

Cloning and Expression of Mouse UDP-GalNAc:Polypeptide N-Acetylgalactosaminyltransferase-T3

Jane Zara, Fred K. Hagen, Kelly G. Ten Hagen,
Brian C. Van Wuyckhuyse, and Lawrence A. Tabak¹

*Departments of Dental Research and Biochemistry, School of Medicine and Dentistry,
University of Rochester, 601 Elmwood Avenue, Rochester, New York 14642*

Received September 24, 1996

A novel isoform of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase, designated ppGaNTase-T3, has been cloned from a mouse testis cDNA library and expressed in COS7 cells. ppGaNTase-T3 displayed 64 and 59% amino acid identity with ppGaNTase-T1 and ppGaNTase-T2, respectively, and 96% amino acid identity with the recently reported human form of ppGaNTase-T3. The ppGaNTase-T3 transcript is abundant in the major salivary glands, gastrointestinal tract and both the male and female reproductive systems. ppGaNTase-T3 and ppGaNTase-T1 display overlapping substrate preferences *in vitro*, although mapping studies of O-glycosylated peptides suggests that certain hydroxyamino acids are preferentially glycosylated by each isoform. This suggests that more than one isoform of ppGaNTase may be required to complete the O-glycosylation of endogenous substrates. © 1996 Academic Press, Inc.

UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase (EC 2.4.1.41) is the enzyme which catalyzes the reaction $\text{UDP-GalNAc} + \text{polypeptide-(Ser/Thr)-OH} \rightarrow \text{GalNAc } \alpha\text{-O-Ser/Thr-polypeptide} + \text{UDP}$, thereby initiating O-glycosylation of serine and threonine residues on an array of glycoproteins (1, 2). Recent studies have demonstrated that there are multiple ppGaNTases and that these isoforms vary in both their pattern of expression and *in vitro* substrate specificity (1-5).

A human expressed sequence tag (hEST, T11328) was reported which displayed regions of sequence similarity which are common among divergent species of the ppGaNTase, suggesting that the hEST encoded a novel isoform of ppGalNAc (3). While this work was in progress, a human salivary gland cDNA corresponding to this hEST was isolated, expressed and shown to encode a novel form of the transferase, termed ppGaNTase-T3 (5). In this report, we have cloned, sequenced and expressed the mouse cDNA corresponding to this hEST, and demonstrate that recombinant mouse ppGaNTase-T3 preferentially glycosylates certain hydroxyamino acids in multi-site substrates, *in vitro*.

MATERIALS AND METHODS

Northern blot analysis. Following electrophoresis, mouse RNA samples (6) were transferred to Hybond-N membranes (Amersham). A probe consisting of nt1-822 of hEST-T11328 (3) was used to detect transcript. The anti-sense strand of this fragment was labeled by asymmetric PCR using the 3' antisense oligonucleotide 5'ACGAGACCTTGA-GCAGCAT3', (Universal DNA, Inc.) as described previously (7). Antisense 18S ribosomal subunit oligonucleotide 18S anti-5'TATTGGAGCTGGAATTACCGCGGCTGCTGG3' was end-labeled as described (6) and used to normalize loading of samples. All hybridizations were performed in 5× SSPE/50% formamide at 42°C with two final washes in 2× SSC/0.1% SDS at 65°C for 20 min.

Cloning and sequencing the mouse homologue of hEST T11328. A mouse testis cDNA library (Unizap XR;

¹ To whom correspondence should be addressed.

Abbreviations: ppGaNTase: UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase; UDP-GalNAc: uridine diphospho-N-acetyl-D-galactosamine (EC 2.4.1.41); PTH: phenylthiohydantoin.

TABLE 1
Summary of Substrates Glycosylated in Vitro with ppGaNTase-T3 and -T1

Peptide Name	Sequence	Derived from	Reference	T3	T1
				fmoles	GalNAc/ μ g/min ^a
EA2	PTTDSTTPAPTTK	Rat submandibular mucin	(9)	172 \pm 19	3598 \pm 97
Muc 1a	APPAHGVTSA PDTRPAPGC	Muc 1 mucin type glycoprotein	(7)	113 \pm 5	1670 \pm 529
Muc 1b	PDTRPAPGSTAPPAC	Muc 1 mucin type glycoprotein	(7)	29 \pm 1	125 \pm 3
Muc 2	PTTTPISTTTMTPTPTPTC	Human intestinal mucin	(8)	310 \pm 9	2969 \pm 187
AWN1a	AIPPLNLSCGKE	Porcine spermadhesin AWN-1	(10)	—	—
MCP1	LPPSSTKPPALSHS	Membrane cofactor protein	(11)	59 \pm 3	339 \pm 34
MCP2	STSSSTTKSPASSAS	Membrane cofactor protein	(11)	157 \pm 14	1061 \pm 142
MCP3	GRPTYKPPVSNYP	Membrane cofactor protein	(11)	25 \pm 2	257 \pm 19

^a Transfections for each construct were performed simultaneously and enzyme assays performed in triplicate.

Stratagene) was screened (1×10^6 plaques) using the manufacturer's protocol and the hEST probe described above. Bi-directional sequencing was performed by a combination of fluorescence (Ladderman enzyme, Pan Vera; with a Model 4000L LiCor Automated DNA Sequencer) and radioactive methods (Taq-Trac; Perkin Elmer).

Subcloning the mouse homologue of hEST T32595. An amino truncated form of the putative glycosyltransferase, lacking the first 49 amino acids, was generated by PCR using primers 5' CACACGCGTCAAAGATGGAGAGGAAGTGA AAA and 5' CACGGATCCGTATTCTAGTTGCTGTGCTTTC to obtain a secreted form of the enzyme from COS7 cells. Amplification was performed at 95°C, 5 min; 53°C, 1.5 min; and 72°C, 3 min for 20 cycles, and the product inserted into pSVL (Pharmacia) containing an insulin secretion signal adjacent to the amino terminus of truncated mouse cDNA by utilizing engineered BamHI and MluI sites (2).

Cell culture and transient transfections. COS7 cells were maintained in DMEM (Gibco/BRL) 10% fetal calf serum in 5% CO₂ at 37°C. Cells were transiently transfected using DEAE dextran (6). Twelve to 18 h after, transfection cells were grown at 30°C and harvested 3d post-transfection. Secreted recombinant protein was harvested from cell supernatants (2). Sham transfections were performed with pSVL lacking transferase insert.

In vitro glycosylation assays. ppGaNTase assays were performed as described previously (2) using COS7 cell culture supernatant as the source of recombinant enzyme. For comparison, recombinant mouse ppGaNTase-T1 (F.K. Hagen and L.A. Tabak, unpublished) was expressed and assayed in an identical fashion.

A panel of synthetic peptide substrates, synthesized by Fmoc chemistry using a Milligene 9050 synthesizer, were used to assay for ppGaNTase activity (Table 1). The integrity of each peptide was confirmed by Edman degradation and mass spectroscopy; peptide concentrations were verified by amino acid analysis of acid hydrolyzed sample. Substrate glycosylation was compared by incubating 400 μ M of each peptide substrate (Table I), 50 μ M ¹⁴C-UDP-GalNAc (16.4 mCi/mmol). Reactions were stopped by addition of 30 mM EDTA. Glycosylated products were resolved from unincorporated ¹⁴C-UDP-GalNAc with anion-exchange (BioRad AG 1X-8) spin columns after verifying results by direct measurement of peptide isolated by reverse phase HPLC (12). Protein concentrations of culture supernatants were determined by the method of Lowry (13).

Peptide mapping. Recombinant enzyme (8 μ l of transfected culture supernatant [2.7 or 4.4 mg/ml for ppGaNTase-T1 and T3, respectively]) was added to 2 mM UDP-GalNAc and 200 μ M of either EA2 or MCP-2 (in 100 μ l). Following incubation (6-8h at 37°C), an additional 4 μ l of recombinant enzyme was added and then incubated (8-10h at 37°C). Resultant products were covalently coupled to polyvinylidene difluoride membranes (Sequelon; Millipore). Edman degradation was performed with modified cycles for trifluoroacetic acid cleavage and extractions. Glycosylated effluents from each cleavage cycle were counted for ¹⁴C-GalNAc incorporation and nonglycosylated residues were converted to PTH amino acids and quantitated as described previously (12).

RESULTS AND DISCUSSION

Given the abundance of ppGaNTase transcript in testis as determined by a preliminary northern blot analysis, we screened a mouse testis cDNA library. Of the eighteen positive clones isolated, one contained the entire open reading frame of ppGaNTase-T3, flanked by 222 nt of 5'UTR and 754 nucleotides of 3'UTR. The aligned nt sequence of mouse and human ppGaNTase-T3 and the deduced amino acid sequences are shown in Fig. 1. The sequence

40

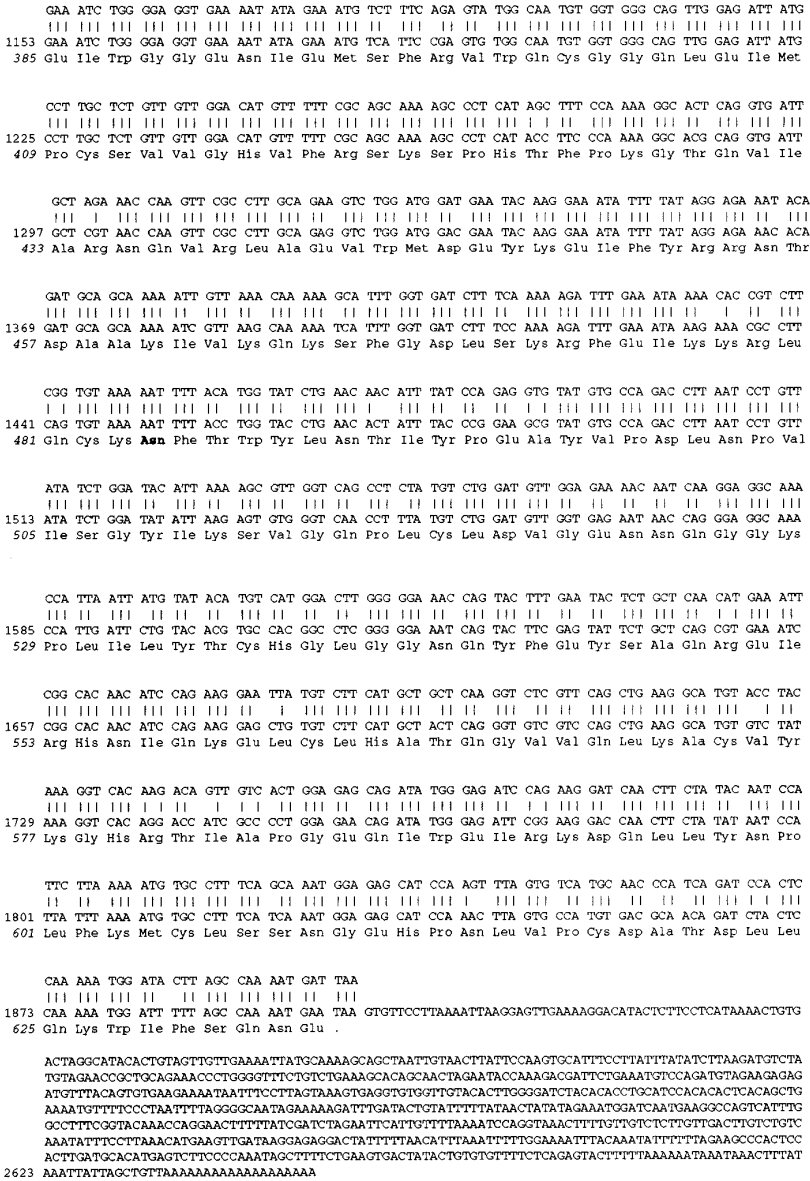


FIG. 1. Nucleotide and deduced amino acid sequences of mouse ppGaNTase-T3. The nucleotide sequence of mouse T3 is shown with corresponding nucleotide numbers on the left. The human T3 nucleotide sequence recently reported by Bennett et al (5) is shown above the mouse T3 sequence and homologous nucleotide sequences between human and mouse T3 are depicted by vertical lines. The deduced amino acid sequence of mouse T3 is also shown (italicized numbers on the left), where the potential N-glycosylation residues are denoted in bold and the purported membrane spanning region is underlined.

similarity of mouse ppGaNTase-T3 with other known isoforms of the transferase is shown in Table 2.

The message begins with a 222 nt 5' UTR that contains 3 termination codons, which are in frame with the AUG codon at position 223 (+1 in Fig. 1). A putative translation initiation site at position 223 agrees with the Kozak consensus sequence for eukaryotic translation

TABLE 2
Comparison of Mouse ppGaNTase-T3 with Other ppGaNTases

Isoform	Nucleotide Identity, %	AA Similarity, %
Human T1 ^a	51	64
Human T2	49	59
Human T3	87	96
Mouse T1 ^b	62	65
<i>C. elegans</i> (ZK688) ^a	40	47

^a The GeneBank accession numbers for human T1, human T2, human T3, and *C. elegans* are X85018, X85019, X92689, and L16621, respectively.

^b Mouse T1 sequence, F. K. Hagen and L. A. Tabak, unpublished.

initiation in 6 out of 9 nt (14). The deduced amino acid sequence of mouse ppGaNTase-T3 reveals a type II membrane protein of 633 aa residues, with a predicted membrane spanning region of 15 aa residues and a cytoplasmic tail of 22 aa residues.

A single transcript of 3.2 Kb encoding ppGaNTase-T3 was detected by northern blot analysis of RNAs derived from a number of mouse tissues (Fig. 2A). The transcript was highly prevalent in the reproductive tract, principally in the testis and uterus, and to a lesser degree in the cervix with only trace levels detected in the ovary. ppGaNTase-T3 message was also highly abundant in sublingual gland, stomach and colon, with more moderate amounts present in the submandibular and parotid gland as well as the kidney.

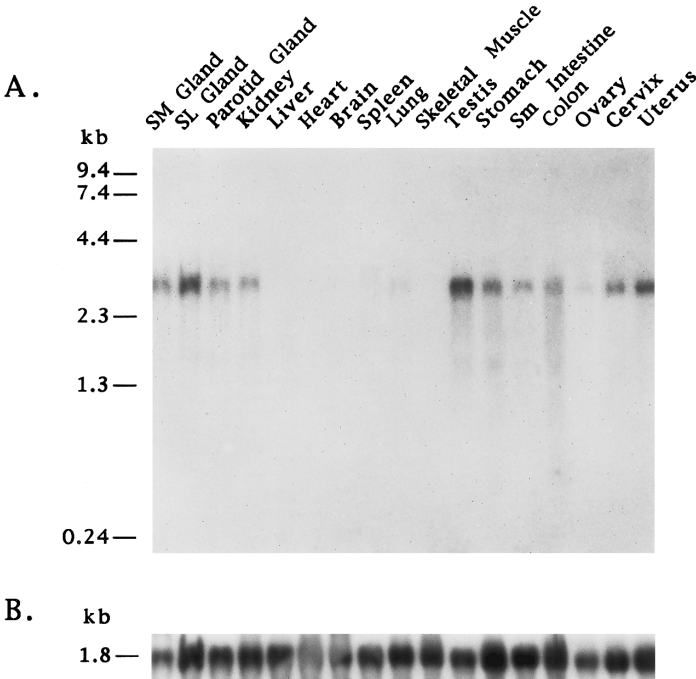


FIG. 2. Northern analysis of T3 expression in mouse tissues. Total RNA (from Balb/C mice) was extracted from the glands and organs listed above panel A. After electrophoresis on 1% formaldehyde-agarose gel and transfer to Hybond-N membranes, RNA was hybridized with a T3 specific probe (Pane A) and an 18s anti-rRNA probe (Panel B) as a control for RNA integrity. Each lane contains 15 μ g of total RNA. Size markers are indicated on the left.

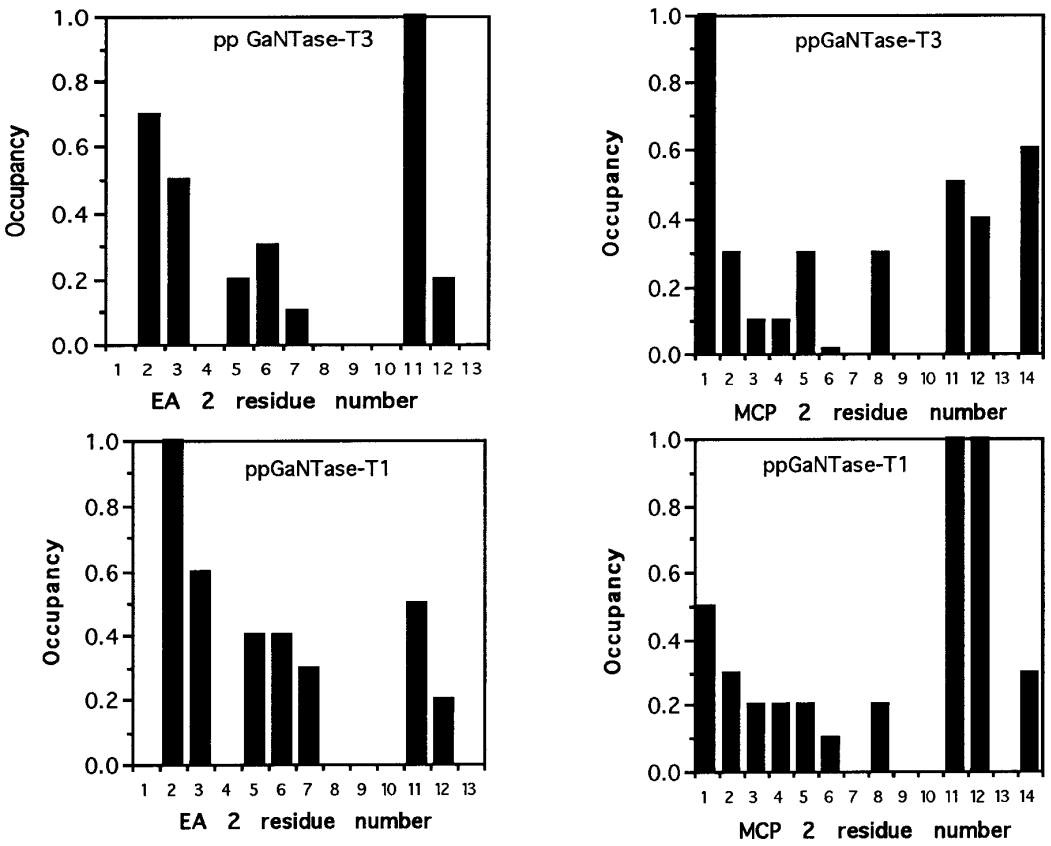


FIG. 3. Occupancy of O-glycosylated residues on peptides EA2 and MCP2. The left-hand panels show differences in occupancy of residues on the substrate EA2 generated by ppGaNTase-T3 (upper left panel) and ppGaNTase-T1 (lower left panel). The right-hand panels illustrate differences in occupancy of residues on the peptide MCP2 generated by ppGaNTase-T3 (upper right panel) and ppGaNTase-T1 (lower right panel). The most highly glycosylated residue in each column was set to 1.0. Actual occupancies of residues were as follows: For the T1 isoform, Thr2 of EA2 had an occupancy of 30% and Ser12 of MCP2 had an occupancy of 39%. For the T3 isoform, Thr11 of EA2 had an occupancy of 10% and Ser1 of MCP2 had an occupancy of 47%.

mouse tissues, we find that ppGaNTase-T3 message is more widely expressed than suggested by the work of Bennett et al. (5). However, this result may be due to species-specific differences in expression between mouse and human.

ppGaNTase activity was detected in the culture supernatants of COS7 cells transfected with the mouse homologue of hEST T11328, verifying that this cDNA encodes a ppGaNTase. Peptide sequences derived from Muc 1 (Muc 1a and 1b), human intestinal (Muc 2) and rat salivary (EA 2) mucins were each glycosylated by recombinant ppGaNTase-T1 and ppGaNTase-T3 in vitro (Table 1), suggesting that one or both of these forms could play a role in the biosynthesis of these mucins. Similarly, several peptides based on membrane cofactor protein (MCP), a protein thought to play a role in reproductive functions in both the male and female reproductive tracts (11), were also glycosylated by both isoforms of transferase in vitro. In contrast, the peptide derived from testis associated spermadhesin protein (AWN) (10) was not glycosylated by either isoform of ppGaNTase (Table 1); perhaps this domain must be presented within a larger context of the spermadhesin molecule, or, alternatively, another isoform of ppGaNTase may be required to modify these sites.

The percent of each hydroxyamino acid occupied by GalNAc in O-glycosylated EA 2 and MCP 2 is summarized in Fig. 3. There are clear differences when each substrate is glycosylated with the two isoforms in vitro. We speculate that these differences suggest that multiple forms of ppGaNTase are required to optimally O-glycosylate multi-site substrates in vivo, as has been suggested for O-mannosylation in *S. cerevisiae* (15).

ACKNOWLEDGMENTS

This work was supported in part by NIH grant DE-08108. J.Z. was supported by T32 DE07202. The nucleotide sequences reported in this paper have been submitted to the GenBank with accession number U70538. We thank Pat Noonan for her help in preparing this manuscript.

REFERENCES

1. Homa, F. L., Hollander, T., Lehman, D. J., Thomsen, D. R., and Elhammer, Å. P. (1993) *J. Biol. Chem.* **268**, 12609–12616.
2. Hagen, F. K., VanWuyckhuysse, B. C., and Tabak, L. A. (1993) *J. Biol. Chem.* **268**, 18960–18965.
3. Hagen, F. K., Gregoire, C. A., and Tabak, L. A. (1995) *Glycoconj. J.* **12**, 901–909.
4. Sørensen, T., White, T., Wandall, H. H., Kristensen, A. K., Roepstorff, P., and Clausen, H. (1995) *J. Biol. Chem.* **270**, 24166–24173.
5. Bennett, E. P., Hassan, H., and Clausen, H. (1996) *J. Biol. Chem.* **271**, 17006–17012.
6. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
7. 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.
8. Gum, J. R., Jr., Hicks, J. W., Toribara, N. W., Siddiki, B., and Kim, Y. S. (1994) *J. Biol. Chem.* **269**, 2440–2446.
9. Albone, E. F., Hagen, F. K., VanWuyckhuysse, B. C., and Tabak, L. A. (1994) *J. Biol. Chem.* **269**, 16845–16852.
10. Sanz, L., Calvete, J. J., Mann, K., Schafer, W., Schmid, E. R., Amselgruber, W., Sinowatz, F., Ehrhard, M., and Topfer-Petersen, E. (1992) *FEBS Lett.* **300**, 213–218.
11. Liszewski, M. K., Post, T. W., and Atkinson, J. P. (1991) *Annu. Rev. Immunol.* **9**, 431–455.
12. O'Connell, B. C., Hagen, F. K., and Tabak, L. A. (1992) *J. Biol. Chem.* **267**, 25010–25018.
13. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275.
14. Kozak, M. (1984) *Nuc. Acids. Res.* **12**, 857–873.
15. Gentzsch, M., Immervoll, T., and Tanner, W., (1995) *FEBS Lett.* **377**, 128–130.