Characterization of a novel human UDP-GalNAc transferase, pp-GalNAc-T10[1]

Lamei Cheng[a], Kouichi Tachibana[a], Yan Zhang[a], Jian-ming Guo[a], Kahori Kahori Tachibana[a], Akihiko Kameyama[a,b], Han Wang[a], Toru Hiruma[a,c], Hiroko Iwasaki[a,b], Akira Togayachi[a,d], Takashi Kudo[a,d], Hisashi Narimatsu[a,*]

[a] *Glycogene Function Team, Research Center for Glycoscience, National Institute of Advanced Industrial Science and Technology (AIST), Central-2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan*
[b] *Amersham Biosciences KK, 3-25-1, Hyakunincho, Shinjuku-ku, Tokyo 169-0073, Japan*
[c] *Fujirebio Inc., Diagnostics Research Laboratories, 51 Komiyacho, Hachioji, Tokyo 192-0031, Japan*
[d] *New Energy and Industrial Technology Development Organization (NEDO), Sunshine 60 Building, 3-1-1 Higashi Ikebukuro, Toshima-ku, Tokyo 170-6028, Japan*

Received 10 July 2002; revised 6 September 2002; accepted 6 September 2002

First published online 3 October 2002

Edited by Felix Wieland

Abstract **A novel member of the human UDP-*N*-acetyl-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase (pp-GalNAc-T) gene family was cloned as a homolog of human pp-GalNAc-T7, and designated pp-GalNAc-T10. pp-GalNAc-T10 transcript was found in the small intestine, stomach, pancreas, ovary, thyroid gland and spleen. In a polypeptide GalNAc-transfer assay, recombinant pp-GalNAc-T10 transferred GalNAc onto a panel of mucin-derived peptide substrates. Furthermore, pp-GalNAc-T10 demonstrated strong transferase activity with glycopeptide substrates.**

Key words: Glycosylation; Polypeptide *N*-acetylgalactosaminyltransferase; *O*-Glycosylation; *O*-Glycan; Mucin

1. Introduction

O-Glycosylation is an important post-translational modification found in many secreted and membrane bound glycoproteins. A family of UDP-GalNAc:polypeptide *N*-acetylgalactosaminytransferases (pp-GalNAc-T, EC2.4.1.41) is responsible for the initial step in the synthesis of mucin-type oligosaccharides by transferring GalNAc from UDP-GalNAc to the hydroxyl group of either a serine or threonine residue on the polypeptide acceptor. To date, ten human pp-GalNAc-T isoforms (T1, T2, T3, T4, T6, T7, T8, T9, T11, and T12) of this gene family have been cloned and characterized [1–9]. pp-GalNAc-Ts have also been identified and cloned in other vertebrates and nematodes [10–17]. Analysis of mammalian pp-GalNAc-Ts has demonstrated that different isoforms have different kinetic properties and unique substrate specificities for substrates such as tandem repeat regions of mucins [3,12,15,18], although there are substantial overlaps, too. Evidence is accumulating that the substrate specificity of pp-GalNAc-T is the major factor determining the sites and density of *O*-glycan attachments on peptide backbones [18]. In addition to having different catalytic properties, each isoform has a different expression profile. The differential expression of pp-GalNAc-Ts in cells and tissues may control the *O*-glycosylation mechanisms. However, the exact molecular processes governing the specificity and the kinetics of *O*-glycosylation remain unclear. The cloning and characterization of novel pp-GalNAc-Ts will help our understanding of the *O*-glycosylation process.

In the present study, we report the cloning and characterization of another member of this polypeptide GalNAc-transferase family, pp-GalNAc-T10. During this study, rat pp-GalNAc-T9, which shares high homology with human pp-GalNAc-T10, was cloned and analyzed [16]. However, unlike rat pp-GalNAc-T9, which exhibited no GalNAc-transferase activity toward non-glycosylated peptides [16], human pp-GalNAc-T10 apparently displayed catalytic activity toward non-glycosylated peptides in addition to strong activity for glycosylated peptides.

2. Materials and methods

2.1. Cloning of pp-GalNAc-T10 cDNA

A BLAST search of expressed sequence tag databases identified five cDNAs (GenBank accession numbers: BG772195, BF359671, BF359677, BF359680, and AK023782) as homologs of human pp-GalNAc-T7 (AJ002744). A full-length open reading frame (ORF) of a novel polypeptide GalNAc transferase was identified by database gene walking. The cDNA encoding the full-length ORF was obtained by polymerase chain reaction (PCR) using the Expand™ High Fidelity PCR system (Roche), the cDNA of a human colon cancer cell line (Colo205) as a template, and the primers 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAGATAGAACCATGAGGCGGAAGGAGAAGCGGCTC-3′ (forward) and 5′-GG GGACCACTTTGTACAAGAAAGCTGGGTCTCAGTTCTATTGAATTTTT-

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

[1] The nucleotide sequence of human pp-GalNAc-T10 reported in this paper has been deposited in the DDBJ/EMBL/GenBank databases with accession number AB078145.

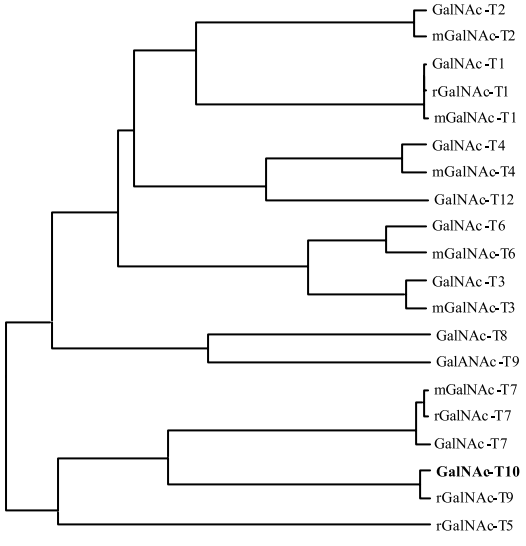
Abbreviations: pp-GalNAc-T, UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase; PCR, polymerase chain reaction; HPLC, high-performance liquid chromatography; FAM, 5-carboxyfluorescein succinimidyl ester; TFA, trifluoroacetic acid; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight

0014-5793/02/$22.00
PII: S0014-5793(02)03399-9

A

```
ATGAGGCGGAAGGAGAAGCGGCTCCTGCAGGCGGTGGCGCTGGTGCTGGCGGCCCTGGTCCTCCTGCCCAACGTGGGGCTTTGGGCGCTGTACCGCGAGCGGCAGCCCGACGGCACCCCT 121
M R R K E K R L L Q A V A L V L A A L V L L P N V G L W A L Y R E R Q P D G T P

GGGGGATCGGGGGCGGCGGTGGCGCCGGCGGCGGGACAGGGCTCACACAGTCGACAAAAGAAAACGTTTTTCTTGGGAGATGGGCAGAAGCTGAAGGACTGGCATGACAAGGAGGCCATC 241
G G S G A A V A P A A G Q G S H S R Q K K T F F L G D G Q K L K D W H D K E A I

CGGAGGGACGCTCAGCGCGTAGGAAATGGAGAACAAGGAAGACCTTACCCCATGACCGATGCTGAGAGAGTGGATCAGGCATACCGAGAAAATGGATTTAACATCTACGTCAGTGATAAA 361
R R D A Q R V G N G E Q G R P Y P M T D A E R V D Q A Y R E N G F N I Y V S D K

ATCTCCTTGAATCGCTCTCTCCCAGATATCCGGCACCCAAACTGCAACAGCAAGCGCTACCTGGAGACACTTCCCAACACAAGCATCATCATCCCCTTCCACAACGAGGGCTGGTCCTCC 481
I S L N R S L P D I R H P N C N S K R Y L E T L P N T S I I I P F H N E G W S S

CTCCTCCGCACCGTCCACAGTGTGCTCAATCGCTCGCCTCCAGAGCTGGTCGCCGAGATTGTACTGGTCGACGACTTCAGTGATCGAGAGCACCTGAAGAAGCCTCTTGAAGACTACATG 601
L L R T V H S V L N R S P P E L V A E I V L V D D F S D R E H L K K P L E D Y M

GCCCTTTTCCCCAGTGTGAGGATTCTTCGAACCAAGAAACGGGAAGGGCTGATAAGGACCCGAATGCTGGGGGCCTCAGTGGCAACTGGGGATGTCATCACATTCTTGGATTCACACTGT 721
A L F P S V R I L R T K K R E G L I R T R M L G A S V A T G D V I T F L D S H C

GAAGCCAATGTCAACTGGCTTCCCCCCTTGCTTGACCGCATTGCTCGGAACCGCAAGACCATTGTGTGCCCGATGATTGATGTAATTGACCATGACGACTTTCGGTACGAGACACAGGCA 841
E A N V N W L P P L L D R I A R N R K T I V C P M I D V I D H D D F R Y E T Q A

GGGGATGCCATGCGGGGAGCCTTTGACTGGGAGATGTACTACAAGCGGATCCCGATCCCTCCAGAACTGCAGAAAGCTGACCCCAGCGACCCATTTGAGTCTCCCGTGATGGCCGGTGGA 961
G D A M R G A F D W E M Y Y K R I P I P P E L Q K A D P S D P F E S P V M A G G

CTGTTCGCCGTGGATCGGAAGTGGTTCTGGGAACTCGGCGGGTATGACCCAGGCTTGGAGATCTGGGGAGGGGAGCAGTATGAAATCTCCTTCAAGGTGTGGATGTGTGGGGGCCGCATG 1081
L F A V D R K W F W E L G G Y D P G L E I W G G E Q Y E I S F K V W M C G G R M

GAGGACATCCCCTGCTCCAGGGTGGGCCATATCTACAGGAAGTATGTGCCCTACAAGGTCCCGGCCGGAGTCAGCCTGGCCCGGAACCTTAAGCGGGTGGCCGAAGTGTGGATGGATGAG 1201
E D I P C S R V G H I Y R K Y V P Y K V P A G V S L A R N L K R V A E V W M D E

TACGCAGAGTACATTTACCAGCGCCGGCCTGAATACCGCCACCTCTCCGCTGGGGATGTCGCAGTCCAGAAAAAGCTCCGCAGCTCCCTTAACTGCAAGAGTTTCAAGTGGTTTATGACG 1321
Y A E Y I Y Q R R P E Y R H L S A G D V A V Q K K L R S S L N C K S F K W F M T

AAGATAGCCTGGGACCTGCCCAAATTCTACCCACCCGTGGAGCCCCCGGCTGCAGCTTGGGGGGAGATCCGAAATGTGGGCACAGGGCTGTGTGCAGACACAAAGCACGGGGCCTTGGGC 1441
K I A W D L P K F Y P P V E P P A A A W G E I R N V G T G L C A D T K H G A L G

TCCCCACTAAGGCTAGAGGGCTGCGTCCGAGGCCGTGGGGAGGCTGCCTGGAACAACATGCAGGTATTCACCTTCACCTGGAGAGAGGACATCCGGCCTGGAGACCCCCAGTACACCAAG 1561
S P L R L E G C V R G R G E A A W N N M Q V F T F T W R E D I R P G D P Q Y T K

AAGTTCTGCTTTGATGCCATTTCCCACACCAGCCCTGTCACGCTGTACGACTGCCACAGCATGAAGGGCAACCAGCTGTGGAAATACCGCAAAGACAAGACCCTGTACCACCCTGTCAGT 1681
K F C F D A I S H T S P V T L Y D C H S M K G N Q L W K Y R K D K T L Y H P V S

GGCAGCTGCATGGACTGCAGTGAAAGTGACCATAGGATCTTCATGAACACCTGCAACCCATCCTCTCTCACCCAGCAATGGCTGTTTGAACACACCAACTCAACAGTCTTGGAAAAATTC 1801
G S C M D C S E S D H R I F M N T C N P S S L T Q Q W L F E H T N S T V L E K F

AATAGGAACTGA 1812
N R N *
```

B



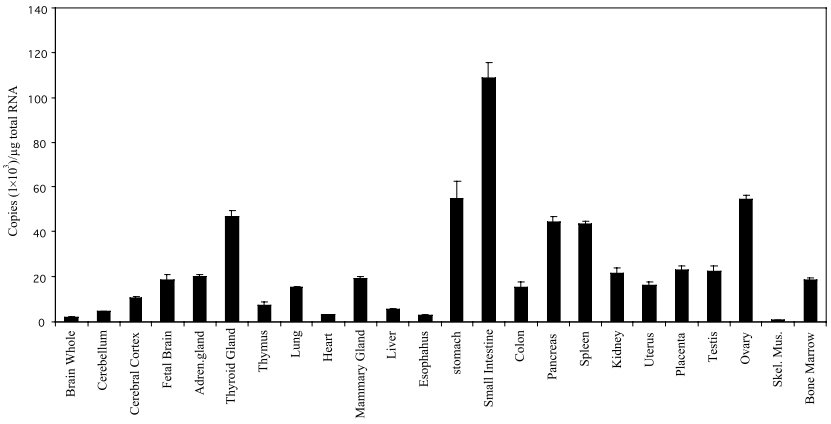


Fig. 2. Quantitative analysis of pp-GalNAc-T10 transcripts in various human tissues by real-time PCR. Expression levels of transcripts are shown as a bar chart. Values are demonstrated as copy numbers of the target gene in 1 μg of total RNA. Data were obtained from triplicate experiments and are indicated as the mean ± S.D.

CCAA-3′ (reverse). These primers have an attB1 or attB2 overhang to create recombination sites. The amplified fragment was subcloned into the vector pDONR201™ using the BP CLONASE Enzyme Mix (Invitrogen). The sequences of subcloned cDNA were determined using the DYEnamic™ DNA sequencing kit (Amersham Pharmacia Biotech).

2.2. Quantitative analysis of the pp-GalNAc-T10 transcripts in human tissues by real-time PCR

Marathon ready® cDNAs of various human tissues were purchased from CLONTECH. The primer set and the probe for pp-GalNAc-T10 were as follows: the forward primer, 5′-CGCGTAGGAAATGGAGAACAA-3′; the reverse primer, 5′-TCGGTATGCCTGATCCACTCT-3′; and the probe, 5′-AGACCTTACCCCATGACCGATGCTGA-3′. Standard curves were generated by the methods described previously [19]. PCR was carried out using the TaqMan Universal PCR Master Mix (Applied Biosystems), primers, probe, and cDNAs. PCR products were continuously measured with the ABI PRISM 7700 Sequence Detection System (Applied Biosystems). The relative amount of pp-GalNAc-T10 transcript was normalized to the amount of human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) transcript in each cDNA.

2.3. Production and purification of recombinant pp-GalNAc-T10 protein with FLAG peptide tag

The cDNA fragment encoding the putative catalytic domain of pp-GalNAc-T10 (amino acids 35–604) was amplified by PCR using the primers 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCAGCCCGACGGCACCCCTGGGGGA-3′ (forward), 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGTTCCTATTGAATTTTTCCAA-3′ (reverse). The amplified fragment was cloned into pDONR201™, and subcloned into the expression vector pFBIF using the LR CLONASE enzyme mix (Invitrogen) [9,19]. Bacmid DNA was isolated from the $DH10_{BAC}$ competent cells (the BAC to BAC Baculovirus Expression Systems, Invitrogen) that were transformed with the pFBIF containing the pp-GalNAc-T10 catalytic domain. The bacmid was transfected into Sf21 cells (Pharmingen, San Diego, CA, USA) to yield conditioned medium containing recombinant pp-GalNAc-T10 protein. The secreted recombinant pp-GalNAc-T10 protein was purified using anti-FLAG M1 antibody-conjugated resin (Sigma). The purified pp-GalNAc-T10 protein was separated by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE), and analyzed by Western blotting using anti-FLAG M2 (Sigma) and the enhanced chemiluminescence-plus (ECL^+) Western blot detection kit (Amersham Biosciences).

The soluble forms of recombinant pp-GalNAc-T2 and -T6, containing amino acid residues 52-571(T2) and 83-622(T6), respectively [1,4], were also generated, purified, and quantified by the same methods as for pp-GalNAc-T10.

2.4. Analysis of pp-GalNAc-T10 activity

Standard assays for pp-GalNAc-T10 activity were performed by the incubation of 20 μl of total reaction mixture containing 25 mM Tris–HCl (pH7.4), 5 mM $MnCl_2$, 0.2% Triton X-100, 0.25 mM UDP-GalNAc (Sigma), 50 pmol acceptor substrate peptide labeled with 5-carboxyfluorescein succinimidyl ester (FAM), and the adequate amount of purified recombinant pp-GalNAc-T10 or -T6. After incubation for various periods at 37°C, the reaction was stopped by boiling samples for 3 min. Then, the reaction mixture was injected onto a C18 reverse column (water $5C_{18}$–AR, 4.6×250 mm) in high-performance liquid chromatography (HPLC; Shimadzu, Kyoto, Japan). The reaction products were eluted with a gradient (0–50%) of acetonitrile containing 0.05% trifluoroacetic acid (TFA) at a flow rate of 1.0 ml/min at 40°C, and monitored with a fluorescence spectrophotometer. FAM-labeled peptide substrates: Muc1a (AHGVTSAPDTR), Muc2(PTTTPITTTTTVTPTPTPTGTQTK), Muc5AC (GTTPSPVPTTSTTSA), Muc7 (PTPSATTPAPPSSSAPPETTAAK), and EA2 (PTTDSTTPAPTTK), which were derived from the tandem repeat domains of corresponding mucins [20–24], were purchased from Sawady (Tokyo, Japan). The mono-glycopeptide substrates were prepared by the incubation of the EA2 or Muc5AC peptide with pp-GalNAc-T2, followed by the purification with HPLC, and were analyzed by mass spectrometry.

2.5. Mass spectrometry

Mass spectrometry was performed by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF). To determine the number of GalNAcs that were transferred to the EA2 peptides by pp-GalNAc-T10, or by -T6, the reaction products were isolated and purified by HPLC. The purified fractions were dried, redissolved in distilled water, mixed with the same amount of matrix solution (α-cyano-4-hydroxycinnamic acid, saturated in 30% acetonitrile containing 0.1% TFA), and subjected to MALDI-TOF mass spectrometry using REFLEX™ IV (Bruker Daltonics, Tsukuba, Japan).

←

Fig. 1. A: Nucleotide and predicted amino acid sequences of pp-GalNAc-T10. Numbering of the cDNA begins with the initiation codon. The predicted amino-acid sequence was presented under its putative nucleotide coding sequence. The triangle represents the starting position of the truncated form. The predicted transmembrane domain is underlined with a single solid line, the GT1 motif is boxed, the Gal/GalNAc-T motif is underlined with a dashed line, and the lectin-like domain is represented by a double line. B: A phylogenetic tree of pp-GalNAc-Ts. A phylogenetic tree of pp-GalNAc-Ts was constructed by the neighbor-joining method [34], based on the amino acid sequences. The branch length indicates the evolutionary distance.

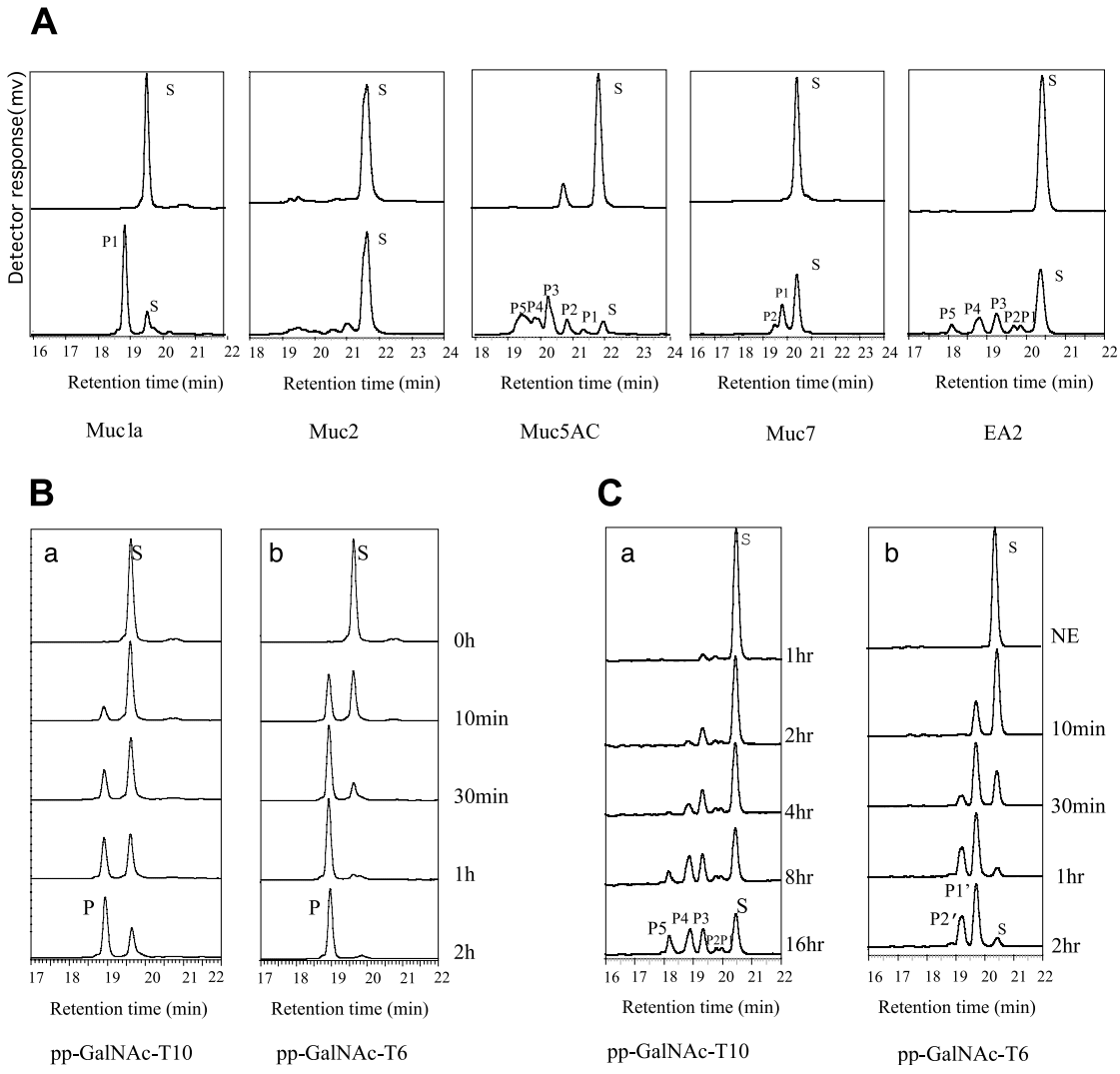


Fig. 3. A: HPLC analysis of in vitro *O*-glycosylation toward a panel of peptide substrates by pp-GalNAc-T10. The elution profiles of the original peptide substrates (*S*) and the reaction products (*P*), which were generated by pp-GalNAc-T10. B: HPLC analysis of the in vitro *O*-glycosylation toward Muc1a by pp-GalNAc-T10 and -T6. Panels a and b show the elution profiles of the Muc1a peptide glycosylated by pp-GalNAc-T10 and -T6, respectively. Peak *S* corresponds to the original substrate. Peak *P* corresponds to the products of the enzyme reaction. The incubation time is indicated on the right side of the panel. C: HPLC analysis of the in vitro *O*-glycosylation toward the EA2 peptide by pp-GalNAc-T10 and -T6. Panels a and b show the elution profiles of the EA2 peptide glycosylated by pp-GalNAc-T10 and -T6, respectively. Peak *S* corresponds to the original substrate, EA2 peptide. Peaks P1–P5 correspond to the products generated by pp-GalNAc-T10. Peaks P1′ and P2′ correspond to the products generated by the pp-GalNAc-T6. The incubation time is indicated on the right sides of the panels. D: HPLC and MALDI-TOF (MS) profiles of EA2 peptide substrate and reaction products generated by pp-GalNAc-T10 and -T6. Panel a shows the elution profiles of the parent EA2 peptide (*S*) and reaction products generated by pp-GalNAc-T10 (P1 to P5) and -T6 (P1′ to P2′). Panel b shows MALDI-TOF mass spectra for each isolated peak. Mono, di, tri, tetra represent one, two, three, four GalNAc-glycopeptide, respectively. E: HPLC analysis of the glycosylation of EA2 peptide and the mono-GalNAc-EA2 glycopeptide generated by pp-GalNAc-T10. *S*″ corresponds to the mono GalNAc-EA2 that was used as a glycopeptide substrate. P1″, P2″ and P3″ correspond to the products generated by pp-GalNAc-T10 from the mono-GalNAc-EA2.

3. Results and discussion

In an effort to further understand the diverse pp-GalNAc-T family, we cloned a novel isoform and designated it pp-GalNAc-T10. As shown in Fig. 1A, the cDNA contains an 1812-bp ORF encoding a type II membrane protein with 603 amino acids. Within the catalytic domain, pp-GalNAc-T10 contains the conserved GT1 and Gal/GalNAc-T motifs, which are putative binding sites for both an UDP-sugar and an acceptor. At the C-terminus, the cloned isozyme contains a lectin-like domain, a potential carbohydrate recognition domain. A lectin-like domain is observed only in pp-GalNAc-Ts among glycosyltransferase families [3,25,26].

In the phylogenetic tree (Fig. 1B), pp-GalNAc-T10 is positioned in the subfamily of five pp-GalNAc-Ts, including rat pp-GalNAc-T9, and human, rat and mouse pp-GalNAc-T7. pp-GalNAc-T10 shares 94% amino-acid homology with rat pp-GalNAc-T9, suggesting that rat pp-GalNAc-T9 is an orthologous isoform of pp-GalNAc-T10.

The expression levels of the pp-GalNAc-T10 transcript in various human tissues were examined using quantitative real-time PCR. pp-GalNAc-T10 was expressed in all tissues tested,

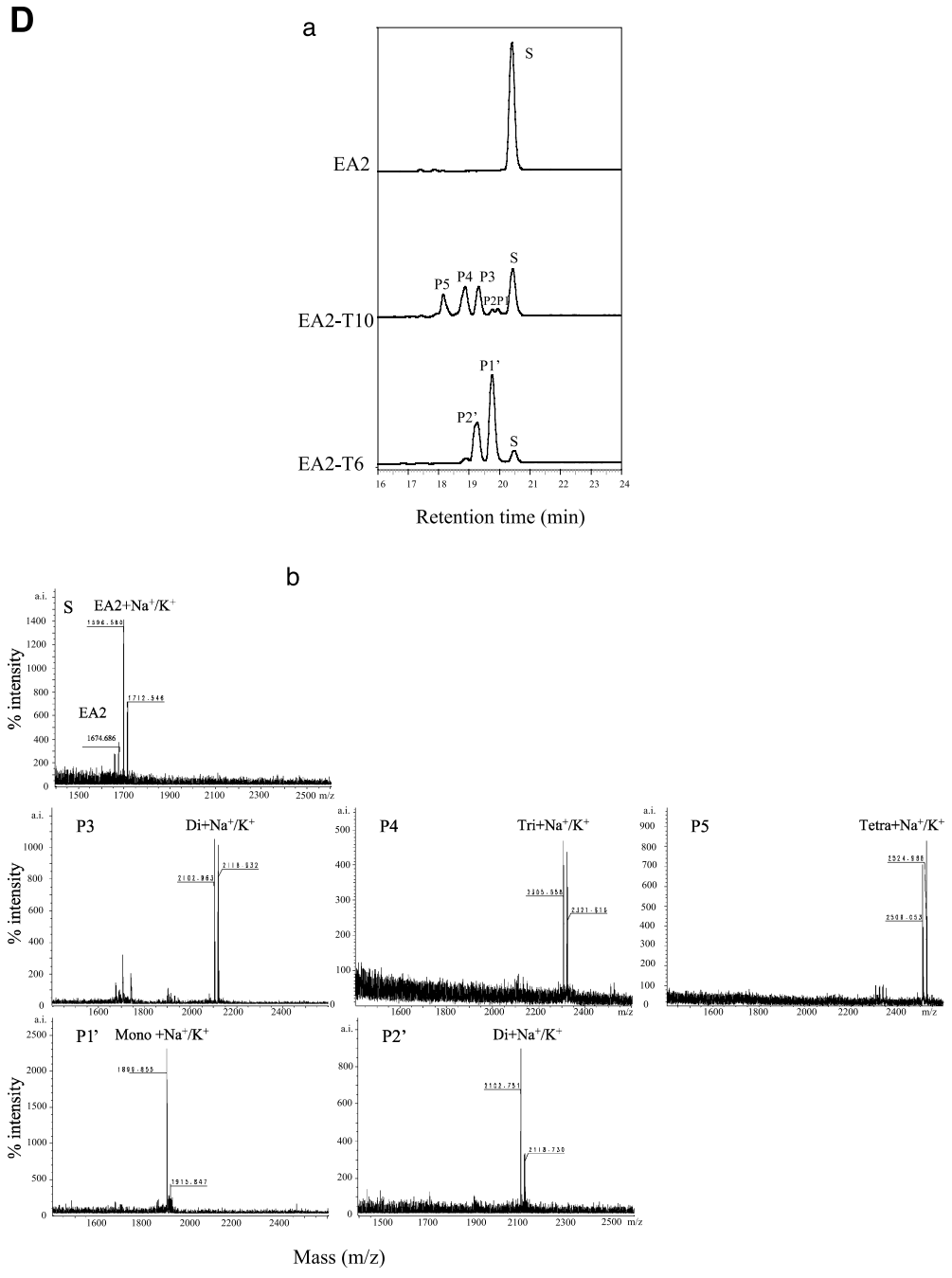


Fig. 3 (*Continued*).

although the expression levels were different depending on the tissues (Fig. 2). The highest level of transcript was found in the small intestine, while intermediate levels were seen in the stomach, pancreas, ovary, thyroid gland and spleen. Lower levels to trace amounts were observed in other tissues examined. Human pp-GalNAc-T isoforms, i.e. pp-GalNAc-T1, -T2, -T4, -T7 and -T8, were also ubiquitously expressed [1,3,5,6]. The reasons why multiple pp-GalNAc-Ts are expressed within a single tissue are still unclear. Because pp-GalNAc-Ts have specific target substrates [18,27,28], there may be competition and coordination among pp-GalNAc-Ts for the appropriate glycosylation of acceptor substrates in a cell. Thus, the expression of multiple pp-GalNAc-Ts within a tissue may be necessary for the serial glycosylation of substrate proteins by pp-GalNAc-Ts, or may provide redundancy to ensure protein glycosylation.

The expression level of the pp-GalNAc-T10 transcript was low in most tumor cell lines examined (data not shown), and we could not find a specific expression pattern among cell lines.

To analyze pp-GalNAc-T10 activity in vitro, the recombinant pp-GalNAc-T10 fused with the FLAG-epitope peptide was purified from the conditioned medium of the transfected Sf21 cells by affinity chromatograph with anti-FLAG M1-conjugated resin. The purified protein was quantified by Western blotting with anti-FLAG mAb, and used for enzyme as-

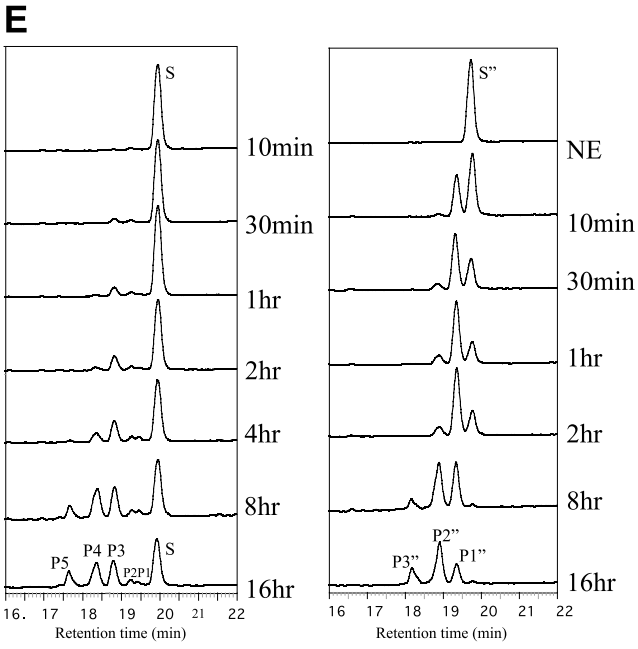


Fig. 3 (*Continued*).

says. The purified recombinant pp-GalNAc-T10, UDP-GalNAc, and a panel of acceptor substrates were incubated for various time periods, and the reaction products were analyzed by HPLC.

First, we analyzed the catalytic activities of pp-GalNAc-T10 toward peptide substrates derived from Muc1, Muc2, Muc5AC, Muc7, and EA2. As demonstrated in Fig. 3A, pp-GalNAc-T10 generated product peaks (P) that were shifted from substrate peaks (S) with the Muc1a, Muc5AC, Muc7, and EA2 peptide, but negligible activity toward Muc2. These data indicate that pp-GalNAc-T10 has significant pp-GalNAc-T activity toward the mucin-derived peptides.

To assess the strength of the catalytic activity of pp-GalNAc-T10, we incubated the Muc1a peptide with pp-GalNAc-T10 or -T6 for varying periods and analyzed the reaction mixture by HPLC. As shown in Fig. 3B, both pp-GalNAc-T10 and -T6 effectively utilized the Muc1a peptide (shown as 'S') and yielded one product peak, 'P', even after extended incubation. The amount of the substrate remaining after 1 h incubation with pp-GalNAc-T10 was almost the same as that of 10 min incubation with pp-GalNAc-T6. Thus, in this experiment, the activity of pp-GalNAc-T10 toward the Muc1a peptide was one-sixth of that of pp-GalNAc-T6, which has strong GalNAc-transfer activity with the Muc1a peptide [4].

The EA2 peptide, which contains seven potential glycosylation sites, was derived from the tandem repeats of the rat submandibular gland mucin. The incubation of the EA2 peptide with pp-GalNAc-T10 generated multiple product peaks with a unique pattern on HPLC profile (panel a in Fig. 3C). The product peaks P1 and P2, which were the closest to the original substrate peak, were very small in any incubation period, whereas peaks P3, P4 and P5, reflecting more heavily glycosylated peptides, gradually appeared but became higher than P1 and P2. In contrast, a product peak, P1′, that was generated by pp-GalNAc-T6, was the only peak in the initial reaction period, and then P1′ was gradually consumed and the subsequent product peak, P2′, appeared as incubation extended (panel b in Fig. 3C). The retention time of P2 was the same as that of P1′, while the retention time of P3 was the same as that of P2′ (also demonstrated in panel a in Fig. 3D).

Because these results suggested that pp-GalNAc-T10 transferred multiple GalNAcs to the EA2 peptide, we wanted to determine the numbers of GalNAcs transferred to the EA2 peptide by pp-GalNAc-T10 or -T6. For this purpose, we purified product peaks by HPLC and subjected the purified fractions to the mass spectrometry. As shown in panel b of Fig. 3D, the determined molecular weights of the products corresponding to peaks P3, P4, P5 (generated by T10), P1′, and P2′ (generated by T6) matched the expected molecular weights of the FAM-labeled EA2 with di-, tri-, tetra-, and mono-, di-GalNAc residues, respectively. Unfortunately, we could not fractionate sufficient amounts of peaks P1 and P2 for mass spectrometry. However, because the retention times of P2 and P1′ were the same, the results strongly suggested that both P1 and P2 peaks represented mono-GalNAc-attached EA2 peptides, presumably at different sites. Taken together, our results indicated that pp-GalNAc-T10 first transfers one GalNAc to a Thr or Ser site in the EA2 peptide in a less efficient fashion, and then transfers another GalNAc immediately to the other sites, suggesting that GalNAc-T10 had strong pp-GalNAc-T activity toward the glycosylated EA2 and relatively weak pp-GalNAc-T activity toward the EA2 peptide.

To further analyze the substrate specificity of pp-GalNAc-T10, we prepared the mono-GalNAc EA2 glycopeptide and used it as an acceptor substrate. As shown in Fig. 3E, a product peak P1″ appeared soon after a short incubation with GalNAc-T10, while the substrate peak S″ was rapidly consumed. Then, P1″ was gradually consumed and the other product peaks, P2″ and P3″, appeared as incubation extended. The amount of the EA2 substrate remaining after the GalNAc-T10 enzyme reaction for 8 h was almost the same as that of the GalNAc-EA2 substrate remaining after the enzyme

reaction for only 10 min, indicating that pp-GalNAc-T10 showed more preferential activity for the mono-GalNAc-EA2 glycopeptide than for the EA2 peptide. Several studies have demonstrated relative effects, negative as well as positive, of adjacent and distant *O*-glycans on the kinetics of subsequent in vitro *O*-glycosylation [16,29,30]. The addition of a GalNAc to the EA2 peptide may induce conformational changes of the peptide that influence the accessibility of particular acceptor sites for pp-GalNAc-T10. Previous publications have identified the effects of *O*-glycosylation on peptide backbone conformation near the glycosylation sites as well as at distant sites [31–33]. Another possible mechanism is that the lectin domain of pp-GalNAc-T10 binds to the peptide-attached GalNAc and contributes to the glycosylation of the GalNAc-EA2 glycopeptide by pp-GalNAc-T10, which is a similar mechanism found in human pp-GalNAc-T4 [25].

pp-GalNAc-T10 exhibited strong GalNAc-transferase activity toward the mono-GalNAc-Muc5AC glycopeptide, too (data not shown).

pp-GalNAc-T10 shares 94% amino acid sequence homology with rat pp-GalNAc-T9, suggesting human pp-GalNAc-T10 is an orthologous isoform of rat pp-GalNAc-T9. However, human pp-GalNAc-T10 and rat pp-GalNAc-T9 are different in substrate specificities. It was reported that rat pp-GalNAc-T9 did not transfer a GalNAc to unmodified peptide substrates, including Muc1, EA2, Muc2, and Muc5AC[16], but human pp-GalNAc-T10 showed clear catalytic activity toward a panel of unmodified polypeptide substrates. Both pp-GalNAc-T10 and rat pp-GalNAc-T9 demonstrated strong pp-GalNAc-T activity toward EA2 and Muc5AC glycopeptides.

In conclusion, the present study provides evidence that human pp-GalNAc-T10 is a new member of the pp-GalNAc-T family. The broad expression and unique acceptor substrate specificity suggest that pp-GalNAc-T10 may play important roles in *O*-glycosylation.

References

[1] White, T., Bennett, E.P., Takio, K., Sorensen, T., Bonding, N. and Clausen, H. (1995) J. Biol. Chem. 270, 24156–24165.
[2] Bennett, E.P., Hassan, H. and Clausen, H. (1996) J. Biol. Chem. 271, 17006–17012.
[3] 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.
[4] 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.
[5] Bennett, E.P., Hassan, H., Hollingsworth, M.A. and Clausen, H. (1999) FEBS Lett. 460, 226–230.
[6] White, K.E., Lorenz, B., Evans, W.E., Meitinger, T., Strom, T.M. and Econs, M.J. (2000) Gene 246, 347–356.
[7] Toba, S., Tenno, M., Konishi, M., Mikami, T., Itoh, N. and Kurosaka, A. (2000) Biochim. Biophys. Acta 1493, 264–268.
[8] 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.
[9] Guo, J.M., Zhang, Y., Cheng, L., Iwasaki, H., Wang, H., Kubota, T., Tachibana, K. and Narimatsu, H. (2002) FEBS Lett. 524, 211–218.
[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., Ten Hagen, K.G., Beres, T.M., Balys, M.M., Van Wuyckhuyse, B.C. and Tabak, L.A. (1997) J. Biol. Chem. 272, 13843–13848.
[13] Ten Hagen, K.G., Tetaert, D., Hagen, F.K., Richet, C., Beres, T.M., Gagnon, J., Balys, M.M., Van Wuyckhuyse, B., Bedi, G.S., Degand, P. and Tabak, L.A. (1999) J. Biol. Chem. 274, 27867–27874.
[14] Hagen, F.K., Gregoire, C.A. and Tabak, L.A. (1995) Glycoconjugate J. 12, 901–909.
[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., 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.
[17] Hagen, F.K. and Nehrke, K. (1998) J. Biol. Chem. 273, 8268–8277.
[18] 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.
[19] 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.
[20] Gendler, S.J., Lancaster, C.A., Taylor-Papadimitriou, J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E.N. and Wilson, D. (1990) J. Biol. Chem. 265, 15286–15293.
[21] 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.
[22] 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), 211–9.
[23] Bobek, L.A., Tsai, H., Biesbrock, A.R. and Levine, M.J. (1993) J. Biol. Chem. 268, 20563–20569.
[24] Albone, E.F., Hagen, F.K., Van Wuyckhuyse, B.C. and Tabak, L.A. (1994) J. Biol. Chem. 269, 16845–16852.
[25] Hassan, H., Reis, C.A., Bennett, E.P., Mirgorodskaya, E., Roepstorff, P., Hollingsworth, M.A., Burchell, J., Taylor-Papadimitriou, J. and Clausen, H. (2000) J. Biol. Chem. 275, 38197–38205.
[26] Hazes, B. (1996) Protein Sci. 5, 1490–1501.
[27] Clausen, H. and Bennett, E.P. (1996) Glycobiology 6, 635–646.
[28] Sorensen, T., White, T., Wandall, H.H., Kristensen, A.K., Roepstorff, P. and Clausen, H. (1995) J. Biol. Chem. 270, 24166–24173.
[29] 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.
[30] Brockhausen, I., Toki, D., Brockhausen, J., Peters, S., Bielfeldt, T., Kleen, A., Paulsen, H., Meldal, M., Hagen, F. and Tabak, L.A. (1996) Glycoconjugate J. 13, 849–856.
[31] Kirnarsky, L., Prakash, O., Vogen, S.M., Nomoto, M., Hollingsworth, M.A. and Sherman, S. (2000) Biochemistry 39, 12076–12082.
[32] Huang, X., Barchi Jr., J.J., Lung, F.D., Roller, P.P., Nara, P.L., Muschik, J. and Garrity, R.R. (1997) Biochemistry 36, 10846–10856.
[33] Huang, X., Smith, M.C., Berzofsky, J.A. and Barchi Jr., J.J. (1996) FEBS Lett. 393, 280–286.
[34] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.