignments. While incomplete exoglycosidase treatment allowed discrimination between sites of *O*-Man and *O*-GalNAc initiation, it limited mapping of intact glycan structures at many sites (Table 2-2 and Supplemental Table 1). In several cases, there was not sufficient fragmentation information (i.e. fragments containing glycans) to

map the exact site of attachment but we could assign the glycan to a particular peptide or subset of residues in the peptide (Table 2-2 and Supplemental Table 2-S1). Of note, stretches of Thr residues can be particularly problematic for mapping sites of GalNAc attachment since the molecular weight of two adjacent Thr residues is almost identical to the weight of a HexNAc.

In total, we observed 91 glycopeptides in our analyses that allowed us to assign 16 specific *O*-glycosylated residues within α-DG (Fig. 2-1c). In addition, another 16 sites of modification were restricted to a small subset of possible residues (Fig. 2-1c). As expected, for many of the sites of modification, we saw microheterogeneity; glycopeptides with different glycan structures on the identical residues were observed (Supplemental Table 2-S1). We also observed that 2 sites of glycosylation (S475 and T478) could accept *O*-Man and *O*-GalNAc initiated glycan structures. This suggests that *O*-mannosylation, at least on these sites, is substoichiometric since the enzymes for *O*-Man attachment are localized in the ER and likely precede the *O*-GalNAc machinery that is localized to the cis-Golgi and/or ERGIC[7, 9]. In total, we observed 24 sites of glycosylation on rabbit skeletal muscle α-DG including 9 sites of *O*-mannosylation.

Recently, an *in vitro* study using recombinant POMT1/2 enzymes and synthetic peptides derived from α-DG performed by Manya *et al*, concluded that mammalian *O*-mannosylation is dependent upon a consensus sequence (I*X*PT(P/*X*)T*X*P*XXXX*PT*X*(T/*X*)*XX*)[22]. When we compared the 9 sites we identified as *O*-mannosylated with the proposed consensus site, 6 of our defined sites do not fit this model. However, in our studies, there are three other unresolved sites of *O*-Man within a single peptide, containing eight potential sites of modification, that are consistent with their model. The sites reported in the *in vitro* study, T404, T406, and T414, potentially overlap with the three sites of *O*-Man modification localized between residues 404-424 on endogenously glycosylated α-DG. Breloy and colleagues in 2008, relying on data generated via overexpression of fragments of human α-DG in epithelial cells, argued that *O*-mannosylation was regulated in a much more complex manner than a simple local primary sequence[23]. Alignment of our *O*-Man sites generated from endogenous rabbit skeletal muscle α-DG provided no obvious local consensus site for attachment (Supplemental Table 2-S2) and thus our findings are in

agreement with Breloy and colleagues. Therefore, we conclude that *O*-mannosylation of particular residues is not regulated solely by a local consensus sequence. Further work is needed to determine the mechanism by which residues for *O*-Man addition are selected and to elucidate the effect of *O*-Man modification on further modification of the glycoprotein by *O*-GalNAc-initiated structures.

Campbell and colleagues recently demonstrated that O-mannosyl phosphorylation was present on α-DG and was required for laminin binding[24]. Furthermore, that study placed phosphomannose at Thr-379 on a human α-DG construct isolated from cell lines with minimal LARGE activity. Of particular interest to this study, we were unable to observe the analogous rabbit peptide (374-389) by our methods. Presumably, this is because of an unknown LARGE-dependent modification of the phosphorylated *O*-Man structure. Without knowledge of the complete chemical nature of the LARGE-dependent modification, peptides bearing this structure would be missed using our pseudo-neutral loss triggered methods. For completeness, it should also be noted that Thr-381 and Thr-388 of this peptide were also observed to be *O*-Man modified in the human overexpression study[24] and they are in a stretch of sequence similar, but not identical, to the proposed consensus sequence of Endo and colleagues [24].

Interestingly, beyond this missed peptide that contains the phospho-mannose containing trisaccharide extended with an unknown large-dependent modification[24], we did not detect 4 other predicted tryptic peptides of greater than 4 amino acids in length. Two of these peptides are contiguous and represent the extreme N-terminus of the fully processed polypeptide. Given that the N-terminus of the mature protein was apparently blocked based on results of automated Edman sequencing, it is not surprising that the extreme N-terminus peptide was not observed as the nature of the moiety blocking the N-terminus is unknown. This leaves us with 3 peptides that we failed to detect, all of which contain Ser/Thr residues. Two of these peptides (338-359 and 550-572) are quite large and contains 6 and 5 potential sites of glycosylation, respectively. Thus, it is likely that if multiple sites were utilized for glycosylation that these peptide would exceed the upper mass limit of the instrument (2000 *m/z*). The remaining unexplained absent peptide is 583-597. This peptide only contains one potential site of *O*-linked glycosylation (Ser-586). Given the modest size of this peptide, its lack of multiple sites of

glycosylation, and the fact that both peptides flanking this sequence were assigned, failure to detect this peptide is difficult to explain unless like peptide 374-389 this peptide also contains the large-modified phospho-mannose trisaccharide at Ser-586. Given that we chose to map sites on fully functional α-DG, it is not surprising that we were unable to map the phospho-mannose containing trisaccharide peptide(s) that we had previously assigned in a Large-deficient cell line[24]. However, based off the absence of detection in this study, we speculate that other peptides, including 583-597 and possibly 338-359 and 550-572 may indeed be modified in a Large-dependent manner as well.

In conclusion, we have developed and implemented a work flow that enabled us to assign defined *O*-glycan structures to specific residues on the polypeptide by utilizing both glycomics and glycopeptidomics. The resulting site map describing both *O*-Man and *O*-GalNAc initiated glycosylation on α-DG isolated from rabbit skeletal muscle provides a framework for elucidating structure/function relationships for this complex glycoprotein. It should also facilitate a greater understanding of the interplay between *O*-GalNAcylation and *O*-mannosylation, two glycosylation pathways that theoretically are competing for the same sites of modification. Given that *O*-mannosylation is defective in multiple forms of Congenital Muscular Dystrophy, is required for α-DG function, and is likely found on other yet-to-be-identified mammalian proteins, the work presented here lays an essential groundwork for future functional studies.

REFERENCES

1. Jaeken, J.; Hennet, T.; Freeze, H. H.; Matthijs, G., On the nomenclature of congenital disorders of glycosylation (CDG). *J Inherit Metab Dis* **2008**.

2. Aebi, M.; Helenius, A.; Schenk, B.; Barone, R.; Fiumara, A.; Berger, E. G.; Hennet, T.; Imbach, T.; Stutz, A.; Bjursell, C.; Uller, A.; Wahlstrom, J. G.; Briones, P.; Cardo, E.; Clayton, P.; Winchester, B.; Cormier-Dalre, V.; de Lonlay, P.; Cuer, M.; Dupre, T.; Seta, N.; de Koning, T.; Dorland, L.; de Loos, F.; Kupers, L.; et al., Carbohydrate-deficient glycoprotein syndromes become congenital disorders of glycosylation: an updated nomenclature for CDG. First International Workshop on CDGS. *Glycoconj J* **1999,** 16, (11), 669-71.

3. Cohn, R. D., Dystroglycan: important player in skeletal muscle and beyond. *Neuromuscul Disord* **2005,** 15, (3), 207-17.

4. Haltiwanger, R. S.; Lowe, J. B., Role of glycosylation in development. *Annu Rev Biochem* **2004,** 73, 491-537.

5. Barresi, R.; Campbell, K. P., Dystroglycan: from biosynthesis to pathogenesis of human disease. *J Cell Sci* **2006,** 119, (Pt 2), 199-207.

6. Martin, P. T., Congenital muscular dystrophies involving the O-mannose pathway. *Curr Mol Med* **2007,** 7, (4), 417-25.

7. Ten Hagen, K. G.; Fritz, T. A.; Tabak, L. A., All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology* **2003,** 13, (1), 1R-16R.

8. Chai, W.; Yuen, C. T.; Kogelberg, H.; Carruthers, R. A.; Margolis, R. U.; Feizi, T.; Lawson, A. M., High prevalence of 2-mono- and 2,6-di-substituted manol-terminating sequences among O-glycans released from brain glycopeptides by reductive alkaline hydrolysis. *Eur J Biochem* **1999,** 263, (3), 879-88.

9. Endo, T.; Manya, H., O-mannosylation in mammalian cells. *Methods Mol Biol* **2006,** 347, 43-56.

10. Ervasti, J. M.; Campbell, K. P., A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin. *J Cell Biol* **1993,** 122, (4), 809-23.

11. Wells, L.; Vosseller, K.; Cole, R. N.; Cronshaw, J. M.; Matunis, M. J.; Hart, G. W., Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. *Mol Cell Proteomics* **2002,** 1, (10), 791-804.

12. Combs, A. C.; Ervasti, J. M., Enhanced laminin binding by alpha-dystroglycan after enzymatic deglycosylation. *Biochem J* **2005,** 390, (Pt 1), 303-9.

13. Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M., Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels. *Analytical Chemistry* **1996,** 68, (5), 850-858.

14. Colognato, H.; Winkelmann, D. A.; Yurchenco, P. D., Laminin polymerization induces a receptor-cytoskeleton network. *J Cell Biol* **1999,** 145, (3), 619-31.

15. Ciucanu, I.; Kerek, F., A Simple and Rapid Method for the Permethylation of Carbohydrates. *Carbohydrate Research* **1984,** 131, (2), 209-217.

16. Aoki, K.; Porterfield, M.; Lee, S. S.; Dong, B.; Nguyen, K.; McGlamry, K. H.; Tiemeyer, M., The diversity of O-linked glycans expressed during Drosophila melanogaster development reflects stage- and tissue-specific requirements for cell signaling. *J Biol Chem* **2008**.

17. Aoki, K.; Perlman, M.; Lim, J. M.; Cantu, R.; Wells, L.; Tiemeyer, M., Dynamic developmental elaboration of N-linked glycan complexity in the Drosophila melanogaster embryo. *J Biol Chem* **2007,** 282, (12), 9127-42.

18. Kanagawa, M.; Saito, F.; Kunz, S.; Yoshida-Moriguchi, T.; Barresi, R.; Kobayashi, Y. M.; Muschler, J.; Dumanski, J. P.; Michele, D. E.; Oldstone, M. B.; Campbell, K. P., Molecular recognition by LARGE is essential for expression of functional dystroglycan. *Cell* **2004,** 117, (7), 953-64.

19. Finne, J.; Krusius, T.; Margolis, R. K.; Margolis, R. U., Novel mannitol-containing oligosaccharides obtained by mild alkaline borohydride treatment of a chondroitin sulfate proteoglycan from brain. *J Biol Chem* **1979,** 254, (20), 10295-300.

20. Smalheiser, N. R.; Haslam, S. M.; Sutton-Smith, M.; Morris, H. R.; Dell, A., Structural analysis of sequences O-linked to mannose reveals a novel Lewis X structure in cranin (dystroglycan) purified from sheep brain. *J Biol Chem* **1998,** 273, (37), 23698-703.

21. Alvarez-Manilla, G.; Troupe, K.; Fleming, M.; Martinez-Uribe, E.; Pierce, M., Comparison of the substrate specificities and catalytic properties of the sister N-acetylglucosaminyltransferases, GnT-V and GnT-Vb (IX). *Glycobiology* 20, (2), 166-74.

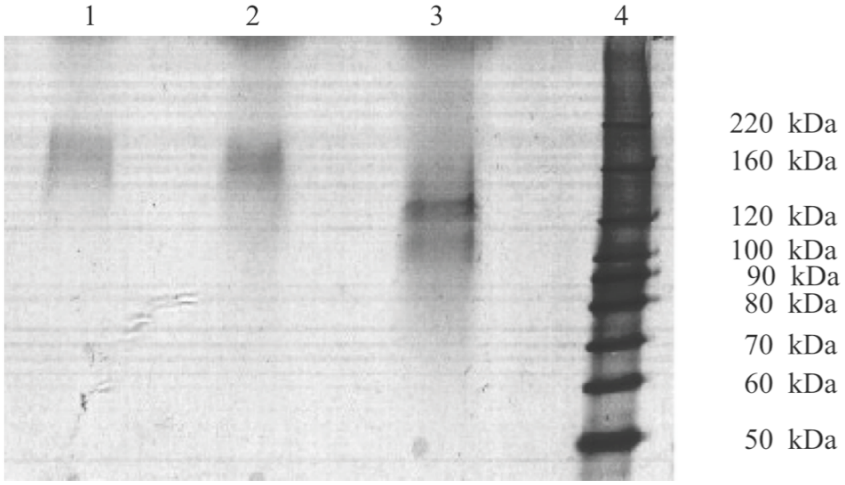
22. Manya, H.; Suzuki, T.; Akasaka-Manya, K.; Ishida, H. K.; Mizuno, M.; Suzuki, Y.; Inazu, T.; Dohmae, N.; Endo, T., Regulation of Mammalian Protein O-Mannosylation: PREFERENTIAL AMINO ACID SEQUENCE FOR O-MANNOSE MODIFICATION. *J Biol Chem* **2007,** 282, (28), 20200-6.

23. Breloy, I.; Schwientek, T.; Gries, B.; Razawi, H.; Macht, M.; Albers, C.; Hanisch, F. G., Initiation of mammalian O-mannosylation in vivo is independent of a consensus sequence and controlled by peptide regions within and upstream of the alpha-dystroglycan mucin domain. *J Biol Chem* **2008,** 283, (27), 18832-40.

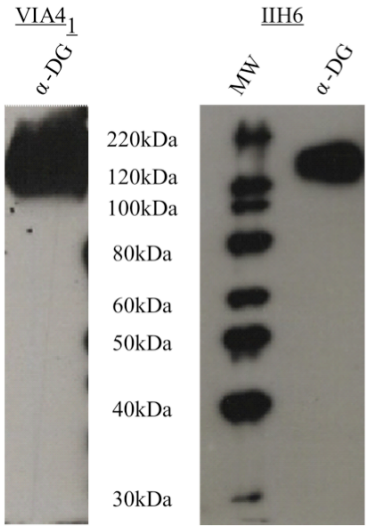
24. Yoshida-Moriguchi, T.; Yu, L.; Stalnaker, S. H.; Davis, S.; Kunz, S.; Madson, M.; Oldstone, M. B.; Schachter, H.; Wells, L.; Campbell, K. P., O-mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. *Science* 327, (5961), 88-92.

Figure 2-1. Purified, functionally glcosylated α-DG from rabbit skeletal muscle.
(a) Silver staining following SDS-PAGE of purified α-DG (lane 1). Lane 2 and 3 represent mock or glycoside (N-glycosidase F, sialidase, endo-O-glycosidase, β(1-4)galactosidase, and β-N-acetylglucosaminidase)-treated α-DG . (b) Wetern blot analysis following SDS-page of purified α-DG with the glycan-dependent anti-α-DG monoclonal $VIA4_1$ and IIH6, which specifically recognizes fully glycosylated, functionally active α-DG. (c) Protein sequence derived from the dystroglycan gene with the capitalized, bolded sequence representing the predicted mature α-DG protein. Peptides assigned by tandem mass spectrometry are underlined. Sites detected to be modified by GalNAc are highlighted in blue; sites of O-mannosylation are highlighted with red; residues observed to be modified by both GalNAc and mannose are green. Sites of potential modification are highlighted similarly and distinguished by striking through the modified residue. (d) Untreated α-DG (a), α-DG treated with b-galactosidase and sialidase (a+b), or glycosidases alone (b, enzymes without α-DG present) binding to immobilized laminin-1 as measured by surface plasmon resonance.

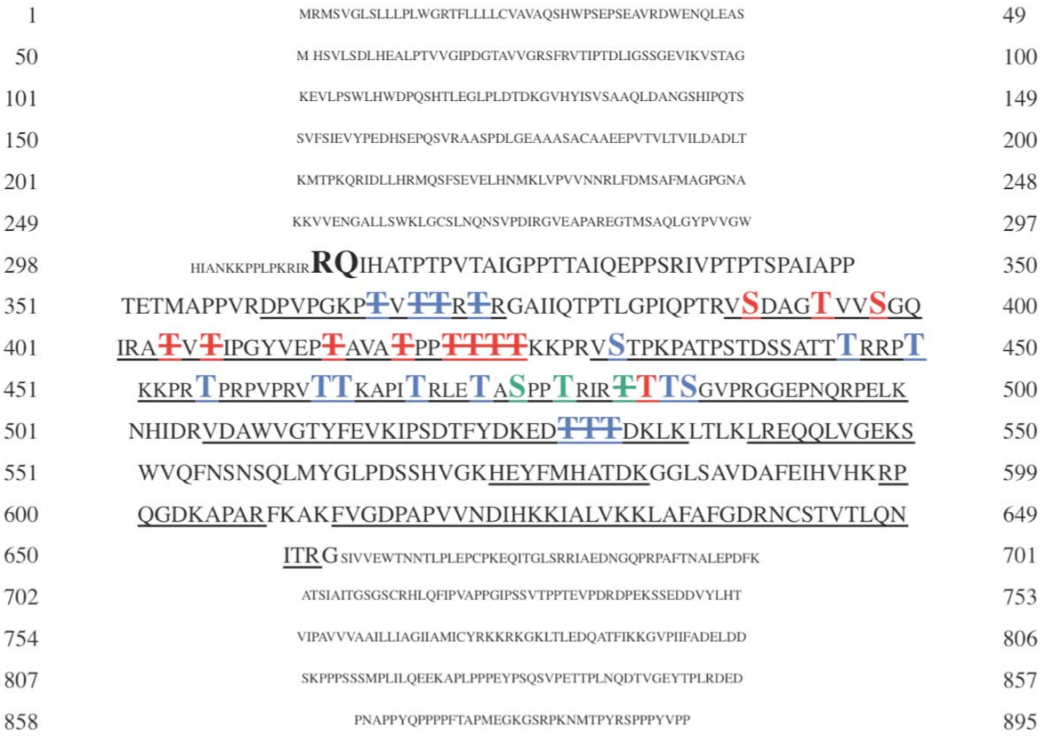
a)


1
2
3
4
220 kDa
160 kDa
120 kDa
100 kDa
90 kDa
80 kDa
70 kDa
60 kDa
50 kDa

b)


VIA4_1
IIH6
α-DG
MW
α-DG
220kDa
120kDa
100kDa
80kDa
60kDa
50kDa
40kDa
30kDa

c)


1 MRMSVGLSLLLPLWGRTFLLLLCVAVAQSHWPSEPSEAVRDWENQLEAS 49
50 M HSVLSDLHEALPTVVGIPDGTAVVGRSFRVTIPTDLIGSSGEVIKVSTAG 100
101 KEVLPSWLHWDPQSHTLEGLPLDTDKGVHYISVSAAQLDANGSHIPQTS 149
150 SVFSIEVYPEDHSEPQSVRAASPDLGEAAASACAAEEPVTVLTVILDADLT 200
201 KMTPKQRIDLLHRMQSFSEVELHNMKLVPVVNNRLFDMSAFMAGPGNA 248
249 KKVVENGALLSWKLGCSLNQNSVPDIRGVEAPAREGTMSAQLGYPVVGW 297
298 HIANKKPPLPKRIRRQIHATPTPVTAIGPPTTAIQEPPSRIVPTPTSPAIAPP 350
351 TETMAPPVRDPVPGKPTVTTRTRGAIIQTPTLGPIQPTRVSDAGTVVSGQ 400
401 IRATVTIPGYVEPTAVATPPTTTTKKPRVSTPKPATPSTDSSATTTRRPT 450
451 KKPRTPRPVPRVTTKAPITRLETASPPTRIRTTTSGVPRGGEPNQRPELK 500
501 NHIDRVDAWVGTYFEVKIPSDTFYDKEDTTTDKLKLTLKLREQQLVGEKS 550
551 WVQFNSNSQLMYGLPDSSHVGKHEYFMHATDKGGLSAVDAFEIHVHKRP 599
600 QGDKAPARFKAKFVGDPAPVVNDIHKKIALVKKLAFAFGDRNCSTVTLQN 649
650 ITRGSIVVEWTNNTLPLEPCPKEQITGLSRRIAEDNGQPRPAFTNALEPDFK 701
702 ATSIAITGSGSCRHLQFIPVAPPGIPSSVTPPTEVPDRDPEKSSEDDVYLHT 753
754 VIPAVVVAAILLIAGIIAMICYRKKRKGKLTLEDQATFIKKGVPIIFADELDD 806
807 SKPPPSSSMPLILQEEKAPLPPPEYPSQSVPETTPLNQDTVGEYTPLRDED 857
858 PNAPPYQPPPPFTAPMEGKGSRPKNMTPYRSPPPYVPP 895

d)

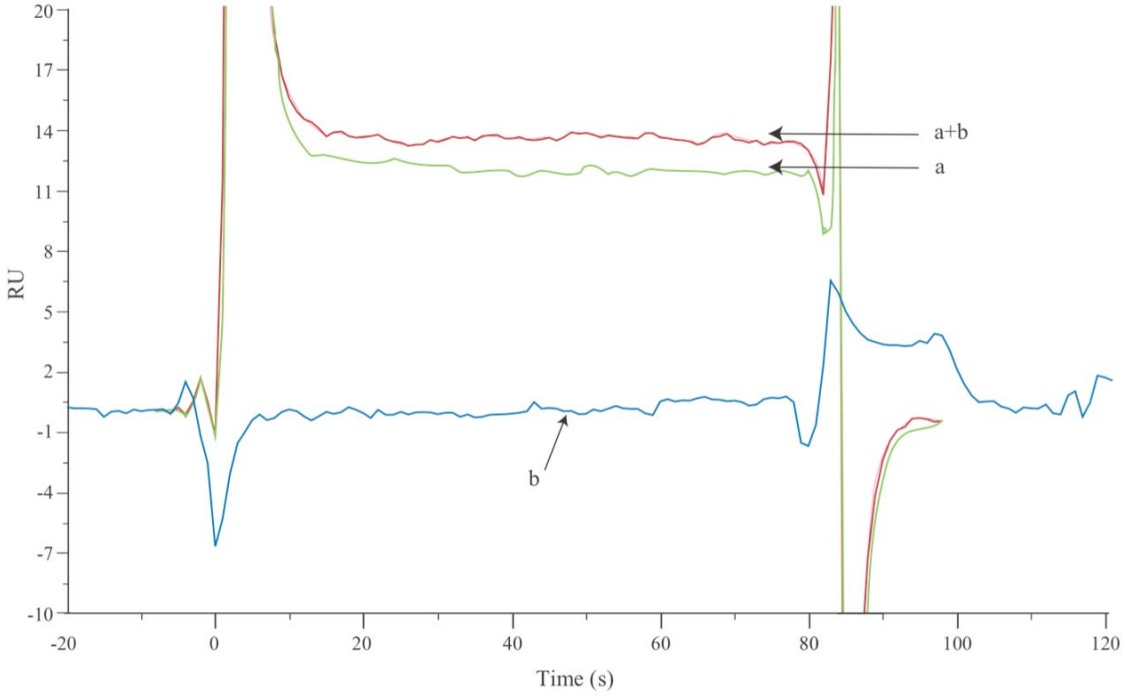
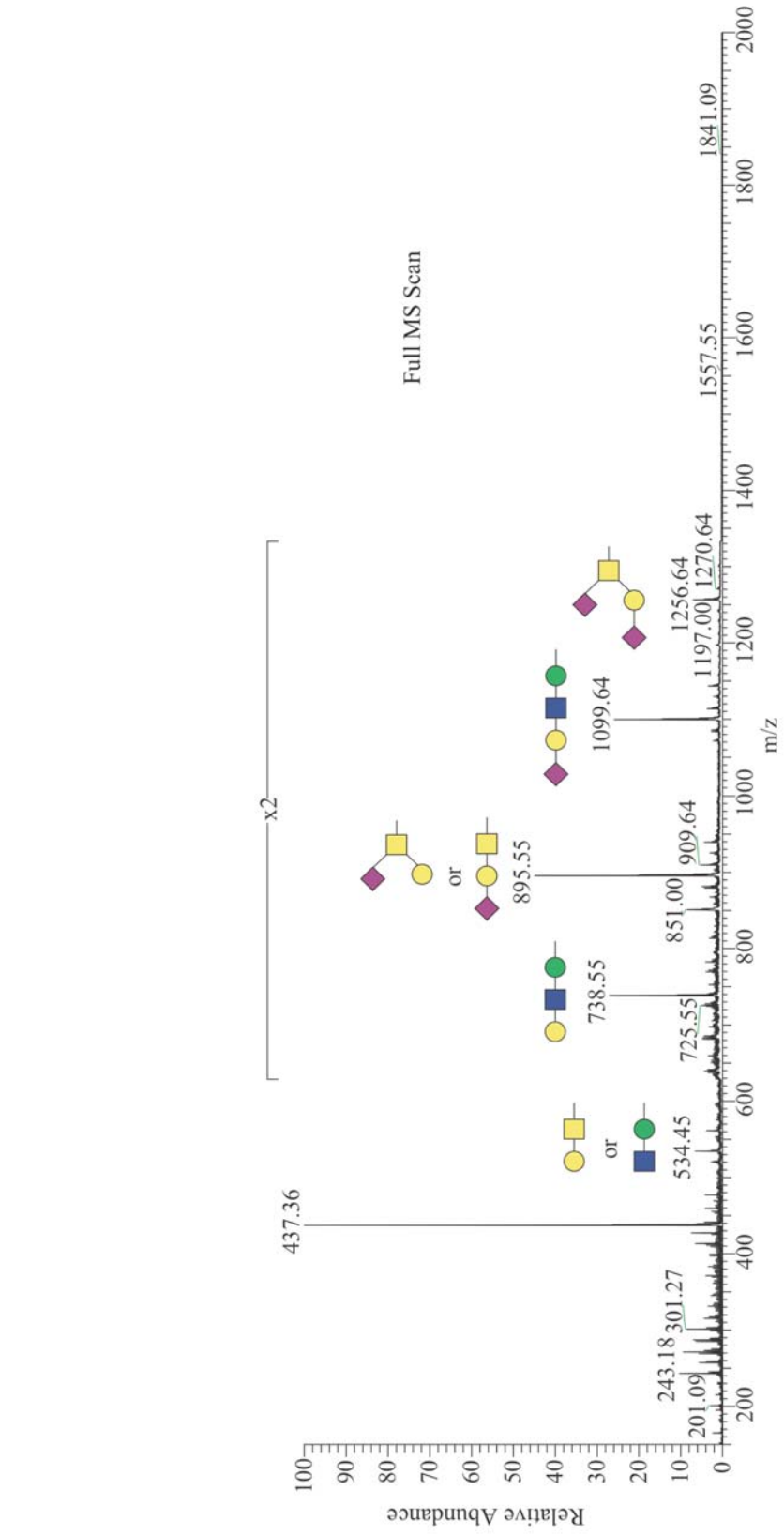
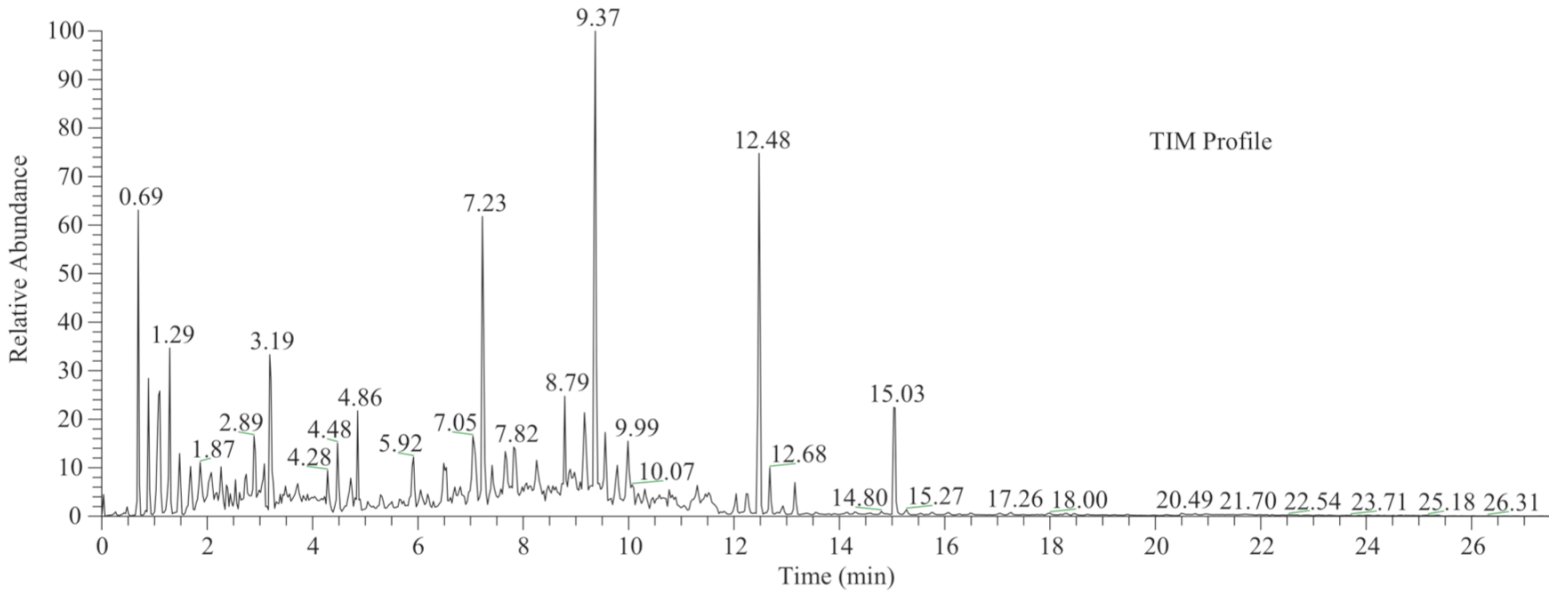


Figure 2-2. Purified, functionally glcosylated α-DG from rabbit skeletal muscle.
(a) Silver staining following SDS-PAGE of purified α-DG (lane 1). Lane 2 and 3 represent mock or glycoside (N-glycosidase F, sialidase, endo-O-glycosidase, *β*(1-4)galactosidase, and *β*-N-acetylglucosaminidase)- treated α-DG. **(b)** Wetern blot analysis following SDS-page of purified α-DG with the glycan-dependent anti-α-DG monoclonal VIA41 and IIH6, which specifically recognizes fully glycosylated, functionally active α-DG. **(c)** Protein sequence derived from the dystroglycan gene with the capitalized, bolded sequence representing the predicted mature α-DG protein. Peptides assigned by tandem mass spectrometry are underlined. Sites detected to be modified by GalNAc are highlighted in blue; sites of *O*-mannosylation are highlighted with red; residues observed to be modified by both GalNAc and mannose are green. Sites of potential modification are highlighted similarly and distinguished by striking through the modified residue. **(d)** Untreated α-DG (a), α-DG treated with β-galactosidase and sialidase (a + b), or glycosidases alone (b, enzymes without α-DG present) binding to immobilized laminin-1 as measured by surface plasmon resonance.

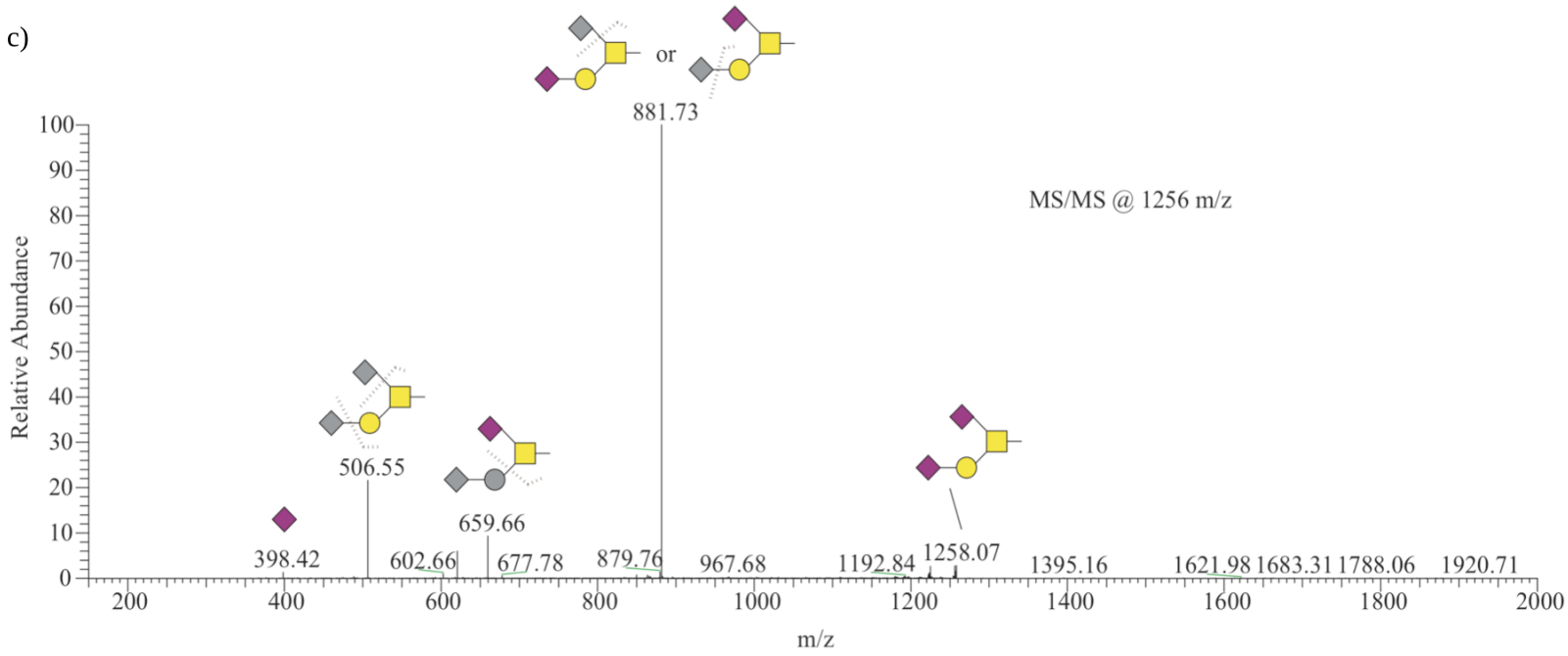
a)


Full MS Scan
Relative Abundance
m/z
x2
437.36
534.45
725.55
738.55
851.00
895.55
909.64
1099.64
1197.00
1256.64
1270.64
1557.55
1841.09
243.18
301.27
201.09
or
or

b)


TIM Profile
Relative Abundance
Time (min)
0.69
1.29
1.87
2.89
3.19
4.28
4.48
4.86
5.92
7.05
7.23
7.82
8.79
9.37
9.99
10.07
12.48
12.68
14.80
15.03
15.27
17.26
18.00
20.49
21.70
22.54
23.71
25.18
26.31

c)


or
881.73
MS/MS @ 1256 m/z
Relative Abundance
100
90
80
70
60
50
40
30
20
10
0
506.55
659.66
398.42
602.66
677.78
879.76
967.68
1192.84
1258.07
1395.16
1621.98
1683.31
1788.06
1920.71
200
400
600
800
1000
1200
1400
1600
1800
2000
m/z

d)

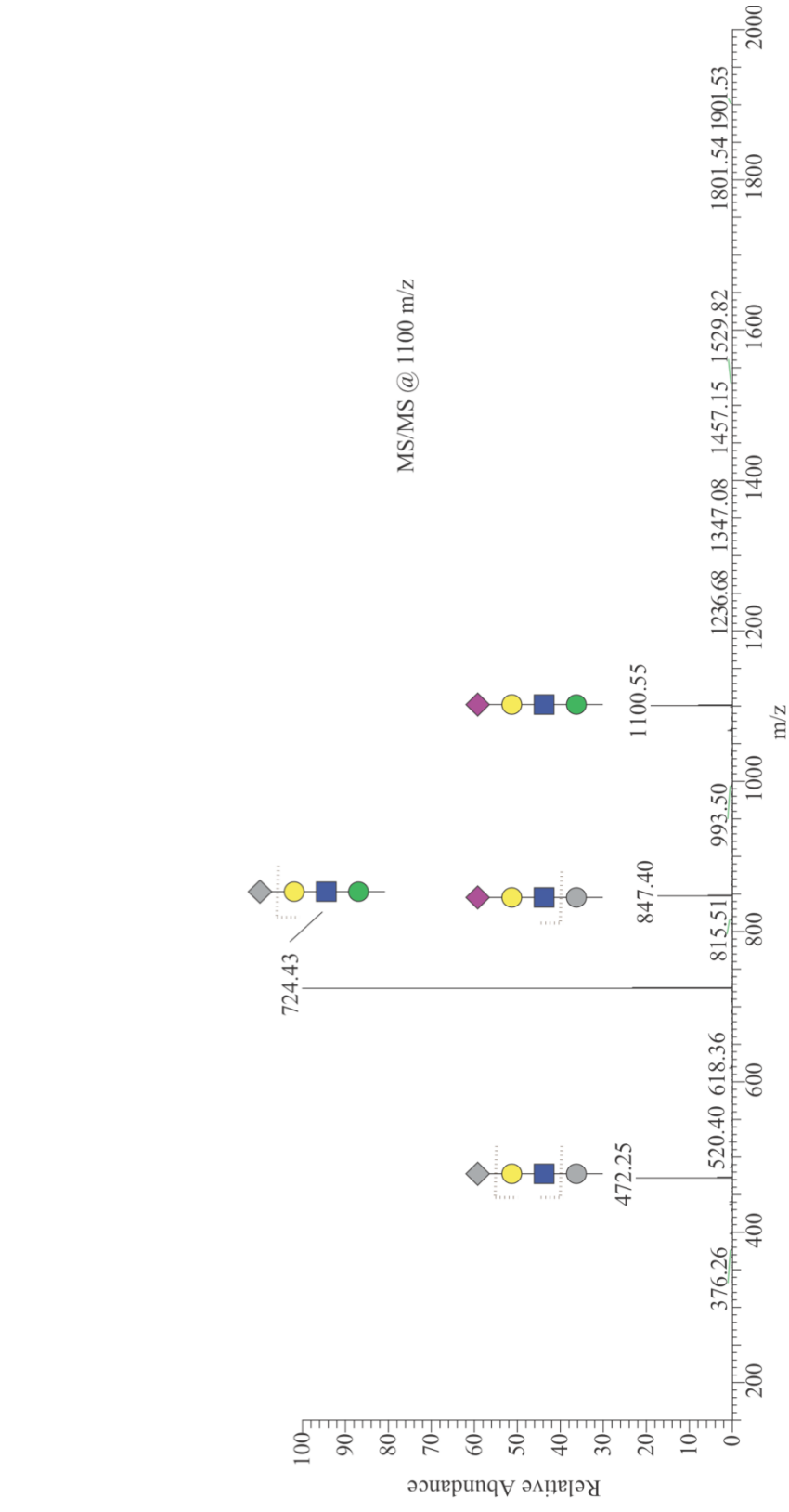
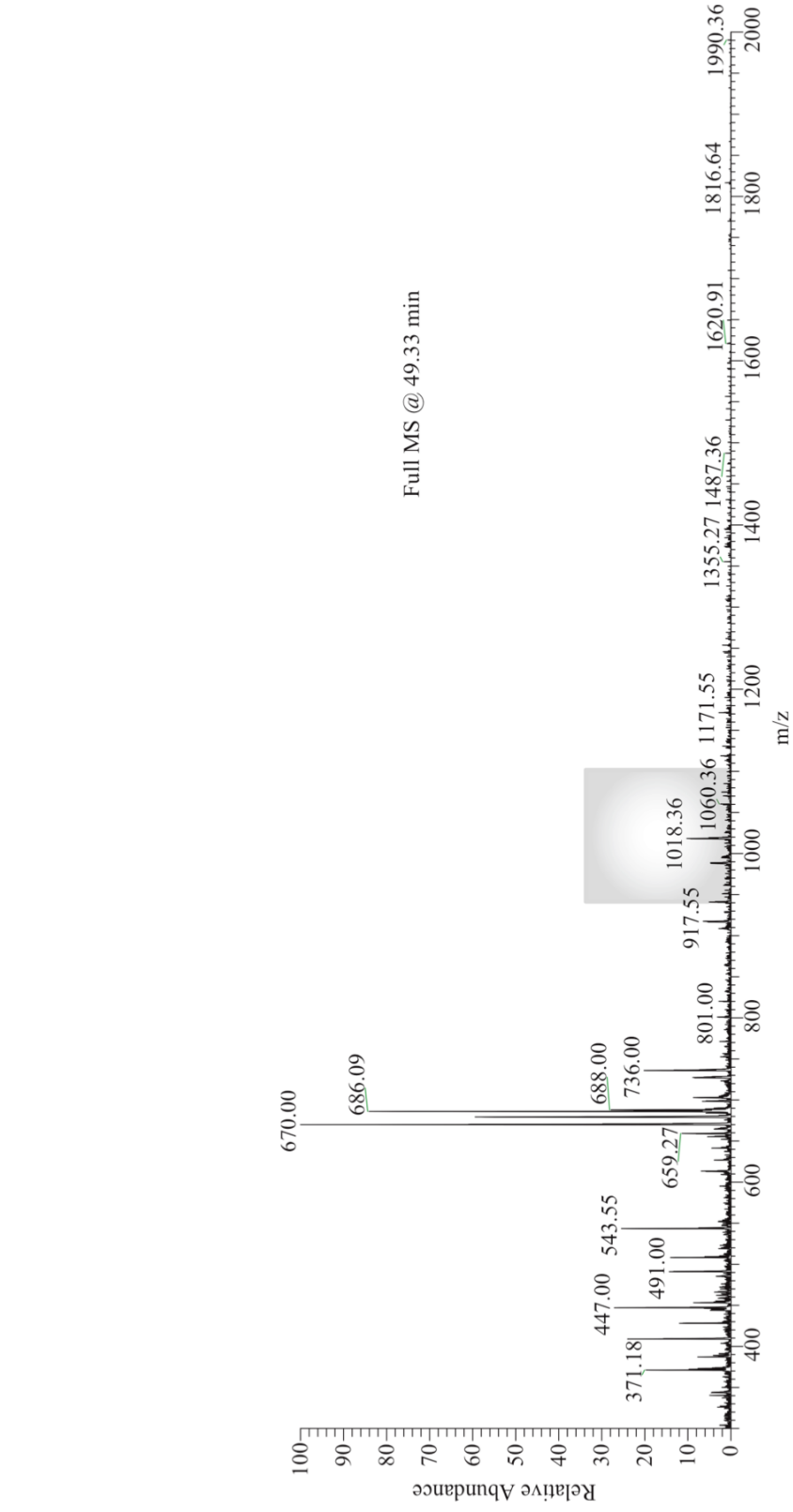
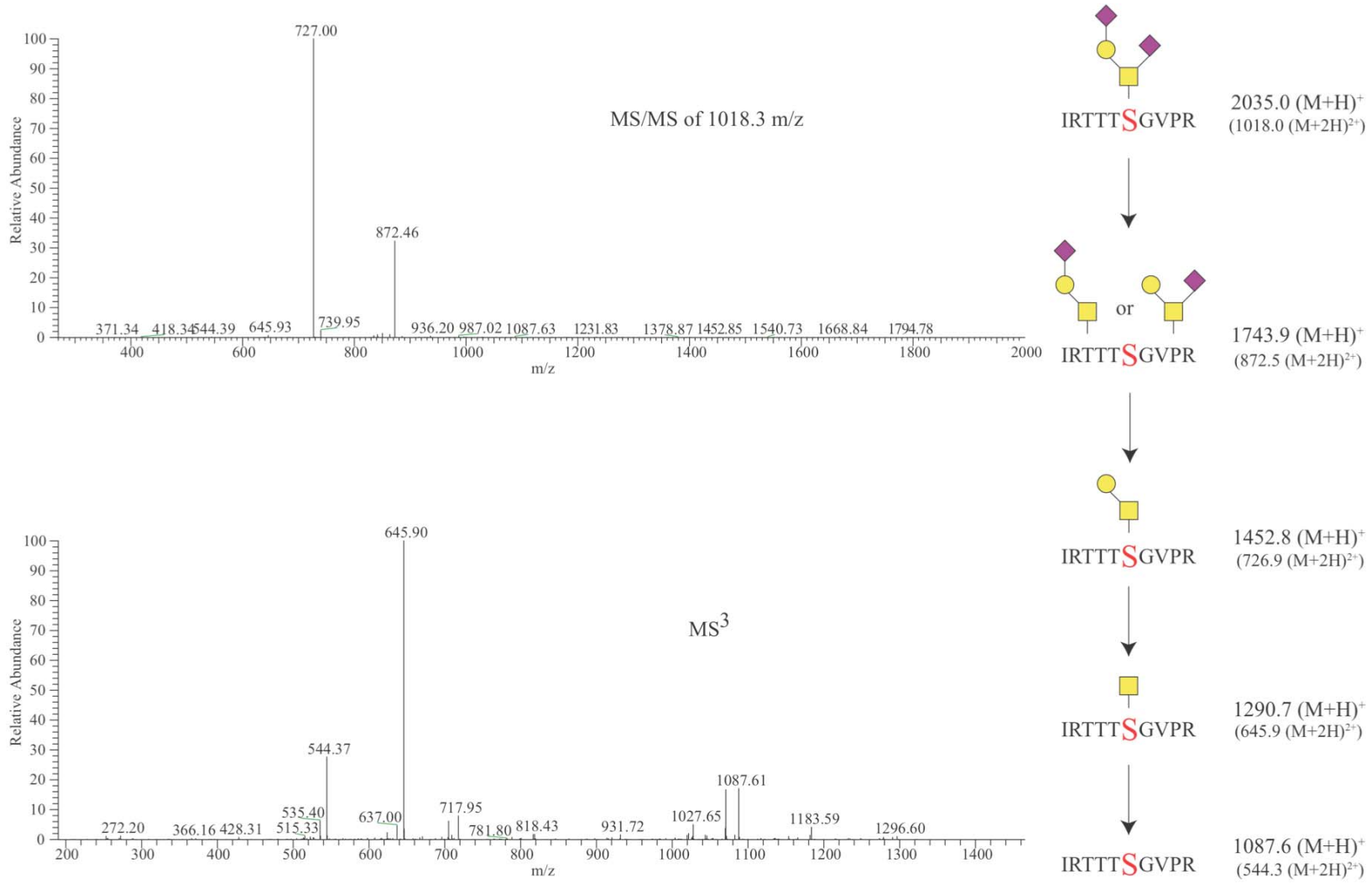

MS/MS @ 1100 m/z
Relative Abundance
m/z
100
90
80
70
60
50
40
30
20
10
0
200
400
600
800
1000
1200
1400
1600
1800
2000
376.26
472.25
520.40
618.36
724.43
815.51
847.40
993.50
1100.55
1236.68
1347.08
1457.15
1529.82
1801.54
1901.53

Figure 2-3. Assignment of an O-GalNAc α-DG glycopeptide.
From the full scan acquired at 49.33 minutes **(a)**, a peak at 1018.3 m/z was selected for fragmentation **(b)** The resulting MS/MS of 1018.3 m/z yielded the neutral loss of two terminal SA residues, which was then followed by MS^3 fragmentation indicating the loss of a Gal residue followed by a reducing end GalNAc. The combined glycan structure was determined to belong to a peptide with 1087.6 m/z. From examining the MS/MS spectra (not shown) and the neutral loss triggered MS^3 spectra **(c)**, the site of post-translational modification to the serine within the peptide IRTTTSVGPR is assigned.

a)


Full MS @ 49.33 min
Relative Abundance
m/z
100
90
80
70
60
50
40
30
20
10
0
400
600
800
1000
1200
1400
1600
1800
2000
371.18
447.00
491.00
543.55
659.27
670.00
686.09
688.00
736.00
801.00
917.55
1018.36
1060.36
1171.55
1355.27
1487.36
1620.91
1816.64
1990.36

b)


MS/MS of 1018.3 m/z
Relative Abundance
m/z
727.00
872.46
371.34
418.34
544.39
645.93
739.95
936.20
987.02
1087.63
1231.83
1378.87
1452.85
1540.73
1668.84
1794.78
IRTTTSGVPR
2035.0 (M+H)+
(1018.0 (M+2H)2+)
or
IRTTTSGVPR
1743.9 (M+H)+
(872.5 (M+2H)2+)
MS3
645.90
544.37
1087.61
272.20
366.16
428.31
515.33
535.40
637.00
717.95
781.80
818.43
931.72
1027.65
1183.59
1296.60
IRTTTSGVPR
1452.8 (M+H)+
(726.9 (M+2H)2+)
IRTTTSGVPR
1290.7 (M+H)+
(645.9 (M+2H)2+)
IRTTTSGVPR
1087.6 (M+H)+
(544.3 (M+2H)2+)

c)

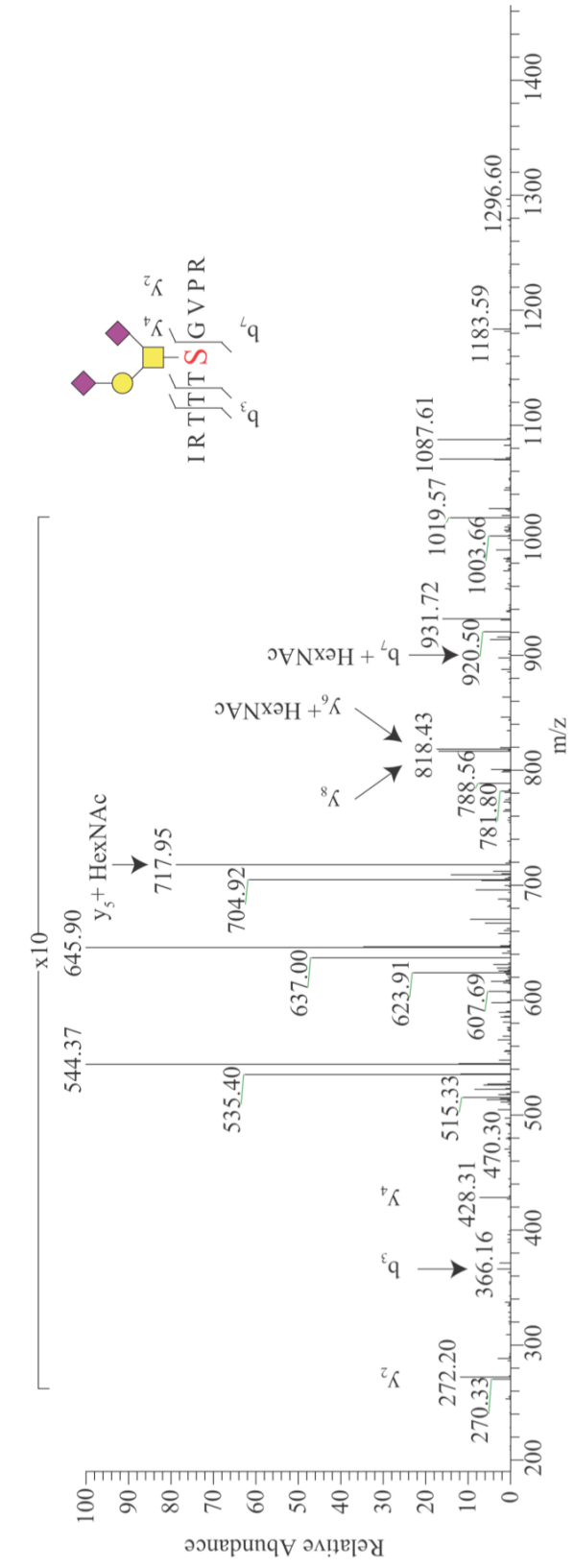
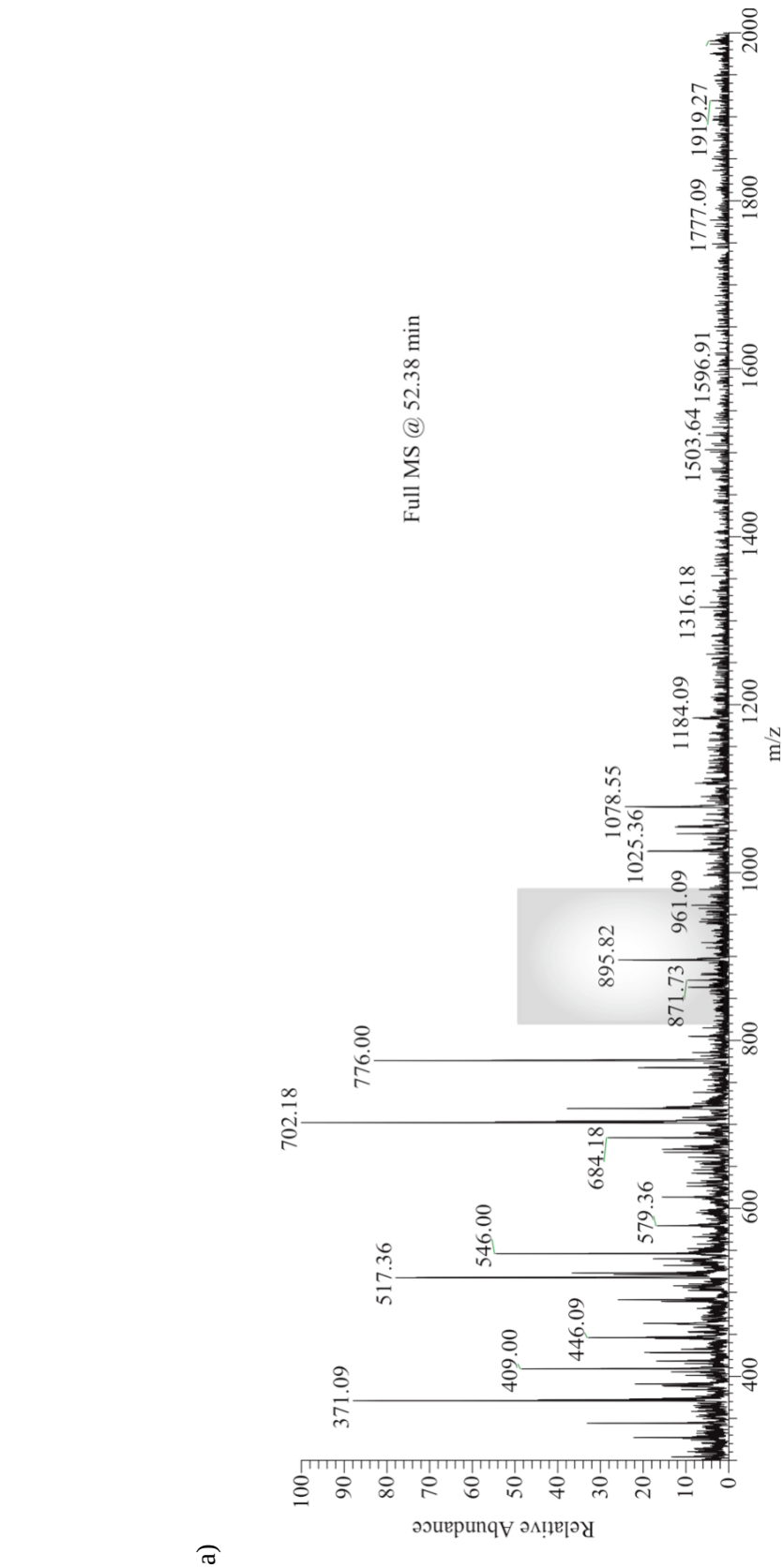
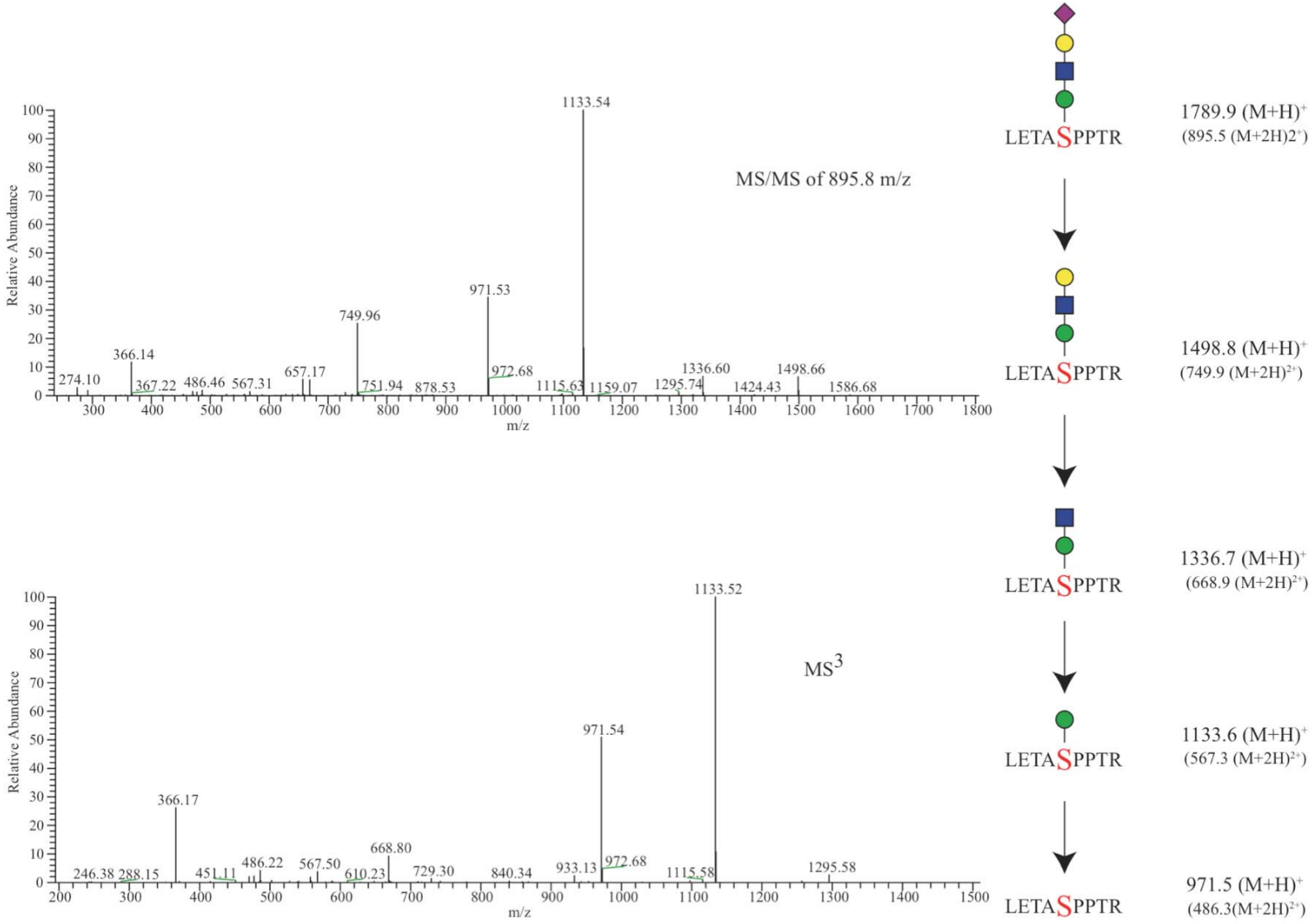

Relative Abundance
m/z
x10
544.37
645.90
y_5+ HexNAc
717.95
704.92
637.00
623.91
607.69
535.40
515.33
470.30
428.31
366.16
272.20
270.33
y_2
b_3
y_4
y_8
818.43
y_6+ HexNAc
788.56
781.80
931.72
b_7 + HexNAc
920.50
1003.66
1019.57
1087.61
1183.59
1296.60
I R T T T S G V P R
b_3
y_4
b_7
y_2

Figure 2-4. Assignment of an O-Man α-DG glycopeptide.
From the full scan acquired at 52.38 minutes **(a)**, a peak at 895.8 m/z was selected for fragmentation **(b)** The resulting MS/MS of 895.8 m/z yielded the neutral loss of terminal SA residue, which was then followed by MS3 fragmentation indicating the neutral loss of a Gal-GalNAc residue followed by a reducing end Man. The combined glycan structure was determined to belong to a peptide with 971.5 m/z. From examining the MS/MS spectra (data not shown) and the neutral loss triggered MS^3 spectra **(c)**, the site of posttranslational modification to the serine within the peptide LETASPPTR is assigned.

a)


Full MS @ 52.38 min
Relative Abundance
m/z
100
90
80
70
60
50
40
30
20
10
0
400
600
800
1000
1200
1400
1600
1800
2000
371.09
409.00
446.09
517.36
546.00
579.36
684.18
702.18
776.00
871.73
895.82
961.09
1025.36
1078.55
1184.09
1316.18
1503.64
1596.91
1777.09
1919.27

b)


MS/MS of 895.8 m/z
Relative Abundance
m/z
1133.54
971.53
749.96
366.14
274.10
367.22
486.46
567.31
657.17
751.94
878.53
972.68
1115.63
1159.07
1295.74
1336.60
1424.43
1498.66
1586.68
LETASPPTR
1789.9 (M+H)+
(895.5 (M+2H)2+)
LETASPPTR
1498.8 (M+H)+
(749.9 (M+2H)2+)
LETASPPTR
1336.7 (M+H)+
(668.9 (M+2H)2+)
MS3
1133.52
971.54
366.17
246.38
288.15
451.11
486.22
567.50
610.23
668.80
729.30
840.34
933.13
972.68
1115.58
1295.58
LETASPPTR
1133.6 (M+H)+
(567.3 (M+2H)2+)
LETASPPTR
971.5 (M+H)+
(486.3(M+2H)2+)

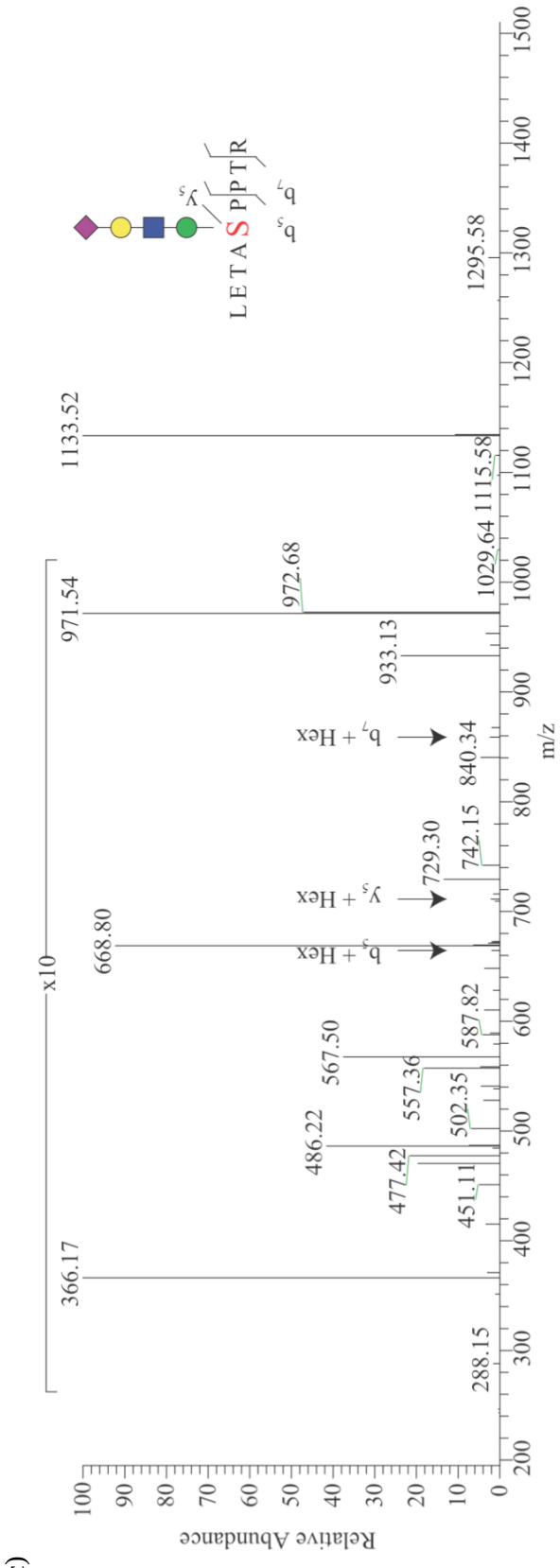

c)
Relative Abundance
m/z
x10
366.17
288.15
668.80
971.54
1133.52
1295.58
1115.58
1029.64
972.68
933.13
840.34
742.15
729.30
587.82
567.50
557.36
502.35
486.22
477.42
451.11
b5 + Hex
y5 + Hex
b7 + Hex
L E T A S P P T R
b5
b7
y5

Table 2-1. Quantification of O-Glycans released from α-DG purified from rabbit skeletal muscle.

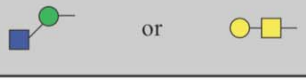
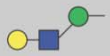
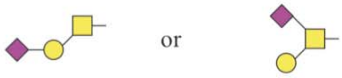
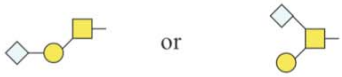
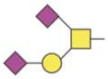
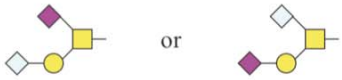
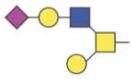
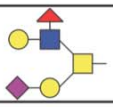
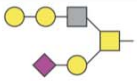
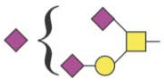
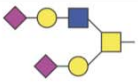
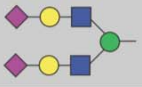
	Structure	Theoretical	Observed	Δ(The.-Obs.)	% total
1		534.7	534.5	0.2	9.5
2		575.7	575.5	0.3	1.9
3		738.9	738.6	0.4	19.7
4		779.9	779.4	0.5	0.1
5		896.0	895.6	0.5	34.1
6		926.0	925.6	0.4	1.7
7		937.1	937.6	-0.5	0.4
8		984.1	984.6	-0.5	< 0.1%
9		1100.3	1099.6	0.7	25.2
10		1141.3	1141.2	0.1	0.3
11		1257.4	1256.6	0.9	6.7
12		1287.4	1287.8	-0.4	0.2
13		1345.5	1345.7	-0.2	< 0.1%
14		1519.7	1519.6	0.1	< 0.1%
15		1549.7	1549.7	0.0	< 0.1%
16		1618.8	1618.2	0.6	< 0.1%
17		1706.9	1706.1	0.8	< 0.1%
18		1911.2	1911.5	-0.3	< 0.1%

Table 2-2. Summary of sites of modification and modifying residue(s).

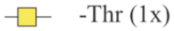
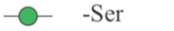
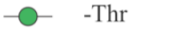
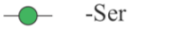
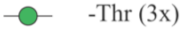
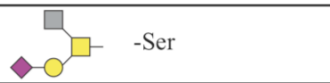
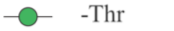
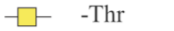
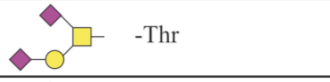
RESIDUE	STRUCTURE
T367/T369/T370/T372	
S391	
T395	
S398	
T404 / T406 / T414 / T418 / T421-4	
S430 #	
T446	
T450	
T455 #	
T463-4	
T469	
T473	
S475	
S475	
S475	
T478	
T478	
T483	
T484 #	
S485	
T529/T530/531	

Table 2-S1. Observed glycopeptides, glycans, and site(s) of modification.

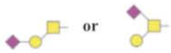
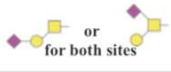
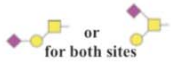
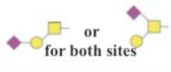
	Experimental Total Mass	AA Sequence	Location of Modification	MS[1] Glycans	Proposed Glycan Structure	Ions observed to assign glycosylation site	Obs. #b & y ions/# total ions	AA#	Scan #	Details of Sample Digestion
Full Structures	1317.9	(R)VSDAGTVVSGQIR(A)	2 sites total: S°391/T°395/S°398	SA-Hex- HexNAc- Hex, Hex-HexNAc-Hex			0:23	390 402	15891, 15898	Trypsin
	594.8	(R)T**PRPVPR(V)	T□455**	Hex-HexNAc		b2	10:11	455 461	574, 669, 684, 718, 732, 759, 767, 815,	Trypsin
	740.3	(R)T**PRPVPR(V)	T□455**	SA- Hex- HexNAc		b2	11:11	455 461	485, 493, 532, 539, 579, 587, 597, 603, 642, 649, 690, 697, 737, 744,785, 792, 834	Trypsin
	1150.7	(R)VTTKAPITR(L)	2 sites total: T□463/T□464/ T□469	SA-Hex-HexNAc, SA-Hex-HexNAc			2:15	462 470	15012, 15019	Trypsin
	859.6	(R)VTTKAPITR(L)	T□463/T□464, T□469	Hex-HexNAc, Hex-HexNAc		y5	3:15	462 470	561, 572, 583, 592, 622, 630, 675, 682, 724, 779, 799, 808, 853, 856, 899, 906, 917, 963	Trypsin
	1150.4	(R)VTTKAPITR(L)	T□463/T□464, T□469	SA-Hex-HexNAc, SA-Hex-HexNAc		y5	5:15	462 470	3493, 3500	Trypsin
	1004.7	(R)VTTKAPITR(L)	2 sites total: T□463/T□464/ T□469	SA-Hex-HexNAc, Hex-HexNAc			2:15	462 470	14491, 14500, 14604	Trypsin
	669.2	(R)LETASPPTR(I)	T□473	Hex-HexNAc		b3, y7	8:15	471 479	15050, 15063	Trypsin
	997.6	(R)LETASPPTR(I)	T□473, S□ 475	SA-Hex-HexNAc, Hex-HexNAc		b5, y5	6:15	471 479	15396	Trypsin
	1017.6	(R)LETASPPTR(I)	T□473, S□475, T□478	HexNAc[2], SA-Hex-HexNAc			2:15	471 479	15446, 15454	Trypsin
	1011.1	(R)LETASPPTR(I)	S□475	SA-HexNAc-Hex-HexNAc		b5, y5	3:15	471 479	15475	Trypsin
	1061.7	(R)LETASPPTR(I)	2 sites total: T□473/S□475/ T□478	SA-HexNAc-Hex, SA-HexNAc			1:15	471 479	15500, 15507	Trypsin
	791.1	(R)LETASPPTR(I)	T□473, S□ 475, T□478	HexNAc[3]		b4, y5	9:15	471 479	3626	Trypsin
	852.1	(R)LETASPPTR(I)	T□473, S□ 475	Hex-HexNAc, Hex-HexNAc		b3, y5	9:15	471 479	3497, 3502	Trypsin
	895.6	(R)LETASPPTR(I)	S°475	SA-Hex-HexNAc-Hex		b5	7:15	471 479	4058	Trypsin
	915.9	(R)LETASPPTR(I)	T□473, S□ 475	SA-Hex-HexNAc, HexNAc		b3, y5	5:15	471 479	3959, 3974	Trypsin
	1018.0	(R)LETASPPTR(I)	T□473, S□ 475, T□478	SA-Hex-HexNAc, HexNAc, HexNAc		b3, y5	11:15	471 479	3952, 3957	Trypsin
	1142.7	(R)LETASPPTR(I)	T□473, S□ 475	SA-Hex-HexNAc, SA-Hex-HexNAc		b3, y5	8:15	471 479	4184, 4190, 4228, 4235	Trypsin
	770.6	(R)LETASPPTR(I)	T□ 473, T□ 478	SA-Hex-HexNAc, HexNAc		b2, y2, y4	11:15	471 479	3577, 3584	Trypsin

Table 2-S1. Continued.

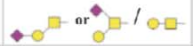
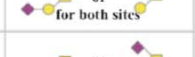
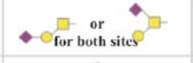
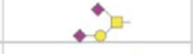
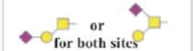
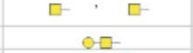
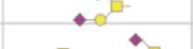
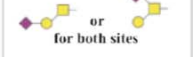
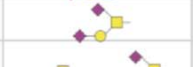
Full Structures	770.6	(R)LETASPPTR(I)	T□473, S□ 475/T□478	Hex-HexNAc, HexNAc		b3	4:15	471 479	10256, 10368, 10375	Lys-C &Trypsin
	851.6	(R)LETASPPTR(I)	2 sites total: T□473/S□475/ T□478	Hex-HexNAc, Hex-HexNAc		y6, y7	3:15	471 479	9716, 9724, 9825, 9831, 9936	Lys-C &Trypsin
	997.6	(R)LETASPPTR(I)	S□ 475, T□ 473/T□ 478	SA-Hex-HexNAc, Hex-HexNAc		b5	4:15	471 479	10671, 10678, 10786, 10889, 10895, 11112	Lys-C &Trypsin
	1018.0	(R)LETASPPTR(I)	T□473, S□ 475, T□478	HexNAc^2, SA-Hex-HexNAc			6:15	471 479	10824	Lys-C &Trypsin
	1993.6	(R)LETASPPTR(I)	2 sites total: T□473/S□475/ T□478	SA-Hex-HexNAc, Hex-HexNAc			2:15	471 479	10743, 10761, 10763, 10791	Lys-C &Trypsin
	1288.7	(R)LETASPPTR(I)	2 sites total: T□473/S□475/ T□478	SA^2-Hex-HexNAc, SA-Hex-HexNAc			3:15	471 479	11082, 11097	Lys-C &Trypsin
	1142.8	(R)LETASPPTR(I)	2 sites total: T□473/S□475/ T□478	SA-Hex-HexNAc, SA-Hex-HexNAc			1:15	471 479	11307, 11327, 11437, 11445	Lys-C &Trypsin
	1200.7	(R)IRTTT**SGVPR(G)	T□484**, S□485	SA-Hex-HexNAc, SA-Hex-HexNAc		y5	3:17	480 489	15087, 15102, 15293, 15304	Trypsin
	1018.3	(R)IRTTT**SGVPR(G)	S□485	SA^2-Hex-HexNAc		y5	8:17	480 489	3827, 3835	Trypsin
	872.7	(R)IRTTT**SGVPR(G)	S□485	SA-Hex-HexNAc		y5	12:17	480 489	11072, 11093, 11099, 12942, 12956	Lys-C &Trypsin
	1018.2	(R)IRTTT**SGVPR(G)	S□485	SA^2-Hex-HexNAc		b6, y5	11:17	480 489	10516, 10523, 10618, 10625	Lys-C &Trypsin
	1200.8	(R)IRTTT**SGVPR(G)	T□484**, S□485	SA-Hex-HexNAc, SA-Hex-HexNAc		b7, y6		480 489	9673, 9691, 10448	Lys-C &Trypsin
	727.6	(R)IRTTT**SGVPR(G)	S□485	Hex-HexNAc		y5	10:17	480 489	11209, 11222	Lys-C &Trypsin
	747.8	(R)IRTTT**SGVPR(G)	T□484**, T□482/T□483/S□485	HexNAc^2			3:17	480 489	12526	Lys-C &Trypsin
	592.6	(R)TTT**SGVPR(G)	S□485	Hex-HexNAc		y5	6:13	482 489	351, 430, 443	Trypsin
	819.4	(R)TTT**SGVPR(G)	T*483	SA-Hex-HexNAc-Hex		b2	3:13	482 489	2967	Trypsin
	883.7	(R)TTT**SGVPR(G)	S□485	SA^2-Hex-HexNAc		b4	3:13	482 489	3410, 3415, 3470, 3479, 3487, 3531, 3546, 3552	Trypsin
	1066.2	(R)TTT**SGVPR(G)	T□484**, S□ 485	SA-Hex-HexNAc, SA-Hex-HexNAc		b3, y6	4:13	482 489	3456, 3466	Trypsin
	592.5	(R)TTT**SGVPR(G)	T□484**	Hex-HexNAc		y6	7:13	482 489	2481, 2581, 2589, 3027, 3062, 3070, 3164, 3171, 3361, 3369, 3471, 3479, 3577, 3689, 3696, 3797, 3819, 3937, 3952, 4675, 4904	Lys-C &Trypsin
	819.3	(R)TTT**SGVPR(G)	T*483	SA-Hex-HexNAc-Hex		b2	2:13	482 489	8148	Lys-C &Trypsin
	883.4	(R)TTT**SGVPR(G)	T□484**	SA^2-Hex-HexNAc		y6	4:13	482 489	9494, 9508, 9604	Lys-C &Trypsin
	1474.5	(R)TTT**SGVPR(G)	2 sites total: T□483/T□484**/S□ 485	SA-Hex-HexNAc, SA-Hex-HexNAc			1:13	482 489	8625, 8637	Lys-C &Trypsin

Table 2-S1. Continued.

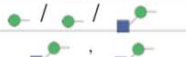
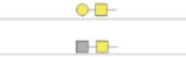
Trimmed Structures	864.7	(R)DPVPGKPTVTTRTR(G)	T□367/T□369/T□370/T□372	HexNAc		y7	5:25	360 373	21632	Glycosidases were added for 3 consecutive days; trypsin.
	945.6	(R)DPVPGKPTVTTRTR(G)	T□369/T□370/T□372	Hex-HexNAc		y5	4:25	360 373	13641	Glycosidases added O/N, trypsin
	1070.9	(R)VSDAGTVVSGQIR(A)	S°391, T°395, S°398	Hex, Hex, Hex-HexNAc-Hex		y5	8:24	390 402	11691, 11696	Glycosidases added O/N, trypsin
	1010.1	(R)VSDAGTVVSGQIR(A)	2 sites total: S°391/T°395/S°398	HexNAc-Hex, HexNAc-Hex			14:24	390 402	11789, 11796	Glycosidases added O/N, trypsin
	888.4, 807.0	(R)VSDAGTVVSGQIR(A)	S°391,T°395,S°398	Hex, Hex, Hex			16:23	390 402	12085	Glycosidases were added for 3 consecutive days; trypsin.
	807.0	(R)VSDAGTVVSGQIR(A)	S°391, T°395/S°398	Hex, Hex		b2	18:23	390 402	12305, 12314, 12421, 12427, 12931, 12937, 13041, 13261, 13265, 13987, 13993	Glycosidases were added for 3 consecutive days; trypsin.
	1070.6	(R)VSDAGTVVSGQIR(A)	S°391,T°395,S°398	Hex, Hex, Hex-HexNAc-Hex		y5	5:24	390 402	13091	Glycosidases were added for 3 consecutive days; trypsin.
	726.0	(R)VSDAGTVVSGQIR(A)	S°391	Hex		b3	17:23	390 402	13638, 13643, 13754, 14752, 14758	Glycosidases were added for 3 consecutive days; trypsin.
	807.1	(R)VSDAGTVVSGQIR(A)	S°391, T°395/S°398	Hex, Hex		b5	19:24	390 402	11095, 11104, 11204, 11211, 11591, 11597, 11631, 11653, 11667, 11763, 11881, 11979, 11998, 12009, 12497, 12503, 12600, 12608, 12612	Glycosidases added O/N, trypsin
	908.9	(R)VSDAGTVVSGQIR(A)	S°391/T°395/S°398	Hex, HexNAc-Hex			9:24	390 402	11969, 13164, 13270	Lys-C, Trypsin, Glycosidases
	909.0	(R)VSDAGTVVSGQIR(A)	S°391/T°395/S°398	Hex-HexNAc-Hex			9:24	390 402	12083	Lys-C, Trypsin, Glycosidases
	1091.6	(R)VSDAGTVVSGQIR(A)	S°391,T°395,S°398	Hex, Hex, HexNAc-Hex			12:24	390 402	12100	Lys-C, Trypsin, Glycosidases
	1010.8	(R)VSDAGTVVSGQIR(A)	S°391/T°395, S°398	HexNAc-Hex, HexNAc-Hex		y5	10:24	390 402	12315, 12431	Lys-C, Trypsin, Glycosidases
	1816.0	(R)VSDAGTVVSGQIR(A)	S°391/T°395, S°398	Hex, HexNAc-Hex		y7	2:24	390 402	12408, 12522, 12528, 12633, 12641	Lys-C, Trypsin, Glycosidases
	1401.8	(R)ATVTIPGYVEPTAVATPPTTTTK(K)	3 sites total: T°404/T°406/T°414/T°418/T°421/T°422/T°423/T°424	Hex, Hex, Hex		13:43	12:43	403 425	18548, 18649	Glycosidases added O/N, trypsin
	1318.3	(R)VS**TPKPATPSTDSSATTTR(R)	S□430**, T□446	Hex-HexNAc, Hex-HexNAc		y2, b2	20:35	429 447	11583, 11593	Glycosidases were added for 3 consecutive days; trypsin.
	706.4	(R)RPTKKPR(T)	T□450	Hex-HexNAc-HexNAc			1:15	428 454	9350, 9365	Lys-C, Trypsin, Glycosidases
	595.0	(R)TPRPVPR(V)	T□455**	Hex-HexNAc			1:11	455 461	1246, 1253, 1281	Lys-C, Trypsin, Glycosidases
	615.6	(R)TPRPVPR(V)	T□455**	HexNAc-HexNAc				455 461	1340, 1347	Lys-C, Trypsin, Glycosidases
	822.6	(R)VTTKAPITR(L)	T□463/T□464/T□469	SA-Hex-HexNAc			1:15	462 470	9915	Glycosidases were added for 3 consecutive days; trypsin.
	676.6	(R)VTTKAPITR(L)	T□463/T□464/T□469	Hex-HexNAc		b8	6:15	462 470	9011, 9817	Glycosidases were added for 3 consecutive days; trypsin.

Table 2-S1. Continued.

Trimmed Structures	859.9	(R)VTTKAPITR(L)	2 sites total: T□463/T□464/T□469	Hex-HexNAc, Hex-HexNAc			3:15	462 470	7479, 7576, 7582, 7678, 7685, 7787, 7989, 7996, 8100, 8201, 8630, 8685	Lys-C, Trypsin, Glycosidases
	1041.9	(R)VTTKAPITR(L)	T□463, T□464, T□469	Hex-HexNAc, Hex-HexNAc, Hex-HexNAc		y5	5:15	462 470	8703, 8711, 8905, 8910	Lys-C, Trypsin, Glycosidases
	462.2	(K)APITR(L)	T□469	Hex-HexNAc		y3	1:09	466 470	2756, 2765, 2864, 2873, 2977, 2985, 3089, 4728, 4753, 4775, 4782, 4887, 4896, 5000, 5100, 5106, 5209, 5215, 5313, 5320, 5424, 5432, 5533, 5540, 5641, 5648, 5750, 5757, 5862	Lys-C, Trypsin, Glycosidases
	462.2	(K)APITR(L)	T□469	Hex-HexNAc		y2	1:09	466 470	2640	Glycosidases added O/N, trypsin
	851.7	(R)LETASPPTR(I)	T□473, S□ 475	Hex-HexNAc, Hex-HexNAc		b4, y5	3:15	471 479	9611, 9897, 10062, 10139, 10147, 10275, 10280, 10383, 10393	Glycosidases were added for 3 consecutive days; trypsin.
	770.9	(R)LETASPPTR(I)	S□ 475/T□478	Hex-HexNAc-HexNAc		y5	2:15	471 479	9947	Glycosidases were added for 3 consecutive days; trypsin.
	770.9	(R)LETASPPTR(I)	T□473, S□ 475	Hex-HexNAc, HexNAc		b3, y5	5:15	471 479	9120, 9244	Glycosidases added O/N, trypsin
	851.7	(R)LETASPPTR(I)	T□473, S□ 475/T□478	Hex-HexNAc, Hex-HexNAc		b4, y7	4:15	471 479	7900, 7934, 7987, 8005, 8106, 8114, 8217, 8225, 8325, 8332, 8435, 8443, 8542, 8549, 8653, 8763, 8768, 8867, 8876, 8979, 8988, 9087, 9190, 9198, 9298, 9306, 9406, 9413, 9516, 9523, 9621, 9630	Glycosidases added O/N, trypsin
	851.7	(R)LETASPPTR(I)	S□ 475, T□478	Hex-HexNAc, Hex-HexNAc		b5, y4	4:15	471 479	7831	Glycosidases added O/N, trypsin
	669.0	(R)LETASPPTR(I)	T□473	Hex-HexNAc		b3	12:15	471 479	8525, 8535, 8633, 8640, 8741, 8748, 8849, 8961, 8966, 9065, 9072, 9169, 9176, 9274, 9283, 9386, 9393, 9493, 9502, 9603, 9612, 9712, 9719, 9823, 9831, 9933, 9940, 10045, 10053	Glycosidases added O/N, trypsin
	568.0	(R)LETASPPTR(I)	T*473/S*475	Hex		b7	14:15	471 479	10131, 10238, 10345, 10354, 11045, 11625, 11638	Glycosidases were added for 3 consecutive days; trypsin.
	588.1	(R)LETASPPTR(I)	T□478	HexNAc		y2	10:15	471 479	10839, 11012	Glycosidases were added for 3 consecutive days; trypsin.
	669.1	(R)LETASPPTR(I)	T□473	Hex-HexNAc		b3, y7	12:15	471 479	9358, 9378, 9385, 9495, 9502, 9614, 9621, 9848, 9855, 9963, 9970, 10079, 10088, 10195, 10202, 10308, 10315, 10436, 10440, 10552, 10561, 10672, 10679, 10788, 10795, 10909, 10916	Glycosidases were added for 3 consecutive days; trypsin.
	851.6	(R)LETASPPTR(I)	2 sites total: T□473/S□475/ T□478	Hex-HexNAc, Hex-HexNAc			1:15	471 479	9922, 9927, 10036, 10143, 10252	Lys-C, Trypsin, Glycosidases
	669.4	(R)LETASPPTR(I)	S□475	Hex-HexNAc		b5, y5	8:15	471 479	10030, 10046, 10314, 10534, 10541, 10646, 10654	Lys-C, Trypsin, Glycosidases
	872.6	(R)LETASPPTR(I)	T□473, S□475, T□478	Hex-HexNAc, HexNAc, HexNAc		b3	3:15	471 479	10189, 10200	Lys-C, Trypsin, Glycosidases
	770.9	(R)LETASPPTR(I)	S□475	Hex-HexNAc-HexNAc		y5	6:15	471 479	10282, 10294, 10395, 10402, 10422, 10430	Lys-C, Trypsin, Glycosidases
	750.4	(R)LETASPPTR(I)	S□475, T*478	Hex-HexNAc, Hex		b5, y4, y6	8:15	471 479	10369, 10385	Lys-C, Trypsin, Glycosidases
	791.6	(R)LETASPPTR(I)	T□473, S□475, T□478	HexNAc, HexNAc, HexNAc			5:15	471 479	10476	Lys-C, Trypsin, Glycosidases

Table 2-S1. Continued.

Trimmed Structures	727.3	(R)IRTTT**SGVPR(G)	T□484**	Hex-HexNAc			11:17	480 489	9234, 9241, 12357, 12359	Glycosidases added O/N, trypsin
	727.3	(R)IRTTT**SGVPR(G)	S□485	Hex-HexNAc		y5	15:17	480 489	9876, 9882, 9988, 9993	Glycosidases were added for 3 consecutive days; trypsin.
	909.9	(R)IRTTTSGVPR(G)	2 sites total: T□482/T□483/T□484**, S□485	Hex-HexNAc, Hex-HexNAc		y5	4:15	480 489	8774, 8782	Lys-C, Trypsin, Glycosidases
	775.3	(R)TTTSGVPR(G)	2 sites total:	Hex-HexNAc, Hex-HexNAc			2:15	482 489	2924, 2929, 3038, 3142, 3149, 3253, 3365, 3371, 3768, 3783, 3883, 3889, 3992, 4000, 4104, 4115, 4211, 4219, 4358, 5485, 5526, 5588, 5594, 5693, 5699, 5803, 5812, 5913, 6022, 6029, 6132, 6139, 6240, 6248, 6347, 6356, 6458, 6464, 6566, 6574, 7353, 7471, 8041, 8322	Lys-C, Trypsin, Glycosidases
	673.6	(R)TTTSGVPR(G)	T□484**, T*482/T*483/S*485	Hex-HexNAc, Hex				482 489	3647, 4612, 4660, 4668, 4772, 4885, 4893, 4994, 5002	Lys-C, Trypsin, Glycosidases
	775.3, 613.1	(R)TTTSGVPR(G)	2 sites total: T□482/T□483/S□485, T□484**	Hex-HexNAc, Hex-HexNAc		b3	2:15	482 489	4326, 6686, 7306, 7500, 7507, 7603, 7611, 7796, 7801, 7915, 8314	Lys-C, Trypsin, Glycosidases
	592.8	(R)TTTSGVPR(G)	T□484**	Hex-HexNAc		y6	5:15	482 489	4371, 4389, 4493, 4499, 5874, 5884, 5983, 6094, 6110, 6211, 6237, 6263, 6269, 6363, 6368, 6375, 6476, 6483, 6585, 6592, 6618, 6691, 6698, 6798, 6805, 6902, 6909, 7004, 7010, 7017, 7118, 7124, 7222, 7229, 7233, 7328, 7335, 7345, 7437, 7444, 7467, 7542, 7549, 7647, 7653, 7750, 7756, 7857, 7865, 7969, 8074, 8082, 8182, 8239, 8245, 8301, 8340, 8348, 8452, 8460, 8601	Lys-C, Trypsin, Glycosidases
	1183.8 (1+)	(R)TTTSGVPR(G)	T□484**	Hex-HexNAc		y6	2:15	482 489	6930, 7031, 7138, 7146, 7251, 7258	Lys-C, Trypsin, Glycosidases
	612.9	(R)TTTSGVPR(G)	2 sites total: T□482/T□483/S□485, T□484**	HexNAc, HexNAc		y6		482 489	8722	Lys-C, Trypsin, Glycosidases
	708.7	(K)EDTTTDKLK(L)	T□529/T□530/T□531	Hex-HexNAc		y7	3:17	527 535	9531, 9590	Glycosidases added O/N, trypsin

Table 2-S2. Alignment of O-Man Sites Mapped for Determination of Consensus Sequence.

CHAPTER 3

GLYCOMIC ANALYSIS OF MOUSE MODELS OF CONGENITAL MUSCULAR DYSTROPHY

Hammond S, Aoki K, Lim JM, Porterfield M, Liu M, Buskirk S, Campbell KP, Hu H, Live D, Tiemeyer M, Wells L.
To be submitted to *Journal of Biological Chemistry.*

ABSTRACT

Alpha-dystroglycan (α-DG) is an extensively *O*-glycosylated extracellular matrix binding protein and the main component of the dystrophin-glycoprotein complex (DGC). Previous studies have shown α-DG to be post-translationally modified by both *O*-GalNAc and *O*-mannose initiated glycan structures. Mutations in several genes leading to hypoglycosylation of α-DG, which is associated with a loss of ligand binding activity, are causal for various forms of Congenital Muscular Dystrophy (CMD). In this study we sought to perform glycomic analysis on *O*-linked glycan structures released from three different mouse brain models of CMD (POMGnT1, LARGE, and α-DG -/-). Using mass spectrometry approaches we were able to identify 9 *O*-mannose initiated and 25 *O*-GlaNAc initiated glycan structures in the wild-type littermate control mouse brains. Through our analysis, we were able to confirm that POMGnT1 is essential for the extension of all observed *O*-mannose glycan structures with β-1,2 linked GlcNAc. Loss of LARGE had no observable effect on the *O*-mannose initiated glycan structures characterized here. Interestingly, we also determined that similar amounts of *O*-mannose initiated glycan structures are present on mouse brain proteins from the α-DG -/- animals compared to wild type, indicating that there must be additional proteins that are *O*-mannosylated in the mammalian brain. Our findings illustrate that elongation and β-1,6 GlcNAc branching of *O*-mannose glycan structures is dependent upon the POMGnT1 enzyme, and that *O*-mannosylation is not solely limited to α-DG in the brain.

INTRODUCTION

Congenital muscular dystrophy (CMD) is a heterogeneous group of inherited neuromuscular disorders characterized by severe muscle weakness, ocular and neuronal migration abnormalities, and variable mental retardation [1]. Within recent years, it has become increasing clear through genetic studies that hypoglycosylation of the protein dystroglycan (DG) is a commonality between various phenotypes of CMD. DG is post-translationally cleaved into an extracellular α-DG subunit and a transmembrane β-DG subunit [2]. α-DG is the central component of the dystrophin-glycoprotein complex (DGC), and serves as a link between the cytoskeleton of the muscle cell and the extracellular matrix by binding to proteins such

as laminin[3]. Interaction between α-DG and its extracellular ligands require α-DG to be properly post-translationally modified through the addition of *O*-linked oligosaccharides, specifically *O*-mannose [4, 5]. To date, mutations in six genes that encode determined or predicted glycosyltransferases have been shown to result in varying forms of CMD in which the post-translational processing of α-DG is affected[4-6]. The six mutated genes and their original resulting form of CMD are: *protein O-mannosyltransferase 1*(*POMT1*) and *2 (POMT2*)-Walker-Warburg Syndrome (WWS) [7, 8], *protein O-linked mannose β1,2 N-acetylglucosaminyltransferase* (*POMGnT1*)- Muscle–Eye-Brain disease (MEB) [9], *Fukutin*-Fukuyama congenital muscular dystrophy (FCMD) [10], *Fukutin-related protein* (*FKRP*)- congenital muscular dystrophy 1C (MDC1C) [11], and *LARGE*- congenital muscular dystrophy (MDC1D) [12]. Recent work has demonstrated that selected mutations in some of these genes can cause various forms of CMD likely dependent on the severity of the mutation on enzymatic activity[13]. Abnormal glycosylation of α-DG appears to be a commonality among all of the aforementioned forms of CMD. While expression of α-DG appears to not be grossly affected, the ability of α-DG to be recognized by monoclonal antibodies IIH6 and $VIA4_1$ is eliminated, as well as the ability of α-DG to properly bind its ligands [14].

α–DG is comprised of a central mucin-like region which is extensively heterogeneously glycosylated with glycan chains that are initiated by both *O*-GalNAc and *O*-mannose [15-17]. POMT1 and 2 and POMGnT1 have been shown to catalyze the first two steps involved in the synthesis of a sialyated *O*-mannosyl tetrasaccharide structure [9, 18]. The previously identified glycan structure, Neu5Ac(α2–3)Gal(β1–4)GlcNAc(β1–2)Man, was identified on α-DG isolated from both brain and muscle and was previously thought to be required for ligand binding activity. However, more recent studies have suggested that laminin binding is most likely not dependent on this structure, as treating α-DG with a series of glycosidases resulting in the trimming of this *O*-mannosyl tetrasaccharide structure does not abolish laminin binding activity, but rather enhances the ability of α-DG to bind to laminin [19]. Additionally, recent work by Yoshida-Moriguchi *et. al.*, has shown that LARGE, a putative glycosyltransferase, is involved in the synthesis of a rare extended phosphorylated *O*-mannosyl trisaccharide core structure on α-

DG that is required for laminin binding [20]. The roles that the other putative glycosyltransferases, fukutin and FKRP, have in the post-translational processing of α-DG have yet to be determined.

Given the importance of post-translational processing of α-DG as related to CMD and the abundance of *O*-mannosyl glycan structures in the brain (~30% of *O*-glycan structures are *O*-mannose initiated [21]), we sought to identify and characterize the relative abundance of *O*-glycan structures released from the mouse brain proteins of CMD disease models and their comparative wild type controls. To do this we implemented strategies that had been previously developed and described [22]. Glycomic analysis of the various mouse models of CMD consisted of chemically releasing all *O*-glycans by b-elimination, which was followed by permethylation, and analysis by mass spectrometry [23]. From the implemented workflow we were able to fully characterize *O*-glycans released from mouse brain proteins of selected mouse models of CMD and compare their relative abundances between the control and disease states. This study highlights the relative abundance of *O*-mannose and *O*-GalNAc initiated glycan structures characterized from the three CMD mouse brain models studied as well as the role of crucial enzymes and substrates in the *O*-mannosylation process.

EXPERIMENTAL PROCEDURES

Preparation of whole animal and tissue specific knockout mouse models of CMD

The whole body and tissue specific knockout mouse models of congenital muscular dystrophy were prepared as reported previously [24-26].

Preparation of Brain Protein Powder

Protein powder was prepared from various mouse models of congenital muscular dystrophy as described previously [23]. Briefly, to prepare protein powder from mouse brains, frozen brains were homogenized and delipidated in a solvent mixture with a final ratio of chloroform/methanol/water equal to that of 4:8:3. The extracted material was allowed to incubate for 6 hours at room temperature. The precipitated protein material was collected by centrifugation, and the resulting protein pellet was re-extracted with fresh solvent. The precipitated protein pellet was washed with ice cold 20% acetone and then dried under a gentle nitrogen stream at 45ºC.

Release of O-linked Glycans

O-glycans were released using reductive β-elimination as reported previously [16]. Briefly, protein powder was resuspended with 1M $NaBH_4$ in 50mM NaOH and incubated at 45°C for 18 hrs. Following incubation, the reaction mixture was neutralized by adding 10% acetic acid dropwise while vortexing. The completely neutralized samples were desalted by loading onto an AG-50W-X8 (Bio-Rad) column (1mL bed volume). The released oligosaccharides were eluted from the column with 3 volumes of 5% acetic acid and dried to dryness using a SpeedVac. Borate was removed by adding 10% acetic acid in methanol and drying under a nitrogen stream at 45°C.

Permethylation and Analysis of O-glycans by Nanospray Ionization Mass Spectrometry (NSI-MS^n)

Released oligosaccharides were permethylated to facilitate their analysis by mass spectrometry as reported previously [27]. For analysis by NSI-MS^n, permethylated glycans were dissolved in 1mM NaOH in 50% methanol and directly infused into a linear ion trap mass spectrometer (LTQ, Thermo Scientific) at a flow rate of 0.4 mL/min. As described previously, the total ion mapping (TIM) function of the Xcallibur software package was used to detect and quantify the prevalence of all detected glycan structures [22]. Glycan nomenclature and representation of all glycan structures shown in figures and tables is in agreement with the guidelines proposed by the consortium of functional glycomics with unidentified HexNAc residues displayed in grey.

Characterization of O-mannitol by High-Pressure Liquid Chromatography

As described previously [23], monosaccharides were released from both POMGnT1 +/+ and -/- mouse brain proteins and the abundance of *O*-mannitol was analyzed as described previously [23]. Briefly, *O*-glycans were released from protein powder prepared from both POMGnT1 +/+ and -/- brains by reductive β-elimination, which were separated on a Dionex CarboPac MA-1 column (250 X 4 mm) and detected by HPAEC-PAD. The detected monosaccharide peaks were later quantified based on monosaccharide alditol standards.

Western Blotting

Mouse brain proteins from all three phenotypes were separated by SDS-PAGE on a 7.5% Tris-HCl precast gel purchased from Bio-Rad Laboratories. Following separation proteins were transferred to PVDF membranes using semi-dry transfer and probed using monoclonal antibody IIH6 (1:2000 dilution) (Millipore) and ERK-2 (C-14) (1:5000 dilution) (Santa Cruz). Proteins were detected using SuperSignal West Pico chemiluminescent substrate (Thermo Scientific).

NMR Studies

A sample of < 1mg of Ac-YVEP(GlcNAc-β-1,2-Man-α-)TAV-NH_2 was dissolved in 90μL of D_2O in a 3mm Shigemi NMR microcell (Shigemi, Inc., Allison Park, PA). NMR spectra were obtained on a Varian NMRS 600 MHz NMR spectrometer with a 3mm HCN coldprobe. ^{1}H DQCOSY, ^{1}H TOCSY, ^{1}H-^{13}C gradient-HMQC and ^{1}H-^{13}C gradient-HMBC data were recorded using the standard pulse sequences in the VnmrJ software and were processed on the spectrometer. From these data, all resonances were assigned for the disaccharide glycan, and connectivities across the glycosidic linkages established with the HMBC experiment.

Synthesis of Ac-Tyr-Val-Glu-Pro-Thr(α-D-Man)-Ala-Val-NH_2

Peptide synthesis and preparation of Fmoc-Thr(Ac_4-α-D-Man)-OH generally follows methods reported previously. Peptide chain assembly starting with Fmoc-PAL-PEG-PS resin (0.3 g, 0.18 mmol/g) (Applied Biosystems, Foster City, CA) was carried out manually. A CEM Discover microwave reactor (CEM Corporation, Mathews, NC) was used for accelerated deprotection and coupling reactions at elevated temperature. Side-chain protection employed was ^{t}Bu for Tyr and O^{t}Bu for Glu. All reactions were carried out in a plastic vessel (25 ml) with a porous polypropylene frit, and no stirring was used. Fmoc removal was achieved with piperidine-DMF (1:4) for 4 min at 75^{o}C under 30 W microwave irradiation, couplings (2.5 equiv of Fmoc-amino acids) were mediated by HCTU (2.5 equiv)-HOBt (2.5 equiv)-DIEA (5 equiv) for 9.5 min at 75^{o}C under 25 W microwave irradiation, and *N*-acetyl capping was carried out with acetic anhydride-DMF (1:4) for 4 min at 75^{o}C under 30 W microwave irradiation. The coupling of Fmoc-Thr (Ac_4-α-D-Man)-OH (1.5 equiv) was mediated by HCTU (1.5 equiv) -HOBt (1.5

equiv) -DIEA (3 equiv) for 9.5 min at 75°C under 25 W microwave irradiation. Double couplings of Fmoc-Thr (Ac_4-α-D-Man)-OH (0.5 equiv), Fmoc-Pro-OH (2.5 equiv) and Fmoc-Glu (O^{t}Bu)-OH (2.5 equiv) were carried out under the same microwave coupling conditions, and with corresponding amount of HCTU, HOBt and DIEA. Approximately half of the peptide resin was treated with TFA-H_2O (19:1, 5 ml) in a glass flask (25 ml) with stirring for 2 h at 25°C to cleave the Ac-YVEPT(Ac_4-α-D-Man)AV-NH_2 from the resin, and concurrently remove ^{t}Bu and O^{t}Bu side-chain protecting groups. The cleaved peptide-resin mixture was then washed extensively with TFA (3×1 ml), and the *O*-acetylated glycopeptide was precipitated by adding cold anhydrous ether (100 ml) to the combined filtrates. The ether-treated material was kept overnight at 4°C, and was then collected by centrifugation and washed further with cold ether (2×10 ml). The *O*-acetylated glycopeptide was purified by semi-preparative RP-HPLC, and the fractions with the desired products were pooled and lyophilized. The lyophilized *O*-acetylated glycopeptide was then dissolved in methanol (~4 ml/mg) for removal of *O*-acetyl groups. The pH of the methanol solution was adjusted to ~9 (as detected by wet litmus paper) by adding a solution of NaOMe in methanol. The *O*-deacetylation reaction was carried out with mild stirring, and monitored by analytical RP-HPLC, going to completion after 5 h. Powdered CO_2 (dry ice) was then added carefully to reach a pH of 6, and the pure glycopeptide was obtained after semi-preparative RP-HPLC followed by lyophilization of the appropriate fractions. Amount: 15 mg. Overall yield based on initial substitution of resin: 57%. ESI-MS: 981.4 $[M+H]^+$, 1003.4 $[M+Na]^+$.

Enzymatic Synthesis of Ac-YVEP(GlcNAc-β-1,2-Man-α-)TAV-NH_2

Ac-YVEP(GlcNAc-β-1,2-Man-α-)TAV-NH_2 was prepared from Ac-YVEP(Man-α)TAV-NH_2 using the POMGnT1 enzyme. The DNA sequence for human POMGnT1 (provided by Dr. Huaiyu Hu, State University of New York) was cloned into a pSecTAG2B vector and stably transfected into HEK-293 cells. For POMGnT1 purification, culture media was collected from the HEK-POMGnT1 cells and combined with a mixture of protease inhibitors. The media were then chromatographed through a Nickel chelating column and the active glycosyltransferase was partially purified by elution with 300 mM

imidazole at pH of 6.8. To avoid aggregation and stabilize the recombinant enzyme in solution, bovine serum albumin was usually added and the resulting mixture was dialyzed (to remove the imidazole), and then concentrated. To assay the enzyme activity, the incorporation of radiolabeled GlcNAc from UDP-[^{3}H]-GlcNAc into Manα1-*O*-benzyl was measured over time and the reaction product was purified by solid phase extraction on a SepPack C-18 cartridge. In a typical purification procedure, 2 to 5 mL of a solution with partially pure enzyme could be obtained with an activity of 190 milliunits/mL. 0.8 mg of Ac-YVEP(Man-α-)TAV-NH_2 was treated with recombinant POMGnT1 and UDP-GlcNAc and the process driven to completion over 72 hours using buffer conditions as previously described in the presence of alkaline phosphatase to remove to prevent the accumulation of the glycosyltransferase inhibitor UDP. The monosaccharide and disaccharide substituted peptides were separated by analytical RP-HPLC on a C_{18} column (4.6×250 mm), with detection at 220 nm, and elution with 0.1% aq. TFA (buffer A)-0.1% TFA in CH_3CN (buffer B), linear gradient from 0 to 10% buffer B, at a flow rate of 1 ml/min for 40 min.

RESULTS

POMGnT1 is Required for the Extension of O-mannose by β-1,2 Linked GlcNAc

O-linked glycans were released from both POMGnT1+/+ and -/- mouse brain proteins by β-elimination and then permethylated to facilitate their analysis by NSI-MS/MS. The acquired Full MS scans (*m/z* 700-1500) (Fig. 3-1a and b) allowed for the detection of *O*-glycan structures released from both POMGnT1 +/+ and -/- mouse brain proteins. The structural assignment of *O*-glycans observed in the full MS scans was based upon subsequent MS/MS fragmentation data. In order to detect all of the released *O*-glycan structures, we used total ion mapping (TIM) methodologies as previously described [22]. By using TIM, we were able to acquire MS/MS fragmentation profiles allowing for the detection of *O*-glycans over the entire scanned *m/z* range. In Figure 3-1c and d, we show the MS/MS fragmentation profile of disialyated T antigen (*m/z* 1256) detected in both the POMGnT1 +/+ and -/-. Table 3-1 contains a list of all glycan structures detected in both the POMGnT1 +/+ and -/-. Fragmentation of extended *O*-

mannose structures was not observed in the POMGnT1 -/- (the glycan prevalence is indicated as 'ND' (not detected) Table 3-1 for all extended O-Man structures). Additionally, mouse brain proteins from both POMGnT1 +/+ and -/- were separated by SDS-PAGE and probed with monoclonal antibody IIH6 (Fig. 3-1e), which detects fully glycosylated, functionally active α-DG. From the Western blot we observed a loss of reactivity of IIH6 with proteins from the POMGnT1 -/-, indicating that α-DG is hypoglycosylated and no longer functionally active. Through HPAEC-PAD analysis of β-eliminated, reduced monosaccharide alditols from both POMGnT1 +/+ and -/- mouse brain proteins, Man-ol was shown to be elevated ~2.4 fold in the POMGnT1 -/- compared to the POMGnT1 +/+ littermate control suggesting that POMT1/2 activity was unaffected.

The assignments of the resonance positions for all NMR ^{1}H and ^{13}C signals arising from the GlcNAc-Man disaccharide glycan on the glycopeptide Ac-YVEP(GlcNAc-β-1,2-Man-α-)TAV-NH_2 (residues 410-416 of αDG), prepared by recombinant POMGnT1 enzymatic addition of GlcNAc to the Man glycopeptide, were made using homonuclear two-dimensional ^{1}H DQ-COSY, and TOCSY experiments and heteronuclear ^{1}H-^{13}C HMQC and HMBC two-dimensional correlations. The latter experiment shows correlations between ^{1}H and ^{13}C nuclei with two or three intervening bonds,[28] allowing explicit connections to be established across the glycosidic linkages. The initial identification of the Man-α and the GlcNAc-β anomeric sites could be readily made based on the distinct ^{1}H and ^{13}C carbon shift positions associated with their respective stereochemistry[29, 30]. Further, the $^{3}J_{H1-H2}$ couplings for the anomeric protons were as predicted for the linkage stereochemistry, with a large coupling ~ 8.5Hz in the ^{1}H 1-dimensional spectrum clearly showing the GlcNAc linkage to be β[30]. The location of the Man anomeric signals were also in agreement with that observed for the synthetic substrate glycopeptide Ac-YVEP(Man-α)TAV-NH_2 (data not shown). Using correlations that could be traced back to the assigned anomeric resonances, shifts for nuclei at the additional sites in each residue were determined with information derived from the experiments above (Supplemental Figure 3-1). Based on these assignments, HMBC cross peaks arising from ^{1}H and ^{13}C connectivities in both directions across the glycosidic

linkages between the sugar residues, and from the Man to the Thr β (Fig. 3-2) were identified. These results unequivocally establish the formation of a β1-2 linkage between the GlcNAc and the Man residues mediated by POMGnT1.

O-glycans released and detected from both LARGE +/+ vs. -/- are identical

O-linked glycans were released from LARGE +/+ and -/- mouse brain proteins as described previously. Full MS scans (*m/z* 700-1500) (Fig. 3-3a and b) of the released *O*-glycans indicate that the most abundant glycan structures found in both the LARGE +/+ and -/- samples to be similar. Structural analysis of the released *O*-glycans was based upon the acquired MS/MS fragmentation data. In Figure 3-3c and d, we show that synthesis of the *O*-mannosyl tetrasaccharide structure remains unaffected in the LARGE -/- based upon structural assignment of characteristic fragments. Prevalence of all the characterized glycan structures were quantified by using TIM and no statistically relevant changes were observed (Table 3-2). Mouse brain proteins from both LARGE +/+ and -/- were separated by SDS-PAGE and probed with monoclonal antibody IIH6 (Fig. 3-3e). From the Western blot we observed a loss of reactivity of IIH6 with proteins from the LARGE -/-, indicating that α-DG is hypoglycosylated and no longer functionally active.

O-mannose initiated glycan structures are still detectable in the α-DG -/- brain

As previously described O-glycans were released from α-DG +/+ and -/- mouse brain proteins. Full scans were acquired (*m/z* 700-1500) of released *O*-glycans from both α-DG +/+ and -/- mouse brain proteins with the most abundant glycan structures shown (Fig. 3-4a and b). Through the MS/MS fragmentation spectra we were able to confirm the presence of a fucosylated *O*-mannose trisaccharide isolated from the mouse brain proteins of both the α-DG +/+ and -/-. Mouse brain proteins from both LARGE +/+ and -/- were separated by SDS-PAGE and probed with monoclonal antibody IIH6 (Fig. 3-3e). From the Western blot we observed a loss of reactivity of IIH6 with proteins from the α-DG -/-, indicating the absence of functionally active α-DG

DISCUSSION

Hypoglycosylation of α-DG through genetic mutations, has been shown to result in various phenotypes of CMD in which the post-translational processing of α-DG is affected, preventing binding to proteins of the ECM such as laminin, perlecan, and agrin [31]. To date, quantitative comparative studies of O-glycans released from models of CMD have yet to be completed. Thus we sought to characterize and quantify the structure and prevalence of the various *O*-glycans released from mouse brain proteins of three models of CMD.

O-glycans were chemically released from the mouse brain proteins of POMGnT1 -/-, LARGE-/-, and α-DG -/- brains as well as their comparative wild type littermates by β-elimination. We started by comparing the full MS scans of *O*-glycans released from knockout mouse brain to that of their wild type littermates (Fig. 3-1a, 3-1b, 3-3a, 3-3b, 3-4a, 3-4b). Previous studies have shown that mutations in *POMGnT1* prevent the extension of *O*-mannosyl glycans with GlcNAc residues [32]. We were able to confirm the absence of extended *O*-mannosyl glycans in the POMGnT1 -/- as shown in Table 3-1. Interestingly, no β-1-6 branching was observed suggesting that the activity of POMGnT1 precedes and is necessary for the activity of GnTV-b that is involved in branching of *O*-Man structures[33]. However, the presence of *O*-Mannitol was elevated approximately 2.4 fold in the brain tissue of the POMGnT1 -/- over that of the POMGnT1 +/+ suggesting no defect in POMT1/2 activity. The disialyated T antigen (*m/z* 1256) was observed to be the most abundant glycan structure found in both the POMGnT1 +/+ and -/- brains. Structural assignment was made based upon acquired MS/MS fragmentation spectra (Fig. 3-1c and d). Using IIH6 antibody, which specifically recognizes functionally active, fully glycosylated α-DG, we detected the absence of functional α-DG in the POMGnT1 -/- brain.

Although POMGnT1 had previously been reported to extend *O*-mannosyl glycans with a β1,2 linked GlcNAc residue [34], the presence of *O*-mannosyl glycans on α-DG extended with β-1,4 GlcNAc has been recently proposed[20]. Prior work completed to determine the linkage of the GlcNAc concluded that the GlcNAc residue was linked to the C2 position of the mannose based upon treatment of released

glycans with jack bean β-N-acetylhexosaminidase, which cleaves at the β-1,2 linkage. Additionally, analysis by reversed phase HPLC was used to confirm the extension of the *O*-mannose by β-1,2 linked GlcNAc [9]. However, recent work by Yoshida-Moriguchi *et al.*, indicates the presence of a novel phosphorylated structure in which the *O*-mannose is extended by a β-1,4 linked GlcNAc [20]. Therefore we definitively determined the linkage at which GlcNAc was added to a synthetic *O*-mannosyl peptide with recombinant POMGnT1 using NMR. From our analysis we concluded the POMGnT1 is indeed responsible for extending *O*-mannose with a β-1,2 linked GlcNAc (Fig. 3-2) as previously reported[34]. This suggests that there is as a yet undetermined GlcNAc transferase capable of acting on O-Man residues to introduce a β-1,4 linked GlcNAc moiety.

Through our analysis of *O*-glycans released from LARGE +/+ and -/- mouse brain proteins, we were unable to detect any differences in the relative abundance of observed *O*-glycan structures. However, previous work has shown α-DG to be hypoglycosylated in the LARGE deficient myodystrophy (Large^{myd}) mouse, resulting in a reduction in ligand binding activity[14]. In work previously reported by Sutton-Smith *et. al.*, *O*-glycan profiles from both a normal and Large^{myd} mouse brain were compared. In their study, they observed 3 *O*-GalNAc and 2 *O*-mannose initiated glycan structures of similar abundance in both the normal and Large^{myd} mouse brains [35]. Here, we corroborate their findings with more in depth of analysis with our assignment of 25 *O*-GalNAc structures and 9 *O*-mannose structures in both the LARGE +/+ and -/- brains with similar abundances. Recently, Campbell and colleagues observed a phosphorylated *O*-mannosyl glycan structure, which is LARGE dependent, present on α-DG that is required for laminin binding [20]. However, during our analysis of *O*-glycans released from LARGE -/- mouse brain proteins we were unable to observe this glycan structure. Most likely the absence of this glycan structure in the LARGE -/-brains can be attributed to its minor contribution to the complete pool of *O*-glycans and perhaps may reflect that it is on a limited set of proteins including α-DG.

Having fully characterized O-glycans from two mouse models of CMD, we proceeded to characterize *O*-glycans released from α-DG +/+ and -/- mouse brain proteins, since α-DG is a primary

protein associated with CMD and the only well characterized *O*-Man modified mammalian protein. Similar to the other instances in which we analyzed the released *O*-glycans, we proceeded to compare the relative abundance of *O*-glycans detected from both the wild type and knockout mouse brain proteins. Surprisingly from the generated full MS scans (Fig. 3-4a and b) there was not a discernable difference in the prevalence of *O*-mannose initiated glycan structures (Table 3-3) observed in the α-DG -/- brain compared to those detected from the wild type. The presence of *O*-mannosyl glycans was further confirmed upon analysis of the MS/MS fragmentation spectra as shown in Figures 4 C and D, where the presence of a fucosylated *O*-mannosyl glycan was confirmed in both the wild type and α-DG -/- brain. Consequently, we observed a significant reduction in the amount of α-DG detectable by western blotting with IIH6 against proteins from the α-DG -/- brain (Figure 3-4 e). Our results mimic those that were obtained by Campbell and colleges where minute amounts of α-DG (less than 5% wild type levels) could be detected in the vascular smooth muscle of cerebral blood vessels in the α-DG -/- brain [26]. Comparison of the relative amounts of *O*-mannosyl glycans expressed on α-DG +/+ and -/- mouse brain proteins resulted in an inability to detect a significant difference in the amount of *O*-mannose structures attached to brain proteins though α-DG expression was greatly reduced. This strongly suggests that there must be additional *O*-mannose modified proteins in the mammalian brain to account for the presence *O*-mannose structures that were determined to be on the α-DG -/- mouse brain proteins.

In conclusion, we were able to develop and implement workflows that allowed us to perform glycomic analysis on *O*-glycans released from the mouse brain proteins from select mouse models of CMD. Collectively, we were able to fully characterize 34 different glycan structures from the three phenotypes and the wild type controls we investigated (Tables 3-1, 3-2, and 3-3). Through our analysis we were able to identify 9 *O*-Man and 25 *O*-GalNAc (including 6 isobaric species) initiated *O*-glycan structures. By characterizing these *O*-glycan structures, we are able to provide some insight into the role of the known and putative glycosyltransferases POMGnT1 and LARGE, respectively, have on the *O*-mannosyltion pathway. Additionally, through the detection of *O*-mannosyl glycan structures in the α-DG -/-, we were able to provide evidence for additional *O*-mannosylated proteins in the mammalian brain.

REFERENCES

1. Lisi, M. T.; Cohn, R. D., Congenital muscular dystrophies: new aspects of an expanding group of disorders. *Biochim Biophys Acta* **2007,** 1772, (2), 159-72.

2. Ibraghimov-Beskrovnaya, O.; Ervasti, J. M.; Leveille, C. J.; Slaughter, C. A.; Sernett, S. W.; Campbell, K. P., Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature* **1992,** 355, (6362), 696-702.

3. Ervasti, J. M.; Campbell, K. P., A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin. *J Cell Biol* **1993,** 122, (4), 809-23.

4. Barresi, R.; Campbell, K. P., Dystroglycan: from biosynthesis to pathogenesis of human disease. *J Cell Sci* **2006,** 119, (Pt 2), 199-207.

5. Martin, P. T., Congenital muscular dystrophies involving the O-mannose pathway. *Curr Mol Med* **2007,** 7, (4), 417-25.

6. Endo, T.; Manya, H., O-mannosylation in mammalian cells. *Methods Mol Biol* **2006,** 347, 43-56.

7. Beltran-Valero de Bernabe, D.; Currier, S.; Steinbrecher, A.; Celli, J.; van Beusekom, E.; van der Zwaag, B.; Kayserili, H.; Merlini, L.; Chitayat, D.; Dobyns, W. B.; Cormand, B.; Lehesjoki, A. E.; Cruces, J.; Voit, T.; Walsh, C. A.; van Bokhoven, H.; Brunner, H. G., Mutations in the O-mannosyltransferase gene POMT1 give rise to the severe neuronal migration disorder Walker-Warburg syndrome. *Am J Hum Genet* **2002,** 71, (5), 1033-43.

8. 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.; Huynen, M. A.; Verrips, A.; Walsh, C. A.; Barth, P. G.; Brunner, H. G.; van Bokhoven, H., POMT2 mutations cause alpha-dystroglycan hypoglycosylation and Walker-Warburg syndrome. *J Med Genet* **2005,** 42, (12), 907-12.

9. Yoshida, A.; Kobayashi, K.; Manya, H.; Taniguchi, K.; Kano, H.; Mizuno, M.; Inazu, T.; Mitsuhashi, H.; Takahashi, S.; Takeuchi, M.; Herrmann, R.; Straub, V.; Talim, B.; Voit, T.; Topaloglu, H.; Toda, T.; Endo, T., Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. *Dev Cell* **2001,** 1, (5), 717-24.

10. Kobayashi, K.; Nakahori, Y.; Miyake, M.; Matsumura, K.; Kondo-Iida, E.; Nomura, Y.; Segawa, M.; Yoshioka, M.; Saito, K.; Osawa, M.; Hamano, K.; Sakakihara, Y.; Nonaka, I.; Nakagome, Y.; Kanazawa, I.; Nakamura, Y.; Tokunaga, K.; Toda, T., An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy. *Nature* **1998,** 394, (6691), 388-92.

11. Brockington, M.; Yuva, Y.; Prandini, P.; Brown, S. C.; Torelli, S.; Benson, M. A.; Herrmann, R.; Anderson, L. V.; Bashir, R.; Burgunder, J. M.; Fallet, S.; Romero, N.; Fardeau, M.; Straub, V.; Storey, G.; Pollitt, C.; Richard, I.; Sewry, C. A.; Bushby, K.; Voit, T.; Blake, D. J.; Muntoni, F., Mutations in the fukutin-related protein gene (FKRP) identify limb girdle muscular dystrophy 2I as a milder allelic variant of congenital muscular dystrophy MDC1C. *Hum Mol Genet* **2001,** 10, (25), 2851-9.

12. Longman, C.; Brockington, M.; Torelli, S.; Jimenez-Mallebrera, C.; Kennedy, C.; Khalil, N.; Feng, L.; Saran, R. K.; Voit, T.; Merlini, L.; Sewry, C. A.; Brown, S. C.; Muntoni, F., Mutations in the human LARGE gene cause MDC1D, a novel form of congenital muscular dystrophy with severe mental retardation and abnormal glycosylation of alpha-dystroglycan. *Hum Mol Genet* **2003,** 12, (21), 2853-61.

13. Jimenez-Mallebrera, C.; Torelli, S.; Feng, L.; Kim, J.; Godfrey, C.; Clement, E.; Mein, R.; Abbs, S.; Brown, S. C.; Campbell, K. P.; Kroger, S.; Talim, B.; Topaloglu, H.; Quinlivan, R.; Roper, H.; Childs, A. M.; Kinali, M.; Sewry, C. A.; Muntoni, F., A comparative study of alpha-dystroglycan glycosylation in dystroglycanopathies suggests that the hypoglycosylation of alpha-dystroglycan does not consistently correlate with clinical severity. *Brain Pathol* **2009,** 19, (4), 596-611.

14. Michele, D. E.; Barresi, R.; Kanagawa, M.; Saito, F.; Cohn, R. D.; Satz, J. S.; Dollar, J.; Nishino, I.; Kelley, R. I.; Somer, H.; Straub, V.; Mathews, K. D.; Moore, S. A.; Campbell, K. P., Post-translational disruption of dystroglycan-ligand interactions in congenital muscular dystrophies. *Nature* **2002,** 418, (6896), 417-22.

15. Ervasti, J. M.; Campbell, K. P., Membrane organization of the dystrophin-glycoprotein complex. *Cell* **1991,** 66, (6), 1121-31.

16. Stalnaker, S. H.; Hashmi, S.; Lim, J. M.; Aoki, K.; Porterfield, M.; Gutierrez-Sanchez, G.; Wheeler, J.; Ervasti, J. M.; Bergmann, C.; Tiemeyer, M.; Wells, L., Site-mapping and characterization of O-glycan structures on alpha-dystroglycan isolated from rabbit skeletal muscle. *J Biol Chem*.

17. Nakamura, N.; Stalnaker, S. H.; Lyalin, D.; Lavrova, O.; Wells, L.; Panin, V. M., Drosophila Dystroglycan is a target of O-mannosyltransferase activity of two protein O-mannosyltransferases, Rotated Abdomen and Twisted. *Glycobiology* 20, (3), 381-94.

18. Manya, H.; Chiba, A.; Yoshida, A.; Wang, X.; Chiba, Y.; Jigami, Y.; Margolis, R. U.; Endo, T., Demonstration of mammalian protein O-mannosyltransferase activity: coexpression of POMT1 and POMT2 required for enzymatic activity. *Proc Natl Acad Sci U S A* **2004,** 101, (2), 500-5.

19. Combs, A. C.; Ervasti, J. M., Enhanced laminin binding by alpha-dystroglycan after enzymatic deglycosylation. *Biochem J* **2005,** 390, (Pt 1), 303-9.

20. Yoshida-Moriguchi, T.; Yu, L.; Stalnaker, S. H.; Davis, S.; Kunz, S.; Madson, M.; Oldstone, M. B.; Schachter, H.; Wells, L.; Campbell, K. P., O-mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. *Science* 327, (5961), 88-92.

21. Chai, W.; Yuen, C. T.; Kogelberg, H.; Carruthers, R. A.; Margolis, R. U.; Feizi, T.; Lawson, A. M., High prevalence of 2-mono- and 2,6-di-substituted manol-terminating sequences among O-glycans released from brain glycopeptides by reductive alkaline hydrolysis. *Eur J Biochem* **1999,** 263, (3), 879-88.

22. Aoki, K.; Perlman, M.; Lim, J. M.; Cantu, R.; Wells, L.; Tiemeyer, M., Dynamic developmental elaboration of N-linked glycan complexity in the Drosophila melanogaster embryo. *J Biol Chem* **2007,** 282, (12), 9127-42.

23. Aoki, K.; Porterfield, M.; Lee, S. S.; Dong, B.; Nguyen, K.; McGlamry, K. H.; Tiemeyer, M., The diversity of O-linked glycans expressed during Drosophila melanogaster development reflects stage- and tissue-specific requirements for cell signaling. *J Biol Chem* **2008**.

24. Liu, J.; Ball, S. L.; Yang, Y.; Mei, P.; Zhang, L.; Shi, H.; Kaminski, H. J.; Lemmon, V. P.; Hu, H., A genetic model for muscle-eye-brain disease in mice lacking protein O-mannose 1,2-N-acetylglucosaminyltransferase (POMGnT1). *Mech Dev* **2006,** 123, (3), 228-40.

25. Kanagawa, M.; Saito, F.; Kunz, S.; Yoshida-Moriguchi, T.; Barresi, R.; Kobayashi, Y. M.; Muschler, J.; Dumanski, J. P.; Michele, D. E.; Oldstone, M. B.; Campbell, K. P., Molecular recognition by LARGE is essential for expression of functional dystroglycan. *Cell* **2004,** 117, (7), 953-64.

26. Moore, S. A.; Saito, F.; Chen, J.; Michele, D. E.; Henry, M. D.; Messing, A.; Cohn, R. D.; Ross-Barta, S. E.; Westra, S.; Williamson, R. A.; Hoshi, T.; Campbell, K. P., Deletion of brain dystroglycan recapitulates aspects of congenital muscular dystrophy. *Nature* **2002,** 418, (6896), 422-5.

27. Ciucanu, I.; Kerek, F., A Simple and Rapid Method for the Permethylation of Carbohydrates. *Carbohydrate Research* **1984,** 131, (2), 209-217.

28. Bax, A.; Summers, M., H-1 and C-13 Assignments from Sensitivity-Enhanced Detection of Heteronuclear Multiple-Bond Connectivity by 2d Multiple Quantum NMR. *J Am Chem Soc* **1986,** 108, (8), 2093-2094.

29. Podlasek, C.; Wu, J.; Stripe, W.; Bondo, P.; Serianni, A. S., [C-13]-Enriched Methyl Aldopyranosides-Structural Interpretations of C-13-1h Spin-Coupling Constants and H-1 chemical-Shifts. *J Am Chem Soc* **1995,** 117, (33), 8635-8644.

30. Bubb, W. A., NMR Spectroscopy in the Study of Carbohydrates: Characterizing the Structural Complexity. *Concepts in Magnetic Resonance Part A* **2003,** 19A, (1), 1-19.

31. Michele, D. E.; Campbell, K. P., Dystrophin-glycoprotein complex: post-translational processing and dystroglycan function. *J Biol Chem* **2003,** 278, (18), 15457-60.

32. Endo, T., O-mannosyl glycans in mammals. *Biochim Biophys Acta* **1999,** 1473, (1), 237-46.

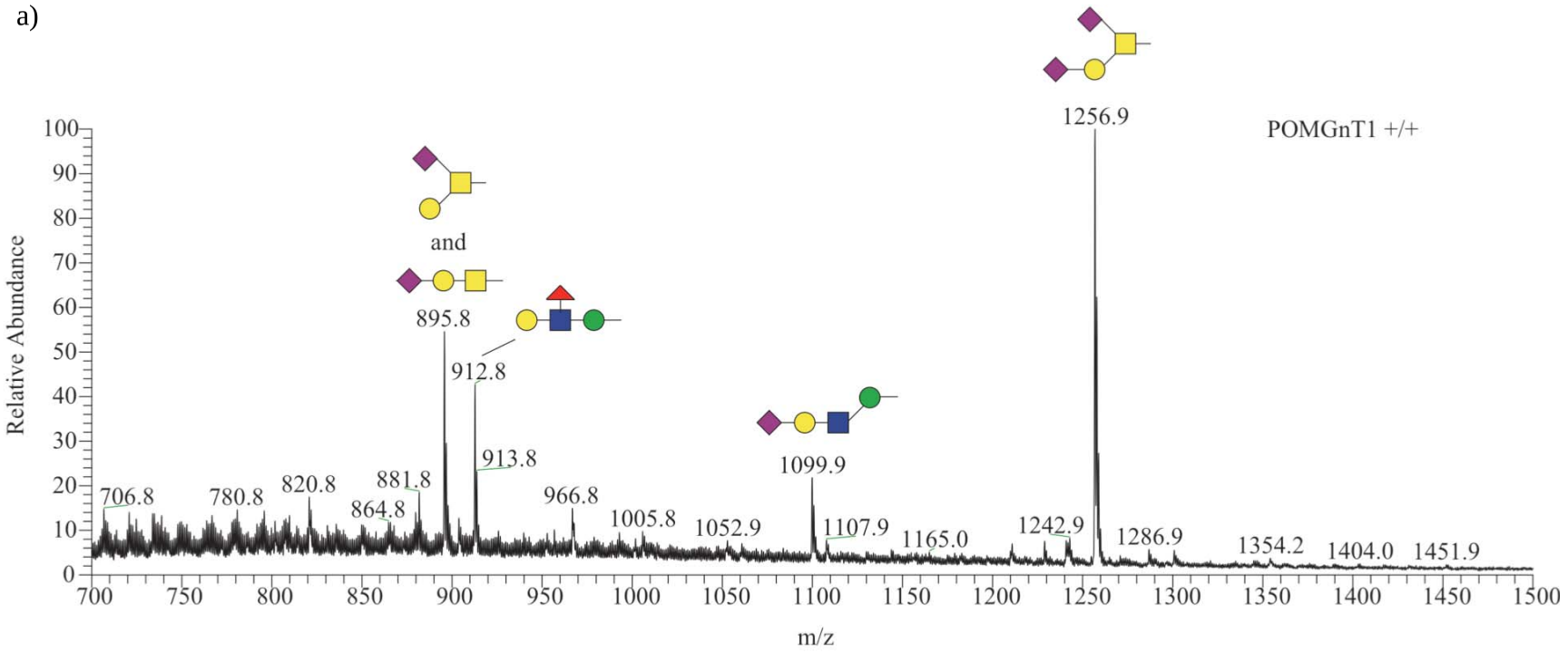
33. Alvarez-Manilla, G.; Troupe, K.; Fleming, M.; Martinez-Uribe, E.; Pierce, M., Comparison of the substrate specificities and catalytic properties of the sister N-acetylglucosaminyltransferases, GnT-V and GnT-Vb (IX). *Glycobiology* 20, (2), 166-74.

34. Chiba, A.; Matsumura, K.; Yamada, H.; Inazu, T.; Shimizu, T.; Kusunoki, S.; Kanazawa, I.; Kobata, A.; Endo, T., Structures of sialylated O-linked oligosaccharides of bovine peripheral nerve alpha-dystroglycan. The role of a novel O-mannosyl-type oligosaccharide in the binding of alpha-dystroglycan with laminin. *J Biol Chem* **1997,** 272, (4), 2156-62.

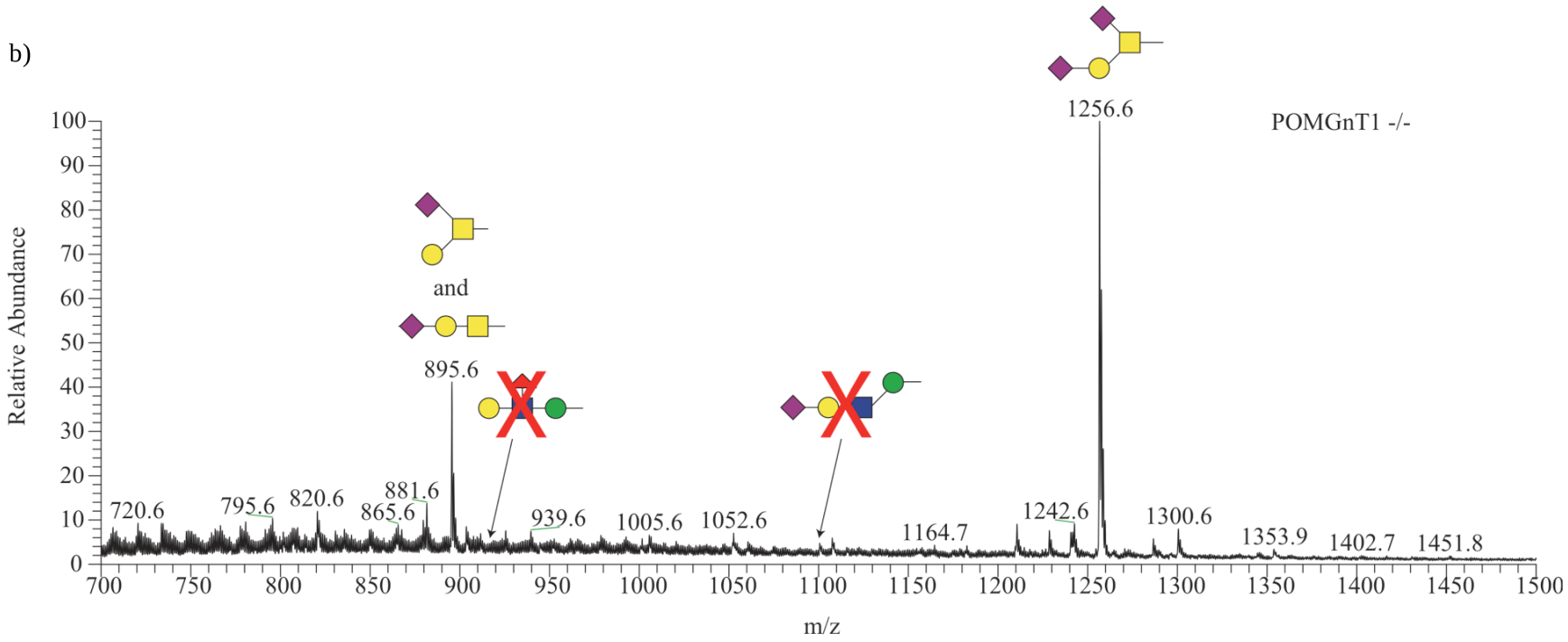
35. Sutton-Smith, M.; Morris, H. R.; Grewal, P. K.; Hewitt, J. E.; Bittner, R. E.; Goldin, E.; Schiffmann, R.; Dell, A., MS screening strategies: investigating the glycomes of knockout and myodystrophic mice and leukodystrophic human brains. *Biochem Soc Symp* **2002**, (69), 105-15.

FIGURE 3-1. *O*-glycans released from POMGnT1 +/+ and -/- mouse brain proteins.
a,b) *O*-glycans were released from protein powder made from POMGnT1 +/+ and -/- mouse brain proteins by β-elimination. Through comparison of the full MS scans of +/+ and -/- we were able to observe the absence of prominent *O*-mannose structures in the -/-. **c,d)** From the MS/MS scan *m/z* 1256 the disialyated Tn antigen, is unaffected in both the POMGnT1 +/+ and -/-. **e)** Antibody IIH6, which recognizes fully glycosylated, functionally active form of α-DG, shows absence of functionally active α-DG in POMGnT1 -/- brains. **f)** Individual monosaccharides were released from both POMGnT1 +/+ and -/- mouse brains protein and the relative abundance of *O*-mannitol compared. From the two data sets, proteins carrying non-extended *O*-mannose structures were observed to be enriched ~2.4 fold in the POMGnT1 -/- animals.

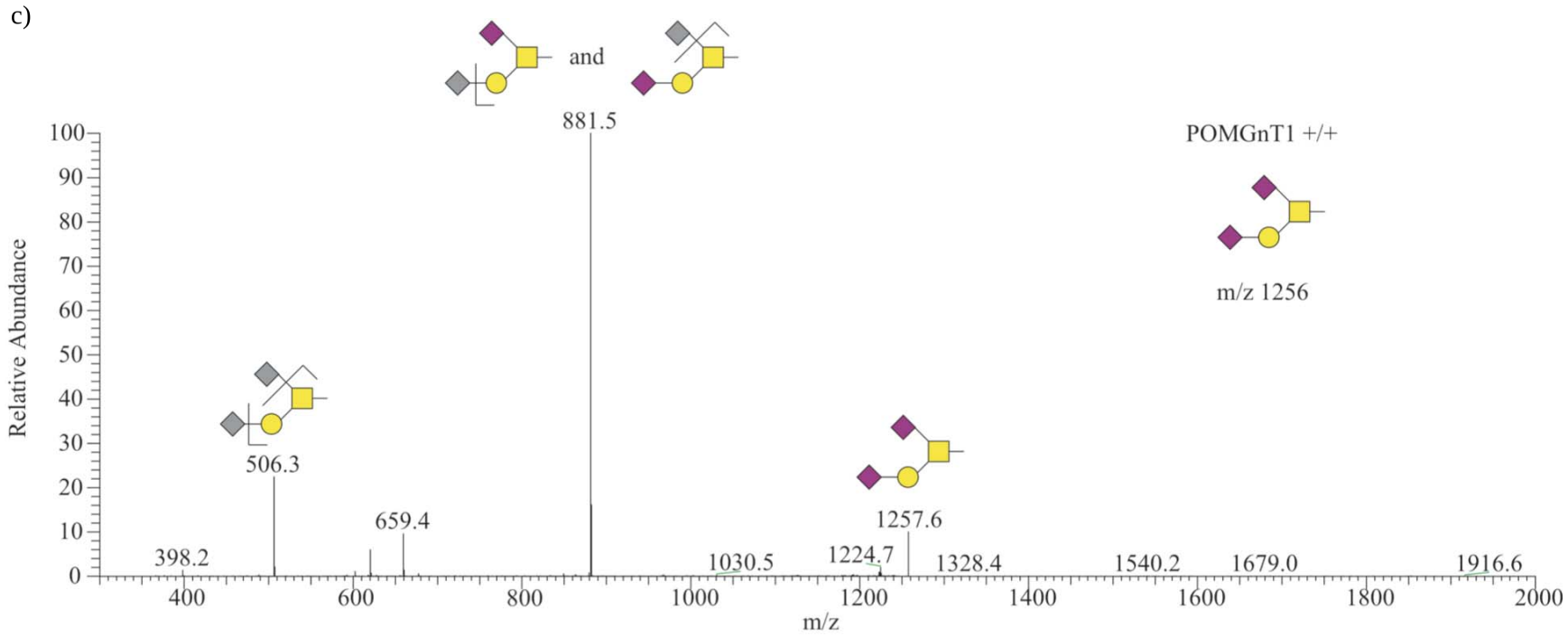
a)


POMGnT1 +/+
Relative Abundance
m/z
and
706.8
780.8
820.8
864.8
881.8
895.8
912.8
913.8
966.8
1005.8
1052.9
1099.9
1107.9
1165.0
1242.9
1256.9
1286.9
1354.2
1404.0
1451.9

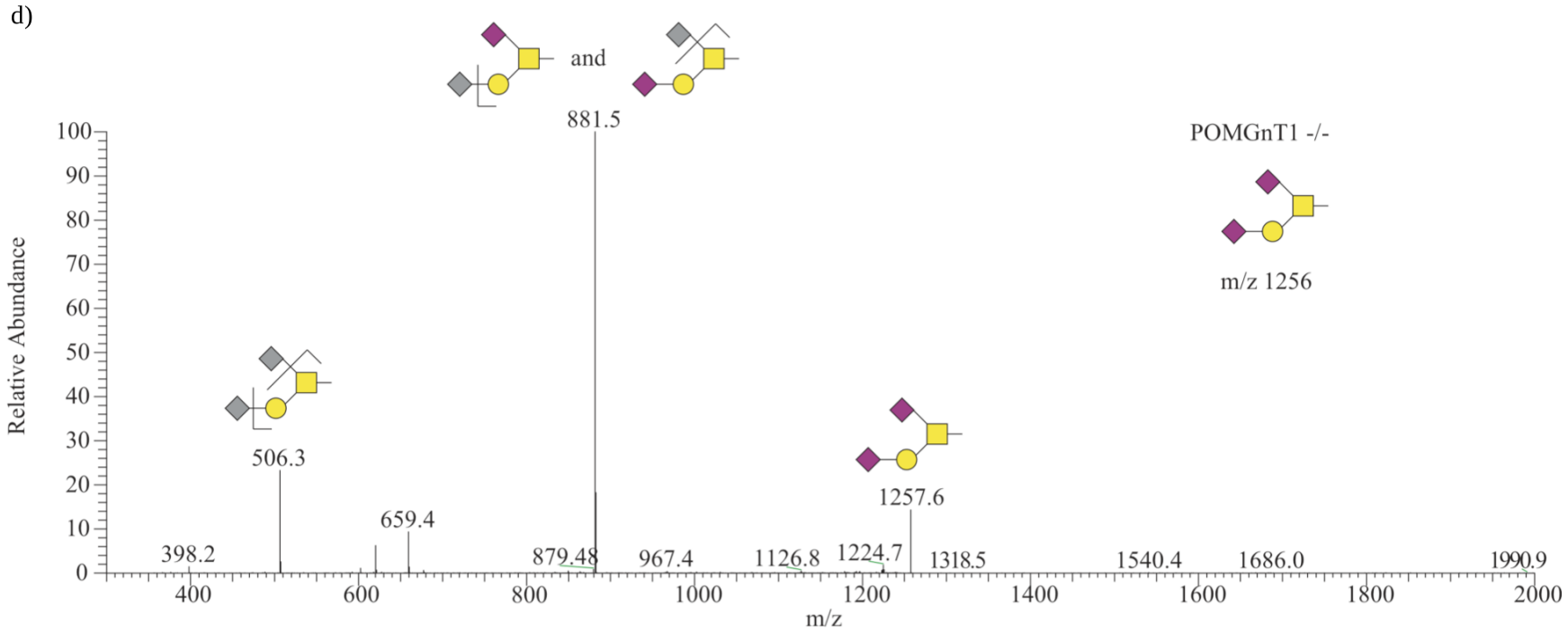
b)


POMGnT1 -/-
Relative Abundance
m/z
and
1256.6
895.6
720.6
795.6
820.6
881.6
865.6
939.6
1005.6
1052.6
1164.7
1242.6
1300.6
1353.9
1402.7
1451.8

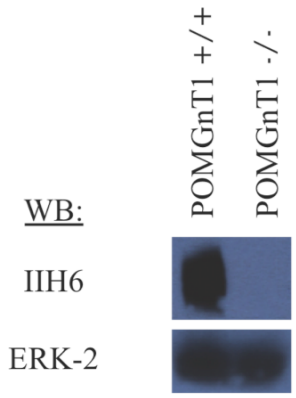
c)


and
881.5
POMGnT1 +/+
m/z 1256
506.3
659.4
398.2
1257.6
1030.5
1224.7
1328.4
1540.2
1679.0
1916.6
Relative Abundance
m/z
100
90
80
70
60
50
40
30
20
10
0
400
600
800
1000
1200
1400
1600
1800
2000

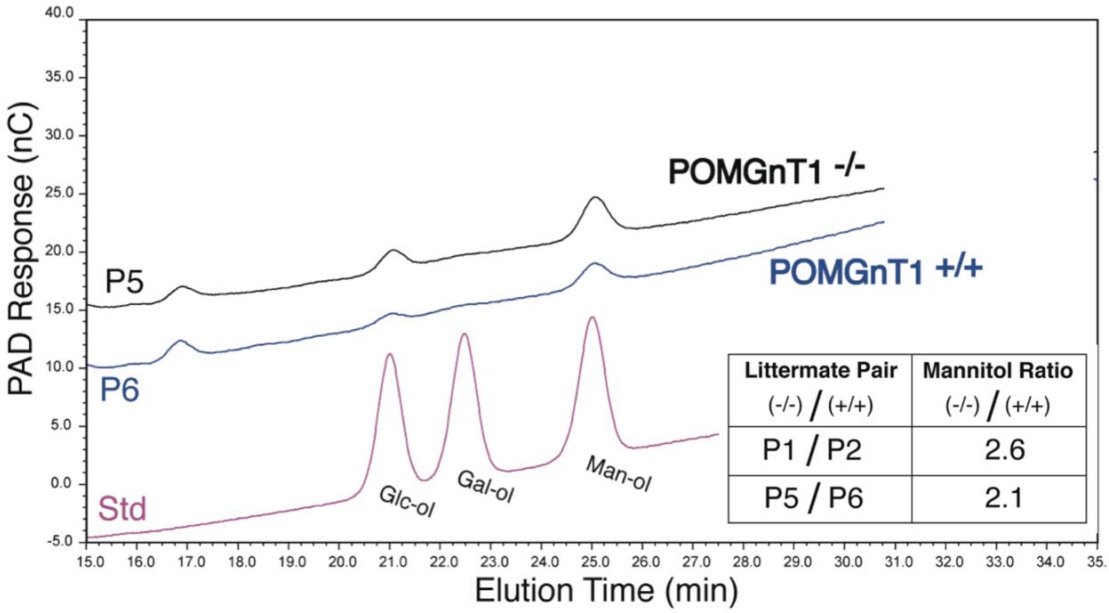
d)


and
881.5
POMGnT1 -/-
m/z 1256
506.3
659.4
1257.6
398.2
879.48
967.4
1126.8
1224.7
1318.5
1540.4
1686.0
1990.9
Relative Abundance
m/z
100
90
80
70
60
50
40
30
20
10
0
400
600
800
1000
1200
1400
1600
1800
2000

e)



f)


PAD Response (nC)
Elution Time (min)
POMGnT1 -/-
POMGnT1 +/+
P5
P6
Std
Glc-ol
Gal-ol
Man-ol
Littermate Pair (-/-) / (+/+)
Mannitol Ratio (-/-) / (+/+)
P1 / P2
2.6
P5 / P6
2.1

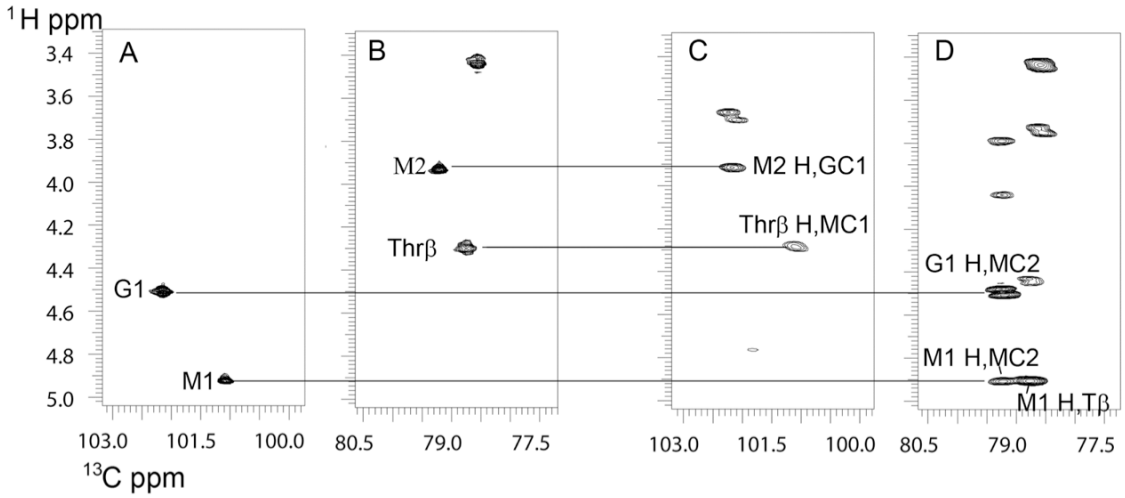
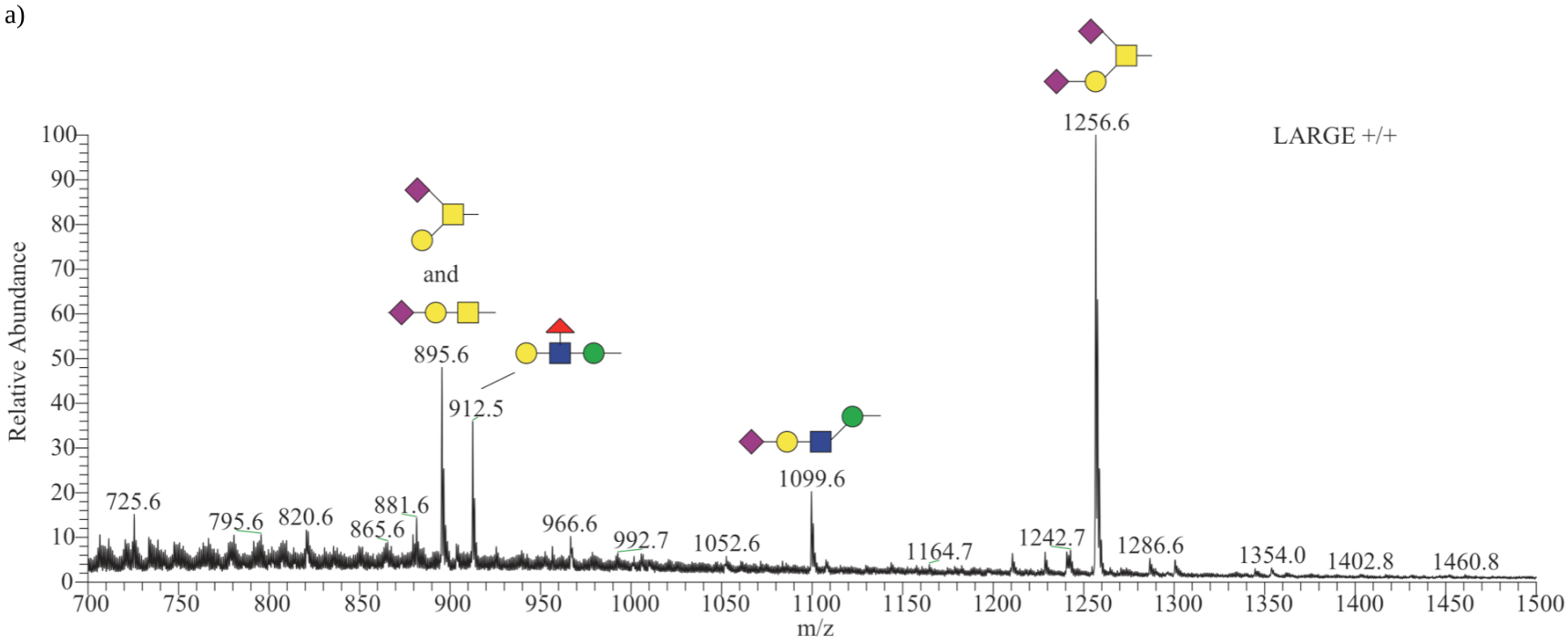


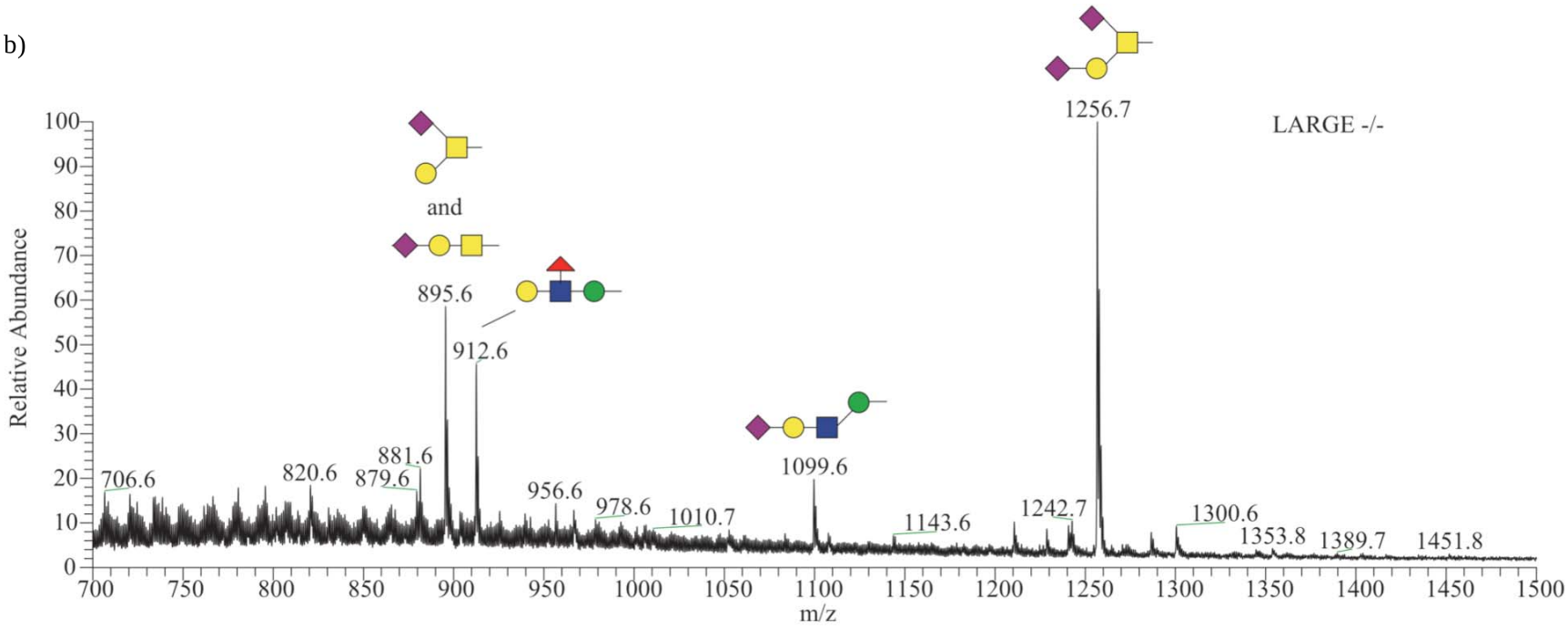
FIGURE 3-2. Linkage analysis of recombinant POMGnT1 by NMR.
^{1}H-^{13}C 2-dimensional correlation spectra showing signals derived from the carbohydrate portion of the Ac-YVEP(GlcNAc-β-1,2-Man-α-)TAV-NH_2 glycopeptide prepared from POMGnT1 action on the mannosylated peptide. Panels A and B are sections from an HMQC spectrum where the crosspeaks are at the coordinates of pairs of directly bonded protons and carbons, and C and D are similar regions from the HMBC spectrum where the peaks are at the coordinates of a proton and carbons two or three bonds away. The lines identify peaks in the spectra that are associated with the same protons associated with sites of glycosidic linkage. Peaks in C show connections between the H-2 proton of Man and the C1 of GlcNAc, establishing the GlcNAc C1 to Man C2 glycosidic linkage, and additionally the linkage from the Thr β proton and the Man C1. These are confirmed in the other direction across the linkages by the peaks in panel D.

FIGURE 3-3. *O*-glycans released from *Large* +/+ and -/- mouse brains proteins.
a,b) *O*-glycans were released from protein powder made from both *Large+/+ and -/-* mouse brain proteins. Through comparison of the full MS scans we observed no difference in the prominent *O*-glycan structures that were detected from +/+ and -/- brain proteins. **c,d)** MS/MS fragmentation spectra of m/z 1100 indicates the presence of the classical *O*-man tetrasaccharide in both the *Large* +/+ and -/- mouse brains. **e)** Antibody IIH6, which recognizes the fully glycosylated, functionally active form of α-DG, shows absence of functionally active α-DG in *Large* -/- brains.

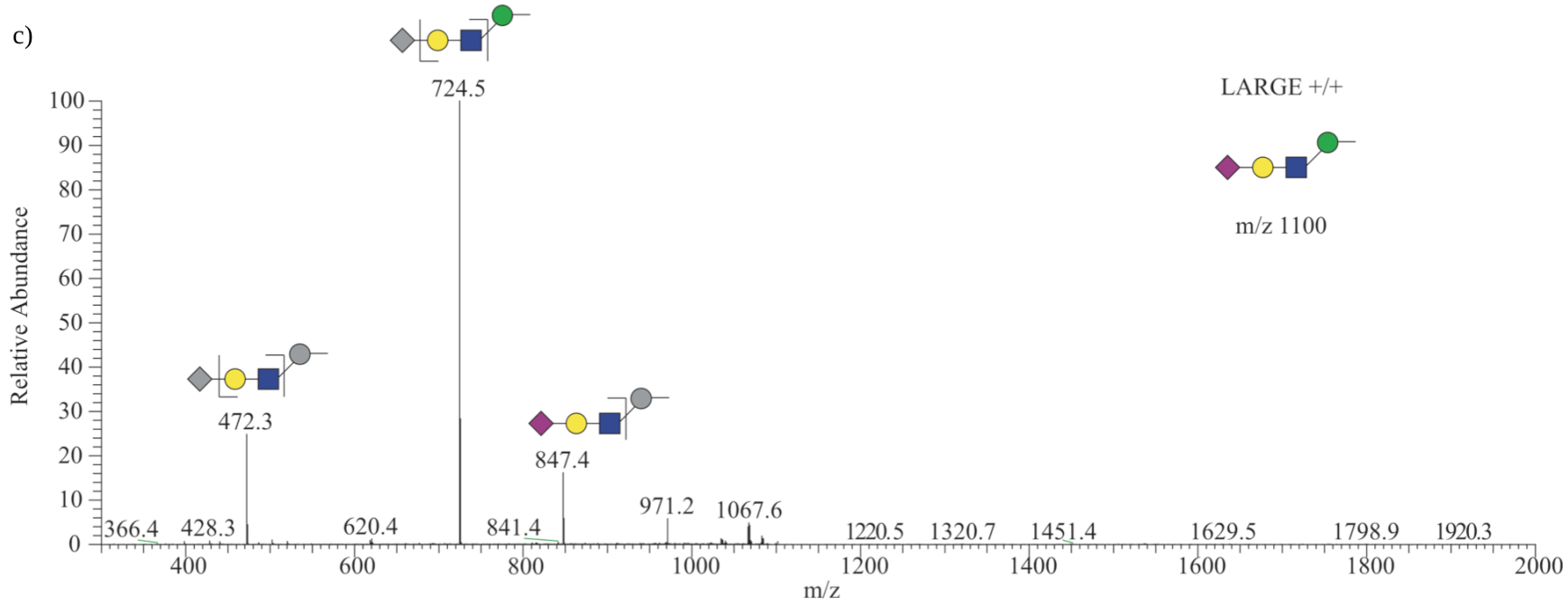
a)


LARGE +/+
Relative Abundance
m/z
and
725.6
795.6
820.6
865.6
881.6
895.6
912.5
966.6
992.7
1052.6
1099.6
1164.7
1242.7
1256.6
1286.6
1354.0
1402.8
1460.8

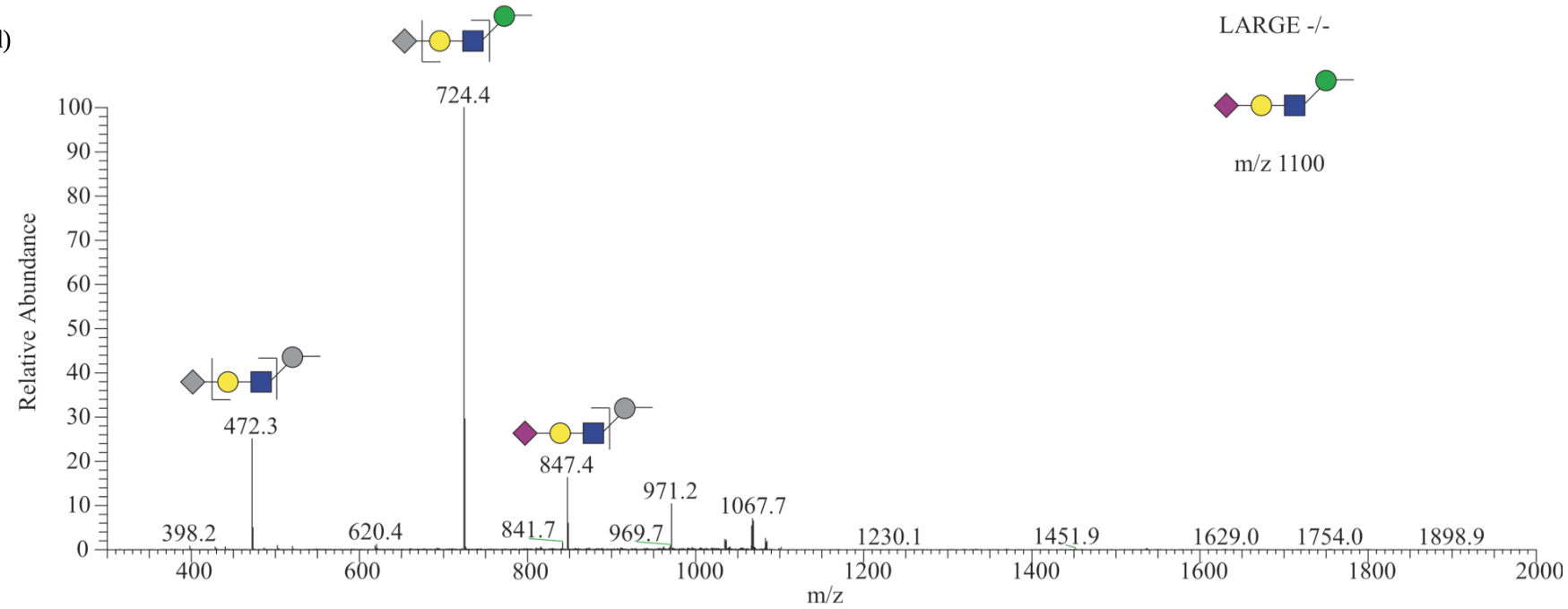
b)


LARGE -/-
Relative Abundance
m/z
and
706.6
820.6
879.6
881.6
895.6
912.6
956.6
978.6
1010.7
1099.6
1143.6
1242.7
1256.7
1300.6
1353.8
1389.7
1451.8

c)


724.5
LARGE +/+
m/z 1100
472.3
847.4
971.2
1067.6
366.4
428.3
620.4
841.4
1220.5
1320.7
1451.4
1629.5
1798.9
1920.3
Relative Abundance
m/z
100
90
80
70
60
50
40
30
20
10
0
400
600
800
1000
1200
1400
1600
1800
2000

d)


LARGE -/-
m/z 1100
724.4
472.3
847.4
971.2
1067.7
398.2
620.4
841.7
969.7
1230.1
1451.9
1629.0
1754.0
1898.9
Relative Abundance
m/z
100
90
80
70
60
50
40
30
20
10
0
400
600
800
1000
1200
1400
1600
1800
2000

e)

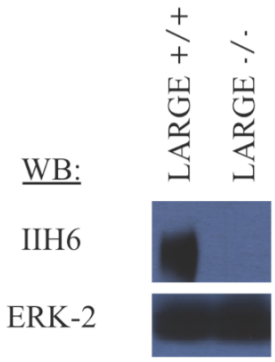
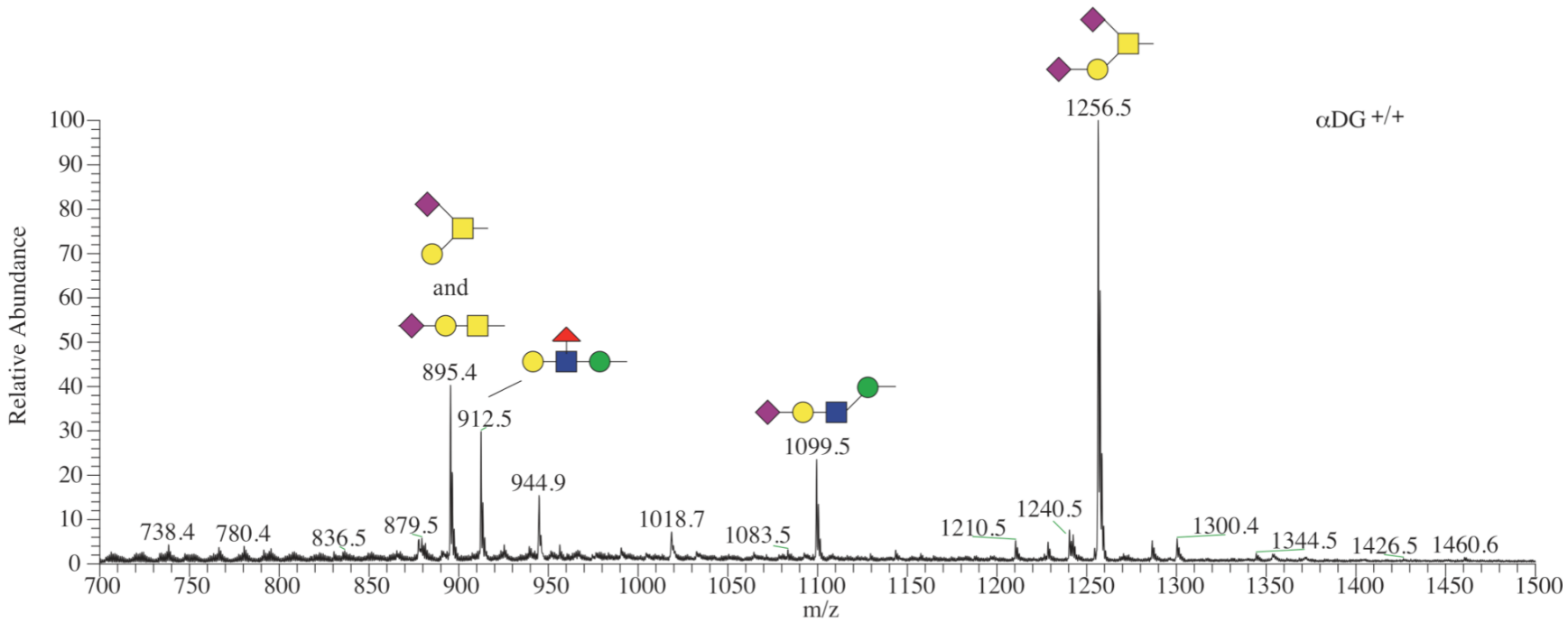
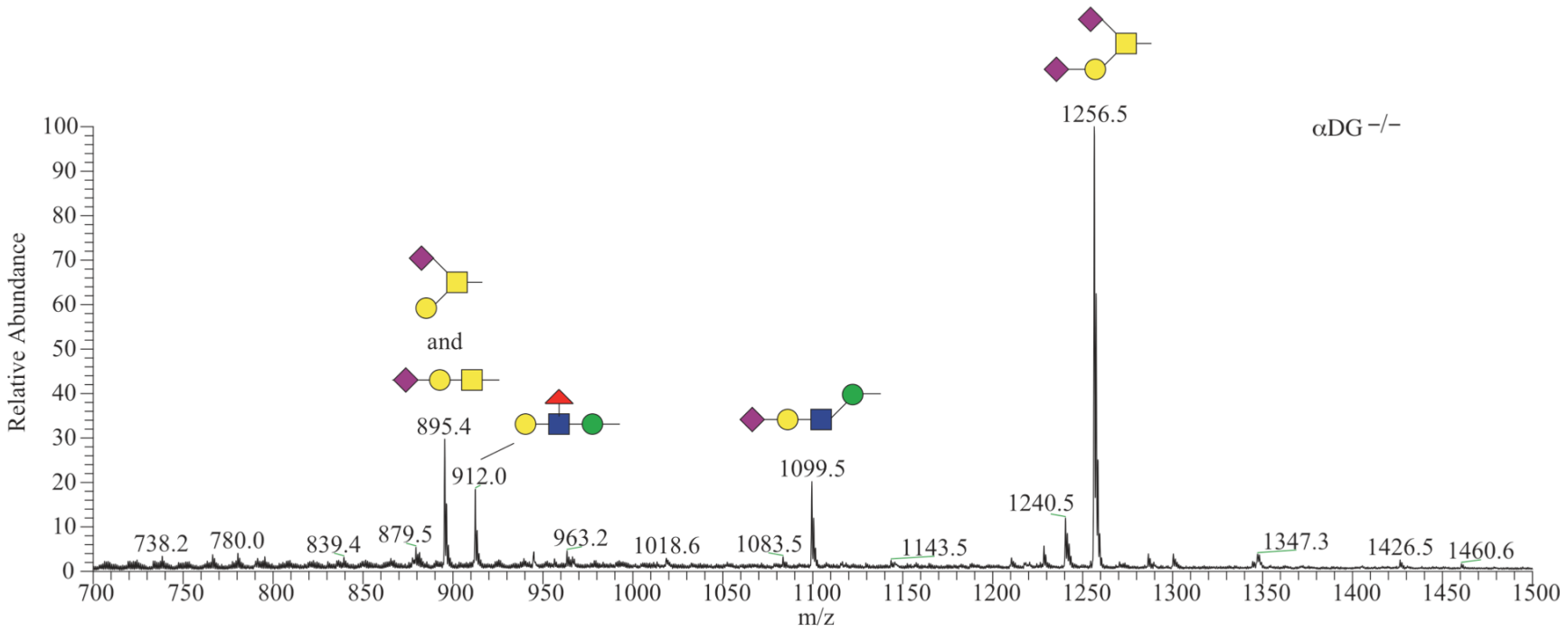


FIGURE 3-4. O-glycans released from α-DG +/+ and -/- mouse cerebrum proteins.
a,b) *O*-glycans were released from protein powder made α-DG +/+ and -/- cerebrum. Through comparison of the full scan we observed no difference in the amount of prominent *O*-mannose initiated glycan structures detected in the α-DG +/+ and -/- proteins. **c,d)** MS/MS fragmentation spectra of *m/z* 912 indicates the presence of the fucosylated *O*-mannose trisaccharide structure in both the α-DG +/+ and -/- mouse brains. **e)** Antibody IIH6, which recognizes the fully glycosylated, functionally active form of α-DG indicates a significant decrease in the amount of α-DG levels present in the α-DG -/- brain compared to that of the +/+.

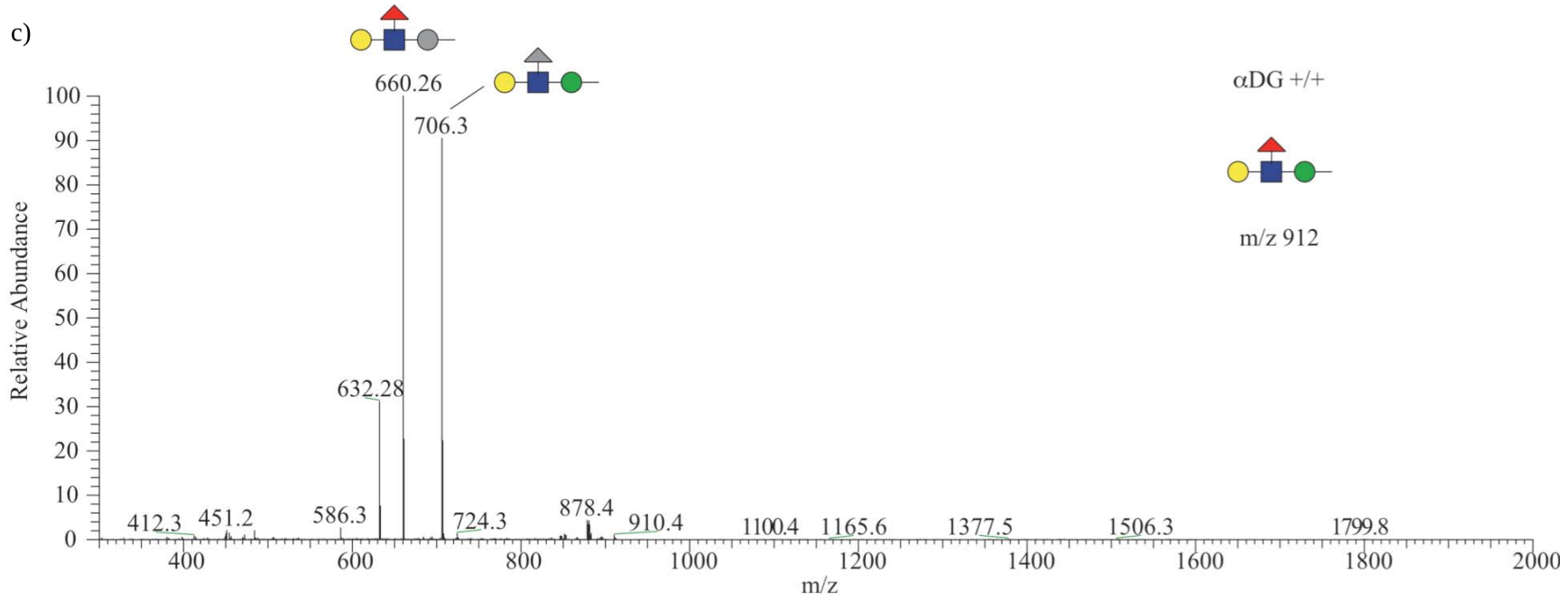
a)


αDG +/+
Relative Abundance
m/z
and
895.4
912.5
944.9
1099.5
1256.5
738.4
780.4
836.5
879.5
1018.7
1083.5
1210.5
1240.5
1300.4
1344.5
1426.5
1460.6

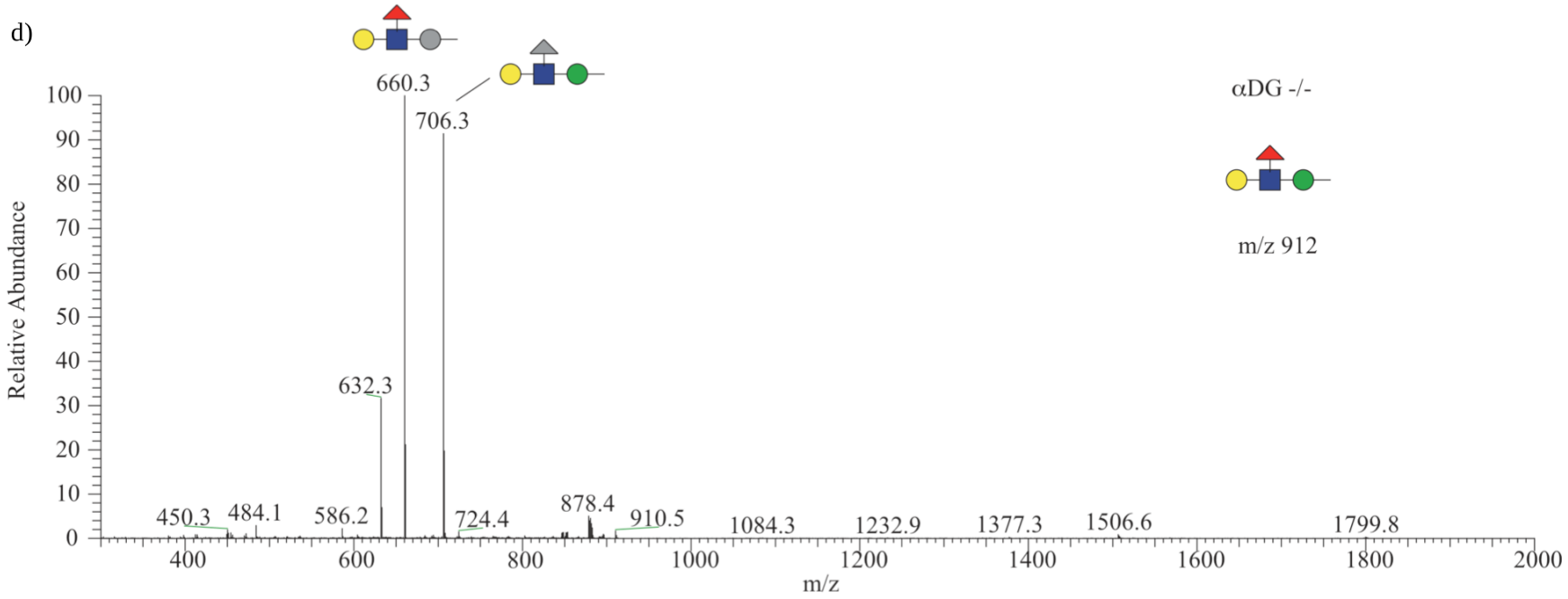
b)


αDG −/−
Relative Abundance
m/z
and
738.2
780.0
839.4
879.5
895.4
912.0
963.2
1018.6
1083.5
1099.5
1143.5
1240.5
1256.5
1347.3
1426.5
1460.6

c)


αDG +/+
m/z 912
Relative Abundance
m/z
660.26
706.3
632.28
412.3
451.2
586.3
724.3
878.4
910.4
1100.4
1165.6
1377.5
1506.3
1799.8

d)


αDG -/-
m/z 912
Relative Abundance
m/z
660.3
706.3
632.3
450.3
484.1
586.2
724.4
878.4
910.5
1084.3
1232.9
1377.3
1506.6
1799.8

e)

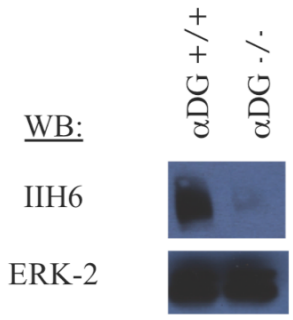


Table 3-1. *O*-glycans Released from POMGnT1 +/+ and -/- Mouse Brain Proteins.

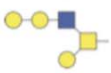
No.	Structure		MW	POMGnT1 +/+	POMGnT1 -/-
1	GlcNAc-Man		534.3	11.9% ± 1.1%	ND
2	Gal-GalNAc	and 			13.9% ± 4.9%
3	HexNAc-GalNAc		575.3	1.7% ± 0.5%	2.5% ± 1.1%
4	Gal-GlcNAc-Man		738.4	2.1% ± 0.4%	ND
5	Neu-Gal-GalNAc		867.5	2.0% ± 1.0%	3.9% ± 3.5%
6	Gal-(Neu)-GalNAc	and 			
7	NeuAc-Gal-GalNAc		895.5	15.2% ± 0.2%	17.1% ± 4.6%
8	Gal-(NeuAc)-GalNAc	and 			
9	Gal-Fuc-GlcNAc-Man		912.5	7.5% ± 3.8%	ND
10	NeuGc-Gal-GalNAc		925.5	2.1% ± 0.5%	2.8% ± 0.9%
11	Gal-(NeuGc)-GalNAc	and 			
12	Gal-(Gal-GlcNAc)-GalNAc		983.5	0.9% ± 0.2%	1.3% ± 0.5%
13	NeuAc-Gal-GlcNAc-Man		1099.6	5.8% ± 1.6%	ND
14	NeuGc-Gal-GlcNAc-Man		1129.6	0.7% ± 0.1%	ND
15	Gal-(Gal-Fuc-GlcNAc)-GalNAc		1157.6	0.7% ± 0.04%	0.8% ± 0.2%
16	Gal-Gal-(Gal-GlcNAc)-GalNAc		1187.6	0.9% ± 0.1%	0.9% ± 0.3%
17	Gal-(Gal-Gal-GlcNAc)-GalNAc	and 			

Table 3-1. Continued.

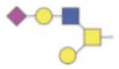
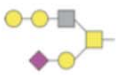
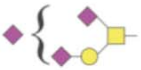
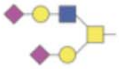
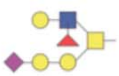
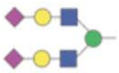
18	Neu-Gal-(NeuAc)-GalNAc	 and 	1228.6	2.5% ± 2.3%	8.4% ± 7.1%
19	NeuAc-Gal-(Neu)-GalNAc				
20	NeuAc-Gal-(NeuAc)-GalNAc		1256.6	40.0% ± 5.3%	41.7% ± 10.3%
21	NeuGc-Gal-(NeuAc)-GalNAc	 and 	1286.6	1.6% ± 0.5%	2.1% ± 0.3%
22	NeuAc-Gal-(NeuGc)-GalNAc				
23	NeuAc-Gal-(Fuc-GlcNAc)-GalNAc		1314.7	0.9% ± 1.1%	0.5% ± 0.1%
24	NeuGc-Gal-(NeuGc)-GalNAc		1316.7	0.4% ± 0.04%	0.5% ± 0.2%
25	NeuAc-Gal-GlcNAc-(GlcNAc)-Man	 and 	1344.7	0.9% ± 0.2%	ND
26	Gal-(NeuAc-Gal-GlcNAc)-GalNAc				1.1% ± 0.2%
27	NeuAc-(NeuAc-Gal-GlcNAc)-GalNAc		1501.8	0.1% ± 0.01%	0.3% ± 0.04%
28	NeuAc-Gal-GlcNAc-(Gal-GlcNAc)-Man	 and 	1548.8	0.3% ± 0.1%	ND
29	NeuAc-Gal-(Gal-Gal-HexNAc)-GalNAc				0.6% ± 0.3%
30	NeuAc-NeuAc-Gal-(NeuAc)-GalNAc		1617.8	1.1% ± 0.3%	1.1% ± 0.3%
31	NeuAc-Gal-(NeuAc-Gal-GlcNAc)-GalNAc		1705.9	0.2% ± 0.02%	0.3% ± 0.2%
32	NeuAc-Gal-GlcNAc-(Gal-Fuc-GlcNAc)-Man	 and 	1722.9	0.2% ± 0.02%	ND
33	NeuAc-Gal-Gal-(Gal-Fuc-GlcNAc)GalNAc				0.1% ± 0.04%
34	NeuAc-Gal-GlcNAc-(NeuAc-Gal-GlcNAc)-Man		1910.0	0.2% ± 0.02%	ND

Table 3-2. *O*-glycans Released from LARGE +/+ and -/- Mouse Brain Proteins.

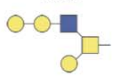
No.	Structure		MW	LARGE +/+	LARGE -/-
1	GlcNAc-Man		534.3	24.9% ± 12.8%	23.2% ± 13.4%
		and			
2	Gal-GalNAc				
3	HexNAc-GalNAc		575.3	2.0% ± 0.4%	1.7% ± 0.4%
4	Gal-GlcNAc-Man		738.4	2.8% ± 0.2%	2.6% ± 0.2%
5	Neu-Gal-GalNAc		867.5	1.5% ± 0.4%	2.1% ± 1.2%
		and			
6	Gal-(Neu)-GalNAc				
7	NeuAc-Gal-GalNAc		895.5	15.9% ± 0.1%	15.6% ± 0.6%
		and			
8	Gal-(NeuAc)-GalNAc				
9	Gal-Fuc-GlcNAc-Man		912.5	9.8% ± 0.9%	11.4% ± 1.9%
10	NeuGc-Gal-GalNAc		925.5	2.3% ± 0.6%	2.6% ± 0.3%
		and			
11	Gal-(NeuGc)-GalNAc				
12	Gal-(Gal-GlcNAc)-GalNAc		983.5	0.8% ± 0.2%	0.9% ± 0.2%
13	NeuAc-Gal-GlcNAc-Man		1099.6	5.6% ± 1.8%	6.2% ± 1.9%
14	NeuGc-Gal-GlcNAc-Man		1129.6	0.6% ± 0.1%	0.8% ± 0.4%
15	Gal-(Gal-Fuc-GlcNAc)-GalNAc		1157.6	0.8% ± 0.3%	0.8% ± 0.3%
16	Gal-Gal-(Gal-GlcNAc)-GalNAc		1187.6	0.6% ± 0.02%	0.6% ± 0.3%
		and			
17	Gal-(Gal-Gal-GlcNAc)-GalNAc				

Table 3-2. Continued.

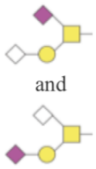
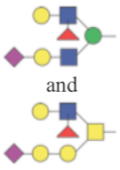
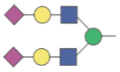
18	Neu-Gal-(NeuAc)-GalNAc		1228.6	1.6% ± 0.6%	1.7% ± 0.5%
19	NeuAc-Gal-(Neu)-GalNAc				
20	NeuAc-Gal-(NeuAc)-GalNAc		1256.6	26.2% ± 9.7%	25.6% ± 8.0%
21	NeuGc-Gal-(NeuAc)-GalNAc		1286.6	1.3% ± 0.4%	1.5% ± 0.4%
22	NeuAc-Gal-(NeuGc)-GalNAc				
23	NeuAc-Gal-(Fuc-GlcNAc)-GalNAc		1314.7	0.2% ± 0.1%	0.2% ± 0.1%
24	NeuGc-Gal-(NeuGc)-GalNAc		1316.7	0.4% ± 0.1%	0.4% ± 0.2%
25	NeuAc-Gal-GlcNAc-(GlcNAc)-Man		1344.7	0.9% ± 0.2%	0.9% ± 0.2%
26	Gal-(NeuAc-Gal-GlcNAc)-GalNAc				
27	NeuAc-(NeuAc-Gal-GlcNAc)-GalNAc		1501.8	0.1% ± 0.02%	0.1% ± 0.07%
28	NeuAc-Gal-GlcNAc-(Gal-GlcNAc)-Man		1548.8	0.3% ± 0.1%	0.2% ± 0.1%
29	NeuAc-Gal-(Gal-Gal-HexNAc)-GalNAc				
30	NeuAc-NeuAc-Gal-(NeuAc)-GalNAc		1617.8	0.7% ± 0.4%	0.6% ± 0.03%
31	NeuAc-Gal-(NeuAc-Gal-GlcNAc)-GalNAc		1705.9	0.2% ± 0.1%	0.2% ± 0.07%
32	NeuAc-Gal-GlcNAc-(Gal-Fuc-GlcNAc)-Man		1722.9	0.2% ± 0.1%	0.2% ± 0.1%
33	NeuAc-Gal-Gal-(Gal-Fuc-GlcNAc)GalNAc				
34	NeuAc-Gal-GlcNAc-(NeuAc-Gal-GlcNAc)-Man		1910.0	0.2% ± 0.1%	0.2% ± 0.1%

Table 3-3. *O*-glycans Released from α-DG +/+ and -/- Mouse Brain Proteins.

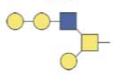
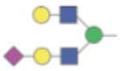
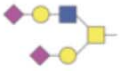
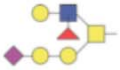
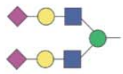
No.	Structure		MW	αDG +/+	αDG -/-
1	GlcNAc-Man		534.3	9.4%	11.4%
		and			
2	Gal-GalNAc				
3	HexNAc-GalNAc		575.3	1.9%	2.1%
4	Gal-GlcNAc-Man		738.4	0.9%	1.5%
5	Neu-Gal-GalNAc		867.5	0.5%	0.9%
		and			
6	Gal-(Neu)-GalNAc				
7	NeuAc-Gal-GalNAc		895.5	12.9%	12.6%
		and			
8	Gal-(NeuAc)-GalNAc				
9	Gal-Fuc-GlcNAc-Man		912.5	7.5%	6.9%
10	NeuGc-Gal-GalNAc		925.5	1.0%	1.1%
		and			
11	Gal-(NeuGc)-GalNAc				
12	Gal-(Gal-GlcNAc)-GalNAc		983.5	0.4%	0.5%
13	NeuAc-Gal-GlcNAc-Man		1099.6	8.5%	8.2%
14	NeuGc-Gal-GlcNAc-Man		1129.6	0.4%	0.7%
15	Gal-(Gal-Fuc-GlcNAc)-GalNAc		1157.6	0.6%	0.6%
16	Gal-Gal-(Gal-GlcNAc)-GalNAc		1187.6	0.4%	0.6%
		and			
17	Gal-(Gal-Gal-GlcNAc)-GalNAc				

Table 3-2. Continued.

18	Neu-Gal-(NeuAc)-GalNAc	 and 	1228.6	1.7%	1.9%
19	NeuAc-Gal-(Neu)-GalNAc				
20	NeuAc-Gal-(NeuAc)-GalNAc		1256.6	48.1%	45.4%
21	NeuGc-Gal-(NeuAc)-GalNAc	 and 	1286.6	1.7%	1.6%
22	NeuAc-Gal-(NeuGc)-GalNAc				
23	NeuAc-Gal-(Fuc-GlcNAc)-GalNAc		1314.7	0.2%	0.2%
24	NeuGc-Gal-(NeuGc)-GalNAc		1316.7	0.3%	0.2%
25	NeuAc-Gal-GlcNAc-(GlcNAc)-Man	 and 	1344.7	0.7%	0.7%
26	Gal-(NeuAc-Gal-GlcNAc)-GalNAc				
27	NeuAc-(NeuAc-Gal-GlcNAc)-GalNAc		1501.8	0.2%	0.2%
28	NeuAc-Gal-GlcNAc-(Gal-GlcNAc)-Man	 and 	1548.8	0.3%	0.4%
29	NeuAc-Gal-(Gal-Gal-HexNAc)-GalNAc				
30	NeuAc-NeuAc-Gal-(NeuAc)-GalNAc		1617.8	1.3%	1.2%
31	NeuAc-Gal-(NeuAc-Gal-GlcNAc)-GalNAc		1705.9	0.2%	0.2%
32	NeuAc-Gal-GlcNAc-(Gal-Fuc-GlcNAc)-Man	 and 	1722.9	0.3%	0.3%
33	NeuAc-Gal-Gal-(Gal-Fuc-GlcNAc)GalNAc				
34	NeuAc-Gal-GlcNAc-(NeuAc-Gal-GlcNAc)-Man		1910.0	0.7%	0.6%

CHAPTER 4

CONCLUSIONS

Next to the identification of proteins, studies implementing glycoproteomic techniques to focus on the analysis of post-translational modifications (PTMs) are becoming increasingly popular. Through the use of advanced mass spectrometry approaches we investigated the glycosylation of α-DG purified from rabbit skeletal muscle. Previous studies had reported α-DG to be modified by both *O*-GalNAc and *O*-mannose initiated glycan structures, the latter of which is rarely observed on mammalian proteins. Following chemical release, 4 *O*-mannose and 17 *O*-GalNAc initiated glycan structures from α-DG were identified using tandem mass spectrometry. Additionally, we developed a pseudo-neutral loss-triggered MS^3 method for the analysis of glycopeptides from α-DG to facilitate mapping sites of PTMs. By combining data from our analysis of released glycans and glycopeptides from α-DG we were able to identify 21 sites of modification.

Furthermore, we performed complete glycomic analysis of *O*-glycans released from mouse models of CMD. From our analysis we were able to characterize 9 *O*-mannose and 25 *O*-GalNAc initiated glycan structures from the mouse brain proteins of POMGnT1 -/-, LARGE -/-, and α-DG -/- and their equivalent wild type brains. From this study we were able to conclude that POMGnT1 is essential for the extension of *O*-mannose with a β-1,2 linked GlcNAc. Also through our analysis of *O*-glycans released from both LARGE +/+ and -/-, we observed 7 *O*-mannose structures and 22 *O*-GalNAc structures in addition to *O*-glycan structures previously observed from normal and $Large^{myd}$ mouse brains. From our analysis of *O*-glycan structures in the α-DG -/- brain, we were able to provide evidence for the presence of *O*-mannosyl glycans, confirming previously stated hypotheses that there are additional *O*-mannosylated proteins besides α-DG in mammalian systems.

Future experiments will implement similar glycoproteomic workflows discussed here allowing for the identification of additional *O*-mannosylated proteins from the mammalian brain using chemoenzymatic strategies. Following identification of *O*-mannosylated proteins, pseudo-neutral loss-triggered MS^3 method will be applied mapping modification sites. Additional experiments will need to be completed to expand upon collaborative work with Kevin Campbell's group at the University of Iowa. By continuing this collaboration attempts would be made to map any additional sites modified by the phosphorylated *O*-mannosyl glycan structure, identify the LARGE dependent modification, as well as determine the enzymes involved in the synthesis of this novel glycan structure. Work will also focus on the biological implications of *O*-glycan structures on that modify α-DG, and the overall effect that site occupancy of various glycans has on the post-translational processing of this protein.

APPENDIX A

O-MANNOSYL PHOSPHORYLATION OF ALPHA-DYSTROGLYCAN IS REQUIRED FOR LAMININ BINDING

CONTRIBUTION: Using MS^3 data dependent neutral loss methodologies I was able to map the attachment site of a novel phosphorylated O-mannosyl glycan structure on α-DG.

Yoshida-Moriguchi T, Yu L, Hammond S, Davis S, Kunz S, Madson M, Oldstone MBA, Schachter H, Wells L, Campbell KP. 2010. *Science*. 327:88-92.
Reprinted here with permission of publisher.

ABSTRACT

Alpha-dystroglycan (α-DG) is a cell-surface glycoprotein that acts as a receptor for both extracellular matrix proteins containing laminin-G domains and certain arenaviruses. Receptor binding is thought to be mediated by a posttranslational modification, and defective binding with laminin underlies a subclass of congenital muscular dystrophy. Using mass spectrometry– and nuclear magnetic resonance (NMR)–based structural analyses, we identified a phosphorylated O-mannosyl glycan on the mucin-like domain of recombinant α-DG, which was required for laminin binding. We demonstrated that patients with muscle-eye-brain disease and Fukuyama congenital muscular dystrophy, as well as mice with myodystrophy, commonly have defects in a postphosphoryl modification of this phosphorylated O-linked mannose, and that this modification is mediated by the like-acetylglucosaminyltransferase (LARGE) protein. These findings expand our understanding of the mechanisms that underlie congenital muscular dystrophy.

Diverse posttranslational modifications influence the structure and function of many proteins. Dystroglycan (DG) is a membrane protein that requires extensive posttranslational processing in order to function as an extracellularmatrix receptor. It is composed of an extracellular α-DG subunit and a transmembrane α–DG subunit[1]. α-DG serves as a receptor for extracellular matrix laminin G domain–containing ligands such as laminin[1] and agrin[2] in both muscle and brain, and these interactions depend on an unidentified posttranslational α-DG modification. α-DG is also the cellular receptor for lymphocytic choriomeningitis virus (LCMV), Lassa fever virus (LFV), and clade C New World arenaviruses[3, 4]. Although the binding sites for LCMV and LFV on α-DG have not yet been identified, they are thought to overlap with the modification recognized by laminin[5, 6].

Glycosyltransferase-mediated glycosylation is one type of posttranslational modification with two main forms in mammals: N- and O-glycosylation, which are distinguished by how the oligosaccharide moiety links to the amino acid. Mutations in six known or putative glycosyltransferase genes—*protein O-mannosyl transferase 1 (POMT1)*[7], *POMT2*[8], *protein O-mannose beta-1,2-N-acetylglucosaminyltransferase 1 (POMGnT1)*[9], *fukutin*[10], *fukutin-related protein (FKRP)*[11], and *LARGE*[12] —have been identified in patients with congenital muscular dystrophy (CMD). These disorders affect the brain, eye, and skeletal muscle to different extents, the most severe being Walker-Warburg syndrome [WWS; Online Mendelian Inheritance in Man (OMIM) identification number (ID) 236670], with less severe phenotypes seen in muscle-eye-brain disease (MEB; OMIM ID 253280) and Fukuyama CMD (FCMD; OMIM ID 253800). In these diseases, the ability of α-DG to bind laminin is markedly reduced[13], suggesting that these (putative) glycosyltransferases participate in the posttranslational modification that enables α-DG to bind laminin. Whereas the molecular functions of LARGE, fukutin, and FKRP remain unclear, POMT1 and -2[14] and POMGnT1[9] are known to catalyze two steps in the biosynthesis of an *O*-mannosyl tetrasaccharide (NeuNAc-α-2,3-Gal-β-1,4-GlcNAc-β-1,2-Man) that is found in high abundance on both brain and muscle α-DG[15, 16]. However, this glycan itself is probably not the laminin binding moiety, because glycosidase-mediated removal of the glycan does not reduce α-DG binding to laminin[17].

To determine which posttranslational modification is necessary for the α-DG /laminin interaction, we processed wheat germ agglutin enriched proteins (glycoproteins) from C57BL/6J (wild-type, WT) muscle using various enzymatic and chemical treatments. Treatment with cold aqueous hydrofluoric acid (HFaq), which specifically cleaves phosphoester linkages[18], resulted in the reduction of the α-DG relative molecular mass (Mr) from 150 to 70 kD, the loss of IIH6 immunoreactivity and laminin binding (Fig. A-1a), and the loss of binding to LFV and LCMV (Fig. A-1b). Because the *M*r of *N*-glycosylated β-DG did not change (Fig. A-1a), these effects were not caused by the degradation of either peptide or glycosyl linkages. A quantitative solid-phase assay revealed a 97% reduction in total high affinity binding to laminin (Fig. A-1c). HFaq treatment also abolished the laminin-receptor activity of α-DG in the heart, brain, and kidney. We next tested whether *N*-glycan and/or the two O-glycans known to modify the laminin-binding form of α-DG —Core1 *O*-glycan and the *O*-mannosyl tetrasaccharide (in either the sialylated or fucosylated form)[15, 16] —are sensitive to HFaq treatment. Immunoblotting of WT muscle glycoproteins treated with several cocktails of glycosidases that degrade these three glycans showed that the glycosidase-mediated reduction in α-DG glycosylation was impervious to HFaq treatment (Fig. A-1d). A similar experiment using muscle glycoproteins from the CMD-model mouse *LARGE*^{myd}, in which a mutation in LARGE prevents the α-DG modification necessary for laminin-binding[19], revealed that HFaq treatment did not significantly reduce the *M*r of α-DG (Fig. A-1d). Thus, HFaq specifically degrades the laminin-binding moiety on α-DG. Further, functional modification of α-DG appeared to involve an internal phosphoryl linkage rather than a monoester-linked phosphate, because digesting WT muscle glycoproteins with alkaline phosphatase did not reduce the laminin binding ability.

To verify that α-DG is phosphorylated, we labeled human embryonic kidney (HEK293) cells expressing Fc-tagged α-DG recombinants (Fig. A-2a) that are secreted into the medium with [32P]-orthophosphate. Phosphor imaging showed that secreted DGFc4, which contains only the mucin-like region of α-DG[20], was phosphorylated (Fig. A-2b). Hydrolysis of [32P]-DGFc4 under conditions that are conducive to the dissolution of polypeptide and phosphoester linkages to carbohydrates, but not to linkages to amino acids[21], generated inorganic phosphate but not phospho–amino acids, suggesting that

phosphorylation does not occur directly on the peptide. To test whether the phosphorylation depends on glycosylation, we expressed DGFc5 in [32P]-orthophosphate–labeled human cells derived from *POMT1*-mutated WWS, *POMGnT1*-mutated MEB, *fukutin*-mutated FCMD, and control cells, as well as in fibroblasts from *Large*^{myd} and WT mice (Fig. A-2c). All except the *POMT1*-mutated WWS cells secreted [32P]-phosphorylated DGFc5 into the medium, strongly suggesting that phosphorylation occurs on the *O*-linked mannose of α-DG. By measuring inorganic phosphate after acid hydrolysis, we confirmed that native α-DG purified from rabbit skeletal muscle is phosphorylated at 4.7 mol of phosphate per mole of protein (SD = 0.22, n = 3 trials).

To assess whether CMD cells that synthesize phosphorylated α-DG can further modify the phosphate residue, we immunoprecipitated glycoproteins from mouse *Large*^{myd} and WT muscle, as well as from human *POMGnT1*-mutated MEB, *fukutin*-mutated FCMD, and control muscle, by using immobilized metal affinity chromatography (IMAC) beads that bind to monoester-linked, but not diester-linked, phosphorylated compounds. Only *fukutin*-mutated FCMD and *Large*^{myd} muscle α-DG were captured by the beads, revealing that the phosphate residue does not undergo further modification in these CMD cells (Fig. A-2d). This finding suggests that fukutin and LARGE participate in a common pathway to assemble the laminin-binding moiety onto the phosphorylated *O*-linked mannose. This speculation is compatible with the fact that α-DG prepared from *Large*^{myd} muscle that was rescued by adenovirus-mediated expression of LARGE regains laminin-receptor activity and concomitantly loses its affinity for IMAC beads. In the case of *POMGnT1*-mutated MEB patient muscle, several forms of α-DG were observed; the majority of these were captured by the beads, although a certain amount of α-DG with laminin-binding activity was detected in the void fraction (Fig. A-2d). This finding suggests that a defect in POMGnT1 partially inhibits modification on the phosphoryl branch chain of the *O*-mannosyl glycan on α-DG.

DGFc4 that was produced by HEK293 cells bound to IMAC beads and gained laminin-binding activity when it was coexpressed with LARGE; such a gain in activity has also been observed in FCMD, MEB, and *Large*^{myd} cells, both in this study and elsewhere[22]. To determine the structure of the

phosphorylated *O*-mannosyl glycan that was necessary to assemble the laminin-binding moiety, we prepared *O*-glycans from HEK293-expressed DGFc4 by reductive β elimination, and we isolated the phosphorylated O-glycan using IMAC beads. Linear trap quadrupole (LTQ) mass spectrometry–based analyses detected prominent ions at mass-to-charge ratios *m/z* = 667 ($[M-H]^-$) and *m/z* = 333 ($[M-2H]^{2-}$) that are assigned as a phosphorylated trisaccharide composed of $HexNAc_2Hexitol_1$, and analysis by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) revealed the compositional sugars to be GlcNAc, GalNAc, and mannitol. Homo- and heteronuclear NMR techniques were used to assign the $^{13}C/^{1}H$ heteronuclear multiple quantum coherence (HMQC) spectrum of the reduced *O*-glycan (Fig. A-3A). The GlcNAc (subunit B) was assigned using double quantum-filtered correlation spectroscopy (DQF-COSY) and total correlation spectroscopy (TOCSY) spectra with a series of mixing times. The GalNAc (subunit C) was partially assigned based on a selective TOCSYHSQC spectrum (Fig. A-3b). The GalNAc (subunit C) is linked via a β1-3 linkage to the C3 position on GlcNAc (subunit B), which is in turn connected via a β1-4 linkage to the C4 position on mannitol (subunit A), as evidenced by the observed heteronuclear multiple-bond correlation (HMBC) cross peaks CH1/BC3 and BH1/AC4 (Fig. A-3c). The phosphate group is attached to the C6 position of the mannitol (subunit A) as determined from the cross peaks $^{31}P/AH6'$ and $^{31}P/A6H''$ detected in the $^{31}P/^{1}H$ COSY spectrum (Fig. A-3d). The complete NMR resonance assignments of the reduced *O*-glycan and its inter-residue correlations detected in the nuclear Overhauser effect spectroscopy (NOESY) and HMBC spectra are summarized (data not shown), and the determined structure is shown in Fig. A-3e. To verify that this phosphorylated trisaccharide modifies the mucin-like domain of DGFc4, we enriched the trypsinized peptides using *Wisteria floribunda* agglutinin-lectin and analyzed the GalNAc-terminated peptides by liquid chromatography–mass spectrometry (LC-MS)/MS. MS/MS fragmentation patterns at *m/z* = 1318.63 (Fig. A-4a), 1420.17 (fig. A-S2), and 1501.19 (fig. A-S3) identified a peptide (amino acids 374 to 389 of α-DG; GenBank ID CAA45732) bearing these modifications: the phosphorylated trisaccharide in conjunction with Hex-HexNAc-Hex, HexNAc- Hex, or Hex. The presence of nonphosphorylated mannose-initiated structures on y3, y4, and y10 ions revealed that Thr^{379} is modified

by the phosphorylated trisaccharide in all cases (Fig. A-4, a and b, and figs. A-S2 and A-S3). Additional studies showed that α-DG is phosphorylated within the Golgi complex and that this phosphorylation occurs independently from the mannose-6-phosphate synthetic pathway that is required for lysosomal protein modification[23]; fibroblasts derived from patients with mucolipidosis II (OMIMID 252500), which have a defect in GlcNAc-1-phosphotransferase, can synthesize the laminin-binding form of α-DG (Fig. A-4c).

We demonstrated that MEB, FCMD, and *Large*^{myd} cells, which have genetically distinct abnormalities, show a similar defect in postphosphoryl modification on the *O*-mannosyl glycan. These convergent mechanisms to pathology offer an explanation for previous reports that forced expression of LARGE can circumvent defects in α-DG modification in these CMD cells[22]. We speculate that LARGE, a putative glycosyltransferase with catalytic domains sharing homology with β-1,3-N-acetylglucosaminyltransferase and bacterial glycosyltransferase[19] participates in postphosphoryl glycosylation, because the forced expression increases the affinity of the cell surface for both the IIH6 antibody and the *Vicia villosa* lectin. To our knowledge, we provide the first evidence that a vertebrate non– glycosylphosphatidylinositol- anchored glycoprotein is modified by a phosphodiester linkage. Glycoproteins in the cell walls of yeasts and fungi bear phosphodiester-linked glycans that are generated by a process involving phosphorylation on the C6 hydroxyl of mannose[24]. α-DG, which is well conserved as an epithelial cell-surface protein in species ranging from lower vertebrates to mammals, is likewise modified by this ancient type of glycosylation. A recent study has shown that the most severe form of CMD—WWS—is a genetically heterogeneous disease. Moreover, only 40% of WWS cases are explained by mutations in known CMD-causative genes[25]. Thus, a defect in the phosphorylation of an *O*-linked mannose may be responsible for severe CMD, indicating that the discovery of mutations in new genes responsible for WWS may not be far off.

REFERENCES

1. Ibraghimov-Beskrovnaya, O.; Ervasti, J. M.; Leveille, C. J.; Slaughter, C. A.; Sernett, S. W.; Campbell, K. P., Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature* **1992,** 355, (6362), 696-702.

2. Gee, S. H.; Montanaro, F.; Lindenbaum, M. H.; Carbonetto, S., Dystroglycan-alpha, a dystrophin-associated glycoprotein, is a functional agrin receptor. *Cell* **1994,** 77, (5), 675-86.

3. Cao, W.; Henry, M. D.; Borrow, P.; Yamada, H.; Elder, J. H.; Ravkov, E. V.; Nichol, S. T.; Compans, R. W.; Campbell, K. P.; Oldstone, M. B., Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus. *Science* **1998,** 282, (5396), 2079-81.

4. Spiropoulou, C. F.; Kunz, S.; Rollin, P. E.; Campbell, K. P.; Oldstone, M. B., New World arenavirus clade C, but not clade A and B viruses, utilizes alpha-dystroglycan as its major receptor. *J Virol* **2002,** 76, (10), 5140-6.

5. Kunz, S.; Sevilla, N.; McGavern, D. B.; Campbell, K. P.; Oldstone, M. B., Molecular analysis of the interaction of LCMV with its cellular receptor [alpha]-dystroglycan. *J Cell Biol* **2001,** 155, (2), 301-10.

6. Kunz, S.; Rojek, J. M.; Perez, M.; Spiropoulou, C. F.; Oldstone, M. B., Characterization of the interaction of lassa fever virus with its cellular receptor alpha-dystroglycan. *J Virol* **2005,** 79, (10), 5979-87.

7. Beltran-Valero de Bernabe, D.; Currier, S.; Steinbrecher, A.; Celli, J.; van Beusekom, E.; van der Zwaag, B.; Kayserili, H.; Merlini, L.; Chitayat, D.; Dobyns, W. B.; Cormand, B.; Lehesjoki, A. E.; Cruces, J.; Voit, T.; Walsh, C. A.; van Bokhoven, H.; Brunner, H. G., Mutations in the O-mannosyltransferase gene POMT1 give rise to the severe neuronal migration disorder Walker-Warburg syndrome. *Am J Hum Genet* **2002,** 71, (5), 1033-43.

8. 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.; Huynen, M. A.; Verrips, A.; Walsh, C. A.; Barth, P. G.; Brunner, H. G.; van Bokhoven, H., POMT2 mutations cause alpha-dystroglycan hypoglycosylation and Walker-Warburg syndrome. *J Med Genet* **2005,** 42, (12), 907-12.

9. Yoshida, A.; Kobayashi, K.; Manya, H.; Taniguchi, K.; Kano, H.; Mizuno, M.; Inazu, T.; Mitsuhashi, H.; Takahashi, S.; Takeuchi, M.; Herrmann, R.; Straub, V.; Talim, B.; Voit, T.; Topaloglu, H.; Toda, T.; Endo, T., Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. *Dev Cell* **2001,** 1, (5), 717-24.

10. Kobayashi, K.; Nakahori, Y.; Miyake, M.; Matsumura, K.; Kondo-Iida, E.; Nomura, Y.; Segawa, M.; Yoshioka, M.; Saito, K.; Osawa, M.; Hamano, K.; Sakakihara, Y.; Nonaka, I.; Nakagome, Y.; Kanazawa, I.; Nakamura, Y.; Tokunaga, K.; Toda, T., An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy. *Nature* **1998,** 394, (6691), 388-92.

11. Brockington, M.; Yuva, Y.; Prandini, P.; Brown, S. C.; Torelli, S.; Benson, M. A.; Herrmann, R.; Anderson, L. V.; Bashir, R.; Burgunder, J. M.; Fallet, S.; Romero, N.; Fardeau, M.; Straub, V.; Storey, G.; Pollitt, C.; Richard, I.; Sewry, C. A.; Bushby, K.; Voit, T.; Blake, D. J.; Muntoni, F., Mutations in the fukutin-related protein gene (FKRP) identify limb girdle muscular dystrophy 2I as a milder allelic variant of congenital muscular dystrophy MDC1C. *Hum Mol Genet* **2001,** 10, (25), 2851-9.

12. Longman, C.; Brockington, M.; Torelli, S.; Jimenez-Mallebrera, C.; Kennedy, C.; Khalil, N.; Feng, L.; Saran, R. K.; Voit, T.; Merlini, L.; Sewry, C. A.; Brown, S. C.; Muntoni, F., Mutations in the human LARGE gene cause MDC1D, a novel form of congenital muscular dystrophy with severe mental retardation and abnormal glycosylation of alpha-dystroglycan. *Hum Mol Genet* **2003,** 12, (21), 2853-61.

13. Michele, D. E.; Barresi, R.; Kanagawa, M.; Saito, F.; Cohn, R. D.; Satz, J. S.; Dollar, J.; Nishino, I.; Kelley, R. I.; Somer, H.; Straub, V.; Mathews, K. D.; Moore, S. A.; Campbell, K. P., Post-translational disruption of dystroglycan-ligand interactions in congenital muscular dystrophies. *Nature* **2002,** 418, (6896), 417-22.

14. Manya, H.; Chiba, A.; Yoshida, A.; Wang, X.; Chiba, Y.; Jigami, Y.; Margolis, R. U.; Endo, T., Demonstration of mammalian protein O-mannosyltransferase activity: coexpression of POMT1 and POMT2 required for enzymatic activity. *Proc Natl Acad Sci U S A* **2004,** 101, (2), 500-5.

15. Sasaki, T.; Yamada, H.; Matsumura, K.; Shimizu, T.; Kobata, A.; Endo, T., Detection of O-mannosyl glycans in rabbit skeletal muscle alpha-dystroglycan. *Biochim Biophys Acta* **1998,** 1425, (3), 599-606.

16. Chiba, A.; Matsumura, K.; Yamada, H.; Inazu, T.; Shimizu, T.; Kusunoki, S.; Kanazawa, I.; Kobata, A.; Endo, T., Structures of sialylated O-linked oligosaccharides of bovine peripheral nerve alpha-dystroglycan. The role of a novel O-mannosyl-type oligosaccharide in the binding of alpha-dystroglycan with laminin. *J Biol Chem* **1997,** 272, (4), 2156-62.

17. Combs, A. C.; Ervasti, J. M., Enhanced laminin binding by alpha-dystroglycan after enzymatic deglycosylation. *Biochem J* **2005,** 390, (Pt 1), 303-9.

18. Ilg, T.; Stierhof, Y. D.; Craik, D.; Simpson, R.; Handman, E.; Bacic, A., Purification and structural characterization of a filamentous, mucin-like proteophosphoglycan secreted by Leishmania parasites. *J Biol Chem* **1996,** 271, (35), 21583-96.

19. Grewal, P. K.; Holzfeind, P. J.; Bittner, R. E.; Hewitt, J. E., Mutant glycosyltransferase and altered glycosylation of alpha-dystroglycan in the myodystrophy mouse. *Nat Genet* **2001,** 28, (2), 151-4.

20. Kanagawa, M.; Saito, F.; Kunz, S.; Yoshida-Moriguchi, T.; Barresi, R.; Kobayashi, Y. M.; Muschler, J.; Dumanski, J. P.; Michele, D. E.; Oldstone, M. B.; Campbell, K. P., Molecular recognition by LARGE is essential for expression of functional dystroglycan. *Cell* **2004,** 117, (7), 953-64.

21. Ilg, T.; Overath, P.; Ferguson, M. A.; Rutherford, T.; Campbell, D. G.; McConville, M. J., O- and N-glycosylation of the Leishmania mexicana-secreted acid phosphatase. Characterization of a new class of phosphoserine-linked glycans. *J Biol Chem* **1994,** 269, (39), 24073-81.

22. Barresi, R.; Michele, D. E.; Kanagawa, M.; Harper, H. A.; Dovico, S. A.; Satz, J. S.; Moore, S. A.; Zhang, W.; Schachter, H.; Dumanski, J. P.; Cohn, R. D.; Nishino, I.; Campbell, K. P., LARGE can functionally bypass alpha-dystroglycan glycosylation defects in distinct congenital muscular dystrophies. *Nat Med* **2004,** 10, (7), 696-703.

23. Dahms, N. M.; Lobel, P.; Kornfeld, S., Mannose 6-phosphate receptors and lysosomal enzyme targeting. *J Biol Chem* **1989,** 264, (21), 12115-8.

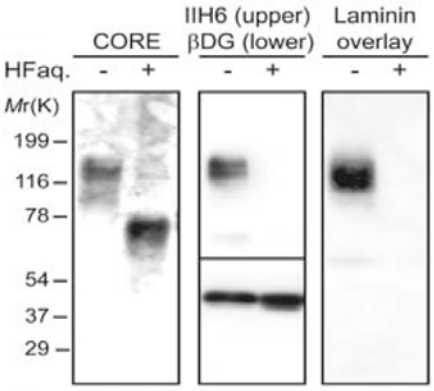
24. Gander, J. E., Fungal cell wall glycoproteins and peptido-polysaccharides. *Annu Rev Microbiol* **1974,** 28, (0), 103-19.

25. Manzini, M. C.; Gleason, D.; Chang, B. S.; Hill, R. S.; Barry, B. J.; Partlow, J. N.; Poduri, A.; Currier, S.; Galvin-Parton, P.; Shapiro, L. R.; Schmidt, K.; Davis, J. G.; Basel-Vanagaite, L.; Seidahmed, M. Z.; Salih, M. A.; Dobyns, W. B.; Walsh, C. A., Ethnically diverse causes of Walker-Warburg syndrome (WWS): FCMD mutations are a more common cause of WWS outside of the Middle East. *Hum Mutat* **2008,** 29, (11), E231-41.

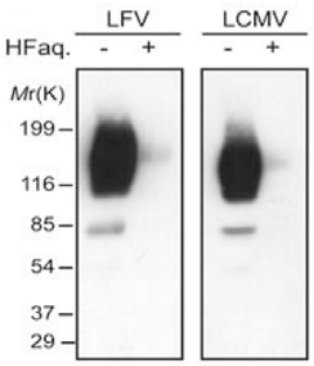
Figure. A-1. Chemical dephosphorylation by HFaq treatment abolishes laminin and virus binding to α-DG in tissues from WT mice.

(**a** and **b**) Treated glycoproteins prepared from WT muscle were subjected to the following: (**a**) immunoblotting with antibodies against the α-DG core protein (CORE) or the laminin-binding form α-DG epitope (IIH6) and β-DG, or to laminin overlay assay; (**b**) virus overlay assays with γ-inactivated LFV or LCMV cl-13. (**c**) WT muscle glycoproteins with and without HFaq treatment were subjected to a solid-phase laminin-binding assay (n = 3). Open circles, treated; solid circles, untreated. Error bars indicate SD. (**d**) Muscle glycoproteins from WT and *Large*^{myd} (Myd) mice were digested with cocktails of glycosidases that degrade sialylated and/or fucosylated *N*-glycan, Core 1 *O*-glycan, and *O*-mannosyl glycan, before (first four lanes) and after (last four lanes) HFaq treatment. The products were subjected to either immunoblotting with CORE antibody or laminin overlay assay.

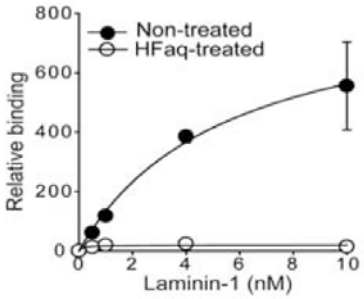
a)


CORE
IIH6 (upper)
βDG (lower)
Laminin
overlay
HFaq.
- + - + - +
Mr(K)
199
116
78
54
37
29

b)


LFV
LCMV
HFaq.
- + - +
Mr(K)
199
116
85
54
37
29

c)


Non-treated
HFaq-treated
Relative binding
800
600
400
200
0
0
2
4
6
8
10
Laminin-1 (nM)

d)

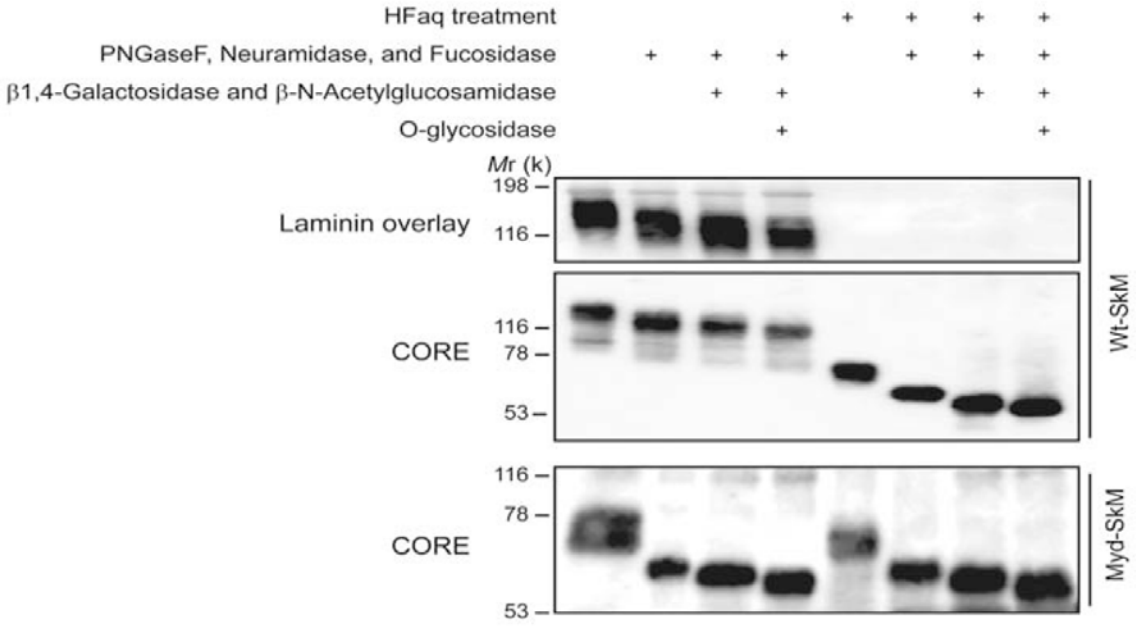
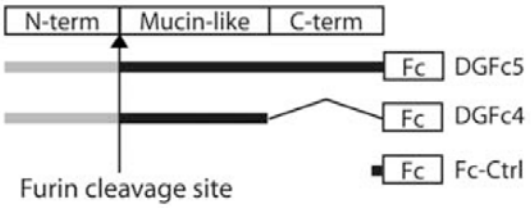

HFaq treatment
PNGaseF, Neuramidase, and Fucosidase
β1,4-Galactosidase and β-N-Acetylglucosamidase
O-glycosidase
Mr (k)
198
116
Laminin overlay
116
78
53
CORE
Wt-SkM
116
78
53
CORE
Myd-SkM

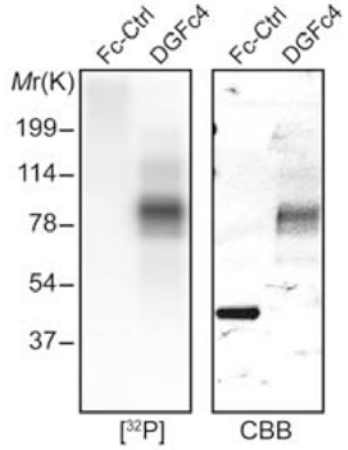
Figure A-2. The mucin-like domain of α-DG is phosphorylated in an *O*-mannosylation–dependent manner.

(**a**) Structures of recombinant α-DG constructs used in the study. (**b** and **c**) [^{32}P]-orthophosphate labeling of (**b**) Fc-Ctrl– or DGFc4- expressing HEK293 cells and (**c**) DGFc5-expressing cultured cells from CMD patients (WWS, MEB, or FCMD) and control humans, and from WT and *Large*^{myd} (Myd) mice. Fc-tagged recombinant α-DG was isolated from the culture medium with protein-A agarose, separated by SDS–polyacrylamide gel electrophoresis, stained with Coomassie brilliant blue (CBB), and analyzed by phosphorimaging ([^{32}P]). Phosphorylation of α-DG required prior *O*-mannosylation. Asterisks indicate contaminating proteins derived from fetal bovine serum. (**d**) IMAC-binding assay testing glycoproteins from WT and *Large*^{myd} mice, and from FCMD, MEB, and control human muscle (SkM).

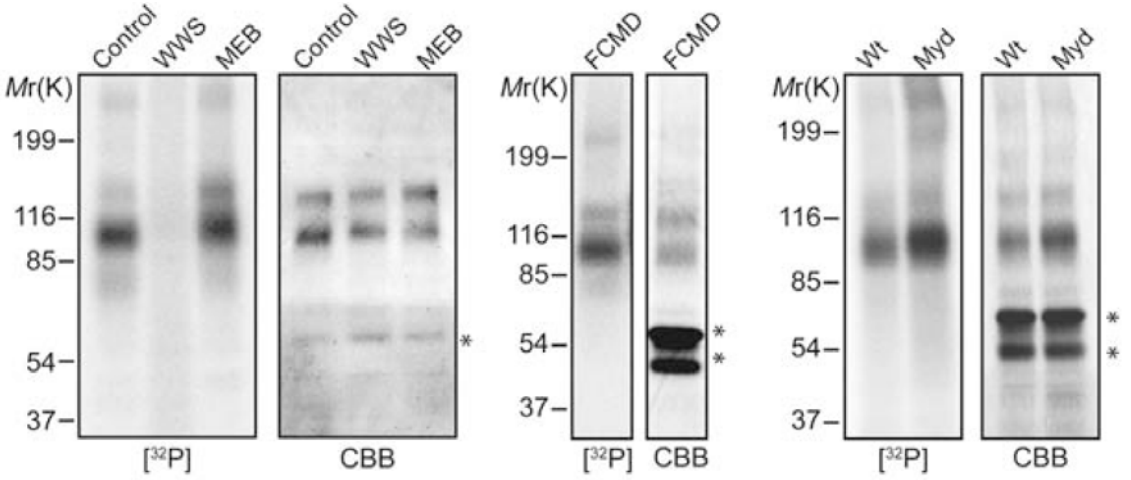
a)


N-term
Mucin-like
C-term
Fc
DGFc5
Fc
DGFc4
Fc
Fc-Ctrl
Furin cleavage site

b)


Fc-Ctrl
DGFc4
Fc-Ctrl
DGFc4
Mr(K)
199
114
78
54
37
[32P]
CBB

c)


Control
WWS
MEB
Control
WWS
MEB
Mr(K)
199
116
85
54
37
*
[32P]
CBB
FCMD
FCMD
Mr(K)
199
116
85
54
37
*
*
[32P]
CBB
Wt
Myd
Wt
Myd
Mr(K)
199
116
85
54
37
*
*
[32P]
CBB

d)

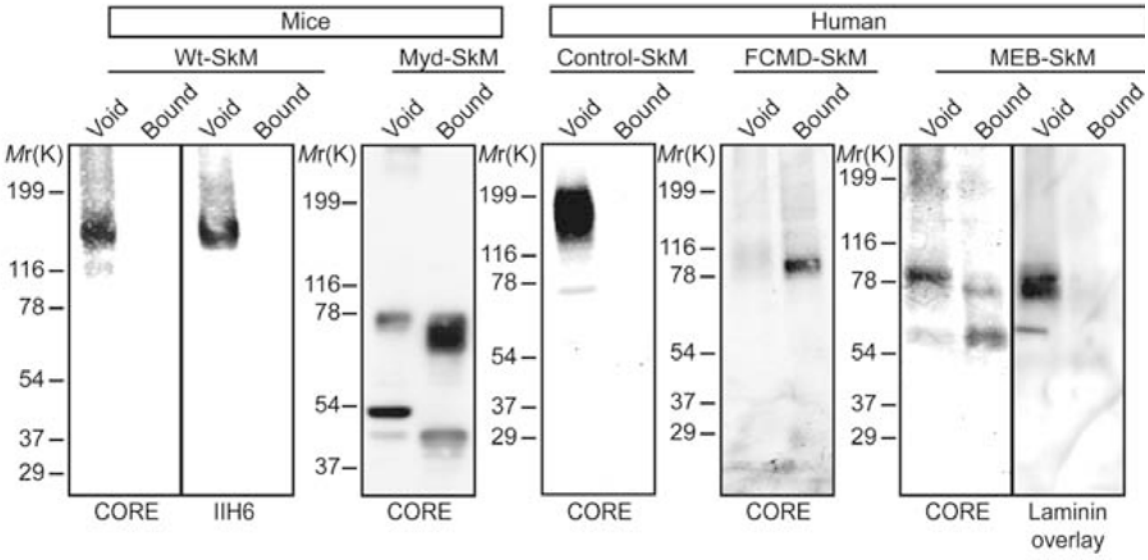

Mice
Human
Wt-SkM
Myd-SkM
Control-SkM
FCMD-SkM
MEB-SkM
Void
Bound
Mr(K)
199
116
78
54
37
29
CORE
IIH6
Laminin
overlay

Figure A-3. NMR analysis of phosphorylated O-glycan on HEK293-produced DGFc4.
(**a**) HMQC spectrum where the assigned cross peaks are labeled with a letter for the subunit designated in (**e**) and a number for the position on that subunit. The folded cross peaks are indicated in blue, and the cross peaks derived from sample impurities are marked by asterisks. (**b**) TOCSY-HSQC spectrum obtained using a selective excitation pulse at the subunit CH1 proton and a selective TOCSY mixing time of 113 ms. (**c**) HMBC (red) and HMQC (black and blue) spectra for the assignment of interglycoside linkages. (**d**) $^{31}P/^{1}H$ COSY spectrum. (**e**) Structure of the O-glycan, with the sugar subunits labeled A to C. ppm, parts per million.

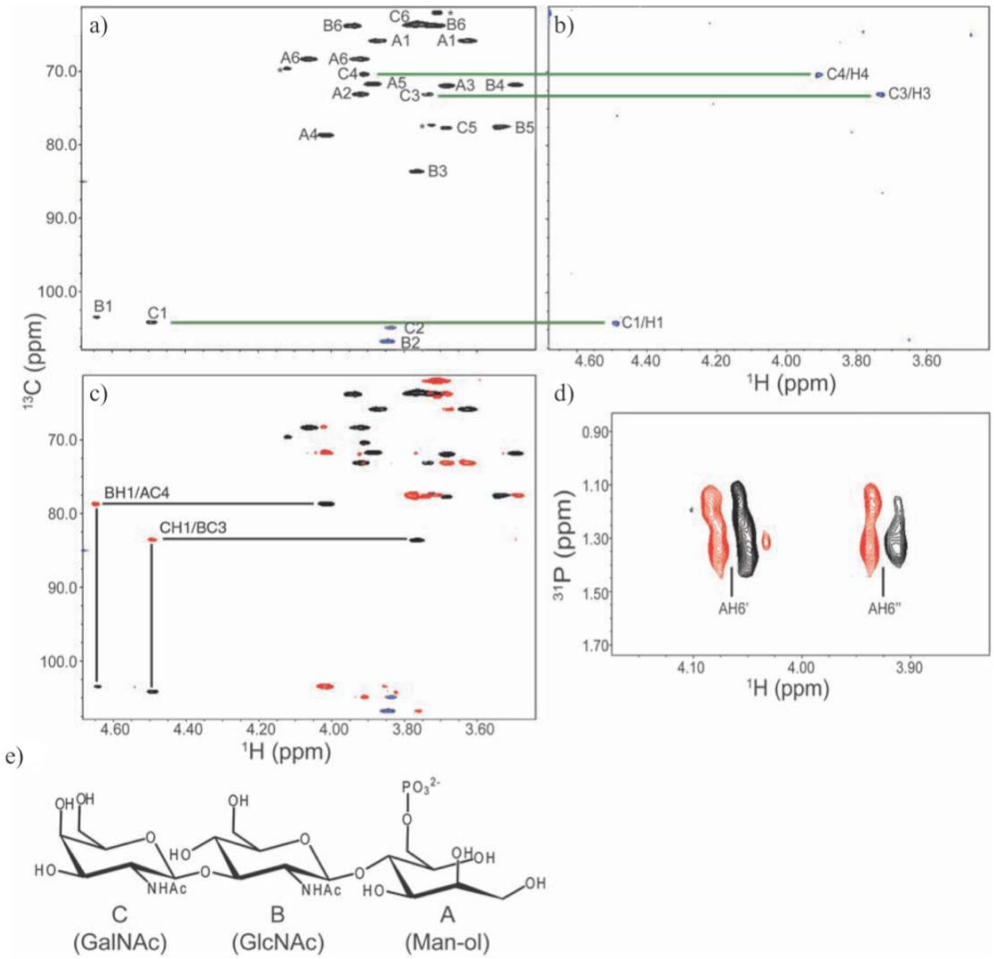
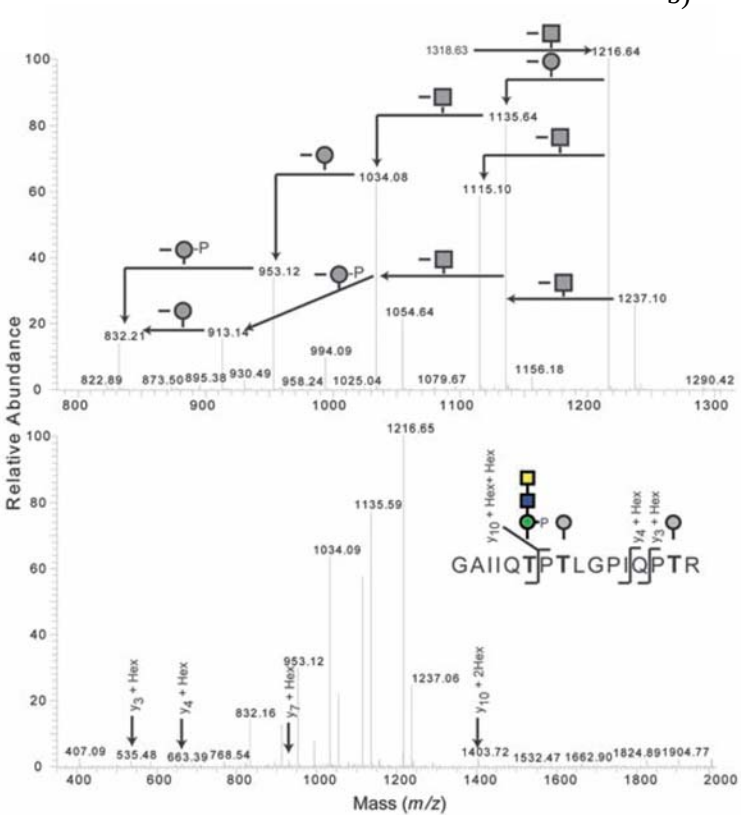
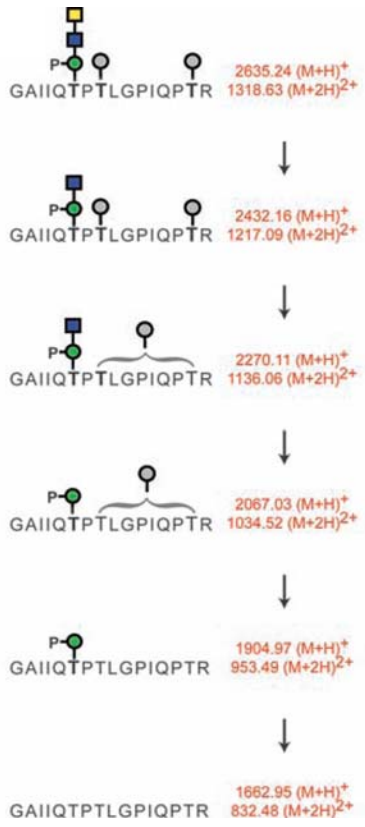

a)
b)
c)
d)
e)
^{13}C (ppm)
^{1}H (ppm)
^{31}P (ppm)
B6
C6
A1
A6
C4
A5
A3
B4
A2
C3
C5
B5
A4
B3
B1
C1
C2
B2
C4/H4
C3/H3
C1/H1
BH1/AC4
CH1/BC3
AH6'
AH6''
NHAc
C
(GalNAc)
B
(GlcNAc)
A
(Man-ol)

Figure A-4. Mapping of phosphorylated trisaccharide on HEK293-produced DGFc4 and characterization of mannosyl phosphorylation.
(a) Collision-induced dissociation (CID)–MS/MS spectra from 780 to 1320 *m/z* (upper panel) and 375 to 2000 *m/z* (lower panel), revealing a neutral loss pattern (upper panel) and peptide-derived b and y ions (lower panel) of the selected precursor ions at *m/z* = 1318.63. The full Fourier transform mass spectrum is shown in fig. S8. Squares, HexNAc; circles, hexose. (b) Peptide and monosaccharide unit identification based on fragmentation of the phosphorylated glycopeptide. (c) Glycoproteins prepared from cell lysates of fibroblasts derived from mucolipidosis II patients were subjected to immunoblotting with CORE antibody and laminin overlay assay.

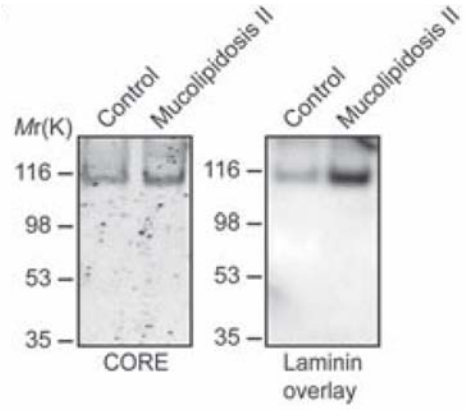
a)


Relative Abundance
Mass (m/z)
GAIIQTPTLGPIQPTR
1318.63
1216.64
1135.64
1034.08
1115.10
953.12
1237.10
1054.64
832.21
913.14
994.09
1216.65
1135.59
1034.09
1237.06
y3 + Hex
y4 + Hex
y7 + Hex
y10 + 2Hex
y10 + Hex+ Hex

b)


GAIIQTPTLGPIQPTR
2635.24 (M+H)+
1318.63 (M+2H)2+
2432.16 (M+H)+
1217.09 (M+2H)2+
2270.11 (M+H)+
1136.06 (M+2H)2+
2067.03 (M+H)+
1034.52 (M+2H)2+
1904.97 (M+H)+
953.49 (M+2H)2+
1662.95 (M+H)+
832.48 (M+2H)2+

c)


Control
Mucolipidosis II
Mr(K)
116
98
53
35
CORE
Control
Mucolipidosis II
116
98
53
35
Laminin
overlay

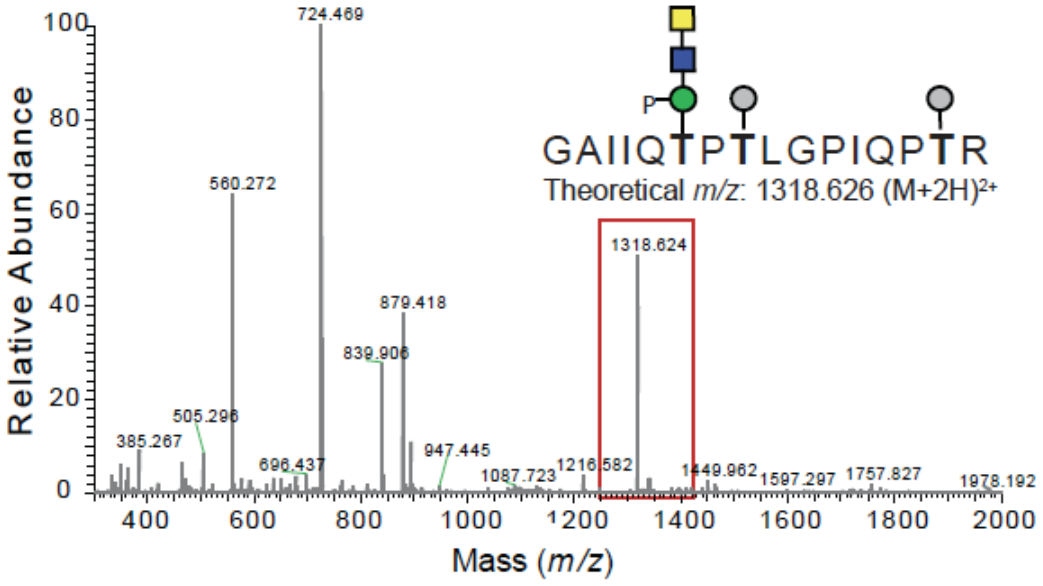
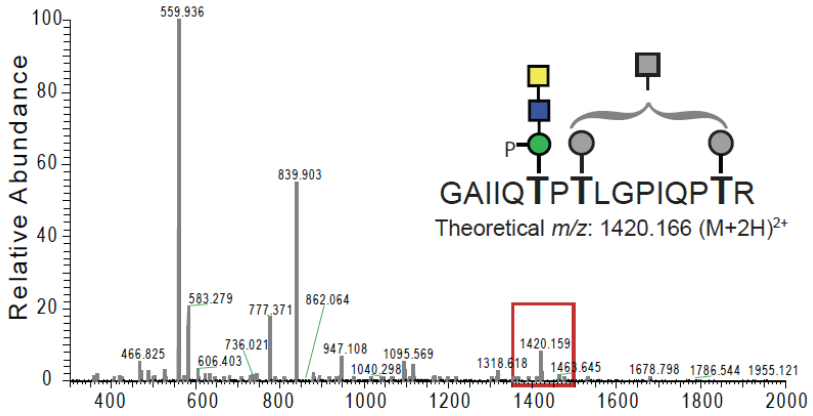


Figure A-S1. The full FT mass spectrum of the peptide (*m/z* 1318.624) modified by the phosphorylated trisaccharide.

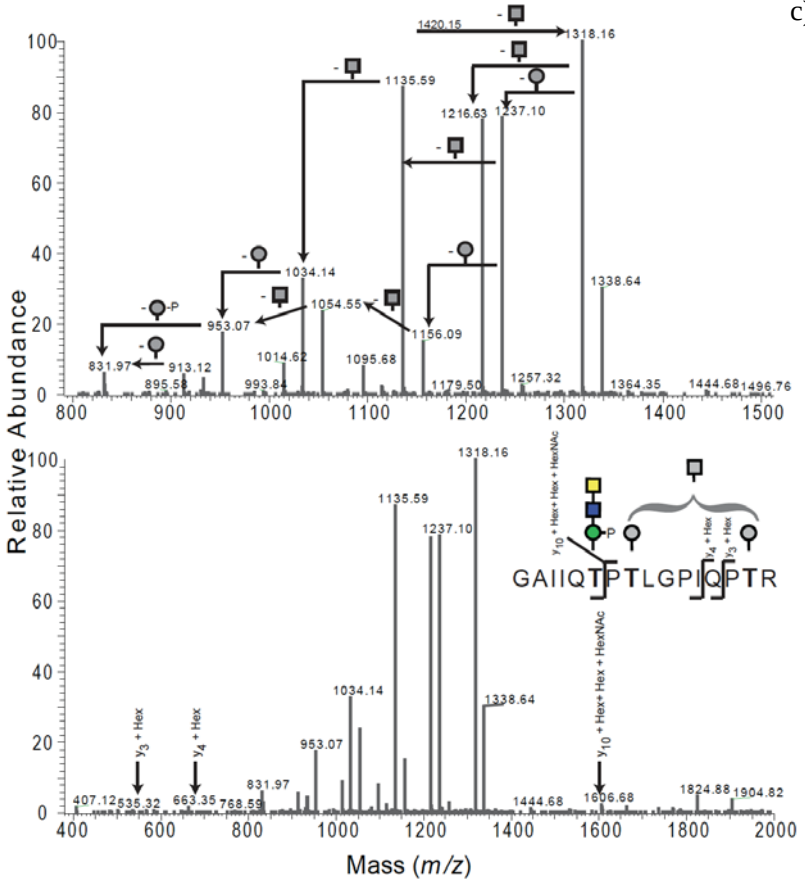
The full mass spectrum yielded a mass accuracy of 1.4 ppm for the precursor ions at *m/z* 1318.624. The CID-MS/MS spectrum is shown in Figure A-4.

Figure A-S2. Mass analysis of the peptide modified by phosphorylated trisaccharide (*m/z* 1420.166).
(a) The full FT mass spectrum of the peptide containing the phosphorylated trisaccharide. The full mass spectrum yielded a mass accuracy of 4.4 ppm for the precursor ion at *m/z* 1420.166. **(b)** The CID-MS/MS spectra from 780-1320 *m/z* (upper panel) and 375-2000 *m/z* (lower panel) show the neutral loss pattern (upper panel) and the peptide-derived b and y ions (lower panel) of the selected precursor ions at *m/z* 1420.17. Square: HexNAc; circle: Hexose. **(c)** The fragmentation of the phosphorylated glycopeptides allowed for identification of the peptide and mononsaccharide units as illustrated. The observation of ions corresponding to the non-phosphorylated hexoses on Thr_{381} and Thr_{388} revealed that Thr_{379} was modified by the phosphorylated trisaccharide and that Thr_{381} and Thr_{388} were occupied by a combination of HexNAc-Hex and Hex in this peptide.

a)


Relative Abundance
GAIIQTPTLGPIQPTR
Theoretical m/z: 1420.166 (M+2H)2+
559.936
839.903
583.279
777.371
862.064
736.021
947.108
1095.569
1420.159
466.825
606.403
1040.298
1318.618
1463.645
1678.798
1786.544
1955.121
Mass (m/z)

b)


Relative Abundance
Mass (m/z)
GAIIQTPTLGPIQPTR
1420.15
1318.16
1135.59
1237.10
1216.63
1034.14
1338.64
1054.55
953.07
1156.09
831.97
913.12
1014.62
1095.68
1179.50
1257.32
1364.35
1444.68
1496.76
y3 + Hex
y4 + Hex
y10 + Hex + HexNAc

c)

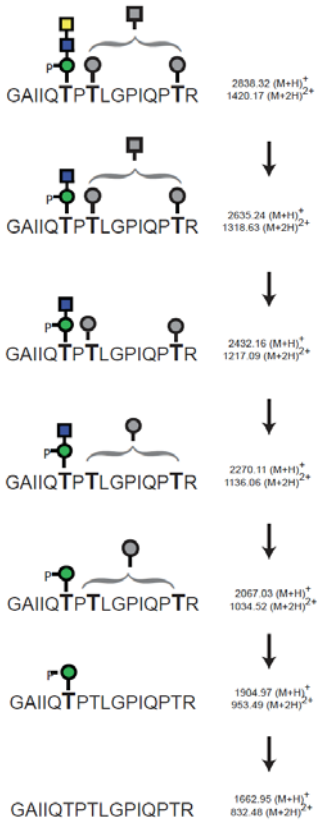
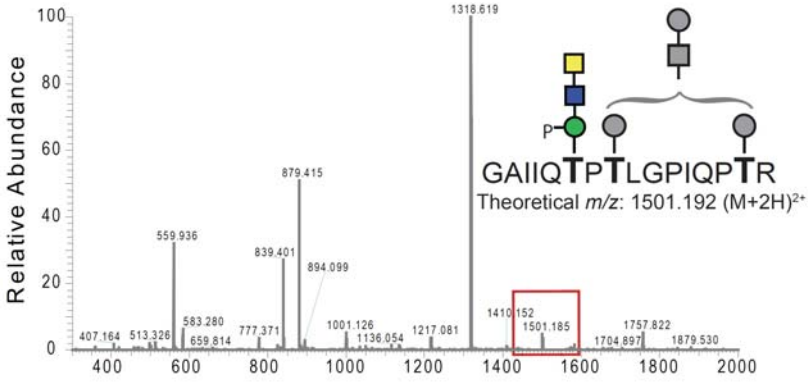

GAIIQTPTLGPIQPTR
2838.32 (M+H)+
1420.17 (M+2H)2+
2635.24 (M+H)+
1318.63 (M+2H)2+
2432.16 (M+H)+
1217.09 (M+2H)2+
2270.11 (M+H)+
1136.06 (M+2H)2+
2067.03 (M+H)+
1034.52 (M+2H)2+
GAIIQTPTLGPIQPTR
1904.97 (M+H)+
953.49 (M+2H)2+
GAIIQTPTLGPIQPTR
1662.95 (M+H)+
832.48 (M+2H)2+

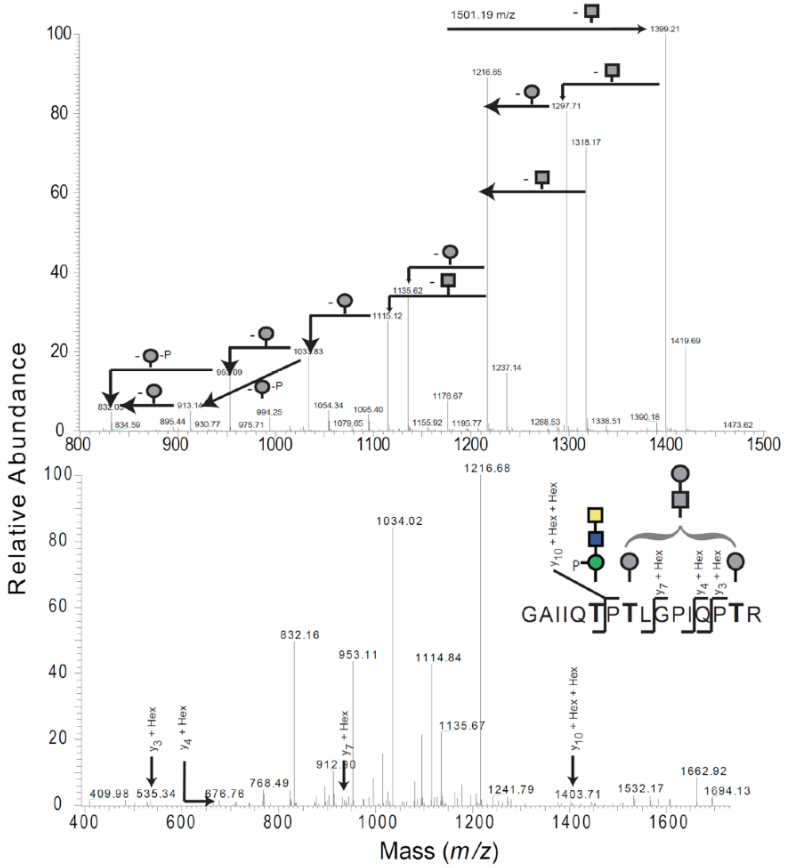
Figure A-S3. Mass analysis of the peptide modified by phosphorylated trisaccharide (*m/z* 1501.192).

(a) The full FT mass spectrum of the peptide containing the phosphorylated trisaccharide. The full mass spectrum yielded a mass accuracy of 4.5 ppm for the precursor ions at *m/z* 1501.192. **(b)** The CID-MS/MS spectra from 780-1320 *m/z* (upper panel) and 375-2000 *m/z* (lower panel) show the neutral loss pattern (upper panel) and the peptide-derived b and y ions (lower panel) of the selected precursor ions at *m/z* 1501.19. Square: HexNAc; circle: Hexose. **(c)** The fragmentation of the phosphorylated glycopeptide allowed for identification of the peptide and mononsaccharide units as illustrated. The observation of ions corresponding to the non-phosphorylated hexoses on Thr_{381} and Thr_{388} revealed that Thr_{379} was modified by the phosphorylated trisaccharide and that Thr_{381} and Thr_{388} were occupied by a combination of Hex-HexNAc-Hex and Hex in this peptide. The Hex-HexNAc-Hex is likely to be Gal-β-1,4-GlcNAc-β-1,2-Man, which was previously shown to modify native and recombinant α-DG in brain, muscle and HEK293 cells.

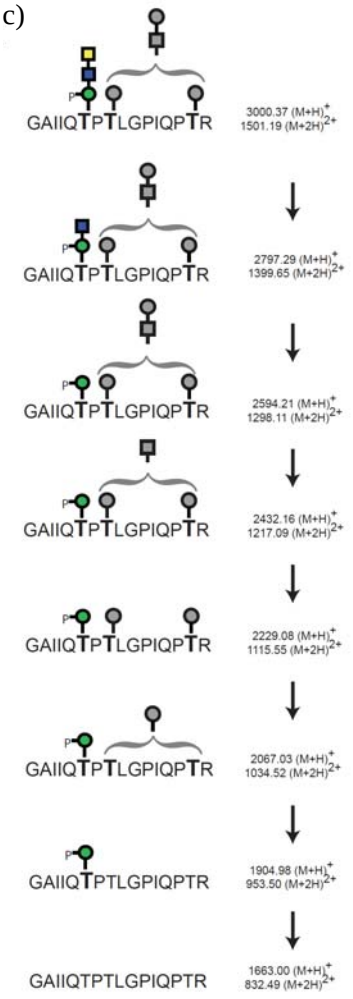
a)


Relative Abundance
GAIIQTPTLGPIQPTR
Theoretical m/z: 1501.192 (M+2H)2+

b)


Relative Abundance
Mass (m/z)
1501.19 m/z
GAIIQTPTLGPIQPTR

c)


GAIIQTPTLGPIQPTR
3000.37 (M+H)+
1501.19 (M+2H)2+
2797.29 (M+H)+
1399.65 (M+2H)2+
2594.21 (M+H)+
1298.11 (M+2H)2+
2432.16 (M+H)+
1217.09 (M+2H)2+
2229.08 (M+H)+
1115.55 (M+2H)2+
2067.03 (M+H)+
1034.52 (M+2H)2+
GAIIQTPTLGPIQPTR
1904.98 (M+H)+
953.50 (M+2H)2+
GAIIQTPTLGPIQPTR
1663.00 (M+H)+
832.49 (M+2H)2+

APPENDIX B

DROSOPHILA DYSTROGLYCAN IS A TARGET OF *O*-MANNOSYLTRANSFERASE ACTIVITY

OF TWO PROTEIN *O*-MANNOSYLTRANSFERASES, ROTATED ABDOMEN AND TWISTED

CONTRIBUTION: Used MS^3 methodologies to identify glycopeptides and map sites of post-translational modification on DG.

Nakamura N, Hammond S, Lyalin D, Lavrova O, Wells L, Panin VM
2010. *Glycobiology*. 20(3): 381–394.
Reprinted here with permission of publisher.

ABSTRACT

Recent studies highlighted an emerging possibility of using *Drosophila* as a model system for investigating the mechanisms of human congenital muscular dystrophies, called dystroglycanopathies, resulting from the abnormal glycosylation of α-dystroglycan. Several of these diseases are associated with defects in *O*-mannosylation, one of the most prominent types of α-dystroglycan glycosylation mediated by two protein *O*-mannosyltransferases. *Drosophila* appears to possess homologs of all essential components of the mammalian dystroglycan-mediated pathway; however, the glycosylation of *Drosophila* Dystroglycan (DG) has not yet been explored. In this study, we characterized the glycosylation of *Drosophila* DG using a combination of glycosidase treatments, lectin blots, trypsin digestion, and mass spectrometry analyses. Our results demonstrated that DG extracellular domain is *O*-mannosylated in vivo. We found that the concurrent in vivo activity of the two *Drosophila* protein *O*-mannosyltransferases, Rotated Abdomen and Twisted, is required for *O*-mannosylation of DG. While our experiments unambiguously determined some *O*-mannose sites far outside of the mucin-type domain of DG, they also provided evidence that DG bears a significant amount of *O*-mannosylation within its central region including the mucin-type domain, and that *O*-mannose can compete with *O*-GalNAc glycosylation of DG. We found that Rotated Abdomen and Twisted could potentiate in vivo the dominant-negative effect of DG extracellular domain expression on crossvein development, which suggests that *O*-mannosylation can modulate the ligand-binding activity of DG. Taken together these results demonstrated that *O*-mannosylation of Dystroglycan is an evolutionarily ancient mechanism conserved between *Drosophila* and humans, suggesting that *Drosophila* can be a suitable model system for studying molecular and genetic mechanisms underlying human dystroglycanopathies.

INTRODUCTION

Dystroglycan, a highly glycosylated protein of mammalian muscle cells, is a central component of the dystrophin–glycoprotein complex (DGC) that provides structural stability to the sarcolemma during muscle contraction. Mammalian dystroglycan undergoes posttranslational cleavage into separate α- and β-subunits[1]. Proper glycosylation of α-dystroglycan (α-DG) has been proven to be essential for interaction

with the extracellular matrix (ECM) ligands, such as laminin, agrin, and perlecan, thus providing functionality for the DGC at the sarcolemma. α-DG has a complex pattern of abundant glycosylation, including the presence of both *N*- and *O*-linked glycans. While not all structures of these glycans have been fully characterized, the *O*-glycans are reported to be initiated by *O*-GalNAc and *O*-mannose, the latter of which is rarely observed on other mammalian proteins[1-3]. The presence of the *O*-mannose-linked glycans is thought to be particularly important for ligand binding activity of α-DG[1, 4]. Several human congenital muscular dystrophies (CMDs) were found to be caused by genetic defects in glycosyltransferases involved in the biosynthesis of the *O*-mannose-linked carbohydrates. These CMDs are associated with hypoglycosylation of α-DG and classified as dystroglycanopathies; they include Walker–Warburg syndrome (WWS), which is caused by mutations in protein *O*-mannosyltransferase genes, POMT1 and POMT2, responsible for the addition of *O*-mannose onto the protein backbone, and muscle-eye-brain disease (MEB) which results from defects in POMGnT1, the glycosyltransferase that elongates *O*-linked mannose with GlcNAc[5-7]. Although substantial progress has been made in understanding the molecular and genetic bases of *O*-mannosylation of α-DG[8, 9] (reviewed in[1, 4]), the complexity of mammalian glycosylation pathways along with limitations of genetic approaches indicate that a suitable experimentally amenable model system would be a useful tool for studying biological mechanisms of *O*-mannosylation and its involvement in human pathologies. Several recent studies have hinted at *Drosophila* as a potential model organism for such studies.

Drosophila genome encodes two protein *O*-mannosyltransferases, Rotated Abdomen (RT) and Twisted (TW) (aka DmPOMT1 and DmPOMT2, respectively), along with counterparts of all essential components of the mammalian DGC, including Dystroglycan. However, in *Drosophila*, unlike in mammals, DG appears not to be cleaved into α- and β-subunits upon maturation[10], while alternative splicing is predicted to produce three different DG isoforms, DG-A, -B, and -C[11]. Out of these three isoforms, only DG-C includes a predicted mucin-type domain with the potential for extensive *O*-glycosylation, sharing this feature with mammalian α-DG (Figure B-1a,[11, 12]). *Drosophila Dg* is required for apicobasal polarity in epithelial cells and antero-posterior polarity in the oocyte, while the

downregulation of *Dg* expression in larvae and adult flies causes neuromuscular junction synaptic defects, muscle defects and degeneration[11-16]. The similarity of defects caused by Dystroglycan abnormalities in *Drosophila* and mammals has led to the hypothesis that DG functions are evolutionary conserved between *Drosophila* and humans[11, 13-17]. Potential *O*-mannosylation of *Drosophila* DG and its putative role in DG regulation were discussed in several studies. Similar to their mammalian counterparts, RT and TW proteins appear to function as a heterocomplex[18, 19], and they can also add *O*-mannose to a fragment of mammalian α-DG in vitro[18]. Mutations in *rt* and *tw* were found to cause defects in larval NMJ similar to the defects found in *Dg* mutants, and these genes showed genetic interactions[13, 16]. Moreover, Western blot detection revealed the absence of a high-molecular-mass band of DG in *rt* mutants, which was interpreted as an evidence of the involvement of RT in DG glycosylation[16]. However, the glycosylation of *Drosophila* DG has not been examined, and the presence of *O*-mannose modifications on the DG protein remains obscure. In this work, we analyzed the glycosylation of *Drosophila* DG using both in vivo and in vitro approaches. We found that the DG protein is a target of *O*-mannosyltransferase activity and that both RT and TW are simultaneously required for the modification of DG with *O*-linked mannose in vivo. Our results demonstrate that *O*-mannosylation of DG is an evolutionarily conserved mechanism and suggest that it plays an important role in the regulation of DG function in *Drosophila*. These data also indicate that *Drosophila* can be used as a model organism to study molecular and genetic mechanisms of CMDs.

RESULTS

Drosophila DG is a substrate for protein O*-mannosylation activity of RT and TW in vitro*

We tested if *Drosophila* DG can be used as a substrate for *O*-mannosylation by RT and TW in vitro. Previous experiments revealed that the mucin-type domain of mammalian α-DG is a target of *O*-mannose modification[20, 21]. Thus, we focused our analysis on *Drosophila* DG-C isoform that, similarly to its mammalian counterpart, also includes a mucin-type domain, a potential target of *O*-mannosylation. A region of mucin-type domain of the DG-C isoform (Figure B-1a) was expressed in *Escherichia coli*, purified and tested as a substrate in an in vitro *O*-mannosylation assay using microsomal membrane

fraction prepared from *Drosophila* larvae as a source of RT-TW activity (see *Material and methods*). We also tested a corresponding fragment of DG-A extracellular domain that spans the predicted region of mucin-type *O*-glycosylation (Figure B-1a). As a positive control, we used the region of mucin-type domain of rabbit α-DG that was previously shown to be a substrate for RT-TW in vitro[18]. Our results indicated that, similarly to mammalian α-DG, DG-C could serve as a substrate for in vitro *O*-mannosylation, with the DG-C fragment being apparently a better substrate for RT-TW activity in vitro than the region of mucin-type domain of rabbit α-DG. Incorporation of mannose into DG-A was not significantly above the background (Figure B-1b).

Extracellular domain of DG is properly folded and trafficked in Drosophila cells with or without RT-TW activity

We expressed the entire extracellular part of DG-C tagged with 3xFLAG epitope (designated below as ExDG) as a transgenic construct in *Drosophila*. To confirm that ExDG is properly folded and trafficked in *Drosophila* cells, we verified by Western blot that the ExDG protein was efficiently secreted in a diffusible form into cell medium when expressed in *Drosophila* S2 cultured cells (Figure B-2a). Next, we expressed ExDG in vivo within wing imaginal disk epithelium using a UAS-GAL4 expression system. We tested three different genetic backgrounds: *rt* mutant, wildtype, and *rt-tw* coexpression. The immunofluorescent detection of in vivo expressed ExDG revealed that it was similarly delivered to the cell surface and no difference in subcellular localization of ExDG could be detected between these backgrounds (Figure B-2b). Thus, we concluded that the trafficking and subcellular localization of ExDG was not affected by RT-TW activity in vivo.

The concurrent in vivo activity of RT and TW is necessary and sufficient to generate the high-molecular-mass form of DG

The ExDG protein was expressed in *Drosophila* with different genetic backgrounds corresponding to varied levels of RT and TW, and then it was analyzed by Western blotting using an anti-FLAG antibody. We found a drastic difference in the pattern of high-molecular-mass bands of ExDG expressed in *rt* and/or *tw* mutant, wildtype, and RT-TW coexpression backgrounds (Figure B-3a and b).

In wild-type flies, we observed two major bands of estimated sizes of 175 kDa (designated as S (small) band) and 215 kDa (designated as L (large) band) present at approximately equal amounts. The L band was undetectable in *rt* mutants, *tw* mutants, as well as in *rt* and *tw* doublemutants. We also found that the relative amount of the L band was significantly increased in flies coexpressing both RT and TW. At the same time, the overexpression of one of the *O*-mannosyltransferases, RT or TW, resulted in no significant increase of the relative amount of the L band, as compared to a wild-type background. These results indicated that RT and TW are simultaneously required in vivo for producing the high-molecular-mass form of ExDG.

Interestingly, we did not observe a band corresponding to the ~110 kDa band that was detected for endogenously expressed DG with antibodies raised against its intracellular part, but not with the antibody specific for DG-C[12, 16], which confirmed previous conclusion that this smaller band represents only DG-A/B isoforms[12].

Glycosidase treatments and lectin blots revealed O*-mannosylation of high-molecular-mass form of DG*

We analyzed the glycosylation of DG-C isoforms represented by L and S bands using treatments with specific glycosidases. In agreement with *N*-linked site prediction for DG-C (Figure B-1a), the removal of *N*-linked glycans with PNGaseF resulted in a gel shift of ExDG, with the shift being similar for both L and S bands (Figure B-4a). The S band was similarly decreased in size for both *rt* mutant and RT-TW coexpression backgrounds. The shift of ExDG bands after PNGaseF treatment, estimated as ~10 kDa, corresponds to approximately six oligomannose-type *N*-glycans (typical N-linked structures present on *Drosophila* glycoproteins[22-24]). This result is consistent with the presence of total eight putative N-linked sites within the extracellular part of DG-C, with five of them predicted to have a significant potential for N-glycosylation (Figure B-1a). We also treated purified ExDG with α-mannosidase that cleaves off α-linked mannose residues. This treatment should both remove O-linked mannose[18, 25] and trim oligomannose *N*-linked glycans down to the *4*-β-mannosyl chitobiose (Manβ*1–4*GlcNAcβ*1–4*GlcNAc. . .) of the core structure. α-Mannosidase treatment alone resulted in a significantly greater shift

of the L band as compared to the S band, which suggested that the high-molecular-mass form is extensively modified with *O*-linked mannose. This conclusion was confirmed by sequential digestion with PNGaseF and α-mannosidase, which showed an additional decrease in the mass of the L band as compared to PNGaseF digestion alone. A similar comparison between PNGaseF-treated and double PNGaseF /α-mannosidase-treated S bands did not reveal a significant difference, thus suggesting that the S band represents a glycoform without extensive *O*-mannosylation.

The presence of α-linked mannose on ExDG was examined by concanavalin A (Con A) lectin blot (Figure B-4b). Con A strongly reacted with the L band, while showing weaker reactivity with the S band from *rt* mutant and RT-TW coexpression backgrounds. The ConA reactivity of the S band from *rt* mutants was eliminated by PNGaseF treatment, indicating that this reactivity was solely due to mannose structures of *N*-linked glycans and providing further evidence that ExDG was not modified with *O*-mannose in *rt* mutants. At the same time, PNGaseF treatment did not eliminate the reactivity of the L band, suggesting that *O*-mannose is present on this glycoform. Interestingly, the S glycoform from RT-TW coexpression also showed Con A reactivity after PNGaseF treatment, suggesting the presence of some *O*-linked mannose on this isoform as well. We also found that double PNGaseF/α-mannosidase-treated ExDG from RT-TW coexpression retains some residual reactivity for Con A, which suggest the presence of mannosidase-resistant *O*-mannose on ExDG (Figure B-4b). This conclusion is consistent with the fact that *O*-linked mannose is a relatively poor substrate for α-mannosidase[25], while there is also a possibility that some modification present on *O*-mannose may inhibit α-mannosidase.

We also examined the glycosylation of ExDG using several other lectins. The presence of *O*-linked GalNAc on purified ExDG was analyzed using *Vicia villosa* (VVA) lectin (Figure B-5a). Only S band showed strong VVA staining, while the reactivity of the L band was barely above the background. Interestingly, the S band from *rt* mutants had a significantly stronger VVA reactivity than that from RT-TW coexpression, suggesting more extensive *O*-GalNAc modification of ExDG in the absence of *O*-linked mannose. *O*-GalNAc can be extended with β1–3 linked galactose, which produces the core 1 structure, one of the most abundant *O*-linked modification in *Drosophila*[21, 26]. Thus, we tested a

possibility that *O*-GalNAc of the L band was masked from VVA recognition because of β1–3 galactose extension by lectin blot using Peanut agglutinin (PNA) that specifically recognizes this disaccharide. However, neither L nor S band showed PNA reactivity (data not shown), indicating the absence of core 1 structure on ExDG.

We also performed analysis of ExDG glycosylation by *Wisteria floribunda* agglutinin (WFA) and *Hippeastrum hybrid* lectin (HHL) that recognize terminal GalNAc and mannose residues, respectively. To exclude possible reactivity with *N*-linked glycans, purified ExDG was first treated with PNGaseF in these experiments. Similar to VVA, WFA strongly recognized the S band from *rt* mutants, had a weaker reactivity with the S band from RT-TW coexpression, and could barely detect the L band (Figure B-5b). On the other hand, HHL staining had a pattern complementary to that of WFA and VVA: HHL strongly recognized the L band, showed weak reactivity with the S band from RT-TW coexpression, and did not react at all with the S band from *rt* mutants (Figure B-5b). Together, these results further supported our conclusion that ExDG from RT-TW coexpression has *O*-mannose modification, with the L band representing more extensively *O*-mannosylated glycoform and the S band probably including only limited *O*-mannose modification. At the same time, the extent of *O*-linked GalNAc modification appears to be complementary to that of *O*-mannosylation, with the S band from *rt* mutants being most prominently modified with *O*-GalNAc, the amount of *O*-GalNAc being significantly decreased on the S band from RT-TW coexpression and even further reduced on the L glycoform.

Mass spectrometry of purified extracellular domain of DG identified O-*mannose and* O-*GalNAc modifications*

To further elucidate the glycosylation of ExDG, we performed mass spectrometry analyses of the L and S bands of ExDG from RT-TW coexpression, and the S band from *rt* mutants. To this end, samples of purified and gel-separated L and S bands were subjected to digestion either by trypsin or by a combination of trypsin and elastase. The peptide coverage was very similar for all three analyzed samples (see supplementary Table), revealing no difference in polypeptide sequences. The central region of ExDG including mucin-type domain (a.a. 309–552) was not amenable to the mass spectrometry analysis, which

was likely because of its resistance to protease digestion due to a peculiar amino acid sequence (including only 4 trypsin sites) and putative extensive glycosylation. The rest of the protein was almost completely covered by mass spectrometry analysis. Glycopeptides were analyzed by a combination of BEMAD and a pseudo-neutral loss method on a linear ion trap (see *Material and methods* and[27]), which allowed us to not only map the site of modification but also determine the identity of the *O*-linked glycan structure (Figure B-6a and b). In total, we identified nine *O*-linked modifications on the L glycoform, five modifications on the S glycoform from RT-TW coexpression background, and two *O*-linked sites on the S glycoform from *rt* mutants (Table B-1, Figure B-7, supplementary Figure B-S1). We found that the L and S glycoforms from RT-TW coexpression have six and three *O*-mannose sites outside of the mucin region, respectively. No *O*-mannose was found on ExDG purified from *rt* mutants, even though we specifically looked for the presence of glycopeptides observed in RT-TW coexpression background. Two *O*-GalNAc sites were identified on each of the analyzed glycoforms, with some of them overlapping with *O*-linked mannose sites on ExDG coexpressed with RT and TW. These results are consistent with our conclusions from glycosidase treatment and lectin blot analyses about the absence of *O*-mannose in *rt* mutants and the reciprocal distribution of *O*-GalNAc and *O*-mannose on the L and S glycoforms. These data also suggested that the extensive *O*-mannose modification predicted by glycosidase and lectin analyses is confined to the central region of ExDG including mucin-type domain.

ExDG modified with O*-mannose is resistant to protease digestion*

We further characterized properties of purified ExDG from RT-TW coexpression by analyzing its sensitivity to trypsin. Our mass spectrometry analysis suggested a possibility that the central region of ExDG may be resistant to proteolytic cleavage because of extensive glycosylation. We tested this hypothesis by analyzing products of trypsin digestion of denatured PNGaseF treated ExDG. Since it has been reported that Coomassie and silver staining methods often fail to efficiently detect highly glycosylated peptides, we used PAS and PAS-silver staining of SDS–PAGE-separated tryptic digestion to detect glycopeptides[28, 29]. A ~120 kDa band was detected on the gel by PAS staining indicating the

presence of a glycosylated peptide resistant to proteolysis (Figure B-8). We analyzed this glycopeptide by binding to lectin affinity agarose beads. The glycopeptide was quantitatively bound by Con A beads; however, it did not bind to WFA or VVA agarose (Figure B-8). This result suggested that the glycopeptide has significant *O*-mannose modification, while revealing no evidence for the presence of GalNAc. *O*-Mannosylation of this glycopeptides was further corroborated by treatment with α-mannosidase which resulted in a significant shift of the glycopeptide on the gel (Figure B-8). The glycopeptide band also became less intensely stained after α-mannosidase treatment, which is consistent with lower sensitivity of PAS-silver staining for proteins with reduced amount of glycosylation[29]. Taken together, these results provided further evidence that ExDG contains a trypsin-resistant region with extensive *O*-mannosylation, suggesting that this region corresponds to the central part of ExDG not amenable to mass spectrometry analysis and including mucin-type domain.

Function of ExDG is modulated by RT-TW activity in vivo

In order to test the functional influence of *O*-mannosylation on *Drosophila* DG, we developed an in vivo assay for ExDG activity. Intracellular truncation of membrane proteins involved in cell adhesion or signaling processes is known to often result in dominant-negative forms that could antagonize functions of their endogenous wild-type forms by participating in nonproductive molecular interactions with ligands[30, 31]. Since ExDG construct includes only the extracellular domain of DG and thus should be mostly expressed as a secreted soluble protein (Figure B-2a), we expected that it could function as a dominant-negative form by outcompeting endogenous DG-ligand interactions on the cell surface. Indeed, we found that the ubiquitous expression of ExDG in developing wing imaginal disks results in defective crossveins, causing the loss of crossvein tissue (Figure B-9a), with the phenotype penetrance of ~12% (Figure B-9b). This phenotype is characteristic for loss-of-function mutations of *Dg* and *detached*, the *Drosophila* ortholog of vertebrate *dystrophin*, which indicates the involvement of the DGC in crossvein development[32]. Thus, ExDG can function as a dominant-negative form antagonizing endogenous DG activity. We found that the penetrance of the ExDG phenotype was significantly increased (~2.5 times)

by coexpression with RT and TW (Figure B-9b), which suggests that the activity of ExDG is potentiated by *O*-mannosylation in vivo.

DISCUSSION

Here we demonstrated for the first time, to our knowledge, that *Drosophila* Dystroglycan is *O*-mannosylated and serves as a substrate for two *Drosophila* protein *O*-mannosyltransferases, RT and TW, in vitro and in vivo. Our study was focused on the extracellular part of the DG-C isoform that includes a mucin-type domain and most closely resembles the structure of mammalian α-DG. For in vivo analyses, the entire extracellular domain of DG-C, ExDG, was expressed in flies with genetically varied levels of RT and TW. When characterized by Western blots, a characteristic pattern of two major bands (L and S) of ExDG was observed for the genetic backgrounds with wild-type or elevated levels of RT and TW, while only the lower-molecular-mass band (S) was detected when the activity of either *rt* or *tw* genes was compromised, as well as in double *rt/tw* mutants. These results are consistent with the recently reported absence of high-molecular-mass band of endogenous DG in *rt* mutants[16]. Most importantly, we found that the relative amount of the L form was significantly elevated only by a concurrent increase of RT and TW, while increasing the level of either RT or TW alone had no significant effect on the amount of the L band (Figure B-3 a and b). Together, these results indicate that RT and TW concurrent activity is required for generating the high-molecular-mass form of ExDG in vivo, suggesting that RT and TW work as a heterocomplex in vivo and providing further support for similar conclusions of in vitro[18] and genetic[19] experiments. Since the expression of constructs was driven ubiquitously in these experiments, the fact that the elevated level of either RT or TW alone does not increase the amount of the L glycoform also suggests that the levels and patterns of endogenous RT and TW expression largely coincide throughout larval and pupal stages. This conclusion is consistent with in situ hybridization data that revealed the overlapping expression of *rt* and *tw* at embryonic and larval stages ([19]; Lavrova and Panin, unpublished data).

To elucidate the nature of the L and S forms of ExDG, we investigated their glycosylation using a combination of approaches, including specific glycosidase treatments, lectins, tryptic digestion, and mass

spectrometry. Treatment with PNGaseF suggested that there is no difference in *N*-linked glycosylation between ExDG species represented by L and S bands from *rt* mutant and RT-TW coexpression backgrounds, which indicates that RT-TW activity has no effect on *N*-linked glycosylation of ExDG. At the same time, α-mannosidase treatments of purified ExDG suggest that the L band represents a glycoform extensively modified with *O*-linked mannose, while the S band corresponds to the glycoform without significant *O*-mannosylation (Figure B-4). This conclusion is consistent with the fact that α-mannosidase has similar effect on S bands from *rt* mutant and RT-TW coexpression backgrounds (Figure B-3). Further evidence of *O*-mannosylation of the L glycoform was obtained by lectin blots with mannose-recognizing lectins, Con A and HHL (Figures B-4b and B-5b). The estimated decrease in the L glycoform molecular mass upon α-mannosidase digestion of PNGaseF-treated ExDG is ≥10 kDa (Figure B-4a), which implies the presence of more than 50 *O*-mannose residues attached to ExDG. Interestingly, our mass spectrometry analysis identified only six *O*-linked mannoses outside of a central ~250 a.a. region including a mucin-type domain (Table B-1, Figure B-7, supplementary Figure B-S1), suggesting that this central region of ExDG bearsmost of the predicted *O*-mannosylation. Mass spectrometry results also suggested that this region may be resistant to proteases due to its extensive glycosylation. This possibility was further supported by the analysis of trypsin digestion of ExDG purified from RT-TW coexpression background that revealed the presence of~120 kDa trypsin-resistant glycopeptide. Moreover, the specific binding of this glycopeptide to Con A beads suggested that it is *O*-mannosylated (Figure B-8). Digestion with α-mannosidase resulted in an estimated ~8 kDa shift of this glycopeptide on the gel (Figure B-8), which corresponds to he loss of ~44 mannose residues. This result is consistent with the extensive α-mannosylation of the central region of ExDG suggested by mass spectrometry and glycosidase analyses of the full-length protein.

Our mass spectrometry analysis revealed a similar peptide coverage for purified L and S bands from both analyzed genotypes, *rt* mutant and RT-TW coexpression (supplementary Table B-S1 and Figure B-1). Notably, the most-*N*-terminal peptide recovered for all three analyzed bands is the same, and

it is located only 29 amino acids downstream of the ExDG N-terminus and two residues downstream of the predicted cleavage of the signal peptide. Since the integrity of ExDG C-terminus was conferred by anti-FLAG reactivity and purification, and a hypothetical cleavage of 29 N-terminal amino acids (corresponding to ~3.5 kDa) could not account for the remaining ~10 kDa mass difference between the L and S bands after a combined PNGaseF/α-mannosidase treatment, we concluded that this difference reflects a posttranslational modification of ExDG. At least in part, this difference is explained by the presence of *O*-mannose residues resistant to α-mannosidase since the L glycoform retains residual Con A reactivity after combined PNGaseF/α-mannosidase treatment (Figure B-4). Thus, the L and S bands indeed most likely represent distinct glycoforms of ExDG that are primarily different in *O*-mannosylation, which, however, does not exclude a possible contribution of other types of modifications triggered or influenced by *O*-mannosylation.

Staining with VVA lectin that recognizes terminal GalNAc was previously found to correlate with the level of DG expressed in *Drosophila* muscles and NMJ[13], suggesting that *Drosophila* DG may bear this modification. Thus, besides *O*-mannose, the L and S glycoforms may be modified with *O*-linked GalNAc. Interestingly, we found that the reactivity of ExDG bands to VVA and WFA lectins that recognize *O*-linked GalNAc was essentially complementary to that of Con A and HHL that bind *O*-mannose, with the S band from *rt* mutants being most prominently stained with both VVA and WFA, the S band from RT-TW coexpression showing significantly weaker reactivity, and the L band being barely detectable with these lectins (Figure B-5a and b). Thus, *O*-GalNAc appears to be present on ExDG glycoforms in a reverse proportion to *O*-mannose. This result is consistent with our finding that two of the identified *O*-mannose sites could also be modified by *O*-linked GalNAc (Table B-1, Figure B-7). Since RT and TW function in the ER[19]upstream in the secretory pathway to *O*-GalNAc-transferases, PGANTs, that function in the Golgi[33], this complementary pattern of glycosylation suggest that *O*-mannosylation can compete with *O*-GalNAc modification of DG. Thus, these data provide further support for our conclusion that the majority of *O*-mannosylation occurs within the region of mucin-type domain, where

O-mannose could directly compete with the bulk of predicted *O*-GalNAc modification. This possible competition may underlie a regulatory mechanism that modulates DG function, for instance, via changing the ligand-binding activity of DG.

Importantly, the mass spectrometry analysis of ExDG purified from RT-TW co-expression background unequivocally detected *O*-mannosylation sites outside of the mucin-type region of ExDG (Table B-I, Figure B-7, supplementary Figure B-S1). To our knowledge, this is the first report of animal *O*-mannosylation outside of mucin-type domain. This result has implication for understanding the substrate specificity of animal protein *O*-mannosyltransferases. The identified *O*-mannose sites on *Drosophila* DG neither agree with a consensus sequence proposed for mammalian *O*-mannosylation based on in vitro assays[20], nor conform to the sequence preferences proposed for in vivo *O*-mannose modification of α-DG[21]. It is worth noting that previous studies were focused on the modification α-DG mucin domain and used relatively short peptides or fragments of α-DG for assaying the substrate specificity of *O*-mannosylation. In our experiments, we analyzed *O*-mannosylation of a complete extracellular domain of DG. Thus, the mentioned difference in sequences of identified *O*-mannose sites can suggest that the substrate recognition for *O*-mannosylation in vivo occurs differently in a context of a full-length protein as opposed to a peptide substrate, which is in general agreement with the previous conclusion that structural elements distant from modification sites can determine the substrate recognition for *O*-mannosylation in vivo[21]. Alternatively, this result can suggest that *Drosophila* and mammalian *O*-mannosylation may be controlled by different structural properties of substrates. Interestingly, three out of the six identified *O*-mannose sites are located within the region that is shared by all three *Drosophila* DG isoforms produced by alternative splicing (Figures B-1 and B-7), which suggests that all of them, not just DG-C, may be targeted by *O*-mannosylation.

Intriguingly, the generation of the L glycoform by RT-TW activity appears to be an “all-or-nothing” process with the absence of intermediate forms. This intriguing observation suggests a possible mechanism of *O*-mannosylation that requires recognition of some determinants that results in quantitative modification of nearly all available sites. These determinants may belong to DG itself, such as some

posttranslational modifications, or they may represent certain factors that regulate the function of the RT-TW complex within the cell. The first scenario includes an interesting possibility that recognition by RT-TW outside of the mucin domain may serve as a determinant that initiates processive *O*-mannosylation of DG within its mucin domain region. This possibility is consistent with the difference in *O*-mannose sites outside of the mucin region on the L and S glycoforms from RT-TW coexpression, as well as with the evidence that a region N-terminal to mucin-type domain plays a role in *O*-mannosylation initiation of human α-DG mucin domain[21]. In any case, these determinants, or factors that influence them, may provide an additional level of DG regulation and modulate the ratio of S and L glycoforms in a cell-specific manner. This hypothesis predicts that the glycoform ratio may vary between different cells independently from the level of RT-TW activity, and this interesting possibility will be addressed in future experiments.

Protein *O*-mannosylation in yeast has been implicated in secretory protein sorting, protein stability, and degradation via ERAD-mediated pathway[34-36]. However, in our experiments, the immunofluorescent detection and optical sectioning did not reveal significant differences in the level of expression or subcellular localization of ExDG within the wing imaginal disk epithelium in the presence or absence of RT-TW activity. At the same time, we found that the dominant-negative effect of ExDG on the crossvein formation is significantly potentiated by RT-TW activity (Figure B-9), which suggests that the in vivo activity of ExDG is modulated by *O*-mannosylation and that the outcome of DG function depends on the balance between the amounts of L and S glycoforms. Thus, we think that *O*-mannosylation does not affect the stability or trafficking of *Drosophila* DG, but rather it changes its functional properties, e.g. by modifying interactions with extracellular ligands. These results suggest that *O*-mannosylation can function as a molecular regulator of the DGC-mediated pathway via changing the ligand-binding activity of DG.

In conclusion, the present study has unambiguously demonstrated that the extracellular domain of *Drosophila* DG-C is a target of *O*-mannosylation. We found that *O*-mannosylation is an abundant modification of ExDG that requires concurrent in vivo activity of two *Drosophila* protein *O*-

mannosyltransferases, RT and TW. A combination of glycosidase treatments, lectin blotting, trypsin digestion, and mass spectrometry analyses provided strong evidence that the majority of *O*-mannose is located within a central 250 a.a. region of the extracellular part of DG that includes a mucin-type domain. Our experiments revealed that subcellular localization and trafficking of ExDG are not affected by RT and TW, while ExDG activity can be modulated by RT–TW activity in vivo ExDG, thus suggesting that *O*-mannosylation regulates ligand-binding activity of *Drosophila* DG. These results highlight the evolutionary conservation of mechanism of *O*-mannosylation between *Drosophila* and mammals and suggest that *Drosophila* can be a suitable model system to study molecular and genetic mechanisms of mammalian α-DG *O*-mannosylation. In addition, our experiments revealed an interplay between two *O*-glycosylation pathways that modify DG with *O*-mannose and *O*-GalNAc, which suggests that *O*-mannose can function as a molecular switch that regulates other types of glycosylation. In addition, we found that *O*-mannose sites can be located far outside of the mucin-type domain, which suggests that proteins without mucin-type domains may also be targets of *O*-mannosylation in vivo.

MATERIALS AND METHODS

Drosophila *strains and cDNA*

The following *Drosophila* mutant alleles and transgenic insertions were obtained from the Bloomington *Drosophila* Stock Center, Indiana University: *tw1*, *rtp*, *rt2*, *Df(1)su(s)83* (*tw* deficiency), *Dp(1;Y)y2sc* (*tw* duplication), *Act5C-GAL4–25*, *tubPGAL4 (LL7)*, *ptc-GAL4*. *UAS-rt*, and *UAS-tw* transgenes were previously described (Lyalin et al. 2006). *Drosophila Dg-C* cDNA was obtained from Dr. Ruohola-Baker (University of Washington, Seattle).

Constructs and proteins for in vitro O-*mannosylation assays*

Templates for engineering expression constructs were plasmids with cDNA sequences of *Dg-A* (obtained from DGRC, Indiana University), *Dg-C* (obtained from Hannele Ruohola-Baker, Seattle), and *rabbit* α-*Dg* (*DGFc5*, a gift from Michael Oldstone and Stephan Kunz from Scripps Institute, La Jolla, supported by NIH grant AI09484[37]). Fragments of *Dg- A, Dg-C*, and rabbit *Dg* cDNAs encoding protein regions with predicted abundant *O*-glycosylation (amino acids 204–399 for DG-A, 315–592 for DG-C,

and 314–487 for rabbit α-DG) were PCR-amplified and cloned into the *pET41a* vector (Novagen, EMD USA, Gibbstown, NJ) between *Nco*I and *Eag*I cloning sites and in frame with the N-terminal GST tag. The GST-fused proteins were expressed and purified from *E. coli* BL21(DE3) cells according to manufacturer's protocol (Novagen). Briefly, the protein expression was induced with 0.4 mMIPTG for 18 h, and then cells were harvested and lysed in the wash buffer (PBS, 0.1% NP-40, 0.1% PMSF) by sonication. The lysates were precleared by centrifugation and 0.45 μm membrane filtering, and incubated with GST-beads on a nutator at 4◦C overnight. Then, beads were extensively washed with the wash buffer and purified proteins were eluted with 100mMglutathione in 0.5M Tris, pH 8.0. The eluted proteins were dialyzed against 2 mM EDTA, 0.1% PMSF, 20 mM Tris pH 8.0, concentrated by Millipore centrifugal filters (15 kDa cut-off), and used as substrates in the in vitro *O*-mannosylation assay.

In vitro O-*mannosylation assays*

The assays were performed essentially as described previously[18]. Briefly, a reaction mixture (20 μL total volume) contained 20 mM Tris, pH 8.0, 100 nM [^{3}H]- mannosylphosphoryldolicol (ARC, Inc.), 2 mM 2-mercaptoethanol, 10 mM EDTA, 0.5% n-octyl-β-D-thioglucoside, 10 μg GST-tagged acceptor protein, and 80 μg of microsomal membrane fraction as a source of *O*-mannosyltransferase activity. The microsomal membrane fraction was prepared from third instar *Drosophila* larvae using previously published protocol[18]. The concentration of proteins in microsomal fraction was determined by the Bradford assay. After 1 h incubation at 24ºC, mannosyltransferase reactions were stopped by adding 200 μL PBS with 1% Triton X-100, the mixtures were centrifuged at 10,000 x *g* for 10 min at 4ºC, and the supernatant was bound to pre-washed GST-beads on a nutator. Then, the beads were washed three times with PBS containing 0.5% Triton X-100, and the incorporation of radioactive mannose was measured in dpm using a liquid scintillation counter. Background control was determined by a mock assay following the same protocol but with 10 μg of BSA instead of a real acceptor. The results of the assay were presented as the ratio of radioactive mannose incorporation for Dystroglycan acceptors to that for BSA as a control.

ExDg construct

The *ExDg* construct was generated by in-frame fusion of cDNA region encoding the first 1048 amino acids of DG-C isoform (the entire predicted extracellular domain of DG-C) with the fragment encoding the 3xFLAG tag (Sigma) using standard molecular biology techniques. Details on molecular cloning of *ExDg* are available on request. The ExDG protein encoded by the construct lacks transmembrane domain and is predicted to be a secreted protein.

Expression of ExDG in Drosophila *S2 cells*

S2 tissue culture cells were maintained and transfected as previously described[38]. For tissue culture expression, *ExDg* was subcloned into the *pMK33* vector with metallothionein promoter using standard molecular biology techniques (details are available upon request). The expression of ExDG was induced in transiently transfected cells by 0.7 mM $CuSO_4$.

In vivo expression and purification of ExDG

ExDg was cloned into the *pUAST* vector for in vivo expression, and transgenic *Drosophila* strains were obtained by P-elementmediated germline transformation. In vivo expression of the *UAS-ExDg* transgene was induced using the UAS-GAL4 system[39]. The specificity of ExDG detection by Western blot was confirmed using 'negative control' samples without ExDG expression. ExDG was expressed in vivo in the following genetic backgrounds: *rt−*, *rtP/rt2*; *tw−, tw1/ Df(1)su(s)83* (where *Df(1)su(s)83* is a *tw* deficiency); *rt+tw+, UAS-rt,UAS-tw/Act5C-GAL4*; *rt+, UAS-rt/Act5C-GAL4*; *tw+, UAS-tw/Act5C-GAL4*. ExDG protein was purified from pupae with anti-FLAG M2 agarose (Sigma). For each purification experiment, 15–80 larvae or pupae were collected and lysed in 300 μL–1.5 mL of lysis buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% Triton X100) including cocktail of protease inhibitors (Complete, Roche). Following centrifugation at 18,000 x *g* for 20 min at 4ºC, the supernatant was added to 10–40 μL of FLAG agarose beads and incubated for 2 h at 4ºC with nutation. The agarose beads were then washed four times with the lysis buffer. The purified ExDG bound to beads was directly used in later assays.

Glycosidase treatments

ExDG protein purified from 3–4 pupae was incubated with 500 units of PNGaseF (NEB) in 5mM sodium phosphate buffer (pH 7.5) for 1 h at 37ºC (mock control was without PNGaseF). It was later treated with 0.4 milliunits of α-mannosidase from Jack beans (Sigma) (or without α-mannosidase for control) in the 50 mM sodium citrate buffer (pH 4.5) for 3 h at 37ºC. Reactions were stopped by adding 2 x SDS loading buffer and used in Western or lectin blot analyses.

Western and lectin blot analyses

Analyses were performed according to standard protocols. Briefly, the tissue lysates normalized by the Bradford method for protein amount, or purified ExDG from pupae, were run on 5% SDS–PAGE gel and then the separated proteins were transferred onto a nitrocellulose membrane. For Western blotting, mouse anti-FLAG M2 primary (Sigma) and rabbit anti-mouse HRP-conjugated secondary (Jackson Laboratories) antibodies were used for detection by chemiluminescence (SuperSignal, Thermo Scientific). For lectin blots, proteins on the membrane were blocked with 2% BSA (Roche) in TBST (10mMTris-HCl, pH 8.0, 150mMNaCl, 0.05% Tween 20) and then incubated with biotinylated lectins (2.5μg/mL of ConA, or 10 μg/mL of VVA, WFA, HHL, and PNA, all from Vector) for 1 h at room temperature followed by detection with a Vectastain ABC kit (Vector). The specificity of lectin staining was confirmed by incubation with lectin in the presence of corresponding inhibiting sugars (i.e., 0.2M GalNAc for VVA and WFA, 0.2M galactose for PNA, 0.2M methyl α-D-mannopyranoside for ConA and HHL). Blotswere quantified using the ChemiDoc XRS system (Bio-Rad).

In-gel digestion of purified ExDG

ExDG was purified from ~60 pupae and separated by 5% SDS– PAGE. Protein bands visualized by Coomassie G250 staining were cut from the gel. In-gel digestion was performed as follows: Coomassie G250 was removed from the selected bands by repeated swelling and shrinking with 40 mM ammonium bicarbonate and acetonitrile, respectively. After destaining, the proteins within the gel bands were reduced by swelling in 40mM ammonium bicarbonate containing 10mM DTT for 1 h at 56ºC, followed by carbamidomethylation with 55mM iodoacetomide (Sigma). Bands were swelled, dehydrated,

and dried to dryness using SpeedVac. Both sequence grade trypsin (Promega) and elastase (Worthington Biochemical) were reconstituted using 40mM ammonium bicarbonate. The dried gel pieces were swelled on ice by adding 5µg of both trypsin and elastase. After swelling for a period of 45 min on ice, the bands were then subjected to overnight digestion at 37ºC. The resulting glycopeptides were extracted three times for 15 min using 5% formic acid in 25%, 50%, and 75% acetonitrile. The extracted glycopeptides were dried down using a SpeedVac. Sequencing of the mucin region of DG proved to be difficult possibly because of the extensive glycosylation and/or the lack of suitable amino acids preventing enzymatic cleavage of the protein backbone.

Mass spectrometry analyses: BEMAD

β-Elimination through Michael addition with DTT (BEMAD) was carried out essentially as described previously[27]. Briefly, lyophilized peptides were β-eliminated by resuspending them in 1% triethylamine and 0.1% sodium hydroxide. The peptides were then subjected to Michael addition with DTT. The reaction was incubated at 52ºC for 2 h. After completion of incubation, the reaction was quenched with trifluoroacetic acid (final concentration ~1%) and the peptides were dried in a Speed Vac. BEMAD allowed us to identify sites of glycosylation by detecting a mass increase of 136.2 Daltons due to DTT addition. However, the identity of the glycan present at those sites could not be distinguished. To reveal the identity of glycosylation, the BEMAD data were combined with the results of the pseudo-neutral loss method on a linear ion trap.

Nano-LC-MS^3

Extracted ExDG glycopeptides were also analyzed using Nano-LC-MS^3 on a linear ion trap mass spectrometer (LTQ; ThermoFisher). The glycopeptides were resuspended in 0.5 µL of solvent B (0.1% formic acid/80% acetonitrile) and 19.5 µL of solvent A (0.1% formic acid) and loaded on a 75 mm x 8.5 cm C18 reverse phase column (packed in-house, YMC GEL ODSAQ120ÅS-5) using a nitrogen bomb. Peptides were eluted using a linear gradient increasing from 5% to 50% B over 85min. Each full MS scan from 300 to 2000 *m*/*z* yielded five MS/MS scans of the top five most intense peaks with a dynamic exclusion of 2 for 30 s. Data-dependent MS^3 scans were triggered if a neutral loss was observed equal to

the singly or doubly charged mass of Hexose, HexNAc, Fucose, or Neu5Ac (sialic acid) within the top three peaks from the MS/MS scan.

Data analysis

The resulting data were searched against a nonredundant *Drosophila* database (Flybase), as well as a database only containing the *Drosophila melanogaster* protein sequences for α-DG (NCBI) added to the common contaminants database (Thermo Fisher), using the TurboSequest algorithm (Bio-Works 3.3, Thermo Fisher). For identifying glycopeptides, a mass increase of 162.1, 203.1, and 365.2 Daltons was allowed on serines and threonines looking for the addition of mannose, GalNAc, and Hex-HexNAc. Additionally, when searching the peptides that underwent the BEMAD procedure, amass increase of 136.2 Daltons on both serines and threonines was allowed dynamically. All peptides indicating a posttranslational modification were confirmed through manual inspection after automatically deleting all peptides matching back to α-DG with Sf and P scores below 0.3 and 20.

Analysis of trypsin-resistant glycopeptides of ExDG

ExDG purified from ~80 *Drosophila* pupae was treated with 1000 units of PNGaseF as described above. After incubation in 4 mM DTT /6.4 M urea at ~100ºC for 20 min, denatured ExDG was incubated with 0.5 μg of trypsin (TPCK treated, Sigma) in 0.9M urea 43mM Tris, pH 8.0, overnight. Trypsin was inactivated by boiling and the sample was dialyzed against 37.5mM NaCl 12.5mM Tris-HCl, pH 7.4, for binding to lectin affinity beads, or against 10mM sodium citrate, pH 4.5, for α-mannosidase treatment (see below). The dialyzed ExDG peptides were incubated with ConA, WFA, and VVA agarose beads (Vector) in 50mM Tris-HCl, pH 7.4, 150mM NaCl, 0.1mM CaCl2, 0.01mM MnCl2, and 0.05% Tween-20 for 2 h at room temperature. The beads were washed 3 times with the incubation buffer, and the bound fraction was analyzed by SDS– PAGE and PAS-silver staining. Treatment with α-mannosidase was performed in 45mM sodium citrate, pH 4.5, at 37ºC overnight using 2.2 μg (0.05 milliunits) of α-mannosidase from jack beans (Sigma). In the control experiment, ExDG peptides were incubated using this protocol with the same amount of heat-inactivated α-mannosidase. The samples were analyzed by SDS–PAGE and PAS/PAS-silver staining.

PAS and PAS-silver staining

PAS (periodic acid-Schiff) and PAS-silver staining was performed essentially as described before with some modifications[28, 29]. Briefly, for PAS staining, gels were incubated in 1% periodic acid for 15 min at 4ºC following rinse with 7.5% acetic acid for 5 min. After six washes with H_2O (5 min each), gels were incubated in Schiff reagent for 15 min at 4ºC and then washed with 0.5% sodium metabisulfite three times for 10 min followed by extensive washes with H_2O. For PAS-silver staining, gels were PAS-stained and then incubated in 0.2% AgNO3/0.03% formaldehyde for 15 min. The staining was developed in 2.3% Na2CO3/0.01% formaldehyde/0.001% sodium thiosulfate for 5–10 min, and the development of staining was stopped by of 2.5% acetic acid.

Immunostaining and microscopy

Expression of the *pUAST-ExDg* construct was induced in wing imaginal disks using the *ptc-GAL4* driver. Third instar imaginal disks were dissected, fixed, and stained as described previously[40]. The following antibodies and corresponding dilutions were used for immunostaining: primary mouse anti-FLAG (1:2,000) (Sigma) and secondary donkey anti-mouse Cy3 (1:250) (Jacksons Laboratories). Digital images were obtained using a Zeiss Axioplan 2 fluorescent microscope with the ApoTome module for optical sectioning. Z-sections were reconstructed using Zeiss AxioVision software.

Bioinformatic analyses

Prediction of DG glycosylation and signal peptide cleavage was performed by NetNGlyc, NetOGlyc, and Signal IP software at the Center for Biological Sequence Analysis site, DTU, Denmark (http://www.cbs.dtu.dk/services/)[41, 42].

REFERENCES

1. Barresi, R.; Campbell, K. P., Dystroglycan: from biosynthesis to pathogenesis of human disease. *J Cell Sci* **2006,** 119, (Pt 2), 199-207.

2. Endo, T., O-mannosyl glycans in mammals. *Biochim Biophys Acta* **1999,** 1473, (1), 237-46.

3. Martin, P. T., Mechanisms of disease: congenital muscular dystrophies-glycosylation takes center stage. *Nat Clin Pract Neurol* **2006,** 2, (4), 222-30.

4. Martin, P. T., Congenital muscular dystrophies involving the O-mannose pathway. *Curr Mol Med* **2007,** 7, (4), 417-25.

5. Yoshida, A.; Kobayashi, K.; Manya, H.; Taniguchi, K.; Kano, H.; Mizuno, M.; Inazu, T.; Mitsuhashi, H.; Takahashi, S.; Takeuchi, M.; Herrmann, R.; Straub, V.; Talim, B.; Voit, T.; Topaloglu, H.; Toda, T.; Endo, T., Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. *Dev Cell* **2001,** 1, (5), 717-24.

6. Beltran-Valero de Bernabe, D.; Currier, S.; Steinbrecher, A.; Celli, J.; van Beusekom, E.; van der Zwaag, B.; Kayserili, H.; Merlini, L.; Chitayat, D.; Dobyns, W. B.; Cormand, B.; Lehesjoki, A. E.; Cruces, J.; Voit, T.; Walsh, C. A.; van Bokhoven, H.; Brunner, H. G., Mutations in the O-mannosyltransferase gene POMT1 give rise to the severe neuronal migration disorder Walker-Warburg syndrome. *Am J Hum Genet* **2002,** 71, (5), 1033-43.

7. 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.; Huynen, M. A.; Verrips, A.; Walsh, C. A.; Barth, P. G.; Brunner, H. G.; van Bokhoven, H., POMT2 mutations cause alpha-dystroglycan hypoglycosylation and Walker-Warburg syndrome. *J Med Genet* **2005,** 42, (12), 907-12.

8. Willer, T.; Prados, B.; Falcon-Perez, J. M.; Renner-Muller, I.; Przemeck, G. K.; Lommel, M.; Coloma, A.; Valero, M. C.; de Angelis, M. H.; Tanner, W.; Wolf, E.; Strahl, S.; Cruces, J., Targeted disruption of the Walker-Warburg syndrome gene Pomt1 in mouse results in embryonic lethality. *Proc Natl Acad Sci U S A* **2004,** 101, (39), 14126-31.

9. Liu, J.; Ball, S. L.; Yang, Y.; Mei, P.; Zhang, L.; Shi, H.; Kaminski, H. J.; Lemmon, V. P.; Hu, H., A genetic model for muscle-eye-brain disease in mice lacking protein O-mannose 1,2-N-acetylglucosaminyltransferase (POMGnT1). *Mech Dev* **2006,** 123, (3), 228-40.

10. Greener, M. J.; Roberts, R. G., Conservation of components of the dystrophin complex in Drosophila. *FEBS Lett* **2000,** 482, (1-2), 13-8.

11. Deng, W. M.; Schneider, M.; Frock, R.; Castillejo-Lopez, C.; Gaman, E. A.; Baumgartner, S.; Ruohola-Baker, H., Dystroglycan is required for polarizing the epithelial cells and the oocyte in Drosophila. *Development* **2003,** 130, (1), 173-84.

12. Schneider, M.; Khalil, A. A.; Poulton, J.; Castillejo-Lopez, C.; Egger-Adam, D.; Wodarz, A.; Deng, W. M.; Baumgartner, S., Perlecan and Dystroglycan act at the basal side of the Drosophila follicular epithelium to maintain epithelial organization. *Development* **2006,** 133, (19), 3805-15.

13. Haines, N.; Seabrooke, S.; Stewart, B. A., Dystroglycan and protein O-mannosyltransferases 1 and 2 are required to maintain integrity of Drosophila larval muscles. *Mol Biol Cell* **2007,** 18, (12), 4721-30.

14. Shcherbata, H. R.; Yatsenko, A. S.; Patterson, L.; Sood, V. D.; Nudel, U.; Yaffe, D.; Baker, D.; Ruohola-Baker, H., Dissecting muscle and neuronal disorders in a Drosophila model of muscular dystrophy. *EMBO J* **2007,** 26, (2), 481-93.

15. Bogdanik, L.; Framery, B.; Frolich, A.; Franco, B.; Mornet, D.; Bockaert, J.; Sigrist, S. J.; Grau, Y.; Parmentier, M. L., Muscle dystroglycan organizes the postsynapse and regulates presynaptic neurotransmitter release at the Drosophila neuromuscular junction. *PLoS One* **2008,** 3, (4), e2084.

16. Wairkar, Y. P.; Fradkin, L. G.; Noordermeer, J. N.; DiAntonio, A., Synaptic defects in a Drosophila model of congenital muscular dystrophy. *J Neurosci* **2008,** 28, (14), 3781-9.

17. Kucherenko, M. M.; Pantoja, M.; Yatsenko, A. S.; Shcherbata, H. R.; Fischer, K. A.; Maksymiv, D. V.; Chernyk, Y. I.; Ruohola-Baker, H., Genetic modifier screens reveal new components that interact with the Drosophila dystroglycan-dystrophin complex. *PLoS One* **2008,** 3, (6), e2418.

18. Ichimiya, T.; Manya, H.; Ohmae, Y.; Yoshida, H.; Takahashi, K.; Ueda, R.; Endo, T.; Nishihara, S., The twisted abdomen phenotype of Drosophila POMT1 and POMT2 mutants coincides with their heterophilic protein O-mannosyltransferase activity. *J Biol Chem* **2004,** 279, (41), 42638-47.

19. Lyalin, D.; Koles, K.; Roosendaal, S. D.; Repnikova, E.; Van Wechel, L.; Panin, V. M., The twisted gene encodes Drosophila protein O-mannosyltransferase 2 and genetically interacts with the rotated abdomen gene encoding Drosophila protein O-mannosyltransferase 1. *Genetics* **2006,** 172, (1), 343-53.

20. Manya, H.; Suzuki, T.; Akasaka-Manya, K.; Ishida, H. K.; Mizuno, M.; Suzuki, Y.; Inazu, T.; Dohmae, N.; Endo, T., Regulation of Mammalian Protein O-Mannosylation: PREFERENTIAL AMINO ACID SEQUENCE FOR O-MANNOSE MODIFICATION. *J Biol Chem* **2007,** 282, (28), 20200-6.

21. Breloy, I.; Schwientek, T.; Gries, B.; Razawi, H.; Macht, M.; Albers, C.; Hanisch, F. G., Initiation of mammalian O-mannosylation in vivo is independent of a consensus sequence and controlled by peptide regions within and upstream of the alpha-dystroglycan mucin domain. *J Biol Chem* **2008,** 283, (27), 18832-40.

22. North, S. J.; Koles, K.; Hembd, C.; Morris, H. R.; Dell, A.; Panin, V. M.; Haslam, S. M., Glycomic studies of Drosophila melanogaster embryos. *Glycoconj J* **2006,** 23, (5-6), 345-54.

23. Aoki, K.; Perlman, M.; Lim, J. M.; Cantu, R.; Wells, L.; Tiemeyer, M., Dynamic developmental elaboration of N-linked glycan complexity in the Drosophila melanogaster embryo. *J Biol Chem* **2007,** 282, (12), 9127-42.

24. Koles, K.; Lim, J. M.; Aoki, K.; Porterfield, M.; Tiemeyer, M.; Wells, L.; Panin, V., Identification of N-glycosylated proteins from the central nervous system of Drosophila melanogaster. *Glycobiology* **2007,** 17, (12), 1388-403.

25. Manya, H.; Chiba, A.; Yoshida, A.; Wang, X.; Chiba, Y.; Jigami, Y.; Margolis, R. U.; Endo, T., Demonstration of mammalian protein O-mannosyltransferase activity: coexpression of POMT1 and POMT2 required for enzymatic activity. *Proc Natl Acad Sci U S A* **2004,** 101, (2), 500-5.

26. Aoki, K.; Porterfield, M.; Lee, S. S.; Dong, B.; Nguyen, K.; McGlamry, K. H.; Tiemeyer, M., The diversity of O-linked glycans expressed during Drosophila melanogaster development reflects stage- and tissue-specific requirements for cell signaling. *J Biol Chem* **2008**.

27. Wells, L.; Vosseller, K.; Cole, R. N.; Cronshaw, J. M.; Matunis, M. J.; Hart, G. W., Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. *Mol Cell Proteomics* **2002,** 1, (10), 791-804.

28. Zacharius, R. M.; Zell, T. E.; Morrison, J. H.; Woodlock, J. J., Glycoprotein staining following electrophoresis on acrylamide gels. *Anal Biochem* **1969,** 30, (1), 148-52.

29. Gradilone, S. A.; Arranz, S. E.; Cabada, M. O., Detection of highly glycosylated proteins in polyacrylamide gels. *Anal Biochem* **1998,** 261, (2), 224-7.

30. Rebay, I.; Fehon, R. G.; Artavanis-Tsakonas, S., Specific truncations of Drosophila Notch define dominant activated and dominant negative forms of the receptor. *Cell* **1993,** 74, (2), 319-29.

31. Sun, X.; Artavanis-Tsakonas, S., The intracellular deletions of Delta and Serrate define dominant negative forms of the Drosophila Notch ligands. *Development* **1996,** 122, (8), 2465-74.

32. Christoforou, C. P.; Greer, C. E.; Challoner, B. R.; Charizanos, D.; Ray, R. P., The detached locus encodes Drosophila Dystrophin, which acts with other components of the Dystrophin Associated Protein Complex to influence intercellular signalling in developing wing veins. *Dev Biol* **2008,** 313, (2), 519-32.

33. Rottger, S.; White, J.; Wandall, H. H.; Olivo, J. C.; Stark, A.; Bennett, E. P.; Whitehouse, C.; Berger, E. G.; Clausen, H.; Nilsson, T., Localization of three human polypeptide GalNAc-transferases in HeLa cells suggests initiation of O-linked glycosylation throughout the Golgi apparatus. *J Cell Sci* **1998,** 111 (Pt 1), 45-60.

34. Lommel, M.; Bagnat, M.; Strahl, S., Aberrant processing of the WSC family and Mid2p cell surface sensors results in cell death of Saccharomyces cerevisiae O-mannosylation mutants. *Mol Cell Biol* **2004,** 24, (1), 46-57.

35. Proszynski, T. J.; Simons, K.; Bagnat, M., O-glycosylation as a sorting determinant for cell surface delivery in yeast. *Mol Biol Cell* **2004,** 15, (4), 1533-43.

36. Hirayama, H.; Fujita, M.; Yoko-o, T.; Jigami, Y., O-mannosylation is required for degradation of the endoplasmic reticulum-associated degradation substrate Gas1*p via the ubiquitin/proteasome pathway in Saccharomyces cerevisiae. *J Biochem* **2008,** 143, (4), 555-67.

37. Kunz, S.; Sevilla, N.; McGavern, D. B.; Campbell, K. P.; Oldstone, M. B., Molecular analysis of the interaction of LCMV with its cellular receptor [alpha]-dystroglycan. *J Cell Biol* **2001,** 155, (2), 301-10.

38. Koles, K.; Irvine, K. D.; Panin, V. M., Functional characterization of Drosophila sialyltransferase. *J Biol Chem* **2004,** 279, (6), 4346-57.

39. Brand, A. H.; Manoukian, A. S.; Perrimon, N., Ectopic expression in Drosophila. *Methods Cell Biol* **1994,** 44, 635-54.

40. Panin, V. M.; Papayannopoulos, V.; Wilson, R.; Irvine, K. D., Fringe modulates Notch-ligand interactions. *Nature* **1997,** 387, (6636), 908-12.

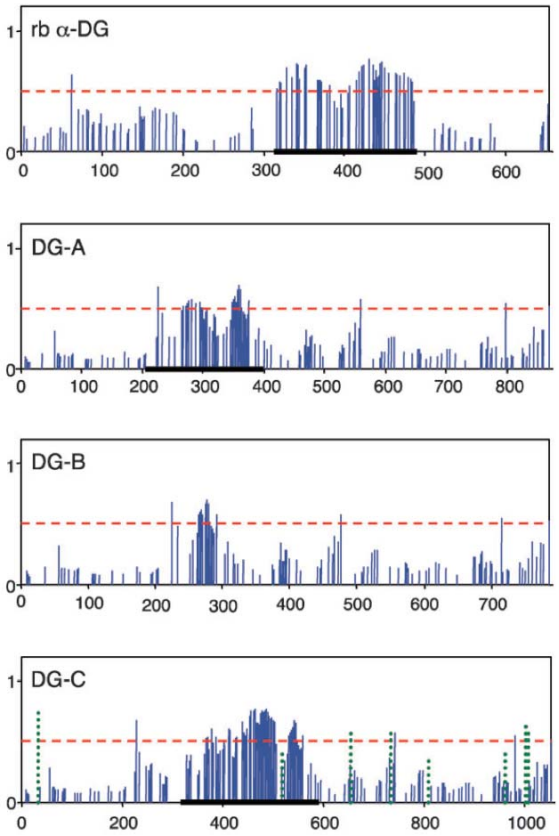
41. Bendtsen, J. D.; Nielsen, H.; von Heijne, G.; Brunak, S., Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **2004,** 340, (4), 783-95.

42. Julenius, K.; Molgaard, A.; Gupta, R.; Brunak, S., Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* **2005,** 15, (2), 153-64.

Figure B-1. Predicted glycosylation and in vitro *O*-mannosylation of *Drosophila* Dystroglycan.

(**a**) Prediction of mucin-type *O*-linked glycosylation of the extracellular domains of rabbit α-DG and *Drosophila* DG isoforms A, B, and C. Solid vertical bars show the G-score of corresponding S/T residues with respect to the glycosylation potential[41]; the predicted glycosylation sites have bars above the threshold (horizontal dashed line). DC-C panel also includes the prediction of *N*-linked glycosylation (dotted vertical bars). Regions of rabbit α-DG, DG-A, and DG-C proteins used in vitro *O*-mannosylation assays are underlined. (**b**) *O*-Mannosylation assay of *Drosophila* DG-A and DG-C proteins. Purified fragments of extracellular domain of DG-A and DG-C isoforms were used as substrates in vitro *O*-mannosylation assays with microsomal fraction from *Drosophila* larvae as a source of RT-TW mannosyltransferase activity and [^{3}H]-mannosyl phosphoryl dolicol as a sugar donor. Incorporation of mannose is shown as the ratio of incorporated radioactivity for a substrate to that for BSA as a mock control. Error bars indicate SEM calculated from six independent assays. ** and * – indicate significant differences from the mock control with *t*-test $P < 0.01$ and $P < 0.05$, respectively. Note that the results for DG-C and rabbit α-DG were also significantly different from each other (*t*-test, $P < 0.05$), while incorporation of *O*-mannose in DG-A was not significantly different from the background control (*t*-test, $P > 0.4$).

a)


rb α-DG
DG-A
DG-B
DG-C

b)

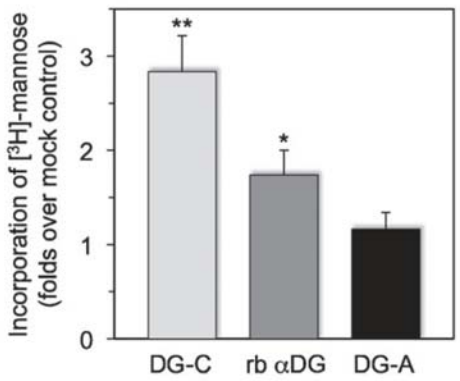
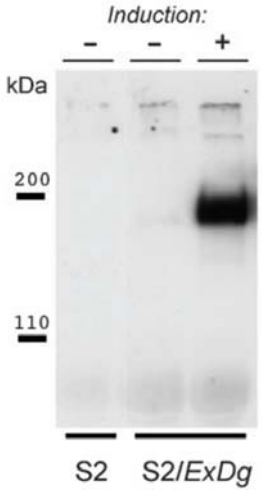

Incorporation of [3H]-mannose
(folds over mock control)
**
*
DG-C
rb αDG
DG-A

Figure B-2. Secretion and subcellular localization of the ExDG protein.
(**a**) Anti-FLAG Western blot of tissue culture media from S2 cells (control) or from S2 cells transfected with ExDG-expressing construct (with or without induction of expression). The results show that ExDG is efficiently secreted in a diffusible form outside of the cell. (**b**) Expression of ExDG in the third instar larval wing imaginal disks with a *patched-GAL4* driver using the UAS-GAL4 system. Genotype of the disks: left disk (*ptc::ExDg rt*−) – *ptc-GAL4 UAS-GFP/UAS-ExDg; rt*P/*rt2*; middle disk (*ptc::ExDg*) – *ptc-GAL4 UAS-GFP/+; UAS-ExDg/+*; right disk (*ptc::ExDg+rt+tw*) – *ptc-GAL4 UAS-GFP/UAS-rt UAS-tw; UAS-ExDg/+*. ExDG expression is detected by immunofluorescent staining (red), while GFP signal (green) highlights the pattern of the *ptc* driver. Yellow dashed line indicates the position of Z cross-sections reconstructed for each disk in panel Z. Z cross-sections: no accumulation of the ExDG protein can be detected inside the columnar epithelium cells (asterisks), while the protein is efficiently delivered to the basal (arrows) or apical (arrowheads) surfaces of the disk epithelium. There is no significant difference in the subcellular localization of ExDG between the three genotypes. For all panels: dorsal is to the top; for the frontal-view sections: anterior is to the left; for Z cross-sections: basal is to the left.

a)


Induction:
–
–
+
kDa
200
110
S2
S2/ExDg

b)

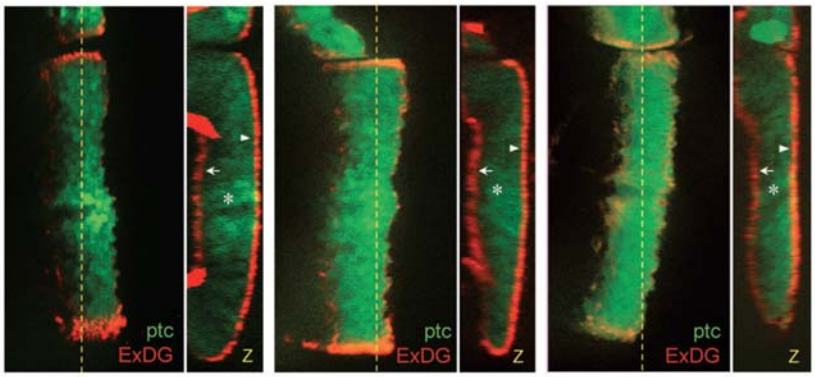
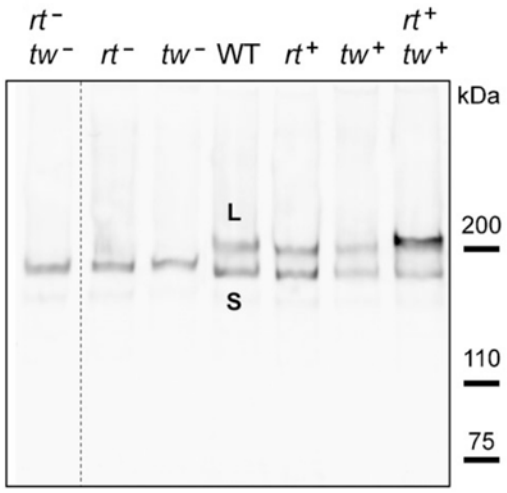

ptc
ExDG
z
ptc::ExDg rt⁻
ptc::ExDg
ptc::ExDg+rt+tw

Figure B-3. Western and lectin blot analyses of in vivo expressed ExDG.
(**a**) Western blot detection of ExDG expressed in *rt-tw* double mutants (*rt– tw–*), *rt* mutants (*rt–*), *tw* mutants (*tw–*), wild-type background (WT), and backgrounds with ubiquitous ectopic expression of RT (*rt+*), TW (*tw+*), or RT-TW co-expression (*rt+ tw+*). The L band is the top band present in wild-type background and backgrounds with RT and TW expression, but absent in *rt* and/or *tw* mutants; the S band is present in all genetic backgrounds analyzed. (**b)** The ratio between the intensities of L and S bands was quantified for *rt* mutant, *tw* mutant, wild type, and RT-TW co-expression backgrounds. Error bars represent SD.

a)


rt⁻ tw⁻
rt⁻
tw⁻
WT
rt⁺
tw⁺
rt⁺ tw⁺
kDa
L
S
200
110
75

b)

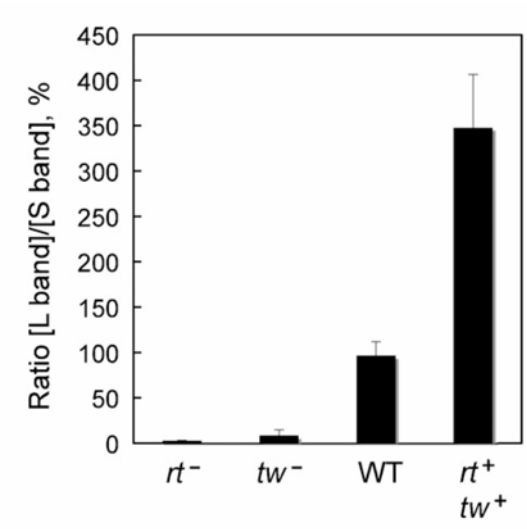
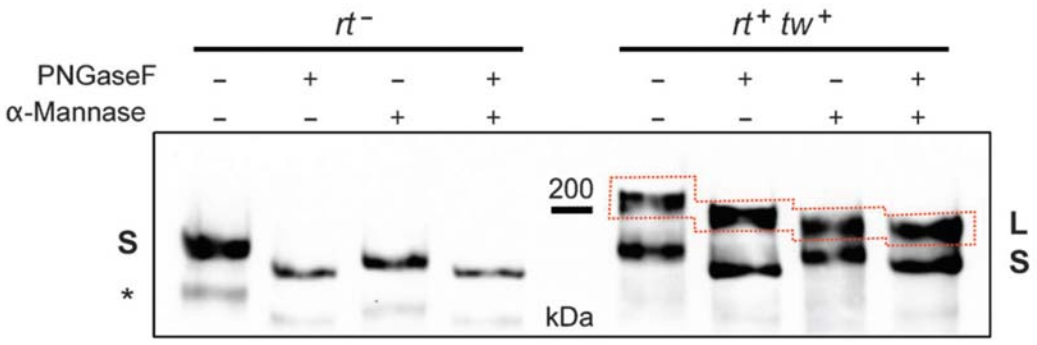

Ratio [L band]/[S band], %
450
400
350
300
250
200
150
100
50
0
rt⁻
tw⁻
WT
rt⁺ tw⁺

Figure B-4. Analysis of ExDG glycosylation by glycosidase treatments.
(**a**) left panel: ExDG purified from *rt* mutants was treated with PNGaseF or α-mannosidase, or with both glycosidases sequentially. No additional shift of the ExDG band (S band) is detected in double PNGaseF/α-mannosidase treatment as compared to PNGaseF treatment alone, suggesting that ExDG has no *O*-mannose modification. Right panel: glycosidase treatments of ExDG purified from RT-TW coexpression background. Top (L) band shows significant loss of mass (≥10 kDa) when PNGaseF-treated ExDG was digested with α-mannosidase, which suggests the presence of abundant *O*-mannose modifications. No such loss of mass was detected for the lower (S) band, suggesting that *O*-mannosylation of ExDG in this band is not significant. (**b**) Con A reactivity of purified ExDG after treatments with PNGaseF and α-mannosidase. The S glycoform purified from *rt* mutant background loses Con A reactivity either after the removal of *N*-linked glycans by PNGaseF or after treatment with α-mannosidase removing α-linked mannose residues (top panel, left), suggesting the absence of *O*-mannose modification and efficient removal of oligomannose structures either by trimming *N*-linked branches with α-mannosidase or by complete elimination of *N*-linked glycans. At the same time, the L glycoform purified from RT-TW co-expression background retains Con A reactivity after treatment with PNGaseF, α-mannosidase, or both glycosidases (top panel, right), suggesting that L glycoform is *O*-mannosylated, and that α-mannosidase does not remove *O*-mannose completely. The bottom panel shows anti-FLAG western control corresponding to the lectin blot. Red dashed line outlines the region of the L glycoform on the blots. Asterisk "*****" indicates an additional minor band sometimes detected by FLAG Western blots that probably represents ExDG proteolytic degradation.

a)


rt⁻
rt⁺ tw⁺
PNGaseF
α-Mannase
S
*
200
kDa
L
S

b)

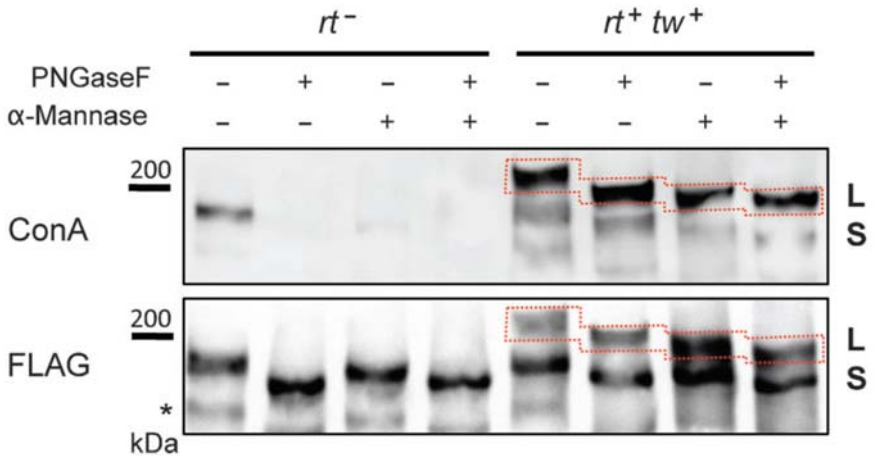
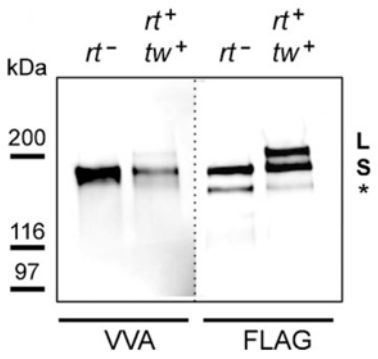

rt⁻
rt⁺ tw⁺
PNGaseF
α-Mannase
200
ConA
L
S
200
FLAG
*
kDa
L
S

Figure B-5. Analysis of ExDG glycosylation by lectin blots.
ExDG was purified from *rt* mutant or RT-TW coexpression backgrounds. (**a**) VVA lectin blot (left two lanes) and anti-FLAG western control for protein amount (right two lanes). The S band from RT-TW overexpression has significantly weaker VVA reactivity than that from *rt* mutants, while the reactivity of the L band is even further reduced to a nearly undetectable level. (**b**) Anti-FLAG western control (left two lanes), WFA and HHL lectin blots (middle and right two lanes, respectively). While the WFA reactivity of ExDG bands shows similar pattern to VVA staining in (**a**), the reactivity to HHL is complementary to VVA and WFA stainings. Only the L band reveals strong HHL staining, with the S band from RT-TW coexpression showing much weaker staining, and the S band from *rt* mutants having no HHL reactivity at all. ExDG was treated with PNGaseF (de-*N*-glycosylated) in (**b**). Asterisk indicates an additional minor band representing some proteolytic degradation.

a)


kDa
rt⁻
rt⁺ tw⁺
rt⁻
rt⁺ tw⁺
200
116
97
L
S
*
VVA
FLAG

b)

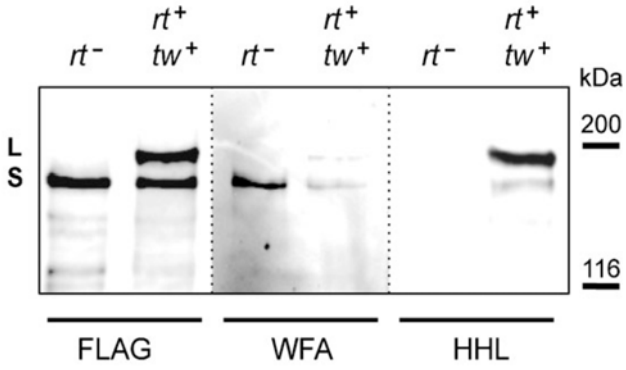
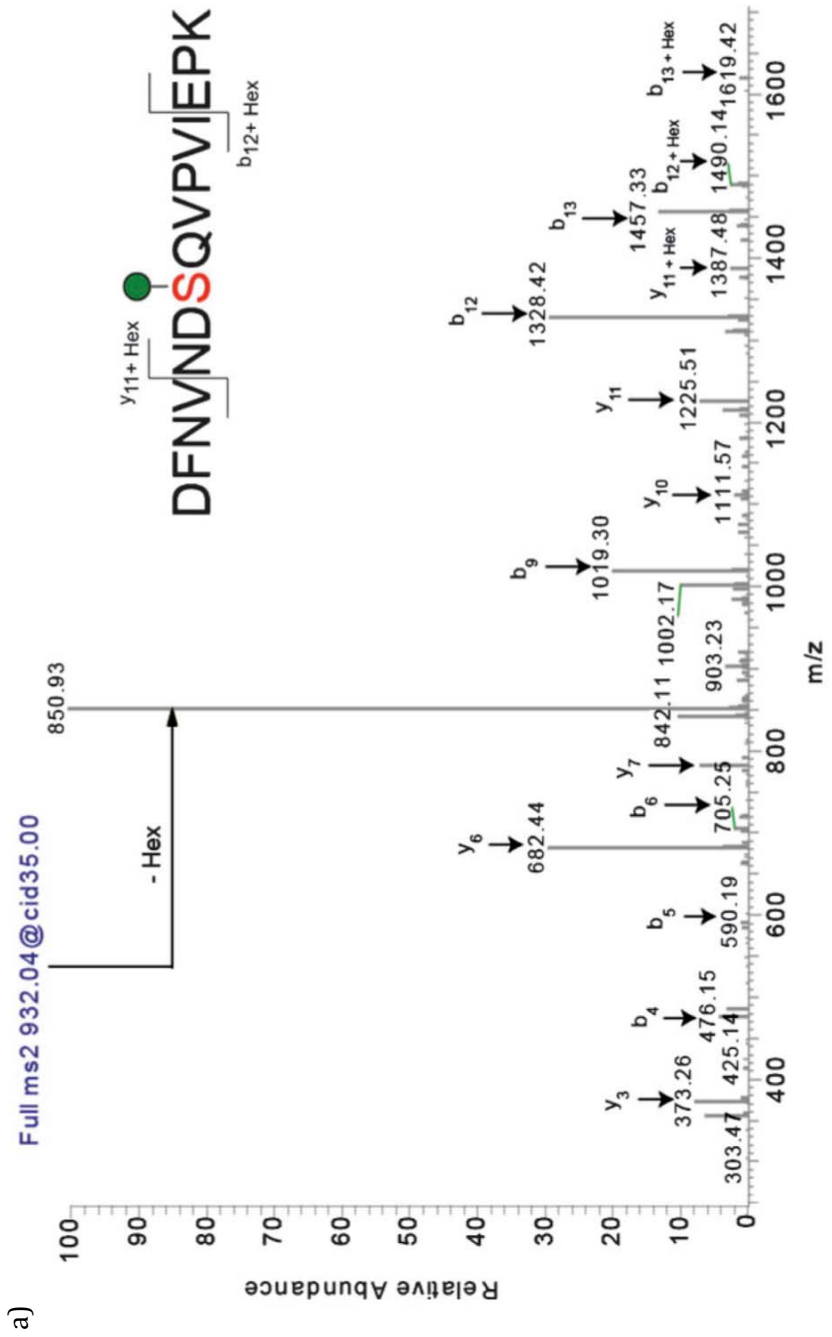

rt⁻
rt⁺ tw⁺
rt⁻
rt⁺ tw⁺
rt⁻
rt⁺ tw⁺
kDa
L
S
200
116
FLAG
WFA
HHL

Figure B-6. Identification of *O*-linked sites on ExDG glycopeptides using mass spectrometry.
The pseudo-neutral loss method was applied using LC-MS/MS allowing for data-dependent MS^3 fragmentation upon witnessing a neutral loss of a glycan mass in the MS^2 spectra. Through the application of the pseudo-neutral loss method, it was possible to confirm the assignment of the glycosylated peptide that was identified using TurboSequest. Panel **a** shows the MS^2 of an *O*-mannose-modified peptide, while panel **b** shows the MS^2 of an *O*-GalNAc-modified peptide. Note that in both cases the major fragment is the loss of the glycan. Using the MS^2 profile (shown) and the MS^3 profile of the neutral-loss species (the unmodified peptide, not shown) allowed us to assign the peptide, the glycan, and the site of modification in an automated fashion using TurboSequest followed by manual validation. Sites of modification were manually validated by looking for the addition of Hex and HexNAc to the theoretical *m*/*z* of b- and y-ions for the glycopeptide.


a)
Full ms2 932.04@cid35.00
- Hex
DFNVNDSQVPVIEPK
y11+ Hex
b12+ Hex
Relative Abundance
m/z
850.93
842.11
1002.17
1019.30
1328.42
1457.33
1225.51
1111.57
682.44
705.25
903.23
373.26
425.14
476.15
590.19
303.47
1387.48
1490.14
1619.42

b)

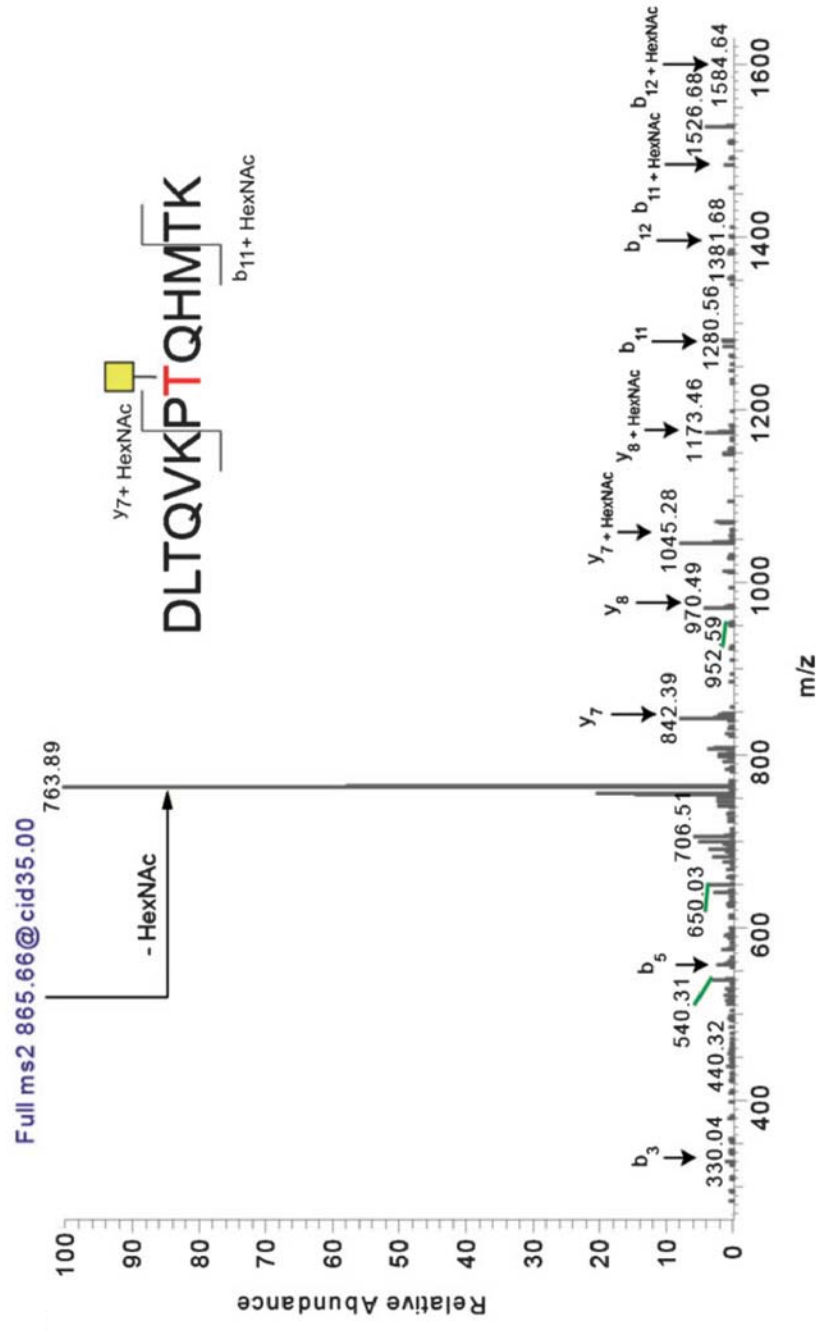

Full ms2 865.66@cid35.00
- HexNAc
763.89
DLTQVKPTQHMTK
y7+ HexNAc
b11+ HexNAc
Relative Abundance
m/z
b3
330.04
440.32
b5
540.31
650.03
706.51
y7
842.39
y8
952.59
970.49
y7 + HexNAc
1045.28
y8 + HexNAc
1173.46
b11
1280.56
b12
1381.68
b11 + HexNAc
1526.68
b12 + HexNAc
1584.64

Figure B-7. Glycosylation of the DG-C extracellular domain.
Proposed glycosylation profiles of L and S glycoforms of the DG-C extracellular domain based on mass spectrometry analyses, glycosidase treatments, and lectin reactivity. Orange box indicates the mucin-type domain of DG-C; hatched region shows the part of the full-length DG-C protein that was truncated in ExDG (including the transmembrane domain shown in blue). N–C termini orientation is from left to right.

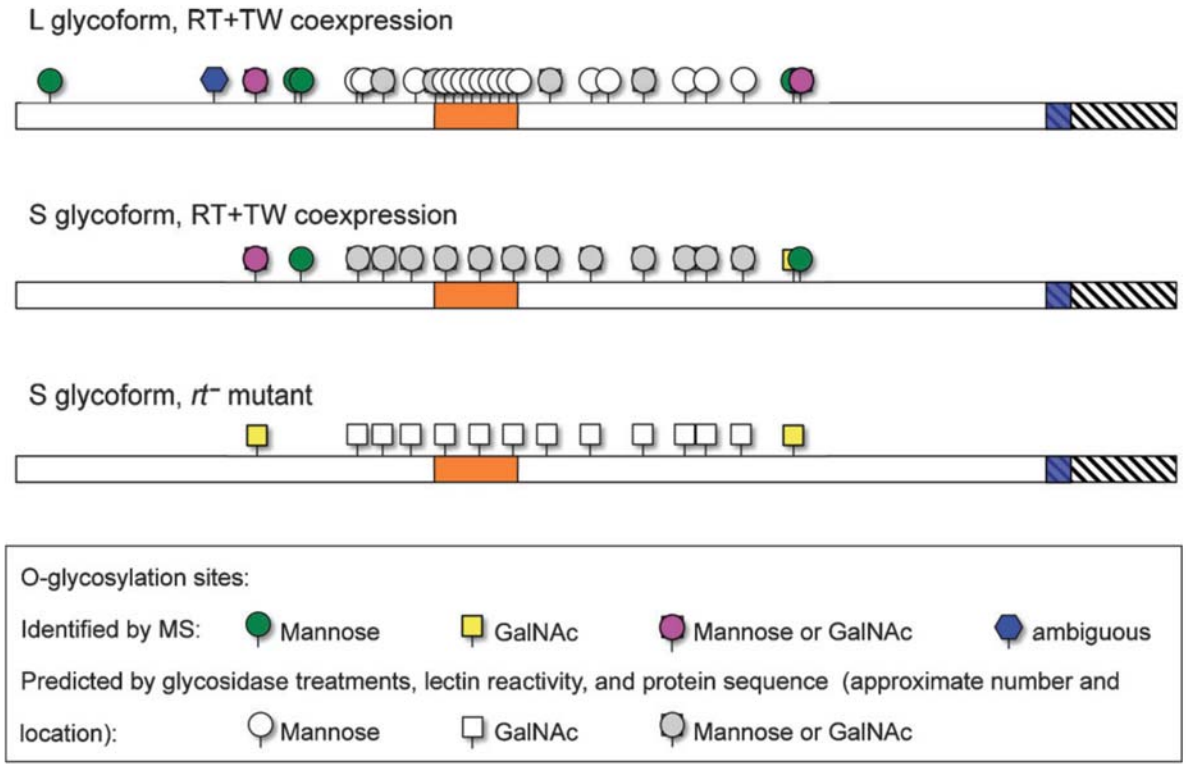

L glycoform, RT+TW coexpression
S glycoform, RT+TW coexpression
S glycoform, rt⁻ mutant
O-glycosylation sites:
Identified by MS:
Mannose
GalNAc
Mannose or GalNAc
ambiguous
Predicted by glycosidase treatments, lectin reactivity, and protein sequence (approximate number and location):
Mannose
GalNAc
Mannose or GalNAc

Figure B-8. Analysis of trypsin-digested ExDG.

Trypsin digestion of ExDG purified from RT-TW coexpression generates a ~120 kDa trypsin-resistant glycopeptide that can be detected with PAS or PAS-silver staining. The glycopeptide quantitatively binds to Con A agarose and changes its gel mobility upon α-mannosidase treatment. PAS, PAS-staining of ExDG after trypsin digestion. Other lanes show PAS-silver staining: Input, amount control for the ExDG tryptic glycopeptide used for binding to lectin agarose beads; ConA, WFA, VVA, the glycopeptide bound to Con A, WFA, and VVA beads, respectively; α-Mannase "-" and "+", α-mannosidase treatment of the glycopeptide with heat-inactivated (control) and active α-mannosidase, respectively. Note that PNGaseF-treated ExDG was used.

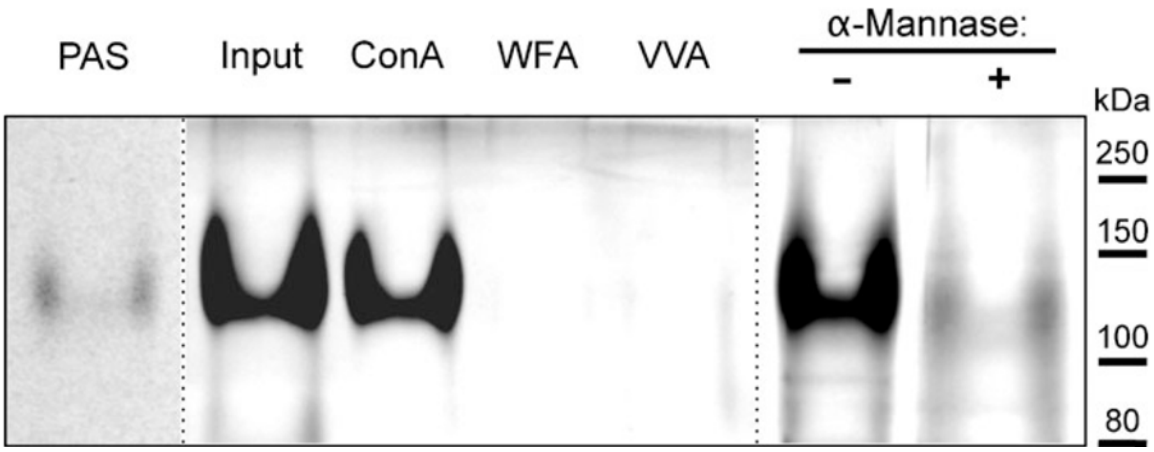
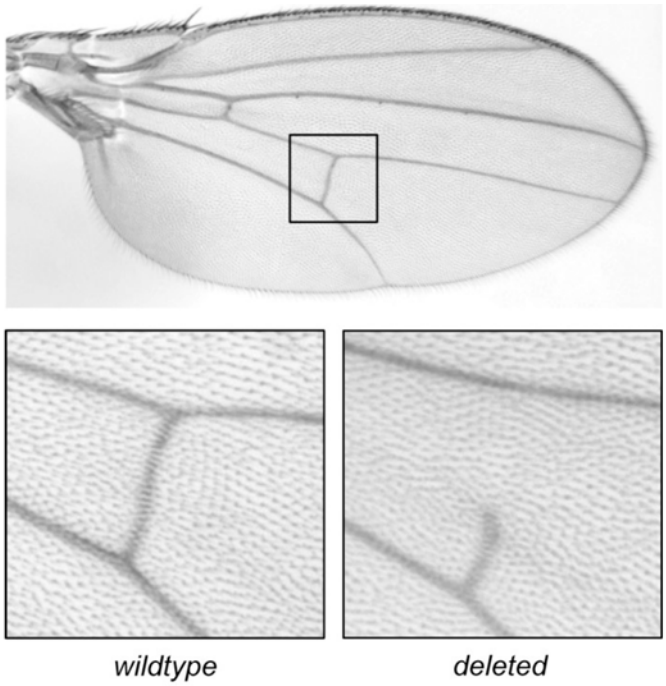

PAS
Input
ConA
WFA
VVA
α-Mannase:
-
+
kDa
250
150
100
80

Figure B-9. Effect of ExDG ectopic expression on the posterior crossvein development.
(**a**) Top panel, wild-type *Drosophila* wing with the region of posterior crossvein shown by a rectangle. Bottom panels show enlarged examples of a wild-type and a defective crossvein (with deleted anterior part) as a result of ExDG ectopic expression using an *Act-GAL4* driver. The phenotype varied from a severe 'deletion' of the crossvein (shown) to a less extreme but always obvious loss of crossvein tissue. (**b**) Penetrance of the crossvein defect is enhanced by RT-TW activity. *N* indicates the number of flies scored for each genotype. ExDG encompasses the complete extracellular domain of DG-C (a.a. 1–1048).

a)


wildtype
deleted

b)

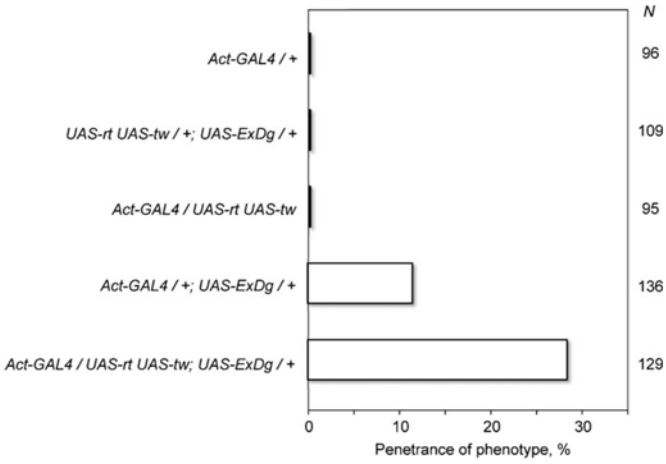

N
Act-GAL4 / +
96
UAS-rt UAS-tw / +; UAS-ExDg / +
109
Act-GAL4 / UAS-rt UAS-tw
95
Act-GAL4 / +; UAS-ExDg / +
136
Act-GAL4 / UAS-rt UAS-tw; UAS-ExDg / +
129
0
10
20
30
Penetrance of phenotype, %

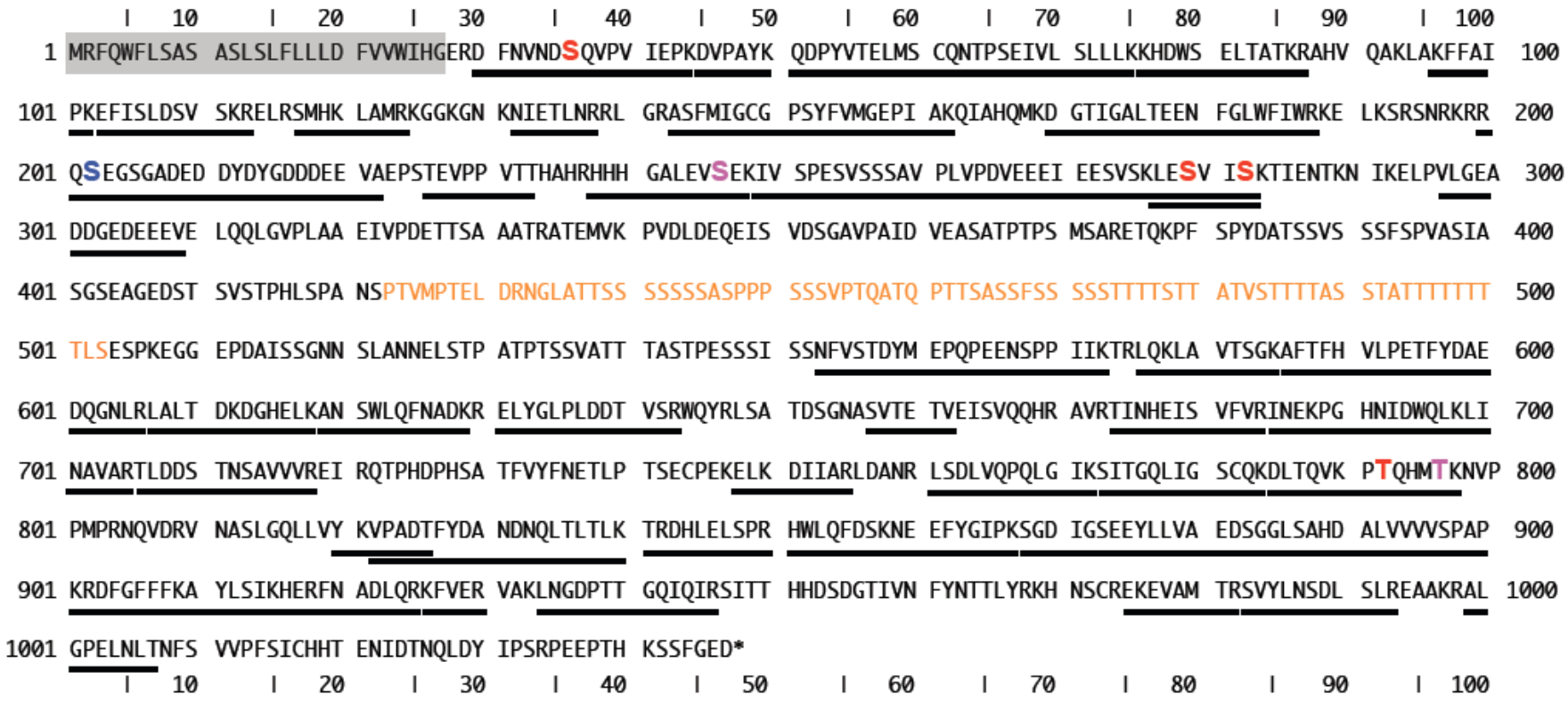
Table B-1. Summary of ExDG glycopeptides with *O*-linked mannose and GalNAc modifications identified by mass spectrometry.

Glycoform S *rt* mutants	Glycoform S RT-TW coexpression RT-TW coexpression	Glycoform L RT-TW coexpression
Glycopeptides identified	*Glycopeptides identified*	*Glycopeptides identified*
With *O*-mannose:	With *O*-mannose:	With *O*-mannose:
		30 DFNVND**S**QVPVIEPK 44
	238 HHHGALEV**S**EK 248	238 HHHGALEV**S**EK 248
—	277 LESVI**S**K 283	277 LE**S**VI**S**K 283
	785 DLTQVKPTQHM**T**K 797	785 DLTQVKP**T**QHMTK 797
		785 DLTQVKPTQHM**T**K 797
With O-GalNAc:	With O-GalNAc:	With O-GalNAc:
238 HHHGALEVSEK 248	238 HHHGALEV**S**EK 248	238 HHHGALEV**S**EK 248
785 DLTQVKPTQHMTK 797	785 DLTQVKP**T**QHMTK 797	785 DLTQVKPTQHM**T**K 797
Other/ambiguous:	Other/ambiguous:	Other/ambiguous:
—	—	200 RQ**S**EGSGADEDDYDYGDDDEEVA 222

Table B-S1. Peptide coverage for purified ExDG glycoforms analyzed by mass spectrometry.

Glycoform S, *rt* mutants			Glycoform S, RT-TW coexpression			Glycoform L, RT-TW coexpression		
Protein Coverage (all peptides) 46.8%			Protein Coverage (all peptides) 44.8%			Protein Coverage (all peptides) 47.2%		
Peptides Identified	**Start**	**End**	**Peptides Identified**	**Start**	**End**	**Peptides Identified**	**Start**	**End**
DFNVNDSQVPVIEPK	30	44	DFNVNDSQVPVIEPK	30	44	DFNVNDSQVPVIEPK	30	44
DVPAYK	45	50	DVPAYK	45	50	DVPAYK	45	50
QDPYVTELMSCQNTPSEIVLSLLLK	51	75	QDPYVTELMSCQNTPSEIVLSLLLK	51	75	QDPYVTELMSCQNTPSEIVLSLLLK	51	75
KHDWSELTATKR	76	87	KHDWSELTATKR	76	87	KHDWSELTATKR	76	87
KFFAIPK	96	102	KFFAIPK	96	102	KFFAIPK	96	102
EFISLDSVSKR	103	113	EFISLDSVSKR	103	113	EFISLDSVSKR	103	113
-	-	-	-	-	-	SMHKLAMR	117	124
NIETLNR	132	138	NIETLNR	132	138	NIETLNR	132	138
ASFMIGCGPSYFVMGEPIAK	143	162	ASFMIGCGPSYFVMGEPIAK	143	162	ASFMIGCGPSYFVMGEPIAK	143	162
DGTIGALTEENFGLWFIWR	170	188	DGTIGALTEENFGLWFIWR	170	188	DGTIGALTEENFGLWFIWR	170	188
-	-	-	-	-	-	RQSEGSGADEDDYDYGDDDEEVA	200	222
TEVPPVTT	226	233	TEVPPVTT	226	233	TEVPPVTT	226	233
HHHGALEVSEK	238	248	HHHGALEVSEK	238	248	HHHGALEVSEK	238	248
IVSPESVSSSAVPLVPDVEEEIEESVSKLESVISK	249	283	IVSPESVSSSAVPLVPDVEEEIEESVSKLESVISK	249	283	IVSPESVSSSAVPLVPDVEEEIEESVSKLESVISK	249	283
LGEADDGEDEEEV	297	309	LGEADDGEDEEEV	297	309	LGEADDGEDEEEV	297	309
AAEIVPDET	319	327	-	-	-	-	-	-
-	-	-	EIVPDET	321	327	-	-	-
ETQKPFSPYDATS	375	387	-	-	-	-	-	-
STPHLSPA	413	420	-	-	-	-	-	-
ESPKEGGEPDAI	504	515	ESPKEGGEPDAI	504	515	-	-	-
STDYMEPQPEENSPPIIK	556	573	NFVSTDYMEPQPEENSPPIIK	553	573	NFVSTDYMEPQPEENSPPIIK	553	573
LAVTSGK	579	585	LQKLAVTSGK	576	585	LQKLAVTSGK	576	585
AFTFHVLPETFYDAEDQGNLR	586	606	AFTFHVLPETFYDAEDQGNLR	586	606	AFTFHVLPETFYDAEDQGNLR	586	606
LALTDKDGHELK	607	618	LALTDKDGHELK	607	618	LALTDKDGHELK	607	618
ANSWLQFNADKR	619	630	ANSWLQFNADK	619	629	ANSWLQFNADK	619	629
ELYGLPLDDTVSR	631	643	ELYGLPLDDTVSR	631	643	ELYGLPLDDTVSR	631	643
-	-	-	-	-	-	SVTETV	657	662
TINHEISVFVR	674	684	TINHEISVFVR	674	684	TINHEISVFVR	674	684
INEKPGHNIDWQLK	685	698	INEKPGHNIDWQLK	685	698	INEKPGHNIDWQLK	685	698
LINAVAR	699	705	LINAVAR	699	705	LINAVAR	699	705
TLDDSTNSAVVVR	706	718	TLDDSTNSAVVVR	706	718	TLDDSTNSAVVVR	706	718
ELKDIIAR	748	755	ELKDIIAR	748	755	ELKDIIAR	748	755
LSDLVQPQLGIK	761	772	LSDLVQPQLGIK	761	772	LSDLVQPQLGIK	761	772
SITGQLIGSCQK	773	784	SITGQLIGSCQK	773	784	SITGQLIGSCQK	773	784
DLTQVKPTQHMTK	785	797	DLTQVKPTQHMTK	785	797	DLTQVKPTQHMTK	785	797
VNASLGQLLVYK	810	821	-	-	-	-	-	-
-	-	-	-	-	-	YKVPADT	820	826
VPADTFYDANDNQLTLTLK	822	840	VPADTFYDANDNQLTLTLK	822	840	VPADTFYDANDNQLTLTLK	822	840
TRDHLELSPR	841	850	TRDHLELSPR	841	850	TRDHLELSPR	841	850
HWLQFDSKNEEFYGIPK	851	867	HWLQFDSKNEEFYGIPK	851	867	HWLQFDSKNEEFYGIPK	851	867
SGDIGSEEYLLVAEDSGGLSAHDALVVVVSPAPK	868	901	SGDIGSEEYLLVAEDSGGLSAHDALVVVVSPAPK	868	901	SGDIGSEEYLLVAEDSGGLSAHDALVVVVSPAPK	868	901
RDFGFFFK	902	909	RDFGFFFK	902	909	RDFGFFFK	902	909
-	-	-	AYLSIKHER	910	918	AYLSIKHER	910	918
FNADLQR	919	925	FNADLQR	919	925	FNADLQR	919	925
-	-	-	KFVER	926	930	KFVER	926	930
LNGDPTTGQIQIR	934	946	LNGDPTTGQIQIR	934	946	LNGDPTTGQIQIR	934	946
EKEVAMTR	975	982	-	-	-	EKEVAMTR	975	982
SVYLNSDLSLR	983	993	SVYLNSDLSLR	983	993	SVYLNSDLSLR	983	993
ALGPELNLT	999	1007	ALGPELNLT	999	1007	ALGPELNLT	999	1007

Figure B-S1.Peptide coverage and O-glycosylation sites identified by mass spectrometry on L glycoform of ExDG protein purified from RT-TW co-expression background.
Detected peptides are underlined. Sites of O-mannosylation are highlighted in red; residues carrying either O-mannose or O-GalNAc are highlighted in pink; light blue residue indicates the site carrying an O-linked sugar which was not unambiguously identified. Grey box shows the signal peptide that is predicted to be cleaved. The location of mucin-type domain is highlighted in orange. Asterisk indicates the position of 3xFLAG epitope tag fusion.


10 20 30 40 50 60 70 80 90 100
1 MRFQWFLSAS ASLSLFLLLD FVVWIHGERD FNVNDSQVPV IEPKDVPAYK QDPYVTELMS CQNTPSEIVL SLLLKKHDWS ELTATKRAHV QAKLAKFFAI 100
101 PKEFISLDSV SKRELRSMHK LAMRKGGKGN KNIETLNRRL GRASFMIGCG PSYFVMGEPI AKQIAHQMKD GTIGALTEEN FGLWFIWRKE LKSRSNRKRR 200
201 QSEGSGADED DYDYGDDDEE VAEPSTEVPP VTTHAHRHHH GALEVSEKIV SPESVSSSAV PLVPDVEEEI EESVSKLESV ISKTIENTKN IKELPVLGEA 300
301 DDGEDEEEVE LQQLGVPLAA EIVPDETTSA AATRATEMVK PVDLDEQEIS VDSGAVPAID VEASATPTPS MSARETQKPF SPYDATSSVS SSFSPVASIA 400
401 SGSEAGEDST SVSTPHLSPA NSPTVMPTEL DRNGLATTSS SSSSSASPPP SSSVPTQATQ PTTSASSFSS SSSTTTTSTT ATVSTTTTAS STATTTTTTT 500
501 TLSESPKEGG EPDAISSGNN SLANNELSTP ATPTSSVATT TASTPESSSI SSNFVSTDYM EPQPEENSPP IIKTRLQKLA VTSGKAFTFH VLPETFYDAE 600
601 DQGNLRLALT DKDGHELKAN SWLQFNADKR ELYGLPLDDT VSRWQYRLSA TDSGNASVTE TVEISVQQHR AVRTINHEIS VFVRINEKPG HNIDWQLKLI 700
701 NAVARTLDDS TNSAVVVREI RQTPHDPHSA TFVYFNETLP TSECPEKELK DIIARLDANR LSDLVQPQLG IKSITGQLIG SCQKDLTQVK PTQHMTKNVP 800
801 PMPRNQVDRV NASLGQLLVY KVPADTFYDA NDNQLTLTLK TRDHLELSPR HWLQFDSKNE EFYGIPKSGD IGSEEYLLVA EDSGGLSAHD ALVVVVSPAP 900
901 KRDFGFFFKA YLSIKHERFN ADLQRKFVER VAKLNGDPTT GQIQIRSITT HHDSDGTIVN FYNTTLYRKH NSCREKEVAM TRSVYLNSDL SLREAAKRAL 1000
1001 GPELNLTNFS VVPFSICHHT ENIDTNQLDY IPSRPEEPTH KSSFGED*
10 20 30 40 50 60 70 80 90 100

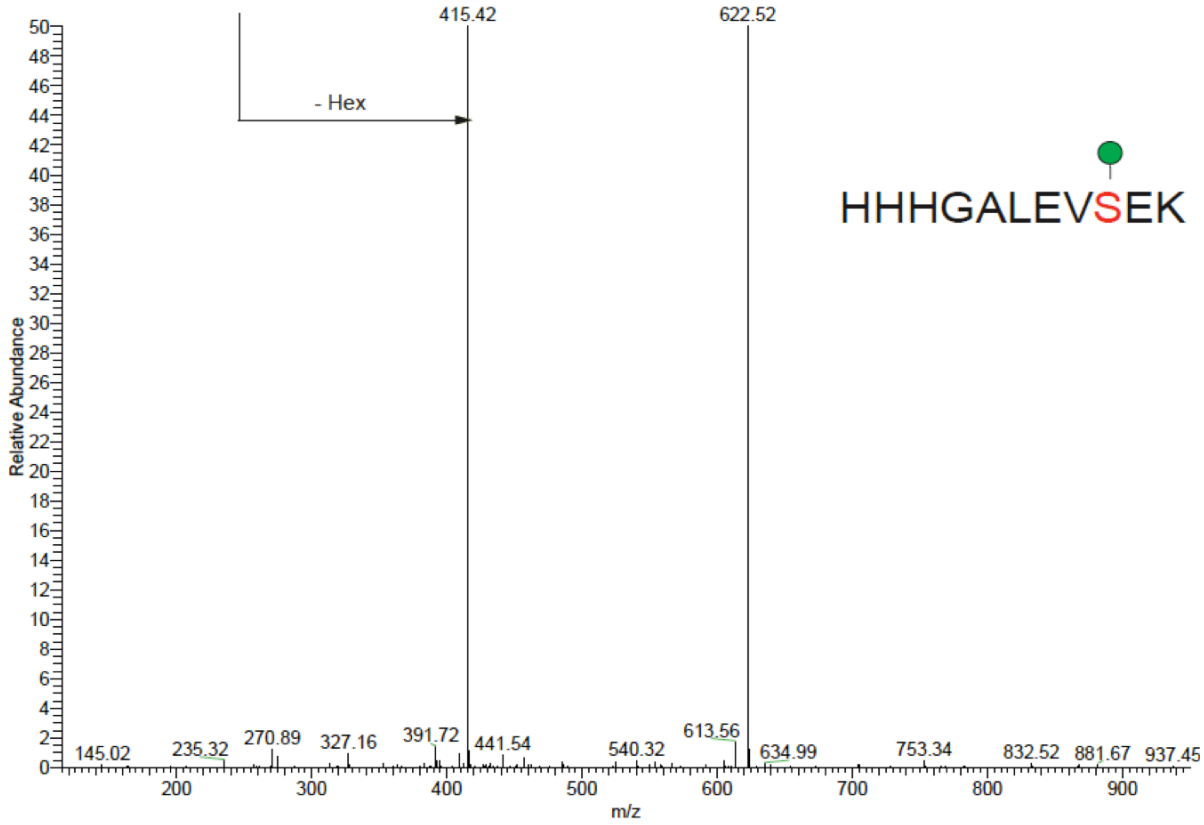


Figure B-S2. The spectrum of the MS/MS profile of an O-mannosylated peptide observed from glycoform S. This peptide was also observed in glycoform L.

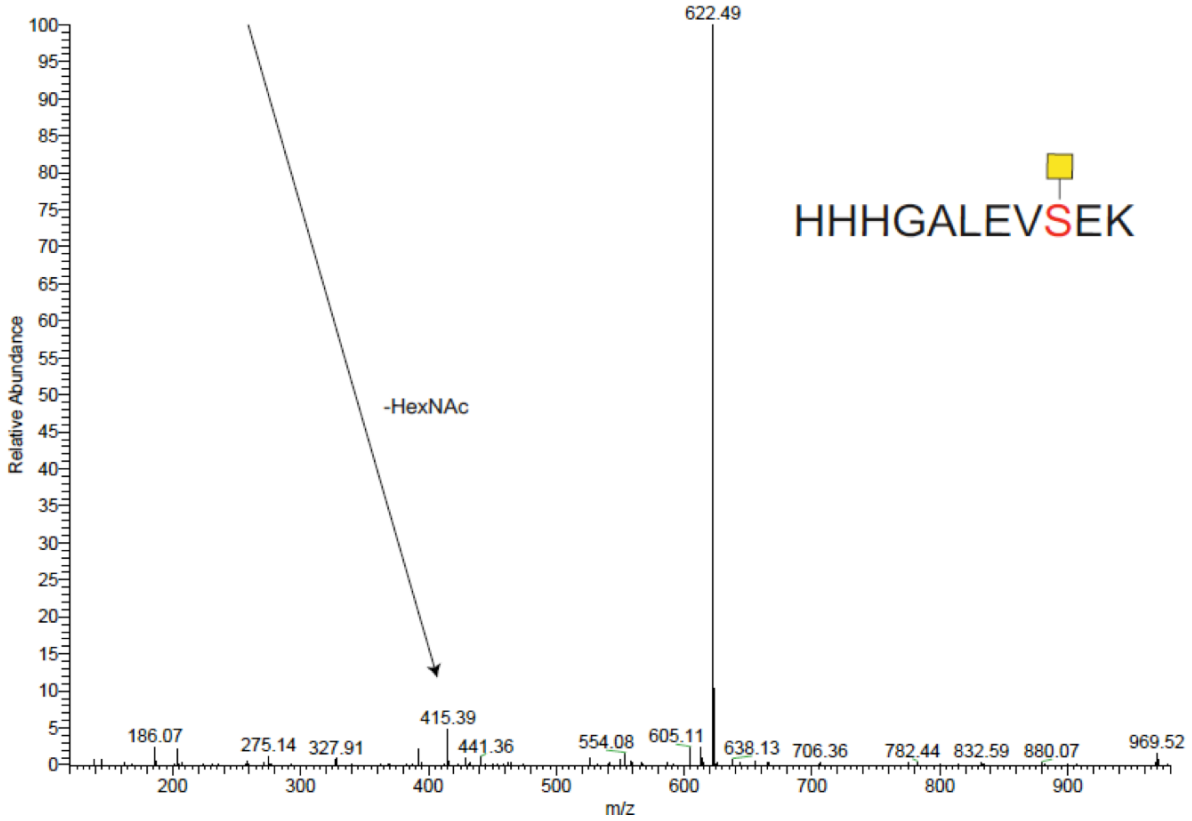


Figure B-S3. The spectrum of the MS/MS profile of an O-GalNAc modified peptide that was observed in both the S and L glycoforms as well as the *rt* mutants.

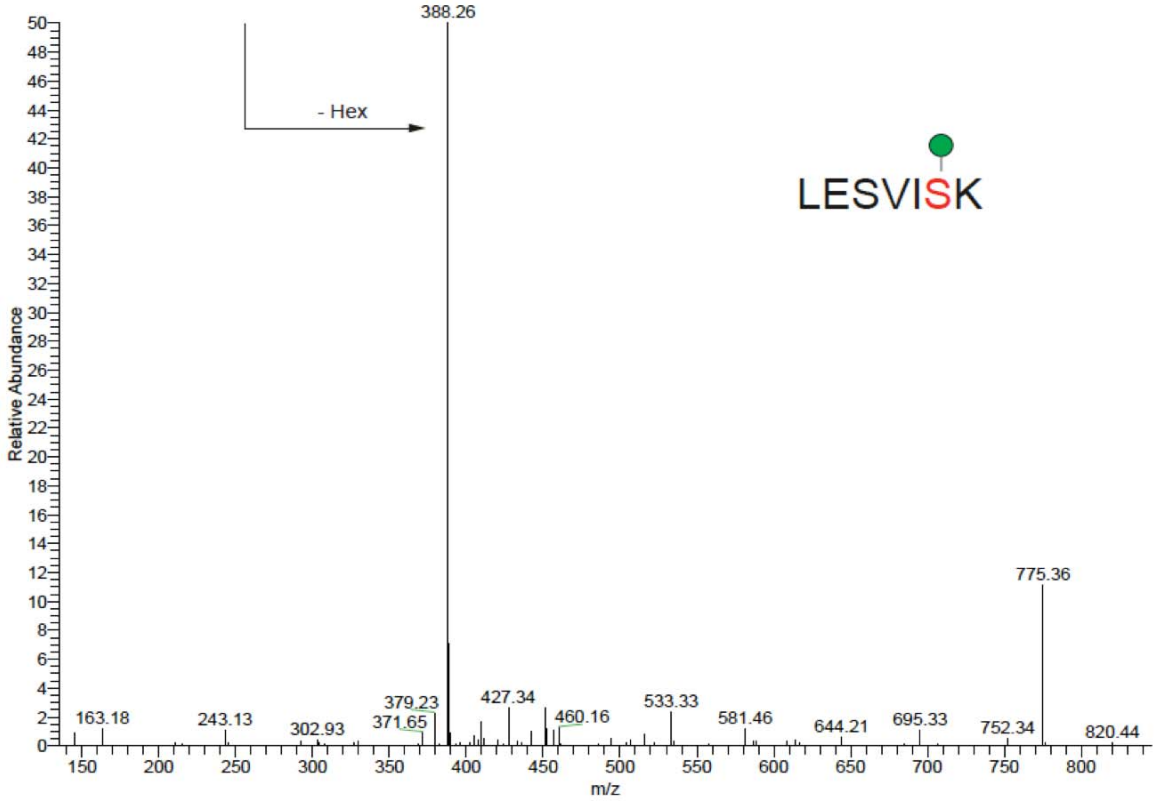


Figure B-S4. The spectrum of the MS/MS profile of an O-mannosylated peptide observed from glycoform S. This peptide was also observed in glycoform L.

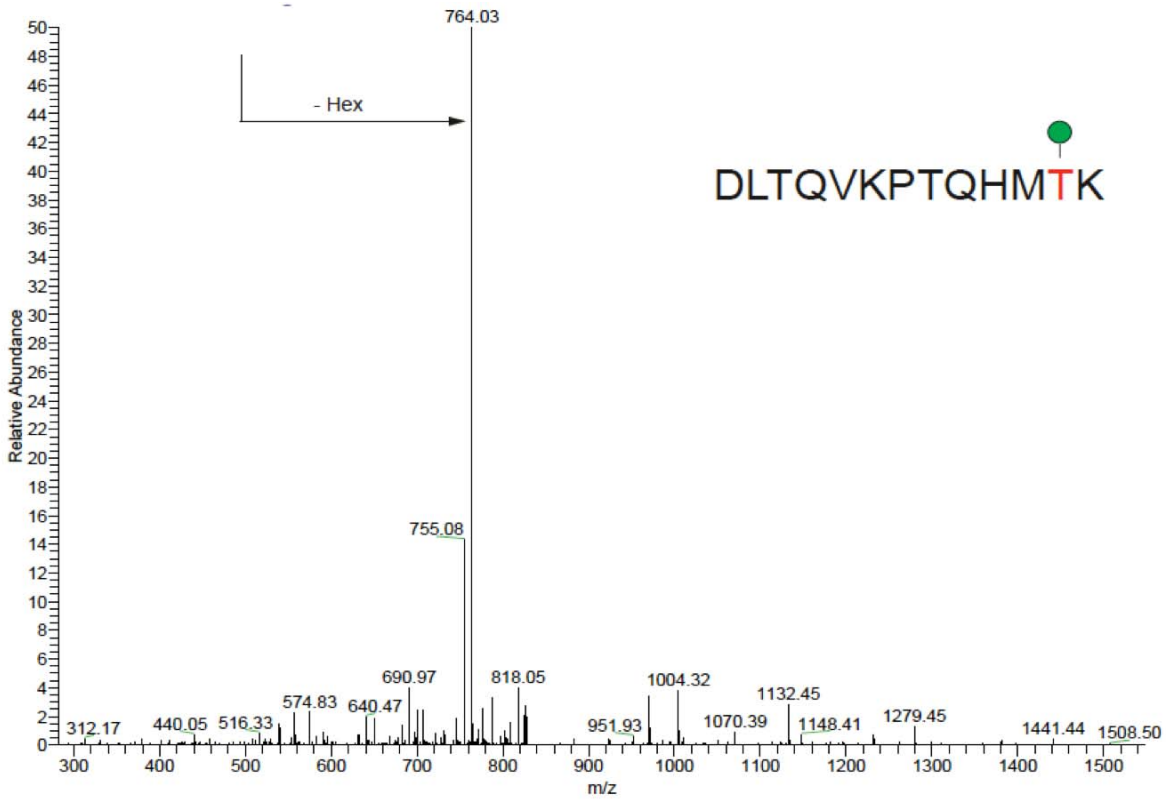


Figure B-S5. The spectrum of the MS/MS profile of an O-mannosylated peptide observed from glycoform S. This peptide was also observed in glycoform L.

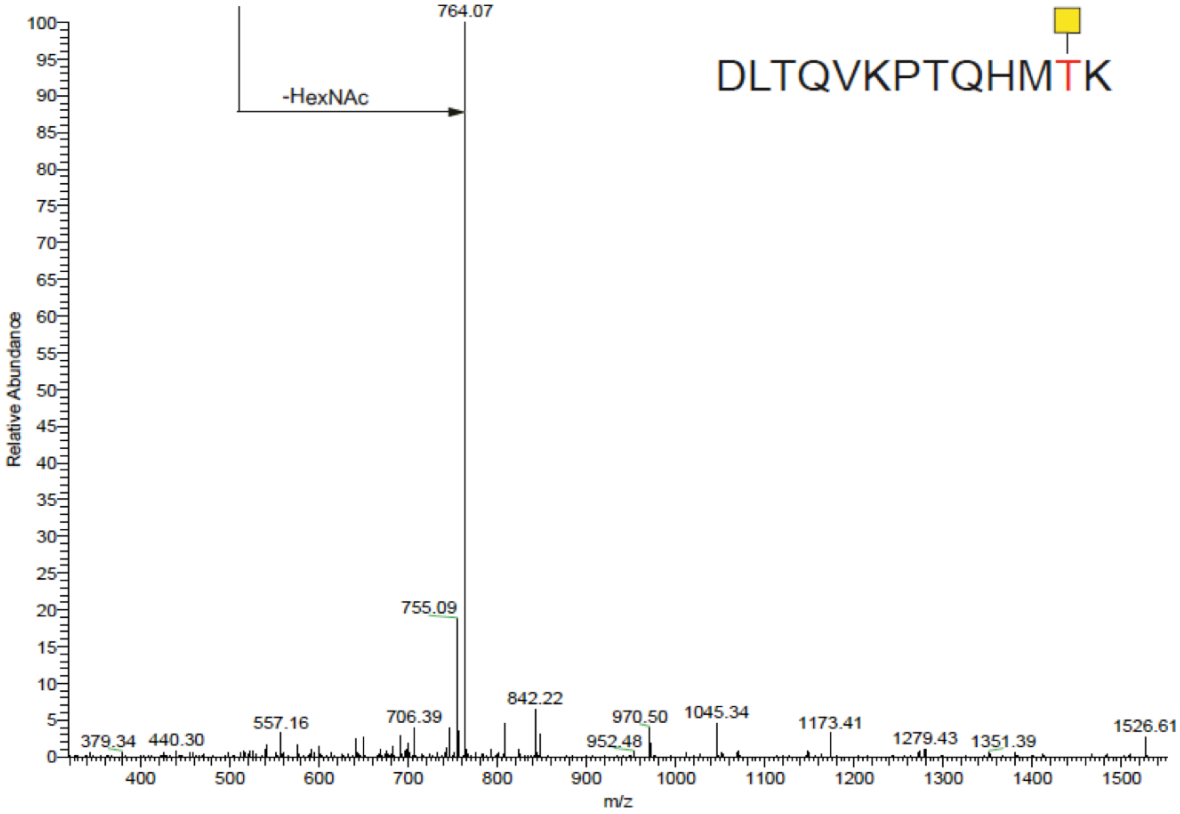


Figure B-S6. The spectrum of the MS/MS profile of an O-mannosylated peptide observed from glycoform L.

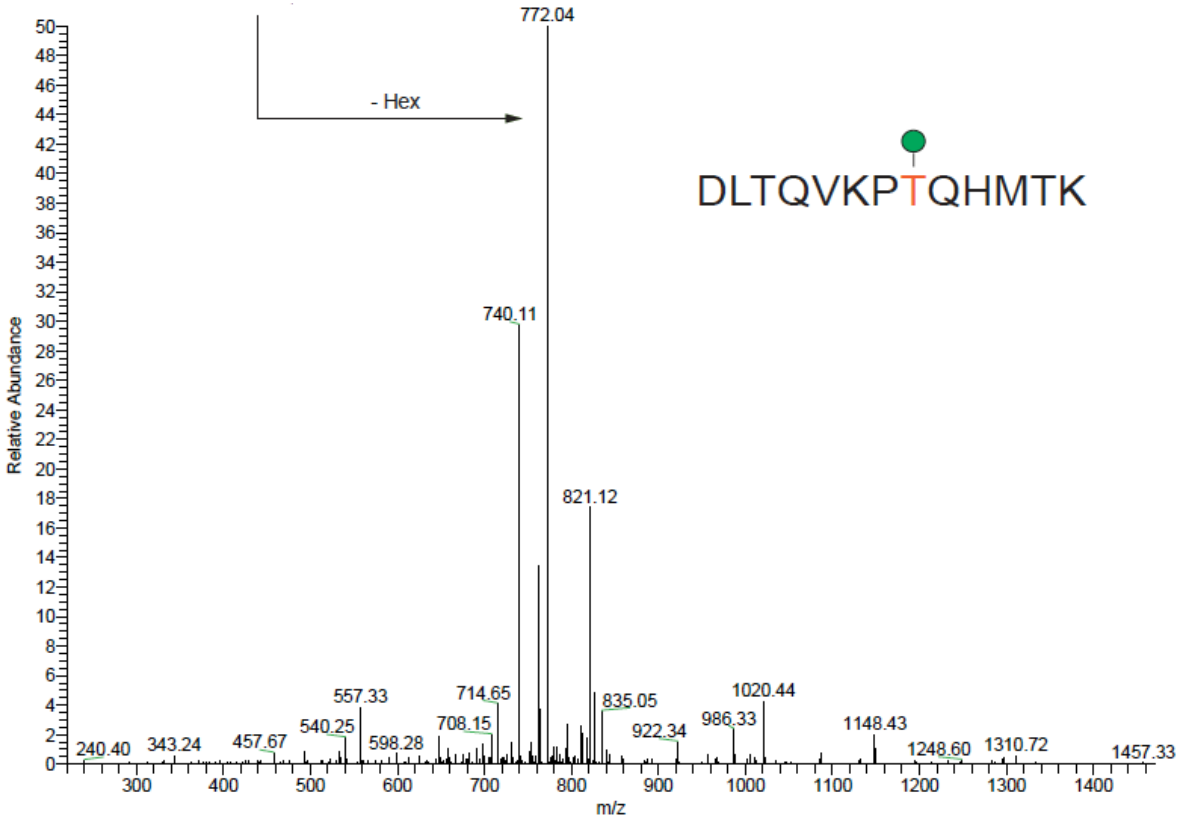


Figure B-S7. The spectrum shows the MS/MS profile of an O-mannosylated peptide observed from glycoform L.

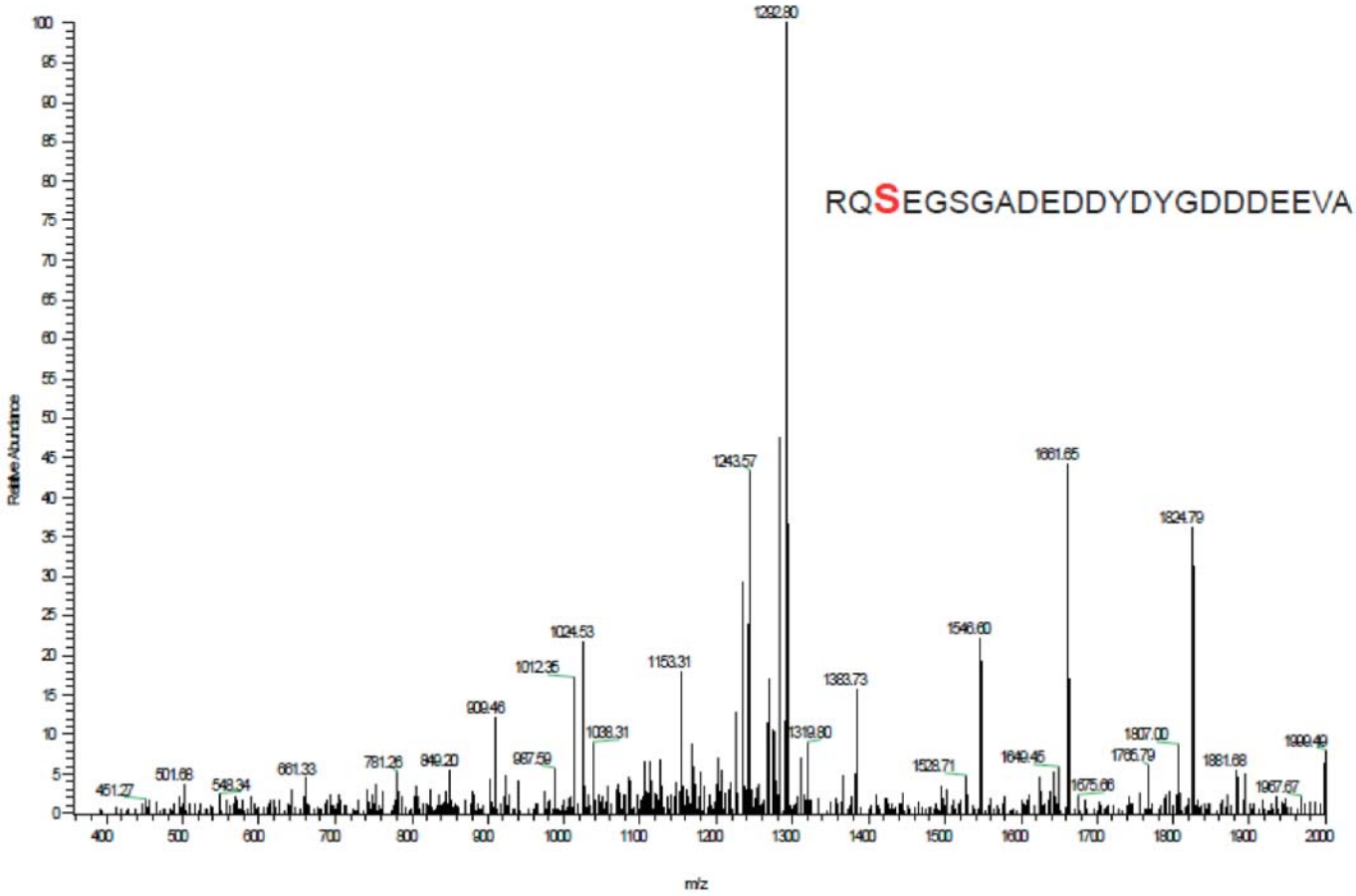


Figure BS-8. The spectrum of the MS/MS profile of a glycopeptide that was identified by BEMAD. This peptide was observed from glycoform L.