# Identification of Key Functional Residues in the Active Site of Human $\beta$1,4-Galactosyltransferase 7

Ibtissam Talhaoui, Catherine Bui, Rafael Oriol, Guillermo Mulliert, Sandrine Gulberti, Patrick Netter, Michael W H Coughtrie, Mohamed Ouzzine, Sylvie Fournel-Gigleux

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 48, pp. 37342–37358, November 26, 2010


# Identification of Key Functional Residues in the Active Site of Human β1,4-Galactosyltransferase 7

## *A MAJOR ENZYME IN THE GLYCOSAMINOGLYCAN SYNTHESIS PATHWAY**[S]

Received for publication, June 8, 2010, and in revised form, September 14, 2010 Published, JBC Papers in Press, September 15, 2010, DOI 10.1074/jbc.M110.151951

**Ibtissam Talhaoui[‡], Catherine Bui[‡1], Rafael Oriol[§], Guillermo Mulliert[¶], Sandrine Gulberti[‡], Patrick Netter[‡], Michael W. H. Coughtrie[‖], Mohamed Ouzzine[‡], and Sylvie Fournel-Gigleux[‡2]**

*From the [‡]Faculté de Médecine, UMR 7561 CNRS-Université de Nancy I, BP 184, 54505 Vandoeuvre-lès-Nancy, France, the [§]Institut André Lwoff, U1004 INSERM, Université Paris Sud, 94807 Villejuif, France, the [¶]Faculté des Sciences et Techniques, UMR 7036 CNRS-Université de Nancy I, BP 184, 54505 Vandoeuvre-lès-Nancy, France, and the [‖]Division of Medical Sciences, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, Scotland, United Kingdom*

**Glycosaminoglycans (GAGs) play a central role in many pathophysiological events, and exogenous xyloside substrates of β1,4-galactosyltransferase 7 (β4GalT7), a major enzyme of GAG biosynthesis, have interesting biomedical applications. To predict functional peptide regions important for substrate binding and activity of human β4GalT7, we conducted a phylogenetic analysis of the β1,4-galactosyltransferase family and generated a molecular model using the x-ray structure of *Drosophila* β4GalT7-UDP as template. Two evolutionary conserved motifs, $^{163}$DVD$^{165}$ and $^{221}$FWGWGREDDE$^{230}$, are central in the organization of the enzyme active site. This model was challenged by systematic engineering of point mutations, combined with *in vitro* and *ex vivo* functional assays. Investigation of the kinetic properties of purified recombinant wild-type β4GalT7 and selected mutants identified Trp$^{224}$ as a key residue governing both donor and acceptor substrate binding. Our results also suggested the involvement of the canonical carboxylate residue Asp$^{228}$ acting as general base in the reaction catalyzed by human β4GalT7. Importantly, *ex vivo* functional tests demonstrated that regulation of GAG synthesis is highly responsive to modification of these key active site amino acids. Interestingly, engineering mutants at position 224 allowed us to modify the affinity and to modulate the specificity of human β4GalT7 toward UDP-sugars and xyloside acceptors. Furthermore, the W224H mutant was able to sustain decorin GAG chain substitution but not GAG synthesis from exogenously added xyloside. Altogether, this study provides novel insight into human β4GalT7 active site functional domains, allowing manipulation of this enzyme critical for the regulation of GAG synthesis. A better understanding of the mechanism underlying GAG assembly paves the way toward GAG-based therapeutics.**

Glycosaminoglycans (GAGs),[3] the polysaccharide chains of proteoglycans, are distributed at the surface of most cells and in extracellular matrices of virtually every tissue. GAGs are implicated as regulators in many biological events related to intracellular signaling, cell-cell interactions, and tissue morphogenesis. Thus, the pathophysiological roles of GAGs are highly diversified, ranging from mechanical support to intricate effects on various cellular processes such as cell adhesion, differentiation, proliferation, and motility (1). These vital functions depend on the interactions of the GAG chains with a variety of molecules, including growth factors, cytokines, and their receptors, enzymes, including matrix proteases and coagulation factors as well as extracellular matrix proteins (2).

The biosynthesis of GAG chains, which governs the expression of these functions, involves the ordered stepwise addition of a series of monosaccharides derived from UDP-sugars onto the polypeptide backbone of a proteoglycan through the action of specific glycosyltransferases (3). This multistep process starts by the addition of Xyl to a serine residue of the core protein followed by the addition of two Gal and one GlcUA to form the tetrasaccharide linkage region (GlcUAβ1,3Galβ1,3Galβ1,4Xylβ1) common to most GAGs. This tetrasaccharide structure serves as a primer for the formation of heparin/heparan sulfate and chondroitin/dermatan sulfate chains, which is initiated by the attachment of GlcNAc or GalNAc, respectively. The glucosaminoglycan (heparin/heparan sulfate) and the galactosaminoglycan (chondroitin/dermatan sulfate) chains then assemble by the alternating addition of GlcUA and GlcNAc or GlcUA and GalNAc, respectively. GAG chains are further modified by the cooperative action of epimerases and sulfotransferases that add considerable complexity and functionality to the polysaccharide backbone (4, 5).

The enzyme β4GalT7 (xylosylprotein β1,4-galactosyltransferase, EC 2.4.1.133) catalyzes the transfer of a Gal residue provided by UDP-α-D-Gal (UDP-Gal) onto Xyl, a key step in the synthesis of the linkage region of GAG chains (6). Accordingly, we found that phosphorylation of Xyl, preventing β4GalT7

* This work was supported by Agence Nationale de la Recherche, GAG Network Grant ANR-08-PCVI-0023-01, INSERM-University of Dundee International Collaboration Contract C2I, Royal Society International Joint Grant (to M. W. H. C. and S. F.-G.), and the Région Lorraine. Part of this work was performed under the auspices of a European-associated laboratory between CNRS-UHP Nancy I and the University of Dundee.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S2 and Tables S1–S4.

[1] Present address: Catherine Bui's, Institute of Cellular Medicine, Newcastle University, Medical School, Newcastle upon Tyne NE2 4HH, United Kingdom.

[2] To whom correspondence should be addressed. Tel.: 33-3-83-68-39-72; Fax: 33-3-83-68-39-59; E-mail: sfg@medecine.uhp-nancy.fr.

[3] The abbreviations used are: GAG, glycosaminoglycan; β4GalT7, β1,4-galactosyltransferase 7; MN-Xyl, 7-methoxy-2-naphthyl-β-D-xylopyranoside; MN-Xyl-2P, 7-methoxy-2-naphthyl-β-D-xylopyranoside-2-phosphate; 4-MUX, 4-methylumbelliferone-β-D-xylopyranoside; 4-NP-Xyl, 4-nitrophenol-β-D-xylopyranoside; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; PDB, Protein Data Bank.

activity, may act as a key regulatory event in the initiation of GAG synthesis (7). Furthermore, both in vertebrates and invertebrates, defects in this enzyme, which disrupt GAG synthesis, result in severe biological consequences. In humans, point mutations in the *B4GALT7* gene are associated with the progeroid form of Ehlers-Danlos syndrome, characterized by aged appearance, developmental delay, dwarfism, and several connective tissue disorders (8). In the invertebrate *Drosophila melanogaster*, a reverse genetic approach proved β4GalT7 to be essential for viability (9). The structure of the fly protein has recently been solved by x-ray crystallography (10). In *Caenorhabditis elegans*, defective β4GalT7 provokes serious morphological abnormalities during embryonic development (11).

Interestingly, the requirement for a Xyl polypeptide as a starting block to build a GAG chain can be bypassed *in vitro* and *in vivo* by xyloside derivatives consisting of Xyl linked to a hydrophobic aglycone (12). These xylosides serve as substrates for β4GalT7, thus promoting initiation and polymerization of GAG chains. Such GAG precursors hold promise as antithrombotic drugs (13) and anti-amyloid agents (14). On the other hand, inhibitors of GAG synthesis have been proposed as chemotherapeutic agents (15–17). Thus, an understanding of the structure and function of β4GalT7, a pivotal enzyme in GAG biosynthesis, would represent a major step toward the development of GAG-based therapeutics.

The β4GalT7 enzyme belongs to the human β1,4-galactosyltransferase (β4GalT) family, containing seven members that are involved in the formation of Galβ1,4GlcNAc or Galβ1,4Glc linkages in different glycoconjugates and free saccharides (18). Analysis of their substrate specificity revealed major physiological functions and substantial impact in the pathogenesis of diseases. β4GalT1 (lactose synthase), one of the first glycosyltransferases to be cloned and characterized, acts on a nonreducing terminal GlcNAc residue as acceptor substrate, whereas in the presence of α-lactalbumin, the enzyme prefers Glc to GlcNAc (19). β4GalT2 and β4GalT3 catalyze the formation of Galβ1,4GlcNAc bonds in several glycoproteins and specific glycolipids. β4GalT5 and β4GalT6 are involved in the synthesis of lactosylceramide, which plays a major role in the regulation of cell proliferation, adhesion, migration, and angiogenesis. Recently, the importance of β4GalT4 in the construction of keratan sulfate chains, essential in the maintenance of corneal matrix structure, has been discovered (20). Interestingly, the β4GalT genes show sequence homology with those coding for β1,4-*N*-acetylgalactosaminyltransferases (β1,4GalNAcTs), among them several enzymes are involved in chondroitin sulfate GAG chains synthesis.

Several short conserved peptide motifs, possibly involved in the donor- and acceptor-binding sites, have been identified in the human enzymes and among β4GalT orthologs in various animal species (21). Interestingly, these motifs are also conserved between members of the β1,4-glycosyltransferase family, including seven β1,4-galactosyltransferases (β1,4GalTs), two β1,4GalNAcTs, and four chondroitin synthases (19). Human β4GalT7 presents 73% protein sequence similarity with its *Drosophila* ortholog (10), and the seven human β4GalT proteins exhibit 25–55% sequence homology (19). The β4GalT7 is the most distant of the seven members of the β4GalT family (22).

Unlike β4GalT1, which has been subjected to extensive investigation for more than 20 years, our understanding of the mechanism of action of human β4GalT7 is limited. To gain further insight into the structure and function of this enzyme, we conducted a phylogenetic analysis of β1,4GalTs, which identified 48 related sequences containing 8 highly conserved peptide motifs. Using the structural coordinates of *Drosophila* β4GalT7, we modeled the human β4GalT7 structure, predicting that the peptide motifs $^{163}$DVD$^{165}$ and $^{221}$FWGWRGEDDE$^{230}$ are likely to be important for the organization of the active site. Finally, using site-directed mutagenesis and a combination of *in vitro* and *ex vivo* activity assays, we analyzed the functional implications of these β4GalT hallmarks. Our results suggest the involvement of a carboxylate residue in the reaction catalyzed by human β4GalT7 and highlight the central function of Trp$^{224}$ in donor and acceptor substrate binding.

## EXPERIMENTAL PROCEDURES

*Chemicals and Reagents*— 4-Methylumbelliferyl-β-D-xylopyranoside (4-MUX), 4-nitrophenol-β-D-xylopyranoside (4-NP-Xyl), UDP-Gal, UDP-α-D-glucose (UDP-Glc), UDP-α-D-*N*-acetylglucosamine (UDP-GlcNAc), UDP-α-D-*N*-acetylgalactosamine (UDP-GalNAc), UDP-α-D-glucuronic acid (UDP-GlcUA), UDP-α-D-xylose (UDP-Xyl), and anti-mouse IgG-alkaline phosphatase antibodies were provided from Sigma. UDP[$^{14}$C]Gal, UDP[$^{14}$C]GlcUA, UDP[$^{14}$C]Xyl, UDP[$^{14}$C]GalNAc, UDP[$^{14}$C]GlcNAc, and $Na_2[S^{35}]SO_4$ were purchased from PerkinElmer Life Sciences, and UDP[$^{14}$C]Glc was from Amersham Biosciences. 7-Methoxy-2-naphthyl-β-D-xylopyranoside (MN-Xyl) and 7-methoxy-2-naphthyl-β-D-xylopyranoside-2-phosphate (MN-Xyl-2P) were synthesized as described previously (7). Anti-Myc antibodies were from Invitrogen. Cell culture media were provided by Invitrogen. Restriction enzymes and T4 DNA ligase were from New England Biolabs. The eukaryotic expression vector pcDNA3.1(+) and competent One Shot® Top10 *Escherichia coli* cells were from Invitrogen. The bacterial expression vector pET-41a(+) and *E. coli* BL21(DE3) cells were from Novagen-EMD4Biosciences, and the QuikChange site-directed mutagenesis kit was from Stratagene.

*In Silico Retrieval of β4GalT7 Sequences*—Only eukaryote sequences were considered in this study. Twenty nine β4GalT7 homologous sequences from different animal species were retrieved from the NCBI Data Bank by PSI-BLAST (24) with default parameters using B4GT7_HUMAN as seeding sequence (25). Other β4GalT7 homologous sequences were searched (26) through exploration of all genomic and expression sequence tags available and sequences from general (NCBI, DDBJ, and ENSEMBL) or specialized databases like the Joint Genome Institute (Department of Energy) for *Branchiostoma floridae,* using BLASTN and TBLASTN with default parameters (e-value cut off at 0.01). Within each animal species, the expression sequence tag contigs for each new β4GalT7-like sequence were compiled with CAP3 (23). Splice site prediction analysis was achieved at the Berkeley Drosophila genome pro-

ject. The structure of each gene, in terms of exon/intron boundaries, was deduced from several nonexclusive strategies as follows: (i) comparison of the boundaries proposed by Genscan (MIT Server), (ii) comparison of EST sequences to genomic assemblages (scaffolds or contigs), and (iii) comparison of predicted boundaries to those present in known genes. All nucleotide sequences allowing generation of a complete protein were considered, and the assignation of the new sequences to the β4GalT7 family was defined by the presence of the eight peptide motifs (Fig. 1). The 19 new sequences of β4GalT7 reconstructed by these approaches were submitted to EMBL with the following accession numbers: AM231255 *Ciona savigny*; AM231256 *Drosophila ananassae*; AM231257 *Drosophila erecta*; AM231258 *Drosophila mojavensis*; AM231259 *Drosophila sechellia*; AM231260 *Drosophila simulans*; AM231261 *Drosophila yakuba*; AM231262 *Gallus gallus*; AM231263 *Oryzias latipes*; AM231264 *Pan troglodytes*; AM231265 *Schistosoma mansoni*; AM231266 *Tetraodon nigroviridis*; AM231267 *Takifugu rubripes*; FN568101 *Bos taurus*; FN568102 *Glossina morsitans*; FN568103 *Ovis aries*; FN568104 *Rana catesbeiana*; FN568105 *Sus scrofa*; and FN568106 *Squalus acanthias*.

*Phylogenetic Analysis and Sequence Alignment*—Multiple alignments were performed using ClustalW (27) and saved in Pir format. The Pir alignments were used for the selection of informative positions by G-Blocks (28). By this computerized method, 220 informative amino acid positions in eight G-blocks were selected for the phylogeny analyses. Phylogeny was carried out with Phylowin® software (29) using pairwise gap removal, BIONJ, Poisson correction, and 500 bootstrap replicates.

*Molecular Modeling*—A molecular model of the catalytic domain of human β4GalT7 was initially built based on the structure of monomer A of bovine β4GalT1 complexed with UDP-Gal (PDB code entry 1O0R) (31) using Modeler version 8v2 (30). The resulting model was then minimized using Amber simulation program version 8 (32) with 500 iterations of steepest descent and 150 of conjugate gradient. Docking of UDP-Gal-$Mn^{2+}$ molecules in the human β4GalT7 structure was performed manually based on the position of the donor substrate in the structure of bovine β4GalT1. Charges of the atoms of UDP-Gal were calculated using the GAUSSIAN94 package (33) and the HF/6–31 G* basis set. Atom-centered charges were fitted with the antechamber of Amber 8 software package (32). $Mn^{2+}$ parameters available from the Amber website were used to build the model of β4GalT7-UDP-Gal-$Mn^{2+}$ complex by energy minimization (500 iterations of steepest descent and 1500 of conjugate gradient). This model structure was next solvated using an octahedron box. The energy of the water molecules was minimized keeping those of the protein, UDP-Gal and $Mn^{2+}$, fixed, with 1000 iterations of steepest descent and 1500 of conjugate gradients. Subsequently, the whole system was submitted to molecular dynamics, by the mean of an equilibration phase of 20 ps allowing the temperature rising from 0 to 300 K followed by a production phase of 600 ps at 300 K. A molecular model of the catalytic domain of human β4GalT7 was also generated based on the structure of *Drosophila* β4GalT7 linked to an N-terminal peptide of bovine β4GalT1 (PDB code entry 3LW6) using Modeler version 8v2 (30). The resulting model was then minimized using Amber simulation program version 11 (34), as described above. Protein sequence alignments used to generate the models are provided as supplemental material F1.

*Expression Vector Construction*—The human β4GalT7 sequence (GenBank® nucleotide sequence accession number NM_007255) was cloned by PCR amplification from a placenta cDNA library (Clontech), as described previously (7). For the heterologous expression of the human β4GalT7 in mammalian cell lines, the full-length cDNA sequence was modified by PCR to include a KpnI site and a Kozak consensus sequence at the 5′ end (5′-CGGTACCACCATGTTCCCCTCGCGGAGGAAAGCGGGCGC-3′), and a sequence encoding a Myc tag (EQKLIEEDL) and an XbaI site at the 3′ end (5′-GTCTAGATCACAGATCCTCTTCAGAGATGAGTTTCTGCTCGCTGAATGTGCACCAGGGTTGGC-3′). The modified cDNA sequence was then subcloned into the KpnI-XbaI sites of the eukaryotic expression vector pcDNA3.1(+) to produce pcDNA-β4GalT7.

For bacterial expression, a truncated form of β4GalT7 was expressed as a fusion protein with glutathione *S*-transferase (GST). The sequence lacking the codons of the first 60 N-terminal amino acids was amplified from the full-length cDNA using the oligonucleotide sense (5′-CCATGGTCAGGGGACAAGGGCAG-3′) and antisense primers (5′-GCGGCCGCTCACCTGAATGTGCACCA-3′), including NcoI and NotI sites, respectively, and subcloned into the corresponding sites of pET-41a(+) to produce the plasmid pET-β4GalT7. Final vector constructs were sequenced to verify that no errors had been introduced.

Mutations were constructed using the QuikChange site-directed mutagenesis kit, employing pcDNA-β4GalT7 or pET-β4GalT7 as template, and sense and antisense primers are listed in supplemental Table 1. Mutants were systematically checked by sequencing.

The human decorin cDNA sequence (GenBank® accession number NM_001920.3) was cloned by PCR amplification from a placenta cDNA library (Clontech) using sense (5′-ATGAAGGCCACTATCATCCTCCTTCTG-3′) and antisense primers (5′-CTACTTATAGTTTCCGAGTTGAATGGC-3′). For the heterologous expression of the recombinant decorin in eukaryotic cells, the full-length cDNA sequence was modified by PCR to include an AflII site, a Kozak consensus sequence at the 5′ end, a sequence encoding a $His_5$ tag, and an XhoI site at the 3′ end. This sequence was subcloned into pcDNA3.1(+) to produce pcDNA-decorinHis.

*Heterologous Expression of Wild-type and Mutant β4GalT7 in Eukaryotic Cells*—HeLa cells and CHO pgsB-618 cells purchased from the American Type Culture Collection (ATCC) were cultured in Dulbecco's modified Eagle's medium (DMEM) and F-12/DMEM (1:1), respectively, supplemented with 10% (v/v) fetal bovine serum (FBS, Invitrogen), penicillin (100 units/ml)/streptomycin (100 mg/ml), and 1 mM glutamine.

Wild-type and mutant pcDNA-β4GalT7 plasmids were individually transfected into HeLa or CHO pgsB-618 cells grown to 80% confluency, using ExGen 500 reagent (Euromedex, Souffelweyersheim, France) according to the manufacturer's recommendations. The cells were harvested in phosphate-buffered

saline (PBS) 48 h after transfection and centrifuged for 10 min at 5,000 × *g* at 4 °C. The pellet was resuspended in 0.25 M sucrose, 5 mM HEPES buffer (pH 7.4) and sonicated for 5 s. Protein concentration was determined by the method of Bradford (35) prior to SDS-PAGE or activity analyses. Alternatively, transfected cells were transferred to labeling medium prior to GAG extraction, as described below. CHO pgsB-618 cells were transfected with human pcDNA-decorinHis using geneticin G-418 (Sigma) for selection of stable transfectants, as described previously (36), and then used for transient transfection with wild-type and mutant pcDNA-β4GalT7 plasmids.

*Expression and Purification of a Soluble Form of β4GalT7*—To express wild-type and mutant GST-β4GalT7, *E. coli* BL21(DE3) cells were transformed with pET-β4GalT7 plasmids and selected by kanamycin resistance. Recombinants were inoculated in fresh Luria-Bertani (LB) medium and incubated at 37 °C on an orbital shaker (160 rpm) overnight. Each sample was then inoculated into 200 ml of fresh medium at a ratio of 1:100 and incubated in a shaking incubator at 37 °C until an absorbance (*A*) at 600 nm of 0.6–0.8 was reached. Isopropyl β-D-thiogalactopyranoside (Sigma) was then added to a final concentration of 1 mM, and cells were further cultured at 20 °C for 16 h. Bacterial cells were collected by centrifugation and lysed by sonication in buffer containing 50 mM sodium phosphate (pH 7.4), 1 mM phenylmethylsulfonyl fluoride, 1 mM EDTA, 5% (v/v) glycerol, and supplemented with Protease Inhibitor Mixture tablets (Roche Diagnostics). Cell lysates were centrifuged for 5 min at 3,000 × *g* to eliminate cell debris, and the resulting supernatant was centrifuged 15 min at 12,000 × *g*. The supernatant and pellet from this centrifugation were then subjected to 10% (w/v) SDS-PAGE, and gels were stained with Coomassie Brilliant Blue. The supernatant from the last centrifugation containing soluble proteins was used for affinity chromatography purification of GST-β4GalT7 using glutathione-Sepharose™ 4B packed columns (GE Healthcare), following recommendations of the manufacturer, and the amount of purified protein was quantified using Quant-iT™ assay kit and Qubit™ spectrofluorimeter.

*Western Blot Analysis*—Protein samples from HeLa or CHO pgsB-618 cells transfected with wild-type or mutant pcDNA-β4GalT7 plasmids were separated on SDS-10% (w/v) polyacrylamide gels and electrophoretically transferred to Immobilon-P® polyvinylidene difluoride membranes (Millipore). The blot was developed using monoclonal anti-Myc antibodies and alkaline phosphatase-conjugated anti-mouse IgG as secondary antibodies, as described previously (37). The amount of recombinant wild-type and mutant protein expressed in cells was evaluated from immunoblot analysis using a calibration curve established with 15–50 ng of recombinant GST-Myc protein (Pierce) run on the same gel, as described previously (7). Scanning densitometry was performed using Scion 1.63 Image software.

*Determination of Galactosyltransferase Activity*—The *in vitro* assay of β4GalT7 activity was performed as described previously (7). Briefly, reactions were performed in 100 mM sodium cacodylate buffer (pH 7.0) containing 10 mM $MnCl_2$, 1 mM UDP-Gal, 0.05 μCi of UDP-[U-$^{14}$C]Gal, 5 mM 4-MUX, and 40 μg of total cell protein or 0.2 μg of purified protein. Incubations were carried out at 37 °C for 1 h (or 30 min in the case of the purified protein) in a total volume of 50 μl and terminated by the addition of 5 μl of 6 N HCl. A reaction without 4-MUX was used as a control. The incubation mixture was centrifuged at 10,000 × *g* for 10 min at 4 °C, and the supernatant was analyzed by HPLC with a reverse phase $C_{18}$ column (Apollo, 4.6 × 150 mm, 4 μm, Grace Davison Discovery Sciences, Belgium) using a Waters equipment (Alliance Waters e2695) equipped with a Berthold FLowStar radioactivity monitor. The mobile phase was composed of 14% (v/v) acetonitrile and 0.02% (v/v) trifluoroacetic acid in water. Radioactivity associated with the reaction product was quantified by scintillation counting using Ultima Flo™ scintillant mixture (PerkinElmer Life Sciences).

For the determination of glycosyltransferase activity toward various donor (UDP-Glc, UDP-Xyl, UDP-GlcUA, UDP-GalNAc, and UDP-GlcNAc) or acceptor substrates (4MU-Glc, 4-MU-GlcNAc, 4-MU-GalNAc, 4-MU-Gal, 4-NP-Xyl, MN-Xyl, and MN-Xyl-2P), incubations were performed in similar conditions, and reaction products were separated by thin layer chromatography as described previously (38, 39). Briefly, following protein precipitation, the supernatant was loaded onto plates (LK6DF silica gel, 250 mm, Whatman). The plates were developed with *n*-butanol, acetone, acetic acid, aqueous ammonia (28% v/v), and water (70:50:18:1.5:60, v/v). The radioactivity associated with the reaction product was visualized by autoradiography with X-Omat Kodak film (Sigma) for 3 days. The silica gel areas of the reaction product were scraped off and quantified by scintillation counting using Ultima Gold™ scintillant mixture on an LKB spectrophotometer (Packard Instrument Co.).

*Determination of Kinetic Parameters*—Kinetic constants $K_m$ and $V_{max}$ or $k_{cat}$ were determined toward 4-MUX as acceptor substrate by incubating total HeLa cell proteins (40 μg) or purified GST-β4GalT7 (0.2 μg) with increasing concentrations of 4-MUX (0–10 mM), in the presence of a fixed concentration of UDP-Gal (1 mM). The kinetic constants $K_m$ and $V_{max}$ or $k_{cat}$ toward UDP-Gal donor substrate were determined by incubating cell or purified proteins with increasing concentrations of UDP-Gal (0–10 mM), in the presence of a fixed concentration of 4-MUX (5 mM). Kinetic parameters were determined by nonlinear least squares regression analysis of the data fitted to Michaelis-Menten rate equation using the curve-fitter program of Sigmaplot 9.0 (Erkraft, Germany).

*Analysis of GAG Synthesis by $Na_2[^{35}S]SO_4$ Incorporation*—To analyze GAG chain synthesis initiated from exogenously added xyloside, CHO pgsB-618 cells transfected with wild-type, mutant pcDNA-β4GalT7, or empty vector were incubated for 12 h in Fischer's medium containing 10 μCi/ml $Na_2[^{35}S]SO_4$ in the presence or absence of 4-MUX (5 μM). 1 ml of medium was applied to a G-50 column to separate nonincorporated radiolabeled sulfate, and radiolabeled GAGs present in the eluates were quantified by scintillation counting or resolved by SDS-PAGE using Criterion Precast gels (4–12% BisTris, Bio-Rad) and visualized by autoradiography on Amersham Biosciences Hyperfilm MP. For analysis of GAG chain synthesis on decorin core protein, CHO pgsB-618 cells stably expressing human recombinant decorin were transiently transfected with pcDNA-β4GalT7 or mutant plasmids and incubated for 12 h in Fischer's

medium containing 10 μCi/ml [$^{35}$S]$SO_4$, as described above. The recombinant His-tagged decorin was purified from 1 ml of culture medium using the MagneHis Protein Purification System according to the recommendations of the manufacturer (Promega). The radiolabeled purified decorin was analyzed by SDS-PAGE using Criterion Precast gels (4–12% BisTris). After incubation in $H_2O$/methanol/acetic acid (5:1:1) solution, then in Amplify reagent (Amersham Biosciences), the gel was dried and subjected to autoradiography.

For analysis of total GAG synthesis, CHO pgsB-618 cells, 48 h after transfection, were incubated for 8 h in Fischer's medium supplemented with 10% (v/v) fetal bovine serum without sulfate containing 10 μCi/ml $Na_2$[$^{35}$S]$SO_4$, and proteoglycan extraction was performed as described previously (37). Briefly, 1 ml of culture medium was collected and treated with papain (1 mg/ml) for 3 h at 60 °C. The digestion was stopped by heating at 100 °C for 10 min, and the protein was precipitated by centrifugation at 8,000 × *g* for 15 min. Three volumes of 5% potassium acetate (w/v in absolute ethanol) were added to 1 volume of supernatant, and after mixing, the suspension was stored overnight at 4 °C and then centrifuged for 30 min at 8,000 × *g*. The precipitate was resuspended in 500 μl of 0.2 M NaCl and centrifuged for 30 min at 8,000 × *g*. The supernatant was incubated for 2–3 h at 37 °C with 100 μl of cetylpyridinium chloride (5%, w/v) in the presence of 100 μg of shark chondroitin sulfate to facilitate GAG precipitation. The precipitate was collected by centrifugation for 30 min at 8,000 × *g* and washed three times in the cetylpyridinium chloride (5%, w/v), 0.2 M NaCl solution and collected by centrifugation for 10 min at 8,000 × *g*. The precipitate was dissolved in 200 μl of 2.5 M NaCl, 5 volumes of ethanol, and centrifuged for 30 min at 10,000 × *g*, and the final precipitate was dissolved in 200 μl of Tris-HCl (0.2 M, pH 8.0). The radioactivity associated to GAGs was determined by scintillation counting.

*Statistical Analyses*—All data are presented as means ± S.D. of three determinations of at least three sets of experiments. For comparison among multiple groups, one-way analysis of variance followed by Bonferroni post hoc correction test was used. Statistical analyses were performed using GraphPad Prism software (version 4.0).

## RESULTS

To gain insight into potential functional domains of β4GalT7 proteins, we performed a multiple alignment of 54 selected sequences, including the 48 animal β4GalT7 sequences plus six human β4GalTs (B4GT1 to B4GT6), taken as an outgroup (Fig. 1). The sequence comparison highlighted a high degree of similarity, with eight conserved peptide motifs distributed all along the alignment and separated by intermotif amino acid stretches of similar length (Fig. 1). The size of the N-terminal domain ahead of the first conserved motif was longer in the β4GalT1–6 outgroup compared with the 48 β4GalT7 proteins. Furthermore, 17 positions were identical only in the β4GalT7 sequences (Fig. 1, indicated by #) and less well conserved in the outgroup, whereas 32 amino acid positions (Fig. 1, indicated by *) were identical, both in the outgroup and in the β4GalT7 family.

Highly conserved β4GalT amino acids were present in the first seven motifs (Fig. 1) but were particularly abundant in "motif 3" that contains a typical D*X*D signature sequence ($^{163}$DVD$^{165}$ in human β4GalT7) and in "motif 5" containing a second predicted functional motif $^{221}$FWGWGREDDE$^{230}$ (numbered in human β4GalT7). In the crystal structure of bovine β4GalT1 and *Drosophila* β4GalT7, these two polypeptides have been found to be located in the vicinity of the donor substrate UDP-Gal that is common to the β4GalT enzymes (19). Amino acids specific to β4GalT7 sequences are found in "motif 6" that contains a new $Mn^{2+}$-binding motif, H*X*H (10). "Motif 8" exhibits 5 out of the 9 positions highly conserved in the β4GalT7 family, including two Cys residues forming a disulfide bond, as shown in *Drosophila* (Cys$^{300}$–Cys$^{308}$, corresponding to Cys$^{316}$–Cys$^{324}$ in human sequence). In *Drosophila* x-ray structure, this motif following an extended β-sheet has been suggested to compensate for the absence of the N-terminal region present in bovine β4GalT1, which plays an important role in the catalytic activity. Further functional analysis of human β4GalT7 is required to test this assumption.

*Phylogeny of β4GalT Enzymes*—To further define the conserved features of the β4GalT7 family, we performed a phylogenetic analysis based on the data available from the various genome sequencing programs. This approach led to important insight into the organization of this large glycosyltransferase family. Fig. 2 shows the phylogenetic tree built with the 54 sequences identified. The β4GalT sequences could be clearly separated into two main groups, one containing the 48 members of the β4GalT7 family and the other containing the six human β4GalT1–6 enzymes, which were considered as an outgroup. The vertebrate β4GalT7 branch is robust with high bootstrap values and constitutes a very homogeneous group in all vertebrate species from fish to man. As expected, the sequences of invertebrate β4GalT7 have longer branch lengths with low bootstrap values. Thus, within the invertebrate branch, some species do not appear in the expected order, as seen in the case of *H. magnipapillata* and *B. floridae* and urochordata and arthropods branches that are inverted. However, the topology of the phylogenetic tree shows a good fit with the expected evolution order. The β4GalT7 protein is probably present in all metazoa, because it was found in cnidaria, trematods, nematods, arthropods, urochordata, and cephalochordata. The enzyme activity of these proteins (EC 2.4.1.133) has only been experimentally tested in some vertebrates, and in *D. melanogaster* and *C. elegans*, but the presence of the eight highly conserved peptide motifs distributed at well preserved distances all along the protein suggests that the other proteins may also have similar enzyme activities.

*Structural Model of Human β4GalT7 Active Site*—Although the structure for human β4GalT7 has not yet been determined, protein sequence alignment identifies several domains of homology with bovine β4GalT1 (supplemental Fig. 1), for which a number of x-ray structures has been established (31, 40). Thus, we initially created a molecular model of human β4GalT7 using bovine β4GalT1-UDP-Gal-$Mn^{2+}$ (PDB entry code 1O0R) as template. The modeled active site revealed the presence of the $^{163}$DVD$^{165}$ and the $^{221}$FWGWGREDDE$^{230}$ functional domains (numbered in human β4GalT7) in the

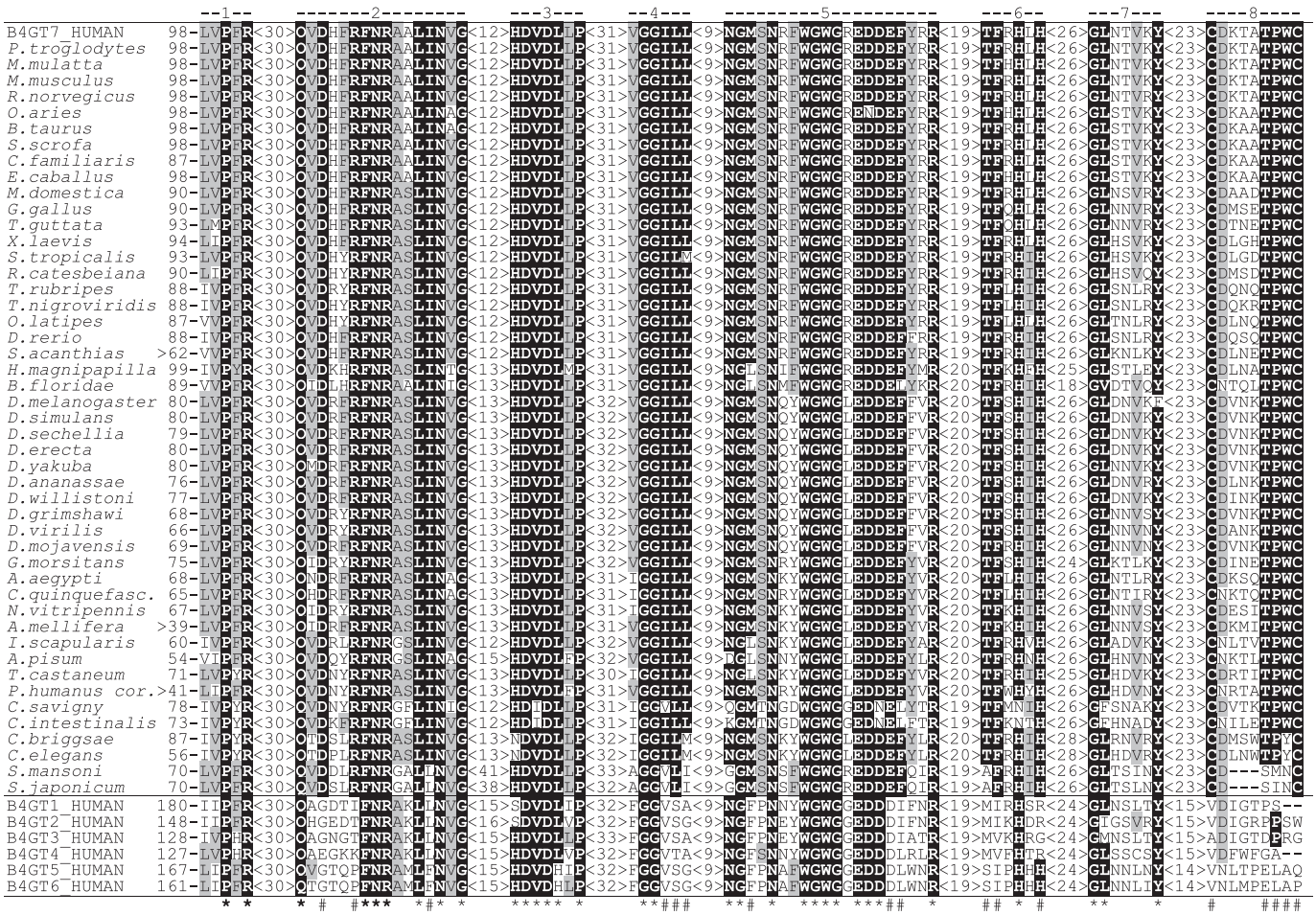

FIGURE 1. **Conserved peptide motifs found in the multiple alignment of the 48 animal β4GalT7 proteins and the human β4GalT1–6 sequences.** Protein and DNA alignments were performed by ClustalW and saved in Pir format. The Pir alignment was used for the selection of 220 informative positions by G-blocks. The eight highly conserved peptide motifs were visually defined within the ClustalW alignment of the 220 selected positions. Positions identical at more than 90% of the sequences are shown as *bold white letters* on a *solid background*. Positions identical at more than 50% are shown as *black letters* on a *gray background*. In the *bottom line*, amino acid positions identical in >90% of all the β4GalT sequences are represented by an *asterisk*, and positions also identical at >90%, but only in the β4GalT7 family are represented by a *number sign*.

vicinity of the donor substrate, UDP-Gal (supplemental Fig. 2). In supplemental Fig. 2.), *panel A* shows predicted interactions between the $^{163}$DVD$^{165}$ motif and UDP-Gal-$Mn^{2+}$, including a hydrogen bond between Asp$^{163}$ and O3 Gal molecule, and a coordination bond between the side chain carboxyl group of Asp$^{165}$ and the $Mn^{2+}$ divalent cations. Hydrogen bond formation between Nϵ1 of Trp$^{224}$ belonging to the $^{221}$FWGWGREDDE$^{230}$ conserved motif and a β-phosphate oxygen atom of UDP is highlighted (*panel B*). We also identified several amino acid residues of human β4GalT7, other than those belonging to the functional peptide regions of interest (listed in supplemental Table 3), which form possible contacts with UDP-Gal-$Mn^{2+}$, including His$^{257}$ in coordination bonding with the divalent cations (supplemental Fig. 2).

Because the crystal of *Drosophila* β4GalT7, which shares 73% protein sequence homology with human β4GalT7 (supplemental Fig. 1), was recently solved (10), we next created a molecular model of the human protein based on this structure. For this purpose, we used the structural coordinates of the catalytic domain of *Drosophila* β4GalT7 in complex with UDP (PDB entry code 3LW6) and proceeded to the docking of UDP-Gal into the active site of the modeled human protein. As shown Fig. 3, the $^{163}$DVD$^{165}$ and the $^{221}$FWGWGREDDE$^{230}$ functional motifs were found in a similar position with regard to the donor substrate UDP-Gal, when compared with the model of human β4GalT7 based on bovine β4GalT1. Asp$^{163}$ was in hydrogen bond distance to O3 of the Gal molecule, and a coordination bonding was predicted between the side chain carboxyl group of Asp$^{165}$ and the $Mn^{2+}$ divalent cation (Fig. 3*A*). Moreover, in addition to the coordination bond between Nϵ2 of His$^{257}$ and $Mn^{2+}$, a coordination bond between Nδ1 of His$^{259}$ and the divalent cation was identified. Indeed, $^{257}$HVH$^{259}$ corresponds to a new metal-binding motif (10), which is strictly conserved among β4GalT7 proteins, as shown by multiple sequence alignment (see Fig. 1). Furthermore, modeling the UDP-Gal-binding site of human β4GalT7 containing the $^{221}$FWGWGREDDE$^{230}$ motif predicted hydrogen bond formation between the Nϵ1 of Trp$^{224}$ and β-phosphate oxygen atom of UDP of the donor substrate molecule and between Glu$^{227}$ and the O4 of the Gal moiety (Fig. 3*B*). Additional bonds between human β4GalT7 and UDP-Gal were also identified (listed in supplemental Table 4). As found in *Drosophila*, Tyr$^{194}$ (corresponding to Tyr$^{277}$) and Arg$^{250}$ (corresponding to Arg$^{266}$) were predicted to form a hydrogen bond with an oxy-

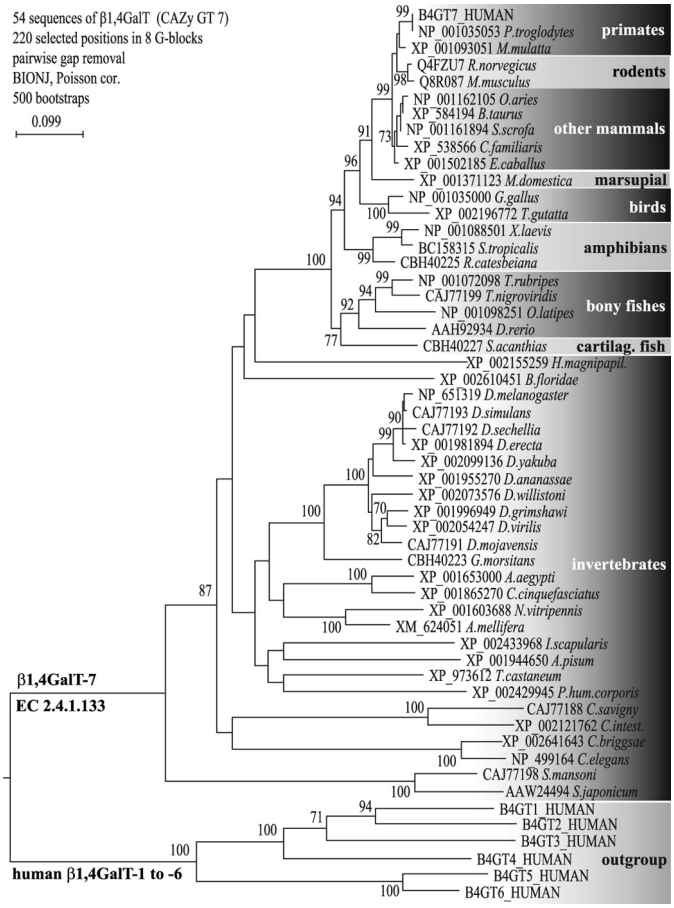


FIGURE 2. **Neighbor-joining phylogenetic tree of the 54 β4GalT sequences.** The robustness of the tree branches was tested by bootstrap calculations from 500 sets of data. Bootstrap values higher than 70% are considered significant and are reported at the *left* of each node. The *scale bar* represents the number of substitutions per site for a unit branch length.

gen atom of the β-phosphate oxygen atom of UDP. In addition, we identified potential interactions between $Tyr^{199}$ and O2 and O3 of Gal that are not predicted in the case of its counterpart ($Phe^{182}$) in *Drosophila*.

*In Vitro Characterization of β4GalT7 Mutants Expressed in Eukaryotic Cells*—To address the functional contribution of two main active site conserved regions of β4GalT7 located in motif 3 and motif 5, we conducted a systematic site-directed mutagenesis approach. Individual residues of these two peptide regions were subjected to conservative and nonconservative substitutions to generate a library of 21 mutants. In a first set of experiments, wild-type and mutant β4GalT7 enzymes were expressed in HeLa cells and tested for *in vitro* galactosyltransferase activity using 4-MUX as acceptor substrate. The recombinant enzymes were expressed as Myc-tagged proteins, to quantify the level of expression using recombinant GST-Myc protein as standard, as described under "Experimental Procedures".

In human β4GalT7, the canonical D*X*D motif ($^{163}$DVD$^{165}$) corresponds to a signature characteristic for many glycosyltransferases and other enzymes that use nucleotide substrates and require a divalent metal ion for activity. To gain insight into the functional importance of the $^{163}$DVD$^{165}$ motif in this enzyme, mutants were engineered in which each residue was replaced by Ala and by Glu in the case of $Asp^{163}$ and $Asp^{165}$. The consequences of these mutations were evaluated on the *in vitro* galactosyltransferase activity of β4GalT7, and the apparent $K_m$ value of wild-type and mutant enzymes was determined toward UDP-Gal and 4-MUX. Immunoblot analysis of HeLa cells expressing the mutants showed that they were produced at a similar level to that of the wild-type protein (Fig. 4*A*). The activity of the recombinant enzymes was normalized to the amount of expressed protein, and values are presented in Fig. 4*B*. The full-length form of human β4GalT7 exhibited high activity up to 4 $\mu$mol·min$^{-1}$·mg protein, although it was undetectable in mock-transfected cells (Fig. 4*B*). On the other hand, replacement of each aspartate of the $^{163}$DVD$^{165}$ motif to Ala completely abolished the *in vitro* enzyme activity, emphasizing the prevalent function of these residues. Interestingly, although substitution of $Asp^{163}$ to Glu did not overcome the loss of activity observed for the Ala mutant, D165E exhibited up to 47% activity compared with wild-type enzyme, allowing the determination of kinetic parameters. The apparent $K_m$ values of the D165E mutant toward UDP-Gal and 4-MUX were about 3.8- and 5.5-fold, respectively, higher than the wild-type enzyme (Table 1). These results indicate that the carboxyl side chain of Glu could replace the negative charge of Asp in position 165, to some extent, although this modification reduced the affinity of the enzyme toward both donor and acceptor substrates. Furthermore, substitution of $Val^{164}$ by an Ala residue did not produce significant changes either in activity or in $K_m$ values. These results are in agreement with the structure of *Drosophila* β4GalT7 and the modeled human β4GalT7 active sites, indicating that the Val residue does not directly or indirectly interact with UDP-Gal, in contrast to the neighboring Asp of the $^{163}$DVD$^{165}$ motif (Fig. 3*A*).

In the next series of experiments, we explored the function of the $^{221}$FWGWGREDDE$^{230}$ region located in motif 5, which exhibits a high degree of conservation among the human β4GalT family members, as well as among β4GalT7-related sequences sampled from distant animal species (see Fig. 1). A series of mutants exhibiting conservative and nonconservative substitutions at each position of the peptide motif were engineered, and immunoblot analysis indicated that the mutants were produced at similar level to that of the wild-type protein in HeLa cells (Fig. 4*A*). Replacement of the first residue of the motif, $Phe^{221}$, by Ala did not affect the *in vitro* galactosyltransferase activity of β4GalT7 (Fig. 4*C*). However, although this mutant presented a similar $K_m$ value to that of the wild-type enzyme toward UDP-Gal, the affinity toward 4-MUX was strongly reduced, as indicated by a 13-fold increase in $K_m$ (Table 1). This result suggests that introduction of Ala at position 221 may affect binding of 4-MUX and/or the organization of the acceptor site but not of the donor substrate-binding site. On the other hand, the conservative substitution $Phe^{221}$ to Tyr led to a mutant that exhibited a similar activity to that of the wild-type, with $K_m$ values toward acceptor and donor substrates in the same range as β4GalT7 (Table 1), indicating that Phe and Tyr residues are interchangeable at this position. Concordantly, either Phe or Tyr was found at a position equivalent to 221 in related β4GalT7 sequences (see Fig. 1). The neighboring $Trp^{222}$ residue is invariant among all β4GalT sequences

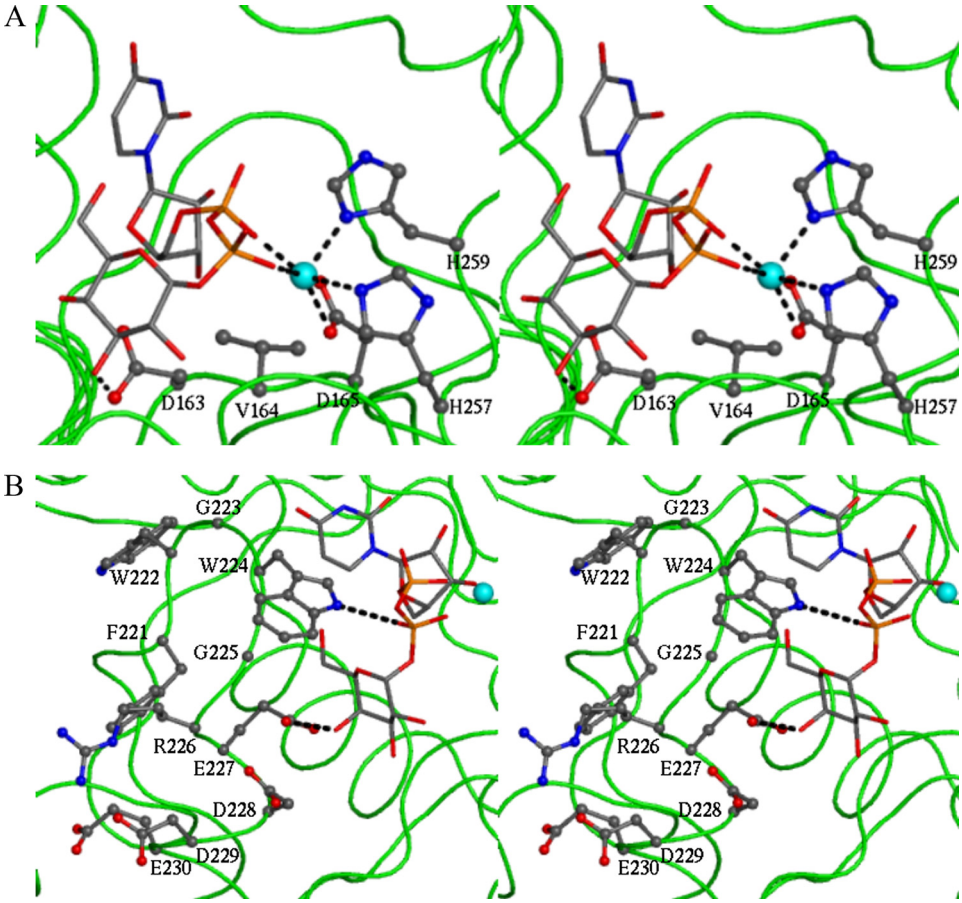


FIGURE 3. **Stereo view of the molecular model of the human β4GalT7 structure in complex with the donor substrate.** The UDP-Gal binding domain was modeled by energy minimization and molecular dynamics calculations based on the structure of *Drosophila* β4GalT7 complexed with UDP (PDB code 3LW6). Representation of the interactions of $^{163}$DVD$^{165}$ and $^{257}$HLH$^{259}$ motifs (*A*) and of $^{221}$FWGWGREDDE$^{230}$ (*B*) motifs with UDP-Gal molecule is shown. The α-carbon trace of the protein is shown in *green*; UDP-Gal is presented in colored "*sticks*," and $Mn^{2+}$ divalent cation is colored in *cyan*. The residues of the $^{163}$DVD$^{165}$, $^{257}$HLH$^{259}$, and $^{221}$FWGWGREDDE$^{230}$ motifs are shown as "*ball* and *sticks*." For simplicity, hydrogen atoms and atoms of the protein main chain other than α-carbons are not shown. *Dashed lines* indicate the predicted interactions of UDP-Gal-$Mn^{2+}$ with the side chain of Asp$^{163}$, Asp$^{165}$, His$^{257}$, and His$^{259}$ residues in *A* and between the β-phosphate group of UDP-Gal and Trp$^{224}$ and between galactose O4 of UDP-Gal and Asp$^{227}$ in *B*. The atom coordinates are those of the average structure obtained during the production phase of molecular dynamics calculation, as indicated under "Experimental Procedures." Figures are drawn with PyMOL molecular visualization system.

analyzed in this study. As shown in Fig. 4*C*, a sharp drop in activity was observed upon replacement of this residue by Ala, whereas the conservative mutation to Phe did not affect the enzyme activity. Furthermore, the W222F mutation had no significant effect on the $K_m$ value toward donor or acceptor substrate. Concordantly, the structural model indicates that the Trp$^{222}$ residue points out of the catalytic center of β4GalT7. Substitution of Gly residues with Ala at positions 223 and 225 produced different effects on the *in vitro* galactosyltransferase activity of β4GalT7. The G223A mutant was about 40% less active than the wild-type enzyme (Fig. 4*C*), and it exhibited over 2-fold increase in apparent $K_m$ values toward 4-MUX (Table 1), whereas the G225A mutant did not display any *in vitro* activity.

Examination of *Drosophila* and modeled human β4GalT7 structures highlighted the role of Trp$^{207}$ and Trp$^{224}$, respectively, as a key active site residue, predicted to interact with the β-phosphate group of the donor substrate (Fig. 3*B*). Thus, we generated a series of substitutions at this position and analyzed the *in vitro* galactosyltransferase activity of the different recombinant human β4GalT7 proteins. Replacement of Trp$^{224}$ by Ala led to a completely inactive protein, suggesting that the presence of a Trp is indeed critical at this position (Fig. 4*C*). Trp$^{224}$ was also converted to Phe or His to investigate the influence of the phenyl or imidazole group of these amino acids, in place of the indole side chain of Trp. Replacement of Trp$^{224}$ by Phe was able to recover about 9% of the wild-type enzyme activity, and W224H activity was about 33% that of β4GalT7. The very low activity observed with the W224F mutant did not allow determination of kinetic parameters. Characterization of the W224H mutant revealed a 4.4-fold decrease in apparent $K_m$ values toward UDP-Gal together with a large increase in $K_m$ values toward the acceptor substrate 4-MUX (Table 1), suggesting that Trp$^{224}$ may be involved not only in donor but also acceptor substrate interactions. Altogether, mutations of the N-terminal hydrophobic amino acids of the $^{221}$FWGWGREDDE$^{230}$ motif revealed a moderate effect of Phe$^{221}$, Trp$^{222}$, and Gly$^{223}$ on enzyme properties, whereas Trp$^{224}$ appears to play an essential role in the organization and function of the catalytic center.

The $^{221}$FWGWGREDDE$^{230}$ motif contains a stretch of four conserved acidic amino acid residues ($^{227}$EDDE$^{230}$). We also investigated the role of this motif by site-directed mutagenesis and kinetic analysis of the resulting recombinant enzymes. Substitution of Glu with Ala at position 227 produced an inactive enzyme, supporting the functional importance of this residue, whereas replacement by an Asp residue resulted in retention of ~50% galactosyltransferase activity compared with the wild-type enzyme (Fig. 4*C*). Furthermore, evaluation of the kinetic properties of E227D showed that this mutant displayed similar apparent $K_m$ values toward both donor and acceptor substrates, compared with the wild-type enzyme (Table 1). Similarly, replacement of Asp$^{228}$ with Ala abolished activity, whereas the D228E mutant expressed in HeLa cells retained about 40% activity of the wild-type enzyme. This mutant exhibited a $K_m$ value toward UDP-Gal similar to the wild-type, whereas an 8.0-fold increase in apparent $K_m$ was observed toward the acceptor substrate (Table 1). In the x-ray structure of bovine β4GalT1 and of *Drosophila* β4GalT7, Asp$^{318}$ and Asp$^{211}$ (equivalent to Asp$^{228}$ in human β4GalT7) are located in the vicinity of the O4 atom of the xylose acceptor substrate and are suggested to act as

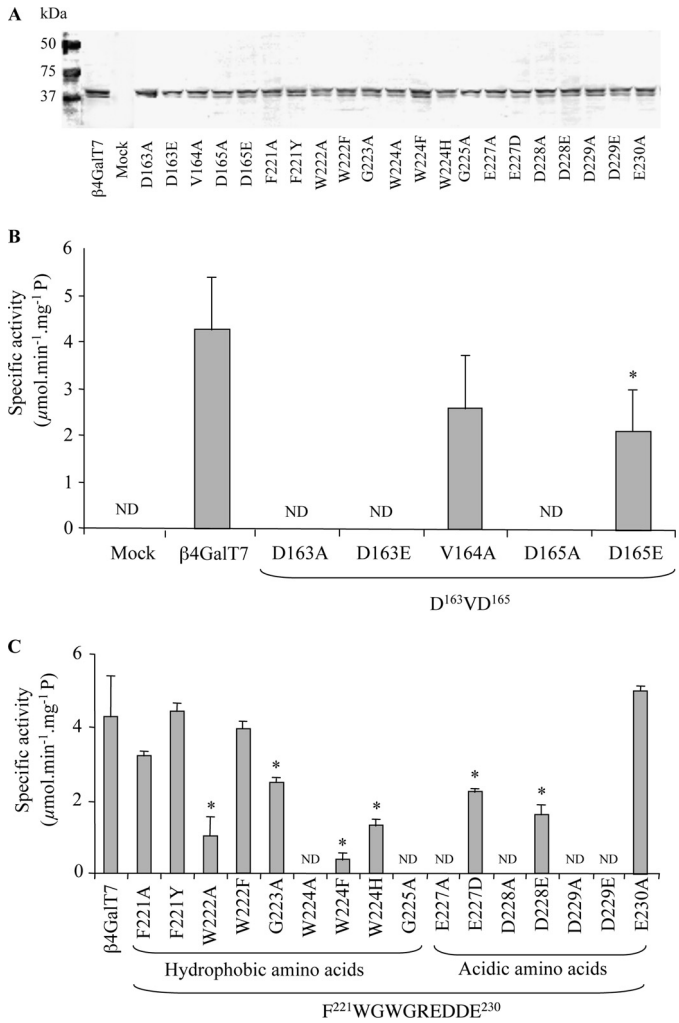


FIGURE 4. **Expression and activity of recombinant wild-type and mutated β4GalT7 expressed in HeLa cells.** *A,* SDS-PAGE and immunoblot analysis of wild-type and mutant β4GalT7 recombinant proteins probed with anti-Myc antibodies. 15 μg of total protein were loaded in each lane. *B* and *C,* galactosyltransferase activity of wild-type and mutant β4GalT7 recombinant enzymes. The reaction was performed using 40 μg of total protein in the presence of 1 mM UDP-Gal and 0.05 μCi of UDP[U-$^{14}$C]Gal as donor substrate and 5 mM 4-MUX as acceptor substrate, as indicated under "Experimental Procedures." The amount of expressed β4GalT7 protein was quantified using a calibration curve established with GST-Myc protein immunostained under the same conditions. Values are expressed as μmol per min per mg of β4GalT7, and results are the mean ± S.D. of three independent experiments (*, $p < 0.05$, *versus* transfected cells with wild-type β4GalT7 cDNA).

**TABLE 1**

**Apparent dissociation constants of wild-type and mutant β4GalT7 expressed in HeLa cells toward UDP-Gal and 4-MUX**

HeLa cells were transiently transfected with wild-type and mutant pcDNA-β4GalT7 plasmids, and enzymatic assays were performed on cell lysates using 40 μg of total protein incubated with increasing concentrations of 4-MUX (0–10 mM) at a constant concentration of UDP-Gal (1 mM) or with increasing concentrations of UDP-Gal (0–10 mM) at a constant concentration of 4-MUX (5 mM). Kinetic parameters were determined by nonlinear least squares regression analysis of the data fitted to the Michaelis-Menten rate equation using Sigmaplot 9.0. The results are the means ± S.D. of three independent assays. ND indicates no kinetic parameters could be determined using excess acceptor or donor substrate.

| Enzyme | $K_m$ | |
|---|---|---|
| | UDP-Gal | 4-MUX |
| | mM | |
| Wild type | 0.31 ± 0.07 | 0.32 ± 0.08 |
| V164A | 0.48 ± 0.10 | 0.53 ± 0.14 |
| D165E | 1.17 ± 0.13* | 1.76 ± 0.25* |
| F221A | 0.36 ± 0.04 | 4.22 ± 0.40* |
| F221Y | 0.33 ± 0.03 | 0.38 ± 0.05 |
| W222A | 0.21 ± 0.01 | ND |
| W222F | 0.45 ± 0.05 | 0.16 ± 0.05 |
| G223A | 0.45 ± 0.06 | 0.73 ± 0.10* |
| W224H | 0.07 ± 0.01* | >4 |
| E227D | 0.47 ± 0.09 | 0.52 ± 0.17 |
| D228E | 0.47 ± 0.08 | 2.57 ± 0.11* |
| D230A | 0.43 ± 0.03 | 0.57 ± 0.08 |

*$p < 0.05$.

the catalytic base (41). Accordingly, the bulkiness of Glu introduced in place of Asp at position 228 in β4GalT7 may produce steric hindrance, contributing to the observed increase in $K_m$ value toward 4-MUX. Replacement of the next Asp residue located at position 229 to either Ala or Glu led to completely inactive enzymes *in vitro*, supporting the critical importance of this residue, although the lack of activity did not allow kinetic studies to be carried out (Fig. 4*C*). However, because this residue was not predicted to be located in the vicinity of the donor substrate, it is not likely to be implicated in direct interactions with UDP-Gal. By contrast, mutation of Glu$^{230}$ to Ala did not impair enzyme activity and had little impact on $K_m$ values (Table 1), indicating a less important function for the last residue of the $^{221}$FWGWGREDDE$^{230}$ peptide region. Although it was found to be conserved among all β4GalT7 family members, this Glu residue was replaced by Asp in the β4GalT1–6 enzymes.

*Expression, Purification, and Kinetic Properties of β4GalT7 Mutants Expressed in E. coli*—In the next series of experiments, we developed an *E. coli* expression system devoted to the characterization of purified β4GalT7 mutants. Because the conserved Trp$^{224}$, Glu$^{227}$, and Asp$^{228}$ were thought to be important for substrate binding and/or catalysis, based on structural predictions and site-directed mutagenesis studies, we conducted detailed kinetic analyses of point mutants at these positions. The wild-type β4GalT7 and selected mutants were produced as truncated fusion proteins lacking the 60 N-terminal amino acids (including the transmembrane domain and part of the stem region) linked to GST. The recombinant enzymes were all purified by affinity chromatography using the same conditions, and expression levels (about 2.5 mg of protein/liter of culture) of the mutants were similar to those of the wild-type (Fig. 5). Kinetic analyses indicated that the wild-type β4GalT7 was highly active with a $k_{cat}$ up to 1.5 $s^{-1}$ and an efficiency of about 5 $mM^{-1}$ $s^{-1}$ (Table 2). Furthermore, the β4GalT7 enzyme purified from *E. coli* cells displayed a $K_m$ value toward either donor or acceptor substrate in the same range as the membrane-bound enzyme expressed in HeLa cells (see Table 1). This indicates that expression of the catalytic domain of β4GalT7 with the GST fusion partner did not substantially affect substrate binding and catalytic properties of this enzyme.

Structural data combined with our systematic site-directed mutagenesis study of the $^{221}$FWGWGREDDE$^{230}$ motif suggested a key role of Trp$^{224}$ in substrate binding, leading us to determine the kinetic properties of purified β4GalT7 enzyme modified at this position and to evaluate possible changes in specificity upon mutation. The W224A and W224F mutants produced and purified from *E. coli* cells were inactive, confirming the functional importance of the conserved Trp$^{224}$ residue. It is noteworthy that the purified W224H mutant displayed a

dramatic change in kinetic behavior compared with the wild-type β4GalT7 enzyme, characterized by a large increase in affinity toward the donor substrate UDP-Gal (15 times lower $K_m$ value) as indicated in Table 2. Because the $k_{cat}$ value for this mutant was reduced, the efficiency of the wild-type enzyme and of the W224H mutant were in the same range (about 5 $mM^{-1}$ $s^{-1}$). On the other hand, the affinity of the purified W224H mutant toward 4-MUX was reduced to such an extent that kinetic parameters toward the acceptor substrate could not be determined.

We next examined whether substitution of Trp at position 224 with His may affect the specificity of β4GalT7 toward either donor or acceptor substrates. To that aim, the activity of purified enzymes was first determined toward a series of UDP-sugar analogs, *i.e.* UDP-Gal, UDP-Glc, UDP-Xyl, UDP-GlcUA, UDP-GalNAc, and UDP-GlcNAc. The wild-type β4GalT7 was able to catalyze the transfer of Glc and Xyl from the corresponding donor substrate onto 4-MUX, although with lower activity (11- and 27-fold, respectively) compared with its galactosyltransferase activity (Fig. 6*A*). As expected from kinetic studies, the W224H mutant exhibited a large decrease in galactosyltransferase activity compared with wild-type (about 10 times). Notably, activity of this mutant was undetectable toward UDP-Glc and UDP-Xyl. The bulkier UDP-sugars, *i.e.* UDP-GlcUA, UDP-GalNAc, and UDP-GlcNAc did not serve as donor substrate for the wild-type enzyme or the W224H mutant (data not shown). These findings indicate that we engineered a W224H variant of β4GalT7 that displays a large increase in affinity toward UDP-Gal compared with the wild-type enzyme and shows no detectable activity toward other UDP-sugars.

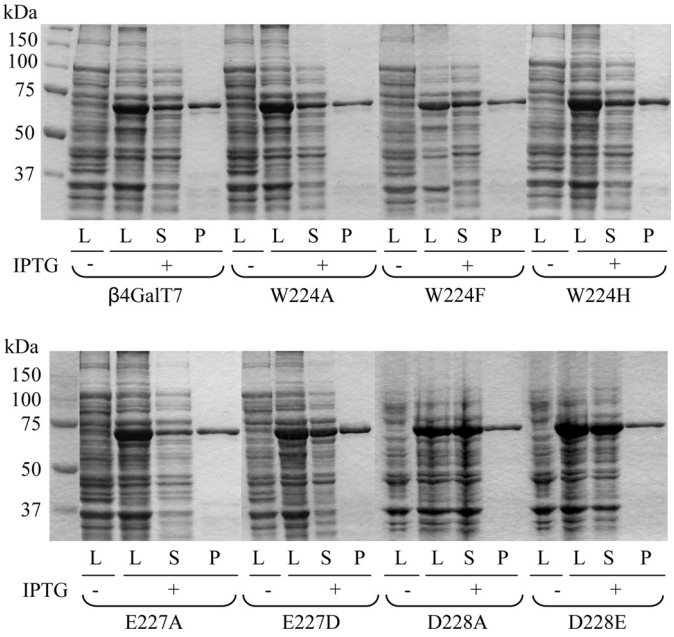


FIGURE 5. **Expression and purification of wild-type and mutated β4GalT7 expressed as GST fusion proteins in *E. coli*.** The wild-type β4GalT7, $Trp^{224}$, $Glu^{227}$, and $Asp^{228}$ mutants were expressed in *E. coli* and purified by glutathione-Sepharose™ 4B affinity chromatography, as described under "Experimental Procedures." Proteins were analyzed by SDS-PAGE on 10% acrylamide gels and stained with Coomassie Brilliant Blue. For each mutant, shown from *left to right*, cell lysate from uninduced and isopropyl β-D-thiogalactopyranoside (*IPTG*)-induced bacteria (*L*, 25 μg), soluble protein from isopropyl β-D-thiogalactopyranoside-induced bacteria (*S*, 12 μg), and purified fusion protein (*P*, 3 μg) are shown. The molecular mass markers are shown in the *1st lane* of each gel.

Furthermore, we analyzed the specificity of wild-type and W224H mutant toward a series of potential acceptor molecules. The wild-type β4GalT7 was inactive toward glycoside derivatives of 4-MU (*i.e.* 4-MU-GalNAc, 4-MU-GlcNAc, 4-MU-Glc, 4-MU-GlcUA, and 4-MU-Gal) tested as acceptor substrates, indicating a strict selectivity toward the xyloside 4-MUX, in agreement with the critical role of this enzyme in GAG chain synthesis (data not shown). As wild-type β4GalT7, the W224H mutant was only active toward 4-MUX. This prompted us to determine the activity of β4GalT7 and the W224H mutant toward xyloside analogs linked to various aglycone moieties. As shown in Fig. 6*B*, wild-type β4GalT7 exhibited higher activity toward MN-Xyl compared with 4-MUX and 4-NP-Xyl and was inactive toward MN-Xyl-2P. The use of a purified form of β4GalT7 showed that the presence of a phosphate group at the position 2 of a xyloside substrate (MN-Xyl-2P) precludes galactosyltransferase activity, in agreement with our previous data (7). Remarkably, the W224H mutant displayed a different substrate profile compared with the wild-type enzyme. Indeed, this mutant exhibited a higher activity toward 4-MUX than for MN-Xyl (Fig. 6*B*). These results suggest that the side chain of $Trp^{224}$ interacts with the aglycone component of the acceptor substrate and that its mutation to His modifies this interaction. These results support the idea that $Trp^{224}$ is a

**TABLE 2**

**Steady-state kinetic parameters of wild-type and β4GalT7 mutants purified from *E. coli***

Kinetic parameters toward donor and acceptor substrates were evaluated from initial velocity values of the reaction using 0.2 μg of purified GST-β4GalT7 protein at a constant amount of 4-MUX (5 mM) in the presence of increasing concentrations of UDP-Gal (0–10 mM) or increasing concentrations of 4-MUX (0–10 mM) in the presence of a fixed concentration of UDP-Gal (1 mM). Kinetic parameters were determined by nonlinear least squares regression analysis of the data fitted to Michaelis-Menten rate equation using Sigmaplot 9.0. The results are the mean values of three independent determinations ± S.D. on assays performed in triplicate. ND indicates no kinetic parameters could be determined using excess acceptor or donor substrate and greater than 0.5 μg of enzyme.

| Enzyme | UDP-Gal | | | 4-MUX | | |
|---|---|---|---|---|---|---|
| | $K_m$ | $k_{cat}$ | $k_{cat}/K_m$ | $K_m$ | $k_{cat}$ | $k_{cat}/K_m$ |
| | mM | $s^{-1}$ | $mM^{-1}s^{-1}$ | mM | $s^{-1}$ | $mM^{-1}s^{-1}$ |
| β4GalT7 | 0.28 ± 0.13 | 1.52 ± 0.52 | 5.43 | 0.33 ± 0.03 | 1.79 ± 0.10 | 5.42 |
| W224A | ND | ND | | ND | ND | |
| W224F | ND | ND | | ND | ND | |
| W224H | 0.02 ± 0.01* | 0.10 ± 0.08* | 5.00 | ND | ND | |
| E227A | ND | ND | | ND | ND | |
| E227D | 0.10 ± 0.01 | 0.46 ± 0.06* | 4.6 | 0.74 ± 0.31 | 0.13 ± 0.03* | 0.17 |
| D228A | ND | ND | | ND | ND | |
| D228E | ND | ND | | ND | ND | |

*$p < 0.05$.

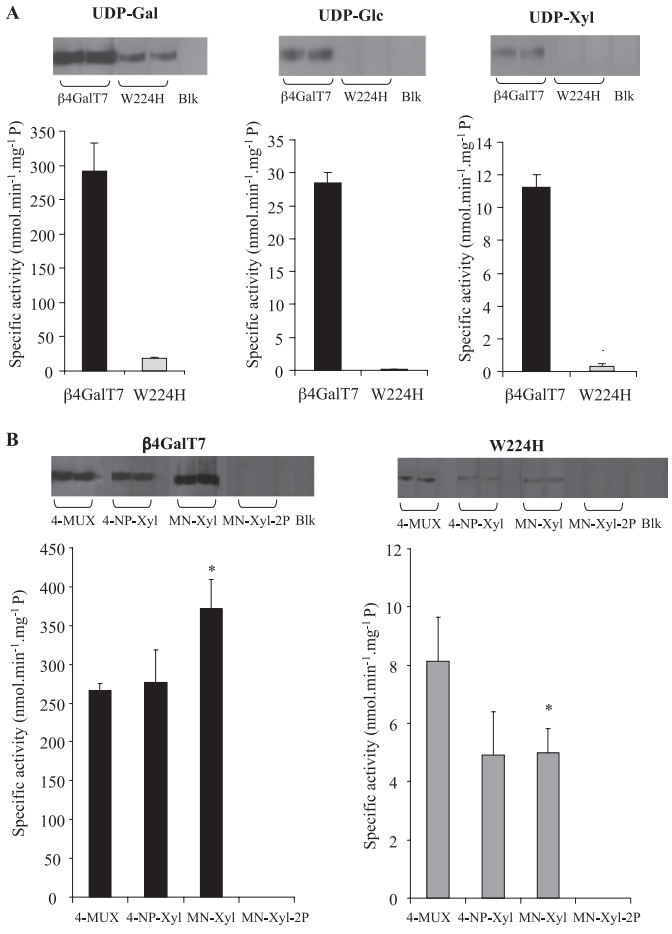


FIGURE 6. **Specificity of wild-type β4GalT7 and the W224H mutant purified from *E. coli*.** Activity of purified wild-type (*WT*) and mutant enzymes was evaluated in standard conditions using 5 mM acceptor substrate and 0.04 mM UDP-sugar (0.06 μCi of UDP-[U-$^{14}$C]sugar) donor substrate in the presence of 0.2 μg of purified protein. The reaction products were separated by thin layer chromatography, visualized by autoradiography (shown in *inset*), and quantified by scintillation counting. The rate values are the means of three experiments. *A,* activity of wild-type β4GalT7 and the W224H mutant toward various UDP-sugars and 4-MUX as acceptor substrate; *B,* activity of wild-type β4GalT7 and the W224H mutant toward various xyloside derivatives using UDP-Gal as donor substrate. Results are the means ± S.D. of three assays on two independent experiments. *B,* *, $p < 0.05$, *versus* activity toward 4-MUX.

key residue for donor but also acceptor substrate binding and that substitution of this residue with His is able to modulate the specificity of β4GalT7 toward acceptor xyloside molecules.

The mechanism of most inverting glycosyltransferases requires one acidic amino acid that activates the acceptor hydroxyl group by deprotonation (42). In *Drosophila* β4GalT7, Asp$^{211}$ (equivalent to Asp$^{228}$ in human β4GalT7) has been suggested to act as the catalytic base (10). In bovine β4GalT1, Asp$^{318}$ and possibly Glu$^{317}$ (equivalent to Asp$^{228}$ and Glu$^{227}$ in human β4GalT7) have been proposed as possible candidates to play such a role (40, 41). On the other hand, Glu$^{317}$ has also been suggested to interact with the Gal residue of the nucleotide sugar donor substrate (41). To further define the role of these carboxylate residues in β4GalT7, Glu$^{227}$ and Asp$^{228}$ were subjected to conservative and nonconservative substitutions, and the resulting recombinant proteins were purified from the *E. coli* expression system, as was the wild-type protein (Fig. 5). Replacement of Glu$^{227}$ with Ala yielded an inactive enzyme, whereas E227D displayed detectable activity. Determination of the kinetic parameters of this mutant revealed about a 3-fold decrease in $k_{cat}$ value with no modification in affinity, as evidenced by no significant change in $K_m$ toward donor or acceptor substrates.

We next examined the activity of recombinant Asp$^{228}$ mutants expressed and purified from *E. coli*. Replacement of this residue with Ala resulted in a complete loss of activity for the purified enzyme. Although a residual activity could be detected for the D228E mutant (data not shown), the high $K_m$ value toward 4-MUX did not allow determination of kinetic parameters. This result highlighted the functional importance of Asp$^{228}$ and indicates that the conservative mutation to Glu also severely affected enzyme activity. Altogether, this set of data confirms that the presence of these acidic residues is essential at the catalytic center.

*Effect of Trp$^{224}$, Glu$^{227}$, and Asp$^{228}$ Mutations on ex Vivo Galactosyltransferase Activity of β4GalT7 toward 4-MUX*—We next designed a series of analyses to determine whether Trp$^{224}$, Glu$^{227}$, and Asp$^{228}$ residues were important for functionality *ex vivo*. For this purpose, CHO pgsB-618 cells, defective in galactosyltransferase activity and exhibiting dramatically impaired GAG synthesis (43), were transfected with wild-type and mutant β4GalT7 cDNAs. In the first instance, we evaluated the capacity of the recombinant wild-type and mutant enzymes to prime GAG chain synthesis on the exogenous β-D-xyloside, 4-MUX, in deficient cells. For this purpose, [$^{35}$S]$SO_4$ incorporation into GAG chains assembled onto 4-MUX was quantified by liquid scintillation counting, and radiolabeled GAGs were subjected to SDS-PAGE analysis. Results shown in Fig. 7*A* indicate that expression of wild-type β4GalT7 in CHO pgsB-618 cells was able to promote GAG chain synthesis from 4-MUX *ex vivo*, as indicated by a 7.5-fold increase in sulfate incorporation compared with untreated or mock-transfected cells. In confirmation of this finding, SDS-PAGE analysis revealed the presence of radiolabeled GAG chains in β4GalT7-transfected cells upon treatment with 4-MUX, whereas no radioactivity could be detected in mock-transfected cells or when addition of the xyloside was omitted (Fig. 7*B*).

To test the functional importance of the Trp$^{224}$ residue in this *ex vivo* model system, we evaluated the ability of the β4GalT7 mutated at this position to initiate GAG synthesis from 4-MUX in CHO pgsB-618 cells. As expected, W224A and W224F mutants that were inactive or exhibited low *in vitro* activity failed to prime GAG synthesis in deficient cells (Fig. 7, *A* and *B*). Interestingly, sulfate incorporation as well as SDS-PAGE analysis of GAG chains in β4GalT7-deficient CHO pgsB-618 cells transfected with W224H showed that this mutant did not initiate GAG synthesis from 4-MUX *ex vivo*. Although this mutant exhibited *in vitro* activity, this observation is consistent with the large increase in the affinity of β4GalT7 toward the acceptor substrate 4-MUX.

We also compared the consequences of mutating the Glu$^{227}$ and Asp$^{228}$ residues on β4GalT7 galactosyltransferase toward 4-MUX in the *ex vivo* and *in vitro* model systems. In agreement with *in vitro* studies, the inactive E227A and D228A mutants were not able to prime GAG synthesis in β4GalT7-deficient cells, confirming that these residues are essential for the func-

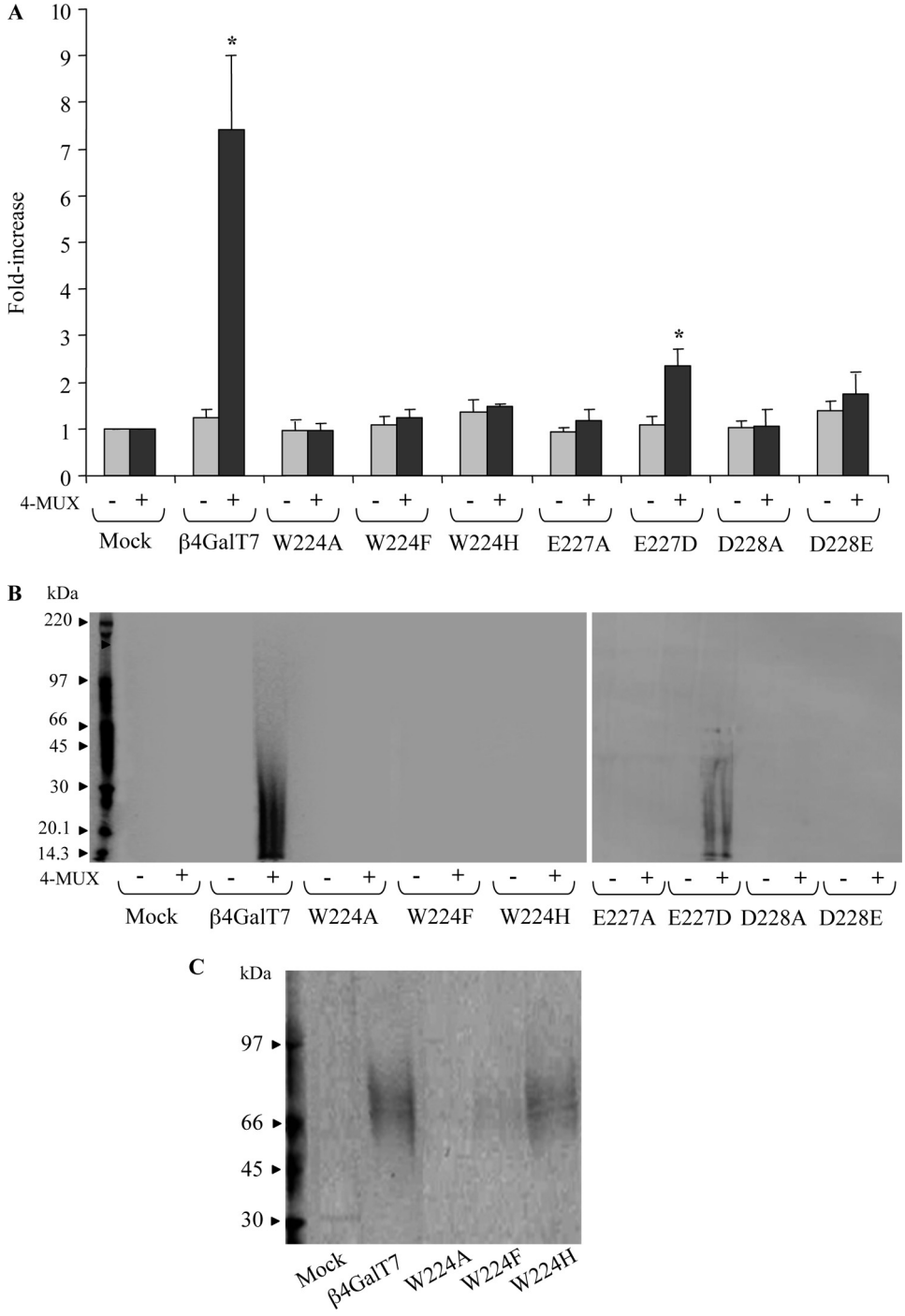


FIGURE 7. **Effect of wild-type and mutated β4GalT7 expression on GAG synthesis initiated from 4-MUX or on decorin core protein in CHO pgsB-618 cells.** For GAG analysis initiated from exogenous xyloside, cells were transiently transfected with wild-type (*WT*) or mutated β4GalT7 cDNA or mock-transfected (empty vector) and then incubated in medium containing $[^{35}S]SO_4$ in the presence (+) or absence (−) of 4-MUX (5 μM) for 12 h prior to GAG analysis. *A,* 1 ml of medium was applied to a G-50 column to separate nonincorporated radiolabeled sulfur, and radiolabeled GAGs present in the eluates were quantified by scintillation counting. Results are the mean ± S.D. of three separate experiments; *, $p < 0.05$, *versus* untreated; *B,* radiolabeled GAG chains were resolved by gel electrophoresis and visualized by autoradiography. *C,* CHO pgsB-618 cells stably expressing the recombinant human His-tagged decorin were transiently transfected with wild-type or mutant β4GalT7 cDNA, or mock-transfected, and incubated in the presence of $[^{35}S]SO_4$ for 12 h prior to analysis. His-tagged decorin was purified on nickel column, analyzed by SDS-PAGE, and radiolabeled sulfated GAG chain of the recombinant protein was visualized by autoradiography.

tion of this enzyme (Fig. 7*A*). On the other hand, results indicated that expression of the mutant in which $Glu^{227}$ was converted to Asp produced a 2-fold increase in sulfate incorporation in the culture medium of transfected cells upon addition of 4-MUX (Fig. 7*A*). As expected, radiolabeled GAGs could be detected by SDS-PAGE analysis in cells transfected by the E227D mutant, although the rate of sulfate incorporation was lower than that produced upon wild-type enzyme expression. Furthermore, substitution of $Asp^{228}$ with Glu, which strongly impaired *in vitro* activity, led to a mutant that was not able to prime GAG synthesis *ex vivo*. Altogether, these *ex vivo* studies are in good agreement with the *in vitro* findings showing the functional importance of these two carboxylate residues for the function of β4GalT7.

*Effect of $Trp^{224}$ Mutations on ex Vivo Initiation of GAG Chain Synthesis onto the Core Protein of Decorin*—We compared the capacity of wild-type and $Trp^{224}$ mutants to assemble GAG chain synthesis onto the core protein of decorin, taken as a model proteoglycan. For this purpose, a His-tagged form of human decorin was stably expressed in CHO pgsB-618 cells defective in β4GalT7. This cell line was transiently transfected with wild-type and $Trp^{224}$ substitution forms of β4GalT7, and the recombinant decorin was purified by nickel chelating chromatography from the culture medium following GAG chain labeling by sulfate incorporation. Gel electrophoresis of purified material from cells expressing wild-type β4GalT7 showed the presence of a GAG-substituted form of decorin migrating as a polydisperse smear centered at about 70 kDa, whereas mock-transfected cells showed no radiolabeling in agreement with the absence of GAG substitution in the β4GalT7-deficient cell line (Fig. 7*C*). When the W224A-substituted β4GalT7 cDNA was transfected into the cell line stably expressing recombinant decorin, no sulfated GAG chains could be detected. $[^{35}S]SO_4$ labeling of decorin was also low when cells were transfected with the W224F mutant, whereas expression of W224H led to a profile of decorin GAG sulfation that was comparable with that of

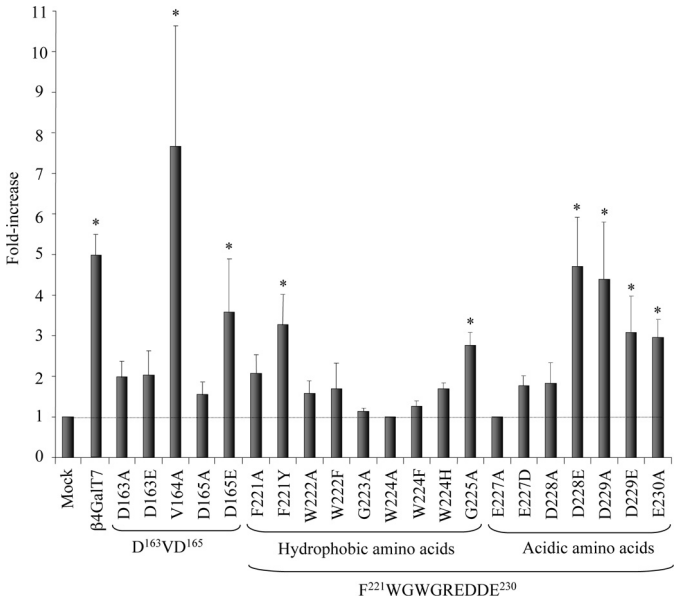


FIGURE 8. **Effect of wild-type and mutated β4GalT7 expression on total proteoglycan synthesis in CHO pgsB-618 cells.** Cells were transiently transfected with wild-type or mutated β4GalT7 cDNAs or mock-transfected (empty vector) for 48 h and placed in $[^{35}S]SO_4$ labeling medium 8 h prior to proteoglycan extraction and GAG quantification, as described under "Experimental Procedures." Results are expressed as the ratio of radioactivity associated with GAGs extracted from wild-type or mutated β4GalT7 transfected cells on radioactivity associated with GAGs extracted from mock-transfected cells. The results are the mean values ± S.D. of at least three independent experiments each performed in triplicate. *, $p < 0.05$, *versus* mock-transfected.

wild-type β4GalT7 (Fig. 7*C*). These results corroborated the importance of Trp at position 224 and indicate that replacement of $Trp^{224}$ with Phe or His yielded enzymes that were able to sustain GAG chain substitution in β4GalT7-deficient cells.

*Effect of Wild-type and Mutant β4GalT7 Expression on GAG Synthesis ex Vivo*—In a final set of experiments, we established the activity profile of the library of β4GalT7 mutants in the *ex vivo* model system and compared it with their *in vitro* activity. For this purpose, we determined the capacity of wild-type and mutated β4GalT7 enzymes to restore GAG synthesis in defective CHO pgsB-618 cells, by radioactive sulfate incorporation. As shown in Fig. 8, transfection of CHO pgsB-618 cells with wild-type β4GalT7 cDNA was able to restore GAG synthesis, producing up to a 5-fold increase in the rate of $[^{35}S]SO_4$ incorporation, compared with mock-transfected cells. Examination of the consequences of mutations within the $^{163}DVD^{165}$ motif indicated that expression of the mutants in which the Asp residues were converted to Ala did not produce a significant increase in sulfate incorporation compared with mock-transfected cells (≤2-fold). Similarly, the D163E mutant (inactive *in vitro*) was not able to support GAG synthesis when expressed in CHO-deficient cells (Fig. 8). By contrast, the GAG synthesis rate following expression of D165E was in the same range as the wild-type enzyme, indicating that Glu at this position was able to sustain significant activity, both *in vitro* and *ex vivo*. Furthermore, substitution of $Val^{164}$ with Ala, which did not affect *in vitro* β4GalT7 activity, resulted in a mutant enzyme able to increase GAG synthesis in defective cells, to about the same extent as that observed for the wild type. Altogether, these GAG synthesis studies are in agreement with our *in vitro* findings showing the functional importance of both Asp residues of the $^{163}DVD^{165}$ motif. Furthermore, in agreement with the *in vitro* assays, our results indicate that substitution of Asp with Glu at position 165, but not position 163, yielded an enzyme able to restore GAG synthesis, as the wild-type enzyme in β4GalT7-deficient cells.

The consequences of mutations within the $^{221}FWGWG-REDDE^{230}$ motif on GAG synthesis were next investigated in CHO pgsB-618-defective cells. Expression of the F221A mutant did not induce a significant increase in the rate of $[^{35}S]SO_4$ incorporation, compared with mock-transfected cells (Fig. 8), indicating that this nonconservative mutation exerts a more severe effect on total GAG synthesis than on *in vitro* galactosyltransferase activity toward the exogenous xyloside 4-MUX. On the other hand, the conservative change of $Phe^{221}$ to Tyr resulted in an enzyme that was able to restore significant GAG synthesis in deficient cells, concordant with the *in vitro* properties of this mutant. The nonconservative mutation of the neighboring $Trp^{222}$ residue to Ala led to reduction of *in vitro* enzyme activity, as well as GAG synthesis *ex vivo* (Fig. 8). Furthermore, the W222F mutant was not able to restore GAG synthesis, although this conservative mutation at position 222 did not affect *in vitro* galactosyltransferase activity toward 4-MUX.

Investigation of the *in vitro* kinetic properties of several mutants of Trp at position 224 suggested a key role for this residue that is located at a strategic position in the β4GalT7 active site. In agreement with the lack or very low galactosyltransferase activity of W224A and W224F, these mutants were not able to restore GAG synthesis when transfected in β4GalT7-deficient cells (Fig. 8). Expression of the W224H mutant did not significantly increase the rate of $[^{35}S]SO_4$ incorporation, compared with mock-transfected cells. The low activity of this mutant appears to be unable to restore the production of significant amounts of total GAGs, but this activity was sufficient to initiate some glycanation of specific proteoglycans, such as decorin, as indicated above. Altogether, analysis of the consequences of substituting $Trp^{224}$ on GAG synthesis in defective cells supports a key functional role of this residue.

The effects of expression of Ala mutants of $Gly^{223}$ or $Gly^{225}$ that surround $Trp^{224}$ to form a short flexible loop were also examined. Substitution of $Gly^{223}$ with Ala, which moderately impaired *in vitro* activity toward 4-MUX, produced a more important effect *ex vivo*, because this mutant was unable to promote GAG synthesis upon transfection in deficient cells (Fig. 8). This observation may be explained by the presence of the bulky side chain of Ala compared with Gly that may be tolerated in the presence of a small acceptor exogenous xyloside, but not when the glycopeptide primer of proteoglycans is the substrate. On the other hand, G225A, which was inactive toward 4-MUX *in vitro* and *ex vivo* (data not shown), was able to restore GAG synthesis when expressed in CHO pgsB-618 cells, although less than the wild-type protein. These results indicate that mutation of Gly to Ala at position 225 affects binding of 4-MUX more than that of the core protein of proteoglycans. In agreement, we found that this mutant was able to initiate decorin glycanation *ex vivo* (data not shown). However, more information on the organization of the acceptor binding site

based on the three-dimensional structure of β4GalT7 in complex with xyloside substrate is required to propose a structure-based explanation of the effects of these mutations.

Examination of the consequences of the mutations of the C-terminal acidic area of the $^{221}$FWGWGREDDE$^{230}$ motif indicated that mutation of Glu$^{227}$ to Ala abolished not only *in vitro* but also *ex vivo* β4GalT7 activity (Fig. 8). Furthermore, transfection of the E227D mutant, which exhibited a reduced *in vitro* galactosyltransferase activity compared with wild-type β4GalT7, was not able to promote GAG synthesis in CHO-defective cells. Mutation of Asp$^{228}$ to Ala strongly impaired *ex vivo* GAG synthetic activity of β4GalT7, whereas transfection of the D228E mutant restored GAG synthesis in deficient cells to the same rate as the wild-type enzyme. On the other hand, examination of the activity of nonconservative or conservative mutations of Asp$^{229}$ in GAG biosynthesis led to unexpected results. Whereas D229A and D229E mutants were found completely devoid of *in vitro* galactosyltransferase activity, both mutants were able to produce an increase in GAG synthesis up to that of the wild-type enzyme in deficient cells. A possible explanation for this discrepancy would be that mutations at this position affect binding of the aglycone part of the exogenous xyloside 4-MUX but not of the core protein of proteoglycans, which may involve additional interactions with β4GalT7 enzyme. In agreement with this hypothesis, we found that both mutants were unable to prime GAG synthesis from 4-MUX *ex vivo* but were able to sustain decorin glycanation (data not shown). However, elucidation of the three-dimensional organization of the acceptor substrate-binding site is required to precisely define the role of Asp$^{229}$. Finally, mutation of Glu$^{230}$, the last residue of the motif, to Ala did not impair GAG synthetic activity, consistent with the effect of this mutation observed on galactosyltransferase activity *in vitro*.

## DISCUSSION

The role of GAGs in a wide range of biological processes has led to considerable interest in understanding the properties of the glycosyltransferases involved in their biosynthesis. β4GalT7 is essential for the formation of the tetrasaccharide linkage region that initiates both glucosamino- and galactosaminoglycan synthesis (6); it therefore plays a key role in the early steps of GAG chain formation built on cellular and extracellular proteoglycans. β4GalT7 has also received much attention because it is able to prime GAG chains from exogenous β-D-xylosides. These compounds have been widely employed for examining the contribution of proteoglycans to biological processes and are currently attracting interest as potential drugs (12, 16, 17).

Here, we have focused on the organization of the functional peptide regions of human β4GalT7 and the mechanisms of Gal transfer from the UDP-Gal donor substrate onto the Xyl residue of exogenous xylosides and of the glycopeptide primer of GAG chains of proteoglycans. To pinpoint potential domains involved in substrate recognition and/or catalysis, we carried out a phylogenetic analysis of the family combined with molecular modeling of the active site of human β4GalT7, based on the crystal structure available for *Drosophila* β4GalT7 (10). The three-dimensional structure of bovine β4GalT1 complexed with UDP-Gal was also considered (31). The canonical motifs, $^{163}$DVD$^{165}$ and $^{221}$FWGWGEDDE$^{230}$ were subjected to systematic site-directed mutagenesis, and we assessed the functional consequences of these in two model systems designed to assess *in vitro* kinetic properties and GAG synthesis *ex vivo*. For the latter, decorin was chosen as the model core protein as it is a member of the small leucine-rich proteoglycan family and an important constituent of interstitial extracellular matrices (44). Decorin is covalently linked with one GAG chain of the chondroitin/dermatan sulfate type, and previous reports showed that genetic defects of β4GalT7 activity affect decorin glycanation with severe clinical consequences linked to the progeroid form of Ehlers-Danlos syndrome (45).

Multiple alignment of 48 sequences of β4GalT7 plus the 6 human β4GalT1 to β4GalT6 highlighted the presence of eight conserved motifs. Among those, motifs 3 and 5 contain two functional domains, *i.e.* the $^{163}$DVD$^{165}$ canonical motif of many glycosyltransferases and the $^{221}$FWGWGREDDE$^{230}$ domain corresponding to the so-called β4GT motif (46). The latter domain encompasses an Asp/Glu catalytic dicarboxylate found in many glycosyltransferases (47, 48). Our investigations showed the functional importance of both aspartate residues of the $^{163}$DVD$^{165}$ domain. In this respect, β4GalT7 exhibits typical features of glycosyltransferases of the GTA-fold group that require $Mn^{2+}$ as cofactor, in which a three-residue motif Asp-Xaa-Asp participates in metal ion binding (42). Examination of the organization of human β4GalT7 active site revealed that Asp$^{165}$ (as its counterpart Asp$^{147}$ in *Drosophila* and Asp$^{254}$ in bovine β4GalT1) is in a position to establish a coordination bond with the divalent metal cation, and we could demonstrate that replacing this residue with Glu maintains some of this function. By contrast, the longer side chain of Glu at position 163, when substituted to Asp, is not in a favorable position to provide interactions with the Gal unit of the donor substrate, as proposed for Asp$^{163}$ in the modeled organization of the human β4GalT7 donor-binding site. Furthermore, the Val$^{164}$ to Ala substitution did not affect β4GalT7 activity and led to a mutant enzyme able to restore GAG synthesis in deficient cells to a similar extent as that observed for the wild-type enzyme, indicating that this residue does not play a prevalent role in UDP-Gal binding. This result is in agreement with the proposed location of this Val residue in the structure of the *Drosophila* and the modeled human β4GalT7, pointing out of the donor-binding site. Altogether, we show that the structure of the $^{163}$DVD$^{165}$ motif of β4GalT7 is similar to that of glycosyltransferases of the GT-A fold group in which the first acidic residue of the sequence interacts directly with the sugar donor and the second acidic residue coordinates the metal divalent cation (21).

The conserved $^{221}$FWGWGREDDE$^{230}$ peptide region can be divided in two subdomains with an N-terminal part consisting of five hydrophobic residues located in a flexible loop and a C-terminal part containing a stretch of acidic amino acids belonging to an α-helix. Conservative and nonconservative substitutions at Phe$^{221}$ and Trp$^{222}$ produced only a moderate impairment of enzyme activity and function and did not change the $K_m$ value toward the donor substrate. Our results suggest that these residues are not directly involved in UDP-Gal bind-

ing, in agreement with their position out the donor substrate binding pocket of the modeled human β4GalT7. Thus, the fact that these residues are highly conserved may indicate a role in maintaining the structural integrity of the β4GalT7 active site. Examination of the consequences of mutating $Trp^{224}$ confirmed that this residue holds key functional importance in β4GalT7. Substitution with Ala resulted in a completely inactive enzyme. No galactosyltransferase activity could be detected when W224A was expressed in eukaryotic cells or when produced and purified from recombinant *E. coli*. Similarly, expression of this mutant was not able to restore GAG synthesis and decorin glycanation in β4GalT7-deficient CHO pgsB-618 cells. Replacement of the large Trp side chain by Phe, a smaller aromatic residue, was able to restore *in vitro* activity to a minor extent, and the low level of activity displayed by the W224F mutant did not allow kinetic parameters to be determined. A most striking finding of this study was revealed by analyzing the consequences of mutating $Trp^{224}$ to His, both on activity and specificity of β4GalT7. Indeed, this substitution produced an enzyme exhibiting a 4.4-fold decrease in $K_m$ value toward UDP-Gal compared with the wild-type enzyme whose affinity toward acceptor 4-MUX was strongly impaired. Characterization of the properties of the purified W224H mutant confirmed a large increase in affinity and showed enhanced selectivity toward UDP-Gal. Furthermore, we showed that the acceptor specificity profile of β4GalT7 was modified when $Trp^{224}$ was replaced with His. Indeed, whereas MN-Xyl was a preferred substrate for the wild-type enzyme, W224H exhibited higher activity toward 4-MUX. Molecular dynamic simulations of the human β4GalT7 structure, based on the x-ray crystal structure of *Drosophila* β4GalT7, indicate that the Nϵ1 atom of the indole side chain of $Trp^{224}$ forms a hydrogen bond with the β-phosphate group of UDP-Gal. Furthermore, in the crystal structure of bovine β4GalT1, it has been proposed that $Trp^{314}$ (equivalent to $Trp^{224}$ in β4GalT7) also aids in keeping the acceptor anchored by means of hydrophobic interactions (49). Thus, it can be hypothesized that in the β4GalT7 W224H mutant, the Nϵ2 atom of the imidazole side chain of His is able to sustain hydrogen bond formation with the pyrophosphate of UDP-Gal. However, the lack of an aromatic moiety in the imidazole side chain of His (compared with indole side chain of Trp) may impair stacking interactions with the acceptor substrate 4-MUX. In line with our results, a key role for a Trp residue at the catalytic center of several glycosyltransferases has previously been reported. For example, $Trp^{314}$ in β4GalT1 (50), $Trp^{284}$ in EXTL2 (51), and $Trp^{314}$ in α3GT (52) have all been shown to be located at a critical position for interacting with both sugar nucleotide and acceptor substrate. Thus, our study, in common with these other examples, highlights the presence of a Trp residue as a critical feature at the catalytic center of several glycosyltransferases. Supporting this idea, a consensus sequence containing the short amino acid stretch WGGE is present in human GalNAc-T4 transferase ($^{334}$WGGE$^{337}$ (53)) and in *Pasteurella multocida* hyaluronan synthase ($^{366}$WGGE$^{369}$, (54)).

Similar to what is observed in the structure of *Drosophila* β4GalT7 and bovine β4GalT1, $Trp^{224}$ (equivalent to $Trp^{207}$ in *Drosophila* and to $Trp^{314}$ in β4GalT1) lies in a flexible loop and is flanked by Gly residues at positions 223 and 225. Mutation of $Gly^{223}$ to Ala had a minor effect on 4-MUX enzyme activity *in vitro* and on decorin GAG synthesis (data not shown), although the G223A substituted enzyme was not able to restore GAG synthesis in deficient cells. On the other hand, replacement of $Gly^{225}$ abolished *in vitro* activity toward 4-MUX, whereas G225A sustained decorin glycanation (data not shown) and, to some extent, GAG synthesis in CHO pgsB-618 cells. Based on structural data and molecular modeling analyses, it is likely that mutation of these residues, which are predicted to allow a rotary motion of the Trp during the conformational change induced by UDP-Gal binding, affects positioning of $Trp^{224}$ within the catalytic center (10, 31). These mutations may also impact the organization of the acceptor-binding site with different consequences on β4GalT7 activity when Xyl is linked to a hydrophobic aglycone such as 4-MU or to the core protein of proteoglycans. Elucidation of the organization of the human β4GalT7 acceptor-binding site by x-ray crystallography is required to assess these assumptions.

Finally, we investigated the functional importance of the C-terminal acidic residues of the $^{221}$FWGWGREDDE$^{230}$ motif, which is highly conserved among all members of the β4GalT7 family. Whereas a crucial role could be demonstrated for $Glu^{227}$, $Asp^{228}$, and $Asp^{229}$, the glutamate residue at position 230 appears much less important in terms of function. Indeed, analysis of the *in vitro* consequences of $Glu^{227}$ mutations showed that Ala substitution completely abrogated enzyme activity, whereas replacement by Asp led to only a partial reduction in β4GalT7 activity, and the enzyme was able to sustain GAG synthesis activity in CHO pgsB-618 cells. Furthermore, our study indicated that a nonconservative mutation of $Asp^{228}$ had a severe impact on the galactosyltransferase activity of β4GalT7. However, replacement of Asp to Glu at that position produced a mutant with significant residual activity *in vitro* and *ex vivo*. Altogether, our data are concordant with a carboxylate residue acting as catalytic base in β4GalT7, facilitating the deprotonation of the acceptor substrate and the nucleophilic attack on the C1 of the sugar residue of UDP-Gal. In the three-dimensional structure of bovine β4GalT1 complexed with UDP-Gal, $Glu^{317}$ was proposed as the catalytic base (40). On the other hand, subsequent analysis of the structure of bovine W314A-β4GalT1 complexed with UDP, $Mn^{2+}$, and GlcNAc predicted that $Asp^{318}$ (equivalent to $Asp^{228}$ in human β4GalT7) should form a hydrogen bond with the hydroxyl group at C4 of GlcNAc and be involved in deprotonating the hydroxyl group (49). When Glc was docked into the acceptor-binding site of *D. melanogaster* β4GalT7-UDP structure, it was predicted to adopt a similar position to xylose, and residue $Asp^{211}$ was predicted to act as catalytic base (10). Molecular modeling of human β4GalT7 active site using *Drosophila* β4GalT7 as template favors the role of $Asp^{228}$ as catalytic base, whereas $Glu^{227}$ is predicted to interact with O4 of Gal. Interestingly, the presence of a functional Glu/Asp motif has been identified in several glycosyltransferases such as the cyclin glucan synthase (47) and the peptidoglycan glycosyltransferase module of class A penicillin-binding protein. The latter contains a conserved dicarboxylate $Glu^{233}$–$Asp^{234}$ (48) and glycan chain elongation catalyzed by this enzyme was proposed to function with the Glu

residue as active site general base to deprotonate the GlcNAc 4-OH group of lipid II acting as an acceptor. Further investigation of the interactions between human β4GalT7 and xylose is underway to decipher the respective role of each residue of the conserved dicarboxylate in the β4GalT7 family.

In conclusion, this work has identified two functional regions, $^{163}$DVD$^{165}$ (in motif 3) and $^{221}$FWGWGREDDE$^{230}$ (in motif 5), critical for the organization of UDP-Gal-binding site of human β4GalT7, and it suggests the existence of two β4GalT7-specific conserved peptide motifs (motifs 6 and 8). We also demonstrated the central role of Trp$^{224}$ in governing interactions with both donor and acceptor substrates by mutation of this amino acid, and we showed that it is thus possible to modulate the activity of β4GalT7 toward exogenous xylosides and proteoglycan core protein. Finally, our results support the critical role for a dicarboxylate motif in the vicinity of the key Trp$^{224}$ in the glycosyl transfer catalyzed by β4GalT7. This study brings us closer to an understanding of the enzymology of human β4GalT7, and it provides useful information toward the design of molecules targeting this enzyme, which have potential therapeutic applications.

*Acknowledgment—Matthieu Chabel is gratefully acknowledged for expert technical assistance.*

## REFERENCES


1. Bishop, J. R., Schuksz, M., and Esko, J. D. (2007) *Nature* **446,** 1030–1037
2. Kreuger, J., Spillmann, D., Li, J. P., and Lindahl, U. (2006) *J. Cell Biol.* **174,** 323–327
3. Prydz, K., and Dalen, K. T. (2000) *J. Cell Sci.* **113,** 193–205
4. Kusche-Gullberg, M., and Kjellén, L. (2003) *Curr. Opin. Struct. Biol.* **13,** 605–611
5. Bui, C., Ouzzine, M., Talhaoui, I., Sharp, S., Prydz, K., Coughtrie, M. W., and Fournel-Gigleux, S. (2010) *FASEB J.* **24,** 436–450
6. Almeida, R., Levery, S. B., Mandel, U., Kresse, H., Schwientek, T., Bennett, E. P., and Clausen, H. (1999) *J. Biol. Chem.* **274,** 26165–26171
7. Gulberti, S., Lattard, V., Fondeur, M., Jacquinet, J. C., Mulliert, G., Netter, P., Magdalou, J., Ouzzine, M., and Fournel-Gigleux, S. (2005) *J. Biol. Chem.* **280,** 1417–1425
8. Quentin, E., Gladen, A., Rodén, L., and Kresse, H. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87,** 1342–1346
9. Nakamura, Y., Haines, N., Chen, J., Okajima, T., Furukawa, K., Urano, T., Stanley, P., Irvine, K. D., and Furukawa, K. (2002) *J. Biol. Chem.* **277,** 46280–46288
10. Ramakrishnan, B., and Qasba, P. K. (2010) *J. Biol. Chem.* **285,** 15619–15626
11. Okajima, T., Yoshida, K., Kondo, T., and Furukawa, K. (1999) *J. Biol. Chem.* **274,** 22915–22918
12. Lugemwa, F. N., Sarkar, A. K., and Esko, J. D. (1996) *J. Biol. Chem.* **271,** 19159–19165
13. Martin, N. B., Masson, P., Sepulchre, C., Theveniaux, J., Millet, J., and Bellamy, F. (1996) *Semin. Thromb. Hemost.* **22,** 247–254
14. Kisilevsky, R., Szarek, W. A., Ancsin, J., Vohra, R., Li, Z., and Marone, S. (2004) *J. Mol. Neurosci.* **24,** 167–172
15. Mani, K., Havsmark, B., Persson, S., Kaneda, Y., Yamamoto, H., Sakurai, K., Ashikari, S., Habuchi, H., Suzuki, S., Kimata, K., Malmström, A., Westergren-Thorsson, G., and Fransson, L. A. (1998) *Cancer Res.* **58,** 1099–1104
16. Belting, M., Borsig, L., Fuster, M. M., Brown, J. R., Persson, L., Fransson, L. A., and Esko, J. D. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99,** 371–376
17. Schuksz, M., Fuster, M. M., Brown, J. R., Crawford, B. E., Ditto, D. P., Lawrence, R., Glass, C. A., Wang, L., Tor, Y., and Esko, J. D. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 13075–13080
18. Amado, M., Almeida, R., Schwientek, T., and Clausen, H. (1999) *Biochim. Biophys. Acta* **1473,** 35–53
19. Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2008) *Curr. Drug Targets* **9,** 292–309
20. Kitayama, K., Hayashida, Y., Nishida, K., and Akama, T. O. (2007) *J. Biol. Chem.* **282,** 30085–30096
21. Breton, C., and Imberty, A. (1999) *Curr. Opin. Struct. Biol.* **9,** 563–571
22. Hennet, T. (2002) *Cell. Mol. Life Sci.* **59,** 1081–1095
23. Huang, X., and Madan, A. (1999) *Genome Res.* **9,** 868–877
24. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410
25. Harduin-Lepers, A., Mollicone, R., Delannoy, P., and Oriol, R. (2005) *Glycobiology* **15,** 805–817
26. Combet, C., Blanchet, C., Geourjon, C., and Deléage, G. (2000) *Trends Biochem. Sci.* **25,** 147–150
27. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
28. Castresana, J. (2000) *Mol. Biol. Evol.* **17,** 540–552
29. Galtier, N., Gouy, M., and Gautier, C. (1996) *Comput. Appl. Biosci.* **12,** 543–548
30. Sali, A., and Blundell, T. L. (1993) *J. Mol. Biol.* **234,** 779–815
31. Ramakrishnan, B., Balaji, P. V., and Qasba, P. K. (2002) *J. Mol. Biol.* **318,** 491–502
32. Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Wang, B., Pearlman, D. A., Crowley, M., Brozell, S., Tsui, V., Gohlke, H., Mongan, J., Hornak, V., Cui, G., Beroza, P., Schafmeister, C., Caldwell, J. W., Ross, W. S., and P.A., K. (2004) *AMBER 8*, University of California, San Francisco
33. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M., Replogle, E. S., and Pople, J. A. (1998) *Gaussian 98,* Revision A 11.2, Gaussian, Inc., Pittsburgh
34. Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M. J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2010) *AMBER 11,* University of California, San Francisco
35. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
36. Fournel-Gigleux, S., Jackson, M. R., Wooster, R., and Burchell, B. (1989) *FEBS Lett.* **243,** 119–122
37. Fondeur-Gelinotte, M., Lattard, V., Oriol, R., Mollicone, R., Jacquinet, J. C., Mulliert, G., Gulberti, S., Netter, P., Magdalou, J., Ouzzine, M., and Fournel-Gigleux, S. (2006) *Protein Sci.* **15,** 1667–1678
38. Bansal, S. K., and Gessner, T. (1980) *Anal. Biochem.* **109,** 321–329
39. Li, D., Fournel-Gigleux, S., Barré, L., Mulliert, G., Netter, P., Magdalou, J., and Ouzzine, M. (2007) *J. Biol. Chem.* **282,** 36514–36524
40. Gastinel, L. N., Cambillau, C., and Bourne, Y. (1999) *EMBO J.* **18,** 3546–3557
41. Ramakrishnan, B., Ramasamy, V., and Qasba, P. K. (2006) *J. Mol. Biol.* **357,** 1619–1633
42. Breton, C., Snajdrová, L., Jeanneau, C., Koca, J., and Imberty, A. (2006) *Glycobiology* **16,** 29R–37R
43. Esko, J. D., Weinke, J. L., Taylor, W. H., Ekborg, G., Rodén, L., Anantharamaiah, G., and Gawish, A. (1987) *J. Biol. Chem.* **262,** 12189–12195
44. Roughley, P. J. (2006) *Eur. Cell. Mater.* **12,** 92–101
45. Seidler, D. G., Faiyaz-Ul-Haque, M., Hansen, U., Yip, G. W., Zaidi, S. H., Teebi, A. S., Kiesel, L., and Götte, M. (2006) *J. Mol. Med.* **84,** 583–594

46. Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2005) *Trends Biochem. Sci.* **30,** 53–62
47. Ciocchini, A. E., Roset, M. S., Briones, G., Iñón de Iannino, N., and Ugalde, R. A. (2006) *Glycobiology* **16,** 679–691
48. Terrak, M., Sauvage, E., Derouaux, A., Dehareng, D., Bouhss, A., Breukink, E., Jeanjean, S., and Nguyen-Distèche, M. (2008) *J. Biol. Chem.* **283,** 28464–28470
49. Ramasamy, V., Ramakrishnan, B., Boeggeman, E., and Qasba, P. K. (2003) *J. Mol. Biol.* **331,** 1065–1076
50. Ramakrishnan, B., and Qasba, P. K. (2001) *J. Mol. Biol.* **310,** 205–218
51. Pedersen, L. C., Dong, J., Taniguchi, F., Kitagawa, H., Krahn, J. M., Pedersen, L. G., Sugahara, K., and Negishi, M. (2003) *J. Biol. Chem.* **278,** 14420–14428
52. Zhang, Y., Swaminathan, G. J., Deshpande, A., Boix, E., Natesh, R., Xie, Z., Acharya, K. R., and Brew, K. (2003) *Biochemistry* **42,** 13512–13521
53. Bennett, E. P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M. A., Merkx, G., van Kessel, A. G., Eiberg, H., Steffensen, R., and Clausen, H. (1998) *J. Biol. Chem.* **273,** 30472–30481
54. DeAngelis, P. L., Jing, W., Drake, R. R., and Achyuthan, A. M. (1998) *J. Biol. Chem.* **273,** 8454–8458