Cloning and Characterization of a Ninth Member of the UDP-GalNAc:Polypeptide N-Acetylgalactosaminylytransferase Family, ppGaNTase-T9*

Received for publication, October 23, 2000, and in revised form, February 1, 2001

Published, JBC Papers in Press, February 6, 2001, DOI 10.1074/jbc.M009638200

Kelly G. Ten Hagen‡, Gurrrinder S. Bedi‡, Daniel Tetaert§, Paul D. Kingsley‡, Fred K. Hagen‡, Marlene M. Balys‡, Thomas M. Beres‡, Pierre Degand§, and Lawrence A. Tabak‡¶

From the ¶Center for Oral Biology, Aab Institute for Biomedical Sciences, University of Rochester, Rochester, New York 14642 and §INSERM Unite 377, Biologie et Physiopathologie de Cellules Mucipares, Place de Verdun, 59045 Lille Cedex, France

We have cloned, expressed and characterized a gene encoding a ninth member of the mammalian UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase (ppGaNTase) family, termed ppGaNTase-T9. This type II membrane protein consists of a 9-amino acid N-terminal cytoplasmic region, a 20-amino acid hydrophobic/transmembrane region, a 94-amino acid stem region, and a 480-amino acid conserved region. Northern blot analysis revealed that the gene encoding this enzyme is expressed in a broadly distributed manner across many adult tissues. Significant levels of 5- and 4.2-kilobase transcripts were found in rat sublingual gland, testis, small intestine, colon, and ovary, with lesser amounts in heart, brain, spleen, lung, stomach, cervix, and uterus. In situ hybridization to mouse embryos (embryonic day 14.5) revealed significant hybridization in the developing mandible, maxilla, intestine, and mesonephric ventricle. Constructs expressing this gene transiently in COS7 cells resulted in no detectable transferase activity in vitro against a panel of unmodified peptides, including MUC5AC (GTTPSPVPTTSTTFAP) and EA2 (PTTDSTTPAAPT). However, when incubated with MUCSAC and EA2 glycopeptides (obtained by the prior action of ppGaNTase-T1), additional incorporation of GalNAc was achieved, resulting in new hydroxamino acid modification. The activity of this glycopeptide transferase is distinguished from that of ppGaNTase-T7 in that it forms a tetra-glycopeptide species from the MUCSAC tri-glycopeptide substrate, whereas ppGaNTase-T7 forms a hexa-glycopeptide species. This isoform thus represents the second example of a glycopeptide transferase and is distinct from the previously identified form in enzymatic activity as well as expression in embryonic and adult tissues. These findings lend further support to the existence of a hierarchical network of differential enzymatic activity within the diversely regulated ppGaNTase family, which may play a role in the various processes governing development.

Mucin type O-linked glycosylation is initiated by the action of a family of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases (ppGaNTase, EC 2.4.1.41), which catalyze the transfer of GalNAc from the nucleotide sugar UDP-GalNAc to the hydroxyl group of either serine or threonine. A number of functional roles for O-glycans have been suggested (reviewed in Ref. 1), including protection from proteolytic degradation (2), alteration of substrate structural conformation (3), aiding in sperm-egg binding during fertilization in mice (4), and coordination of leukocyte rolling along endothelial cells upon inflammation and injury (5). However, the exact biological functions of O-linked glycosylation remain largely unknown, as studies involving chemical/ enzymatic cleavage of sugars and/or maturation of acceptor residues on proteins can result in second ary effects unrelated to sugar removal or absence. Since carbohydrates can only be “mutated” indirectly by modifying the enzymatic activities of the glycosyltransferases responsible for their synthesis, our efforts have focused on the characterization of the enzyme family responsible for the initiation of O-glycan addition.

Thus far, seven distinct mammalian isoforms from this gene family have been identified and functionally characterized: ppGaNTase-T1 (6, 7), -T2 (8), -T3 (9, 10), -T4 (11), -T5 (12), -T6 (13), and -T7 (14, 15). An eighth putative isoform was ablated in mice without any obvious phenotypic effects (16, 17); however, the enzymatic activity and the gene encoding this isoform remain uncharacterized. Whereas some isoforms display a broad range of expression in adult tissues and act on a robust set of substrates (ppGaNTase-T1, -T2, and -T3), others are more restricted in both expression and substrate preference (ppGaNTase-T4, -T5, and -T7). ppGaNTase-T7 (14) has the distinction of being the only transferase identified thus far that requires a GalNAc-containing glycopeptide as a substrate; glycosylation of the peptide substrate by ppGaNTase-T1 is required before ppGaNTase-T7 will further glycosylate additional residues. This result indicates that not all O-linked glycosylation occurs simultaneously and suggests that a hierarchy of action within this family may be responsible for the complex patterns of multisite substrate glycosylation seen in vivo.

Here, we report the cloning of another member of this transferase family, termed ppGaNTase-T9. In common with ppGaNT-
Tase-T7, ppGaNTase-T9 demonstrated no transferase activity against a panel of unmodified peptide substrates in vitro. However, when the MUC5AC peptide substrate was first glycosylated by ppGaNTase-T1, the resultant glycopeptides were readily glycosylated further by ppGaNTase-T9 in a manner distinct from that of ppGaNTase-T7. ppGaNTase-T9 and ppGaNTase-T7 transcript expression patterns differed as well; ppGaNTase-T9 was expressed more widely across adult tissues and exhibited distinct expression patterns within developing mouse embryos. These results suggest that glycosylation of multisite substrates occurs through the specific and hierarchical action of multiple members of this enzyme family, whose expression is uniquely regulated both during development and in adult tissues.

**EXPERIMENTAL PROCEDURES**

**Isolation of ppGaNTase-T9 Probes and Full-length cDNAs**—Previously, the conserved amino acid regions EIWGGEN and VWMDEYK were used to design sense and antisense PCR primers to amplify products from rat sublingual gland (rat SLG) cDNA. These products were cloned, sequenced and used to screen a rat SLG cDNA library as described (12). To generate probes previously used to clone the rat ppGaNTase-T5 cDNA (12) resulted in the detection of additional isoforms when screening an oligo(dT)-primed UniZap XR rat SLG cDNA library according to standard procedures (18). A novel isoform, designated ppGaNTase-T9, was identified by cross-hybridization with the probe derived from positions 2076–2240 of ppGaNTase-T5, previously used to clone the rat ppGaNTase-T5 cDNA (12). All hybridizations were performed in 5× SSPE, 50% formamide against a panel of unmodified peptide substrates based on the 3′-untranslated region FN_37 (15). An oligonucleotide (d(AGACGTTGTGGCCCAGAAAAACGAGGCTC-

**Alignment and Similarity Determinations**—Amino acid sequences were aligned, one pair at a time, using the pairwise ClustalW (1.4) algorithm in MacVector (Oxford Molecular Group). The following alignment modes and parameters were used: slow alignment, open gap penalty = 10, extended gap penalty = 0.1, similarity matrix = blossom, delay divergence = 40%, and no hydrophilic gap penalty. The percentage of amino acid sequence similarity displayed in Tables I and II represents the number of identical or similar amino acids excluding the conserved domains used in Table I begin with the conserved region FNX_37 in the putative catalytic domain (amino acid position 84 in ppGaNTase-T9 (11), 100 in ppGaNTase-T2 (8), 150 in ppGaNTase-T3 (10), 102 in ppGaNTase-T4 (11), 454 in ppGaNTase-T5 (12), 142 in ppGaNTase-T6 (13), 175 in ppGaNTase-T7 (14), and 113 in ppGaNTase-T9). pBSrT7-IS was digested with a conserved amino acid coding region of 0.425 in ppGaNTase-T7, 440 in ppGaNTase-T2, 500 in ppGaNTase-T3, 438 in ppGaNTase-T4, 796 in ppGaNTase-T5, 492 in ppGaNTase-T6, 526 in ppGaNTase-T7, and 451 in ppGaNTase-T9. The segment of conserved sequences is ~340 amino acids in length in the various isoforms and corresponds to the putative catalytic domain based on structural modeling and mutagenesis studies (19). Sequences aligned in Table II consisted of the C-terminal ricin-like lectin motif (19) (amino acids 430–558 in ppGaNTase-T1, 444–569 in ppGaNTase-T2, 505–632 in ppGaNTase-T3, 444–577 in ppGaNTase-T4, 501–929 in ppGaNTase-T5, 498–622 in ppGaNTase-T6, 533–656 in ppGaNTase-T7, and 459–597 in ppGaNTase-T9).

**Northern Blot Analysis**—Total RNA from Wistar rat tissues was extracted according to the single-step isolation method described by Ausubel et al. (20). Following electrophoresis in a 1% formaldehyde-agarose gel, rat total RNA samples were transferred to Hybond-N membranes (Amer sham Pharmacia Biotech) according to Sambrook et al. (18). A 325-bp segment of the ppGaNTase-T9 cDNA region (from nucleotides 1334–1756 of the amino acid coding region was labeled using the Random Primers DNA Labeling system (Life Technologies, Inc.) according to manufacturer’s instructions and used as a probe for ppGaNTase-T9 transcripts. ppGaNTase-T7 and -T1 were detected as described previously (12, 14). Antisense 18 S ribosomal subunit oligonucleotide d(TATTGAGCTC-GAATTACCAGCGGTCTG) was end-labeled as described (21) and used to normalize sample loading by hybridizing with 5× excess of probe. 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. In situ Hybridization—In situ hybridization studies were performed using a modification of procedures described by Wilkinson and Green (21). Mouse embryos were fixed overnight in freshly prepared ice-cold 4% paraformaldehyde in phosphate-buffered saline. The embryos were then washed through ethanol and xylene and dehydrated through ethanol and toluene followed by xylene. Sections (5 μm) were adhered to commercially modified glass slides (Super Frost Plus, VWR), dewaxed in xylene, rehydrated through graded ethanol, and treated with protease K (to enhance probe accessibility) and with acetic anhydride (to reduce nonspecific background). Single-stranded antisense RNA probes were generated by standard techniques with specific activ- ities of 5 × 10^6 dpm/μg. ppGaNTase-T9 was detected using the plasmid pBSrT9-IS as a template for RNA production, ppGaNTase-T7-specific RNA probes were prepared using the plasmid pBSrT7-IS, and ppGaNTase-T1 transcripts were detected using the plasmid pBSmT1-IS. pBSrT9-IS contains nucleotides 199–381 of the rat ppGaNTase-T9 amino acid coding region generated by PCR amplification using the primers mTAIS = (d(ATAGTTCAACAGTCCTGAAACGCGTTAGGA)) and mTAIS = (d(ATAGTTCAACAGTCCTGAAACGCGTTAGGA)). pBSrT7-IS contains a segment of the rat ppGaNTase-T7 (14) amino acid coding region from nucleotide position 1759 to 1946 generated by PCR amplification using the primers mT5IS+ = (d(ATAGTTCAACAGTCCTGAAACGCGTTAGGA)) and mT5IS− = (d(ATAGTTCAACAGTCCTGAAACGCGTTAGGA)).

**Cloning and Expression of ppGaNTase-T9**—Previously, a 325-bp DNA segment from the cDNA clone rTA-0 was amplified using the primers mT5IS+ = (d(ATAGTTCAACAGTCCTGAAACGCGTTAGGA)) and mT5IS− = (d(ATAGTTCAACAGTCCTGAAACGCGTTAGGA)). This amplified product was digested with the enzymes rTA-MluI-S (d(CCTACGCCCTCCTGCGGGGTTCCAG)) and rTA-PCR-AS (d(GTACAACAAAGGGAGGAGGAACCAT)). All of the above mentioned PCR products were cloned into the KpnI and NotI sites of pBluescript KS+ plasmid. Various vectors were transfected into COS7 or subconfluent primary sublingual gland cultures using a calcium phosphate precipitation technique. Recombinant enzymes were assayed and quantitated by a microassay technique. Slides were adhered to commercially modified glass slides (Super Frost Plus, VWR), diced, dewaxed in xylene, rehydrated through ethanol and toluene, followed by xylene. Sections (5 μm) were adhered to commercially modified glass slides (Super Frost Plus, VWR), dewaxed in xylene, rehydrated through ethanol and toluene followed by xylene. Sections (5 μm) were adhered to commercially modified glass slides (Super Frost Plus, VWR), dewaxed in xylene, rehydrated through ethanol and toluene followed by xylene.

**Generation of Secretion Constructs for ppGaNTase-T9—**cDNA clones containing the 1.8-kb coding region of ppGaNTase-T9 were isolated from the rat sublingual gland cDNA library described previously (12). A 325-bp segment of the cDNA clone rTA-0 was cloned into the pBluescript KS+ vector using the primers rTA-MluI-S (d(CCTACGCCCTCCTGCGGGGTTCCAG)) and rTA-PCR-AS (d(GTACAACAAAGGGAGGAGGAACCAT)). This amplified product was digested with MluI and EcoI and cloned into the vector pBS-IMFK3 to create the vector, pBS-rTAmut#7. Sequences were confirmed to verify that no PCR-generated mutations had been sustained in the cloned product. A 660-bp MluI-EcoI (blunt) fragment of pBS-rTAmut#7 was cloned into the site designated mT9 using the plasmid pBSmT1-IS as a template for RNA production, ppGaNTase-T7-specific RNA probes were prepared using the plasmid pBSmT1-IS, and ppGaNTase-T1 transcripts were detected using the plasmid pBSmT1-IS. pBSrT9-IS contains nucleotides 199–381 of the rat ppGaNTase-T9 amino acid coding region generated by PCR amplification using the primers mT9IS+ = (d(ATAGTTCAACAGTCCTGAAACGCGTTAGGA)) and mT9IS− = (d(ATAGTTCAACAGTCCTGAAACGCGTTAGGA)). All of the above mentioned PCR products were cloned into the KpnI and NotI sites of pBluescript KS+ plasmid. Various vectors were transfected into COS7 or subconfluent primary sublingual gland cultures using a calcium phosphate precipitation technique. Recombinant enzymes were assayed and quantitated by a microassay technique. Slides were adhered to commercially modified glass slides (Super Frost Plus, VWR), diced, dewaxed in xylene, rehydrated through ethanol and toluene, followed by xylene. Sections (5 μm) were adhered to commercially modified glass slides (Super Frost Plus, VWR), dewaxed in xylene, rehydrated through ethanol and toluene, followed by xylene.
activity for ppGaNTase-T9 was detected in any of these initial assays. To generate glycopeptide substrates for analysis of ppGaNTase-T9 activity, glycosylated MUC5AC and EA2 were prepared by incubation with Pichia pastoris-derived recombinant ppGaNTase-T1 as described previously (14). Briefly, Pichia-derived ppGaNTase-T1 (0.028 μg) was incubated for 2 h at 37 °C with 1 μg of peptide in a 50 μl reaction volume under the following conditions: 125 mM cacydolate buffer (pH 7.0) containing 0.2% (v/v) Triton X-100, 12.5 mM MnCl₂, 1 mM aprotinin, 1 mM leupeptin, 1 mM EDTA, 1 mM phenylmethanesulfonyl fluoride, 1.25 mM AMP, and 6 mM cold UDP-GalNac. Additional enzyme (0.028 μg) and UDP-GalNac (3 μmol) were added at the first 24-h interval. The di- and tri-glycosylated MUC5AC reaction products and MUC5AC tri-glycopeptide were then incubated with ppGaNTase-T9 for a further 2 h. Reactions were then stopped by the addition of an equal volume of 25 μl of 10 mM EDTA, purified on a Waters 265 HPLC, and analyzed by MALDI-TOF mass spectrometry. 

The purified products of the reaction with Pichia-derived ppGaNTase-T1 were used as substrates in subsequent incubations with COS7 cell-derived ppGaNTase-T9 and ppGaNTase-T7 media to generate the data in Figs. 4-6. Equal relative amounts of each recombinant enzyme were used in each reaction as determined by SDS-PAGE analysis (12). Reactions were carried out in a total volume of 50 μl at the following concentrations: 15 μg of each peptide or glycopeptide, 125 mM cacydolate buffer (pH 7.0) containing 0.2% (v/v) Triton X-100, 12.5 mM MnCl₂, 1 mM aprotinin, 1 mM leupeptin, 1 mM EDTA, 1 mM phenylmethanesulfonyl fluoride, and 1.25 mM AMP. The enzyme samples were preincubated in this reaction mixture for 5 min at 37 °C, and then the reaction was initiated with the addition of UDP-[14C]GalNAc (54.7 mCi/μmol; 2.02 Gbq/μmol) and 20 nmol of cold UDP-GalNac. Reactions were performed for 96 h at 37 °C with additional enzyme and UDP-GalNac (22 nmol) being added after each 24-h interval. Reactions were then stopped by the addition of an equal volume of 10 mM EDTA, purified on a Waters 265 HPLC, and analyzed by capillary electrophoresis and MALDI-TOF as described above.

Periodate Oxidation, Sodium Borohydride Reduction, and Edman Degradation Assays—Purified MUC5AC glycopeptide (100 nmol) and MUC5AC parent peptide (100 nmol) were oxidized with 200 μl of 0.08 M NaIO₄ in 0.05 M acetate buffer (pH 4.5) at 4 °C for 60 h, in the dark (28) in side by side reactions. Excess periodate was destroyed by adding 20 μl of ethylene glycol. The reaction mixtures were adjusted to pH 7.5 with 1 M NaOH. Sodium borohydride was added to a final concentration of 0.2 M and reduced for 4 h at 4 °C. Excess borohydride was destroyed by the addition of 20 μl of glacial acetic acid, and released boric acid was evaporated several times with methanol. The reaction mixtures were purified by HPLC as described above. Periodate-treated and untreated MUC5AC and MUC5AC glycopeptide were then used as substrates in reactions with COS7 cell-derived ppGaNTase-T9, ppGaNTase-T7, or mock-transfected media (pMKPl) as described above to generate the data in Table III. Reactions were performed in duplicate for 24 h at 37 °C. All reactions were stopped by the addition of an equal volume of 10 mM EDTA. Reaction products were passed through AG1-X8 resin and purified by HPLC as described above and analyzed by MALDI-TOF mass spectrometry to determine the mass of each fragment. Fractions were then pooled according to the molecular mass and used for sequence analysis by Edman degradation.

RESULTS

cDNA Cloning and Sequence Analysis of ppGaNTase-T9—A PCR strategy based on conserved regions within the ppGaNTase family was performed on cDNA from rat SLG; the resultant products were purified, cloned, and sequenced to identify the nature of the insert as described previously (12). A PCR product that previously identified the rat ppGaNTase-T5 cDNA (12) resulted in the detection on another novel cDNA, which shared homology to previously identified isoforms. The cDNA clone (rTa-0) contained a 3′ truncation within the coding region. To obtain a full-length clone, an oligonucleotide based on the 3′-most sequence of the partial cDNA clone (d(AGACGTTGTTGCCCAAGAAAAACTCCGAGGCTCC)) was end-labeled and used to screen the cDNA library a second time. A cDNA clone containing a complete open reading frame was isolated (rTa-3), sequenced, and given the designation ppGaNTase-T9. As shown in Fig. 1, the cDNA encoding ppGaNTase-T9 contains a 1812-bp insert encoding a unique 603-amino acid protein. Conceptual translation of this cDNA revealed a type II membrane protein architecture, typical of the ppGaNTase family. The enzyme consists of a 9-amino acid N-terminal cytoplasmic region, a 20-amino acid hydrophobic/transmembrane region, a 94-amino acid stem region, and a 480-amino acid putative catalytic region. Table I summarizes the degree of amino acid similarity between each of the known isoforms within the conserved putative catalytic region. ppGaNTase-T9 displays the greatest degree to similarity within this region to ppGaNTase-T1 and the lowest degree of similarity to ppGaNTase-hT6. Amino acid similarity within the C-terminal ricin-like lectin motif is shown in Table II. This domain displays homology to the carbohydrate binding region of the plant lectin, ricin, and has been hypothesized to be involved in enzyme recognition of carbohydrate moieties on glycopeptide substrates (19, 31). Within this region, ppGaNTase-T9 has the greatest similarity to ppGaNTase-mT3 and the least to ppGaNTase-mT2.

Northern Blot Analysis—Northern blots of rat total RNA were probed with a ppGaNTase-T9 specific probe (Fig. 2) as well as probes specific for the previously characterized ppGaNTase-T7 and -T1 (14). The highest levels of the 5- and 4.2-kb ppGaNTase-T9 message were found in the SLG, testis, small intestine, colon, and ovary. Smaller amounts were detectable in heart, brain, spleen, lung, stomach, cervix, and uterus. ppGaNTase-T7 transcripts were much more restricted in their expression, whereas ppGaNTase-T1 transcripts were much more ubiquitous, as seen previously (14). The 18S ribosomal RNA was used as a control for RNA integrity and loading variations.

Mouse Embyronic in Situ Hybridization Analysis—Given the degree of amino acid conservation and nucleic acid homology for each specific isoform across species as well as similarity of expression patterns seen in adult tissues (14), we examined ppGaNTase-T9 gene expression during mouse development using parasagittal sections of embryos during late organogenesis...
Theiler stage 22–23, embryonic day 14.5. The region of the rat ppGaNTase-T9 gene used as a probe is 96% homologous to the corresponding mouse EST. Sections were hybridized with RNA probes specific for ppGaNTase-T9, -T7 and -T1 and compared with each other (Fig. 3). ppGaNTase-T9 is expressed relatively abundantly compared with ppGaNTase-T7 (Fig. 3, A versus B) and in a more restricted pattern than ppGaNTase-T1 (Fig. 3, C).

A higher magnification view of the developing hindbrain region in these animals (Fig. 3, D–F) shows discrete accumulation of ppGaNTase-T9 transcripts in the rapidly dividing, undifferentiated...
Fig. 2. Northern blot analysis of ppGaNTase-T9, -T7 and -T1. Total RNA from Wistar rats was extracted from glands and organs listed above each lane. After electrophoresis on 1% formaldehyde-agarose gels and transfer to Hybond-N membranes, RNA was hybridized with a ppGaNTase-T9-specific probe (T9), a -T7-specific probe (T7), a -T1-specific probe (T1), and an 18S rRNA probe (18S). The T9 panel and the uppermost 18S panel were from hybridizations of the same blot, whereas the T7 panel, T1 panel, and lowermost 18S panel were from hybridizations of another independent blot. Each lane contains 7.5 µg of total RNA. Size markers are indicated on the left. SL Gland, submandibular gland; SM Gland, sublingual gland; Sm Intestine, small intestine.

Cloning and Expression of ppGaNTase-T9

Functional Expression—The truncated coding region of ppGaNTase-T9 (beginning at amino acid position 39) was cloned downstream of the insulin secretion signal, heart muscle kinase site, and FLAG epitope tag in the vector pIMKF4 to generate the construct pH4-T9. pH4-T9, pF4-T9 as well as similar constructs containing a truncated mouse ppGaNTase-T1 gene (11), a truncated rat ppGaNTase-T7 gene (14), or no insert (pIMKF1) (11) were transfected into COS7 cells as described previously (12). The expressed products from these transfections were harvested from the culture media and used in in vitro glycosylation reactions. Equivalent amounts of each secreted product, as judged by densitometric scanning of Tricine SDS-PAGE gels (data not shown) were used for all enzymatic assays. Initially, in vitro glycosylation activity was seen for ppGaNTase-T9 (data not shown) against a panel of unmodified peptides used previously (14). This result mirrored our initial observations with ppGaNTase-T7, where unmodified peptides would not serve as substrates for this enzyme. Therefore, we prepared glycosylated peptide substrates by incubating the MUC5AC peptide with Pichia pastoris-purified ppGaNTase-T1 enzyme for extended periods of time in the presence of excess nucleotide sugar. The reaction products consisting of di- and tri-substituted glycopeptides were individually purified by HPLC and analyzed by capillary electrophoresis and mass spectrometry to confirm their identity (Fig. 4). Purified di- and tri-glycopeptides (m/z = 1930 and 2133, respectively) were then incubated with equal relative amounts of recombinant ppGaNTase-T9 or ppGaNTase-T7 derived from COS7 cell culture media. The products from these reactions were then analyzed by capillary electrophoresis and mass spectrometry (Fig. 4). ppGaNTase-T9 clearly acts as a glycopeptidase transferase, converting the di-glycopeptide starting material to more extensively glycosylated tri-glycopeptide (m/z = 2133) and tetra-glycopeptide (m/z = 2337) species. ppGaNTase-T7 activity on the di-glycopeptide results in the formation of products distinct from those formed by ppGaNTase-T9, predominantly penta-glycopeptide (m/z = 2540) and hexa-glycopeptide (m/z = 2744) species. When the tri-glycopeptide is used as a substrate, ppGaNTase-T7 activity results in the formation of a tetra-glycopeptide (m/z = 2337), even after extended incubations, whereas ppGaNTase-T7 forms a more heavily glycosylated hexa-glycopeptide (m/z = 2743) (Fig. 4).

The substrates and products of the aforementioned reactions were then sequenced by Edman degradation to determine the sites of GalNAc addition by each enzyme (Fig. 5). Fig. 5A shows the HPLC profiles for residues 1–3 and 9–13 of the MUC5AC parent peptide, the di- and tri-glycosylated species produced by ppGaNTase-T1 and the hexa-glycosylated species produced by ppGaNTase-T7. The * and ** denote the diastereomeric peaks indicative of PTH-Thr-O-GalNAc, and the *** indicates the unresolved doublet peak of PTH-Ser-O-GalNAc. The HPLC profiles indicate that ppGaNTase-T1 glycosylates threonines 3
and 13 in the diglycosylated species and threonines 3, 12, and 13 in the tri-glycosylated species. (Our earlier work indicating that T1 glycosylates serine 5 was in error due to misinterpretation of an additional proline peak in the serine 5 HPLC profile; proline 4 in the MUC5AC sequence gave a peak near the position of PTH-Ser-O-GalNAc, which carried over into the serine 5 HPLC profile and was mistakenly assumed to indicate a glycosylated serine. Since then we have repeated the Edman degradation multiple times to conclusively assign modified positions.)

Upon incubation with the tri-glycopeptide, ppGaNTase-T7 glycosylates threonines 2 and 10 and serine 11 to form the hexa-glycopeptide (Fig. 5A). The sites of GalNAc addition in the hexa-glycopeptide were confirmed by limited proteinase K digestion of this species and analysis of the products by mass spectrometry and Edman degradation (Fig. 5B). The two peaks produced by this analysis correspond to the first 9 residues of MUC5AC substituted with 2 GalNAc residues (m/z 51284) and the last 7 residues substituted with 4 GalNAc residues (m/z 51498). Edman degradation of these fragments confirmed previous sequencing of the unfragmented glycopeptide (data not shown). We recovered insufficient penta-glycopeptide formed by ppGaNTase-T7 to perform sequence analysis.

To further define the requirement of the ppGaNTase-T9 isoform for a GalNAc-containing substrate, we modified GalNAc residues by periodate oxidation and sodium borohydride reduction. The purified glycopeptides obtained from incubation with ppGaNTase-T1, along with the MUC5AC parent peptide, were subjected to mild periodate oxidation followed by sodium borohydride reduction. Periodate-treated and untreated glycopeptides and MUC5AC parent peptide were purified by HPLC, analyzed for integrity by capillary electrophoresis (data not shown), and incubated with COS7 cell-derived ppGaNTase-T1, ppGaNTase-T9, or mock-transfected (pIMKF1) media. Table III compares the counts incorporated into each substrate by each enzyme. The ability of ppGaNTase-T9 to use the glycopeptide as a substrate is clearly reduced upon treatment with periodate and sodium borohydride (compare 17080 cpm incorporated into untreated material to 381 cpm incorporated into treated material) (Table III). However, this reduction in incorporation by ppGaNTase-T9 is not due to the peptide being compromised during periodate treatment, as ppGaNTase-T1 works equally well on both treated and untreated MUC5AC (compare 21545 cpm to 18100 cpm) (Table III). These data suggest that ppGaNTase-T9, like ppGaNTase-T7, requires the presence of intact GalNAc on the MUC5AC peptide for it to be used as a substrate.

To begin to address the activity of these enzyme hierarchies on peptides other than MUC5AC, we incubated P. pastoris-derived ppGaNTase-T1 with the EA2 peptide (m/z 51340) for extended periods of time in the presence of excess UDP-GalNAc as described for MUC5AC. MALDI-TOF analysis and Edman degradation of the product of this incubation indicate that ppGaNTase-T1 produces a mono-glycopeptide (m/z 51543).
with GalNAc at threonine 7 (Fig. 6, A and B). The same mono-glycosylated species is also produced upon incubation of EA2 with ppGaNTase-T2 (data not shown). When this mono-glycosylated species is then used as a substrate in subsequent incubations with ppGaNTase-T9 or -T7, a di-glycosylated species (m/z 51746) is formed by both enzymes (Fig. 6A). (The additional small peak present in the ppGaNTase-T7 sample most likely represents a trace amount of tri-glycosylated species (m/z 51949) that was variably present and in quantities too low to be recovered for further analysis.) Both di-glycosylated species showed an additional GalNAc at threonine 6, indicating that ppGaNTase-T9 and -T7 are transferring GalNAc to the same residue in this glycopeptide, producing the same final product; this is in contrast to their respective activities on the MUC5AC glycopeptides.

**DISCUSSION**

We report the cloning of a novel member of the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family, termed ppGaNTase-T9. ppGaNTase-T9 encodes a type II integral membrane protein similar in structure to previously identified family members. In common with ppGaNTase-T7, ppGaN-
Tase-T9 fails to act on a panel of unmodified peptide substrates, but rather catalyzes the transfer of GalNAc from UDP-GalNAc to a GalNAc-containing peptide substrate. This activity requires the prior activity of another member of the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family (ppGaNTase-T1 or -T2) and results in the modification of an additional hydroxyamino acid within the glycopeptide substrate. Periodate oxidation further demonstrated that ppGaNTase-T9 also requires the presence of an intact GalNAc residue on the glycopeptide substrate. We have determined that ppGaNTase-T9 will act on di- and tri-glycosylated MUC5AC and mono-glycosylated EA2, indicating that ppGaNTase-T9 will recognize different glycoforms of a given peptide as well as more than one type of glycopeptide substrate.

Although both ppGaNTase-T9 and ppGaNTase-T7 require glycosylated substrates, their activities on the MUC5AC glycopeptide substrates are clearly distinct. Whereas ppGaNTase-T7 catalyzes the formation of hexa-glycopeptides from the tri-glycopeptide substrate, ppGaNTase-T9 produces a tetraglycopeptide species, even after extended incubation. Edman degradation revealed that both ppGaNTase-T7 and ppGaNTase-T9 glycosylate threonine 2; however, ppGaNTase-T7 additionally acts upon threonine 10 and serine 11. In contrast to the results with the MUC5AC glycopeptide, ppGaNTase-T9 and -T7 act similarly on mono-glycosylated EA2, indicating that ppGaNTase-T9 will recognize different glycoforms of a given peptide as well as more than one type of glycopeptide substrate.

In both the MUC5AC and EA2 glycopeptides examined, ppGaNTase-T9 adds GalNAc to the position immediately N-terminal to a previously glycosylated residue. ppGaNTase-T7 also transfers to the position immediately N-terminal from a glycosylated threonine in the EA2 glycopeptide and N-terminal from glycosylated threonines in the MUC5AC tri-glycopeptide. Increased glycosylation of sites vicinal to preexisting GalNAc residues has also been observed in vitro using an undefined mixture of transferases present in human milk (32). It is known from the analysis of mucins, that sites of glycosylation tend to be clustered in vivo (33). This clustering may reflect specific GalNAc recognition and subsequent local addition of GalNAc by the glycopeptide transferases. The production of large quantities of specifically designed glycopeptides is necessary to be able to conclusively address the effects of number and position of preexisting GalNAc residues on the activity and subsequent GalNAc addition by the glycopeptide transferases.

Amino acid comparisons of all known functional mammalian ppGaNTases have not uncovered regions of greater conservation between ppGaNTase-T9 and ppGaNTase-T9 relative to the other family members (Tables I and II), including regions within the ricin-like lectin motif. However, a larger panel of glycopeptide-specific enzymes on which to base comparisons...
may aid in deciphering regions involved in the specific recognition of a glycopeptide substrate. Previous work in the nematode, Caenorhabditis elegans, identified nine ppGaNTase isoforms (34), but enzymatic activity was detectable for only five. It is possible that the remaining four isoforms may also require a previously glycosylated peptide as a substrate. One recent study reports that a single amino acid change within the ricin-like lectin motif of ppGaNTase-T4 compromises the glycopeptide transference activity of this enzyme (31). However, ppGaNTase-T4 can act as both a peptide and glycopeptide transferase, and it is unclear what specific affect this mutation had on substrate binding and/or catalytic activity, as kinetic parameters were not investigated.

The gene expression patterns of ppGaNTase-T9 and ppGaNTase-T7, like their enzymatic activities, display some overlap yet are quite distinct. By Northern analysis, ppGaNTase-T9 is broadly expressed across many adult tissues in the rat, including the sublingual gland, digestive tract, female reproductive tract, testis, heart, brain, spleen, and lung. This tissue distribution is more restricted than the near ubiquitous expression seen for ppGaNTase-T1 yet not as specific as that seen for ppGaNTase-T9 and -T7, like their enzymatic activities, display some overlap and -T4 can act as both a peptide and glycopeptide transferase, perhaps uninformative phenotypes. There exist a number of examples where deletions of single genes from other glycosyltransferase families have resulted in viable, fertile mice without out distinguishing phenotypes (39). Therefore, the ablation of multiple isoforms displaying similar enzymatic activity (e.g. ppGaNTase-T7 and -T9) and/or expression profiles may be necessary. Our current efforts to characterize spatial and temporal expression and activity of each member of this family will aid in making informed choices as to which combination of gene ablations may provide insightful phenotypes.

REFERENCES

1. Van den Steen, P., Ruddle, P. M., Dreek, R. A., and Opdenakker, G. (1998) Crit. Rev. Biochem. Mol. Biol. 33, 231–268

2. Marinaro, J. A., Neumann, G. M., Russo, V. C., Leeding, K. S., and Bach, L. A. (1999) Eur. J. Biochem. 267, 5378–5386

3. Woodward, H. D., Kingler, N. J., Selvakumar, R., Simek, I. M., Bhuvanand, V. P., and Davidsson, E. A. (1987) Biochemistry 26, 5315–5322

4. Kinloch, R. A., Sakai, Y., and Wasserman, P. M. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 263–267

5. Dankort, D., Alitalo, A., Young, P. E., Lusis, A. J., and Lashk, L. A. (1993) J. Biol. Chem. 268, 4525–4529

6. Hagen, F. K., VanWuychbey, B., and Tabak, L. A. (1993) J. Biol. Chem. 268, 11960–11965

7. Homa, F. L., Hollander, T., Lehman, D. J., Thomsen, D. R., and Elhammer, A. P. (1993) J. Biol. Chem. 268, 12609–12616

8. White, T., Bennett, E. P., Takio, K., Sarenets, T., Bonding, N., and Clausen, H. (1999) J. Biol. Chem. 274, 23451–23454

9. Bennett, E. P., Hassan, H., and Clausen, H. (1996) J. Biol. Chem. 271, 17006–17012

10. Zara, J., Hagen, F. K., Ten Hagen, K. G., VanWuychbey, B. C., and Tabak, L. A. (1996) Biochem. Biophys. Res. Commun. 228, 38–44

11. Hagen, F. K., Ten Hagen, K. G., Beres, T. M., Balya, M. M., and Tabak, L. A. (1997) J. Biol. Chem. 272, 13848–13848

12. Ten Hagen, K. G., Hagen, F. K., Balya, M. M., Beres, T. M., VanWuychbey, B. C., and Tabak, L. A. (1998) J. Biol. Chem. 273, 27749–27754

13. Bennett, E. P., Hassan, H., Mandel, U., Hollingsworth, M. A., Aikawa, N., Ikematsu, Y., Merkl, G., van Kessel, A. G., Olufson, S., and Clausen, H. (1999) J. Biol. Chem. 274, 25362–25370

14. Ten Hagen, K. G., Tetaart, D., Hagen, F. K., Richet, C., Beres, T. M., Gagnon, J., Balya, M. M., VanWuychbey, B. C., Bedi, G. S., Degand, P., and Tabak, L. A. (1999) J. Biol. Chem. 274, 27867–27874

15. Bennett, E. P., Hassan, H., Hollingsworth, M. A., and Clausen, H. (1999) FEBS Lett. 450, 226–230

16. Hennet, T., Hagen, F. K., Tabak, L. A., and Marsh, J. D. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 12070–12074

17. Marsh, J. D. (1996) Glyobiology 6, 701–705

18. Sambrook, J., Fritial, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

19. Ten Hagen, K. G., Tetaart, D., Hagen, F. K., Richet, C., Beres, T. M., Gagnon, J., Balya, M. M., VanWuychbey, B. C., Bedi, G. S., Degand, P., and Tabak, L. A. (1999) J. Biol. Chem. 274, 16845–16852

20. Allen, A., Hutton, D. A., and Pearson, J. P. (1998) Int. J. Biochem. Cell Biol. 30, 1001–1004

21. Gendler, S. J., Lancaster, C. A., Taylor-Papadimitriu, J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E.-N., and Wilson, D. (1999) J. Biol. Chem. 274, 15368–15373

22. Ohmori, H., Doerman, A. F., Gallup, M., Tsuda, T., Kai, H., Gum, J. R., Jr., Kim, Y. S., and Basbaum, C. B. (1994) J. Biol. Chem. 274, 16845–16852

23. Shekels, L. S., Lylbyg, C., Kieliaszewski, M., Filie, D. G., Kozak, C. A., and Ho, G. (1995) Biochem. J. 299, 775–785

24. Guyonnet-Dupe`rat, V., Audie`, J. P., Debailleul, V., Laine, A., Buisine, M. P., Galiegue-Zouitina, S., Pigny, P., Degand, P., Aubert, J., and Porchet, N. (1999) Biochem. J. 345, 27749–27754

25. Wilkinson, D. G., and Green, J. (1999) in Postimplantation Mammalian Embryos: A Practical Approach (Copp, A. J., and Cockcroft, D. L., eds), pp. 155–171, IRL Press, New York

26. Albone, E. F., Hagen, F. K., VanWuychbey, B. C., and Tabak, L. A. (1994) J. Biol. Chem. 269, 16845–16852

27. Allen, A., Hutton, D. A., and Pearson, J. P. (1998) Int. J. Biochem. Cell Biol. 30, 1371–1381

28. Muller, S., Goetz, S., Packer, N., Gooley, A., Lawaso, A. M., and Hanisch, F. (1997) J. Biol. Chem. 272, 23470–23473

29. Hagen, F. K., Ten Hagen, K. G., Balya, M. M., Zara, J., and Tabak, L. A.,
Cloning and Expression of ppGalTase-T9

36. Bennett, E. P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M. A., Merkus, G., van Kessel, A. G., Eiberg, H., Steffensen, R., and Clausen, H. (1998) J. Biol. Chem. 273, 30472–30481
37. 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
38. Röttger, S., White, J., Wandall, H. H., Olivo, J.-C., Stark, A., Bennett, E. P., Whitehouse, C., Berger, E. G., Clausen, H., and Nilsson, T. (1998) J. Cell Sci. 111, 45–60
39. Hennett, T., and Ellies, L. G. (1999) Biochim. Biophys. Acta 1473, 123–136
Cloning and Characterization of a Ninth Member of the UDP-GalNAc:Polypeptide N-Acetylgalactosaminyltransferase Family, ppGaNTase-T9
Kelly G. Ten Hagen, Gurrinder S. Bedi, Daniel Tetaert, Paul D. Kingsley, Fred K. Hagen, Marlene M. Balys, Thomas M. Beres, Pierre Degand and Lawrence A. Tabak

J. Biol. Chem. 2001, 276:17395-17404.
doi: 10.1074/jbc.M009638200 originally published online February 6, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M009638200

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

This article cites 35 references, 24 of which can be accessed free at http://www.jbc.org/content/276/20/17395.full.html#ref-list-1