O-Mannosyl Phosphorylation of Alpha-Dystroglycan Is Required for Laminin Binding

Takako Yoshida-Moriguchi,[1,2,3,4] Liping Yu,[5] Stephanie H. Stalnaker,[6] Sarah Davis,[1,2,3,4] Stefan Kunz,[7] Michael Madson,[8] Michael B. A. Oldstone,[9] Harry Schachter,[10] Lance Wells,[6] Kevin P. Campbell[1,2,3,4]*

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

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

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

To determine which posttranslational modification is necessary for the α-DG/laminin interaction, we processed wheat germ agglutinin-enriched proteins (glycoproteins) from C57BL/6J (wild-type, WT) muscle using various enzymatic and chemical treatments. Treatment with cold aqueous hydrofluoric acid (HFaq), which specifically cleaves phosphoester linkages (*18*), resulted in the reduction of the α-DG relative molecular mass (M_r) from 150 to 70 kD, the loss of IIH6 immunoreactivity and laminin binding (Fig. 1A), and the loss of binding to LFV and LCMV (Fig. 1B). Because the M_r of *N*-glycosylated β-DG did not change (Fig. 1A), these effects were not caused by the degradation of either peptide or glycosyl linkages. A quantitative solid-phase assay revealed a 97% reduction in total high-affinity binding to laminin (Fig. 1C). HFaq treatment also abolished the laminin-receptor activity of α-DG in the heart, brain, and kidney (fig. S1A). We next tested whether *N*-glycan and/or the two *O*-glycans known to modify the laminin-binding form of α-DG—Core1 *O*-glycan and the *O*-mannosyl tetrasaccharide (in either the sialylated or fucosylated form) (*15*, *16*)—are sensitive to HFaq treatment. Immunoblotting of WT muscle glycoproteins treated with several cocktails of glycosidases that degrade these three glycans showed that the glycosidase-mediated reduction in α-DG glycosylation was impervious to HFaq treatment (Fig. 1D). A similar experiment using muscle glycoproteins from the CMD-model mouse *LARGE*^{myd}, in which a mutation in LARGE prevents the α-DG modification necessary for laminin-binding (*19*), revealed that HFaq treatment did not significantly reduce the M_r of α-DG (Fig. 1D). Thus, HFaq specifically degrades the laminin-binding moiety on α-DG. Further, functional modification of α-DG appeared to involve an internal phosphoryl linkage rather than a monoester-linked phosphate, because digesting WT muscle glycoproteins with alkaline phosphatase did not reduce the laminin-binding ability (fig. S1B).

To verify that α-DG is phosphorylated, we labeled human embryonic kidney (HEK293) cells expressing Fc-tagged α-DG recombinants (Fig. 2A) that are secreted into the medium with [^{32}P]-orthophosphate. Phosphorimaging showed that secreted DGFc4, which contains only the mucin-like region of α-DG (*20*), was phosphorylated (Fig. 2B). Hydrolysis of [^{32}P]-DGFc4 under conditions that are conducive to the dissolution of polypeptide and phosphoester linkages to carbohydrates, but not to linkages to amino acids (*21*), generated inorganic phosphate but not phospho–amino acids (fig. S2), suggesting that phosphorylation does not occur directly on the peptide. To test whether the phosphorylation depends on glycosylation, we expressed DGFc5 in [^{32}P]-orthophosphate–labeled human cells derived from *POMT1*-mutated WWS, *POMGnT1*-mutated MEB, *fukutin*-mutated FCMD, and control cells, as well as in fibroblasts from *Large*^{myd} and WT mice (Fig. 2C). All except the *POMT1*-mutated WWS cells secreted [^{32}P]-phosphorylated DGFc5 into the medium, strongly suggesting that phosphorylation occurs on the *O*-linked mannose

[1]Howard Hughes Medical Institute, University of Iowa Roy J. and Lucille A. Carver College of Medicine, 4283 Carver Biomedical Research Building, 285 Newton Road, Iowa City, IA 52242-1101, USA. [2]Department of Molecular Physiology and Biophysics, University of Iowa Roy J. and Lucille A. Carver College of Medicine, 4283 Carver Biomedical Research Building, 285 Newton Road, Iowa City, IA 52242-1101, USA. [3]Department of Neurology, University of Iowa Roy J. and Lucille A. Carver College of Medicine, 4283 Carver Biomedical Research Building, 285 Newton Road, Iowa City, IA 52242-1101, USA. [4]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 52242-1101, USA. [5]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 52242-1101, USA. [6]Complex Carbohydrate Research Center, University of Georgia, 315 Riverbend Road, Athens, GA 30602, USA. [7]Institute of Microbiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland. [8]Bio Logistics, 2416 North Shore Drive, Clear Lake, IA 50428, USA. [9]The Scripps Research Institute, Department of Immunology and Microbial Science, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. [10]The Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada.

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

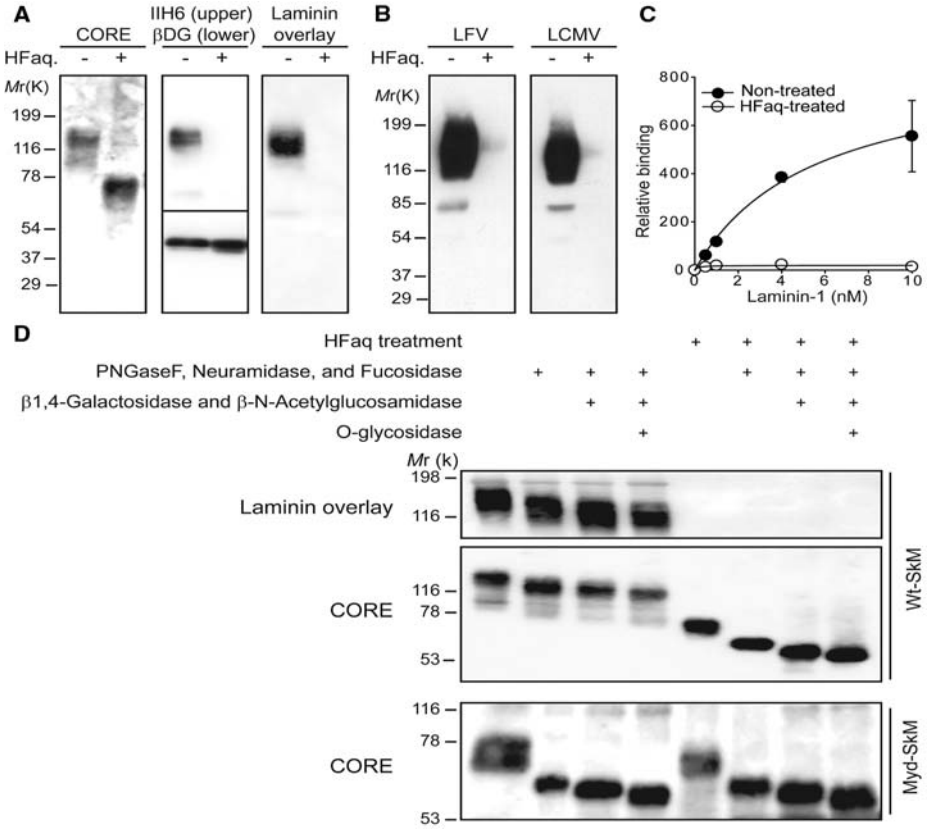

Fig. 1. Chemical dephosphorylation by HFaq treatment abolishes laminin and virus binding to α-DG in tissues from WT mice. (**A** and **B**) Treated glycoproteins prepared from WT muscle were subjected to the following: (A) immunoblotting with antibodies against the α-DG core protein (CORE) or the laminin-binding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

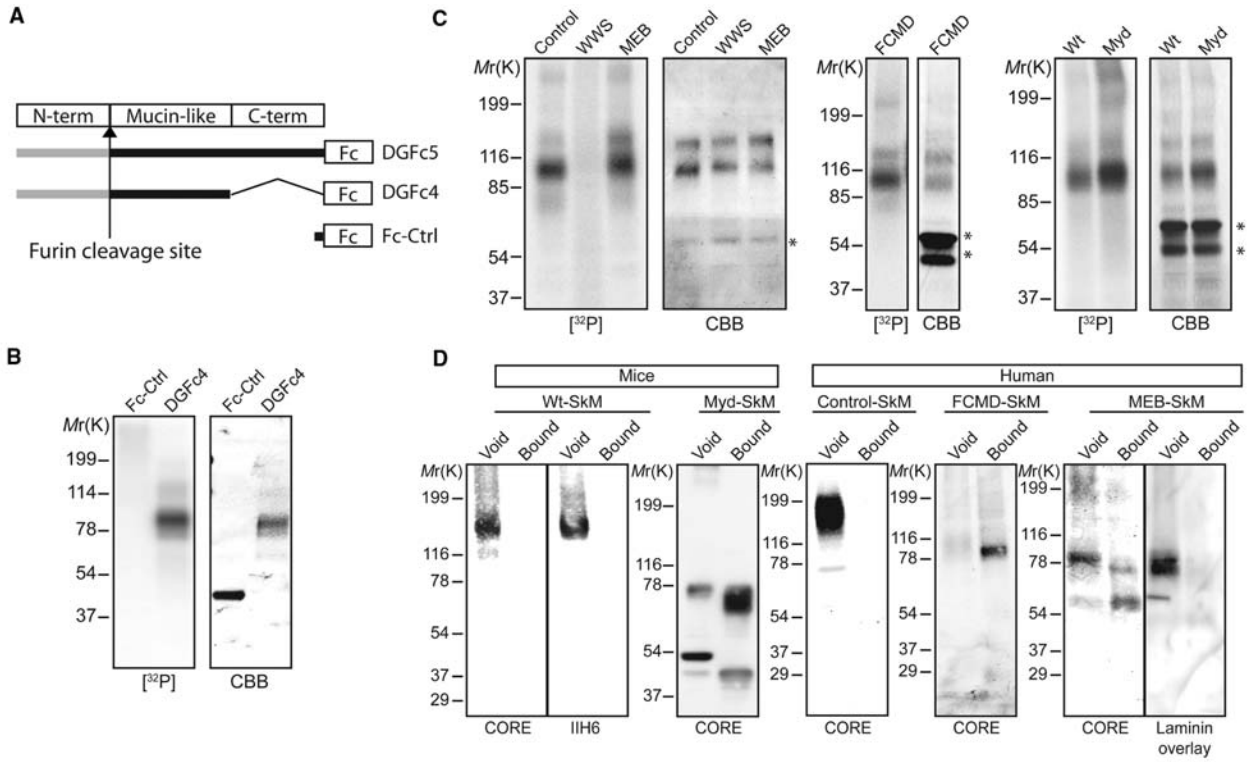

Fig. 2. The mucin-like domain of α-DG is phosphorylated in an *O*-mannosylation–dependent manner. (**A**) Structures of recombinant α-DG constructs used in the study. (**B** and **C**) [^{32}P]-orthophosphate labeling of (B) Fc-Ctrl– or DGFc4-expressing HEK293 cells and (C) DGFc5-expressing cultured cells from CMD patients (WWS, MEB, or FCMD) and control humans, and from WT and *Large*^{myd} (Myd) mice. Fc-tagged recombinant α-DG was isolated from the culture medium with protein-A agarose, separated by SDS–polyacrylamide gel electrophoresis, stained with Coomassie brilliant blue (CBB), and analyzed by phosphorimaging ([^{32}P]). Phosphorylation of α-DG required prior *O*-mannosylation. Asterisks indicate contaminating proteins derived from fetal bovine serum. (**D**) IMAC-binding assay testing glycoproteins from WT and *Large*^{myd} mice, and from FCMD, MEB, and control human muscle (SkM).

of α-DG. By measuring inorganic phosphate after acid hydrolysis, we confirmed that native α-DG purified from rabbit skeletal muscle is phosphorylated at 4.7 mol of phosphate per mole of protein (SD = 0.22, n = 3 trials).

To assess whether CMD cells that synthesize phosphorylated α-DG can further modify the phosphate residue, we immunoprecipitated glycoproteins from mouse *Large*^{myd} and WT muscle, as well as from human *POMGnT1*-mutated MEB, *fukutin*-mutated FCMD, and control muscle, by using immobilized metal affinity chromatography (IMAC) beads that bind to monoester-linked, but not diester-linked, phosphorylated compounds. Only *fukutin*-mutated FCMD and *Large*^{myd} muscle α-DG were captured by the beads, revealing that the phosphate residue does not undergo further modification in these CMD cells (Fig. 2D). This finding suggests that fukutin and LARGE participate in a common pathway to assemble the laminin-binding moiety onto the phosphorylated *O*-linked mannose. This speculation is compatible with the fact that α-DG prepared from *Large*^{myd} muscle that was rescued by adenovirus-mediated expression of LARGE regains laminin-receptor activity and concomitantly loses its affinity for IMAC beads (fig. S3). In the case of *POMGnT1*-mutated MEB patient muscle, several forms of α-DG were observed; the majority of these were captured by the beads, although a certain amount of α-DG with laminin-binding activity was detected in the void fraction (Fig. 2D). This finding suggests that a defect in POMGnT1 partially inhibits its modification on the phosphoryl branch chain of the *O*-mannosyl glycan on α-DG.

DGFc4 that was produced by HEK293 cells bound to IMAC beads (fig. S4A) and gained laminin-binding activity when it was coexpressed with LARGE (fig. S4B); such a gain in activity has also been observed in FCMD, MEB, and *Large*^{myd} cells, both in this study and elsewhere (*22*). To determine the structure of the phosphorylated *O*-mannosyl glycan that was necessary to assemble the laminin-binding moiety, we prepared *O*-glycans from HEK293-expressed DGFc4 by reductive β elimination, and we isolated the phosphorylated *O*-glycan using IMAC beads. Linear trap quadrupole (LTQ) mass spectrometry–based analyses detected prominent ions at mass-to-charge ratios m/z = 667 ($[M\text{-}H]^-$) and m/z = 333 ($[M\text{-}2H]^{2-}$) that are assigned as a phosphorylated trisaccharide composed of $HexNAc_2Hexitol_1$ (fig. S5), and analysis by high-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) revealed the compositional sugars to be GlcNAc, GalNAc, and mannitol (fig. S6). Homo- and heteronuclear NMR techniques were used to assign the $^{13}C/^{1}H$ heteronuclear multiple quantum coherence (HMQC) spectrum of the reduced *O*-glycan (Fig. 3A). The GlcNAc (subunit B) was assigned using double quantum-filtered correlation spectroscopy (DQF-COSY) and total correlation spectroscopy (TOCSY) spectra with a series of mixing times (fig. S7). The GalNAc (subunit C) was partially assigned based on a selective TOCSY-HSQC spectrum (Fig. 3B). The GalNAc (subunit C) is linked via a β1-3 linkage to the C3 position on GlcNAc (subunit B), which is in turn con-

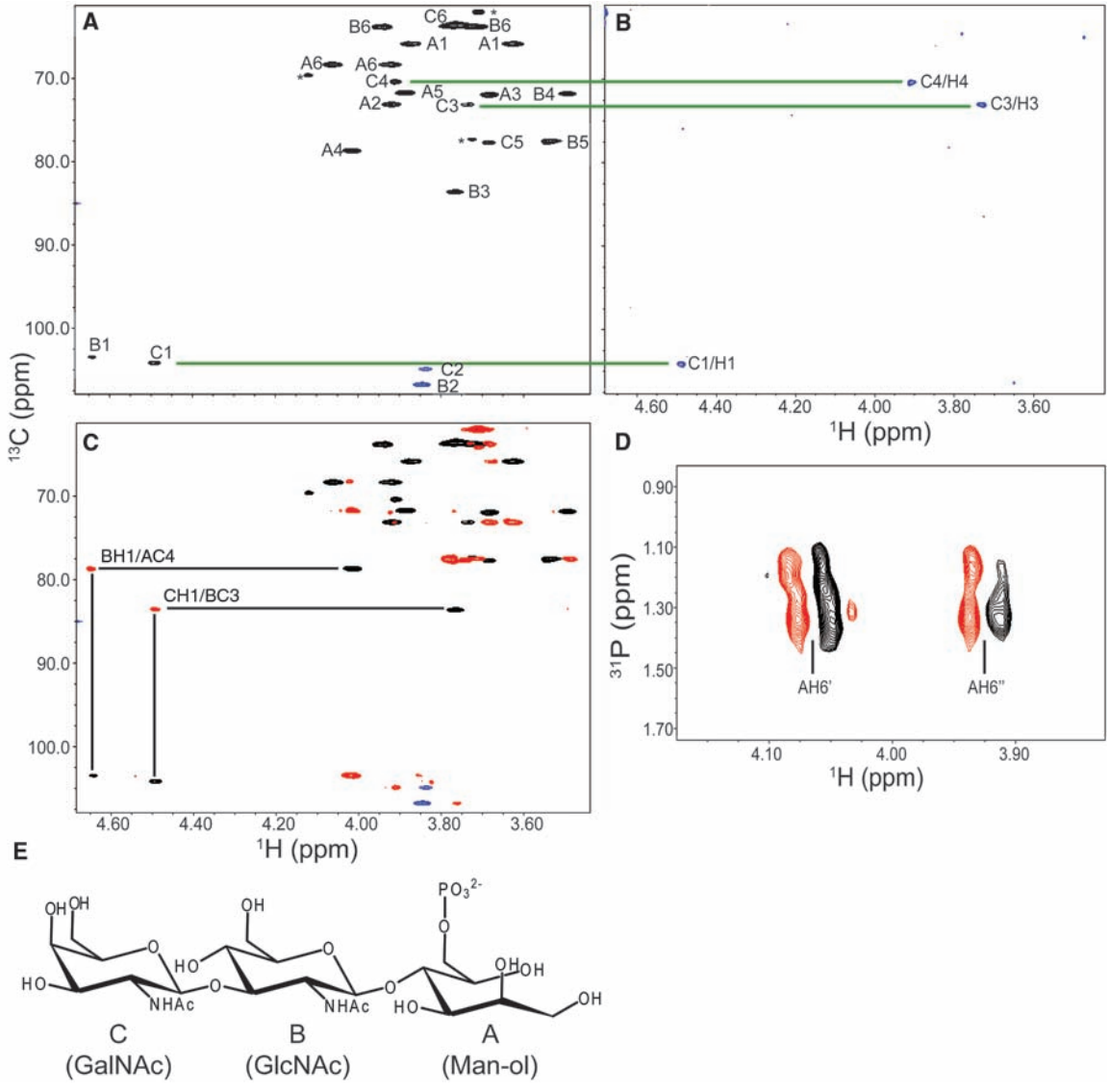

Fig. 3. NMR analysis of phosphorylated *O*-glycan on HEK293-produced DGFc4. (**A**) HMQC spectrum where the assigned cross peaks are labeled with a letter for the subunit designated in (E) and a number for the position on that subunit. The folded cross peaks are indicated in blue, and the cross peaks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

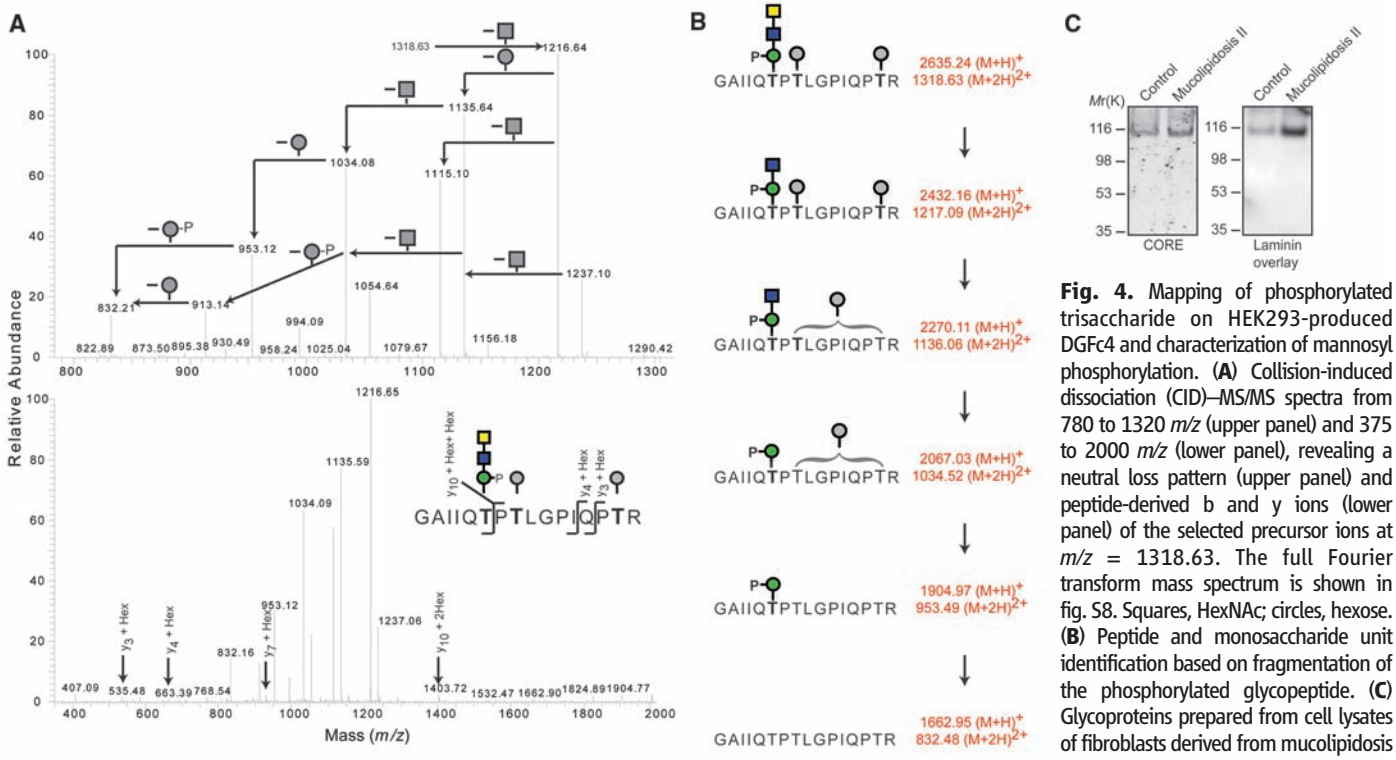

Fig. 4. Mapping of phosphorylated trisaccharide on HEK293-produced DGFc4 and characterization of mannosyl phosphorylation. (**A**) Collision-induced dissociation (CID)–MS/MS spectra from 780 to 1320 *m/z* (upper panel) and 375 to 2000 *m/z* (lower panel), revealing a neutral loss pattern (upper panel)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

nected via a β1-4 linkage to the C4 position on mannitol (subunit A), as evidenced by the observed heteronuclear multiple-bond correlation (HMBC) cross peaks CH1/BC3 and BH1/AC4 (Fig. 3C). The phosphate group is attached to the C6 position of the mannitol (subunit A) as determined from the cross peaks ^{31}P/AH6′ and ^{31}P/A6H″ detected in the ^{31}P/^{1}H COSY spectrum (Fig. 3D). The complete NMR resonance assignments of the reduced *O*-glycan and its inter-residue correlations detected in the nuclear Overhauser effect spectroscopy (NOESY) and HMBC spectra are summarized in table S1, and the determined structure is shown in Fig. 3E. To verify that this phosphorylated trisaccharide modifies the mucin-like domain of DGFc4, we enriched the trypsinized peptides using *Wisteria floribunda* agglutinin-lectin and analyzed the GalNAc-terminated peptides by liquid chromatography–mass spectrometry (LC-MS)/MS. MS/MS fragmentation patterns at *m/z* = 1318.63 (Fig. 4A), 1420.17 (fig. S9), and 1501.19 (fig. S10) identified a peptide (amino acids 374 to 389 of α-DG; GenBank ID CAA45732) bearing these modifications: the phosphorylated trisaccharide in conjunction with Hex-HexNAc-Hex, HexNAc-Hex, or Hex. The presence of nonphosphorylated mannose-initiated structures on y3, y4, and y10 ions revealed that Thr^{379} is modified by the phosphorylated trisaccharide in all cases (Fig. 4, A and B, and figs. S9 and S10). Additional studies showed that α-DG is phosphorylated within the Golgi complex (fig. S11) and that this phosphorylation occurs independently from the mannose-6-phosphate synthetic pathway that is required for lysosomal protein modification (*23*); fibroblasts derived from patients with mucolipidosis II (OMIM ID 252500), which have a defect in GlcNAc-1-phosphotransferase, can synthesize the laminin-binding form of α-DG (Fig. 4C).

We demonstrated that MEB, FCMD, and *Large^{myd}* cells, which have genetically distinct abnormalities, show a similar defect in post-phosphoryl modification on the *O*-mannosyl glycan (fig. S12). These convergent mechanisms to pathology offer an explanation for previous reports that forced expression of LARGE can circumvent defects in α-DG modification in these CMD cells (*22*). We speculate that LARGE, a putative glycosyltransferase with catalytic domains sharing homology with β-1,3-N-acetylglucosaminyltransferase and bacterial glycosyltransferase (*19*) participates in postphosphoryl glycosylation, because the forced expression increases the affinity of the cell surface for both the IIH6 antibody (fig. S13A) and the *Vicia villosa* lectin (fig. S13B). To our knowledge, we provide the first evidence that a vertebrate non– glycosylphosphatidylinositol-anchored glycoprotein is modified by a phosphodiester linkage. Glycoproteins in the cell walls of yeasts and fungi bear phosphodiester-linked glycans that are generated by a process involving phosphorylation on the C6 hydroxyl of mannose (*24*). α-DG, which is well conserved as an epithelial cell-surface protein in species ranging from lower vertebrates to mammals, is likewise modified by this ancient type of glycosylation. A recent study has shown that the most severe form of CMD—WWS—is a genetically heterogeneous disease. Moreover, only 40% of WWS cases are explained by mutations in known CMD-causative genes (*25*). Thus, a defect in the phosphorylation of an *O*-linked mannose may be responsible for severe CMD, indicating that the discovery of mutations in new genes responsible for WWS may not be far off.

References and Notes

1. O. Ibraghimov-Beskrovnaya *et al.*, *Nature* **355**, 696 (1992).
2. S. H. Gee, F. Montanaro, M. H. Lindenbaum, S. Carbonetto, *Cell* **77**, 675 (1994).
3. W. Cao *et al.*, *Science* **282**, 2079 (1998).
4. C. F. Spiropoulou, S. Kunz, P. E. Rollin, K. P. Campbell, M. B. Oldstone, *J. Virol.* **76**, 5140 (2002).
5. S. Kunz, N. Sevilla, D. B. McGavern, K. P. Campbell, M. B. Oldstone, *J. Cell Biol.* **155**, 301 (2001).
6. S. Kunz, J. M. Rojek, M. Perez, C. F. Spiropoulou, M. B. Oldstone, *J. Virol.* **79**, 5979 (2005).
7. D. Beltran-Valero de Bernabe *et al.*, *Am. J. Hum. Genet.* **71**, 1033 (2002).
8. J. van Reeuwijk *et al.*, *J. Med. Genet.* **42**, 907 (2005).
9. A. Yoshida *et al.*, *Dev. Cell* **1**, 717 (2001).
10. K. Kobayashi *et al.*, *Nature* **394**, 388 (1998).
11. M. Brockington *et al.*, *Am. J. Hum. Genet.* **69**, 1198 (2001).
12. C. Longman *et al.*, *Hum. Mol. Genet.* **12**, 2853 (2003).
13. D. E. Michele *et al.*, *Nature* **418**, 417 (2002).
14. H. Manya *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 500 (2004).
15. T. Sasaki *et al.*, *Biochim. Biophys. Acta* **1425**, 599 (1998).
16. A. Chiba *et al.*, *J. Biol. Chem.* **272**, 2156 (1997).
17. A. C. Combs, J. M. Ervasti, *Biochem. J.* **390**, 303 (2005).
18. T. Ilg *et al.*, *J. Biol. Chem.* **271**, 21583 (1996).
19. P. K. Grewal, P. J. Holzfeind, R. E. Bittner, J. E. Hewitt, *Nat. Genet.* **28**, 151 (2001).
20. M. Kanagawa *et al.*, *Cell* **117**, 953 (2004).

21. T. Ilg *et al.*, *J. Biol. Chem.* **269**, 24073 (1994).
22. R. Barresi *et al.*, *Nat. Med.* **10**, 696 (2004).
23. N. M. Dahms, P. Lobel, S. Kornfeld, *J. Biol. Chem.* **264**, 12115 (1989).
24. J. E. Gander, *Annu. Rev. Microbiol.* **28**, 103 (1974).
25. M. C. Manzini *et al.*, *Hum. Mutat.* **29**, E231 (2008).
26. This study was supported in part by a Paul D. Wellstone Muscular Dystrophy Cooperative Research Center grant (1U54NS053672 to K.P.C.). K.P.C. is an Investigator of the Howard Hughes Medical Institute. Additional acknowledgments and funding sources are provided in the supporting online material (SOM) text.

Supporting Online Material
www.sciencemag.org/cgi/content/full/327/5961/88/DC1
Materials and Methods
SOM text
Figs. S1 to S13
Table S1
References

12 August 2009; accepted 20 October 2009
10.1126/science.1180512

The Rate and Molecular Spectrum of Spontaneous Mutations in *Arabidopsis thaliana*

Stephan Ossowski,[1]* Korbinian Schneeberger,[1]* José Ignacio Lucas-Lledó,[2]*† Norman Warthmann,[1] Richard M. Clark,[3] Ruth G. Shaw,[4] Detlef Weigel,[1]† Michael Lynch[2]

To take complete advantage of information on within-species polymorphism and divergence from close relatives, one needs to know the rate and the molecular spectrum of spontaneous mutations. To this end, we have searched for de novo spontaneous mutations in the complete nuclear genomes of five *Arabidopsis thaliana* mutation accumulation lines that had been maintained by single-seed descent for 30 generations. We identified and validated 99 base substitutions and 17 small and large insertions and deletions. Our results imply a spontaneous mutation rate of 7×10^{-9} base substitutions per site per generation, the majority of which are G:C→A:T transitions. We explain this very biased spectrum of base substitution mutations as a result of two main processes: deamination of methylated cytosines and ultraviolet light–induced mutagenesis.

Most of what we know about molecular evolution comes from the comparison of biological sequences that have survived many cycles of natural selection. In order to infer the properties of the original source of variation and to detect the signature of natural selection from such data sets, we need to assume that variants affecting certain types of sites, such as the last base of fourfold redundant codons or pseudogenes, are not subject to natural selection. This pervasive assumption is very rarely tested and difficult to avoid, because of the slow pace of spontaneous mutagenesis. However, with the advent of high-throughput sequencing technologies, some estimates of the rate of spontaneous mutations have begun to appear (*1*–*3*). Here, we report a direct estimate of the spontaneous base substitution rate in *Arabidopsis thaliana*, a plant species with extensive DNA methylation. As a result, we reduce the uncertainty associated with key aspects of the evolutionary history of this species, including the time since divergence from *A. lyrata* and the effect of methylation on the probability of mutation.

We sequenced the genomes of five individuals derived by 30 generations of single-seed descent from the reference strain Col-0 (*4*), for which a high-quality genome was published in 2000 (*5*). We used the Illumina (Illumina, San Diego, CA) Genome Analyzer platform to obtain a depth of sequence coverage of between 23 and 31 in each mutation accumulation (MA) line. Sequencing reads between 36 and 43 base pairs (bp) in length were aligned to the reference genome, from which over 1000 errors had been removed in a previous study (*6*). We identified single-base substitutions using two complementary methods: a "consensus" approach and the single-nucleotide polymorphism (SNP) caller function of SHORE (http://1001genomes.org/downloads/shore.html) (*7*).

In the consensus approach, base substitutions were called if one of the MA lines differed from all others. We estimated a frequency of sequencing errors of ~0.3% per site per read (*8*). We assumed a binomial distribution of errors to derive the probabilities of false positives and false negatives, and we corrected our estimates accordingly (*9*). Because sequencing and mapping errors are not randomly distributed among sites (*3*), we used strict quality filters to exclude from analysis sites suspected to have higher error rates (*8*). Between 93 and 95 million sites out of the 120 million–bp reference genome matched the quality requirements in each line. Across all five lines, 85 single-base substitutions were called by this method, 83 of which were confirmed by Sanger sequencing. In the other two sites, two or three lines had a nonreference base, whereas the rest matched the reference, and we interpret this to be a result of differential fixation of the two alleles present in ancestrally heterozygous sites rather than as parallel mutations, although the latter cannot be ruled out.

In addition to this conservative approach, we used SHORE to detect single-base substitutions, short insertions and deletions (indels) of up to 3 bp, and long deletions. The algorithms implemented in SHORE are more sensitive (*8*), and between 98.8 and 100.9 million sites in each line had sufficient read information for calling either a mutation or the reference base. We detected 99 single-base substitutions (98 of which were confirmed by Sanger sequencing, and 1 was rejected because the reference base was revealed), 9 short deletions (8 confirmed, 1 rejected), 5 short insertions of 1 bp (all confirmed), and 8 long deletions covering 11 to over 5000 bp (4 confirmed, 2 ambiguous, and 2 rejected). A 2-bp deletion was shown to be present in two lines, suggesting that it was heterozygous in the ancestral line. The chromosomal positions of all validated mutations are shown in Fig. 1. Both false positives and false negatives are expected to be absent from the final set of simple base-substitution mutations (*8*). Fifteen sites where all MA lines had a common composition, but different from the reference, including 13 single-base substitutions and two deletions of 1 and 2 bp, were interpreted as fixed mutations in the ancestral line.

We estimated the overall mutation rate to be $5.9 \times 10^{-9} \pm 0.6 \times 10^{-9}$ base substitutions per site per generation according to the consensus approach and $6.5 \times 10^{-9} \pm 0.7 \times 10^{-9}$ according to SHORE. In addition, joint maximum-likelihood estimates of the overall mutation rate and the sequencing error frequency were obtained following a recently developed method (*9*). With this approach, a slightly higher mutation rate of $7.1 \times 10^{-9} \pm 0.7 \times 10^{-9}$ at a slightly lower error frequency of 0.2% was estimated. Mutations were evenly distributed among MA lines (Table 1). Within chromosomes, a significantly higher base-substitution mutation rate for intergenic regions was observed closer to the centromere (within 3.0×10^6 bp, for example) than farther away (Fisher's exact test, P value = 0.01, for nonmethylated sites).

The estimated rates of 1- to 3-bp deletions and insertions are $0.6 \times 10^{-9} \pm 0.2 \times 10^{-9}$ and $0.3 \times 10^{-9} \pm 0.1 \times 10^{-9}$ per site per generation, respectively. Out of the 13 short indels that we observed, 6 were found in complex sequences, corresponding to a mutation rate of $4.0 \times 10^{-10} \pm 1.6 \times 10^{-10}$ indels per site per generation, or about 0.05 ± 0.02 indels per haploid genome per generation, excluding homopolymers and microsatellites. This estimate should be considered a

[1]Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany. [2]Department of Biology, Indiana University, Bloomington, IN 47405, USA. [3]Department of Biology, University of Utah, Salt Lake City, UT 84112, USA. [4]Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN 55108, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: joslucas@indiana.edu (J.I.L.-L.); weigel@weigelworld.org (D.W.)

Supporting Online Material

Materials and Methods

cDNAs. All Fc-tagged α-DG recombinants used in this study (Wt construct DGFc5, deletion construct DGFc4 (DGEKFc4), and Fc-Ctrl (IgG1Fc-pcDNA3)) have been described previously, in (*S1-3*).

Antibodies. Both the AP83 rabbit polyclonal and IIH6 mouse monoclonal antibodies, which recognize β-DG and the unidentified laminin-binding modification of α-DG, respectively, were described previously (*S4*). CORE is from a sheep polyclonal antiserum raised against the entire dystrophin-glycoprotein complex, and was purified against a hypoglycosylated DGFc5 produced in TSA cells. Biotinylated anti-human IgG was obtained from Vector Laboratories (CA, USA).

Protein enrichment, immunoblots, overlay assays, and solid-phase assays. Frozen skeletal muscle and cultured cells were processed as described (*S4*). Immunoblots, laminin overlay assays, and solid-phase assays were conducted as described (*S4*). Virus overlay assays were conducted as described (*S5*). Protein enrichment with IMAC-beads was carried out using PHOS-beads (Sigma) according to the manufacturer's protocol. Both void and bound samples correspond to 100% of input. Prior to the enrichment, samples were desalted using Amicon-Ultra (Millipore) filters.

Treatment with HFaq. Samples were incubated with ice-cold 48% HFaq (Aldrich) at 0°C for 20 h. The reagent was removed under a stream of nitrogen on ice. Control samples were prepared by the same procedure, except that ice-cold water was used instead of 48% HFaq.

Treatment with enzymes. Alkaline phosphatase from calf intestine (NEB), α(1-2, 3, 6)-mannosidase from jack bean (Glyko), β(1-2, 3, 4, 6)-N-acetylhexosaminidase from jack bean (Glyko), α(1-2, 3, 4, 6)-fucosidase from bovine kidney (Sigma), and the Enzymatic Protein Deglycosylation Kit: including PNGaseF, *O*-glycosidase, α-2(3,6,8,9) neuraminidase, β(1-4)-galactosidase, and β-N-Acetylglucosaminidase (Sigma) were used in this study accoding to the manufacturer's protocol. The samples were desalted using Amicon-Ultra (Millipore) filters prior to treatment.

Quantification of phosphate residues on α-DG. α-DG purified from rabbit skeletal muscle as described elsewhere (*S6*) was treated with Antarctic Phosphatase (NEB) or was hydrolyzed by applying 6N HCl at 110°C for 2 h, conditions under which phosphoamino acids remain intact (*S7*). Released inorganic phosphate was quantified using the Malachite Green Phosphate Assay Kit (BioAssay Systems). For controls, α-DG was incubated under the same conditions, with the heat-inactivated phosphatase or in the absence of HCl. We

subtracted the amount of inorganic phosphate in the control from that in the treated samples.

[^{32}P] orthophosphate labelling of cells. Cells were cultured in DMEM with 0.5% penicillin-streptomycin and 2 mM L-glutamate supplemented with fetal bovine serum (FBS) at 10% (HEK293 cells) or 15% (fibroblasts). Prior to ^{32}P-labeling, HEK293 cells were transfected with DGFc4 or Fc-Ctrl using FuGENE6 (Roche Applied Science). Fibroblasts were infected with the E1-deficient recombinant adenovirus Ad5-CMV-DGFc5-WT, which encodes DGFc5 under the CMV-promoter, at an MOI of 1000. At 24 h post-infection/transfection, the cells were starved in phosphate-free DMEM (Gibco) with 10% dialyzed FBS (Gibco) for 1 h, and then labelled with 0.7 mCi/ml of ^{32}P-orthophosphate (Amersham Code PBS11; 10 mCi/ml) in phosphate-free medium for one hour. Cells were cultured in the usual medium for another 4 days, after which recombinant DG was enriched from the medium using Protein-A agarose (Santa Cruz). The beads were boiled with 5x Laemmli sample buffer and applied to 3–15 or 4-12 % SDS-PAGE. The gels were stained with Coomassie brilliant blue (CBB) and then exposed to phosphoimager plates (Fuji).

Phosphoamino acid analyses. HEK293 cell-expressed [^{32}P]-DGFc4 was purified as described above and hydrolyzed by applying 6N HCl at 110°C for 2 h, conditions under which phosphoamino acids remains intact (*S7*). The hydrolysates were cooled at room temperature and dried in a Speedvac (Thermo Savant). The samples were re-suspended in 50 μl of water and dried in the Speedvac. This step was repeated three times. The hydrolysates were finally re-suspended in 10 μl of water and subjected to TLC analysis. The TLC analysis was conducted as described elsewhere (*S8*).

Preparation of *O*-glycan sample. DGFc4 was produced in HEK293H cells, essentially as described elsewhere (*S2*) with minor modification. The suspension cells were cultivated in a CELLine bioreactor (CL1000, Argos). DGFc4 was purified from the medium using a Hitrap Protein-A HP column (GE Healthcare). *O*-glycans were liberated from purified DGFc4 by reductive β-elimination with 0.1 M NaOH and 1 M $NaBH_4$ for 16 h at 45°C. After passage over a Dowex 50W-X8 (H^+) column, the solution was co-evaporated with methanol several times using a Speedvac. Phosphorylated oligosaccharide-alditol was captured using PHOS-beads (Sigma), eluted in modified PBS in which PO_4^{3-} was at 200 mM, and desalted using a porous graphitic carbon column (GlycoClean H Cartridges, Glyko) according to the manufacturer's protocol.

LTQ XL liner ion trap mass spectrometery and HPAEC-PAD analyses. The glycan samples were dissolved in 50% acetonitrile/water (v/v) and directly infused into an LTQ XL liner ion trap mass spectrometer (Thermo Scientific, Inc.) at 5 μl/min, using a syringe pump. The samples were ionized using negative nanospray ionization. Capillary temperature and needle voltage were 200°C and -4.5-5 kV, respectively. Collisions for MS^n were carried out with an isolation window of 3.0 u and a normalized collision energy of 20-35%. Compositional

sugar analysis using HPAEC-PAD was conducted by the Glycotechnology Core Resource at the University of California, San Diego.

NMR study. After exchanging hydroxyl hydrogens of the *O*-glycan for deuteriums by repeated dissolution in 99.9% D_2O and lyophilization, the *O*-glycan sample was dissolved in 99.9% D_2O and used directly in the NMR studies. NMR spectra were recorded at 25°C on a Bruker Avance II 800 MHz spectrometer equipped with a cryoprobe or a Bruker Avance II 500 MHz spectrometer. ^{1}H homonuclear two-dimensional DQF-COSY (*S9*), TOCSY (*S10, 11*), and ROESY (*S11, 12*) experiments and $^{1}H/^{13}C$ two-dimensional heteronuclear HMQC, HMBC, and selective-TOCSY-HSQC experiments (*S13*) were carried out using the 800 MHz NMR spectrometer. The $^{31}P/^{1}H$ COSY spectrum (*S14*) was acquired at the 500 MHz NMR spectrometer. The ^{1}H and ^{31}P chemical shifts were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate and external 2% H_3PO_4 in D_2O (*S15*), respectively. NMR spectra were processed with the NMRPipe package (*S16*), and analyzed using NMRView (*S17*). The 2D $^{13}C/^{1}H$ HMQC spectrum of the reduced O-glycan was assigned using combinations of ^{1}H homonuclear COSY, TOCSY, and ROESY, and heteronuclear HMQC, HMBC, and selective-TOCSY-HSQC techniques. A GlcNAc (subunit B) spin system was assigned using DQF-COSY (Fig. S7A) and TOCSY spectra with a series of mixing times (Fig. S7B-D). At a long mixing time of 120 ms, the GlcNAc (subunit B) H1 proton gave TOCSY cross peaks all the way to the H6 protons. A GalNAc (subunit C) spin system was partially assigned based on a selective-TOCSY-HSQC spectrum. When the GalNAc H1 proton was irradiated with a very selective excitation pulse, selective TOCSY transfer reached to its C3/H3 and C4/H4 cross peaks as detected by an HSQC spectrum (Fig. 3B).

Human and mouse cultured cells and patient biopsies. The MEB fibroblasts and FCMD myoblasts used in this study were described previously (*S18*). POMT1-deficient WWS fibroblasts were derived from a patient as described elsewhere (*S19*). Fibroblasts derived from Mucolipidosis II patients were obtained from the Coriell Institute (GM01586). The GlcNAc-phosphotransferase activity was less than 0.1% of that in normal fibroblasts (specific activity in cell lysate). Wt and *Large*^{myd} mouse fibroblasts were generated from the skin biopsies. The MEB muscle biopsy was obtained from a patient (age 6 months) with two heterozygous mutations in POMGnT1: c. 1539+1G>A, which causes a splice-site defect that affects the processing of exon 17, as reported elsewhere (*S20*); and at c. 704G>A, which is predicted to cause the amino acid substitution G235E. The FCMD muscle biopsy was obtained from a patient (age 7 months) homozygous for retrotransposon insertions in the fukutin gene (*S21*). Control muscle was obtained from a 47 year-old individual. All tissue was obtained and tested in compliance with the Human Subjects Institutional Review Board of the University of Iowa.

Mice and adenovirus infection. *Large*^{myd} mice (JAX) were backcrossed to C57BL/6 mice for 5 generations, and then maintained by the mating of heterozygous pairs. All mice were housed at the University of Iowa Animal Care Unit in accordance with animal usage guidelines. Infectious particles (2×10^{10}) of the E1-deficient recombinant adenovirus Ad5-CMV-LARGE, which encodes human LARGE under the CMV-promoter, were diluted in 0.9% NaCl to a final volume of 100 μl, and were injected percutaneously into quadriceps and calf muscles of *Large*^{myd} mice (12-18 weeks old) using an insulin syringe. At 10 days post-infection, the infected muscle was harvested and processed as described above. All mice were maintained at the University of Iowa Animal Care Unit in accordance with animal usage guidelines.

Assignment of the phosphorylated *O*-Mannose residue by LC-MS/MS. Approximately 2mg of DGFc4 produced in HEK293 cells was reduced, alkylated, and digested with 20 μg of sequence-grade trypsin (Promega), resuspended in 40mM NH_4HCO_3 and incubated at 37°C overnight. The reaction was quenched by adding 1% TFA, resulting in a final concentration of ~0.1% TFA. Enrichment of the digested material was carried out using agarose-bound *Wisteria Floribunda* agglutinin (WFA, Vector Laboratories, CA, USA), essentially as described previously (*S22*). Samples on the WFA-agarose column were eluted using the lectin-column buffer containing 200 mM N-acetylgalactosamine (Toronto Research Chemicals Inc.). Eluted glycopeptides were then desalted using C-18 spin columns (Nest Group) and dried.

The resulting material was reconstituted in 48.5 μl of 0.1% formic acid (mobile phase A) and 1.5 μl of 80% acetonitrile/0.1% formic acid (mobile phase B). The solution was loaded onto a tapered-tip nanospray capillary column (75 μm x 8.5 cm) and separated via a 160 min linear gradient of increasing mobile phase B at a flow rate of ~200nl/min. LC-MS/MS/MS analysis was performed on a LTQ Orbitrap XL mass spectrometer (ThermoFisher) equipped with nanospray. A full mass spectrum at 60,000 resolution was acquired from 400-2000 m/z, followed by three data-dependent MS/MS spectra of the most intense ions. When a neutral loss corresponding to a monosaccharide or phosphate was detected as one of the three most intense ions in the MS/MS spectra, this fragment ion was fragmented again by CID to yield an MS^3 spectrum.

The acquired spectra were searched against a non-redundant rabbit database (NCBI) that included a contaminant database, using BioWorks (3.3.1SP1; ThermoFisher). DTA files were generated for spectra with a threshold of 5 ions, a precursor ion tolerance of 1.4 amu, and a range of MH^+ 400-6000 m/z. The search parameters were set to allow for precursor ion tolerance of 50 ppm and fragmented ion tolerance of 0.5Da with strict trypsin cleavage. Peptides that were assigned as being modified by the O-mannosyl phosphorylated trisaccharide were confirmed manually.

Separation of the ER and Golgi organelles of DGFc4-overexpressing HEK293 cells, and immunoprecipitation of the immature DGFc4 with IMAC-

beads. The ER and Golgi in DGFc4-expressing HEK293 cells were separated by two steps of gradients. The methods were adopted from previous reports (*S23, 24*) with minor modification. The pellet of DGFc4-expressing HEK293 cells was washed one time with PBS and resuspended in homogenization buffer containing 0.25M sucrose, 1mM EDTA, and proteinase inhibitors in 10mM HEPES-NaOH (pH 7.4). The cells were dounce homogenized by hand with 6-7 strokes. The homogenate was spun down at 3,000 x g to obtain the post nuclear supernatant (PNS). The PNS was layered on a sucrose cushion of 8 ml of 1.2M Sucrose, 1mM EDTA, and proteinase inhibitors in 10mM HEPES-NaOH (pH 7.4) in a Beckman ultraclear tube (25 x 89mm). The tubes were spun on a SW 32 Ti rotor at 25,000 rpm (106,750 x g) for 98 minutes. A white band just above the 1.2M sucrose/sample interface was collected through the side of the tube, using a syringe with a 23-gauge needle. The sample was pelleted by diluting 1:1 with ddH_2O and spinning on MLA-80 at 70,000rpm (339,000 x g) for 1 hour. The pellet was resuspended in 1 ml of homogenization buffer, and loaded onto a multi-step discontinuous iodixanol gradient using Optiprep™ (Sigma). A multi-step discontinuous iodixanol gradient was made using 4 ml of 2.5%, 5.0 %, 7.5%, 10.0%, 12.5%, 15%, 17.5 % and 3 ml each of 20% and 30% iodixanol in 42 mM sucrose, 1mM EDTA, 10mm HEPES-NaOH (pH 7.4). The gradients were spun in an SW 32 Ti rotor for 5 hours at 32,000 rpm (174,899 x g) using acceleration and deceleration setting 9. At 5,000 rpm, the brake was taken off and the rotor was left to come to a stop on its own. The tube was punctured at the bottom with a 23-gauge needle, and 3ml fractions were collected, with fraction 1 being 30% iodixanol and fraction 12 being the lightest of 2.55 % iodixanol.

The ER and Goigi-enriched fractions were solubilized in 0.5% CHAPS, and the buffer was exchanged to 0.25 % CHAPS in water using Amicon Ultra (Milipore). The resulting solution was analyzed usign IMAC-beads as described above.

FACS study. HEK293 cells stably expressing LARGE were cloned and analyzed by FACS, using the CORE or IIH6 antibody, or biotinylated lectins (1:200 dilution, Vector Laboratories, CA, USA), as described previously (*S25*).

Supporting text

We thank M. Madson for insightful suggestions; M. E. Anderson and D. Venzke for technical assistance; L. M. Teesch for mass spectrometry, I. Nishino for providing WWS-patient cells, Paul D. Wellstone Muscular Dystrophy Cooperative Research Center Core B for providing patient cells and biopsies; the Gene Transfer Vector Core (UI, supported by NIH/NIDDK P30 DK 54759) for generating adenoviruses; the Glycotechnology Core Resource (UCSD) for HPAEC-PAD analysis; the Integrated Technology Resource for Biomedical Glycomics (UGA, supported by NIH/NCRR P41 RR018502) for tandem mass spectrometry; and past and present members of the Campbell lab for scientific contributions. This study was supported in part by AI55540 (to S.K. and M.B.A.O.).

Supporting figures

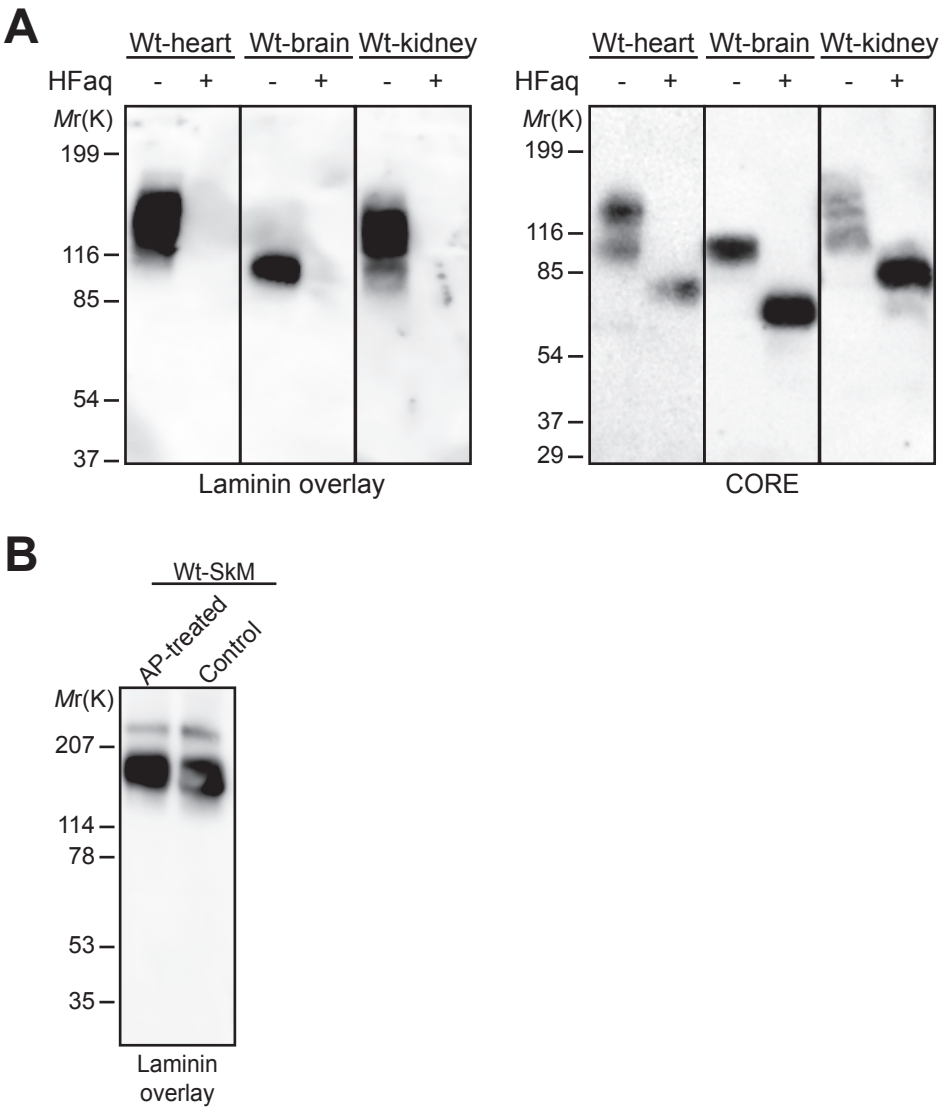

fig. S1. Enzymatic and chemical dephosphorylation of Wt mouse glycoproteins. **(A)** Treated glycoproteins prepared from heart, brain, and kidney of Wt mice were subjected to laminin overlay assay or immunoblotting with CORE antibody. **(B)** Laminin-overlay assay on Wt-muscle α-DG treated with alkaline phosphatase. In contrast to dephosphorylation by HFaq, digestion with alkaline phosphatase did not abolish the laminin-receptor activity of α -DG, suggesting that functional modification of α-DG involves an internal phosphoryl linkage rather than a monoester-linked phosphate.

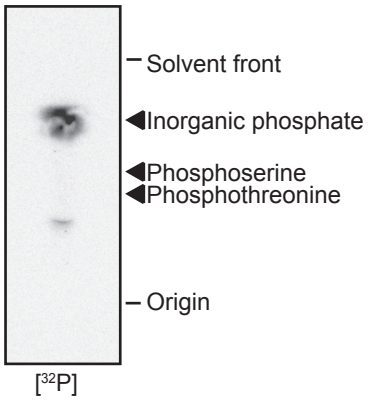

fig. S2. Phospho-amino acid analysis by thin-layer chromatography (TLC). [^{32}P]-labeled DGFc4 produced by HEK293 cells was hydrolyzed under conditions in which phospho-amino acids remain intact. The hydrolysate was separated by cellulose TLC. Arrowheads indicate the spot of unlabeled phospho-amino acid standards. None of [^{32}P]-phospho-amino acid was detected in the hydrolysate, suggesting that phosphorylation does not occur directly on the polypeptide.

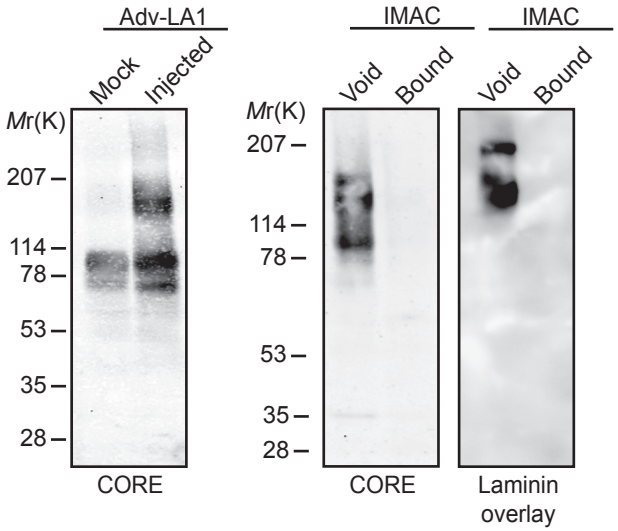

fig. S3. IMAC-binding assay testing glycoproteins from *Large*^{myd} mouse muscle expressing LARGE ectopically (adenovirus-mediated expression). The α-DG regained laminin-receptor activity and concomitantly lost its affinity for IMAC beads, suggesting that LARGE participates in assembly of the laminin-binding moiety on the muscle
α-DG phosphoryl residue.

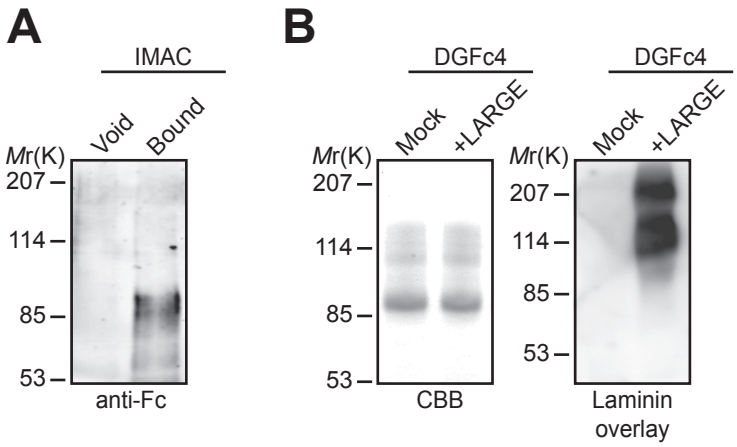

fig. S4. Characterization of DGFc4 produced by HEK293 cells. **(A)** IMAC-binding assay testing DGFc4 produced by HEK293 cells. **(B)** Laminin-overlay assay with DGFc4 produced by HEK293 cells, in the presence and absence of LARGE overexpression.

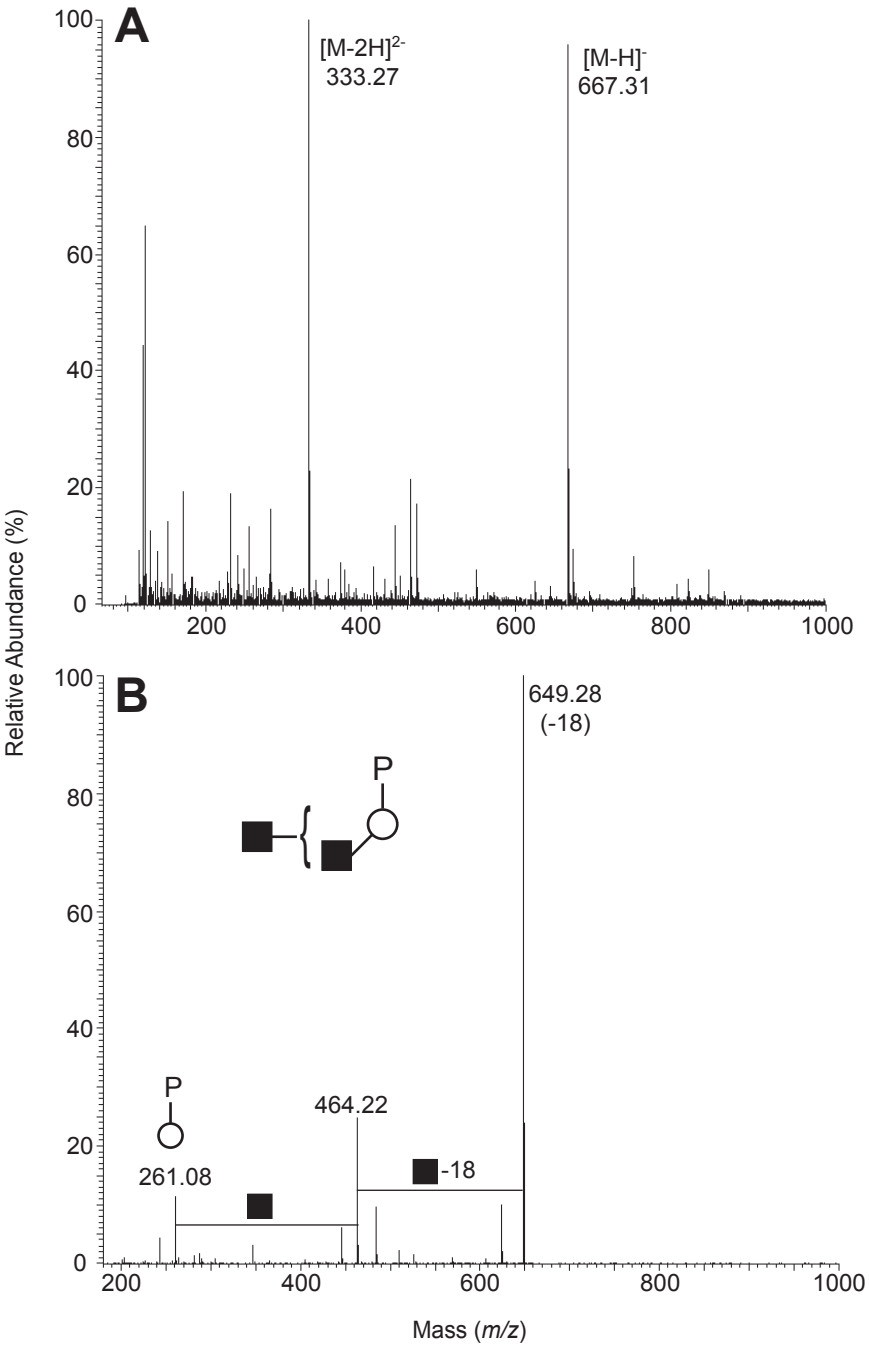

fig. S5. Nanospray ionization-linear ion trap mass spectrometry in negative-ion mode. **(A)** Full mass spectrum of IMAC-enriched *O*-glycan prepared by reductive β-elimination, from HEK293-produced DGFc4. **(B)** CID MS^2 spectrum derived from the precursor ion $[M-H]^-$ (m/z 667.3). Assignments of the fragment ions are shown on the schematic illustration. Filled square: HexNAc; open circle: Hexitol.

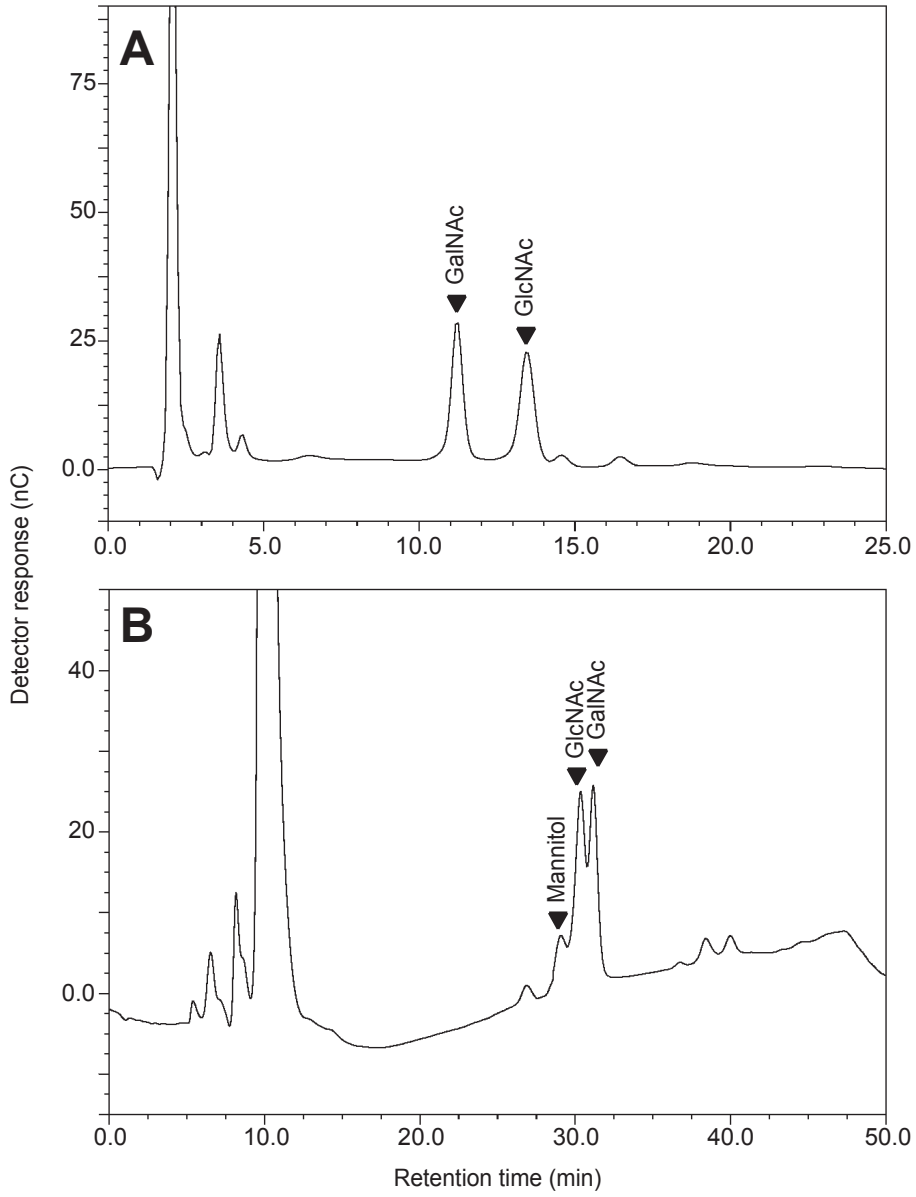

fig. S6. Compositional sugar analysis with HPAEC-PAD. IMAC-enriched *O*-glycan sample prepared by reductive β-elimination from HEK293-produced DGFc4 was acid-hydrolyzed to cleave glycosidic linkages. After hydrolysis, the sample was re-N-acetylated and analyzed by HPAEC-PAD using either a CarboPac PA-1 **(A)** or an MA-1 **(B)** column. GlcNAc, GalNAc and mannitol were detected as the compositional sugars.

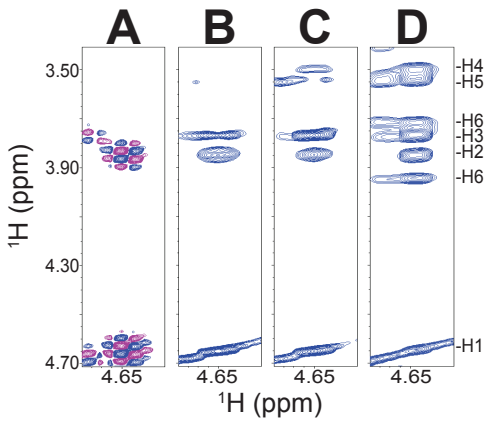

fig. S7. DQF-COSY (A) and TOCSY (B-D) spectra for the assignment of subunit B, with the assigned protons indicated along the right-side panel. TOCSY mixing times of 34, 60, and 120 ms were used for spectra (B), (C), and (D), respectively. At a long mixing time of 120 ms, the GlcNAc (subunit B) H1 proton gave TOCSY cross peaks all the way to the H6 protons.

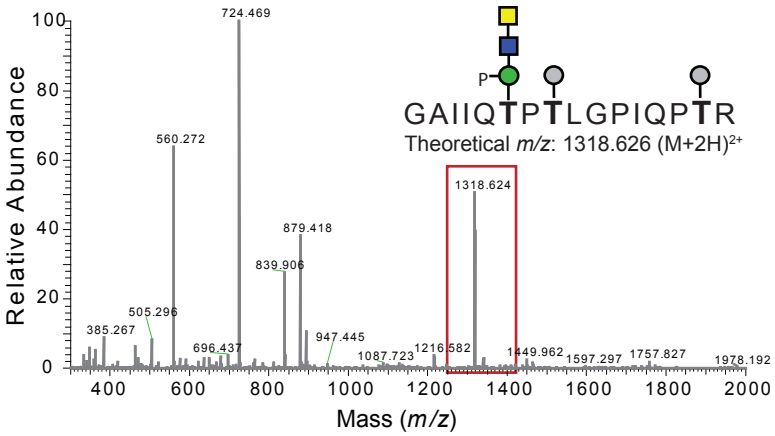

fig. S8. The full FT mass spectrum of the peptide (*m/z* 1318.624) modified by the phosphorylated trisaccharide. The full mass spectrum yielded a mass accuracy of 1.4 ppm for the precursor ions at *m/z* 1318.624. The CID-MS/MS spectrum is shown in Fig. 4A.

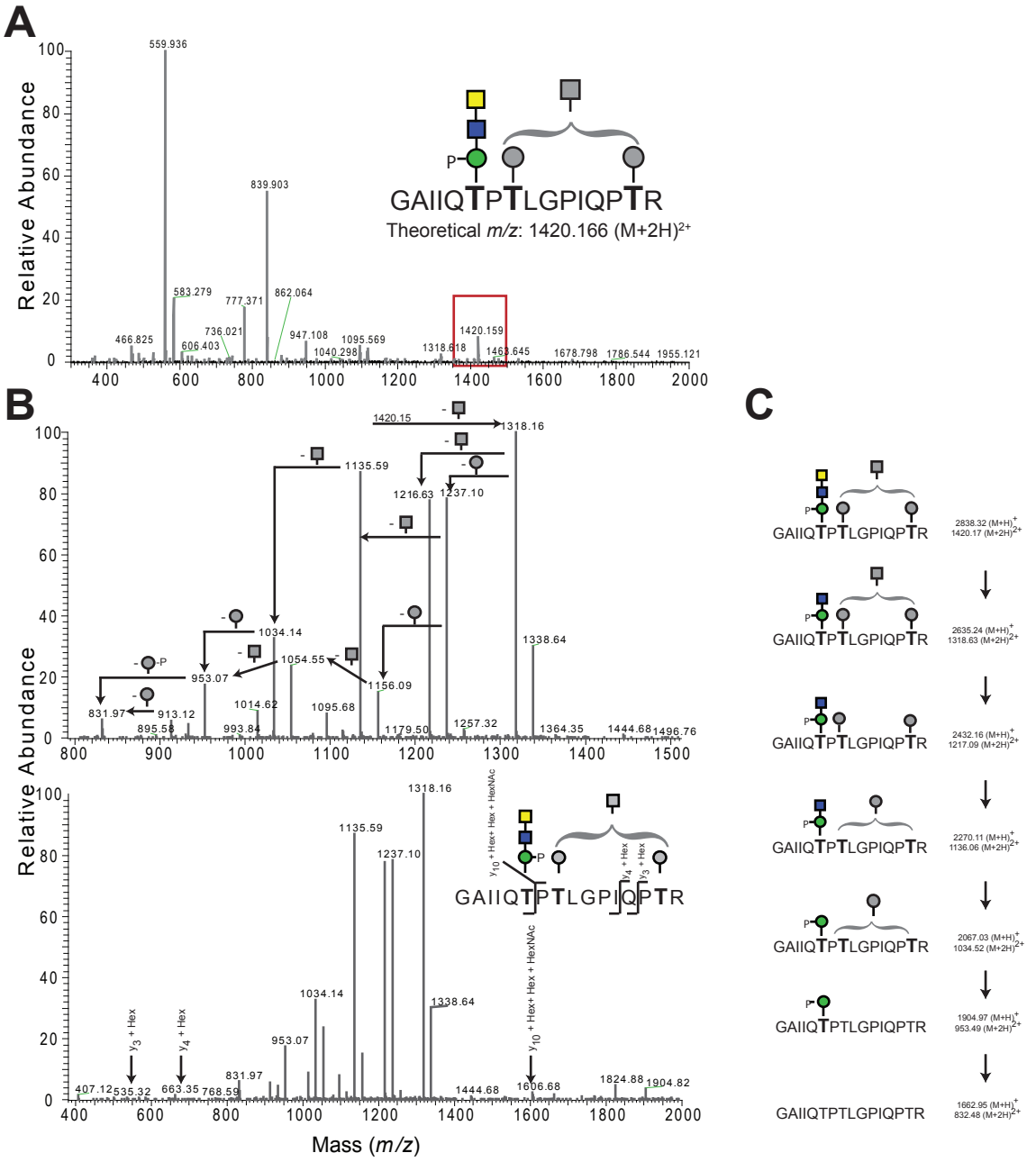

fig. S9. Mass analysis of the peptide modified by phosphorylated trisaccharide (*m/z* 1420.166). **(A)** The full FT mass spectrum of the peptide containing the phosphorylated trisaccharide. The full mass spectrum yielded a mass accuracy of 4.4 ppm for the precursor ion at *m/z* 1420.166. **(B)** The CID-MS/MS spectra from 780-1320 *m/z* (upper panel) and 375-2000 *m/z* (lower panel) show the neutral loss pattern (upper panel) and the peptide-derived b and y ions (lower panel) of the selected precursor ions at *m/z* 1420.17. Square: HexNAc; circle: Hexose. **(C)** The fragmentation of the phosphorylated glycopeptide allowed for identification of the peptide and mononsaccharide units as illustrated. The observation of ions corresponding to the non-phosphorylated hexoses on Thr_{381} and Thr_{388} revealed that Thr_{379} was modified by the phosphorylated trisaccharide and that Thr_{381} and Thr_{388} were occupied by a combination of HexNAc-Hex and Hex in this peptide.

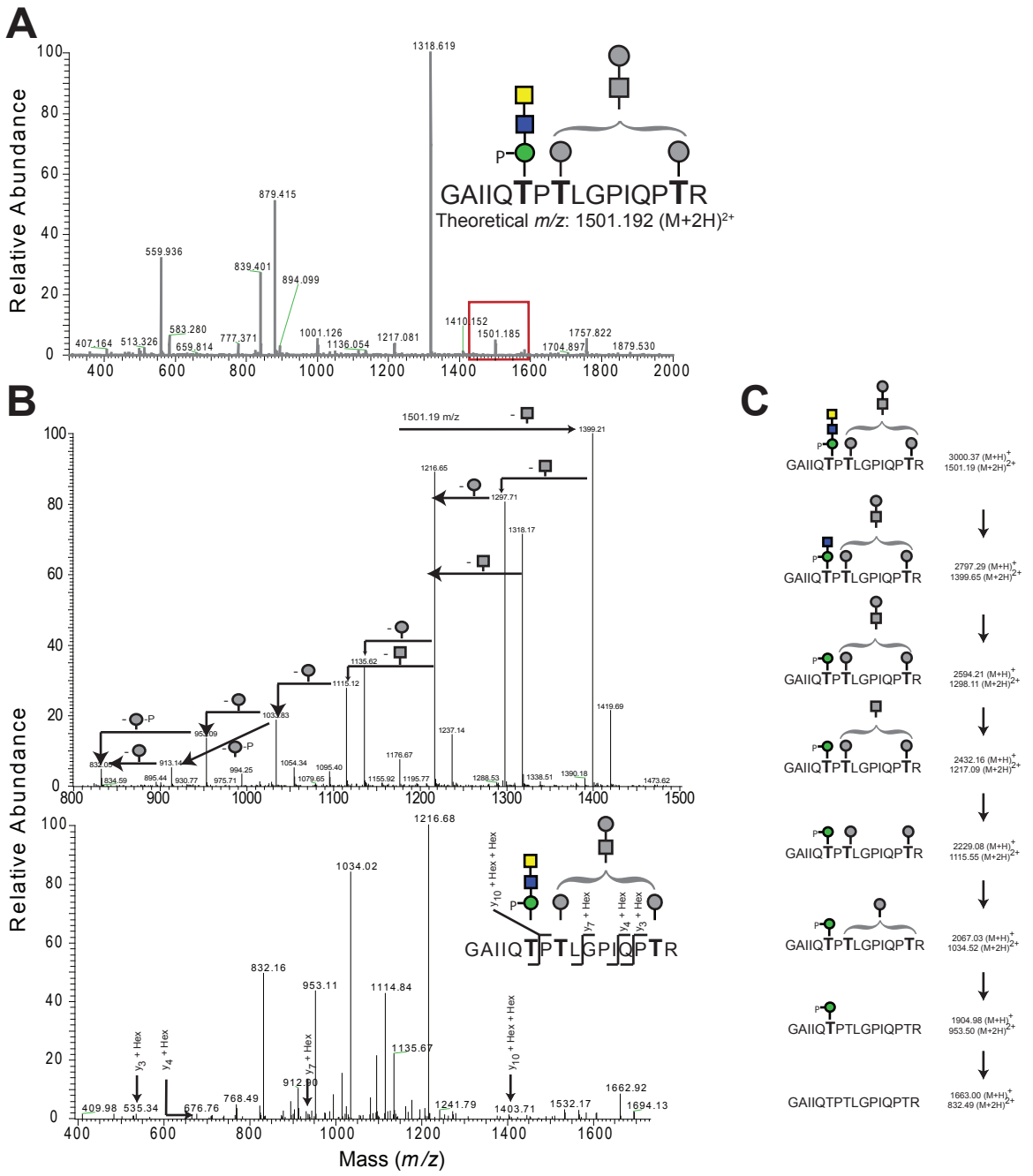

fig. S10. Mass analysis of the peptide modified by phosphorylated trisaccharide (*m/z m/z* 1501.192). **(A)** The full FT mass spectrum of the peptide containing the phosphorylated trisaccharide. The full mass spectrum yielded a mass accuracy of 4.5 ppm for the precursor ions at *m/z* 1501.192. **(B)** The CID-MS/MS spectra from 780-1320 *m/z* (upper panel) and 375-2000 *m/z* (lower panel) show the neutral loss pattern (upper panel) and the peptide-derived b and y ions (lower panel) of the selected precursor ions at *m/z* 1501.19. Square: HexNAc; circle: Hexose. **(C)** The fragmentation of the phosphorylated glycopeptide allowed for identification of the peptide and mononsaccharide units as illustrated. The observation of ions corresponding to the non-phosphorylated hexoses on Thr_{381} and Thr_{388} revealed that Thr_{379} was modified by the phosphorylated trisaccharide and that Thr_{381} and Thr_{388} were occupied by by a combination of Hex-HexNAc-Hex and Hex in this peptide. The Hex-HexNAc-Hex is likely to be Gal-β-1,4-GlcNAc-β-1,2-Man, which was previously shown to modify native and recombinant α-DG in brain, muscle and HEK293 cells (*S26*).

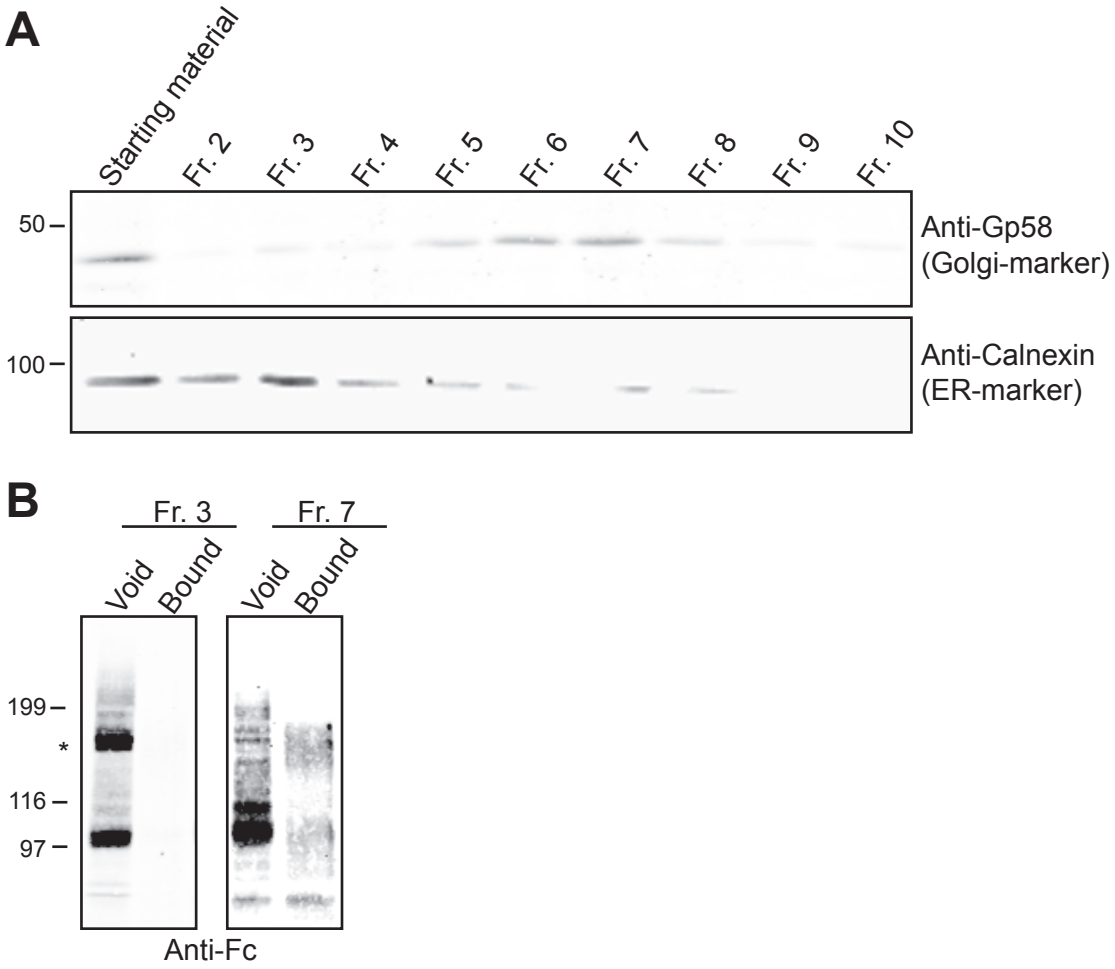

fig. S11. α-DG is phosphorylated in the Golgi. **(A)** Characterization of subcellular fractions prepared by gradient fractionation. HEK293 cells stably expressing DGFc4 were fractionated by discontinuous iodixanol gradient. The fractions were analyzed by immunoblotting with antibodies against a Golgi marker protein (gp58, upper panel) or an ER marker protein (calnexin, lower panel). **(B)** IMAC-binding assay testing immature DGFc4 in the ER and Golgi-enriched fractions. Only the Golgi fraction contained phosphorylated DGFc4, suggesting that phosphorylation occurs in the Golgi. Asterisk indicates dimerized DGFc4 in the Fc-region.

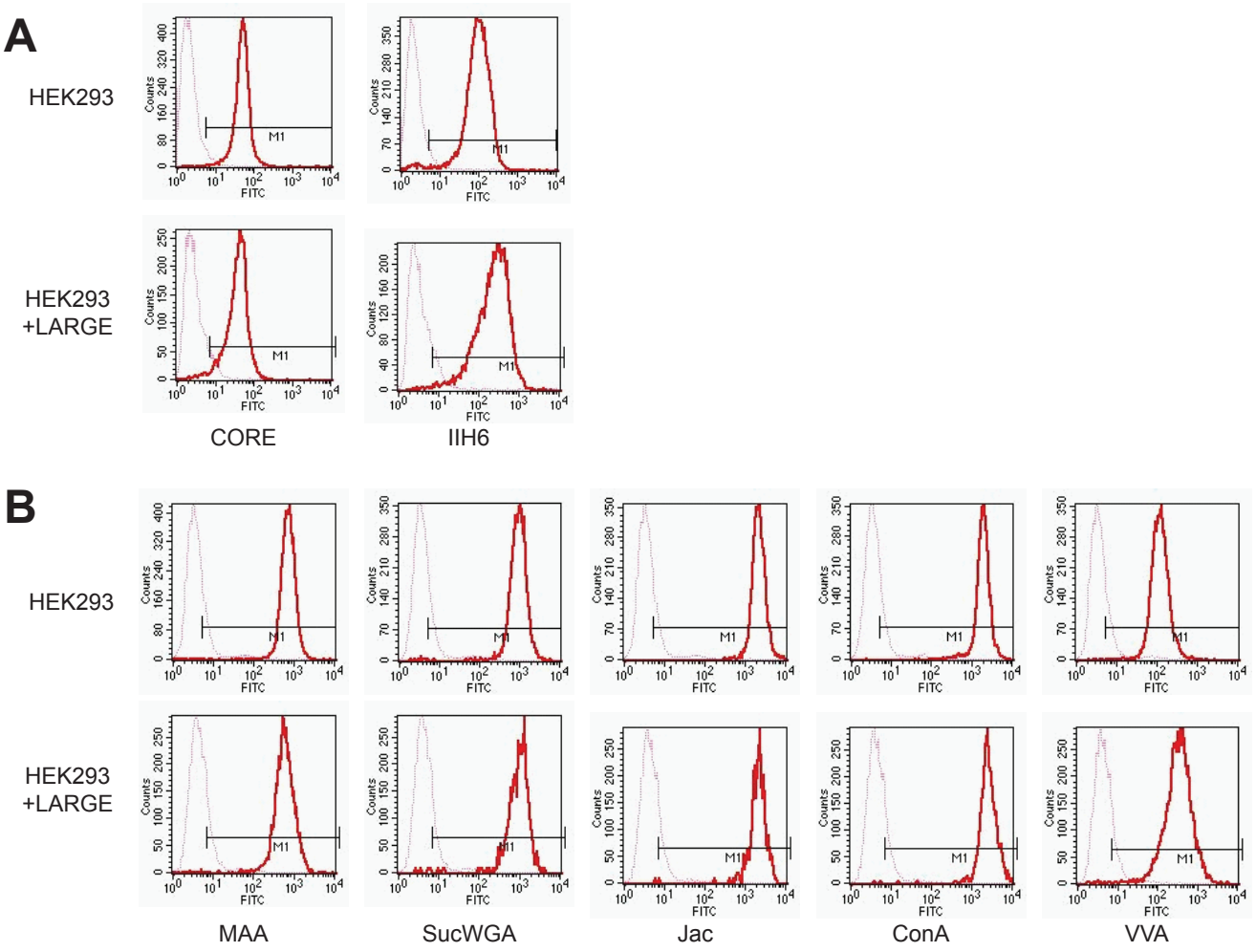

fig. S12. Effects of LARGE expression on glycosylation. HEK293 cells with stable expression of LARGE were generated, and the modified cells were tested by FACS using a CORE or IIH6 antibody **(A)**, or biotinylated lectins **(B)**. LARGE-overexpression increased immunoreactivity of the cell surface against IIH6. Also its overexpression increase the affinity of the cell surface specifically for the VVA lectin, which is known to bind to the laminin-binding form α-DG in muscle and brain (*S27, 28*). Solid line, primary (biotinylated lectins) and secondary antibody (streptavidin); broken line, secondary antibody (streptavidin) only. MAA, Maackia Amurensis Lectin II; SucWGA, succinylated Wheat Germ Agglutinin; Jac, Jacalin; ConA, Concanavalin A; VVA, Vicia Villosa Lectin.

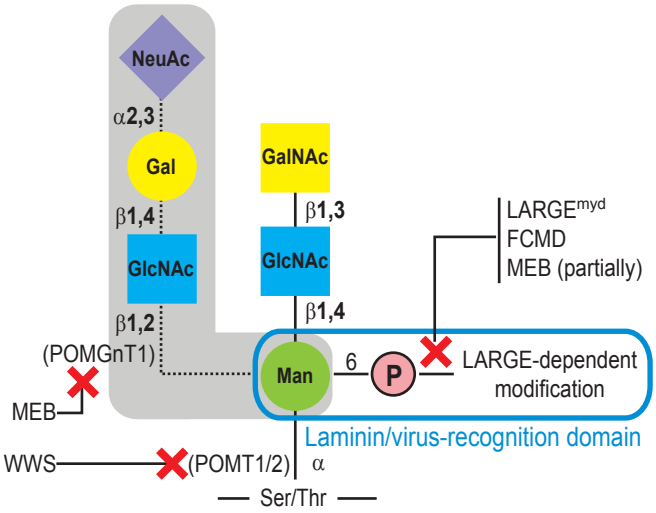

fig. S13. Schematic representation of *O*-mannosyl glycans on α-DG, with proteins involved in its biosynthesis, and diseases in which relevant steps are defective indicated. Shading indicates glycan identified in brain and muscle α-DG.

Supporting table

Table S1. Chemical shifts (ppm) of the signals in the ^{1}H and ^{13}C NMR spectra of the *O*-glycan, and inter-residue correlations from NOESY and HMBC spectra.

Sugar residues	$^{1}H/^{13}C$ (ppm)								
	1	2	3	4	5	6	CH_3	NOE	HMBC
→4)-D-Mannitol-6-P	3.87, 3.63	3.92	3.68	4.02	3.89	4.06, 3.92			
A	65.8	73.1	71.9	78.7	71.7	68.3			
→3)-β-D-GlcpNAc-(1→	4.64	3.85	3.76	3.49	3.54	3.94, 3.71	2.10	**A**, H4	**A**, C4
B	103.5	57.11	83.6	71.8	77.5	63.8	25.0		**A**, H4
β-D-GalpNAc-(1→	4.49	3.84	3.74	3.91	3.69	3.77, 3.77	1.99	**B**, H3	**B**, C3
C	104.2	55.21	73.2	70.4	77.7	63.6	25.0		

Supporting references and notes

S1. M. Kanagawa *et al.*, *Cell* **117**, 953 (2004).
S2. S. Kunz, L. Calder, M. B. Oldstone, *Virology* **325**, 207 (2004).
S3. Y. Chen, T. Maguire, R. M. Marks, *J Virol* **70**, 8765 (1996).
S4. D. E. Michele *et al.*, *Nature* **418**, 417 (2002).
S5. S. Kunz, J. M. Rojek, M. Perez, C. F. Spiropoulou, M. B. Oldstone, *J Virol* **79**, 5979 (2005).
S6. A. C. Combs, J. M. Ervasti, *Biochem J* **390**, 303 (2005).
S7. T. Ilg *et al.*, *J Biol Chem* **269**, 24073 (1994).
S8. I. I. Islas-Flores, C. Oropeza, S. M. Hernandez-Sotomayor, *Plant Physiol* **118**, 257 (1998).
S9. M. Rance *et al.*, *Biochem Biophys Res Commun* **117**, 479 (1983).
S10. L. Braunschweiler, R. R. Ernst, *J. Magn. Reson.* **53**, 521 (1983).
S11. A. Bax, D. G. Davis, *J. Magn. Reson.* **65**, 355 (1985).
S12. J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **71**, 4546 (1979).
S13. H. Sato, Y. Kajihara, *Carbohydr Res* **340**, 469 (2005).
S14. K. V. R. Chary, V. K. Rastogi, G. Govil, *Journal of Magnetic Resonance Series B* **102**, 81 (1993).
S15. U. Olsson, K. Lycknert, R. Stenutz, A. Weintraub, G. Widmalm, *Carbohydr. Res.* **340**, 167 (2005).
S16. F. Delaglio *et al.*, *J. Biomol. NMR* **6**, 277 (1995).
S17. B. A. Johnson, R. A. Blevins, *J. Biomol. NMR* **4**, 603 (1994).
S18. R. Barresi *et al.*, *Nat Med* **10**, 696 (2004).
S19. D. S. Kim *et al.*, *Neurology* **62**, 1009 (2004).
S20. C. Diesen *et al.*, *J Med Genet* **41**, e115 (2004).
S21. K. Kobayashi *et al.*, *Nature* **394**, 388 (1998).
S22. N. Nakata, K. Furukawa, D. E. Greenwalt, T. Sato, A. Kobata, *Biochemistry* **32**, 4369 (1993).
S23. W. Xia *et al.*, *Biochemistry* **37**, 16465 (1998).
S24. V. Malhotra, T. Serafini, L. Orci, J. C. Shepherd, J. E. Rothman, *Cell* **58**, 329 (1989).
S25. D. B. de Bernabe *et al.*, *J Biol Chem* **284**, 11279 (2009).
S26. I. Breloy *et al.*, *J Biol Chem* **283**, 18832 (2008).
S27. E. L. McDearmon, A. C. Combs, K. Sekiguchi, H. Fujiwara, J. M. Ervasti, *FEBS Lett* **580**, 3381 (2006).
S28. J. M. Ervasti, A. L. Burwell, A. L. Geissler, *J Biol Chem* **272**, 22315 (1997).