FEBS 26297

FEBS Letters 524 (2002) 211–218

Molecular cloning and characterization of a novel member of the UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase family, pp-GalNAc-T12[1]

Jian-Ming Guo[a], Yan Zhang[a], Lamei Cheng[a], Hiroko Iwasaki[a,b], Han Wang[a], Tomomi Kubota[a], Kouichi Tachibana[a], Hisashi Narimatsu[a,*]

[a] *Glycogene Function Team, Research Center for Glycoscience, National Institute of Advanced Industrial Science and Technology (AIST), Open Space Laboratory C-2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan*

[b] *Amersham Biosciences KK, 3-25-1, Hyakunincho, Shinjuku-ku, Tokyo 169-0073, Japan*

Received 14 May 2002; revised 13 June 2002; accepted 19 June 2002

First published online 2 July 2002

Edited by Masayuki Miyasaka

Abstract **We cloned in silico a novel human UDP-GalNAc:-polypeptide *N*-acetylgalactosaminyltransferase (pp-GalNAc-T), pp-GalNAc-T12. The deduced amino acid sequence of pp-GalNAc-T12 contains all conserved motifs in pp-GalNAc-T family proteins. Quantitative real time polymerase chain reaction analysis revealed that the pp-GalNAc-T12 transcript was expressed mainly in digestive organs such as stomach, small intestine and colon. The recombinant pp-GalNAc-T12 transferred GalNAc to the mucin-derived peptides such as the Muc1a, Muc5AC, EA2 peptides and the GalNAc-Muc5AC glycopeptide. Since mucins are glycoproteins mainly produced in the digestive organs, our results suggest that pp-GalNAc-T12 plays an important role in the initial step of mucin-type oligosaccharide biosynthesis in digestive organs.** **© 2002 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.**

Key words: Glycosyltransferase; *N*-Acetylgalactosaminyltransferase; Mucin; *O*-Glycosylation; *O*-Glycan

1. Introduction

Mucin-type *O*-linked oligosaccharides are linked to a serine or threonine residue on the acceptor protein via an *N*-acetylgalactosamine residue at their reducing end. The enzyme, UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase (pp-GalNAc-T) (EC 2.4.1.41), catalyzes the initial step of mucin-type oligosaccharide biosynthesis by transferring GalNAc from UDP-GalNAc to a hydroxyl amino acid on a polypeptide acceptor [1]. To date, 11 distinct members have been identified in mammals [2–17]. Each isoform exhibits a different tissue distribution and substrate specificity. However, the substrate specificity of each isoform and the hierarchical network of glycosyltransferases are poorly understood.

A search of the human and murine expressed sequence tag (EST) database revealed numerous new sequence homologues, suggesting that an even larger and more complex pp-GalNAc-T family exists in mammals [12]. Thus, the isolation and characterization of pp-GalNAc-Ts are essential for the further study of *O*-glycosylation.

Here, we report they, pp-GalNAc-T12. The pp-GalNAc-T12 transcript was expressed mainly in digestive organs, and the recombinant pp-GalNAc-T12 transferred GalNAc to tandem repeats of mucins in vitro.

2. Materials and methods

2.1. Cloning of human pp-GalNAc-T12 cDNA

By performing a BLAST search of EST databases, we found three cDNAs (GenBank accession number: AK024865, BE677813, AI800923) homologous to the open reading frame (ORF) of human pp-GalNAc-T4 (Y08564). On searching the human genomic DNA database, we found that a genomic sequence (AL136084) contained a putative initial codon and a transmembrane domain at positions 5286–5397 and aligned to BE677813 at positions 5484–5685. Based on the database gene working, the full ORF of this putative pp-GalNAc-T was predicted to be 1746 bp. The full-length sequence of this putative pp-GalNAc-T was deposited into GenBank with the accession number AB078146.

The cDNA encoding the full-length ORF was obtained by polymerase chain reaction (PCR) using the Expand™ High Fidelity PCR system (Roche), Marathon-Ready™ cDNA of human lung (Clontech) as a template, and primers (forward 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAGATAGAACCATGTGGGGGCGCACGGCGCGGCGG-3′ and reverse 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATAACATGCGCTCTTTGAAGAACCA-3′). PCR was performed according to the manual supplied. The forward and reverse primers have *att*B1 and *att*B2 overhangs, respectively, to create the recombination sites. The amplified fragment was subcloned into the pDONR™201 vector using the BP CLONASE Enzyme Mix (Invitrogen). Finally, an entry clone containing a complete ORF, sharing complete homology with the EST clone, was obtained and given the designation pp-GalNAc-T12.

2.2. Quantitative analysis of the pp-GalNAc-T12 transcripts in human tissues and cell lines by real time PCR

For the quantification of pp-GalNAc-T12 transcripts, we employed the quantitative real time PCR method using TaqMan® Universal PCR Master Mix and ABI PRISM 7700 Sequence Detection System (Applied Biosystems), as described in detail previously [18]. Marathon-Ready® cDNAs of various human tissues were purchased

*Corresponding author. Fax: (81)-298-61 3201.
E-mail address: h.narimatsu@aist.go.jp (H. Narimatsu).

[1] The nucleotide sequence(s) of human pp-GalNAc-T12 reported in this paper has been submitted to the GenBank™/DDBJ/EBI DataBank with the accession number AB078146.

Abbreviations: pp-GalNAc-T, UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase; EST, expressed sequence tag; ORF, open reading frame; PCR, polymerase chain reaction; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; HPLC, high performance liquid chromatography; FAM, 5-carboxyfluorescein succinimidyl ester

PII: S0014-5793(02)03007-7

Table 1
Acceptor substrates of peptides

Non-glycosylated peptide	One GalNAc-glycosylated peptide (by pp-GalNAc-T)	Peptide sequence
Muc1a	GalNAc-Muc1a (T6)	FAM-AHGVT*SAPDTR
Muc2	GalNAc-Muc2 (T2)	PTTTPITTTTTVTPTPTPTGTQTK-FAM
Muc5AC	GalNAc-Muc5AC (T2)	FAM-GTTPSPVPTTSTTSA
Muc7	GalNAc-Muc7 (T13)**	PTPSATT*PAPPSSSAPPETTAAK-FAM
EA2	GalNAc-EA2 (T13)**	PTTDSTT*PAPTTK-FAM

*: Represents the site where one GalNAc residue attached. Unmarked means not identified. All of the glycopeptides were identified by mass.
**: Refers to a novel human pp-GalNAc-T (Zhang and Narimatsu, unpublished).

from Clontech. The cDNAs from various tumor cell lines were generated in our laboratory [19].

The primer set and probe for pp-GalNAc-T12 were as follows: the forward primer, 5′-CGTGCCCGCTTGGAAC-3′, the reverse primer, 5′-TGTTCTGGAGCATCCCGAAG-3′, and the probe, 5′-TGGGGATGTGACAGAGAGGAAGCAGCT-3′ with a minor groove binder [20]. The PCR conditions included one cycle at 50°C for 2 min, one cycle at 95°C for 10 min, and 50 cycles at 95°C for 15 s, and 60°C for 1 min. The relative amount of pp-GalNAc-T12 transcript was normalized by the amount of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) transcript in the same cDNA.

2.3. Production of recombinant pp-GalNAc-T12

To develop the recombinant pp-GalNAc-T12 encoding the putative catalytic domain, the truncated pp-GalNAc-T12 cDNA, starting at amino acid position 60, was generated by PCR using primers (forward 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCGGCGCGAGCCGGTCATGCCGCGGC-3′ and reverse 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATAACATGCGCTCTTTGAAGAACCA-3′). PCR was carried out with these primers, human lung Marathon-Ready™ cDNA as a template, and an Expand™ High Fidelity PCR system. The amplified DNA fragment was subcloned into pDONR™201 using *att*B1 and *att*B2 sequences in primers as described previously [18]. Then, the truncated cDNA fragment of pp-GalNAc-T12 was transferred between the *att*R1 and *att*R2 sites of an expression vector, pFBIF, using the LR CLONASE Enzyme Mix (Invitrogen) from pDONR™201 vector.

Bacmid DNA from pFBIF-GalNAc-T12-transformed $DH10_{BAC}$ competent cells (Invitrogen) was transfected into Sf21 insect cells (BD Pharmingen). Recombinant virus was prepared according to the instruction manual of the BAC to BAC Baculovirus Expression Systems (Invitrogen). Sf21 insect cells were infected with recombinant virus to yield secreted recombinant pp-GalNAc-T12 protein. Then, the culture media were harvested and incubated overnight at 4°C

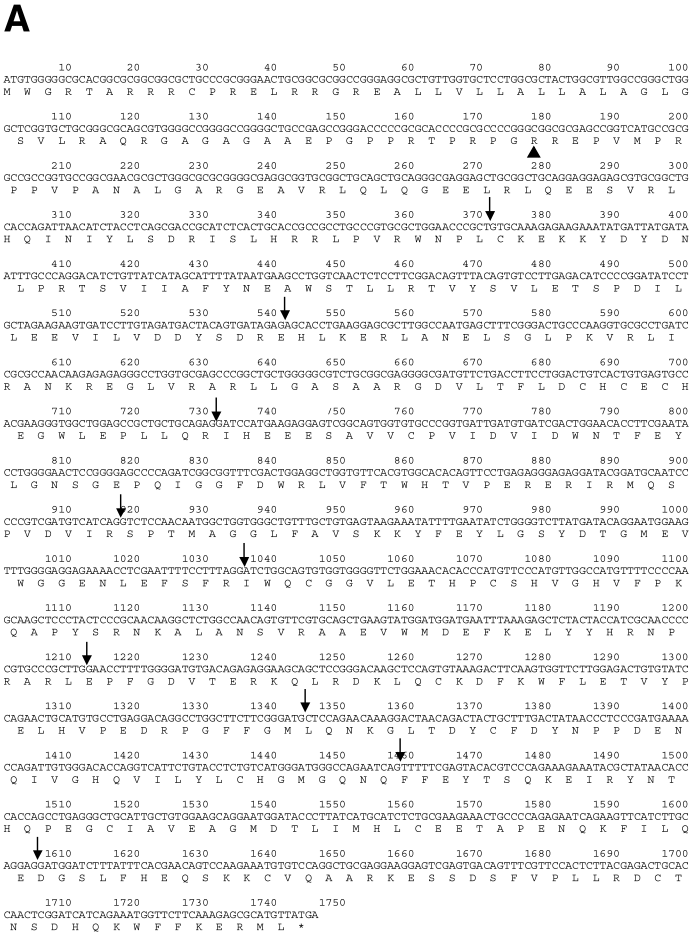

Fig. 1. A: Nucleotide and amino acid sequence of pp-GalNAc-T12. The nucleotides are numbered from the first nucleotide of the initiation codon. The predicted amino acid sequence is presented below the putative nucleotide coding sequence. A triangle represents the truncated position. Positions of exon–intron junctions are indicated by vertical arrows. B: ClustalW alignment for comparison of pp-GalNAc-T12 with the other pp-GalNAc-Ts. Multiple sequence analysis (ClustalW) of the human GalNAc-T. The putative transmembrane domains are underlined. Asterisks indicate the amino acids identical among all proteins. Three motifs of GT1, Gal/GalNAc transferase and lectin motif are indicated above the sequences underlined. Conserved amino acids are shown by dots.

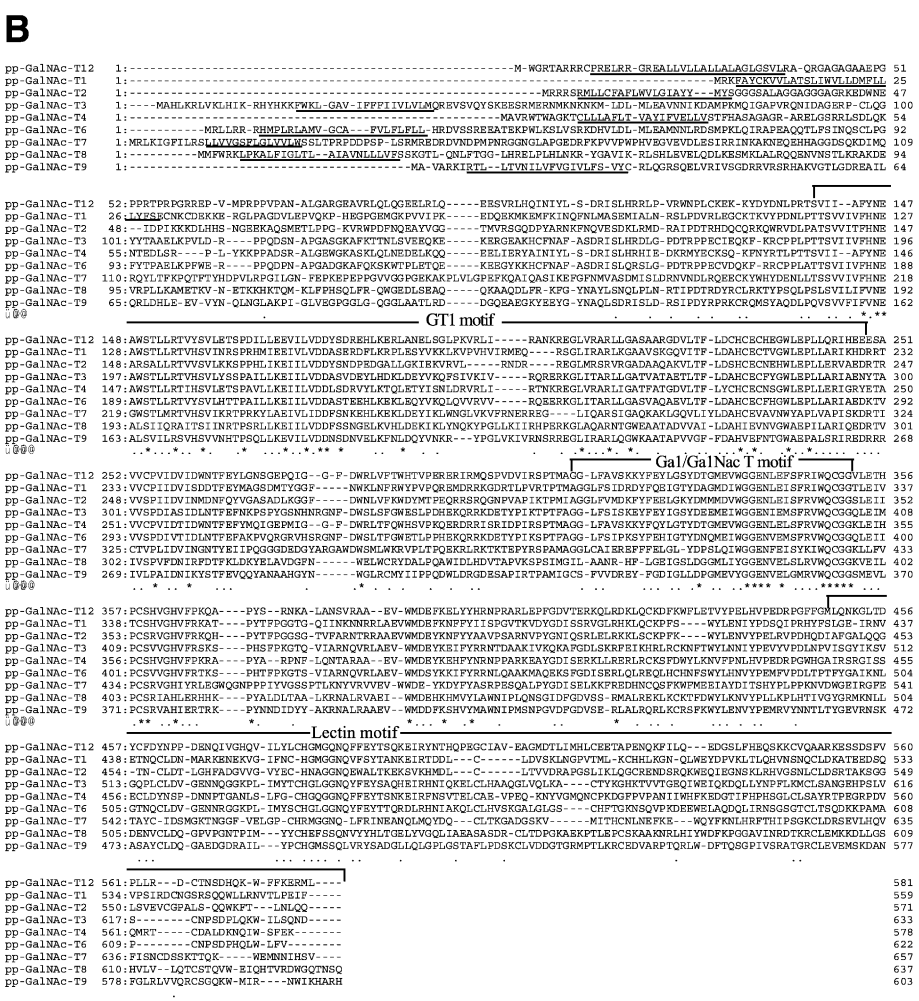

Fig. 1 (*Continued*).

with anti-FLAG M1 antibody-conjugated resin (Sigma). The protein–resin mixtures were washed twice in the buffer containing 50 mM Tris–HCl, pH 7.4, 150 mM NaCl, and 1 mM $CaCl_2$. Then the recombinant protein was eluted with 50 mM Tris-buffered saline containing 1 mM EDTA.

The soluble forms of the truncated human pp-GalNAc-T2 and -T6, containing amino acid residues 52–571 (T2) and 53–622 (T6), respectively [2,5], were also produced by the same method.

The purified recombinant enzymes were detected by Western blotting with anti-FLAG monoclonal antibody after sodium dodecyl sulfate–polyacrylamide gel electrophoresis separation to determine relative amounts. Each purified pp-GalNAc-T was adjusted to the same amount according to the results of Western blotting, and was subjected to pp-GalNAc-T assay.

2.4. Analysis of pp-GalNAc-T activity of recombinant human pp-GalNAc-T12

Reactions for the pp-GalNAc-T assay were performed at 37°C in a 20 μl mixture containing 25 mM Tris–HCl (pH 7.4), 5 mM $MnCl_2$, 0.2% Triton X-100, 0.25 mM UDP-GalNAc (Sigma), an appropriate 5-carboxyfluorescein succinimidyl ester (FAM)-labeled acceptor substrate, and the purified enzyme. 50 pmol of acceptor substrate and 125 ng of enzyme were used for time course reactions (Fig. 4A,B); and 1/10 of the substrates and enzymes were used for primary and secondary enzyme reactions (Fig. 4C). The reaction was terminated by boiling for 3 min, and then subjected to High performance liquid chromatography (HPLC) on a reverse C18 column (Cosmosil $5C_{18}$-AR, 4.6×250 mm) as previously described [18]. FAM-labeled oligopeptides, that represent the tandem repeats of Muc1, Muc2, Muc5AC, Muc7, and EA2 [21–25], respectively, were purchased from Sawady Co., Ltd. (Tokyo, Japan), and were used as acceptor substrates. Peptides of Muc1a, Muc2, Muc5AC, Muc7 and EA2 with one GalNAc residue attached, prepared by glycosylation using pp-GalNAc-T2, -T6, or -T13 (Zhang and Narimatsu, unpublished) followed by purification by HPLC, were used as GalNAc-glycopeptide acceptor substrates for the pp-GalNAc-T assay (Table 1).

2.5. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-TOF) and amino acid sequencing of glycopeptide products

In order to determine the glycosylation site of the product of the pp-GalNAc-T12 reaction, the glycopeptide isolated by HPLC was subjected to MALDI-TOF mass spectrometry using REFLEX™ IV (Bruker Daltonics, Tsukuba, Japan) and Edman degradation on a protein sequencer PPSQ-23A (Shimadzu). The methods for mass spectrometry (MALDI-TOF) and amino acid sequencing were described in our previous paper [18].

3. Results

3.1. Isolation of pp-GalNAc-T12 cDNA

The cDNA of a new member of the pp-GalNAc-Ts family was cloned by the method described in Section 2, and designated human pp-GalNAc-T12. The nucleotide and the deduced amino acid sequences of pp-GalNAc-T12 are shown in Fig. 1A, and an alignment of the amino acid sequences of nine pp-GalNAc-Ts was generated using ClustalW as shown in Fig. 1B. The pp-GalNAc-T12 gene encodes 581

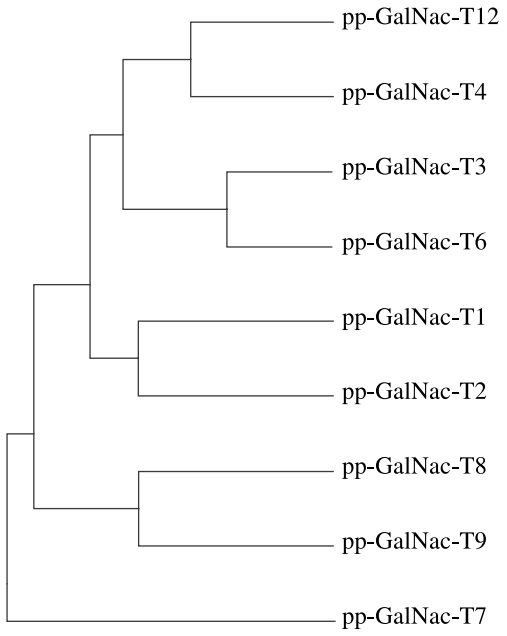

Fig. 2. A phylogenetic tree of pp-GalNAc-Ts. A phylogenetic tree of pp-GalNAc-Ts was constructed by the neighbor joining method based on amino acid sequences. The branch length indicates the evolutionary distance between each member.

amino acids, a typical type II membrane protein, comprising an *N*-terminal cytoplasmic domain, a transmembrane domain, a putative stem region and a catalytic domain containing a GT1 motif, Gal/GalNAc transferase motif and ricin-like lectin motif (Fig. 1B) [26]. These motifs, commonly observed in the pp-GalNAc-T family, were present in the predicted positions in the putative amino acid sequence of pp-GalNAc-T12.

The *pp-GalNAc-T12* gene was localized in a draft genome sequence (GenBank™ accession number AL136084) that was mapped to human chromosome 9 at q22. The ORF of pp-GalNAc-T12 was composed of 10 exons.

The phylogenic tree of pp-GalNAc-Ts was generated by the neighbor joining method [27] based on the amino acid sequences (Fig. 2). The branch length indicates the evolutionary distance between family members. pp-GalNAc-T12 shares approximately 32∼56% homology with the other mammalian pp-GalNAc-Ts in amino acid sequence, and is most homologous to human pp-GalNAc-T4.

3.2. Distribution of pp-GalNAc-T12 transcripts in human tissues and cell lines

The expression levels of the pp-GalNAc-T12 transcript in various human tissues and cell lines were determined by the real time PCR method. pp-GalNAc-T12 was expressed in most tissues (Fig. 3A) and in various cell lines (Fig. 3B), although at different levels. Tissues with a considerably high level of expression were the small intestine, stomach, pancreas, and colon. Moderate expression was observed in the testis, thyroid gland and spleen. The expression levels in the remaining tissues tested were very low or undetectable. pp-GalNAc-T12 was expressed at a relatively low level or undetectable in the cell lines tested. The expression levels of pp-GalNAc-T12 in colon and stomach cancer cell lines (Fig. 3B) were lower than those in colon and stomach (Fig. 3A), respectively.

3.3. Different substrate specificities of pp-GalNAc-T12 from those of the known pp-GalNAc-Ts

To analyze the pp-GalNAc-T activity, the recombinant pp-GalNAc-T12, an approximately 66 kDa secreted protein, was produced in Sf21 cells, purified from culture supernatant, and subjected to pp-GalNAc-T assay. The purified recombinant pp-GalNAc-T12, UDP-GalNAc, and a panel of acceptor substrates were incubated for various periods, and the products generated were analyzed by HPLC.

First, Muc1a-derived peptide (FAM-AHGVTSAPDTR) was incubated with pp-GalNAc-T12, and was analyzed by HPLC. As demonstrated in Fig. 4A, the substrate peak (shown as 'S') was gradually consumed and another peak (shown as 'P') appeared as the incubation time for either GalNAc-T reaction was extended, indicating that pp-GalNAc-T12 transferred GalNAc to Muc1a peptide. Incubation with pp-GalNAc-T12 generated only one shifted peak for the Muc1a peptide, suggesting that only one Ser/Thr residue was glycosylated by T12. The product of pp-GalNAc-T12 was isolated by HPLC and subjected to MALDI-TOF mass spectrometry and protein sequencing. The product was identified to have one GalNAc at the fifth threonine (data not shown).

In contrast to the case of the Muc1a peptide, the reaction products of pp-GalNAc-T12 toward the Muc5AC-derived peptide (FAM-GTTPSPVPTTSTTSA) showed multiple peaks as demonstrated in Fig. 4B. One tall shifted peak appeared on HPLC of the Muc5AC peptide after a short incubation with pp-GalNAc-T12. This peak was gradually consumed and multiple shifted peaks appeared as the incubation time was increased. This result indicates that pp-GalNAc-T12 transferred GalNAc to multiple sites in the Muc5AC peptide [28].

To further identify the catalytic specificity of pp-GalNAc-T12, we assayed peptides derived from EA2, Muc2, and Muc7 (Fig. 4C-1). pp-GalNAc-T12 exhibited enzymic activity toward EA2, but negligible activity toward Muc2 and Muc7.

pp-GalNAc-T12 transferred GalNAc to multiple sites in Muc5AC, suggesting that pp-GalNAc-T12 has a secondary GalNAc-T activity. To elucidate the secondary GalNAc-T activity of pp-GalNAc-T12, we generated mono-GalNAc-transferred mucin peptides and used them for the GalNAc-T assay with pp-GalNAc-T12. As demonstrated in Fig. 4C-2, the incubation with pp-GalNAc-T12 generated shifted peaks only in the GalNAc-Muc5AC peptide, whereas T12 generated no or very few shifted peak in GalNAc-transferred EA2, Muc1a, Muc2, and Muc7. These results indicate that pp-GalNAc-T12 has a different catalytic specificity from pp-GalNAc-T4 and -T7 which have a strong secondary activity as reported by others [4,6].

Thus, pp-GalNAc-T12 exhibited comparable GalNAc-T activity toward non-glycosylated peptides: Muc5AC, Muc1a and EA2, and no detectable activity with Muc2 and Muc7. pp-GalNAc-T12 showed enzymatic activity toward the GalNAc-Muc5AC glycopeptide, but no detectable activity to mono-GalNAc-glycosylated Muc1a, Muc2, Muc7 and EA2. These findings depict a novel and unique catalytic specificity of pp-GalNAc-T12.

4. Discussion

The initiation of *O*-glycosylation is catalyzed by a family of pp-GalNAc-Ts. Studies over the past decade have found a

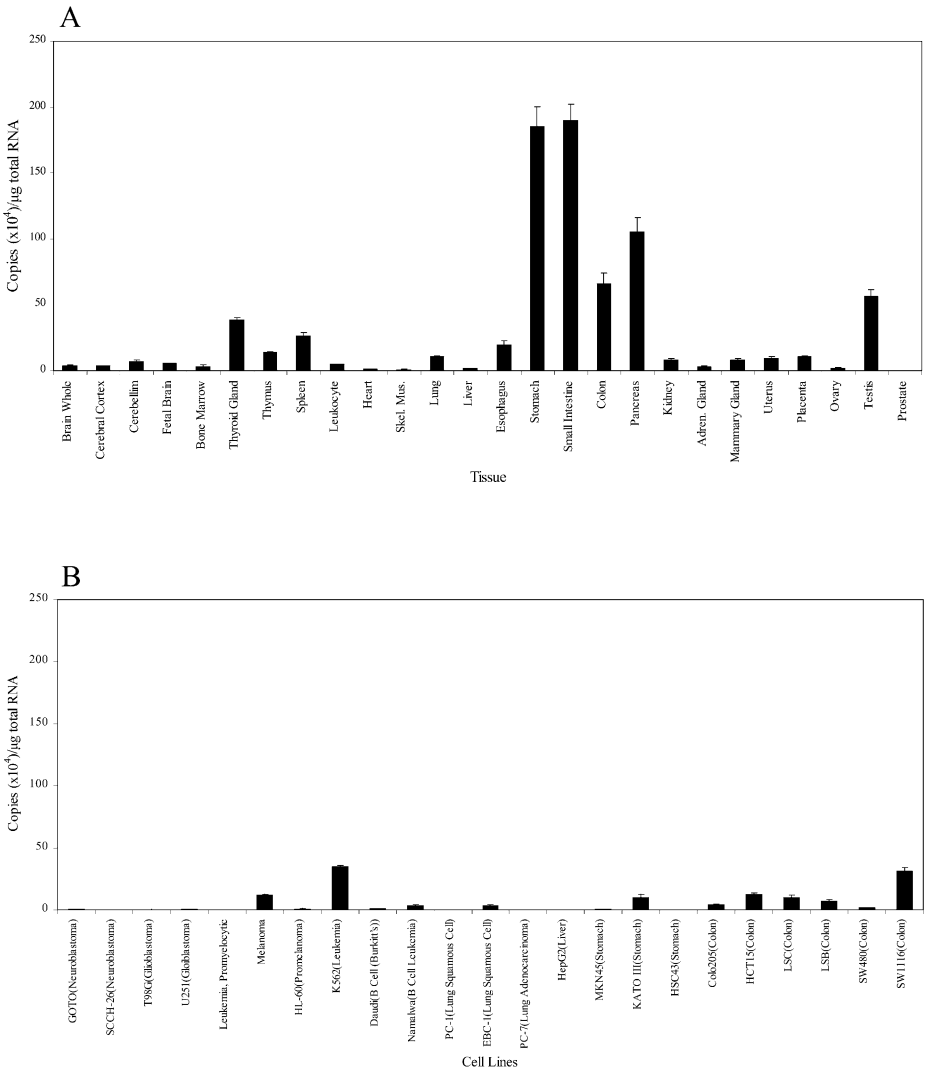

Fig. 3. Quantitative analysis of the pp-GalNAc-T12 transcript in human tissue (A) and tumor cell lines (B) by real time PCR. Standard curves for pp-GalNAc-T12 and GAPDH were generated by serial dilution of each plasmid DNA. The expression level of the pp-GalNAc-T12 transcript was normalized to that of the GAPDH transcript which was measured in the same cDNAs. Values are expressed as copy numbers of the target gene in 1 μg of total RNA. Data were obtained from triplicate experiments and are indicated as the mean ± S.D.

family of pp-GalNAc-Ts that display differences in both their substrate specificity and their patterns of tissue distribution.

In the present study, we have cloned and characterized a novel isoform of the pp-GalNAc-T family, designated pp-GalNAc-T12. Although pp-GalNAc-T12 was expressed in various tissues and cell lines, the expression level of its gene was especially high in digestive organs such as the small intestine, stomach, pancreas, and colon, suggesting that potential substrates of pp-GalNAc-T12 are mucin proteins which are mainly secreted from the digestive tract [29].

Each pp-GalNAc-T displays a unique pattern of expression across human and rodent tissues; pp-GalNAc-T1, -T2, -T3 and -T6 are expressed in a broad range of tissues; while others are more restricted in their expression patterns. For example, the pp-GalNAc-T4 transcript is highly expressed in rodent sublingual gland, stomach, and colon, with lower levels in the small intestine, urogenital track, and lung [14]. The expression of pp-GalNAc-T7 is similar to that of pp-GalNAc-T4 [16].

The expression level of pp-GalNAc-T12 was low in tumor cell lines. Consistent with the result for pp-GalNAc-T12, pp-GalNAc-T1 and -T2 were also expressed at lower levels in cell lines derived from the pancreas, colon, stomach, and breast compared to the original tissues [30].

pp-GalNAc-T isoforms have catalytic specificity toward acceptor substrates, although such specificity overlaps and each enzyme may accommodate a rather broad range of substrates [31]. Prior glycosylation of a specific serine or a threonine in a peptide substrate also influences the acquisition of sugar at vicinal positions [32], strongly suggesting that *O*-glycosylation of multisite substrates proceeds in a specific hierarchical manner. Therefore, one can classify a pp-GalNAc-T through the determination of its catalytic specificity. Synthetic peptides have been commonly used as acceptor substrates for in vitro *O*-glycosylation reactions to determine the substrate specificity of pp-GalNAc-Ts [31].

In the present study, pp-GalNAc-T12 showed primary activity when the peptide (AHGVTSAPDTR) derived from Muc1a was used as a substrate [21,33]. Muc1 is expressed on the apical surface of mucous epithelial cells almost

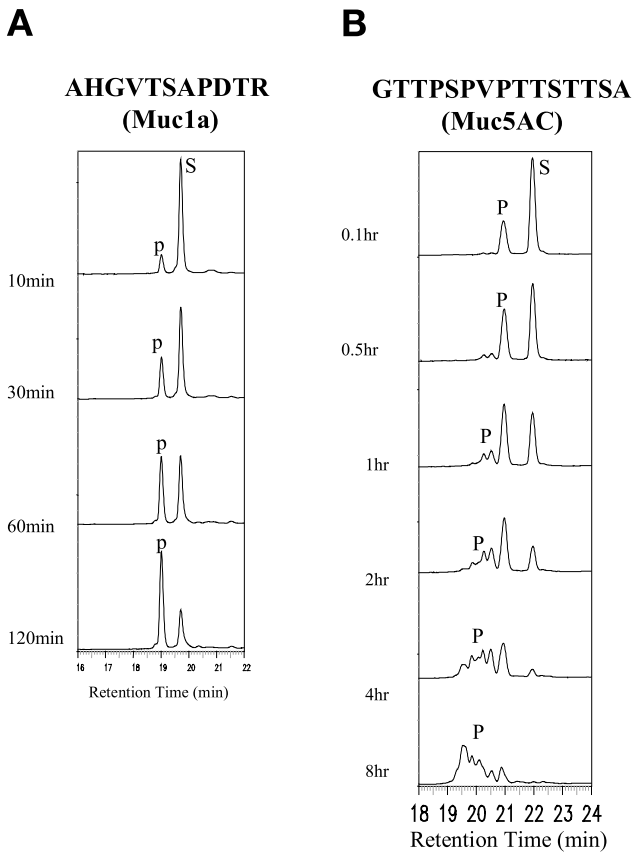

Fig. 4. A: HPLC analysis of in vitro *O*-glycosylation of Muc1a by purified recombinant pp-GalNAc-T12. The panel shows the elution profiles of Muc1a peptide glycosylated by pp-GalNAc-T12. Peak S corresponds to the original substrate, non-glycosylated Muc1a peptide. Peak P corresponds to the products, having a single GalNAc at the fifth threonine, of the enzyme reaction. The incubated time of recombinant enzymes is indicated at the left of the panels. B: HPLC analysis of in vitro *O*-glycosylation of Muc5AC by purified recombinant pp-GalNAc-T12. The panel shows the elution profiles of Muc5AC peptide glycosylated by pp-GalNAc-T12. Peak S corresponds to the original substrate, non-glycosylated peptide. Peak P corresponds to the products of the enzyme reaction. The incubated time of recombinant enzymes is demonstrated. C: HPLC analysis of in vitro *O*-glycosylation of a panel of non-glycosylated peptides and mono-glycosylated glycopeptide by pp-GalNAc-T12. Panels 1 and 2 show the elution profiles of non-glycosylated peptide and mono-glycosylated glycopeptide before and after incubation with pp-GalNAc-T12 to 6 h, respectively. Peak S corresponds to the original substrate, non-glycosylated peptide. Peak P corresponds to the products of the enzyme reaction.

throughout the body [34]. It was reported that the glycosylation of the Muc1a peptide by human GalNAc-T1, -T2, and -T3 resulted in only one glycoform that was glycosylated at Thr in GVTSA as evaluated by MALDI-TOF and amino acid sequencing [21,32]. The remaining sites, Ser in GVTSA and Thr in PDTR, cannot be glycosylated by these enzymes, but it has been shown that the 'fill-up' reactions are catalyzed by pp-GalNAc-T4, which has a preference for peptides partially modified by GalNAc incorporations by other pp-GalNAc-Ts [4,6]. pp-GalNAc-T12 generated only one product from the Muc1a peptide, that is, the same product produced by pp-GalNAc-T1, -T2 and -T3. pp-GalNAc-T12 has only primary activity toward the Muc1a peptide like pp-GalNAc-T1, -T2 and -T3 [21].

The Muc5AC gene is also highly expressed in the gastrointestinal tract, the lungs, the nervous system and bone [35]. The mucin motif peptide substrate (GTTPSPVPTTSTTSA) deduced from the human Muc5AC gene [23] consists of six threonine residues including a XTPXP sequence that was reported as a signal favoring efficient glycosylation. It was reported that pp-GalNAc-T1, -T2, and -T3 transferred two, one and three GalNAcs to Muc5AC, respectively [36]. Our results, accumulation of the first product in Fig. 4B, indicate that pp-GalNAc-T12 transfers GalNAc rapidly to one preferential site in the Muc5AC peptide, and then slowly to the other residues. Because the peaks generated by pp-GalNAc-T12 toward the naked Muc5AC peptide and the mono-GalNAc-Muc5AC were similar, the first glycosylation by pp-GalNAc-T2 did not affect the primary pp-GalNAc-T12's activity. It is suggested that pp-GalNAc-T12 prefers a different residue in the Muc5AC peptide from the residue preferred by pp-GalNAc-T2.

pp-GalNAc-T12 showed comparative levels of activity toward naked and mono-GalNAc-Muc5AC peptides. However, pp-GalNAc-T12 showed very weak or no activity for GalNAc-EA2 or GalNAc-Muc1a, whereas pp-GalNAc-T12 showed relatively high activity toward naked EA2 or Muc1a. These results suggested that the presence of one or two GalNAc-Thr moieties in the substrate significantly reduced this enzymic activity, consistent with the report of a significant reduction in pp-GalNAc-T1 enzymic activity on prior glycosylation of the Muc2 glycopeptide [37].

pp-GalNAc-T12 shares approximately 56% homology with the human pp-GalNac-T4. However, the acceptor substrate

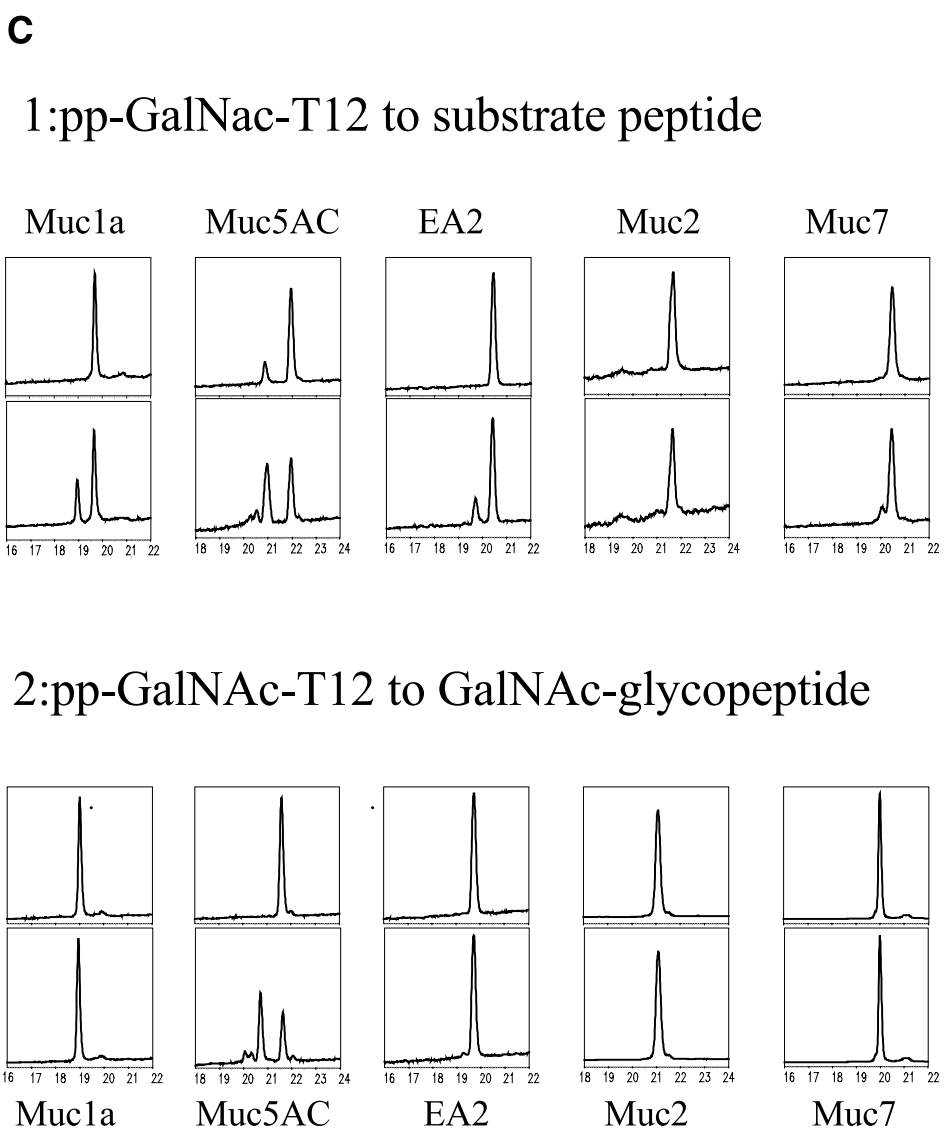

Fig. 4 (*Continued*).

specificities of the two isoforms are totally different, and showed different patterns of activity with these naked or glycosylated peptide substrates. Although pp-GalNAc-T4 showed very weak activity for naked Muc1 and Muc5AC peptides and a weak activity with Muc2 peptide, pp-GalNAc-T4 showed significantly enhanced activity for GalNAc-incorporated peptide substrates, indicating that pp-GalNAc-T4 can utilize the remaining acceptor site [4,6]. In contrast, pp-GalNAc-T12 showed negligible activity or no enhancement of activity with the GalNAc-Muc1a, -Muc2 and -Muc7 glycopeptides. pp-GalNAc-T4 also utilized non-glycosylated Muc7 peptide, but demonstrated no effect of prior glycosylation on the Muc7 peptide, while pp-GalNAc-T12 showed very weak primary activity with Muc7 and no detectable activity for GalNAc-Muc7 glycopeptide [6]. Thus, human pp-GalNAc-T12 is clearly distinguishable from pp-GalNAc-T4.

In summary, pp-GalNAc-T12 has primary pp-GalNAc-T activity toward the peptide substrates Muc1a, Muc5AC and EA2; as well as mono-GalNAc-Muc5AC generated through glycosylation by pp-GalNAc-T2. pp-GalNAc-T12 can also transfer GalNAc to multiple sites in the Muc5AC peptide. In contrast, pp-GalNAc-T12 has very weak or no pp-GalNAc-T activity toward naked and mono-GalNAc-Muc2 and -Muc7, mono-GalNAc-Muc1a and -EA2. The substrate specificity of pp-GalNAc-T12 illustrates the novel catalytic specificity among the pp-GalNAc-T family, and suggests a role for pp-GalNAc-T12 in the hierarchical network of *O*-glycosylation.

Acknowledgements: This work was performed as a part of the R&D Project of the Industrial Science and Technology Frontier Program (R&D for Establishment and Utilization of a Technical Infrastructure for Japanese Industry) supported by the New Energy and Industrial Technology Development Organization (NEDO).

References

[1] Clausen, H. and Bennett, E.P. (1996) Glycobiology 6, 635–646.
[2] White, T., Bennett, E.P., Takio, K., Sorensen, T., Bonding, N. and Clausen, H. (1995) J. Biol. Chem. 270, 24156–24165.
[3] Bennett, E.P., Hassan, H. and Clausen, H. (1996) J. Biol. Chem. 271, 17006–17012.
[4] 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.
[5] Bennett, E.P., Hassan, H., Mandel, U., Hollingsworth, M.A., Akisawa, N., Ikematsu, Y., Merkx, G., van Kessel, A.G., Olofsson, S. and Clausen, H. (1999) J. Biol. Chem. 274, 25362–25370.
[6] Bennett, E.P., Hassan, H., Hollingsworth, M.A. and Clausen, H. (1999) FEBS Lett. 460, 226–230.
[7] White, K.E., Lorenz, B., Evans, W.E., Meitinger, T., Strom, T.M. and Econs, M.J. (2000) Gene 246, 347–356.
[8] Toba, S., Tenno, M., Konishi, M., Mikami, T., Itoh, N. and Kurosaka, A. (2000) Biochim. Biophys. Acta 1493, 264–268.
[9] Schwientek, T.J., Bennett, E.P., Flores, C., Thacker, J., Hollman, M., Reis, C.A., Behrens, J., Mandel, U., Keck, B., Schafer, M.A., Hazelmann, K., Zubarev, R., Roepstorff, P., Hollingsworth, M.A. and Clausen, H. (2002) J. Biol. Chem.
[10] Hagen, F.K., Van Wuyckhuyse, B. and Tabak, L.A. (1993) J. Biol. Chem. 268, 18960–18965.
[11] Homa, F.L., Hollander, T., Lehman, D.J., Thomsen, D.R. and Elhammer, A.P. (1993) J. Biol. Chem. 268, 12609–12616.
[12] Hagen, F.K., Gregoire, C.A. and Tabak, L.A. (1995) Glycoconjug. J. 12, 901–909.
[13] Zara, J., Hagen, F.K., Ten Hagen, K.G., Van Wuyckhuyse, B.C. and Tabak, L.A. (1996) Biochem. Biophys. Res. Commun. 228, 38–44.
[14] Hagen, F.K., Ten Hagen, K.G., Beres, T.M., Balys, M.M., VanWuyckhuyse, B.C. and Tabak, L.A. (1997) J. Biol. Chem. 272, 13843–13848.
[15] Ten Hagen, K.G., Hagen, F.K., Balys, M.M., Beres, T.M., Van Wuyckhuyse, B. and Tabak, L.A. (1998) J. Biol. Chem. 273, 27749–27754.
[16] Ten Hagen, K.G., Tetaert, D., Hagen, F.K., Richet, C., Beres, T.M., Gagnon, J., Balys, M.M., VanWuyckhuyse, B., Bedi, G.S., Degand, P. and Tabak, L.A. (1999) J. Biol. Chem. 274, 27867–27874.
[17] Ten Hagen, K.G., Bedi, G.S., Tetaert, D., Kingsley, P.D., Hagen, F.K., Balys, M.M., Beres, T.M., Degand, P. and Tabak, L.A. (2001) J. Biol. Chem. 276, 17395–17404.
[18] Iwai, T., Inaba, N., Naundorf, A., Zhang, Y., Gotoh, M., Iwasaki, H., Kudo, T., Togayachi, A., Ishizuka, Y., Nakanishi, H. and Narimatsu, H. (2002) J. Biol. Chem. 277, 12802–12809.
[19] Togayachi, A., Akashima, T., Ookubo, R., Kudo, T., Nishihara, S., Iwasaki, H., Natsume, A., Mio, H., Inokuchi, J., Irimura, T., Sasaki, K. and Narimatsu, H. (2001) J. Biol. Chem. 276, 22032–22040.
[20] Kutyavin, I.V., Afonina, I.A., Mills, A., Gorn, V.V., Lukhtanov, E.A., Belousov, E.S., Singer, M.J., Walburger, D.K., Lokhov, S.G., Gall, A.A., Dempcy, R., Reed, M.W., Meyer, R.B. and Hedgpeth, J. (2000) Nucleic Acids Res. 28, 655–661.
[21] Wandall, H.H., Hassan, H., Mirgorodskaya, E., Kristensen, A.K., Roepstorff, P., Bennett, E.P., Nielsen, P.A., Hollingsworth, M.A., Burchell, J., Taylor-Papadimitriou, J. and Clausen, H. (1997) J. Biol. Chem. 272, 23503–23514.
[22] Gum, J.R., Byrd, J.C., Hicks, J.W., Toribara, N.W., Lamport, D.T. and Kim, Y.S. (1989) J. Biol. Chem. 264, 6480–6487.
[23] Guyonnet Duperat, V., Audie, J.P., Debailleul, V., Laine, A., Buisine, M.P., Galiegue-Zouitina, S., Pigny, P., Degand, P., Aubert, J.P. and Porchet, N. (1995) Biochem. J. 305 (Pt. 1), 211–219.
[24] Bobek, L.A., Tsai, H., Biesbrock, A.R. and Levine, M.J. (1993) J. Biol. Chem. 268, 20563–20569.
[25] Albone, E.F., Hagen, F.K., VanWuyckhuyse, B.C. and Tabak, L.A. (1994) J. Biol. Chem. 269, 16845–16852.
[26] Hagen, F.K., Hazes, B., Raffo, R., deSa, D. and Tabak, L.A. (1999) J. Biol. Chem. 274, 6797–6803.
[27] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.
[28] Tetaert, D., Ten Hagen, K.G., Richet, C., Boersma, A., Gagnon, J. and Degand, P. (2001) Biochem. J. 357, 313–320.
[29] Reid, C.J. and Harris, A. (1998) Gut 42, 220–226.
[30] Inoue, M., Takahashi, S., Yamashina, I., Kaibori, M., Okumura, T., Kamiyama, Y., Vichier-Guerre, S., Cantacuzene, D. and Nakada, H. (2001) Cancer Res. 61, 950–956.
[31] Elhammer, A.P., Kezdy, F.J. and Kurosaka, A. (1999) Glycoconjug. J. 16, 171–180.
[32] Hanisch, F.G., Muller, S., Hassan, H., Clausen, H., Zachara, N., Gooley, A.A., Paulsen, H., Alving, K. and Peter-Katalinic, J. (1999) J. Biol. Chem. 274, 9946–9954.
[33] Nishimori, I., Johnson, N.R., Sanderson, S.D., Perini, F., Mountjoy, K., Cerny, R.L., Gross, M.L. and Hollingsworth, M.A. (1994) J. Biol. Chem. 269, 16123–16130.
[34] Gendler, S.J., Spicer, A.P., Lalani, E.N., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Boshell, M. and Taylor-Papadimitriou, J. (1991) Am. Rev. Respir. Dis. 144, S42–S47.
[35] Kingsley, P.D., Hagen, K.G., Maltby, K.M., Zara, J. and Tabak, L.A. (2000) Glycobiology 10, 1317–1323.
[36] Tetaert, D., Richet, C., Gagnon, J., Boersma, A. and Degand, P. (2001) Carbohydr. Res. 333, 165–171.
[37] Brockhausen, I., Toki, D., Brockhausen, J., Peters, S., Bielfeldt, T., Kleen, A., Paulsen, H., Meldal, M., Hagen, F. and Tabak, L.A. (1996) Glycoconjug. J. 13, 849–856.