UNIVERSITY OF COPENHAGEN

Multiple distinct O-Mannosylation pathways in eukaryotes

Larsen, Ida Signe Bohse; Narimatsu, Yoshiki; Clausen, Henrik; Joshi, Hiren J; Halim, Adnan

Multiple distinct O-Mannosylation pathways in eukaryotes

Ida Signe Bohse Larsen, Yoshiki Narimatsu, Henrik Clausen, Hiren J. Joshi, Adnan Halim

Affiliation

Copenhagen Center for Glycomics, Departments of Cellular and Molecular Medicine, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark

Correspondence

Hiren J. Joshi (joshi@sund.ku.dk) or Adnan Halim (halim@sund.ku.dk)

Abstract

Protein O-mannosylation (O-Man), originally discovered in yeast five decades ago, is an important post-translational modification (PTM) conserved from bacteria to humans, but not found in plants or nematodes. Until recently, the homologous family of ER-located protein O-mannosyl transferases (PMT1-7 in yeast; POMT1/POMT2 in humans), were the only known enzymes involved in directing O-Man biosynthesis in eukaryotes. However, recent studies demonstrate the existence of multiple distinct O-Man glycosylation pathways indicating that the genetic and biosynthetic regulation of O-Man in eukaryotes is more complex than previously envisioned. Introduction of sensitive glycoproteomics strategies provided an expansion of O-Man glycoproteomes in eukaryotes (yeast and mammalian cell lines) leading to the discovery of O-Man glycosylation on important mammalian cell adhesion (cadherin superfamily) and signaling (plexin family) macromolecules, and to the discovery of unique nucleocytoplasmic O-Man glycosylation in yeast. It is now evident that eukaryotes have multiple distinct O-Man glycosylation pathways including: i) the classical PMT1-7 and POMT1/POMT2 pathway conserved in all eukaryotes apart from plants; ii) a yet uncharacterized nucleocytoplasmic pathway only found in yeast; iii) an ER-located pathway directed by the TMTC1-4 genes found in metazoans and protists and primarily dedicated to the cadherin superfamily; and iv) a yet uncharacterized pathway found in metazoans primarily dedicated to plexins. O-Man glycosylation is thus emerging as a much more widespread and evolutionary diverse post-translational modification (PTM) with complex genetic and biosynthetic regulation. While deficiencies in the POMT1/POMT2 O-Man pathway underlie muscular dystrophies, the TMTC1-4 pathway appear to be involved in distinct congenital disorders with neurodevelopmental phenotypes. Here, we review and discuss the recent discoveries of the new non-classical O-Man glycosylation pathways, their substrates, functions and roles in disease.

Highlights

- Multiple O-Man glycosylation pathways exist in eukaryotes
- Yeast has a unique nucleocytosolic O-Man glycosylation pathway in lieu of the common eukaryotic O-GlcNAc nucleocytosolic pathway
- Cadherin and plexin superfamilies of proteins are major carriers of O-Man glycans
- Congenital disorders of glycosylation are caused by deficiencies in TMTCs

Introduction

The classical pathway for O-Man glycosylation (Fig. 1) was originally discovered in yeast (1), and following cloning of the initiating protein mannosyltransferases, PMT1-PMT7, homologs of the yeast enzymes (mammalian homologues abbreviated as POMT1/POMT2) were later identified in eukaryotes (except nematodes and plants) and bacteria (2-4). These enzymes are multi-pass transmembrane proteins that utilize dolichol phosphate-β-D-mannose (Dol-P-Man) as donor substrate and catalyze attachment of alpha-linked mannose to serine and threonine residues; they have catalytic domains oriented into the endoplasmic reticulum (ER) lumen and require dimerization for function (5). In yeast, the PMTs are predicted to widely glycosylate proteins trafficking the secretory pathway (6, 7), while recent studies suggest that the homologous POMTs have more restricted functions limited to a few proteins including alpha-dystroglycan (αDG) and the densely O-mannosylated but relatively unexplored protein KIAA1549 (8). Yeast O-Man glycans are extended into poly-mannose structures, while the POMT1/POMT2 O-Man glycans in mammals are further extended into complex glycans through diverse pathways (Fig. 1), including perhaps the most elaborate and unique type of protein glycosylation leading to the important laminin binding epitope known as matriglycan (9-11).

Insight into O-mannosylation in higher eukaryotes has been long underway and our knowledge of the O-Man glycoproteome is still advancing. The first mammalian O-Man glycans were isolated from rat brain in 1979 (12, 13), and αDG was later identified as a mammalian substrate for the classical O-Man pathway (14). The connection between αDG, the *POMT1/POMT2* genes and encoded enzymes, and a subset of congenital muscular dystrophies (15, 16), sparked intense research in this type of O-glycosylation; significant efforts have been devoted to elucidating αDG O-Man structures and functions in the past two decades, unraveling the remarkable complexity and biochemistry of classical O-Man glycans (17-19). Less progress was made in understanding the breadth and depth of the eukaryotic O-Man glycoproteome, *i.e.* the protein substrates (and sites) of O-Man pathways, despite evidence demonstrating that O-Man glycans are not exclusively found on αDG in mammals (20-24), they constitute a major part of the total O-glycans in the brain (25), and they represent the only type of O-glycosylation in yeast (3, 4).

In 2013, two groups independently reported occurrence of O-Man glycans on a new class of proteins (cadherins) in metazoans. Winterhalter *et al.* identified O-Man on T-cadherin derived

from rabbit skeletal muscle, and demonstrated that the O-Man glycan identified was not extended beyond the initial monosaccharide, an unexpected finding considering that classical O-Man glycosylation in mammals are elongated and complex (26) (Fig. 1). Shortly after, the same group demonstrated the occurrence of O-Man glycans on affinity purified E-cadherin and suggested that O-Man fulfils important functions in cadherin-mediated cell-cell adhesion (27). We utilized the SimpleCell gene editing glycoproteomics strategy with a human cell line to explore the O-Man glycoproteome (28), and were able to expand the glycoproteome with over 50 distinct proteins, predominantly identifying members of the cadherin superfamily and the plexin family, including two structurally homologous receptor tyrosine kinases: hepatocyte growth factor receptor (c-MET) and macrophage stimulating 1 receptor (MST1R, also known as RON) (29). Collectively, these studies revealed two unique details that clearly distinguished the new classes of O-Man glycoproteins from the classical αDG-type O-Man glycosylation: i) these O-Man glycosylations were apparently not extended (Man-α-O-Ser/Thr); and ii) the O-Man glycosites were mapped to β-strands of two distinct protein folds – on extracellular cadherin (EC) domains of cadherins (Fig. 2) and on Ig-like, plexin and transcription factor (IPT) domains of plexins, c-MET and RON. The finding that these O-Man glycosites were not extended like the O-Man glycans found on αDG is yet unexplained. In the context of αDG, Man-α-O-Ser/Thr glycans (core M0) are considered to be biosynthetic intermediates that serve as substrates for the β2-GlcNAc transferase POMGnT1 (core M1), for the β6-GlcNAc transferase MGAT5B (core 2) and the β3-GlcNAc transferase POMGnT2 (core M3) (10), however, in the context of EC and IPT domains, Man-α-O-Ser/Thr glycans are apparently not further elongated by any of the enzymes responsible for core M1-3 biosynthesis. The POMT1/POMT2 enzymes initiate O-Man biosynthesis at unfolded regions (mucin-like domain) of αDG, and protein context appears to be important for directing core M1-3 biosynthesis, however, it remains unclear how (and why) distinct protein folds (EC and IPT domains) are targeted by single O-Man monosaccharides.

The finding that O-Man glycosites in EC and IPT domains are uniquely located in folded domains warranted further examination of the O-Man biosynthetic pathways in mammalian cells. We therefore used a knock-out (KO) strategy targeting the classical POMT1 and POMT1/POMT2 enzymes to dissect the biosynthetic regulation of O-Man in mammalian cell lines, and surprisingly we found that the classical POMT1/POMT2 enzymes are not required for O-Man biosynthesis on cadherins and plexins (8). This seeded the hypothesis that mammals have evolved multiple, distinct O-Man biosynthetic pathways targeting different classes of proteins.

Non-classical O-Man glycosylation of cadherins by TMTC1-4

Guided by bioinformatic sequence analyses, we identified a family of four homologous genes, encoding transmembrane and tetratricopeptide repeat containing protein 1-4 (TMTC1-4), as putative protein O-Man transferases. Combinatorial KO of all four genes in human HEK293 cells subsequently demonstrated that O-Man glycosylation was selectively lost on cadherins, but not on αDG and further surprisingly neither on plexins. This identified TMTC1-4 as strong enzyme candidates selectively dedicated to O-Man glycosylation of cadherins (30). Sunryd *et al.*

had previously identified TMTC1 and TMTC2 as ER resident proteins involved in Ca^{2+} regulation (31), and it is reasonable to predict that all four isoforms localize to the ER compartment. Like the POMT1/POMT2 enzymes, the TMTC1-4 proteins are multi-pass transmembrane proteins with conserved DD motifs (predicted catalytic site) in the first loop facing the ER lumen, and they are predicted to share structural homology with the POMTs and the ArnT aminoarabinose-transferase (32). Apart from also being multi-pass transmembrane proteins, the DPY19L1-4 enzymes responsible for C-mannosylation do not appear to share any structural homology with TMTCs (33-35). All four TMTC proteins are further characterized by the presence of a variable number of C-terminal tetratricopeptide repeats (TPR), which is an evolutionary conserved structural scaffold and mediator of protein-protein interactions (36). TPR domains have previously been characterized on the O-GlcNAc transferase (OGT), responsible for nucleocytoplasmic O-GlcNAcylation, and OGT TPRs are known to be important for substrate recognition and recruitment (37, 38). The functional importance of the TPR domains of TMTCs is currently unknown, but it is not unlikely that they serve similar roles, i.e. recruitment of specific substrates (potentially EC domains), which could explain the selective specificity of TMTCs for the cadherin superfamily of adhesion molecules.

A region of the TMTC proteins close to the first loop, and spanning subsequent transmembrane domains is identified as matching with the DUF1732 sequence motif. The presence of both this domain, and variable numbers of TPRs is enough to identify orthologs of this protein family in various organisms (Fig. 3). Members of the TMTC family can be found in nearly all metazoan organisms and extend back to include lower clades such as holozoa (with TMTC members detected in choanoflagellates), and sequence evidence suggests that family members can be found in other protists and in some prokaryotes. Given that this protein family predates the substrates that it glycosylates, as well as the fact that this same DUF1732 domain belongs to the ArnT superfamily domains which include homologues to STT3s and the ALGs, it is likely that there has been a shift in specificity of these enzymes, and the activity of these proteins in lower organisms still needs to be elucidated.

Cadherin proteins constitute a large family of cell-cell adhesion molecules that include the classical cadherins, desmosomal cadherins, protocadherins, and others (Fig. 2) (39). Cadherins are involved in morphogenetic processes including embryonic cell layer separation, cell signaling, and physical homeostasis of mature tissues (40). Consistent with these roles, decreased cadherin expression may result in decreased cellular adhesion, which is a common feature of metastasis (41). Protocadherins expressed on cell surfaces allow neurons to discriminate self from non-self (42), which is indispensable for neuronal wiring and development of the nervous system. Importantly, TMTC mutations/deficiencies associate or directly cause neurological disorders (Fig. 1 and 2C), including schizophrenia (TMTC1), hearing loss (TMTC2 and TMTC4), intellectual impairment (TMTC3), epilepsy (TMTC3) and brain malformation/Cobblestone lissencephaly (TMTC3) (43-48). The etiology of this broad class of disorders remains largely unexplained, however, the identification of TMTCs as O-Man enzymes

specifically targeting cadherin molecules opens for new hypotheses and suggest that the non-classical O-Man glycans on these proteins may serve roles in cell-cell interaction networks.

Cadherins are characterized by the presence of a varying number of EC domains that confer specific interactions required for function (Fig. 2A) (49). The cell-cell binding process is mediated by cooperative formation of distinct EC-EC interactions in *cis* (on the same cell) and/or in *trans* (between different cells) (Fig. 2A). The EC domain has a characteristic immunoglobulin-like (Ig-like) fold consisting of a two-layer β-sheet sandwich with antiparallel β-strands (Fig. 2B). The O-Man glycans are located on evolutionary conserved amino acids on either B- or G-strands (Fig. 2B), but it is currently unclear if they are involved in EC-EC interactions, cadherin folding, secretion and/or fine-tuning of functions. We recently demonstrated that individual TMTC enzymes have specificities for B- or G-strands located on opposite sides of the EC fold (Fig. 2B), suggesting that cadherin function may be tuned by differential O-Man glycosylation on specific EC domains and β-strands. The correlation between different neurological phenotypes and deficiencies in distinct TMTC isoenzymes supports this hypothesis (Fig. 2C), and further indicates that O-Man glycans on distinct EC domains/β-strands are tightly controlled to orchestrate specific functions. Delineating how individual TMTC enzymes operate and direct O-Man glycans on the EC domains may advance the understanding of cellular processes involving cadherin/protocadherin mediated cell adhesion and function. Particular interest may relate to the role of E-cadherin which plays important roles in the malignant phenotypes of cancer. Previous studies have suggested the involvement of POMT-driven O-Man glycosylation in the tumor suppression functions of E-cadherin (50). However, considering the later discovery of TMTCs and their dedicated roles in O-Man glycosylation of cadherins (30), these findings may not be correct and remain unexplained. Nevertheless, it is likely that both N- and O-glycosylation may play roles in E-cadherin function in cancer (51).

Non-classical O-Man glycosylation of IPT domains

The original O-Man glycoproteome study also identified conserved glycosites in IPT domains of plexins, c-MET and RON receptors (29), and surprisingly, we found that the biosynthesis of the O-Man glycans at these sites were independent of both the POMT1/POMT2 and TMTC1-4 glycosylation pathways (8, 30). The biosynthetic enzymes responsible for IPT domain O-Man glycosylation is currently unknown (Fig. 1), although we have identified a candidate gene that is under investigation. Plexins are single-pass transmembrane proteins with large ectodomains composed of a Sema domain, 2-3 plexin-semaphorin (PSI) domains and 3-4 IPT domains (52). Plexins act as the main receptors for a class of signaling factors known as semaphorins, and this signaling circuit is widely recognized as a key regulator of multiple cellular processes in renal, cardiovascular, bone and nervous systems (53); aberrant plexin signaling is associated with multiple disorders, including cancers (54). RON and c-MET tyrosine kinase receptors share sequence- and structural homology with the plexin family of receptors, and incorporate Sema, PSI and IPT domains (55). Both RON and c-MET are essential for embryonic development and organogenesis, but these receptors are also known for their prominent roles in cancer, metastasis

and tumor progression (56). The plexin signaling cascade is initiated by ligand binding to the extracellular Sema domain, leading to plexin dimerization and activation, and although the precise activation mechanism of plexins remains unclear, recent evidence indicates that IPT domains play an important role in this process (57). Furthermore, the structurally homologous relative c-MET is known to harbor the high-affinity binding site for its ligand, hepatocyte growth factor (HGF), at the extracellular IPT domains (58). The biological roles of O-Man on IPT domains remain unclear, but identification of the enzymes responsible for IPT domain O-Man glycosylation may now enable functional studies and potentially identify new congenital disorders linked to dysfunctional O-Man glycosylation on this important class of receptors.

Nucleocytoplasmic O-Man glycosylation in yeast

Another surprising finding derived from our O-Man glycoproteomic analyses was the finding of O-Man glycans widely distributed on nuclear and cytosolic protein in yeast (59). Nucleocytoplasmic O-glycosylation (O-GlcNAcylation) is highly conserved among eukaryotes and is also present in bacteria, however, some protists (*e.g. Leishmania* and *Trypanosoma*) and yeast (*S. cerevisiae or S. pombe*) lack the OGT enzyme responsible for O-GlcNAcylation (60). Yeast and humans share a significant fraction of their functional pathways that control key aspects of eukaryotic cell biology including Ser/Thr phosphorylation to regulate cell signaling (61, 62), however, the conceived lack of a nucleocytoplasmic O-glycosylation mechanism for co-regulation of phosphorylation in yeast remained enigmatic considering the important functional roles of the O-GlcNAc modification in mammals (37, 38). This conundrum may be explained by our recent mass spectrometric analyses of the yeast O-Man glycoproteome that increased the coverage from ~50 to >500 known O-Man glycoproteins (6, 59), and demonstrated that >30% of the identified O-Man glycoproteins were found to originate from mitochondrial and nucleocytoplasmic compartments (59). Specifically, many O-Man glycosites were found on similar proteins and in similar or in some cases identical orthologous positions as O-GlcNAc in mammals, and in several cases the identified yeast O-Man glycosites were known to also be phosphorylated. Given that the classical yeast PMT1-7 have catalytic domains oriented into the ER lumen, these findings suggested that yeast have an unknown glycosylation mechanism for directing O-Man glycosylation on nucleocytoplasmic proteins (Fig. 1). It is interesting to note that ectopic expression of OGT in yeast is toxic and leads to growth defects, which may be alleviated by co-expression of the O-GlcNAc hydrolase (OGA) (63). However, the functions of nucleocytoplasmic O-Man in yeast remain unknown and the enzyme(s) responsible for this type of nucleocytoplasmic O-Man glycosylation have not been identified.

Concluding remarks

Developments in glycoproteomics strategies have provided major progress in our understanding of protein O-Man glycosylation pathways in eukaryotes. It is now evident that O-Man is a widespread PTM controlled by multiple distinct enzymatic pathways. Yeast has two O-Man biosynthetic pathways with the nucleocytoplasmic type likely serving as the substitute for the important nucleocytoplasmic O-GlcNAcylation process found in higher eukaryotes. In

mammalian cells at least two new types of O-Man glycosylation pathways are seemingly dedicated to specific protein folds and functional domains on major classes of adhesion molecules and receptors, leaving a total of three distinct O-Man glycosylation pathways in higher eukaryotic cells serving different classes of proteins. Future studies are needed to explore the different biosynthetic pathways and how the elongation of these O-Man glycans differ as well as the biological roles of these. With respect to POMT1/POMT2 driven O-Man glycosylations, we predict that further studies on the relatively unexplored but densely O-mannosylated protein KIAA1549 will shed new light on the structure-function relationship of O-Man (core M1-M3) glycans. The TMTC genes (TMTC3) already contribute to the growing list of congenital disorders of glycosylation, and further studies are likely to identify additional genes as the underlying cause of previously inexplainable neurological disorders. The lack of knowledge in this area is thus substantial. The recent advances were enabled by the combinatorial use of sensitive lectin enrichment strategies, advanced mass spectrometry-based glycoproteomics and precise genetic engineering, enabling dissection and discovery of novel O-Man biosynthetic pathways and expansion of O-Man glycoproteomes. One important lesson learned from these advances is that we are still many steps away from fully understanding the complexity of glycosylation capacities in eukaryotic cells, and that many new discoveries lie ahead.

Acknowledgments

Figure legends

Figure 1

Graphical depiction of O-Man glycosylation pathways in eukaryotes. Classical O-Man glycosylation is driven by the PMT1-7 (yeast) and POMT1/POMT2 (metazoan) enzymes, which catalyze O-Man transfer from Dol-P-Man to Ser/Thr residues in the ER lumen; classical O-Man glycans are further elongation into poly-mannose (yeast) or complex glycans (mammals) in the Golgi apparatus (3, 10). Non-classical O-Man glycosylation: nucleocytoplasmic proteins (yeast) and major classes of plasma membrane adhesion molecules/receptors (metazoans) are modified by O-Man. The enzyme(s) responsible for nucleocytoplasmic glycosylation in yeast are unknown; TMTC1-4 enzymes specifically target the cadherin superfamily; Plexins, c-MET and RON are O-Man glycosylated on their IPT domains by unknown enzyme(s). Non-classical O-Man glycans in mammals have only been observed as single monosaccharide modifications. Diseases linked to deficiencies in POMT1/2 and TMTC1-4 enzyme families are highlighted in red boxes.

Figure 2

Illustration of structure and O-Man glycosylation of extracellular cadherin (EC) domains. (A) Cadherin domain organization and modes of action; EC1-EC3 of classical cadherins mediate cell-cell adhesion via *cis*- and *trans*-interactions (left). Protocadherin homodimers (cis-interacting EC6) engage EC1-EC4 in head-to-tail *trans*-interactions (right). (B) β-strand topology and organization of EC domains (mouse E-cadherin, Protein Data Bank accession 3Q2V); O-Man is located on opposite sides (B- and G-strands) of EC domains. (C) Hypothetical model for TMTC1-4 specificity based on recent studies (8, 30); disorders linked to deficiencies in individual TMTCs are indicated in red boxes.

Figure 3

Phylogenetic analysis identifies members of the TMTC family in all major classes of Eukaryotes. Proteomes from protists, choanoflagellates, invertebrates and vertebrates were scanned for the DUF1736 domain, that is found in all TMTC proteins. The scans revealed that TMTCs are well conserved down to Holozoa, and are also found in a number of protist species, albeit with varying numbers of TPRs.

References

1. Sentandreu R & Northcote DH (1968) The structure of a glycopeptide isolated from the yeast cell wall. *Biochem J* 109(3):419-432.
2. Strahl-Bolsinger S, Immervoll T, Deutzmann R, & Tanner W (1993) PMT1, the gene for a key enzyme of protein O-glycosylation in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* 90(17):8164-8168.
*3. Loibl M & Strahl S (2013) Protein O-mannosylation: what we have learned from baker's yeast. *Biochim Biophys Acta* 1833(11):2438-2446.
Comprehensive review focused on classical O-mannosylation (PMT pathway) and biological roles of O-Man glycans in yeast.
4. Lommel M & Strahl S (2009) Protein O-mannosylation: conserved from bacteria to humans. *Glycobiology* 19(8):816-828.
5. Manya H*, et al.* (2004) Demonstration of mammalian protein O-mannosyltransferase activity: coexpression of POMT1 and POMT2 required for enzymatic activity. *Proc Natl Acad Sci U S A* 101(2):500-505.
6. Neubert P*, et al.* (2016) Mapping the O-Mannose Glycoproteome in Saccharomyces cerevisiae. *Mol Cell Proteomics* 15(4):1323-1337.
**7. Neubert P & Strahl S (2016) Protein O-mannosylation in the early secretory pathway. *Curr Opin Cell Biol* 41:100-108.
A thorough review discussing primarily PMT- but also POMT-driven O-mannosylation in the endoplasmic reticulum. Special focus is placed on O-Man biosynthesis at the translocon complex in relation to ER protein homeostasis and the unfolded protein response.
8. Larsen ISB*, et al.* (2017) Mammalian O-mannosylation of cadherins and plexins is independent of protein O-mannosyltransferases 1 and 2. *J Biol Chem* 292(27):11586-11598.
9. Praissman JL & Wells L (2014) Mammalian O-mannosylation pathway: glycan structures, enzymes, and protein substrates. *Biochemistry* 53(19):3066-3078.
**10. Sheikh MO, Halmo SM, & Wells L (2017) Recent advancements in understanding mammalian O-mannosylation. *Glycobiology* 27(9):806-819.
An excellent review describing POMT-driven O-Man biosynthesis, enzymatic pathways and structures (core M0-M3) in mammalian systems.
11. Yoshida-Moriguchi T & Campbell KP (2015) Matriglycan: a novel polysaccharide that links dystroglycan to the basement membrane. *Glycobiology* 25(7):702-713.
12. Finne J, Krusius T, Margolis RK, & Margolis RU (1979) Novel mannitol-containing oligosaccharides obtained by mild alkaline borohydride treatment of a chondroitin sulfate proteoglycan from brain. *J Biol Chem* 254(20):10295-10300.
13. Yuen CT*, et al.* (1997) Brain contains HNK-1 immunoreactive O-glycans of the sulfoglucuronyl lactosamine series that terminate in 2-linked or 2,6-linked hexose (mannose). *J Biol Chem* 272(14):8924-8931.
14. Chiba A*, et al.* (1997) 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* 272(4):2156-2162.

15.   Barresi R & Campbell KP (2006) Dystroglycan: from biosynthesis to pathogenesis of human disease. *J Cell Sci* 119(Pt 2):199-207.

16.   Jae LT, et al.* (2013) Deciphering the glycosylome of dystroglycanopathies using haploid screens for lassa virus entry. *Science* 340(6131):479-483.

The authors describe an innovative strategy utilizing Lassa virus for dissection of genes and enzymes responsible for alpha-dystroglycan (O-Man) glycosylations.

**17.   Yoshida-Moriguchi T*, et al.* (2010) O-mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. *Science* 327(5961):88-92.

The authors reveal that phosphorylation of O-Man is required for the formation of the laminin binding epitope on alpha-dystroglycan.

**18.   Kanagawa M*, et al.* (2016) Identification of a Post-translational Modification with Ribitol-Phosphate and Its Defect in Muscular Dystrophy. *Cell Rep* 14(9):2209-2223.

The authors identify ribitol-5-phosphate as a tandem repeat linker connecting core M3 O-Man with the laminin binding moiety on alpha-dystroglycan.

**19.   Praissman JL*, et al.* (2016) The functional O-mannose glycan on alpha-dystroglycan contains a phospho-ribitol primed for matriglycan addition. *Elife* 5.

The authors demonstrate independently that ribitol-5-phosphate is a novel structure linking core M3 glycans with the laminin binding polysaccharide on alpha-dystroglycan.

20.   Martinez T, Pace D, Brady L, Gerhart M, & Balland A (2007) Characterization of a novel modification on IgG2 light chain. Evidence for the presence of O-linked mannosylation. *J Chromatogr A* 1156(1-2):183-187.

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

22.   Bleckmann C*, et al.* (2009) O-glycosylation pattern of CD24 from mouse brain. *Biol Chem* 390(7):627-645.

23.   Pacharra S, Hanisch FG, & Breloy I (2012) Neurofascin 186 is O-mannosylated within and outside of the mucin domain. *J Proteome Res* 11(8):3955-3964.

24.   Pacharra S*, et al.* (2013) The lecticans of mammalian brain perineural net are O-mannosylated. *J Proteome Res* 12(4):1764-1771.

25.   Chai W*, et al.* (1999) High prevalence of 2-mono- and 2,6-di-substituted manol-terminating sequences among O-glycans released from brain glycopeptides by reductive alkaline hydrolysis. *Eur J Biochem* 263(3):879-888.

26.   Winterhalter PR, Lommel M, Ruppert T, & Strahl S (2013) O-glycosylation of the non-canonical T-cadherin from rabbit skeletal muscle by single mannose residues. *FEBS Lett* 587(22):3715-3721.

**27.   Lommel M*, et al.* (2013) Protein O-mannosylation is crucial for E-cadherin-mediated cell adhesion. *Proc Natl Acad Sci U S A* 110(52):21024-21029.

The authors demonstrate that E-cadherin is O-Man glycosylated and further suggest that these modifications are important for cadherin-mediated cell-cell adhesion.

28.   Steentoft C*, et al.* (2014) Precision genome editing: a small revolution for glycobiology. *Glycobiology* 24(8):663-680.

**29. Vester-Christensen MB*, et al.* (2013) Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. *Proc Natl Acad Sci U S A* 110(52):21018-21023.
The authors utilize genetic engineering and advanced mass spectrometry to identify cadherins and plexins as major classes of mammalian cell-surface proteins undergoing O-Man glycosylations.

**30. Larsen ISB*, et al.* (2017) Discovery of an O-mannosylation pathway selectively serving cadherins and protocadherins. *Proc Natl Acad Sci U S A* 114(42):11163-11168.
The authors identify TMTC1-4 as enzymes responsible for O-Man glycosylation on cadherins and protocadherins.

31. Sunryd JC*, et al.* (2014) TMTC1 and TMTC2 are novel endoplasmic reticulum tetratricopeptide repeat-containing adapter proteins involved in calcium homeostasis. *J Biol Chem* 289(23):16085-16099.

32. Petrou VI*, et al.* (2016) Structures of aminoarabinose transferase ArnT suggest a molecular basis for lipid A glycosylation. *Science* 351(6273):608-612.

33. Buettner FF, Ashikov A, Tiemann B, Lehle L, & Bakker H (2013) C. elegans DPY-19 is a C-mannosyltransferase glycosylating thrombospondin repeats. *Mol Cell* 50(2):295-302.

34. Niwa Y, Suzuki T, Dohmae N, & Simizu S (2016) Identification of DPY19L3 as the C-mannosyltransferase of R-spondin1 in human cells. *Mol Biol Cell* 27(5):744-756.

35. Shcherbakova A, Tiemann B, Buettner FF, & Bakker H (2017) Distinct C-mannosylation of netrin receptor thrombospondin type 1 repeats by mammalian DPY19L1 and DPY19L3. *Proc Natl Acad Sci U S A* 114(10):2574-2579.

36. Zeytuni N & Zarivach R (2012) Structural and functional discussion of the tetra-trico-peptide repeat, a protein interaction module. *Structure* 20(3):397-405.

37. Bond MR & Hanover JA (2015) A little sugar goes a long way: the cell biology of O-GlcNAc. *J Cell Biol* 208(7):869-880.

38. Hart GW, Slawson C, Ramirez-Correa G, & Lagerlof O (2011) Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu Rev Biochem* 80:825-858.

39. Nollet F, Kools P, & van Roy F (2000) Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. *J Mol Biol* 299(3):551-572.

40. Halbleib JM & Nelson WJ (2006) Cadherins in development: cell adhesion, sorting, and tissue morphogenesis. *Genes Dev* 20(23):3199-3214.

41. Paredes J*, et al.* (2012) Epithelial E- and P-cadherins: role and clinical significance in cancer. *Biochim Biophys Acta* 1826(2):297-311.

42. Rubinstein R*, et al.* (2015) Molecular logic of neuronal self-recognition through protocadherin domain interactions. *Cell* 163(3):629-642.

43. Farhan SMK*, et al.* (2017) Identification of a novel synaptic protein, TMTC3, involved in periventricular nodular heterotopia with intellectual disability and epilepsy. *Hum Mol Genet* 26(21):4278-4289.

44. Guillen-Ahlers H*, et al.* (2018) TMTC2 variant associated with sensorineural hearing loss and auditory neuropathy spectrum disorder in a family dyad. *Mol Genet Genomic Med*.

****45.   Jerber J*, et al.* (2016) Biallelic Mutations in TMTC3, Encoding a Transmembrane and TPR-Containing Protein, Lead to Cobblestone Lissencephaly. *Am J Hum Genet* 99(5):1181-1189.
The authors establish a link between severe brain malformations and deficiencies in the TMTC3 gene.

46.   Li J*, et al.* (2018) Deletion of Tmtc4 activates the unfolded protein response causing postnatalhearing loss. *J Clin Invest*.

47.   Runge CL*, et al.* (2016) Association of TMTC2 With Human Nonsyndromic Sensorineural Hearing Loss. *JAMA Otolaryngol Head Neck Surg* 142(9):866-872.

48.   Schizophrenia Working Group of the Psychiatric Genomics C (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511(7510):421-427.

49.   Brasch J, Harrison OJ, Honig B, & Shapiro L (2012) Thinking outside the cell: how cadherins drive adhesion. *Trends Cell Biol* 22(6):299-310.

50.   Carvalho S*, et al.* (2016) O-mannosylation and N-glycosylation: two coordinated mechanisms regulating the tumour suppressor functions of E-cadherin in cancer. *Oncotarget* 7(40):65231-65246.

****51.   Pinho SS & Reis CA (2015) Glycosylation in cancer: mechanisms and clinical implications. *Nat Rev Cancer* 15(9):540-555.
A comprehensive review focused on the (dys)functions of glycans in cancer.

52.   Roney K, Holl E, & Ting J (2013) Immune plexins and semaphorins: old proteins, new immune functions. *Protein Cell* 4(1):17-26.

***53.   Worzfeld T & Offermanns S (2014) Semaphorins and plexins as therapeutic targets. *Nat Rev Drug Discov* 13(8):603-621.
A detailed review focused on the plexin family of receptors, their biological functions and their roles in disease.

54.   Angelopoulou E & Piperi C (2018) Emerging role of plexins signaling in glioma progression and therapy. *Cancer Lett* 414:81-87.

55.   Accornero P, Pavone LM, & Baratta M (2010) The scatter factor signaling pathways as therapeutic associated target in cancer treatment. *Curr Med Chem* 17(25):2699-2712.

56.   Chang K, Karnad A, Zhao S, & Freeman JW (2015) Roles of c-Met and RON kinases in tumor progression and their potential as therapeutic targets. *Oncotarget* 6(6):3507-3518.

57.   Kong Y*, et al.* (2016) Structural Basis for Plexin Activation and Regulation. *Neuron* 91(3):548-560.

58.   Basilico C, Arnesano A, Galluzzo M, Comoglio PM, & Michieli P (2008) A high affinity hepatocyte growth factor-binding site in the immunoglobulin-like region of Met. *J Biol Chem* 283(30):21267-21277.

***59.   Halim A*, et al.* (2015) Discovery of a nucleocytoplasmic O-mannose glycoproteome in yeast. *Proc Natl Acad Sci U S A* 112(51):15648-15653.
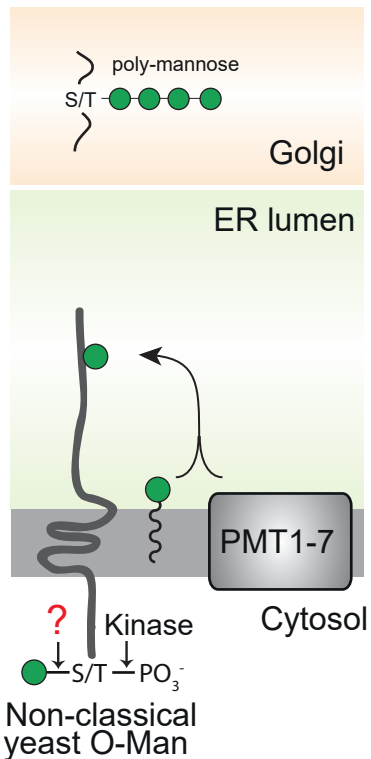The authors demonstrate that yeast (S. cerevisiae and S. pombe) have an O-Man glycosylation pathway targeting nuclear, cytosolic and mitrochondrial proteins.

60.   Banerjee S, Robbins PW, & Samuelson J (2009) Molecular characterization of nucleocytosolic O-GlcNAc transferases of Giardia lamblia and Cryptosporidium parvum. *Glycobiology* 19(4):331-336.

61.   Petranovic D, Tyo K, Vemuri GN, & Nielsen J (2010) Prospects of yeast systems biology for human health: integrating lipid, protein and energy metabolism. *FEMS Yeast Res* 10(8):1046-1059.

62.   Vlastaridis P*, et al.* (2017) The Pivotal Role of Protein Phosphorylation in the Control of Yeast Central Metabolism. *G3 (Bethesda)* 7(4):1239-1249.

63.   Nakanishi H, Li F, Han B, Arai S, & Gao XD (2017) Yeast cells as an assay system for in vivo O-GlcNAc modification. *Biochim Biophys Acta Gen Subj* 1861(5 Pt A):1159-1167.

Figure 1
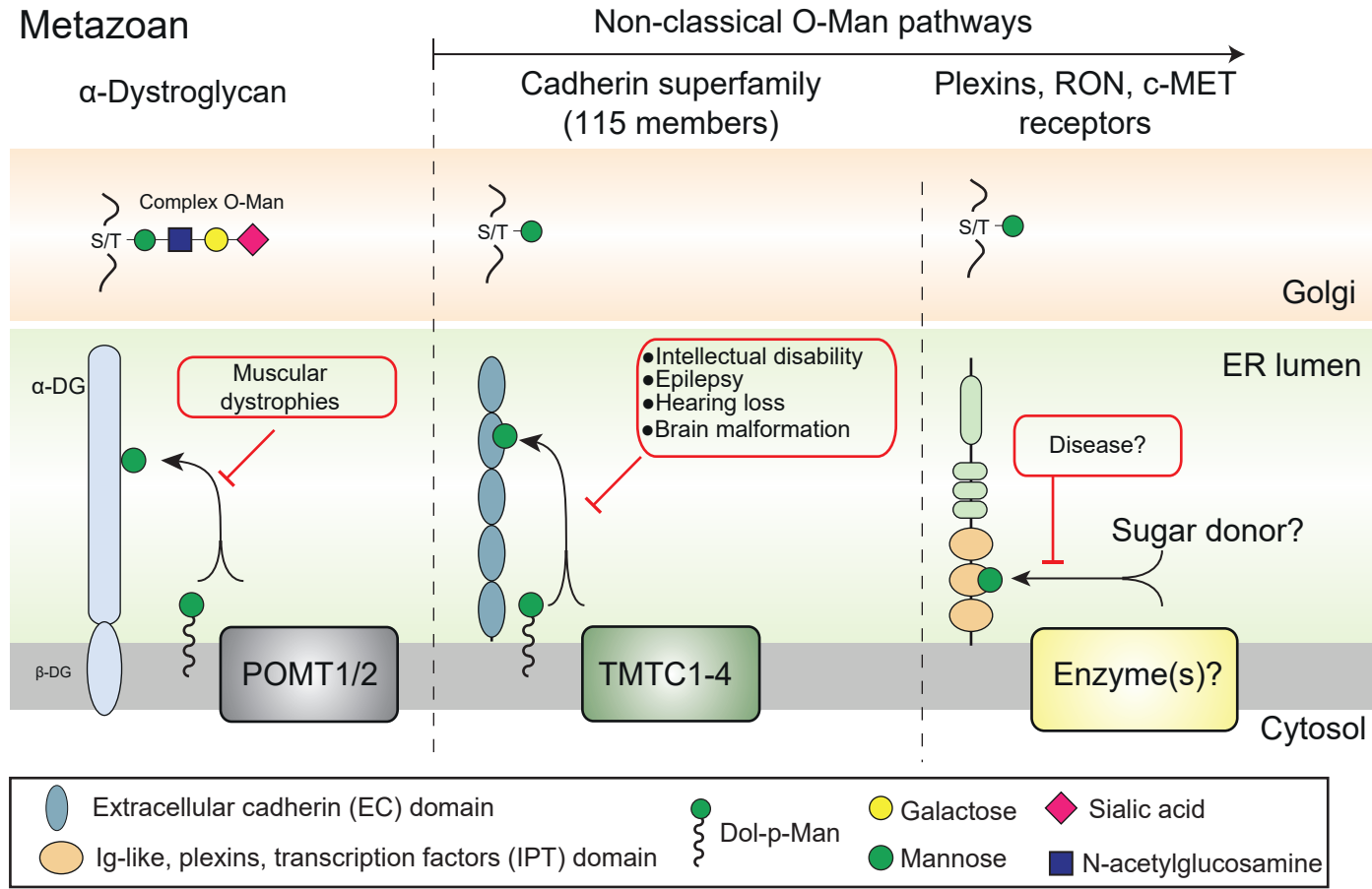
Figure 2

A

Modes of interaction

Classical cadherins

EC5
EC4
EC3
EC2
EC1

Protocadherins

EC6
EC5
EC4
EC3
EC2
EC1

Trans interacting EC domain   *Cis* interacting EC domain

B

EC domain
2-layer β-sandwich

Side view

Top view

90°

O-Man

B
E
A
D
G
F
C

B- and G-strand
O-Man glycosylation

D E **B** A*   A **G** F C
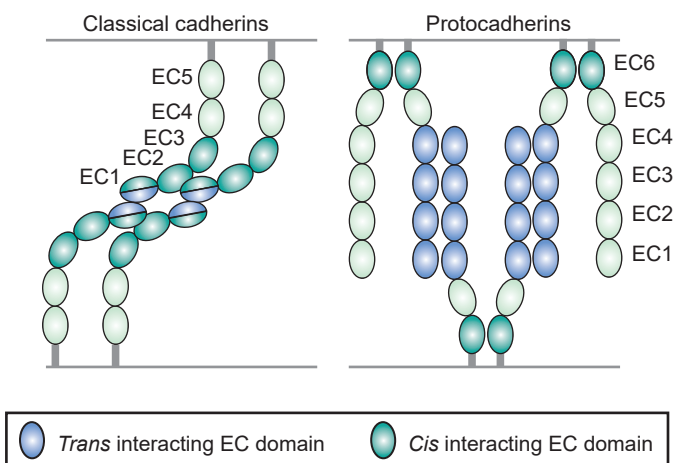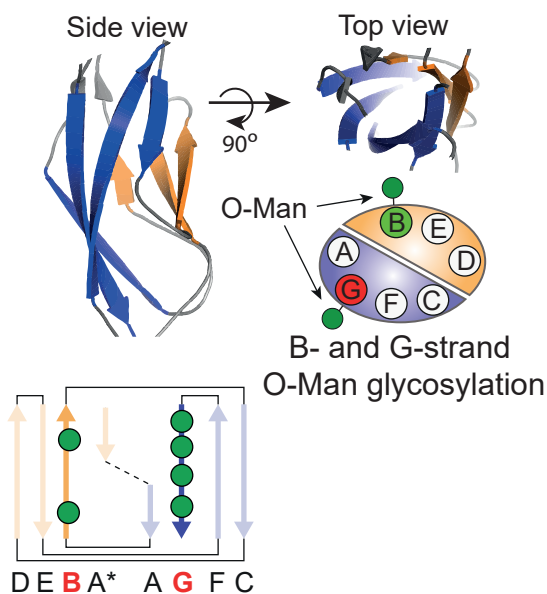
C

TMTC specificity
(hypothetical model)

Disorders connected
to TMTCs

TMTC2   TMTC4

•Hearing loss

B
E
A
D
G
F
C

•Intellectual
impairment
•Epilepsy
•Brain
malformation

TMTC1

TMTC3

•Schizophrenia

Figure 3

Protists

Vertebrates

Invertebrates

Chanoflagellates

Oomycetes

Holozoa

Leishmania

Kinetoplastida

Saprolegnia asitica CBS 223.65
Saprolegnia diclina VS20
Aphanomyces astaci
Aphanomyces invadans
Mus musculus
Homo sapiens
Danio rerio
Drosophila melanogaster
Caenorhabditis elegans
Salpingoeca rosetta
Monosiga brevicollis
Chromera velia CCMP2878
Vitrella brasicaformis CCMP3155
Acanthamoeba castellanii str. Neff

Phytophthora infestans T30
Phytophthora sojae
Phytophthora parasitica INRA
Phytopythium vexans DAOM BR484
Pythium aphanidermatum DAOM BR444
Pythium arrhenomanes ATCC 12531
Pythium irregulare DAOM BR486
Pythium iwayami DAOM BR242034
Pythium ultimum DAOM BR144
Pythium ultimum var. sporangium BR650
Blechomonas ayalai
Trypanosoma theileri
Trypanosoma grayi
Crithidia fasciculata
Endotrypanum monterogeii
Leptomonas seymouri
Leptomonas pyrrhocoris
Leishmania tarentolae
Leishmania panamensis MHOM/COL/81/L13
Leishmania braziliensis MHOM/BR/75/M2904
Leishmania braziliensis MHOM/BR/75/M2903
Leishmania sp. MAR LEM2494
Leishmania turanica
Leishmania arabica
Leishmania gerbilli
Leishmania aethiopica L 147
Leishmania tropica L590
Leishmania enrietti
Leishmania mexicana MHOM/GT/2001/U1103
Leishmania infantum JPCM5
Leishmania donovani BPK282A1
Leishmania major strain LV39c5
Leishmania major strain SD 75.1
Leishmania major strain Friedlin