Molecular Cloning of a Xylosyltransferase that Transfers the Second Xylose to O-Glucosylated Epidermal Growth Factor Repeats of Notch

Maya K. Sethi1, Falk F.R. Buettner1, Angel Ashikov1, Vadim B. Krylov2, Hideyuki Takeuchi3, Nikolay E. Nifantiev2, Robert S. Haltiwanger3, Rita Gerardy-Schahn1, and Hans Bakker1*

From the 1Department of Cellular Chemistry, Hannover Medical School, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany, the 2Laboratory of Glycoconjugate Chemistry, N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Leninsky Prospect 47, 119991 Moscow, Russian Federation, and the 3Department of Biochemistry and Cell Biology, Institute for Cell and Developmental Biology, Stony Brook University, Stony Brook, New York 11794-5215

*To whom correspondence should be addressed: Tel.: 49-511-5325214; E-mail: bakker.hans@mh-hannover.de

Keywords: Notch, glycosylation, xylosyltransferase

Background: Notch EGF repeats are glycosylated with xylose containing O-glycans.

Results: We have identified a human gene encoding an enzyme transferring the second xylose to generate the Xyl-Xyl-Glc trisaccharide on Notch EGF repeats. Genes encoding all glycosyltransferase activities involved in the O-glucose linked modification are now known.

Conclusion: Identification of the responsible genes allows elucidating the biological role of Notch xylosylation.

SUMMARY

The extracellular domain of Notch contains epidermal growth factor (EGF) repeats that are extensively modified with different O-linked glycans. O-Fucosylation is essential for receptor function and elongation with N-acetylgalcosamine, catalyzed by members of the Fringe family, modulates Notch activity. Only recently, genes encoding enzymes involved in the O-glucosylation pathway have been cloned. In the Drosophila mutant rumi, characterized by a mutation in the protein O-glucosyltransferase, Notch signaling is impaired in a temperature dependent manner, and a mouse knock-out leads to embryonic lethality. We have previously identified two human genes, GXLYT1 and GXLYT2, encoding glucoside-xylosyltransferases responsible for the transfer of xylose to O-linked glucose. The identity of the enzyme further elongating the glucosyl group to generate the final trisaccharide xylose-xylose-glucose, however, remained unknown. Here, we describe that the human gene C3ORF21 encodes a UDP-xylose: α1,3xylosyltransferase, acting on xylosea1,3glycoseβ1-containing acceptor structures.

We have, therefore, renamed it XXYL1T (xyloside-xylosyltransferase). XXYL1T cannot act on a synthetic acceptor containing an α-linked xylose alone, but requires the presence of the underlying glucose. Activity on Notch EGF repeats was proven by in vitro xylosylation of a mouse Notch1 fragment recombinantly produced in Sf9 insect cells, a bacterially expressed EGF repeat from mouse Notch2 modified in vitro by Rumi and Gxylt2, and in vivo by co-expression of the enzyme with the Notch1 fragment. The enzyme was shown to be a typical type II membrane bound glycosyltransferase localized in the endoplasmic reticulum.

Notch signaling is based on the interaction between transmembrane ligands of the Delta and Jagged/Serrate family and the extracellular domain of Notch receptors. This interaction enables two subsequent proteolytic cleavages that release the intracellular domain of the Notch receptor. After relocation to the nucleus, it binds to nuclear factors, thereby promoting target gene expression. The extracellular domains of Notch receptors and their ligands are composed of tandem epidermal growth factor-like (EGF5) repeats (reviewed in (1)) that are modified by unusual O-linked carbohydrates. The EGF repeats harbor consensus sequences containing serine or threonine residues to which O-linked fucose (Fuc), glucose (Glc) or N-acetylgalactosamine (GlcNAc) are transferred (2;3). Modification of O-linked GlcNAc has not been identified yet (3). In contrast, both fucose and glucose are further elongated to form the final tetrasaccharide Siaa2,3/6Galβ1,4GlcNAcβ1,3Fucra1-O-Ser/Thr (2) and trisaccharide Xyla1,3Xyla1,3Glcβ1-O-Ser (2;4-6), respectively. These O-glycans are found on a variety of other proteins including coagulation factor VII, IX and...
Thrombospondin (7;8), but are functionally most renowned as modulators of Notch signaling (9-11).

Mice and flies lacking the protein O-fucosyltransferase-1 (Pofut1) exhibit Notch loss-of-function phenotypes (12;13). Whereas, initial fucosylation seems to be a prerequisite for Notch receptor function, transfer of GlcNAc catalyzed by the enzymes of the Fringe family results in altered receptor ligand interactions. Signaling induced by the Notch ligand Delta is increased whereas signaling induced by Jagged/Serrate is decreased (14-18) as a result of Fringe activity.

In contrast, the importance of the trisaccharide Xyl-Xyl-Glc-O-Ser is less well understood. The initiating protein O-glucosyltransferase (Poglut) has been identified only recently through the Drosophila mutant rumi (19). Rumi deficient flies show a temperature sensitive phenotype, which is most severe and equivalent to the complete loss of Notch signaling in flies grown at 28-30 °C. Experiments performed by Acar et al. provide evidence that the proteolytic cleavage of the Notch receptor is impaired in rumi flies, but that interaction with the ligand Delta is not influenced. Thus, O-glucosylation appears to play a role for conformational stabilization of Notch at high temperatures to enable the proteolytic cleavage necessary for signal transmission (9;11;19;20). A mammalian ortholog of Rumi has been demonstrated to encode an active Poglut and results in an early lethal phenotype when knocked out in mice (21). In contrast to phenotypes of other global regulators of Notch signaling, GXYLT1 and GLT8D1 for recombinant expression via the pFast Bac system (Invitrogen) were reported previously (28). To generate the equivalent plasmid of XXYLT1, the putative luminal C-terminal domain, starting from Ser43 was amplified by PCR from human prostate Marathon-Ready cDNA (Clontech) using the primers 5’atctgaattcaggccgggagaccttctc3´ and 5´atctgaattcagccgggagaccttctc3´. After EcoRI digestion the sequence was cloned into the pFast Bac1 vector (Invitrogen), containing the Honeybee Melittin secretion sequence and the Protein A coding sequence from pProtA (ProtA-XXYLT1). C-terminal Myc/His-tagged mouse Notch EGF1-5, encoding the first five EGF domains of mouse Notch1 cloned into pFast Bac1 was described previously (28;30).

Constructs for enzyme expression in mammalian cells: The cDNAs encoding the luminal domains of mouse Gxylt2, starting from Arg26, or Xxylt1, starting from Ser43 were subcloned into a pSecTag2 vector (Invitrogen) so that the recombinant proteins were generated for the control gene, the human β1,4 galactosyltransferase1 (B4GALT1-HA and Flag-B4GALT1-HA) were generated for the control gene, the human β1,4 galactosyltransferase1 (B4GALT1; NCBI Ref. Seq. NM_001497). To generate an appropriate ER Marker, the Golgi located pAcGFP1-Golgi Vector (Clontech), encoding a GFP-fusion protein located in the Golgi apparatus, was modified by insertion of a c-terminal KDEL signal sequence. After PCR on the original plasmid, the modified insert was re-cloned via BamHI and XbaI restriction sites into the vector. The new disaccharide Xyl-Glc- (28) indicating the requirement for an additional xylosyltransferase. Here we describe the identification of another human member of the GT8 family, only distantly related to the previously analyzed GT8 members, encoding a xyloside-xylosyltransferase (XXYLT), which further elongates the Xyl-Glc-O-Ser disaccharide on Notch EGF repeats.

**EXPERIMENTAL PROCEDURES**

**Plasmid Constructs**—Constructs for baculoviral expression in Sf9 cells: Protein A fusion constructs of GXYLT1 and GLT8D1 for recombinant expression via the pFast Bac system (Invitrogen) were reported previously (28). To generate the equivalent plasmid of XXYLT1, the putative luminal C-terminal domain, starting from Ser43 was amplified by PCR from human prostate Marathon-Ready cDNA (Clontech) using the primers 5’atctgaattcaggccgggagaccttctc3´ and 5´atctgaattcagccgggagaccttctc3´. After EcoRI digestion the sequence was cloned into the pFast Bac1 vector (Invitrogen), containing the Honeybee Melittin secretion sequence followed by the Protein A coding sequence from pProtA (ProtA-XXYLT1). C-terminal Myc/His-tagged mouse Notch EGF1-5, encoding the first five EGF domains of mouse Notch1 cloned into pFast Bac1 was described previously (28;30).

Constructs for enzyme expression in mammalian cells: The cDNAs encoding the luminal domains of mouse Gxylt2, starting from Arg26, or Xxylt1, starting from Ser43 were subcloned into a pSecTag2 vector (Invitrogen) so that the recombinant proteins were expressed with a C-terminal Myc/His-tag in mammalian cells and secreted into the culture media.

Constructs for immunofluorescence and cleavage analysis: The full length open reading frame of XXYLT1 (NCBI Ref. Seq. NM_152531) was cloned into the pcDNA3 vector (Invitrogen), modified with protein tags to generate the single tagged XXYLT1-HA (XXYLT1-SRYPYDVPDYASL) and Flag-XXYLT1 (MDYKDDDDKEF-XXYLTL1), and the doubly tagged Flag-XXYLTL1-HA (MDYKDDDDKEF-XXYLTL1-SRYPYDVPDYASL) plasmids. Similar constructs (B4GALT1-HA and Flag-B4GALT1-HA) were generated for the control gene, the human β1,4 galactosyltransferase1 (B4GALT1; NCBI Ref. Seq. NM_001497). To generate an appropriate ER Marker, the Golgi located pAcGFP1-Golgi Vector (Clontech), encoding a GFP-fusion protein located in the Golgi apparatus, was modified by insertion of a c-terminal KDEL signal sequence. After PCR on the original plasmid, the modified insert was re-cloned via BamHI and XbaI restriction sites into the vector. The new disaccharide Xyl-Glc- (28) indicating the requirement for an additional xylosyltransferase. Here we describe the identification of another human member of the GT8 family, only distantly related to the previously analyzed GT8 members, encoding a xyloside-xylosyltransferase (XXYLT), which further elongates the Xyl-Glc-O-Ser disaccharide on Notch EGF repeats.

**EXPERIMENTAL PROCEDURES**

**Plasmid Constructs**—Constructs for baculoviral expression in Sf9 cells: Protein A fusion constructs of GXYLT1 and GLT8D1 for recombinant expression via the pFast Bac system (Invitrogen) were reported previously (28). To generate the equivalent plasmid of XXYLT1, the putative luminal C-terminal domain, starting from Ser43 was amplified by PCR from human prostate Marathon-Ready cDNA (Clontech) using the primers 5’atctgaattcaggccgggagaccttctc3´ and 5´atctgaattcagccgggagaccttctc3´. After EcoRI digestion the sequence was cloned into the pFast Bac1 vector (Invitrogen), containing the Honeybee Melittin secretion sequence followed by the Protein A coding sequence from pProtA (ProtA-XXYLT1). C-terminal Myc/His-tagged mouse Notch EGF1-5, encoding the first five EGF domains of mouse Notch1 cloned into pFast Bac1 was described previously (28;30).

Constructs for enzyme expression in mammalian cells: The cDNAs encoding the luminal domains of mouse Gxylt2, starting from Arg26, or Xxylt1, starting from Ser43 were subcloned into a pSecTag2 vector (Invitrogen) so that the recombinant proteins were generated for the control gene, the human β1,4 galactosyltransferase1 (B4GALT1-HA and Flag-B4GALT1-HA) were generated for the control gene, the human β1,4 galactosyltransferase1 (B4GALT1; NCBI Ref. Seq. NM_001497). To generate an appropriate ER Marker, the Golgi located pAcGFP1-Golgi Vector (Clontech), encoding a GFP-fusion protein located in the Golgi apparatus, was modified by insertion of a c-terminal KDEL signal sequence. After PCR on the original plasmid, the modified insert was re-cloned via BamHI and XbaI restriction sites into the vector. The new disaccharide Xyl-Glc- (28) indicating the requirement for an additional xylosyltransferase. Here we describe the identification of another human member of the GT8 family, only distantly related to the previously analyzed GT8 members, encoding a xyloside-xylosyltransferase (XXYLT), which further elongates the Xyl-Glc-O-Ser disaccharide on Notch EGF repeats.
construct GFP1-KDEL was tested by immunofluorescence studies in CHO cells. DNA sequences of the constructs were confirmed by sequencing. All the primers are available upon request.

**Enzyme Expression and Purification**—Protein A fusion constructs of soluble secreted human GLT8D1, GXYLT1 and XXYLT1 were expressed in Sf9 cells by baculovirus infection (Bac-to-Bac®; Invitrogen) and purified by IgG-Sepharose-6 Fast Flow beads (GE-Healthcare) as described (28). Soluble secreted mouse Gxylt2 and Xxylt1 were expressed in HEK293T cells and purified from the culture media by Ni-NTA affinity chromatography as previously described (24). Protein concentration was determined by Coomassie staining using BSA or Protein A as standard.

**In Vitro Xylosyltransferase Activity Assays**—Assays were performed in a total volume of 50 µl in 100 mM MOPS pH 7.5, 10 mM MnCl₂, 10 mM ATP as described before (28). IgG-bead coupled Protein A fused enzymes were incubated for 1 h at 37 °C in presence of radiolabeled donor sugars UDP-[6-3H]Gal, UDP-[1-1H]Glc (GE-Healthcare), or UDP-[U-14C]Xyl (Perkin Elmer) at 5 µM with specific activity 4 kBq/nmol for [3H] sugars and 0.75 kBq/nmol for [14C]Xyl, obtained by dilution with cold nucleotide sugars (Sigma and Carbosource Services) and a number of synthetic acceptors (Xyl-Xyl-Glc-R; Xyl-Glc-R, Glc-R (31) or para-nitrophenol [pNP]-linked carbohydrates) at a final concentration of 100 µM. To determine the amount of transferred radiolabeled sugars, donor and acceptor were separated via C18 columns (Sep-Pak® Vac 3cc; Waters Corp.) and samples were counted by liquid scintillation (LS 6500; Beckman Coulter). The acquired enzymatic activity was expressed as nmol xylose transferred per nmol protein per hour, calculated from a Coomassie Blue staining using a Protein A standard as reference. Note that this assay was set up to show specificity, but does not allow drawing conclusions about enzymatic efficiency due to the about equimolar amount of UDP-Xyl and enzyme in the assay.

A single 16th EGF repeat (EGF16) from mouse Notch2 was prepared and modified with Glc-O or Xyl-O by Rumi as previously described (24). In vitro xylosylation of Glc-EGF repeat was performed using mouse Gxylt2 that was produced and purified by nickel-affinity from medium of HEK293T cells. Typically, a 5-ml reaction mixture contained 50 mM HEPES pH 7.0, 10 mM MnCl₂, 5 µM Glc-EGF repeat, 200 µM UDP-Xyl, 0.5% Nonidet P-40, and approximately 10 µg of Gxylt2. Samples were incubated at 37 °C overnight and the product was purified by reversed phase HPLC as previously described (24). Substrate specificity of XXYLT1 on these differently glycosylated EGF16 acceptors was tested using radiolabeled UDP-Xyl in a reaction volume of 10 µl containing 50 mM HEPES pH 7.0, 10 mM MnCl₂, 10 µM acceptor substrate, 10 µM UDP-[14C]Xyl (5 kBq/nmol; American Radiolabeled Chemicals), 0.5% NP-40, and 10 ng of the mouse Xxylt1 protein, purified as described above for Gxyt2. After 20 min incubation at 37 °C the reaction was stopped by adding 900 µl of 100 mM EDTA pH 8.0. Samples were loaded onto C18 cartridges (100 mg, Agilent). After washing with 5 ml of H₂O, the EGF repeats were eluted with 1 ml of 80% methanol. Incorporation of [14C]Xyl into the EGF repeats was determined by liquid scintillation. Reactions without substrates were used as background controls.

**Reverse Phase Chromatography and NMR Analysis after Enzymatic Xylosylation of Xyl-Glc-R**—1 mg of synthetic acceptor Xyl-Glc-R was xylosylated in 500 µl reaction buffer by 100 µl bead coupled XXYLT1 in presence of equimolar amounts of cold UDP-Xyl under standard in vitro activity assay conditions. The sample was taken up in 100 µl of 25% aqueous acetonitrile and analyzed (20 µl) by reverse phase HPLC on a column (25 cm x 4.6 mm) with LC-8 (5 µm) sorbent with elution by 25% aqueous acetonitrile at flow rate of 0.8 ml/min. The elution profile was compared to profiles of reference trisaccharide Xyl-Xyl-Glc-R and disaccharide Xyl-Glc-R acceptor (31).

For linkage confirmation in the product of in vitro xylosylation, 'H NMR (600 MHz, 303 K, D₂O, Bruker Avance 600), natural abundance [13C] NMR (125 MHz, 303 K, D₂O, Bruker Avance 600), and 2D Heteronuclear single quantum correlation (HSQC) NMR (600 MHz, 303 K, D₂O, Bruker Avance 600) of the reaction mixture were acquired and compared with spectral data for the synthetic reference compounds (see Fig. 3 and supplemental figures).

**Notch EGF1-5 Expression, Purification and in vitro Xylosylation**—C-terminal Myc/His-tagged mouse Notch EGF1-5 was expressed in 300 ml Sf9 insect cells cultured in Insect-Xpress medium (Lonza) by baculovirus infection. After 72 h, the secreted protein was purified by nickel affinity chromatography (HisTrap HP, 1ml; GE-Healthcare) as described (28). Protein fractions, reactive with anti-Myc, were pooled and concentrated/desalted to 150 µl using an Amicon Ultra-4 centrifugal devise (Millipore) and stored in 100 mM MOPS, pH 7 at -20 °C. SDS-Page followed by Coomassie staining and immunoblotting with monoclonal antibody 9E10 (anti-Myc) confirmed the protein purification. Under standard activity assay conditions, 10 µl of purified Notch EGF1-5 was...
incubated in the presence of 100 μM cold UDP-Xyl as donor and 10 μl bead coupled enzyme for 4 h, separated by SDS-PAGE and analyzed by mass spectrometry (MS).

LC-MS/MS—Protein bands of Notch EGF1-5 were excised from gel, trypsin digested and peptides were recovered as described (32). Reverse phase chromatography using acetonitrile as eluent was performed on a Waters Aquity nano-UPLC device equipped with an analytical column (Waters, BEH130 C18, 100 μm x 100 mm, 1.7 μm particle size) coupled online to an ESI Q-TOF Ultima (Waters). Spectra were recorded in positive ion mode and peptides were automatically subjected to fragmentation (MS/MS). Unglycosylated or glycosylated versions of EGF16 from mouse Notch2 were analyzed by nano-LC-MS/MS as previously described (24).

In Vivo Enzymatic Activity Assay—Notch EGF1-5 was co-expressed with XXYLT1 or as control the inactive enzyme GLT8D1 (28) in Sf9 cells by baculovirus infection for 72 h. Purification and analysis of peptide glycosylation of the Notch fragment was performed as specified above.

Immunostaining—Localization studies were carried out in CHO cells. Cell transfection was performed on glass cover slips in 24 well plates using Metafectene (Biontex). After 24 h cells were fixed using 4% PFA in PBS, permeabilized with 0.1% Saponin in PBS, and stained with respective primary (rabbit anti-HA [Sigma], mouse anti-Flag M5 [Sigma], or rabbit α-mannosidase II [gift of Dr. K. Moremen, University of Georgia, Athens]) and secondary antibodies (anti-mouse IgG-Cy3 [Sigma] and anti-rabbit IgG-Alexa488 [Invitrogen]) using standard staining conditions. As ER marker the [Sigma] and anti-rabbit IgG-Alexa488 [Invitrogen]) and secondary antibodies (anti-mouse IgG-Cy3 [Sigma] and anti-rabbit IgG-Alexa488 [Invitrogen]) were applied to SDS-PAGE. Protein expression was analyzed by western blotting using primary antibodies rabbit anti-HA and rabbit anti-Flag (both Sigma) followed by visualization with rabbit IRDye 800CW (LI-COR) using the LI-COR Odyssey Imager.

RESULTS

Selection of a Potential Xyloside-Xylosyltransferase (XXYLT) —The human genes GXYLT1 and GXYLT2 transfer xylose in α1,3 to O-linked glucose (Fig. 1B). The gene responsible for the transfer of the second xylose remained unknown. Considering that both enzymatic reactions result in the addition of xylose in α1,3 linkage, the genes were anticipated to be homologous. Still, none of the four initially identified members of the GT8 family (GXYLT1, GXYLT2, GLT8D1, and GLT8D2) that were selected based on homology with UDP-Glc:glycoprotein glucosyltransferase (genes UGGT1 and UGGT2) had xylosyltransferase activity towards Xylα1,3Glcβ1 terminated acceptors (28). Using GXYLT1 in a position-specific iterated (PSI)-Blast, a fifth gene was identified, showing less than 20% overall identity at the amino acid level (Fig. 1A). This gene, named C3ORF21, possessed two conserved DXD motifs and an N-terminal signal sequence or membrane anchor. Therefore, its potential involvement in the synthesis of the Xylα1,3Xylα1,3Glcβ1-O oligosaccharide was investigated.

Xylosyltransferase Activity with Synthetic Acceptors—To test the enzymatic activity of the putative xylosyltransferase, renamed XXYLT1 (xyloside-xylosyltransferase 1), the protein, lacking the predicted membrane bound domain, was fused N-terminally to a Protein A tag and cloned in a baculoviral vector with a signal sequence to promote secretion of the fusion-protein. After expression in Sf9 insect cells, the protein was isolated from the culture media by capture with IgG beads, binding the N-terminal Protein A tag. Bead coupled enzyme was used for in vitro glycosyltransferase activity assays with artificial synthetic compounds that mimic the natural acceptor structures (28;31). Assays carried out with the radiolabeled donor substrate UDP-[14C]Xyl revealed xylosyltransferase activity of XXYLT1 with the Xyl-Glc-R synthetic acceptor (Fig. 2A). Marginal increased values, compared to the negative control, were observed for the Xyl-Xyl-Glc-R acceptor, which can be rather interpreted as due to the presence of degradation products within the acceptor compound, than resulting from generation of a product with three xylose residues. This conclusion is supported by the lack of any signal in the HPLC profile of figure 3B that could represent such product. XXYLT1 showed a clear donor substrate specificity for UDP-Xyl (Fig. 2C). Interestingly no enzymatic activity could be observed with a panel of sugars linked to para-nitrophenol (pNP) as acceptors,
(Fig. 2B) including α-linked Xyl-pNP, suggesting that the enzyme requires the disaccharide Xyla1,3Glcβ1- as minimal acceptor structure.

Analysis of the Reaction Product after in Vitro Xylosylation by XXYLT1—The reaction product generated by XXYLT1 was investigated by high performance liquid chromatography (HPLC) and nuclear magnetic resonance (NMR) studies to confirm the structure of the formed trisaccharide, especially of the direction and configuration of the linkage between the two xylosyl units (Fig. 3). HPLC data of Xyl-Glc-R modified by XXYLT1 in the presence of UDP-Xyl, shown in Fig. 3B, indicated that the enzyme generated a product migrating at the same position as the synthetic reference trisaccharide Xyl-Xyl-Glc-R (Fig. 3C). Moreover, acquired $^{13}$C and $^1$H NMR spectra of the enzymatically modified disaccharide (Fig. 3B; suppl. Fig. S1B) are identical to the spectra of the synthetic reference trisaccharide (Fig. 3C; suppl. Fig. S1C). In particular, the characteristic low-field (80.3 ppm) location of C3'-signal in the $^{13}$C-NMR spectrum unequivocally shows, that the xylosyl-residue is linked through an α1,3 linkage, while the α-configuration of the terminal xylose unit is indicated by the coupling constant $J_{H1,H2}$ of 3.7 Hz (suppl. Fig. S1; suppl. Table S1). Finally the precise nature of the product was also confirmed by a 2D HSQC NMR-spectrum (suppl. Fig. S2).

Xylosylation of Notch EGF Repeats by XXYLT1—To determine if Notch is modified by XXYLT1, an in vitro assay was carried out to test a naturally occurring EGF acceptor. A mouse Notch fragment containing five EGF repeats with consensus sequences for O-glycosylation in repeat 2 and 4 (Notch EGF1-5) was expressed in Sf9 insect cells and purified by Nickel-affinity-chromatography. The purified acceptor presented both Glc-O and Xyl-Glc-O modifications in about a one to one ratio as shown by mass spectral analysis of the glycopeptide fragment from EGF4 (m/z 1166 and 1210, Fig. 4A). Incubation of this acceptor with GXYLT1 resulted in the expected conversion of Glc-O into Xyl-Glc-O, visible by the disappearance of the peak at m/z 1166 (Fig. 4B). In contrast, incubation of Notch EGF1-5 with XXYLT1 resulted in the de novo generation of a product at m/z 1254 (Fig. 4C) and loss of the Xyl-Glс-O peak at m/z 1210. The new peak at m/z 1254 demonstrated the addition of one pentose residue to the Xyl-Glc-O containing peptide. Modification of Notch EGF1-5 with both enzymes, GXYLT1 and XXYLT1, resulted in a complete shift towards the formation of the trisaccharide Xyl-Xyl-Glc-O on EGF repeat 4 (Fig. 4D). The same was observed for the peptide of EGF repeat 2, but overlapping signals from another peptide troubled the picture (data not shown).

Since Rumi was recently shown to be capable of transferring either xylose or glucose to certain EGF repeats (e.g. EGF16 from mouse Notch2), and Xyl-O can be further elongated to a Xyl-Xyl-Xyl-O trisaccharide (24), we wanted to test if XXYLT1 can add xylose to an O-xylosylated EGF repeat. Bacterially expressed EGF16 (unglycosylated) was incubated with Rumi in presence of UDP-Glc or UDP-Xyl to generate the O-glycosylated and O-xylosylated forms, respectively. Reverse phase HPLC analysis showed that the glycosylated EGF repeats eluted slightly earlier than the xylosylated form (Fig. 5A-C). Glc-O-EGF16 was converted to Xyl-Glc-O-EGF16 by incubation with mouse Gxylt1 and UDP-Xyl (Fig. 5D; all structures were confirmed by mass spectrometry, suppl. Fig. S4). Each of these EGF repeats was tested as potential acceptor substrate for mouse Xxylt1, but only Xyl-Glc-O-EGF16 was converted by the enzyme (Fig. 5E, F), showing that Xxylt1 is not able to elongate Xyl-O-EGF16.

In vivo enzymatic activity of XXYLT1 was tested by a co-expression approach. The MS data shown in Fig. 4A already indicated that insect cells endogenously express a glucoside-xylosyltransferase, which partially xylosylates O-glucose of mouse Notch EGF repeats. But so far, no trisaccharide modification has been detected in Sf9 cells. Upon co-transfection of Notch EGF1-5 and XXYLT1 in Sf9 cells, we could detect the trisaccharide modification at m/z 1254 (Fig. 6B). In comparison, samples obtained from co-expression of Notch EGF1-5 and the inactive enzyme GLT8D1 only showed Glc-O and Xyl-Glc-O modifications (Fig. 6A).

Intracellular Location and Cleavage Analysis of XXYLT1—To determine the intracellular localization of the newly identified glycosyltransferase, we transiently expressed N-terminal Flag-tagged and/or C-terminal HA-tagged XXYLT1 in CHO cells followed by immunofluorescence staining. Expression of the three constructs resulted in identical subcellular expression patterns and signals acquired from the doubly tagged Flag-XXYLT1-HA construct clearly matched (Fig. 7A and B). Co-localization of Flag-XXYLT1-HA and an ER retained GFP construct, indicated that XXYLT1 is located in the ER (Fig. 7C). In contrast, no merge between the Golgi marker mannosidase II and XXYLT1 was found (Fig. 7D). The same construct of β4-galactosyltransferase (B4GALT1), which was used as a control, was always located, as expected, in the Golgi apparatus (Fig. 7E) (33).
Both Poglut/Rumi and Pofut1 are soluble enzymes in the ER (19;34), whereas most other glycosyltransferases are type II membrane protein in the Golgi (35). The question, therefore, arises if XXYLT1 is a typical type II glycosyltransferase with a signal anchor or is cleaved after a signal sequence has directed the protein into the ER. Moreover, a prediction program (SignalP 3.0) at www.cbs.dtu.dk/ (36) predicted the protein to have a cleaved signal sequence with a total probability of 0.78 (Fig. 7F). To answer this question, the doubly tagged Flag-XXYLT1-HA and C-terminally tagged XXYLT1-HA proteins were analyzed by western blotting. As control, B4GALT1, which is an established non cleaved type II transmembrane protein (33) was investigated in parallel. Two identical blots were generated and developed with either an anti-Flag or anti-HA antibody (Fig. 7G). In all cases XXYLT1 behaved like B4GALT1. The doubly tagged protein was visible as a single band, and no smaller products reacting only with the anti-HA antibody were observed. As the N-terminal fused Flag-tag might shield a possible cleavage signal at the N-terminus, the XXYLT1-HA was investigated too. Absence of the Flag-tag reduced the molecular weight by ~1.4 kDa. For both B4GALT1 and XXYLT1 a slightly shortened protein, corresponding to this difference, was now visible in the blot developed with the HA antibody. But again, no smaller, potentially cleaved, products were observed. Protein bands running at higher molecular weight most likely represented dimers of XXYLT1. These data indicated that XXYLT1, most likely, is a typical type II transmembrane protein, but localized in the ER.

DISCUSSION

This report describes the identification of an α1,3-xylosyltransferase acting on the Xyl1,3Glc- linked glycan of Notch EGF domains and was accordingly named xyloside-xylosyltransferase 1 (XXYLT1). Like the α1,3-xylosyltransferases GXYLT1 and GXYLT2 involved in the preceding glycosylation step, XXYLT1 belongs to the glycosyltransferase family 8. However, although it catalyses the transfer of a xylose residue in α1,3 linkage as the GXYLTs, XXYLT1 uses a different acceptor substrate and only presents a low identity of less than 20% to these enzymes. XXYLT1 has all properties to encode the xylosyltransferase activity previously detected in HepG2 cells (26) and to be responsible for xylosylation of Notch and other EGF repeat containing proteins. Formal proof that this is the only enzyme in the human genome with this activity, however, still has to be provided.

The second Notch xylosyltransferase

The glycosyltransferase family 8 comprises a broad range of glycosyltransferases, from both bacteria and eukaryotes, involved in the transfer of a variety of sugars. Mammalian members of this family have only been shown to transfer either glucose or xylose. Glucose is the donor sugar for glycogenin (37), the autocatalytic precursor for glycogen, and for UDP-Glc:glycoprotein glucosyltransferase (38). GXYLT1, GXYLT2 and the newly identified XXYLT1 have now been shown to encode xylosyltransferases acting on O-glucosylated EGF repeats. However, the enzymatic activity of the remaining enzymes LARGE1, LARGE2, GLT8D1 and GLT8D2 is still unknown. LARGE proteins comprise two predicted catalytic domains of glycosyltransferases, one of the GT8 family and one of family 49 (39) and are known to act on dystroglycan (40).

Both GXYLT and XXYLT seem to have emerged at the same time with the appearance of metazoans. As for GXYLT1 and -2, there is no ortholog of XXYLT in the unicellular Monosiga brevicollis, which has actually a LARGE-like glycosyltransferase. However, clear orthologs of both GXYLT and XXYLT are found in most metazoans, including Drosophila melanogaster. The latter is rather surprising since we have not observed the product of XXYLT on Notch produced in S9 insect cells (28). Similarly, in O-linked glycans isolated from Drosophila, only the Xyl-Glc disaccharide was detected (41). The expression and activity of Drosophila XXYLT still has to be confirmed.

In contrast to GXYLT1 and -2, XXYLT1 showed no activity with xylose linked to pNP or with Xyl-O-EGF16, a structure recently demonstrated to be formed by Poglut/Rumi (24). The GXYLTs were able to use Glcβ-pNP as substrate, but XXYLT1 was only active with an acceptor containing Xylα1-3Glcβ1. This actually is a logical requirement to prevent the enzyme from using its transferred xylose in α1-3 linkage again as acceptor substrate. Indeed, no further extension beyond the trisaccharide is observed both naturally and in our in vitro assays. Whether the xylosyltransferases recognize the underlying EGF repeat is not known. Recognizing EGF repeats would be required to differentially xylosylate EGF repeats. At least in Drosophila, xylosylation appears to be EGF repeat dependent (42), indicating that the Drosophila xylosyltransferase shows site specificity. On the other hand, mouse Notch1 overexpressed in tissue culture cells is mainly uniformly modified with the trisaccharide Xyl-Xyl-Glc (43).

Human XXYLT1 is predicted to have a cleavable signal sequence rather than an N-terminal signal anchor (Fig. 7). This is in fact the case for many...
The second Notch xylosyltransferase

glycosyltransferases that are in reality non-cleaved type II transmembrane proteins of the Golgi (35). In contrast, Pofut1 and Poglut/Rumi, the enzymes responsible for initial O-fucosylation and O-glucosylation, are soluble proteins in the ER and have a cleavable signal sequence (19;34). Our experiments clearly showed that XXYLT1 is not cleaved and behaves like the archetypal type II membrane bound glycosyltransferase B4GALT1 (33). On the other hand, it does not show the typical Golgi localization of B4GALT1. It is localized primarily to the ER. To avoid influencing the subcellular localization by changing the N- or C-terminal sequence, we have generated both N- and C-terminally tagged constructs. These always showed identical localization as well as a construct that was doubly tagged. We, thus, conclude from these experiments that XXYLT1 is a type II transmembrane protein, typical for glycosyltransferases, but is located in the ER. XXYLT1 behaves very similar to COSMC in these respects. COSMC is an ER localized chaperone specific for the Golgi glycosyltransferase T-synthase and is itself a type II transmembrane protein with homology to the same T-synthase. It resides in the ER due to formation of a homodimer by a disulfide-bond within the transmembrane domain that is essential for ER retention (44). Indeed, we observed dimer formation for XXYLT1 (Fig. 7G) and two cysteine residues, potentially responsible for intermolecular disulfide-bond formation, are present within the transmembrane domain. There are, therefore, strong indications that XXYLT1 is retained in the ER by the same mechanism as COSMC. Moreover, the xylosyltransferase has an AxxxxA motif within the predicted transmembrane domain. Axxxx or GxxxG motifs in other proteins have been shown to promote membrane helix interactions (45;46).

With the previous identification of the glucosyltransferase Rumi, the two GXYLTs, and now the cloning of XXYLT1, all genes encoding the enzymes required for the formation of the Xyl-Xyl-Glc-glycotope are currently known. These important steps will now allow investigating the function of these modifications.

Acknowledgements—This work is supported by funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for the Cluster of Excellence REBIRTH (From Regenerative Biology to Reconstructive Therapy) and by the National Institutes of Health Grant GM61126 (to RSH). UDP-Xyl isolation by the Carbosource Services at the Complex Carbohydrate Research Center, University of Georgia, Athens GA, USA was supported in part by NSF-RCN grant # 0090281.

The abbreviations used are :C3ORF21, chromosome 3 open reading frame 21; CID, collision induced dissociation; EGF, epidermal growth factor; ER, endoplasmic reticulum; Fuc, fucose; Gal, galactose; GALT, galactosyltransferase; Glc, glucose; GlcNAc, N-acetylglucosamine; GLT8D, glycosyltransferase 8 family domain; GT8, Glycosyltransferase 8 (family); GXYLT, glucoside-xylosyltransferase; HPLC, high pressure liquid chromatography; HSQC-NMR, Heteronuclear Single Quantum Coherence NMR; Pofut; protein O-fucosyltransferase; Poglut, protein O-glucosyltransferase; Sf9, Spodoptera frugiperda cells; Sia, sialic acid; TOF, time of flight; UGGT, UDP-Glc: glycoprotein glucosyltransferase; XXYLT, xyloside-xylosyltransferase; Xyl, xylose.

REFERENCES

The second Notch xylosyltransferase

FIGURE 1. Sequence alignment of Notch xylosyltransferases and an overview of enzymes involved in Notch glycosylation. A: Amino acid sequence alignment of XXYLT1 with GXYLT1 and GXYLT2. Sequence conservation in at least two of the three aligned enzymes is indicated in black. The predicted transmembrane domain in XXYLT1 is underlined (dashed) as well as the conserved DXD motifs (solid). B: Overview of enzymes involved in O-fucosylation and O-glucosylation of specific consensus sequences on EGF repeats. Shown is one EGF repeat and the enzymes acting on EGF repeats exclusively. GlcNAc-Fuc-O can be further extended by nonspecific galactosyl- and sialyltransferases. Indicated are the official human gene name abbreviations (www.genenames.org) and, in brackets, the *Drosophila* names or abbreviations.
FIGURE 2. *In vitro* enzymatic activity of XXYLT1. **A**: Activity of XXYLT1 compared to GXYLT1 and GLT8D1 with artificial, synthetic acceptors mimicking the naturally occurring carbohydrate substrates on Notch EGF repeats. **B**: Activity of XXYLT1 with different mono-saccharides linked to para-nitrophenol (pNP) compared to Xyl-Glc-R. **C**: UDP-sugar specificity of XXYLT1 using Xyl-Glc-R as acceptor. **D**: Quantification of 0.5 µl IgG-beads coupled enzymes by a Protein A standard stained with Coomassie Blue. For each Assay of **A-C** 10 µl enzyme-beads were used; enzymatic activity is presented as nmol of transferred sugar per nmol of enzyme per hour and values are the mean ± S.E.M of 3 independent measurements.
FIGURE 3. Analysis of the reaction product of XXYLT1 with synthetic Xyl-Glc-R acceptor. Shown are 125 MHz natural abundance $^{13}$C-nuclear magnetic resonance spectra (left column) and reverse phase chromatography profiles (right column) of acceptors and reaction products. A: The pure synthetic reference disaccharide Xyl-Glc-R (Xylα1,3-Glcβ-O-(CH$_2$)$_3$NHCOC$_2$H$_{15}$), B: the reaction product after incubation of 1 mg Xyl-Glc-R with XXYLT1, and C: the synthetic trisaccharide Xyl-Xyl-Glc-R (Xylα1,3-Xylα1,3-Glcβ-O-(CH$_2$)$_3$NHCOC$_2$H$_{15}$). Product analysis by HPLC revealed the conversion of more than 30% of Xyl-Glc-R to a product running at the position of Xyl-Xyl-Glc-R. $^{13}$C-NMR spectra confirmed the identity of signals between the reaction/substrate mixture and signals obtained from pure synthetic standard of the expected product Xyl-Xyl-Glc-R. Additionally, $^1$H NMR and HSQC NMR are presented in supplemental Fig. S1 and S2; chemical shifts are listed in supplemental table S1.
The second Notch xylosyltransferase

**FIGURE 4.** *In vitro* xylosyltransferase activity on Notch EGF1-5. LC-ESI-MS of mouse Notch EGF1-5, produced in Sf9 insect cells and after *in vitro* incubation with the indicated enzymes. **A:** A peptide of EGF4 showed Glc-O and Xyl-Glc-O modifications (m/z 1166 and 1210). **B:** After incubation with GXYLT1 in presence of UDP-Xyl, the Glc-O signal is strongly reduced. **C:** XXYLT1 extended the Xyl-Glc-O disaccharide by one xylose residue, indicated by the appearing signal at m/z 1254. **D:** Incubation of Notch EGF1-5 with both enzymes GXYLT1 and XXYLT1 almost completely shifted the carbohydrate modification to Xyl-Xyl-Glc-O. Peptide sequencing was done by tandem mass spectrometry (MS/MS), verifying that all signals were derived from the indicated peptide (supplemental Fig. S3).
FIGURE 5. Xyl-O-EGF is not modified by Xxylt. Analysis of EGF16 from mouse Notch2 by reverse phase HPLC before modification (A), or after incubation with Rumi and UDP-Glc (B), or Rumi and UDP-Xyl (C). The product from B (Glc-O-EGF) was purified and incubated with mouse Gxylt2 and UDP-xylose to generate Xyl-Glc-O-EGF. Xyl-Glc-O-EGF was incubated with UDP-xylose in the absence (D) or the presence (E) of mouse Xxylt1 and analyzed by reverse phase HPLC. The presence of the glycans on all forms of EGF16 were confirmed by mass spectrometry (supplemental Figure S4). (F) Each of the purified EGF repeats generated in A-D were incubated with mouse Xxylt1 and UDP-[14C]xylose. The values indicate the mean ± S.E.M of 3 independent measurements.
The second Notch xylosyltransferase

FIGURE 6. **XXYL1** shows in vivo activity on Notch EGF repeats. A: Mass spectrometric analysis of Notch EGF1-5 produced in Sf9 cells that were co-infected with the enzyme GLT8D1 as a negative control showed a mixture of Glc-O and Xyl-Glc-O modifications. B: Co-infection with **XXYL1** resulted in de novo generation of the complete trisaccharide at m/z 1254 at the expense of the disaccharide at m/z 1210. There was no further elongation than the trisaccharide observed. Like for figure 4 peptide identity was confirmed by MS/MS (data not shown).
The second Notch xylosyltransferase

FIGURE 7. Intracellular localization and cleavage analysis of XXYLT1. Images A-E show immunostainings of CHO cells that were transiently transfected with tagged plasmids. A: Co-localization of the single tagged constructs Flag-XXYLT1 and XXYLT1-HA. B: Overlaying signals from the N-terminal Flag- and the C-terminal HA-tag of the Flag-XXYLT1-HA construct. C: Co-localization of XXYLT1 visualized by the Flag-tag and a co-transfected GFP1-KDEL as ER marker. D: No merge was observed after staining of tagged XXYLT1 and the Golgi marker mannosidase II. E: As control, the established Golgi located β1,4-galactosyltransferase (Flag-B4GALT1-HA) was investigated in parallel. F: Signal sequence prediction (SignalP 3.0) indicated a high probability for XXYLT1 to have a cleaved signal sequence, which is not predicted anymore in the Flag-tagged construct. G: Lysates of cells transiently expressing Flag-XXYLT1-HA, the control protein Flag-B4GALT1-HA as well as C-terminally HA-tagged enzymes were analyzed by western blotting using either anti-HA or anti-Flag antibody, showing no indication of cleavage of the protein. XXYLT1 partly runs as an SDS resistant dimer.
Molecular Cloning of a Xylosyltransferase that Transfers the Second Xylose to O -Glucosylated Epidermal Growth Factor Repeats of Notch


J. Biol. Chem. published online November 23, 2011

Access the most updated version of this article at doi: 10.1074/jbc.M111.302406

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2011/11/23/M111.302406.DC1