The glucuronyltransferase B4GAT1 is required for initiation of LARGE-mediated α-dystroglycan functional glycosylation

Tobias Willer[a], Kei-ichiro Inamori[a,2], David Venzke[a], Corinne Harvey[a], Greg Morgensen[a], Yuji Hara[a,3], Daniel Beltrán Valero de Bernabé[a], Liping Yu[b], Kevin M. Wright[c], Kevin P. Campbell[a,1]

[a]Howard Hughes Medical Institute, Department of Molecular Physiology and Biophysics, Department of Neurology, Department of Internal Medicine, University of Iowa Roy J. and Lucille A. Carver College of Medicine, 4283 Carver Biomedical Research Building, 285 Newton Road, Iowa City, IA 52242, USA

[b]Medical Nuclear Magnetic Resonance Facility, University of Iowa Roy J. and Lucille A. Carver College of Medicine, B291 Carver Biomedical Research Building, 285 Newton Road, Iowa City, IA 52242, USA

[c]Vollum Institute, 3181 S.W. Sam Jackson Park Rd., Portland, OR 97239, USA

[1]To whom correspondence should be addressed. E-mail: kevin-campbell@uiowa.edu

[2]Present address: Division of Glycopathology, Institute of Molecular Biomembrane and Glycobiology, Tohoku Pharmaceutical University, 4-4-1 Komatsushima, Aobaku, Sendai, Miyagi 981-8558, Japan

[3]Present address: Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Katsura, Nishikyo-ku, Kyoto 615-8510, Japan

Abstract

Dystroglycan is a cell membrane receptor that organizes the basement membrane by binding ligands in the extracellular matrix. Proper glycosylation of the α-dystroglycan (α-DG) subunit is essential for these activities, and lack thereof results in neuromuscular disease. Currently, neither the glycan synthesis pathway nor the roles of many known or putative glycosyltransferases that are essential for this process are well understood. Here we show that FKRP, FKTN, TMEM5 and B4GAT1 (formerly known as B3GNT1) localize to the Golgi and contribute to the O-mannosyl post-phosphorylation modification of α-DG. Moreover, we assigned B4GAT1 a function as a xylose β1,4-glucuronyltransferase. Nuclear magnetic resonance studies confirmed that a glucuronic acid β1,4-xylose disaccharide synthesized by B4GAT1 acts as an acceptor primer that can be elongated by LARGE with the ligand-binding heteropolysaccharide. Our findings greatly broaden the understanding of α-DG glycosylation and provide mechanistic insight into why mutations in B4GAT1 disrupt dystroglycan function and cause disease.

Impact statement:

Post-phosphoryl modification of α-DG requires B4GAT1 (β1,4 glucuronyltransferase); this enzyme synthesizes the acceptor glycan, which serves as a primer for the glycosyltransferase LARGE to initiate synthesis of the laminin-binding glycan.

Authors contribution: T.W. and K.P.C. designed the research and analyzed the data; T.W., D.V., K.I., C.H., G.M., Y.H. and D.V.B performed biochemistry research and analyzed/interpreted the data; L.Y. performed NMR analysis and interpreted the data; K.M.W isolated *B4gat1*-deficient MEFs, which were essential for the study; T.W. and K.P.C wrote the initial manuscript and all authors commented on and approved the manuscript.

Short title: B4GAT1 synthesizes the LARGE glycan acceptor

Keywords: glycosylation, B4GAT1, B3GNT1, LARGE, α-dystroglycan, basement membrane, arenavirus, cancer, proteoglycan

Introduction

Dystroglycan (DG) is a highly glycosylated basement membrane receptor involved in a variety of physiological processes including maintenance of the skeletal muscle-cell membrane integrity and establishment of the structure and function of the central nervous system (1). DG is composed of a cell-surface α-subunit and a transmembrane β-subunit. α-DG acts as a receptor for laminin-G domain-containing extracellular matrix (ECM) proteins such as laminin, agrin, perlecan and neurexin (1). In addition, it serves as a cellular receptor and entry site for most Old World arenaviruses, including the highly pathogenic Lassa virus (LASV) and Clade C New Word arenaviruses (2). LASV is the causative agent of severe hemorrhagic fever in humans, a disease that has a mortality rate of ~15% resulting in several thousand deaths each year.

α-DG effectiveness as a receptor is dependent on complex post-translational modifications. Besides numerous modifications with *N*-glycans and mucin-type *O*-glycans, a highly complicated series of additions to a phosphorylated *O*-mannosyl glycan moiety in the N-terminal region of the mucin domain are essential for ligand binding (3, 4). Defects in the proper post-translational processing of α-DG result in loss of receptor function, and in a broad spectrum of congenital muscular dystrophies (CMDs) that are accompanied by a variety of brain and eye malformations. Collectively, these dystrophies are classified as dystroglycanopathies (1). To date, over 17 genes have been reported to be directly or indirectly involved in this "functional glycosylation" of α-DG, and have been linked to human disease when mutated (5, 6).

Recent gene discovery efforts revealed several novel dystroglycanopathy genes with unknown function (7-9). In our previous work we were able to assign functions to the POMGNT2 (GTDC2), B3GALNT2 and POMK (SGK196) gene products, which contribute to the synthesis of a phosphorylated, O-mannosyl-linked trisaccharide on α-DG (10) (**Fig. 1**). This so-called Core M3 structure (GalNAc-β3-GlcNAc-β4-Man-α Ser/Thr) is synthesized in the endoplasmic reticulum (ER), and it is thought to be a platform for further functional modification of α-DG as it passes through the secretory pathway. The glycosyltransferase LARGE was shown to synthesize and transfer repeating units of [–3-

xylose–α1,3-glucuronic acid-β1–] to α-DG (11). This heteropolymer is postulated to be the terminal glycan moiety anchored by the Core M3 structure. It resembles the ligand-binding glycan and its length correlates with the affinity of α-DG for its ligands (12). However, how the laminin-binding glycan synthesized by LARGE is attached to the Core M3 structure, and which glycans or other molecules contribute to and form part of this linker structure, remains unknown.

We set out to elucidate the structure and monosaccharide composition of the α-DG post-phosphoryl modification, applying a strategic and multifaceted experimental approach starting from the terminal end synthesized by LARGE. We used glycosylation-deficient cells, *in vitro* enzyme assays, deglycosylation strategies, and NMR-based structure analysis as experimental tools.

This strategy revealed a β1,4 glucuronyltransferase activity for B4GAT1. We present experimental evidence that this enzyme B4GAT1, which was previously described in the literature as B3GNT1 (13) in fact encodes for a β1,4 glucuronyltransferase and not a β1,3 N-acetylglucosaminyltransferase as previously thought. This activity contributes to production of the post-phosphoryl glycan linker by transferring a glucuronic acid (GlcA) residue onto a xylose (Xyl) acceptor. It thereby forms a glucuronyl-β1,4-xylosyl disaccharide, the direct acceptor required by the glycosyltransferase LARGE to initiate formation of the terminal heteropolysaccharide that is involved in ligand binding. B4GAT1 enzymatic activity, of both a recombinant form and the endogenous protein in mouse embryonic fibroblasts (MEFs), was further characterized using a newly developed HPLC-based assay for B4GAT1 activity.

Our findings contribute to the current understanding of α-DG posttranslational processing, providing mechanistic insights regarding the pathomechanism underlying α-DG glycosylation-deficient CMDs and revealing new therapeutic avenues for blocking entry of pathogenic LASV viruses.

Results

α-DG O-mannosyl post-phosphoryl modification occurs in the Golgi

Our previous work showed that the ER-resident enzymes POMGNT2 (GTDC2), B3GALNT2 and POMK (SGK196) contribute to synthesis of the phosphorylated Core M3 trisaccharide on α-DG, a moiety that is required as platform for further modification with the LARGE mediated laminin-binding glycan (10). However, a number of additional genes, namely *FKTN* (14, 15), *FKRP* (16, 17) *TMEM5* (8) and *B4GAT1 (B3GNT1)* (9, 18, 19) are known to be crucial for proper α-DG glycosylation, yet how they contribute has not yet been determined (**Fig. 1**). To investigate if these unassigned genes are involved in the pre- or post-phosphorylation process of Core M3, we expressed Fc-tagged recombinant α-DG (DGFc340) in [^{32}P] orthophosphate-labeled control and glycosylation-deficient cells. DGFc340 is a secreted α-DG deletion construct that contains only the minimal region of the α-DG mucin-like domain (aa 316-340), that is required for its functional glycosylation followed by a C-terminal fusion tag encoding the heavy-chain constant (Fc) moiety of human IgG1 (to enable purification of the secreted recombinant protein) (4). Although only a small subpopulation of the expressed DGFc protein enters the pathway for functional maturation it was demonstrated that this truncated α-DG fusion protein is a valuable tool to study α-DG functional glycosylation (4).

The goal was to test if DGFc340 can be [^{32}P] phosphorylated in fibroblasts with defects in various dystroglycanopathy genes (**Tab. 1**). In our experiment, fibroblasts with defects in *FKTN*, *FKRP*, *TMEM5*, *B4GAT1 (B3GNT1)* and *LARGE*, but not in the phosphate kinase *POMK,* were able to produce radioactively labeled DGFc340 (**Fig. 2A**), indicating that FKTN, FKRP, TMEM5, B4GAT1 and LARGE are involved downstream of POMK in the Core M3 post-phosphorylation process.

Immunofluorescence examination of HEK293T cells stably transfected with Myc-tagged constructs of this set of proteins, revealed that they co-localize with the Golgi-resident marker protein Giantin (20) (**Fig. 2B**). Previously, Golgi localization was also demonstrated for FKRP (21), Fukutin (21, 22), B4GAT1 (B3GNT1) (9) and LARGE (23).

These results indicate that most if not all of the α-DG O-mannosyl post-phosphoryl processing is carried out by Golgi-resident enzymes.

The laminin-binding glycan repeat generated by LARGE is hypothesized to be the terminal glycan structure of the α-DG O-mannosyl post-phosphoryl modification (**Fig. 1**). This would suggest that FKTN, FKRP, TMEM5 and B4GAT1 contribute to a post-phosphoryl linker structure, that can serve as an acceptor for the modification with LARGE. Previous work by Kuga et al., (24) also had indicated that FKTN and FKRP are part of the α-DG O-mannosyl post-phosphoryl modification pathway. To test our hypothesis, we infected a panel of glycosylation-deficient cells with a LARGE expressing adenovirus construct and analyzed the glycosylation status of α-DG and the degree of hyperglycosylation by On-Cell immunoblotting with monoclonal antibody IIH6, which recognizes the α-DG laminin-binding glycan transferred by LARGE (11, 12). As expected, overexpression of LARGE did not produce the IIH6-positive glycan or significantly bypass the glycosylation defect in either *FKTN*-, *FKRP*-, *TMEM5*- or *B4GAT1*-deficient cells (**Fig. 2C**), supporting the notion that these encoded proteins work prior to LARGE in the O-mannosyl post-phosphoryl modification process.

In summary, our data suggest that Golgi-localized putative glycosyltransferases FKTN, FKRP, TMEM5 and B4GAT1 are essential for the synthesis of a linker structure that is connecting the α-DG O-mannosyl phosphate platform with the terminal laminin binding glycan added by LARGE.

Glucuronic acid serves as an acceptor sugar for LARGE polymer initiation with Xylose

To further elucidate the structure of the α-DG post-phosphoryl modification, we examined how synthesis of the terminal LARGE glycan was initiated. Although LARGE is known to be a dual glycosyltransferase that synthesizes repeating units of [–3-xylose–α1,3-glucuronic acid-β1–] on α-DG, the identities of both the initiating sugar and the acceptor sugar for this laminin-binding polymer remained unknown. To determine which sugar is initially transferred by LARGE, we developed an *in vitro* glycosylation assay using a

recombinant soluble (transmembrane domain deleted) form of LARGE (LARGEdTM) and the acceptor protein DGFc340. DGFc340 isolated from the culture medium of *Large*^{myd} (*Large*-deficient) MEFs lacks the LARGE modification on phosphorylated O-mannosyl glycans, and is hypothesized to terminate in a glycan acceptor structure that can be recognized by LARGE (11) (**Fig. 3A**). To determine which of the sugars in the polymer is initially transferred by LARGE we incubated DGFc340 isolated from the *Large*^{myd} MEF culture medium with LARGEdTM and [^{14}C]-labeled UDP-xylose (Xyl) and/or UDP-glucuronic acid (GlcA) radionucleotide sugar donors. The glycosyl-transfer reaction was measured as the transfer of radioactivity onto the DGFc340 acceptor glycoprotein. As negative control we used a DGFc340 mutant construct (T317A/T319A), which lacks the O-mannosylation sites that are the crucial acceptor platform for subsequent synthesis of the laminin-binding glycan (4). When the radionucleotide sugars were tested individually in the LARGEdTM *in vitro* assay, the addition of [^{14}C] UDP-Xyl, but not that of [^{14}C] UDP-GlcA radionucleotides, resulted in radioactive labeling of the DGFc340 acceptor (**Fig. 3B**). However, in the presence of both UDP-Xyl and UDP-GlcA, the transfer of radioactivity was significantly increased consistent with the fact that the LARGE glycan is a heteropolysaccharide (**Fig. 3A/B**). These results indicate that xylose is the initial sugar transferred by LARGE, and that this is followed by the transfer of GlcA to form the repeating [–3-xylose–α1,3-glucuronic acid-β1–] heteropolymer.

Next we wanted to identify the acceptor glycan used by LARGE to initiate formation of the laminin-binding glycan. Since the Xyl-T activity of LARGE has acceptor specificity for β-linked GlcA during heteropolymer formation, we hypothesized that β-linked GlcA might be the initial acceptor for the glycan added by LARGE. To test this, we pre-treated DGFc340 from *Large*^{myd} MEF cells with β-glucuronidase (β-GUS) (**Fig. 3A/C**), and assessed its modification by LARGEdTM in an *in vitro* assay. Subsequent immunoblotting with the LARGE glycan-specific antibody (IIH6) revealed that the pretreatment of *Large*^{myd} DGFc340 with β-glucuronidase resulted in a strong reduction of the IIH6 signal (**Fig. 3C**). These data indicate that LARGE uses a β-linked GlcA residue as an acceptor sugar to initiate synthesis of the polymeric glycan.

205

206 **Xylose is part of the α-DG O-mannosyl post-phosphoryl modification**

207 To determine which monosacharides contribute to synthesis of the O-mannosyl post-
208 phosphoryl acceptor for the LARGE glycan, we performed a LARGEdTM assay with
209 DGFc340 acceptor isolated from a panel of sugar nucleotide-deficient CHO cells (25, 26).
210 LARGEdTM was able to efficiently modify DGFc340 from Pro5 (wild-type), Lec2 (CMP-
211 sialic acid-deficient), Lec8 (UDP-galactose-deficient) and Lec13 (GDP-fucose-deficient)
212 cells, suggesting that sialic acid, galactose and fucose do not contribute to functional
213 glycosylation of α-DG (**Fig. 3D**). ldlD cells deficient for UDP-galactose (UDP-Gal) and
214 UDP-N-acetylgalactosamine (UDP-GalNAc) demonstrated reduced acceptor activity for
215 LARGEdTM, which can be explained by the fact that B3GALNT2 requires UDP-GalNAc
216 for synthesis of the initial Core M3 structure. Similarly, DGFc340 from Lec15 cells, which
217 are deficient for Dol-P-Man synthesis, did not serve as LARGEdTM acceptor because
218 POMT1 and POMT2 require the Dol-P-Man sugar donor to initiate the O-mannosyl Core
219 M3 structure. Most interestingly, in the LARGEdTM *in vitro* assay DGFc340 from UDP-
220 Xyl-deficient pgsI-208 CHO cells was not modified (**Fig. 3D**). It had also been shown that
221 ectopic LARGE expression in pgsI-208 CHO cells did not induce α-DG
222 hyperglycosylation (11, 27), consistent with the fact that UDP-Xyl is essential for synthesis
223 of the LARGE glycan *in vivo*. However, in our LARGEdTM *in vitro* assay, pgsI-208
224 DGFc340 was also not modified by LARGEdTM, despite the presence of both of the
225 required sugar nucleotides, UDP-GlcA and UDP-Xyl. This clearly demonstrated that one or
226 more xylose residues are required on α-DG before it can be functionally glycosylated by
227 LARGE (**Fig. 3A)**.

228

229 **B4GAT1 is a glucuronyltransferase with specificity for the β-xylose acceptor**

230 Having identified a β-linked glucuronic acid as the terminal acceptor saccharide for
231 LARGE and xylose as component of the α-DG post-phosphoryl glycan modification, we
232 next sought to determine which enzyme is responsible for the hypothesized
233 glucuronyltransferase activity. Among the group of unassigned genes (*FKTN, FKRP,*

TMEM5, B4GAT1) only B4GAT1 showed homology to glucuronyl-transferases. In particular it shares 44 % similarity with the LARGE GlcA-T domain (CAZy: GT49; **Fig. 4A**). This designated B4GAT1 as a promising candidate for the GlcA-T transferase upstream of LARGE. To test this hypothesis, we generated a 6xHis-tagged soluble construct of B4GAT1 (transmembrane domain deleted, B4GAT1dTM), expressed it in HEK293T cells and purified the recombinant enzyme from the culture medium (**Fig. 5**). We then conducted a transfer assay with B4GAT1dTM as the enzyme source, UDP-GlcA as the sugar donor and fluorescently labeled β-xyloside (4-methylumbelliferyl-β-D-xyloside, Xyl-β-MU) as the acceptor. The reaction products were separated by high-performance liquid chromatography (HPLC). A unique product peak was detected only when UDP-GlcA was used as donor (**Fig. 4B, Fig. 6A**). We also tested the acceptor specificity, which revealed that B4GAT1dTM GlcA-T activity has low preference for α-linked Xyl, but showed >10 fold higher preference and specificity towards β-linked Xyl acceptors (**Fig. 4C, Fig. 6A**). The fact that the LARGE glycan disaccharide Xyl-α1,3-GlcA-MU was a very weak acceptor for B4GAT1dTM GlcA-transfer suggests that B4GAT1 overexpression does not interfere with LARGE mediated synthesis of the laminin-binding glycan. A characterization of the B4GAT1 GlcA-T activity revealed a metal dependence for manganese (Mn^{2+}) divalent cations (**Fig. 6B**) and a pH-optimum near pH 7.0 (**Fig. 6C**). The product peak obtained from the enzymatic reaction of B4GAT1dTM with β-Xyl-MU acceptor was isolated, and its analysis by NMR revealed that the GlcA residue was β-linked to the 4 position of the xylose β-MU (**Fig. 7, Table 2**). Thus, B4GAT1 possesses xylose β1,4-glucuronyltransferase (GlcA-T) activity and it is specific for the substrate β-linked Xyl.

The substrate specificity of LARGE Xyl-T is not dependent on the glycosidic bond of the β-GlcA acceptor

It had previously been shown that, during synthesis of the LARGE heteropolysaccharide, β1,3-linked GlcA serves as the acceptor for LARGE Xyl-T (11). In the current study we found that a β1,4-linked GlcA transferred by B4GAT1 serves as the acceptor glycan for

initiation of synthesis of the LARGE glycan, via the addition of a xylose. To assess if LARGE can use one or the other glycosidic linkage β-linked GlcA acceptor with higher efficiency, we tested LARGEdTM Xyl-T activity on two disaccharides GlcA-β1,4-Xyl-β-MU and GlcA-β1,3-Xyl-α-MU along with the monosaccharide GlcA-β-MU. As shown in **Fig. 8**, LARGE did not distinguish between GlcA-β1,3-Xyl and GlcA-β1,4-Xyl, as similar activities were measured in the presence of both disaccharide acceptors. However, the length of the acceptor appears to be important, since the disaccharide acceptor showed 25-fold higher activity than the monosaccharide acceptor (**Fig. 8**). In summary, it is currently unknown why the β1,4-linked GlcA acceptor in the initial LARGE acceptor primer and the β1,3-linked GlcA in the terminal LARGE glycan have different linkages, and how each linkage contributes spatially to the overall structure, while LARGE shows similar activity towards both acceptors.

MEFs from *B4gat1*-deficient mice lack endogenous B4GAT1 activity

To elucidate further the role of B4GAT1 *in vivo*, we isolated MEFs from control, *Large*^{myd} (*Large*-deficient) and *B4gat1*-deficient mice (19) (**Tab. 1**) and analyzed the glycosylation status of α-DG. Immunoblotting revealed that whereas control MEFs were positive for functional glycosylation of, and laminin-binding by, α-DG from *Large*^{myd} MEFs completely lacked both features (**Fig. 9**). Also, *B4gat1*-deficient MEFs demonstrated strongly reduced but detectable residual functional glycosylation and laminin binding, and normal levels of hypoglycosylated α-DG core protein (**Fig. 9**). Adenovirus-mediated ectopic expression of *B4GAT1* did not affect the glycosylation status of α-DG in control and *Large*^{myd} MEFs but, as expected, was able to rescue the α-DG glycosylation defect in *B4gat1*-deficient MEFs (**Fig. 9**). As demonstrated previously (11, 28, 29), forced adenovirus-mediated ectopic expression of LARGE in control and *Large*^{myd} MEFs induces α-DG hyperglycosylation. In contrast, *B4gat1*-deficient cells showed only a low level of α-DG hyperglycosylation after LARGE overexpression (**Fig. 9**), suggesting that in the context of mutant B4GAT1, only few acceptor sites for LARGE modification are available. Finally, ectopic co-expression of B4GAT1 and LARGE resulted in α-DG

hyperglycosylation in all three tested MEF lines (**Fig. 9**). This result is consistent with the hypothesis that B4GAT1 acts prior to the glycosyltransferase LARGE.

Next, to determine if endogenous B4GAT1 activity was detectable in MEF cells, we subjected samples from control, *B4gat1* and *Large*^{myd} MEFs to the B4GAT1 enzyme activity assay, using Xyl-β-MU as the acceptor. Whereas control and *Large*^{myd} samples showed comparable B4GAT1 transferase activity, only low residual activity (<3%) was detectable in *B4gat1*-deficient cells, and this loss could be restored by ectopic expression of B4GAT1 (**Fig. 10A/B/E)**. Similarly, when we tested LARGE GlcA-T activity in control and glycosylation-deficient MEFs, only *Large*^{myd} MEFs lacked LARGE GlcA-T activity; the control and *B4gat1*-deficient cells were normal (**Fig. 10C/D/F)**. These results suggest that the enzymatic activities of B4GAT1 and LARGE are independent and that each is un-affected by mutations in the gene product of the other.

B4GAT1 mutations affect the subcellular localization and activity of B4GAT1

To date, several disease-causing *B4GAT1* (formerly termed *B3GNT1*) mutations have been reported in human patients (9, 18), in an N-ethyl-N-nitrosourea (ENU)-induced mutant mouse model (19) and in a genetic screen for modifiers of LASV entry (7). To test how reported *B4GAT1* missense mutations affect the intracellular localization of B4GAT1 and its enzymatic activity, we cloned three mutant B4GAT1-Myc expression constructs (**Fig. 11A**): Mut1 (N390D) represents a mutation identified in a patient with Walker-Warburg syndrome (9); Mut2 (D227N/D229N) is a mutation in the glycosyltransferase signature DXD motif (30, 31); and Mut3 (M155T) mimics a mutant allele identified in a *B4gat1*-deficient mouse model with axon guidance defects (19). Immunoblot analysis confirmed that expression levels were similar for the Myc-tagged B4GAT1 control and all three mutant constructs Mut1-Mut3 in stably expressing HEK293T cells (**Fig. 11B**). Previously, it had been shown that B4GAT1 localizes to the *trans*-Golgi near the TGN (*trans*-Golgi network) (9, 31, 32). In our immunofluorescence analysis, we found both the control and mutant construct Mut1 to exhibit normal Golgi localization and to co-localize with the Golgi marker Giantin (**Fig. 11C)**. *B4GAT1* mutations in constructs Mut2 and Mut3,

however, resulted in a high degree of mislocalization to the ER, as judged by overlap of the signal with the ER marker ERp72 (**Fig. 11D**), indicating that the B4GAT1 mutant proteins are misfolded and retained in the ER. Analysis of B4GAT1 enzyme activity in lysates from cells stably overexpressing these constructs revealed strongly reduced activity in the cases of all three mutants, of less than 5% compared to activity levels in wild-type control (**Fig. 11E**). Similarly, none of the B4GAT1 mutant constructs was able to complement and rescue the α-DG glycosylation defect in *B4gat1*-deficient MEFs (**Fig. 11F**). These findings confirm that the identified *B4GAT1* mutations are pathological and have a direct negative impact on B4GAT1 activity regardless of their subcellular localization. Additionally, the finding that the B4GAT1 DXD motif is essential further supports a role for B4GAT1 as a glycosyltransferase, since the DXD motif is thought to be involved in binding carbohydrate sugar-nucleoside diphosphates in manganese-dependent glycosyltransferases (30).

β-Xylose serves as an endogenous acceptor for B4GAT1

To further characterize the endogenous acceptor for B4GAT1, we first tested if B4GAT1dTM was able to use DGFc340 from control, *Large*^{myd} and *B4gat1*-deficient MEFs as an acceptor. Similar to the LARGE acceptor experiment in **Fig. 3B,** we used radiolabeled [^{14}C] UDP-GlcA sugar donor and measured transfer of [^{14}C] to the protein A-bound DGFc340 acceptor. As expected, DGFc340 isolated from *B4gat1*-deficient cells was the only acceptor that incorporated substantial levels of the radioactive label (**Fig. 12A**). This confirmed that only the DGFc340 acceptor from *B4gat1*-deficient cells resembled the terminal acceptor glycan suitable for B4GAT1dTM to add GlcA. Our B4GAT1dTM *in vitro* enzyme assay demonstrated acceptor specificity for β-linked xylose (**Fig. 4C**). To corroborate the hypothesis that β-linked xylose also serves as the endogenous B4GAT1 α-DG acceptor we pre-treated DGFc340 from *B4gat1*-deficient cells with β-xylosidase and measured the transfer of [^{14}C] GlcA by B4GAT1dTM. After β-xylosidase treatment, the ability of *B4gat1*-deficient DGFc340 to act as an acceptor was strongly reduced; this constitutes indirect evidence that a β-linked xylose is indeed the postulated endogenous

acceptor for B4GAT1 (**Fig. 12B**), and that a yet unidentified xylosyltransferase acts upstream of B4GAT1.

NMR-studies confirm that B4GAT1 synthesizes the acceptor glycan for LARGE

To further corroborate our finding that the glycosyltransferase LARGE utilizes a glucuronic acid-β1,4-xylose-β disaccharide acceptor as a primer to elongate it with its dual glycosyltransferase and polymerizing activity, we performed NMR structural studies. In our approach towards confirming each individual glycosidic linkage, we first synthesized the tetrasacharide GlcA-Xyl-GlcA-Xyl-MU, starting with the monosaccharide acceptor Xyl-β-MU and extending it in a stepwise manner using recombinant B4GAT1dTM and LARGEdTM as enzymes sources (**Fig. 13A**).

The ^{1}H and ^{13}C resonances of the isolated tetrasaccharide product were assigned by using heteronuclear multiple quantum coherence (HMQC), heteronuclear 2-bond correlation (H2BC), and heteronuclear multiple bond coherence (HMBC) spectra (, **Fig. 13B**, **Table 3**). The detection of the interglycosidic cross-peaks of BH1/AC4, CC1/BH3, and DH1/CC3 in the HMBC spectrum (**Fig. 13B**) clearly indicates the presence of a 1→4 interglycosidic linkage between sugar residues B and A, a 1→3 interglycosidic linkage between residues C and B, and a 1→3 interglycosidic linkage between residues D and C, respectively. A strong rotating-frame Overhauser effect (ROE) was observed from the H1 to H3 and H5 protons of residues A, B, and D in the ROE spectroscopy (ROESY) spectrum (**Fig. 13C**), demonstrating that they have a β-configuration. The observed strong ROE from the residue C H1 proton to its own H2, but not to H3 and/or H5 demonstrates that the residue C has an α-configuration. The inter-residue ROEs observed in the ROESY spectrum are also consistent with the interglycosidic linkage assignments determined from the HMBC spectrum. Therefore, the tetrasaccharide has the glycosidic linkage structure GlcA-β1,3-Xyl-α1,3-GlcA-β1,4-Xyl-β-MU (**Fig. 13A**). These studies show that B4GAT1 possesses β1,4 glucuronyltransferase activity, and that LARGE can elongate this primer structure by adding repeating units [-3-Xyl-α1,3-GlcA-β1-] to produce a

377 heteropolysaccharide. To further illustrate the complexity of assembling the functional
378 glycan of α-DG, we summarize the current knowledge about the α-DG sugar structures and
379 the contributing genes/enzymes in **Fig. 14**.
380

Discussion

In this study we used a multidisciplinary approach to investigate how the assembly of the α-DG LARGE glycan is initiated, and found that it requires B4GAT1-dependent synthesis of a novel glucuronyl-xylosyl acceptor primer. We show that B4GAT1 is a xylose β1,4-glucuronyltransferase, and that it is involved in synthesizing the glycan primer that subsequently can be elongated by LARGE with the ligand-binding glycan. B4GAT1 was initially cloned and described by Sasaki et al., (13) as β1,3-N-acetylglucosaminyltransferase (iGnT, β3GNT1 or B3GNT1), which is essential for the synthesis of poly-N-acetyllactosamine. Furthermore, the B3GNT1 enzyme was proposed to contribute to the i antigen synthesis pathway by transferring N-acetylglucosamine onto a β-galactose acceptor with β1,3 linkage (13). In contrast our data reveal a β1,4-glucuronyltransferase activity, which we have designated B4GAT1 (B3GNT1). We tested B4GAT1dTM with UDP-GlcNAc and the proposed Gal-β1,4-GlcNAc-β-MU acceptor, but we were not able to validate any N-acetylglucosaminyltransferase activity (**Fig. 15)**. Therefore, we propose to rename the enzyme B4GAT1, as a new member of the glucuronyltransferase family of proteins. To date, only 2 other enzymes are known to have β1,4-glucuronyltransferase activity. These are EXT1 and EXT2, and both are involved in the synthesis of heparan sulphate proteoglycans (33).

Our findings regarding assembly of the LARGE glycan reveal striking similarities to the unique mechanism underlying the synthesis of proteoglycans. Both glycan polymers consist of repeating disaccharides that are synthesized by glycosyltransferases with dual glycosyltransferase activities (11, 34). Furthermore, in both cases assembly of the terminal heteropolymer glycan is initiated by a disaccharide primer, which is part of a larger glycan linker that anchors the polysaccharide to a protein backbone. Future studies are needed to elucidate the full α-DG glycan structure and determine the roles of the putative glycosyltransferases FKTN, FKRP and TMEM5 in anchoring the α-DG ligand-binding glycan moiety to the phosphorylated Core M3 structure. At this point it cannot be ruled out

that other, currently unidentified, genes also contribute to synthesis of the functional α-DG glycan.

Similar to their counterparts in other dystroglycanopathy genes, *B4GAT1 (B3GNT1)* loss-of-function mutations in human patients result in Walker-Warburg Syndrome (WWS) (9, 18), the most severe condition in a range of clinically defined CMDs that are accompanied by brain and eye malformations. Milder *B4GAT1* mutations with residual enzyme activity are expected to cause a milder Limb Girdle Muscular Dystrophy (LGMD) phenotype, but patients with such mutations have not yet been described.

B4gat1 (B3gnt1)-null mutations in mice result in early embryonic lethality, at ~ E9.5 (19), as is the case for reported null mutations in *Dag1* (35), *Pomt1* (36), and *Fukutin* (37). Proper α-DG glycosylation is essential for early embryonic development in the mouse, including formation of the basement membrane, as defects in the Reichert's membrane are the suspected cause of death in α-DG glycosylation-deficient mice (35-37). However, an ENU-based genetic screen for abnormal CNS axonal tracks identified a viable *B4gat1 (B3gnt1)* dystroglycanopathy mouse model carrying a p.M155T *B4gat1* mutation (19). The majority of compound heterozygous mice with both a *LacZ* ($B4gat1^{LacZ}$) null allele and a hypomorphic p.M155T ($B4gat1^{M155T}$) allele die perinatally, but a few survive and develop a characteristic CMD phenotype. In this study we used MEFs isolated from the $B4gat1^{LacZ/M155T}$ mice and measured endogenous B4GAT1 activity. As expected the *B4gat1*-deficient MEFs were hypomorphic, producing low-level residual B4GAT1 activity (<3% relative to levels in wild-type control) (**Fig. 10E**). This corroborates that our B4GAT1 assay can be a valuable diagnostic tool for measuring endogenous activity in patient cells and tissues. The residual B4GAT1 enzyme activity in the *B4gat1*-deficient MEFs was also reflected when the α-DG glycosylation status was analyzed biochemically, by immunoblotting. Although B4GAT1 endogenous activity was very low, it was sufficient to synthesize low amounts of functionally active α-DG that was capable of binding the ligand laminin (**Fig. 9**). This finding accounts for the difference between the early

435 embryonic lethal phenotype in *B4gat1* null (*B4gat1*$^{LacZ/LacZ}$) mice and the slightly milder
436 phenotype in *B4gat1* hypomorphic (*B4gat1*$^{LacZ/M155T}$) mice (19).

437 It is worth noting that α-DG glycosylation is highly tissue specific as well as highly
438 dependent on the developmental stage of the cells/tissue (1). To date, it is not fully
439 understood what causes the tissue-specific differences in α-DG processing, which are
440 reflected as differences in its molecular weight and its ability to bind laminin (12). *LARGE,*
441 the key contributor to assembly of the terminal laminin-binding glycan, and *B4GAT1* as the
442 upstream priming enzyme are broadly expressed at the RNA level (**Fig. 16**). Although both
443 genes are similarly expressed in most tissues they are strikingly different in heart with
444 *LARGE* expression being high and *B4GAT1* being low. Based on published case reports it
445 does not appear that *LARGE* patients (38-40) are more prone to cardiac defects than other
446 dystroglycanopathy patients. Also lower *B4GAT1* expression in the heart does not present a
447 significant bottleneck for α-DG functional glycosylation as heart α-DG has full ligand
448 binding ability (12). Therefore, the functional consequences of such uncoordinated
449 expression of *B4GAT1* and *LARGE* are currently unknown. It is more likely that other gene
450 products involved in α-DG functional glycosylation can become limiting factors and that
451 the integration of all involved players account for the tissue-specific differences of this
452 complex and highly controlled synthesis pathway. Furthermore, whether α-DG that is not
453 properly glycosylated possesses an as yet unidentified ligand binding activity remains
454 unclear. LARGE has been shown to be highly tunable in the context of cancer, T-cell
455 development and muscle regeneration (12, 41, 42). Repression of *LARGE* expression is
456 responsible for the defects in DG-mediated cell adhesion that are observed in epithelium-
457 derived cancer cells, and point to a defect of its glycosylation as a factor in cancer
458 progression (42). Similarly, it was demonstrated that expression of B4GAT1 (B3GNT1) is
459 absent in a IIH6-negative subpopulation (PC3-L) of an otherwise IIH6-positive human
460 prostate cancer cell line (PC3). The loss of *B4GAT1* expression and laminin-binding by α-
461 DG in these cells was inversely correlated with the observed malignancy and tumor
462 progression of the prostate cancer when these cells were transplanted into SCID mice (31).

In general, these results emphasize that proper α-DG glycosylation plays a critical role in tumor suppression.

Previous data suggested that B4GAT1 (B3GNT1) may be an integral component of various enzyme complexes, working with various glycosyltransferases that are functionally associated and involved in the same biosynthetic pathway. For example, it might work with B4GALT1 (32) in the synthesis of poly-N-acetyllactosamine and with LARGE (31) in synthesis of the α-DG laminin-binding glycan. It was also hypothesized that B4GAT1 may regulate LARGE, as B4GAT1 overexpression promoted formation of the LARGE-generated laminin-binding glycan (31). However, in light of data presented in this study, in particular the finding that endogenous LARGE activity is not affected in *B4gat1*-deficient cells and vice versa, it seems more likely that B4GAT1 and LARGE have independent enzyme activities (**Fig. 10E/F**).

In an effort to provide additional direct evidence and further corroboration of our conclusion that a xylose present in the α-DG O-mannosyl post-phosphoryl glycan linker serves as endogenous acceptor for B4GAT1, we performed radioactive metabolic cell labeling with [3H]-xylose. The goal was to show radioactive labeling of DGFc340 expressed in *B4gat1*-deficient MEFs with [3H]-xylose, which could subsequently be released by β-xylosidase treatment. However, this type of metabolic cell labeling proved to be technically challenging since only an insignificant amount (~ 0.01%) of the total [3H]-xylose radioactivity was incorporated into the secreted DGFc340 fusion protein even after 4 day long-term labeling (data not shown). It is known that xylose uptake from the media into cells is poor (43), which in our case becomes the limiting factor and made this experimental approach not feasible. Nevertheless, we feel confident that the sum of our indirect data including B4GAT1 Xyl-β-MU acceptor specificity (**Fig. 4C)**, pgsI-208 DGFc340 LARGE acceptor test (**Fig. 3D)**, β-xylosidase B4GAT1 acceptor treatment (**Fig. 12B)** and finally the *in vitro* synthesis of a GlcA-β1,3-Xyl-α1,3-GlcA-β1,4-Xyl-β-MU tetrasaccharide by the sequential action of B4GAT1 and LARGE (**Fig. 13**) provide strong and convincing evidence that β-xylose is indeed the endogenous acceptor for B4GAT1.

491 In conclusion, our study has identified B4GAT1 as a xylose β1,4-
492 glucuronyltransferase, and revealed that it contributes to the O-mannosyl post-phosphoryl
493 glycan linker of α-DG by synthesizing a glucuronyl-xylosyl disaccharide. This is the
494 crucial acceptor primer that is targeted by the glycosyltransferase LARGE to initiate
495 formation of a heteropolysaccharide on α-DG that is involved in its binding to ligands. As
496 *B4GAT1*-deficiency was linked to laminin-binding defects of α-DG in a variety of contexts,
497 our new findings will shed light on the mechanism underlying α-DG glycosylation-
498 deficient CMDs (9, 18) and tumors (31), and is expected to also open new therapeutic
499 avenues for blocking the entry of pathogenic arenaviruses, including the hemorrhagic
500 LASV into human cells (7).
501

Materials and methods

Subjects and samples

All tissues and patient cells were obtained and tested according to the guidelines set out by the Human Subjects Institutional Review Board of the University of Iowa; informed consent was obtained from all subjects or their legal guardians (see **Tab. 1**).

Cell cultures

Cells were maintained at 37°C and 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM) plus fetal bovine serum (FBS: 10% in the case of HEK293T cells, 20% in the case of fibroblasts from patient skin) and 2 mM glutamine, 0.5% penicillin-streptomycin (Invitrogen, Carlsbad, CA). Pro5 (wild-type) and the glycosylation-deficient CHO (Lec cells) mutant cell lines termed Lec2 and Lec8 were purchased from ATCC (44). The Lec15.2 (45) and ldlD (26) cell lines were kindly provided by Monty Krieger, the Lec13 (46) cells by Pamela Stanley and the pgsI-208 (47) cells by Jeff Esko. These CHO cells were grown and maintained in F12 nutrition mix medium with 10% fetal bovine serum (Invitrogen, Carlsbad, CA) at 37°C and 5% CO_2. MEFs were generated from E13.5 embryos (**Tab. 1)** as previously described (48) and were maintained in DMEM supplemented with 10% FBS, 2 mM glutamine, and 0.5% penicillin-streptomycin at 37°C in 5% CO_2.

[^{32}P] orthophosphate labeling of cells

Phosphorylation of α-DG in glycosylation-deficient fibroblasts was determined based on the incorporation of [^{32}P] into a secreted Fc-tagged α-DG recombinant protein, as described elsewhere (49).

Adenovirus generation and gene transfer

E1-deficient recombinant adenoviruses (Ad5CMV-DGFc340, Ad5CMV-DGFc340mut (T317A/T319A) and Ad5CMV-*LARGE*/RSVeGFP) were generated by the University of Iowa Gene Transfer Vector Core and have been described previously (28). The constructs used to generate the E1-deficient recombinant adenoviruses Ad5CMV-

DGFc340 and Ad5CMV-DGFc340mut (T317A/T319A) were made from pcDNA3-DGFc340 and DGFc340mut (T317A/T319A) (4). pcDNA3-DGFc340 vectors were digested with KpnI/XbaI, and the resulting fragments were ligated into a KpnI/XbaI-digested pacAd5-CMV-KNpA vector. Similarly, Ad5CMV-*B4GAT1*-V5/RSVeGFP was generated by PCR amplifying a 1.3 kb C-terminal V5-tagged open reading frame fragment corresponding to mouse *B4gat1* (*B3gnt1,* NM_175383) and cloning it into the XhoI/NotI polylinker region of pAd5CMVK-NpA. The following primers were used to amplify, by PCR, *B4gat1*-V5: 6823 forward (5'-aga**ctcgag**accATGcaaatgtcctacgccat-3', XhoI adapter is bolded, start ATG is shown in capital letters) and 6822 reverse (5'-tat**gcggccgc**CTACGTAGAATCGAGACCGAGGAGAGGGTTAGGGATAGGCTTACCg catcggtggggagagttgg-3'; the NotI adapter is bolded and the V5-tag is shown in capital letters). Cultured cells were infected with viral vector for 12 h, at an MOI of 400. We examined cultures 3–5 d after treatment. We used nucleofection as nonviral method for transferring genes into MEF cells. The Human Dermal Fibroblast Nucleofector kit was used according to an optimized protocol provided by the manufacturer (Amaxa Biosystems, Germany).

Glycoprotein enrichment and biochemical analysis

WGA-enriched glycoproteins from frozen samples and cultured cells were processed as previously described (50). Immunoblotting was carried out on polyvinylidene difluoride (PVDF) membranes as previously described (50). Blots were developed with IR-conjugated secondary antibodies (Pierce) and scanned with an Odyssey infrared imaging system (LI-COR Bioscience, Lincoln, NE). Laminin overlay assays were performed as previously described (50).

The monoclonal antibodies to the fully glycosylated form of α-DG (IIH6) (51), and also the polyclonal antibodies rabbit β-dystroglycan (AP83)(52) and anti-LARGE (Rb331)(3) were characterized previously. G6317 (core-αDG) from rabbit antiserum was raised against a keyhole limpet hemocyanin (KLH)-conjugated synthetic peptide of human dystroglycan (29). Mouse monoclonal anti-Myc (clone 4A6) antibodies were purchased from Millipore (Billerica, MA), mouse monoclonal anti-β-Actin (Clone AC-74) antibodies

were purchased from Sigma (St. Louis, MO) and mouse monoclonal anti-V5 antibodies were purchased from Invitrogen (Carlsbad, CA).

Immunofluorescence microscopy

HEK293T cells expressing Myc-tagged glycosyltranferases were fixed with 4% paraformaldehyde in PBS, and then permeabilized with 0.2% Triton X-100 in PBS for 10 min on ice. After blocking with 3% BSA in PBS, the slides were incubated with anti-c-Myc antibody (4A6, Millipore) anti-Giantin (abcam) or anti-ERp72 antibody (Calbiochem) for 18 hours at 4°C. The cells were incubated with an appropriate secondary antibody conjugated to Alexa488 or Alexa555 fluorophore after washing with PBS. 4',6'-Diamidino-2-phenylindole dihydrochloride (DAPI, Sigma) was used for nuclear staining. Images were observed using a Zeiss Axioimager M1 fluorescence microscope (Carl Zeiss, Thornwood, NY).

On-Cell complementation and Western Blot assay

The On-Cell complementation assay was performed as described previously (29). In brief, $2x10^5$ cells were seeded into a 48-well dish. The next day the cells were infected with 200 MOI of Ad5CMV-*LARGE1/eGFP* in growth medium. Three days later, the cells were washed in TBS and fixed with 4% paraformaldehyde in TBS for 10min. After blocking with 3% dry milk in TBS + 0.1% Tween (TBS-T), the cells were incubated with primary antibody (glyco α-DG, IIH6) in blocking buffer overnight. To develop the On-Cell Western blots we conjugated goat anti-mouse IgM (Millipore) with IR800CW dye (LI-COR), subjected the sample to gel filtration, and isolated the labeled antibody fraction. After staining with IR800CW secondary antibody in blocking buffer, we washed the cells in TBS and scanned the 48-well plate using an Odyssey infrared imaging system (LI-COR Bioscience, Lincoln, NE). For cell normalization, DRAQ5 cell DNA dye (Biostatus Limited) was added to the secondary antibody solution.

Cloning of C-terminal Myc-tagged B4GAT1, TMEM5, FKTN, FKRP and LARGE

Open reading frames (ORF) were PCR amplified using the following primer sequences:

m*B4gat1* (1.3 kb), pTW324: forward 5'-aagGGATCCacc**atg**caaatgtcctacgccatccg-3' (BamHI restriction site is shown in capital letters and start ATG is bolded) and reverse 5'-aga**gcggccgc**CTACAAGTCTTCTTCAGAAATAAGTTTTTGTTC**GCTAGC**cccgcatcggtg gggagagttgggg-3'(NotI restriction site is bolded, Myc-tag sequence is underlined and NheI restriction site is shown in capital bold letters)

h*FKTN* (1.4 kb), pTW322: forward 5'-taaAGATCTacc**atg**agtagaatcaataagaacgtggttttg-3' (BglII restriction site is shown in capital letters and start ATG is bolded) and reverse 5'-ttcGCTAGCcccatataactggataacctcatcccactc-3' (NheI restriction site is shown in capital letters)

m*Fkrp* (1.5 kb), pTW323: forward 5'-taaGGATCCacc**atg**cggctcacccgctgctg-3' (BamHI restriction site is shown in capital letters and start ATG is bolded) and reverse 5'-ttcGCTAGCcccaccgcctgtcaagcttaagagtgc-3' (NheI restriction site is shown in capital letters)

m*Tmem5* (1.3 kb) pTW330: forward 5'-taaGGATCCacc**atg**cggctgacgcggacacg -3' (BamHI restriction site is shown in capital letters and start ATG is bolded) and reverse 5'-ttcGCTAGCcccaactttattattaataaaaaatgaactttc -3'(NheI restriction site is shown in capital letters)

m*Large* (2.3 kb), pTW355: forward 5'-taaAGATCTacc**atg**ctgggaatctgcagagggag-3' (BglII restriction site is shown in capital letters and start ATG is bolded) and reverse 5'-ttcGCTAGCcccgctgttgttctcagctgtgagatatttc-3' (NheI restriction site is shown in capital letters)

First a BamHI/NotI digested PCR fragment from m*B4gat1* was cloned into the BamHI/NotI multiple cloning site (MCS) of a pIRES-puro3-derived vector, in which the NheI site in the MCS was deleted. Subsequently all other genes were digested with either BamHI and NheI or BglII and NheI, and subcloned into a BamHI and NheI-digested m*B4gat1*-myc pIRES-puro (pTW324) construct.

Cloning of B4GAT1-Myc Mut 1-3 mutant constructs

To generate the mouse *B4gat1*-Myc Mut1-Mut3 mutant expression constructs, we used the same forward primer A (5'-aagGGATCCaccatgcaaatgtcctacgccatccg-3') and a reverse primer D (5'-agagcggccgcCTACAAGTCTTCTTCAGAAATAAGTTTTTGTTCGCTAGCcccgcatcggtggggagagttgggg-3') that were used to clone m*B4gat1*-Myc (see m*B4gat1*-Myc pTW324 cloning). Primers A and D bind at the 5'-end and 3'-end of the m*B4gat1* coding region. For each mutation we designed overlapping forward (B1-3) and reverse (C1-3) primers that included the respective mutation (shown in bold capital letters):

mB4GAT1-Mut1 (c.1168 A>G, p.N390D): B1 5'-ccaaaaggaggctgaa**G**accagcgcaataagatc-3' and C1: 5'-gatcttattgcgctggt**C**ttcagcctccttttgg-3'

mB4GAT1-Mut2 (c.679/685 G>A, p.D227N/D229N): B2 5'-ggccaactacgccctggtgatt**A**atgtg**A**acatggtgcccagcgaagggc-3' and C2 5'-gcccttcgctgggcaccatgt**T**cacat**T**aatcaccagggcgtagttggcc-3'

mB4GAT1-Mut3 (c.464 T>C, M155T): B3 5'- gcgctagggtcgcca**C**gcacctcgtgtgcccctc-3' and C3 5'- gaggggcacacgaggtgc**G**tggcgaccctagcgc

Using the m*B4gat1*-Myc (pTW324) expression construct as template, we PCR amplified 5'-fragments, using primer pairs A/B1-3 and 3'-fragments using C1-3/D, respectively. The PCR products were isolated and used as the template DNAs in the second round of amplification with primer pair A-D. The 1.3 kb final PCR product was purified and digested with BamHI/NheI and then ligated into pTW324 digested with the same enzymes. The sequence of the insert DNA was confirmed by Sanger sequencing.

Cloning of B4GAT1dTM

The construct expressing B4GAT1 without its transmembrane region was generated by amplifying a 1.1 kb cDNA fragment of mouse *B4gnt1* (*B3gnt1*, acc.# NM_175383) from mB4GAT1-Myc expression vector pTW324, using primer pair #8629 (5'-ggtGAATTCcacggccaggaggagcagg - 3') and #8630 (5'-atgACCGGTatgcatattcaagtcttcttcagaaataagtttttgttcgc -3'). EcoRI and AgeI restriction sites included in the primers are indicated in capital letters. The PCR fragment was digested with

643 EcoRI and AgeI and subcloned to generate construct pCMV3xFLAG-TEV-B4GAT1dTM-
644 Myc6xHIS (pTW351), which expresses a mouse B4GAT1dTM fusion protein (amino acids
645 37-415) tagged with a N-terminal 3xFLAG and C-terminal Myc6xHis.

646 **Generation of cell lines stably expressing B4GAT1dTM proteins**

647 HEK293T cells were transfected with constructs pTW351 (B4GAT1dTM) using
648 FuGENE 6 (Roche Applied Science). The construct contains an IRES-puromycin
649 resistance cassette and stable cell lines were selected in medium containing Puromycin
650 (1μg/ml, InvivoGen). Expression and secretion of B4GAT1dTM protein into the culture
651 medium was confirmed by immunoblotting with anti-Myc antibody 4A6 (Millipore). The
652 stable cell lines obtained in this way were adapted to serum-free medium 293SFMII
653 (Invitrogen) and cultivated in CELLine bioreactors (CL1000, Argos Technologies).

654 **Purification of B4GAT1dTM and LARGEdTM**

655 B4GAT1dTM and LARGEdTM secreted into the culture medium by HEK293T cells
656 were purified using the Talon metal-affinity resin (Clontech) according to the
657 manufacturer's instructions. The purity of the protein was confirmed by SDS-PAGE and
658 Coomassie Brilliant blue (CBB) staining (**Fig. 5B**). For the enzyme assay, the eluate was
659 desalted and concentrated using an Amicon Ultra centrifugal filter unit (Millipore).

660 **DGFc340 *in vitro* LARGEdTM assay**

661 To generate the DGFc340 and DGFc340-mut acceptor proteins we infected control
662 and glycosylation-deficient MEFs and CHO-derived cell lines with Ad5-CMV DGFc340
663 adenoviral vectors at an MOI of 400. At 4 days post-infection the secreted proteins were
664 isolated from the culture medium using Protein A-agarose beads (Santa Cruz). DGFc340
665 bound Protein A-agarose beads were washed three times with TBS and Protein A slurry
666 prebound with ~25 μg DGFc340 was added to the *in vitro* LARGEdTM assay. Enzyme
667 reactions (50 μl) were carried out at 37ºC, with 5 mM UDP-GlcA and 5 mM UDP-Xyl, in
668 0.1 M MES buffer (pH 6.0) supplemented with 10 mM $MnCl_2$, 10 mM $MgCl_2$, 0.2% Triton
669 X-100 and 5 μg purified LARGEdTM protein. The reaction was terminated by adding 5x
670 LSB and boiling for 5 min, The samples were subsequently analyzed by immunoblotting.

DGFc340 [^{14}C] radioactive sugar donor *in vitro* assay

DGFc340 (~25 µg) and DGFc340-mut (~25 µg) bound Protein A-agarose beads were washed with TBS and used in the *in vitro* LARGEdTM assay. 30 µl enzyme reactions were carried out at 37ºC for 20 h, with 0.05 µCi UDP-GlcA [GlcA-^{14}C] (final conc. 5.5 µM) and 0.05 µCi UDP-Xyl [Xyl-^{14}C] (final conc. 6.6 µM), in 0.1 M MES buffer (pH 6.0) supplemented with 10 mM $MnCl_2$, 10 mM $MgCl_2$, 0.2% Triton X-100 and 5 µg purified LARGEdTM protein. The reaction was terminated by adding 25 µl of 0.1 M EDTA. After three washes with TBS the Protein A-agarose-bound DGFc340 samples were analyzed by scintillation counting.

The reactions for B4GAT1dTM activity were carried out similarly. Again, 30µl enzyme reactions were carried out at 37ºC for 20 h and with 0.05 µCi UDP-GlcA [GlcA-^{14}C] (final conc. 5.5 µM), but in this case 0.1 M MOPS buffer (pH 7.0) supplemented with 10 mM $MnCl_2$, 10 mM $MgCl_2$, 0.2% Triton X-100 was used, with 0.25 µg purified B4GAT1dTM protein.

[^{14}C] labeled sugar nucleotides were purchased from ARC (American Radiolabeled Chemicals, St. Louis)

Glycosidase digestion

Recombinant β-glucuronidase from *E.coli* was purchased from Sigma (G8295). Each digest was performed in a 100µl volume at 37ºC for 12 h in 50 mM $NaPO_4$, pH 7.0, 5 mM DTT, 1 mM EDTA, 0.1% Triton X-100 in the presence of 10µg (100 units) β-glucuronidase.

Recombinant β-xylosidase from *E.coli* was purchased from Sigma(X3504). Each digest was performed in a 100µl volume at 70ºC for 60min in 50 mM sodium acetate at pH 5.8 in the presence of 20µg β-xylosidase.

Analysis of enzymatic activities of B4GAT1 and LARGE

The HPLC-based enzyme assays for B4GAT1-Myc (100 µg cell lysates) and B4GAT1dTM (0.25 µg purified protein) were performed using Xyl-β-MU (0.1 mM) (Sigma-Aldrich) as the acceptor. The samples were incubated for 2 h for analytical

699 purposes and 24 h for preparative purposes. 50 µl enzyme reactions were carried out at
700 37ºC, with 5 mM UDP-GlcA, in 0.1 M MOPS buffer (pH 7.0) supplemented with 10 mM
701 $MnCl_2$, 10 mM $MgCl_2$, and 0.2% Triton X-100. The reaction was terminated by adding 25
702 µl of 0.1 M EDTA and boiling for 5 min. The supernatant was analyzed using a LC18
703 column (4.6 x 250 mm Supelcosil LC-18 column (Supelco)) with Buffer A (50 mM
704 ammonium formate pH 4.0) and Buffer B (80% acetonitrile in buffer A), using a 12%B
705 isocratic run at 1 ml/minute using Beckman Gold system. The elution of MU derivatives
706 was monitored by fluorescence detection (325 nm for excitation, and 380 nm for emission).
707 For the assessment of metal dependence, all ions were used at a concentration of 10 mM in
708 0.1 M MOPS pH 7.0. To test pH-dependent activity testing buffers ranging from pH 4.5-
709 8.5 were used: 0.1 M sodium acetate (pH 4.5-5.5), 0.1 M MES (pH 5.5-6.5), 0.1 M MOPS
710 (pH 6.5-7.5) and 0.1 M Tris (pH 7.5-8.5).
711 To assess endogenous B4GAT1 GlcA-T activity in MEFs, we solubilized the cells in TBS
712 1% TX-100. 100 µg total protein from crude lysates were added to each assay. 50 µl
713 enzyme reactions were carried out for 18h at 37ºC, with 5 mM UDP-GlcA, in 0.1 M MOPS
714 buffer (pH 7.0) supplemented with 10 mM $MnCl_2$, 10 mM $MgCl_2$, and 0.2% Triton X-100.
715 For analysis of substrate specificity Xyl-α-MU (Sigma), Xyl-β-MU (Sigma) and Xyl-α1,3-
716 GlcA-β-MU were added to the standard enzyme reaction at a concentration of 0.1 mM.
717 The HPLC-based enzymatic assay for LARGEdTM (5 µg purified protein) and
718 endogenous LARGE was performed using GlcA-β-MU, GlcA-β1,3–Xyl-α-MU and GlcA-
719 β1,4–Xyl-β–MU as the acceptor for Xyl-T activity and Xyl-α1,3-GlcA-β-MU for GlcA-T
720 activity as described previously (11, 53, 54). For the assessment of endogenous LARGE
721 GlcA-T activity in MEF cells, we solubilized the cells in TBS 1% TX-100 and enriched
722 glycoproteins from crude lysates (2mg total protein) using WGA-agarose. N-
723 Acetylglucosamine-eluted glycoproteins from WGA-bound glycoproteins were incubated
724 in a volume of 50 µl for 18 h at 37°C, with 0.1 mM MU-acceptor, 5 mM UDP-GlcA in 0.1
725 M MES buffer pH 6.0, 10 mM $MnCl_2$, 10 mM $MgCl_2$ and 0.2% Triton X-100. The reaction
726 was terminated by adding 25 µl of 0.1 M EDTA and boiling for 5 min, and the supernatant
727 was analyzed with an LC-18 column using a 12%B isocratic run.

Analysis of B4GAT1 GlcNAc-transferase enzyme activity

The test B4GAT1dTM for GlcNAc transferase activity Gal-β1,4-GlcNAc-β-MU (0.1 mM) was used as acceptor . The 50 µl enzyme reactions were carried out as described previously (13) at 37ºC, with 5 mM UDP-GlcNAc in 0.1 M cocodylate buffer (pH 7.0) supplemented with 20 mM $MnCl_2$, 5 mM ATP and 0.25 µg B4GAT1dTM enzyme. The reaction was terminated by adding 25 µl of 0.1 M EDTA and boiling for 5 min. The supernatant was analyzed using a LC18 column (4.6 x 250 mm Supelcosil LC-18 column (Supelco)) with Buffer A (50 mM ammonium formate pH 4.0) and Buffer B (80% acetonitrile in buffer A), using a 16%B isocratic run at 1 ml/minute using Beckman Gold system.

Separation and purification of the disaccharide generated by B4GAT1dTM

A large scale reaction was carried out using B4GAT1dTM purified using a metal-affinity resin as described previously for LARGEdTM (11). B4GAT1dTM was added to 10 mM of UDP-GlcA and Xylose-β−MU in 50 mM MOPS buffer pH 6.0, 10 mM $MgCl_2$, 10 mM $MnCl_2$ and 0.5% TX-100 and incubated for 48 h at 37°C with rotation. The sample was then run over a C18 column (4.6 x 250 mm Supelcosil LC-18 column (Supelco)) with Buffer A (50 mM ammonium formate pH 4.0) and Buffer B (80% acetonitrile in buffer A) using a 16%B isocratic run at 1 ml/minute on a Beckman Gold system. The elution of MU derivatives was monitored by fluorescence detection (325 nm for excitation, and 380 nm for emission). The product in the peak fractions was collected and lyophilized. The dried sample was then brought up in Milli-Q water (500 µl) and lyophilized and this procedure was repeated three times, after which the sample was brought up in Milli-Q water. The product was quantitated based on the standard curve of GlcA-β-MU. This sample was used for NMR studies.

Separation and purification of the tetrasaccharide generated by B4GAT1dTM and LARGEdTM

The GlcA-β1,4-xylose-β-MU disaccharide (B4GAT1 product) was added to 10 mM of UDP-Xyl in 50 mM sodium acetate buffer at pH 5.5 and with 10 mM $MgCl_2$, 10 mM $MnCl_2$, 0.5% TX-100 and LARGEdTM attached to metal-affinity resin and incubated for

48 hours at 37°C with rotation. The sample was then run over a LC18 column (4.6 x 250 mm Supelcosil LC-18 column (Supelco)) with Buffer A (50 mM ammonium formate pH 4.0) and Buffer B (80% acetonitrile in buffer A) using a 16%B isocratic run at 1 ml/minute on a Beckman Gold system. The elution of MU derivatives was monitored by fluorescence detection (325 nm for excitation, and 380 nm for emission). The trisaccharide peak was collected and lyophilized. The lyophilized sample was then brought up in 10 mM UDP-GlcA in 50 mM MOPS buffer pH 6.0, 10 mM $MgCl_2$, 10 mM $MnCl_2$ and 0.5% TX-100 and incubated for 48 hours at 37°C with rotation. It was again run on a C18 column with 16% B isocratic run. The product peak fraction was then collected and lyophilized. The dried sample was brought up in Milli-Q water (500 µl) and lyophilized. This procedure was repeated a total of three times. The last time the sample was brought up in Milli-Q water and the product was quantitated using a standard curve of GlcA-β-MU. This sample was used for NMR studies.

Synthesis and purification of the disaccharide generated by B4GAT1

A large scale reaction was carried out using recombinant human B4GAT1 (purchased from R&D Systems cat# 3609-GT). B4GAT1 (1.5µg) was added to 5 mM of UDP-Gal and 3mM GlcNAc-β−MU in 50 mM Tris buffer pH 7.5, 10 mM $MgCl_2$ and 150 mM NaCl and incubated for 48 h at 37°C with rotation. The sample was then run over a C18 column (4.6 x 250 mm Supelcosil LC-18 column (Supelco)) with Buffer A (50 mM ammonium formate pH 4.0) and Buffer B (80% acetonitrile in buffer A) using a 16%B isocratic run at 1 ml/minute on a Beckman Gold system. The elution of MU derivatives was monitored by fluorescence detection (325 nm for excitation, and 380 nm for emission). Over time in the above reaction a peak was seen that ran about 1.5 minutes after the GlcNAc-β−MU peak at 21.5 minutes. This product peak was collected was and lyophilized. The dried sample was then brought up in Milli-Q water (500 µl) and lyophilized and this procedure was repeated three times, after which the sample was brought up in Milli-Q water. The product was quantitated based on the standard curve of GlcA-β-MU. This sample was used for NMR studies.

NMR analysis

Samples were prepared for NMR by fractionation (using gel filtration and/or LC-18 chromatography) as described above, followed by the exchange of hydroxyl hydrogens by lyophilization and dissolution in 10 mM sodium phosphate buffer pH 6.5, in 100% D_2O. ^{1}H homonuclear two-dimensional DQF-COSY (55), TOCSY (56), and ROESY (57) experiments, and $^{1}H/^{13}C$ heteronuclear two-dimensional HMQC, HMBC, and H2BC experiments (58) were collected using a Bruker Avance II 800 MHz NMR spectrometer equipped with a sensitive cryoprobe. All NMR spectra were recorded at 25°C. The ^{1}H chemical shifts were referenced to 2,2-dimethyl- 2-silapentane-5-sulfonate (DSS). NMR spectra were processed using the NMRPipe software package (59) and analyzed using NMRView software (60).

Acknowledgements

We thank the Gene Transfer Vector Core (UI, supported by NIH/NIDDK P30 DK 54759) for generating adenoviruses; We thank Pamela Stanley, Monty Krieger and Jeff Esko for providing us with CHO mutant cells, David Ginty for providing us with *B4gat1* (*B3gnt1)*-deficient mice, Hans v. Bokhoven for providing us with patient fibroblasts, members of the Campbell laboratory for fruitful discussions; Andrew Crimmins for technical support; Christine Blaumueller for critical reading of the manuscript. This work was supported in part by a Paul D. Wellstone Muscular Dystrophy Cooperative Research Center Grant (1U54NS053672, K.P.C. and T.W.), a MDA grant (238219, K.P.C. and T.W.) and an ARRA Go Grant (1 RC2 NS069521-01, K.P.C. and T.W.). K.P.C. is an investigator of the Howard Hughes Medical Institute.

The authors declare no conflicts of interest.

Abbreviations

CHO, Chinese hamster ovary; DG, dystroglycan; ECM, extracellular matrix; FKTN, Fukutin; FKRP, fukutin-related protein; GlcA-β-MU, 4-methylumbelliferyl-β-D-glucuronide; GlcA-T, glucuronyltransferase; HEK, human embryonic kidney; HPLC, high-performance liquid chromatography; LASV, Lassa virus; MES, 2-(*N*-morpholino)ethanesulfonic acid; MOPS, 3-(*N*-morpholino)propanesulfonic acid; NMR, nuclear magnetic resonance; POMGNT2, protein O-linked mannose N-acetylglucosaminyltransferase 2; POMK, protein O-mannose kinase; POMT, protein *O*-

818 mannosyltransferase; Q/TOF, quadrupole time of flight; WT, wild-type; Xyl-β-MU, 4-

819 methylumbelliferyl-beta-D-xyloside; Xyl-T, xylosyltransferase.

References

1. Barresi R & Campbell KP (2006) Dystroglycan: from biosynthesis to pathogenesis of human disease. *Journal of cell science* 119(Pt 2):199-207.
2. Cao W, *et al.* (1998) Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus. *Science* 282(5396):2079-2081.
3. Kanagawa M, *et al.* (2004) Molecular recognition by LARGE is essential for expression of functional dystroglycan. *Cell* 117(7):953-964.
4. Hara Y, *et al.* (2011) Like-acetylglucosaminyltransferase (LARGE)-dependent modification of dystroglycan at Thr-317/319 is required for laminin binding and arenavirus infection. *Proceedings of the National Academy of Sciences of the United States of America* 108(42):17426-17431.
5. Mercuri E & Muntoni F (2012) The ever-expanding spectrum of congenital muscular dystrophies. *Annals of neurology* 72(1):9-17.
6. Bonnemann CG, *et al.* (2014) Diagnostic approach to the congenital muscular dystrophies. *Neuromuscular disorders : NMD* 24(4):289-311.
7. Jae LT, *et al.* (2013) Deciphering the glycosylome of dystroglycanopathies using haploid screens for lassa virus entry. *Science* 340(6131):479-483.
8. Vuillaumier-Barrot S, *et al.* (2012) Identification of mutations in TMEM5 and ISPD as a cause of severe cobblestone lissencephaly. *Am J Hum Genet* 91(6):1135-1143.
9. Buysse K, *et al.* (2013) Missense mutations in beta-1,3-N-acetylglucosaminyltransferase 1 (B3GNT1) cause Walker-Warburg syndrome. *Human molecular genetics* 22(9):1746-1754.
10. Yoshida-Moriguchi T, *et al.* (2013) SGK196 is a glycosylation-specific O-mannose kinase required for dystroglycan function. *Science* 341(6148):896-899.
11. Inamori K, *et al.* (2012) Dystroglycan function requires xylosyl- and glucuronyltransferase activities of LARGE. *Science* 335(6064):93-96.
12. Goddeeris MM, *et al.* (2013) LARGE glycans on dystroglycan function as a tunable matrix scaffold to prevent dystrophy. *Nature* 503(7474):136-140.
13. Sasaki K, *et al.* (1997) Expression cloning of cDNA encoding a human beta-1,3-N-acetylglucosaminyltransferase that is essential for poly-N-acetyllactosamine synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 94(26):14294-14299.
14. Kobayashi K, *et al.* (1998) An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy. *Nature* 394(6691):388-392.
15. de Bernabe DB, *et al.* (2003) A homozygous nonsense mutation in the fukutin gene causes a Walker-Warburg syndrome phenotype. *Journal of medical genetics* 40(11):845-848.
16. Brockington M, *et al.* (2001) Mutations in the fukutin-related protein gene (FKRP) identify limb girdle muscular dystrophy 2I as a milder allelic variant of congenital muscular dystrophy MDC1C. *Human molecular genetics* 10(25):2851-2859.

17. Beltran-Valero de Bernabe D, *et al.* (2004) Mutations in the FKRP gene can cause muscle-eye-brain disease and Walker-Warburg syndrome. *Journal of medical genetics* 41(5):e61.
18. Shaheen R, Faqeih E, Ansari S, & Alkuraya FS (2013) A truncating mutation in B3GNT1 causes severe Walker-Warburg syndrome. *Neurogenetics* 14(3-4):243-245.
19. Wright KM, *et al.* (2012) Dystroglycan organizes axon guidance cue localization and axonal pathfinding. *Neuron* 76(5):931-944.
20. Linstedt AD & Hauri HP (1993) Giantin, a novel conserved Golgi membrane protein containing a cytoplasmic domain of at least 350 kDa. *Molecular biology of the cell* 4(7):679-693.
21. Esapa CT, *et al.* (2002) Functional requirements for fukutin-related protein in the Golgi apparatus. *Human molecular genetics* 11(26):3319-3331.
22. Xiong H, *et al.* (2006) Molecular interaction between fukutin and POMGnT1 in the glycosylation pathway of alpha-dystroglycan. *Biochem Biophys Res Commun* 350(4):935-941.
23. Brockington M, *et al.* (2005) Localization and functional analysis of the LARGE family of glycosyltransferases: significance for muscular dystrophy. *Human molecular genetics* 14(5):657-665.
24. Kuga A, *et al.* (2012) Absence of post-phosphoryl modification in dystroglycanopathy mouse models and wild-type tissues expressing non-laminin binding form of alpha-dystroglycan. *J Biol Chem* 287(12):9560-9567.
25. Stanley P (1985) Membrane mutants of animal cells: rapid identification of those with a primary defect in glycosylation. *Molecular and cellular biology* 5(5):923-929.
26. Kingsley DM, Kozarsky KF, Hobbie L, & Krieger M (1986) Reversible defects in O-linked glycosylation and LDL receptor expression in a UDP-Gal/UDP-GalNAc 4-epimerase deficient mutant. *Cell* 44(5):749-759.
27. Ashikov A, Buettner FF, Tiemann B, Gerardy-Schahn R, & Bakker H (2012) LARGE2 generates the same xylose and glucuronic acid containing glycan structures as LARGE. *Glycobiology*.
28. Barresi R, *et al.* (2004) LARGE can functionally bypass alpha-dystroglycan glycosylation defects in distinct congenital muscular dystrophies. *Nat Med* 10(7):696-703.
29. Willer T, *et al.* (2012) ISPD loss-of-function mutations disrupt dystroglycan O-mannosylation and cause Walker-Warburg syndrome. *Nat Genet* 44(5):575-580.
30. Wiggins CA & Munro S (1998) Activity of the yeast MNN1 alpha-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proceedings of the National Academy of Sciences of the United States of America* 95(14):7945-7950.
31. Bao X, *et al.* (2009) Tumor suppressor function of laminin-binding alpha-dystroglycan requires a distinct beta3-N-acetylglucosaminyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* 106(29):12109-12114.

32. Lee PL, Kohler JJ, & Pfeffer SR (2009) Association of beta-1,3-N-acetylglucosaminyltransferase 1 and beta-1,4-galactosyltransferase 1, trans-Golgi enzymes involved in coupled poly-N-acetyllactosamine synthesis. *Glycobiology* 19(6):655-664.
33. Lidholt K & Lindahl U (1992) Biosynthesis of heparin. The D-glucuronosyl- and N-acetyl-D-glucosaminyltransferase reactions and their relation to polymer modification. *The Biochemical journal* 287 (Pt 1):21-29.
34. Esko JD, Kimata K, & Lindahl U (2009) Proteoglycans and Sulfated Glycosaminoglycans. *Essentials of Glycobiology*, eds Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, & Etzler MECold Spring Harbor (NY)), 2nd Ed.
35. Williamson RA, *et al.* (1997) Dystroglycan is essential for early embryonic development: disruption of Reichert's membrane in Dag1-null mice. *Human molecular genetics* 6(6):831-841.
36. Willer T, *et al.* (2004) Targeted disruption of the Walker-Warburg syndrome gene Pomt1 in mouse results in embryonic lethality. *Proceedings of the National Academy of Sciences of the United States of America* 101(39):14126-14131.
37. Kurahashi H, *et al.* (2005) Basement membrane fragility underlies embryonic lethality in fukutin-null mice. *Neurobiology of disease* 19(1-2):208-217.
38. Longman C, *et al.* (2003) 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. *Human molecular genetics* 12(21):2853-2861.
39. Meilleur KG, *et al.* (2014) Clinical, pathologic, and mutational spectrum of dystroglycanopathy caused by LARGE mutations. *Journal of neuropathology and experimental neurology* 73(5):425-441.
40. Clarke NF, *et al.* (2011) Congenital muscular dystrophy type 1D (MDC1D) due to a large intragenic insertion/deletion, involving intron 10 of the LARGE gene. *Eur J Hum Genet* 19(4):452-457.
41. Liou LY, *et al.* (2010) Functional glycosylation of dystroglycan is crucial for thymocyte development in the mouse. *PLoS One* 5(3):e9915.
42. de Bernabe DB, *et al.* (2009) Loss of alpha-dystroglycan laminin binding in epithelium-derived cancers is caused by silencing of LARGE. *J Biol Chem* 284(17):11279-11284.
43. Snider MD (2002) Metabolic labeling of glycoproteins with radioactive sugars. *Current protocols in cell biology / editorial board, Juan S. Bonifacino ... [et al.]* Chapter 7:Unit 7 8.
44. Patnaik SK & Stanley P (2006) Lectin-resistant CHO glycosylation mutants. *Methods in enzymology* 416:159-182.
45. Maeda Y, Tomita S, Watanabe R, Ohishi K, & Kinoshita T (1998) DPM2 regulates biosynthesis of dolichol phosphate-mannose in mammalian cells: correct subcellular localization and stabilization of DPM1, and binding of dolichol phosphate. *The EMBO journal* 17(17):4920-4929.

46. Ohyama C, *et al.* (1998) Molecular cloning and expression of GDP-D-mannose-4,6-dehydratase, a key enzyme for fucose metabolism defective in Lec13 cells. *J Biol Chem* 273(23):14582-14587.
47. Bakker H, *et al.* (2009) Functional UDP-xylose transport across the endoplasmic reticulum/Golgi membrane in a Chinese hamster ovary cell mutant defective in UDP-xylose Synthase. *J Biol Chem* 284(4):2576-2583.
48. Xu J (2005) Preparation, culture, and immortalization of mouse embryonic fibroblasts. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* Chapter 28:Unit 28 21.
49. Yoshida-Moriguchi T, *et al.* (2010) O-mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. *Science* 327(5961):88-92.
50. Michele DE, *et al.* (2002) Post-translational disruption of dystroglycan-ligand interactions in congenital muscular dystrophies. *Nature* 418(6896):417-422.
51. Ervasti JM & Campbell KP (1991) Membrane organization of the dystrophin-glycoprotein complex. *Cell* 66(6):1121-1131.
52. Duclos F, *et al.* (1998) Progressive muscular dystrophy in alpha-sarcoglycan-deficient mice. *The Journal of cell biology* 142(6):1461-1471.
53. Inamori K, *et al.* (2013) Xylosyl- and glucuronyltransferase functions of LARGE in alpha-dystroglycan modification are conserved in LARGE2. *Glycobiology* 23(3):295-302.
54. Inamori KI, *et al.* (2014) Endogenous Glucuronyltransferase Activity of LARGE or LARGE2 Required for Functional Modification of alpha-dystroglycan in Cells and Tissues. *J Biol Chem*.
55. Rance M, *et al.* (1983) Improved spectral resolution in cosy 1H NMR spectra of proteins via double quantum filtering. *Biochem Biophys Res Commun* 117(2):479-485.
56. Braunschweiler L & Ernst RR (1983) Coherence Transfer by Isotropic Mixing - Application to Proton Correlation Spectroscopy. *J Magn Reson* 53(3):521-528.
57. Davis DG & Bax A (1985) Separation of Chemical-Exchange and Cross-Relaxation Effects in Two-Dimensional Nmr-Spectroscopy. *J Magn Reson* 64(3):533-535.
58. Nyberg NT, Duus JO, & Sorensen OW (2005) Editing of H2BC NMR spectra. *Magn Reson Chem* 43(12):971-974.
59. Delaglio F, *et al.* (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* 6(3):277-293.
60. Johnson BA & Blevins RA (1994) NMR View: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* 4(5):603-614.
61. Van Reeuwijk J, *et al.* (2010) A homozygous FKRP start codon mutation is associated with Walker-Warburg syndrome, the severe end of the clinical spectrum. *Clinical genetics* 78(3):275-281.
62. Grewal PK, Holzfeind PJ, Bittner RE, & Hewitt JE (2001) Mutant glycosyltransferase and altered glycosylation of alpha-dystroglycan in the myodystrophy mouse. *Nat Genet* 28(2):151-154.

992

Figures

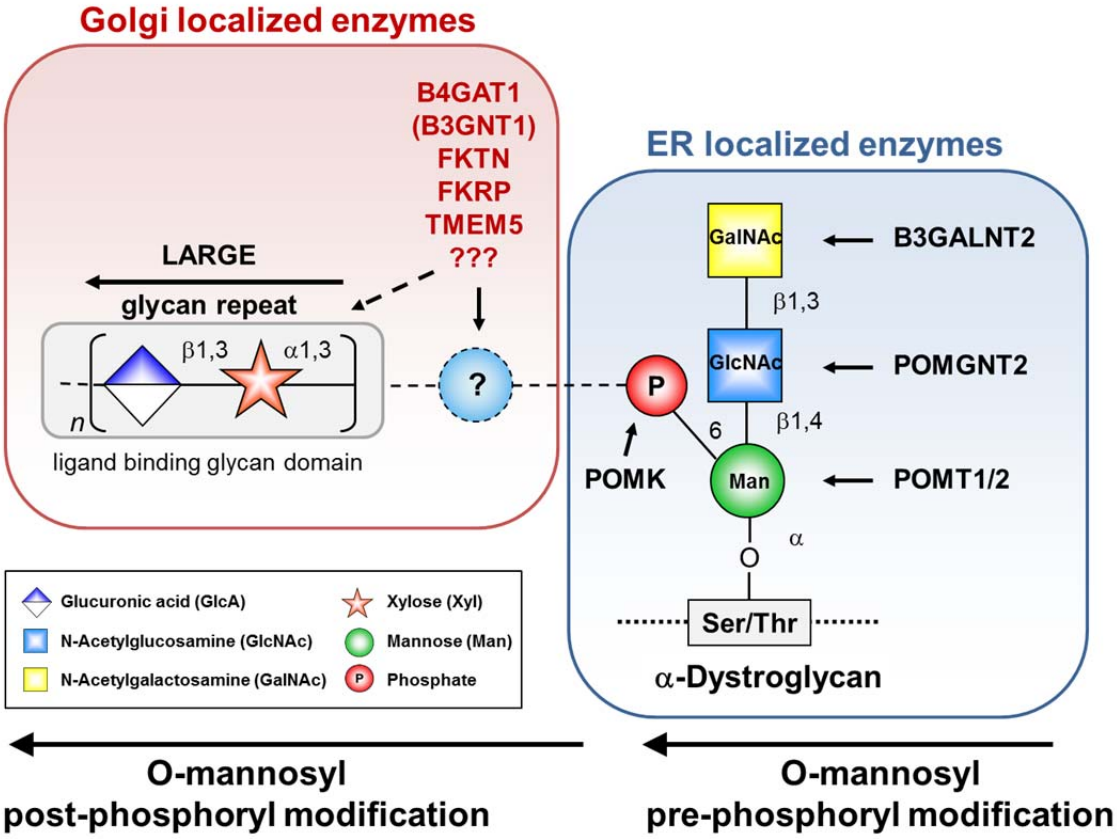
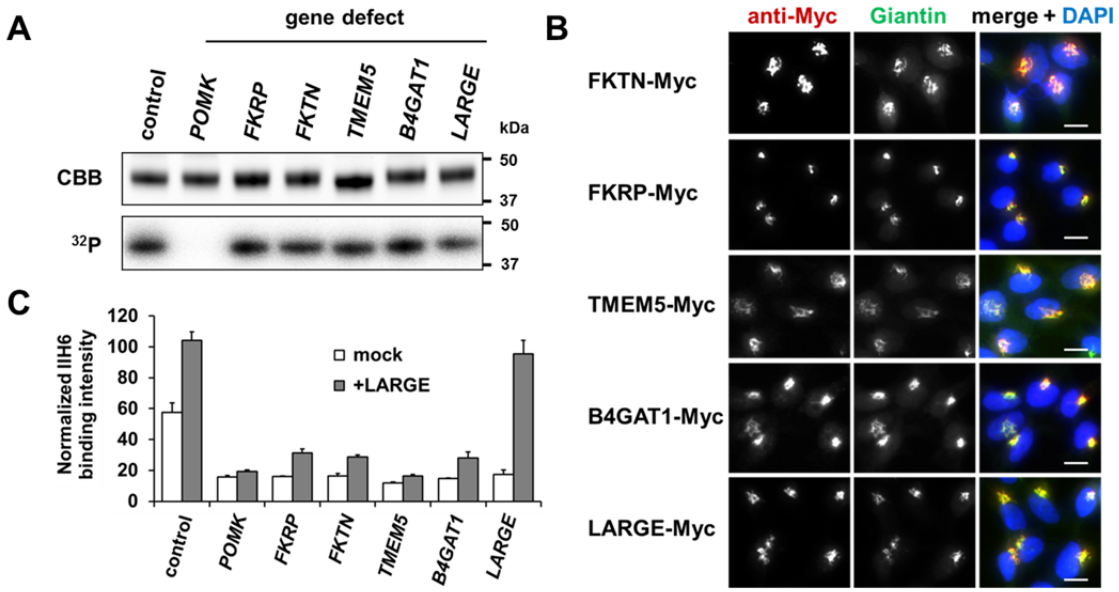


Fig. 1. α-DG functional glycosylation and known proteins contributing to its synthesis. α-DG core M3 functional glycosylation can be divided in 2 major processing steps. O-mannosyl pre-phosphoryl modification which is carried out by enzymes in the endoplasmic reticulum (ER) (highlighted in the blue box) and O-mannosyl post-phosphoryl modification by known or putative glycosyltransferases in the Golgi (highlighted in the red box). Both gene products with known function (black) and gene products with currently unidentified function (red) are indicated. The putative glycosyltransferases B4GAT1 (B3GNT1), FKTN, FKRP and TMEM5 are proposed to act prior to LARGE which adds a GlcA-Xyl heteropolymer that is responsible for ligand binding. However based on current knowledge it cannot be completely ruled out that they are involved in the modification of the LARGE glycan repeat itself to modulate ligand binding.

1007



1008

1009

1010 **Fig. 2. Postulated α-DG modifying enzymes are involved in post-phosphorylation**
1011 **processes in the Golgi prior to LARGE**

1012 (***A***) Phosphorylation of Fc-tagged DGFc340 in the context of α-DG glycosylation defects.
1013 Fc-tagged DGFc340 was produced in [^{32}P] orthophosphate-labeled fibroblasts from control
1014 and glycosylation-deficient patients and mice (**Tab. 1**). The DGFc340 was isolated from
1015 the culture medium by using protein A-agarose and the samples were separated by SDS
1016 PAGE. Gels were stained with Coomassie brilliant blue (CBB) and analyzed by
1017 phosphorimaging (^{32}P).
1018 (*B*) Subcellular localization of α-DG modifying putative glycosyltransferases, as assessed
1019 by immunofluorescence. HEK293T cells stably expressing c-Myc-tagged proteins were
1020 stained with anti-Myc (red), anti-Giantin (Golgi marker, green) and 4´,6-diamidino-2-
1021 phenylindole (DAPI, nuclei, blue). Individual stainings for c-Myc and Giantin are shown in
1022 greyscale and a merged image is shown in color. Scale bars indicate 10µm. (***C***)
1023 Quantitative On-Cell protein blot analysis of LARGE-induced α-DG glycosylation
1024 hyperglycosylation in glycosylation-deficient cells. α-DG glycosylation status was tested

with and without forced LARGE overexpression by adenovirus mediated gene transfer. The On-Cell Western blots were probed with an antibody against the glycosylated form of α-DG (IIH6). IIH6 On-Cell quantitative data were normalized with DRAQ5 cell DNA dyse (n=3). Error bars, s.d.

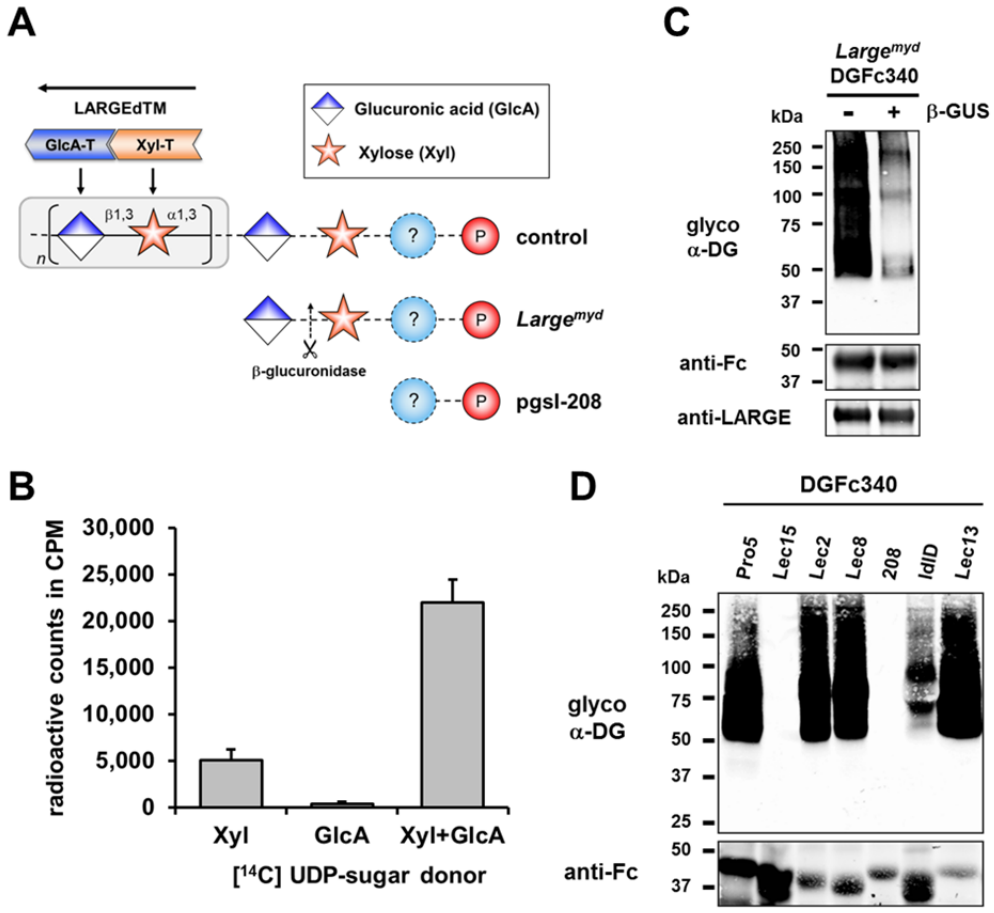


Fig. 3. β-GlcA serves as an acceptor sugar for LARGE modification starting with xylose.

(***A***) Schematic diagram showing the α-DG post-phosphoryl modification in the context of control and glycosylation defects. LARGE adds the ligand-binding glycan that to α-DG via a proposed glucuronic acid (GlcA) acceptor. LARGEdTM catalytic domains Xyl-T (orange) and GlcA-T (blue) are highlighted in color. Depicted are also the hypothesized terminal sugar structures of glycosylation-deficient cell lines *Large*^{myd} (*Large*-deficient) and pgsI-208 (UDP-xylose deficient). Cleavage of terminal β-GlcA by exoglycosidase β-glucuronidase (β-GUS) in *Large*^{myd} is indicated (scissor symbol). (***B***) Transfer of [^{14}C] radiolabeled Xyl and GlcA to DGFc340 by LARGEdTM. Fc-tagged DGFc340 was

produced in *Large*^{myd} (*Large*-deficient) MEF cells and isolated from the culture medium using protein A-agarose. The protein A-bound DGFc340 was used as acceptor in a LARGEdTM reaction with radiolabeled [^{14}C] UDP-Xyl and/or [^{14}C] UDP-GlcA sugar donors. The figure represents the transfer of radiolabeled saccharides onto the donor DGFc340 (n=3). Error bars represent s.d. (***C***) β-Glucuronidase pre-treatment of DGFc340 from *Large*^{myd} deficient cells impairs LARGEdTM modification. Protein A-bound DGFc340 (acceptor) isolated from transfected *Large*^{myd} MEFs was digested with β-glucuronidase (β-GUS) prior to the LARGEdTM (enzyme) reaction, which included UDP-Xyl and UDP-GlcA as sugar (donors). After incubation with LARGEdTM DGFc340 (acceptor protein) was subjected to protein blotting with antibodies against the glycosylated form of α-DG (IIH6), against Fc and against LARGE (Rb331).
(***D***) The ability of LARGEdTM to modify DGFc340 is impaired in the context of sugar donor-deficient CHO mutant cell lines. Fc-tagged DGFc340 was produced in various glycosylation-deficient Lec CHO cells and isolated from the culture medium using protein A-agarose. As in (***C***) protein A-bound DGFc340 acceptor was used in a LARGEdTM reaction and analyzed by protein blotting.

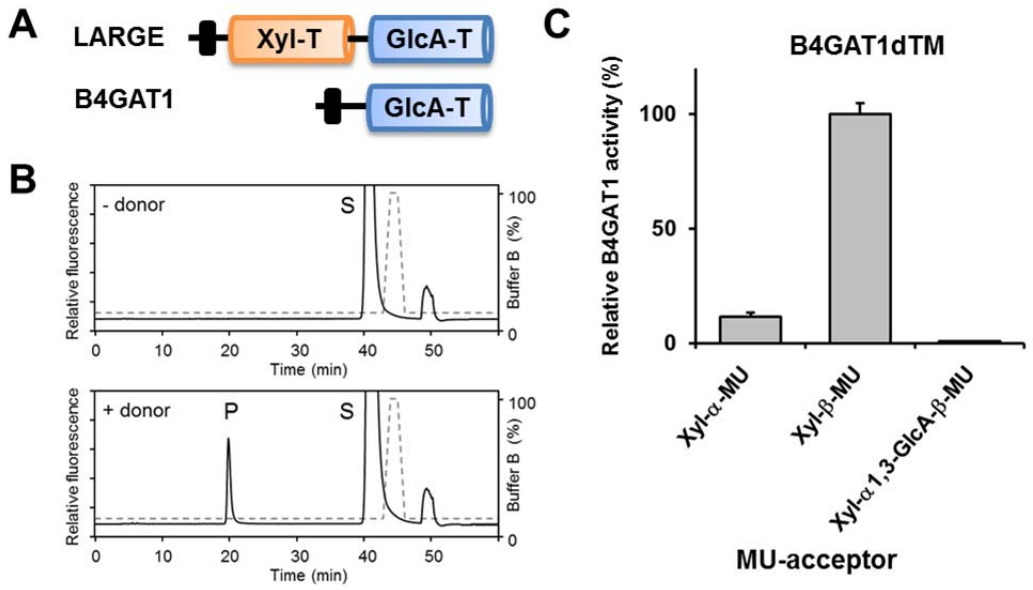


Fig. 4. B4GAT1 has xylose β1,4 glucuronyltransferase activity.

(***A***) Schematic representation of LARGE and B4GAT1 functional domains. GlcA-T (blue), Xyl-T (orange) and transmembrane domain (black) are indicated.

(***B***) Representative HPLC profiles of the reaction product generated in the absence (top) and presence (bottom) of a UDP-GlcA sugar (donor) in a reaction mix containing Xyl-β-MU (acceptor) and B4GAT1dTM (enzyme). Samples were separated on an LC-18 column. P, product. S, unreacted substrate. Dotted line, %B buffer.

(***C***) Comparison of B4GAT1dTM GlcA-T activity with respect to various xylose-MU acceptor sugars. Relative activity (%) with respect to Xyl-β-MU acceptor (specific activity: 0.2 μmol/h/mg) is shown (n=3). Error bars represent s.d.

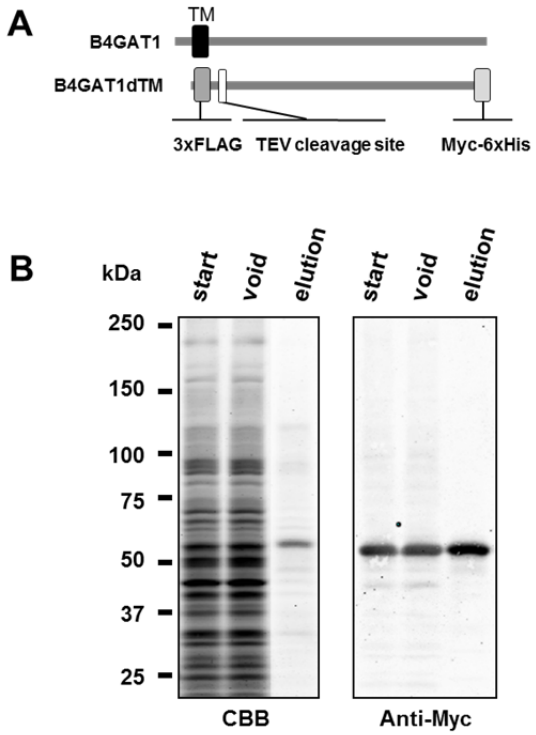


Fig. 5. Purification of B4GAT1dTM.

(***A***) Schematic representation of B4GAT1 and the B4GAT1dTM construct used in the enzymatic activity assay. The transmembrane (TM) sequence was replaced with a 3xFLAG-TEV tag sequence and the C-terminus was modified with a Myc-6xHis-tag. (***B***) Purification of recombinant B4GAT1dTM from bioreactor culture medium. The recombinant protein was expressed in HEK293T cells and purified from the culture medium using Talon metal-affinity resin. The bioreactor medium samples before (start) and after the purification (void) as well as the eluted purified protein (elution) were analyzed by immunoblotting with anti-Myc (4A6) antibody. CBB, stained with Coomassie Brilliant blue.

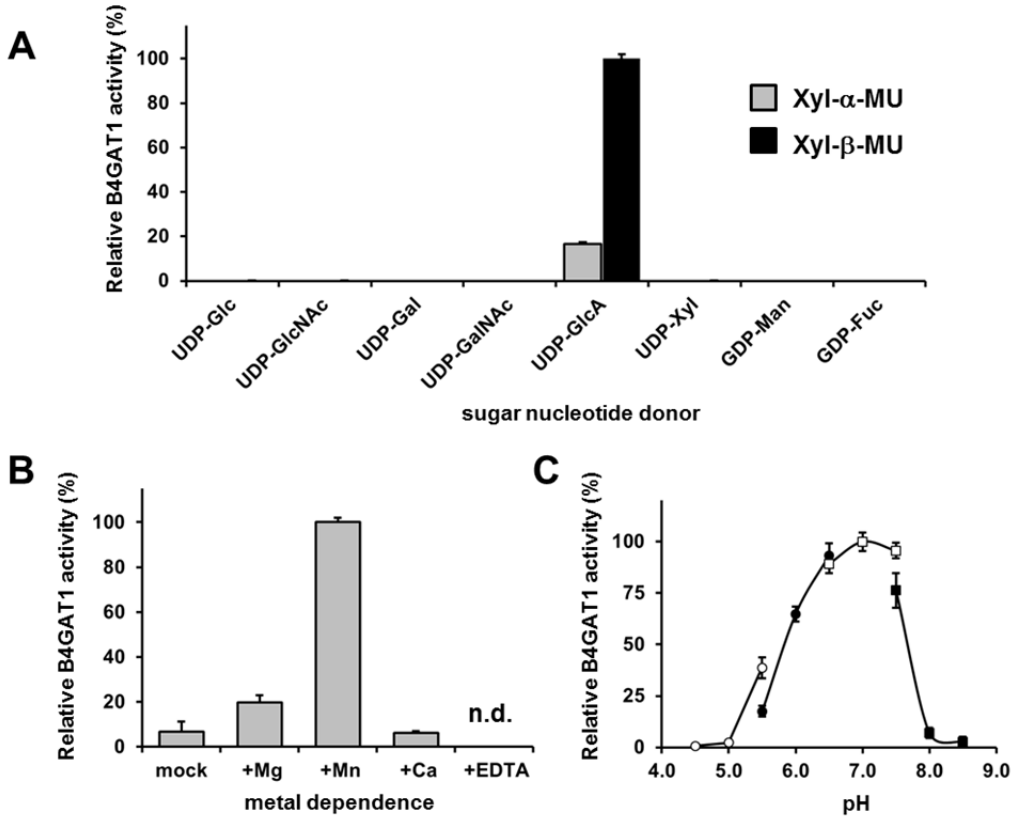


Fig. 6. Basic characterization of the xylose β1,4-glucuronyltransferase activity of B4GAT1

(***A***) Donor sugar specificity of B4GAT1dTM. Representative data from two independent assays, demonstrating relative activity (%) of B4GAT1dTM (enzyme) GlcA-T toward Xyl-α-MU and Xyl-β-MU (acceptor) when tested with various sugar nucleotides (donor). The specific activity set as 100% for acceptor Xyl-β-MU was 0.2 μmol/h/mg. No sugar other than GlcA was transferred to the acceptors to a significant extent.

(***B***) Metal dependence of the B4GAT1dTM GlcA-T activity. Activity assay was carried out in the presence or absence of each metal ion or EDTA (10 mM), and results are shown as relative activity (%). The GlcA-T activity in the presence of Mn^{2+} (specific activity: 0.26 μmol/h/mg) was arbitrarily set at 100% (n=3). n.d., not detected. Error bars represent s.d.

(***C***) pH optimum of B4GAT1dTM GlcA-T activity. Data from three independent experiments are shown as relative activity (%). The highest activity (specific activity: 0.22

μmol/h/mg) in the dataset was arbitrarily set at 100%. The GlcA-T assays were carried out using Xyl-β-MU as acceptor. The buffers used were: acetate for pH 4.5–5.5 (open circle), MES for pH 5.5–6.5 (closed circle), MOPS for pH 6.5–7.5 (open square) and Tris–HCl for pH 7.5–8.5 (closed square). The details of the conditions are presented in the Materials and methods section. Error bars represent s.d.

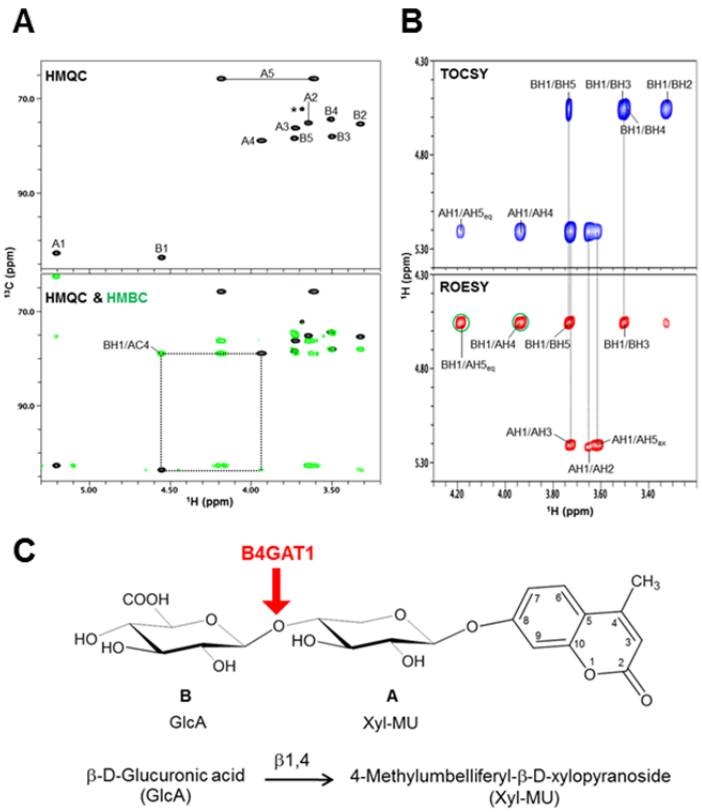


Fig. 7. NMR analysis reveals that B4GAT1 is a β1,4 glucuronyltransferase.

(***A***) HMQC spectrum (top) and overlay of HMQC (black) and HMBC (green) spectra (bottom) for the B4GAT1 enzymatic reaction product. The cross-peaks are labeled with a first letter representing the subunit designated in ***C*** and the rest of the label representing the position on that subunit. The observed interglycosidic cross-peak BH1/AC4 in the HMBC spectrum clearly demonstrates the presence of a 1→4 interglycosidic linkage between the residues B and A. The cross-peak marked with a star represents an impurity.

(***B***) TOCSY (top) and ROESY (bottom) spectra of the B4GAT1 enzymatic reaction product collected with a mixing time of 77 and 300 ms, respectively. The cross-peaks are labeled as in ***A***. The observed interglycosidic ROEs are indicated in green circles. The ROE data indicate that both residues exist in β-configurations.

(***C***) Schematic depiction of the disaccharide structure produced by B4GAT1, with the sugar units labeled A and B.

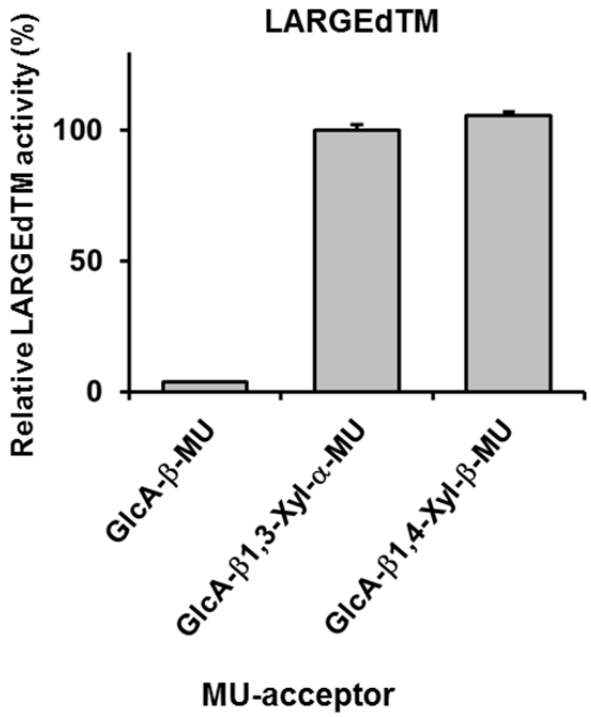


Fig. 8. Substrate specificity of LARGEdTM Xyl-T enzyme activity.

Comparison of LARGEdTM Xyl-T activity with respect to various monosaccharide and disaccharide GlcA-MU acceptor sugars. Relative activity (%) with respect to intrinsic LARGE polymer specific activity with GlcA-β1,3-Xyl-α-MU disaccharide acceptor (0.08 μmol/h/mg) (n=3). Error bars represent s.d.

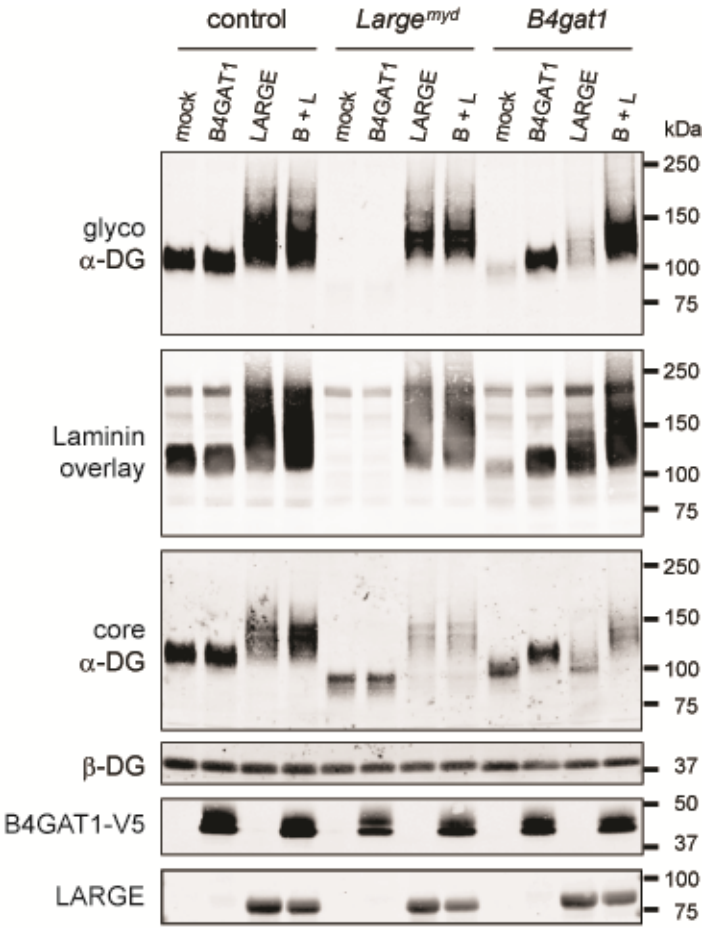


Fig. 9. *B4gat1*-deficient MEFs have impaired α-DG functional glycosylation

Functional glycosylation and complementation analysis of α-DG in wild-type, *Large*- and *B4gat1*-deficient MEFs. Immunoblots and laminin overlay assay of WGA-enriched cell lysates extracted from WT, *Large*^{myd} (*Large*-deficient) and *B4gat1*-deficient MEFs . As indicated MEFs were uninfected (mock) or infected with adenovirus constructs expressing B4GAT1, LARGE or both (B+L). Antibodies used were: glyco α-DG (IIH6), core α-DG, core β-DG (AP83), anti-V5 and anti-LARGE (Rb331).

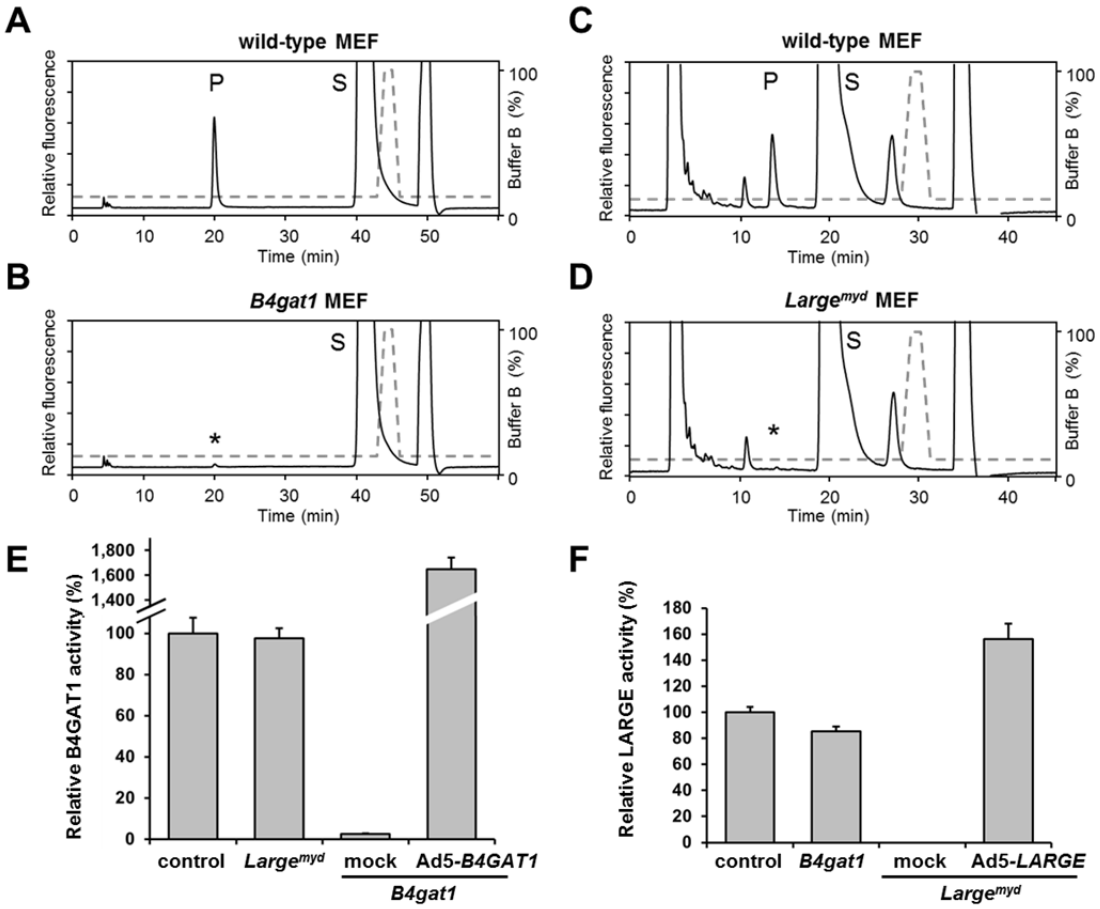


Fig. 10. *B4gat1*-deficient MEFs have impaired endogenous B4GAT1 activity

Representative HPLC profiles of the reaction product are shown. (***A/B***) Endogenous B4GAT1 enzyme activity of cell lysates from wild-type (***A***) and *B4gat1*-deficient (***B***) MEFs . *B4gat1*-deficient MEFs ($B4gat1^{LacZ/M155T}$) show some residual activity < 3% (asterisk). (***C/D***) Endogenous LARGE enzyme activity of WGA-enriched cell lysates from control (***C***) and $Large^{myd}$ (***D***) (*Large*-deficient) MEFs. $Large^{myd}$ MEFs lack LARGE activity and do not show any residual activity (asterisk). The details of the conditions are provided in the Materials and methods section. (***E***) Comparison of endogenous B4GAT1 GlcA-T activity in control, *Large*- and *B4gat1*-deficient MEFs. Additionally, *B4gat1*-deficient MEFs ($B4gat1^{LacZ/M155T}$) complemented with control *B4GAT1* expressing adenovirus (Ad5) were tested. Cell lysates were used as enzyme

source to measure endogenous B4GAT1 activity. Relative activity (%) with respect to control MEFs specific activity (91.6 pmol/h/mg) is shown (n=3). Error bars represent s.d. (***F***) Comparison of endogenous LARGE GlcA-T activity in control, *Large*- and *B4gat1*-deficient MEFs. Additionally, *Large*-deficient MEFs complemented with control *LARGE* expressing adenovirus (Ad5) were tested. WGA enriched glycoprotein samples were used as enzyme source to measure endogenous LARGE activity. Relative activity (%) with respect to control MEFs specific activity (0.52 pmol/h/mg) is shown (n=3). Error bars represent s.d.

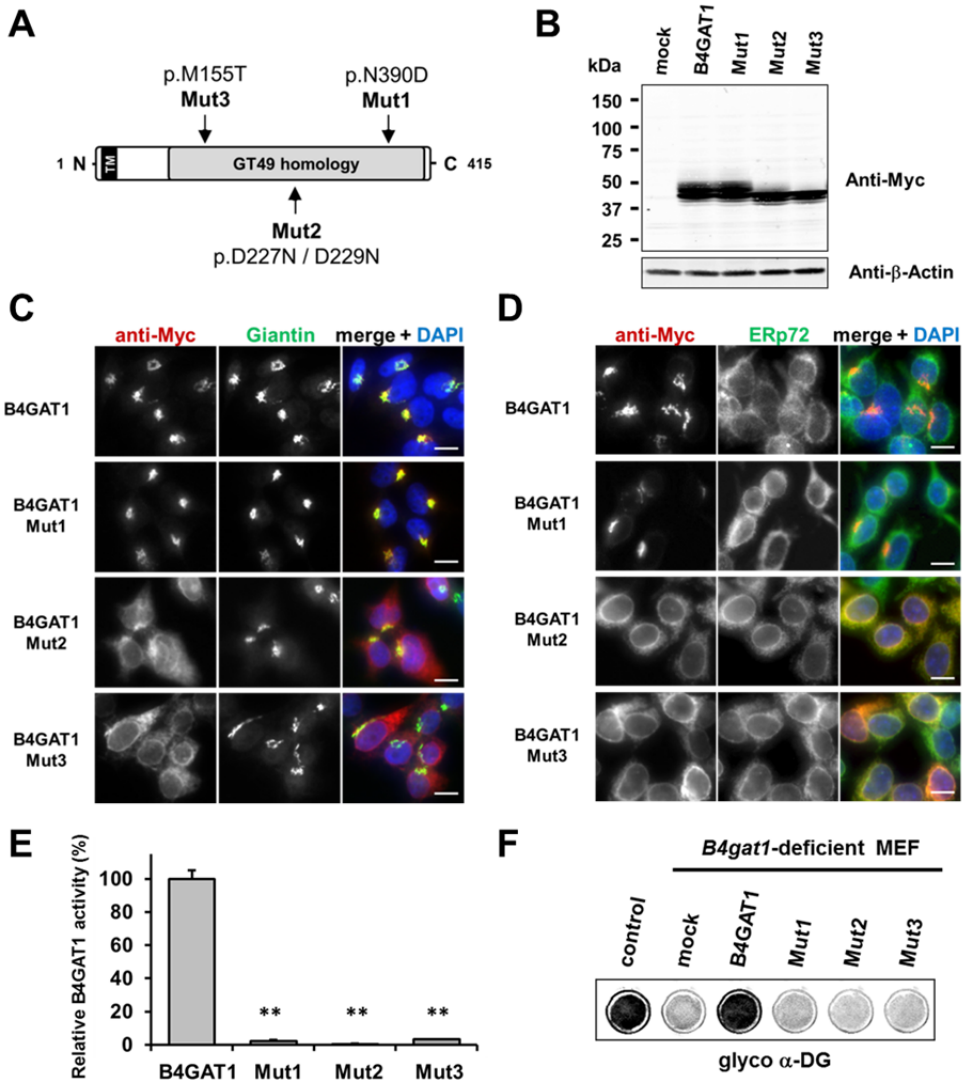


Fig. 11. Expression analysis and GlcA-T enzyme activity of B4GAT1 mutant constructs

(***A***) Schematic presentation shows B4GAT1 enzyme product with functional domains and *B4gat1* mutations Mut1-Mut3 are indicated. (***B***) Expression analysis of B4GAT1-Myc control and mutant constructs in stable HEK293T cells. Immunoblotting of cell lysates from HEK293T cells stably overexpressing wild-type B4GALT1-Myc and mutant constructs (Mut1, Mut2 and Mut3) with anti-Myc antibody and β-Actin (loading control). (***C/D***) Subcellular localization of B4GAT1-Myc control and mutant constructs in stable

HEK293T cells (see ***B***). B4GALT1-Myc constructs were stained with anti-Myc (red), (*C*) anti-Giantin (Golgi marker, green), (*D*) anti-ERp72 (ER marker, green) and 4´,6-diamidino-2-phenylindole (DAPI, nuclei, blue). Individual stainings for c-Myc Giantin and ERp72 are shown in greyscale, and merged images is shown in color. Scale bars indicate 10µm. (***E)*** B4GAT1 enzyme activity in cell lysates from stable HEK293T cells overexpressing B4GALT1-Myc wild-type and B4GALT1-Myc mutant constructs (Mut1-Mut3). Relative activity (%) with respect to B4GALT1wild-type specific activity (19.8 nmol/h/mg) is shown (n=3). Error bars represent s.d., Statistical analyses were performed by two-tail Student's *t* test. **P<0.001. (***F***) Complementation of *B4gat1*-deficient (*B4gat1*$^{LacZ/M155T}$) MEF cells with B4GAT1-Myc control and mutant constructs. *B4gat1*-deficient MEFs were nucleofected with a wild-type or mutant B4GAT1 expression construct. α-DG functional glycosylation was analyzed by On-Cell-Western analysis. α-DG functional glycosylation was detected with α-DG glyco (IIH6) antibody.

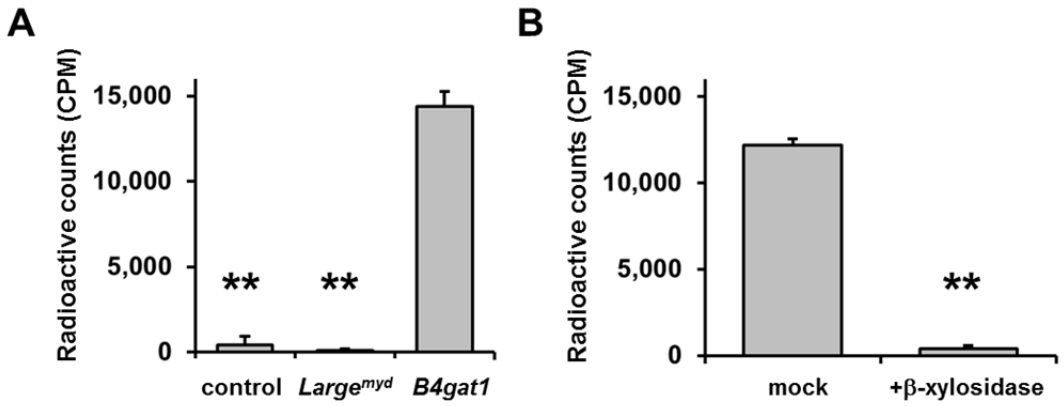


Fig. 12. β-xylose is the endogenous acceptor for B4GAT1.

(***A***) B4GAT1dTM enzymatic transfer of [^{14}C] radiolabeled GlcA to DGFc340. Fc-tagged DGFc340 (acceptor) was produced in control, *Large*^{myd} (*Large*-deficient) and *B4gat1*-deficient MEFs and isolated from the culture medium using protein A-agarose. The protein A-bound Fc340 was used as acceptor in a B4GAT1dTM (enzyme) reaction with radiolabeled [^{14}C] UDP-GlcA sugar (donor). The figure represents the transfer of radiolabeled GlcA onto the donor DGFc340 (n=3). Error bars represent s.d. Statistical analyses were performed by two-tail Student's *t* test. **P<0.001. (***B***) β-Xylosidase pre-treatment impairs B3GAT1dTM transfer of [^{14}C] radiolabeled GlcA. DGFc340 (acceptor) from *B4gat1*-deficient MEFs was digested with β-xylosidase prior to the B4GAT1dTM (enzyme) transfer reaction with [^{14}C] UDP-GlcA sugar (donor). The figure represents the transfer of radiolabeled GlcA onto the donor DGFc340 (n=3). Error bars represent s.d. Statistical analyses were performed by two-tail Student's *t* test. **P<0.001.

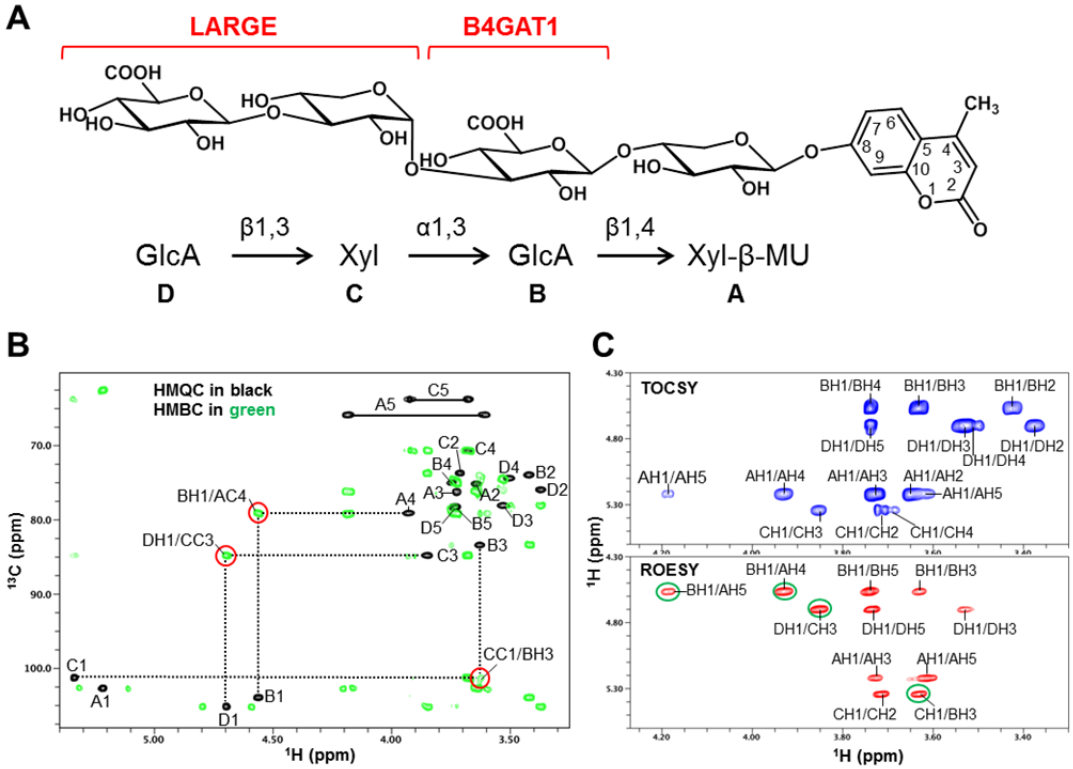


Fig. 13. NMR analyses of the tetrasacharide generated from B4GAT1dTM and LARGEdTM enzymatic reactions.

(***A***) Schematic depiction of the tetrasaccharide structure produced by the sequential reactions of B4GAT1dTM followed by LARGEdTM with the sugar units labeled A-D to indicate the order of their addition. (***B***) Overlay of the HMQC (black) and HMBC (green) spectra of the tetrasaccharide. All cross-peaks in the HMQC spectrum are labeled. Three interglycosidic cross-peaks detected in the HMBC spectrum are also labeled and indicated with red circles. The peaks are labeled with a first letter representing the subunit designated in ***A,*** and the rest of the label representing the position on that subunit.(***C***) TOCSY spectrum (top) and ROESY spectrum (bottom) of the tetrasaccharide. The TOCSY and ROESY spectra were collected with mixing time of 77 and 300 ms, respectively. The cross-peaks are labeled as in ***B***. The observed interglycosidic ROEs are indicated with green circles.

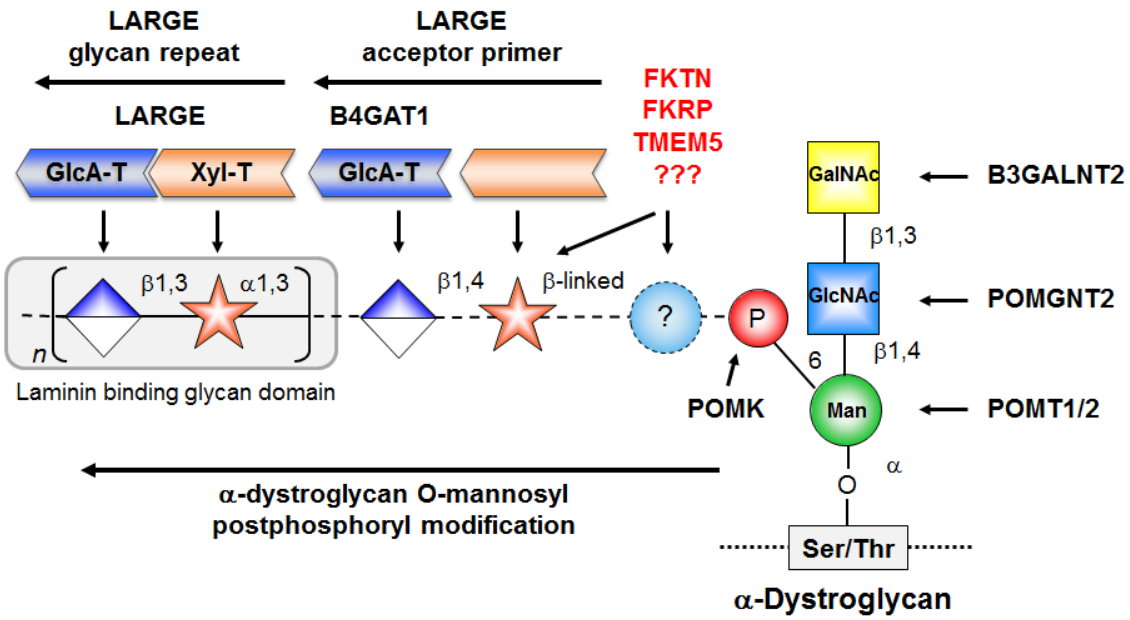


Fig. 14. Model of proposed α-DG O-mannosyl laminin-binding glycan structure and the enzymes that contribute to its synthesis.

Post-phosphoryl modification of α-DG requires B4GAT1 (β1,4 glucuronyltransferase); this enzyme generates the acceptor glycan, which serves as a primer for the glycosyltransferase LARGE to initiate synthesis of the laminin-binding glycan. Both gene products with known function (black) and gene products with currently unidentified function (red) are indicated.

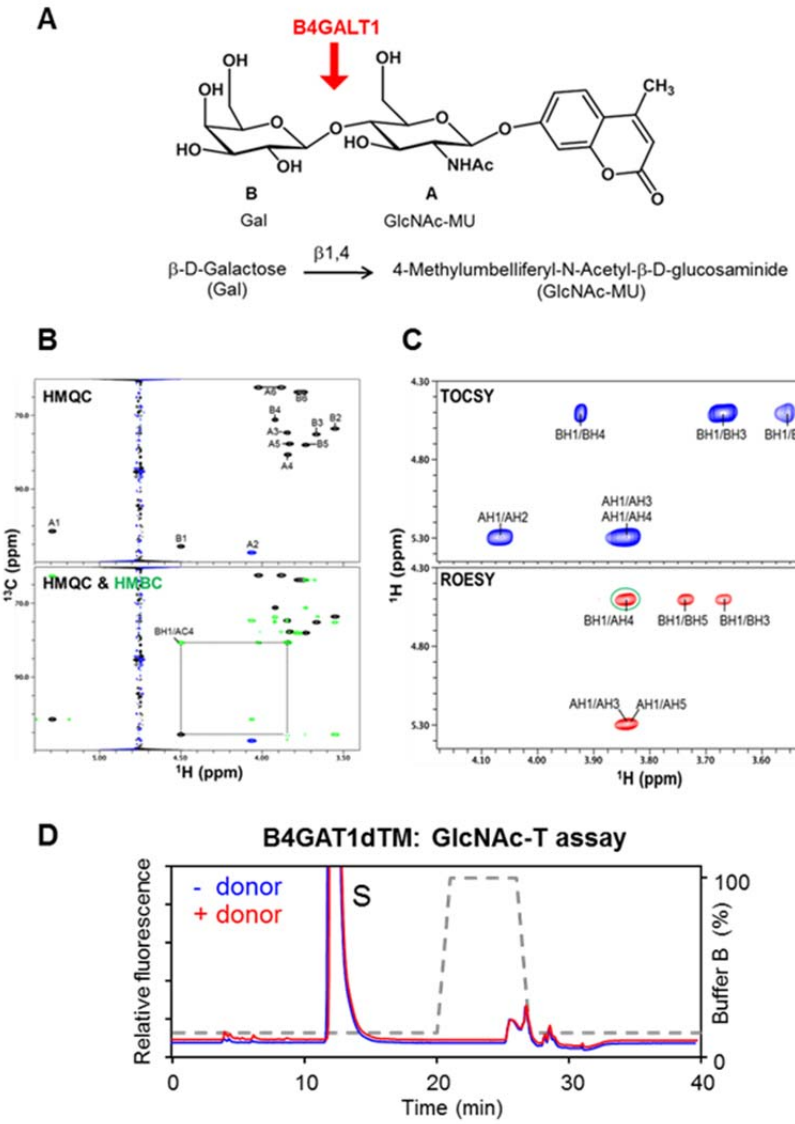


Fig. 15. Test B4GAT1 for GlcNAc transferase activity with iGnT substrate Gal-β1,4-GlcNAc-β-MU

(***A***) Using B4GALT1 we synthesized the hypothesized iGnT substrate Gal-β1,4-GlcNAc-β-MU by transferring a β1,4 Galactose to the acceptor GlcNAc-β-MU. The purified Gal-β1,4-GlcNAc-β-MU disaccharide was further analyzed by NMR. (***B***) HMQC spectrum. The folded peak is shown in blue. The cross peaks are labeled with a first letter representing the subunit as designated in the structure shown above the spectra and the rest of the label representing the position on that subunit. Overlay of HMQC (black and blue)

and HMBC (green) spectra. The strong cross peak labeled as BH1/AC4 was detected in the HMBC spectrum, demonstrating the presence of a 1→4 interglycosidic linkage between Gal and GlcNAc. (***C***) TOCSY spectrum collected with a mixing time of 77 ms. ROESY spectrum collected with a mixing time of 300 ms. The observed strong NOEs from BH1 to BH3 and BH5 (cross peaks labeled as BH1/BH3 and BH1/BH5) demonstrate that the Gal has a β-configuration. Similarly, the observed strong NOEs from AH1 to AH3 and AH5 (cross peaks labeled as AH1/AH3 and AH1/AH5) demonstrate that the GlcNAc has a β-configuration. A strong interglycosidic NOE was observed between BH1 and AH4 (green circle), which is consistent with the 1 to 4 linkage as determined from the HMBC spectrum. (***D***) Representative HPLC profiles of the reaction product generated in the absence (blue) and presence (red) of a UDP-GlcNAc sugar (donor) in a reaction mix containing Gal-β1,4-GlcNAc-β-MU (acceptor) and B4GAT1dTM (enzyme). Samples were separated on an LC-18 column. S, unreacted substrate. Dotted line, %B buffer.

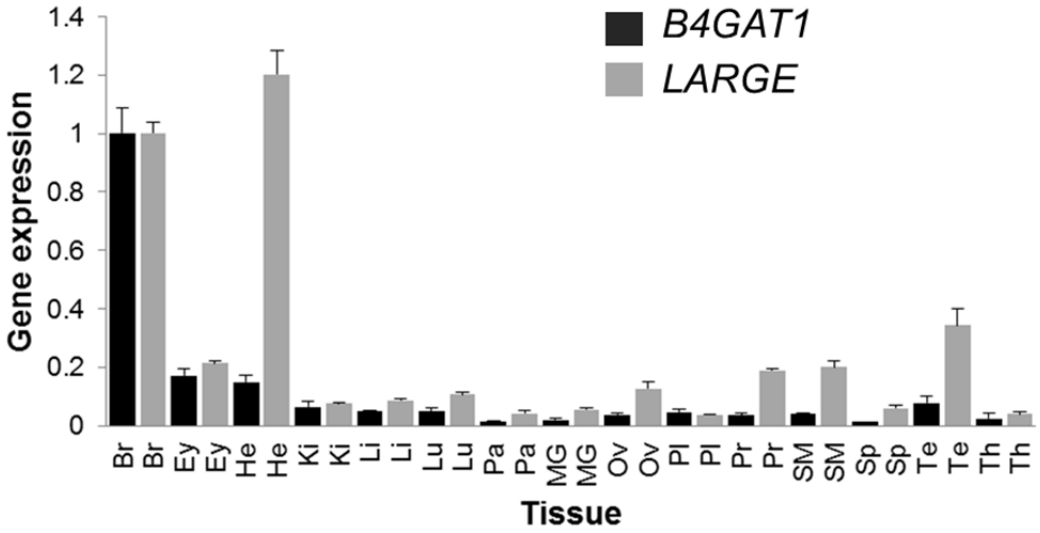


Fig. 16. *B4GAT1 and LARGE* expression in human tissues

qPCR revealing ubiquitous *B4GAT1* and *LARGE* expression in all tissues analyzed, with highest expression of *LARGE* in brain and heart. cDNA was synthesized using random primers and oligo(dT) on commercially available human tissues RNAs. For each tissue, *B4GAT1* and *LARGE* were specifically amplified, in triplicate, in the presence of SYBRgreen, and their expressions was normalized to that of the 28S RNA (normalization control). The expression in each tissue is referenced with respect to that in brain. Analyzed tissues: Br (brain), Ey (eye), He (heart), Ki (kidney), Li (liver), Lu (lung), Pa (pancreas), MG (mammary gland), Ov (ovary), Pl (placenta), Pr (prostate), SM (skeletal muscle), Sp (spleen), Te (testis), Th (thymus). Error bars represent s.d.

Table 1. Summary of features of control and glycosylation-deficient cell lines

Mutant gene	Clinical phenotype	Cell type	Nucleotide variant	Amino acid	reference
Control (human)	none	Human skin fibroblast			CRL-2127 (ATCC)
POMK	WWS/MEB	Human skin fibroblast	14bp homozygous deletion (c.720_733delGCTGGTG AGTGCG)], homozygous	p.Leu241Profs*26	(10)
FKTN	WWS	Human skin fibroblast	c.385delA c.1176C>A, heterozygous	p.I129fsX1 p.Y392X	GM16192 (Coriell Cell Repository)
FKRP	WWS	Human skin fibroblast	c.1A>G, homozygous	p.M1V	(61)
TMEM5	WWS	Human skin fibroblast	c.1101 G>A, homozygous	p.G333R	unpublished
B4gat1 (B3gnt1)	CMD	MEF	c.464 T>C, compound het with LacZ null allele, *B4gat1*$^{LacZ/M155T}$	p.M155T	(19)
***Large*^{myd}**	CMD	MEF	deletion of exons 5-7, homozygous		(62)
Control (mouse)	none	MEF			

Table 2. Chemical shifts (ppm) of the signals in the ^{1}H and ^{13}C NMR spectra of the enzymatic reaction product of GlcA-β1,4-Xyl-β-MU of the glycosyltransferase B4GAT1.

	$^{1}H/^{13}C$ (ppm)[a]										
Products	Sugar						Aromatic Ring[b]				
	1	2	3	4	5	6	3	4-CH_3	6	7	9
→4)-β-D-Xyl-MU	5.21	3.64	3.72	3.93	4.18, 3.61		6.28	2.45	7.76	7.13	7.12
A	102.7	75.2	76.3	78.9	65.8		114.0	20.6	129.4	116.6	106.3
β-D-GlcA-(1→	4.55	3.32	3.50	3.51	3.73						
B	103.6	75.4	78.1	74.4	78.3	178.5[b]					

[a] Chemical shifts at 25°C in 10 mM sodium phosphate, pH 6.5.
[b] Assigned based on the overlay of HMQC and HMBC spectra.

Table 3. Chemical shifts (ppm) of the signals in the ^{1}H and ^{13}C NMR spectra of the tetrasaccharide of GlcA-β1,3-Xyl-α1,3-GlcA-β1,4-Xyl-β-MU produced by B4GAT1 and LARGE.

Products	$^{1}H/^{13}C$ (ppm)[a]										
	Sugar						Aromatic Ring[b]				
	1	2	3	4	5	6	3	4-CH_3	6	7	9
→4)-β-D-Xyl-MU	5.22	3.64	3.73	3.93	4.18, 3.60		6.30	2.47	7.78	7.14	7.15
A	102.7	75.2	76.2	79.1	65.8		114.0	20.6	129.4	116.5	106.4
→3)-β-D-GlcA-(1→	4.56	3.42	3.63	3.74	3.73						
B	103.9	74.0	83.4	75.0	78.3	178.5[b]					
→3)-α-D-Xyl-(1→	5.34	3.71	3.85	3.68	3.68, 3.92						
C	101.3	73.7	84.8	70.8	63.8						
β-D-GlcA-(1→	4.69	3.37	3.53	3.51	3.74						
D	105.2	76.0	78.1	74.4	78.3	178.5[b]					

[a] Chemical shifts at 25°C in 10 mM sodium phosphate, pH 6.5.
[b] Assigned based on the overlay of HMQC and HMBC spectra.