Characterization of a UDP-GalNAc:Polypeptide N-Acetylgalactosaminyltransferase That Displays Glycopeptide N-Acetylgalactosaminyltransferase Activity*

(Received for publication, December 7, 1998, and in revised form, June 15, 1999)

Kelly G. Ten Hagen‡§, Daniel Tetaert‡§, Fred K. Hagen‡, Colette Richet‡, Thomas M. Beres‡, Jean Gagnon, Marlene M. Balys‡, Brian VanWuyckhuyse‡, Gurrinder S. Bedi‡, Pierre Degand‡, and Lawrence A. Tabak‡**

From the ‡Center for Oral Biology, Rochester Institute of Biomedical Sciences, University of Rochester, Rochester, New York 14642, §Unité INSERM 377, Biologie et Physiopathologie de Cellules Mucipares, Place de Verdun, 59045 Lille Cedex, France, and ¶Institut de Biologie Structurale JP EBEL, CEA/CNRS, 41 Avenue de Martyrs, 38027 Grenoble Cedex 1, France

We report the cloning, expression, and characterization of a novel member of the mammalian UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase (ppGaNTase) family that transfers GalNAc to a GalNAc-containing glycopeptide. Northern blot analysis revealed that the gene encoding this enzyme, termed ppGaNTase-T6, is expressed in a highly tissue-specific manner. Significant levels of transcripts were found in rat and mouse sublingual gland, stomach, small intestine, and colon; trace amounts were seen in the ovary, cervix, and uterus. Recombinant constructs were expressed transiently in COS7 cells but demonstrated no transferase activity in vitro against a panel of unmodified peptides, including GTTPSPVPTTS(T/T)SAP (MUC5AC). However, when incubated with the total glycosylated products obtained by action of ppGaNTase-T1 on MUC5AC (mainly G(T/T)GalNAcPSPVPTTS(T/T)GalNAc)SAP, additional incorporation of GalNAc was achieved, resulting in new hydroxyamino acids being modified. The MUC5AC glycopeptide failed to serve as a substrate for ppGaNTase-T6 after modification of the GalNAc residues by periodate oxidation and sodium borohydride reduction, indicating a requirement for the presence of intact GalNAc. This suggests that O-glycosylation of multisite substrates may proceed in a specific hierarchical manner and underscores the potential complexity of the processes that regulate O-glycosylation.

O-Linked glycans are involved in a number of biological functions including leukocyte trafficking (1) and sperm-egg adhesion (2). In addition, clusters of O-linked oligosaccharides impart a “stalk-like” conformation that is common among several membrane receptors (3). In contrast to N-linked glycosylation, O-linked glycans are synthesized stepwise. Thus, the acquisition of GalNAc represents the first step in mammalian (mucin-type) O-glycosylation. A family of UDP-GalNAc:

polypeptide N-acetylgalactosaminyltransferase enzymes (pp-GaNTase, EC 2.4.1.41) is responsible for this initial enzymatic step. Five family members (ppGaNTase-T1 (4, 5), -T2 (6), -T3 (7, 8), -T4 (9), and -T5 (10)) have been identified in mammals thus far and have been shown to have unique expression patterns as well as substrate specificities. However, little is known regarding their respective activities on native substrates as well as potential inter-relationships with one another.

In the present study, we have cloned a novel member of this enzyme family termed ppGaNTase-T6. When recombinant enzyme was expressed as a secreted product from COS7 cells, no ppGaNTase activity was detected in vitro against a panel of unmodified peptides, including the peptide GTTPSPVPTTSTTSAP, which is derived from the human MUC5AC gene sequence (11). However, when this MUC5AC peptide was first glycosylated with ppGaNTase-T1 to yield mainly GTTP(S/P)VPTTS(T/T)SAP (but also mono- and tri-substituted species), the ppGaNTase-T6 isofrom was active toward the glycopeptidic preparation. This suggests that the addition of the initial O-linked sugar may occur in a hierarchical manner with the action of certain ppGaNTases necessary prior to the action of others.

EXPERIMENTAL PROCEDURES

Isolation of ppGaNTase-T6 Probes and Full-length cDNAs—The conserved amino acid regions EIWGGEN and VVMDEYK were used to design sense and antisense PCR primers, dGARATHTGGGGNGN-GARA (321 sense) and d(CTTATYCTCRCTCCANAC) (379 antisense). These were used to perform PCR reactions on rat sublingual gland (rat SLG) cDNA; the resultant 200-base pair PCR products were cloned into M13 vectors and screened as described previously (10). Positively hybridizing M13 clones were sequenced with fluorescent dye-labeled primers on an LI-COR DNA 4000L DNA sequencer. The insert from a unique clone was used to generate an asymmetrically labeled PCR probe using the oligonucotide 379 antisense. This probe was then used to screen $1 \times 10^6$ plaques from an oligo(dT)-primed Uni-Zap XR rat sublingual gland cDNA library (10) according to standard procedures (12). One of the four positive clones obtained was fully sequenced. The N-terminal transmembrane domain was determined by a Kyte-Doolittle hydrophobicity plot. Sequence alignments were performed using the Clustal method of Megalign (DNASTAR) and began at the conserved region FNXXXSXD in the putative luminal domain (amino acid position 84 in ppGaNTase-T1, 100 in ppGaNTase-T2, 150 in pp-

* This work was supported in part by National Institutes of Health Grant DE-08108 (to L. A. T.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF076167 (rat ppGaNTase-T6).

§ The first two authors contributed equally to this work.

** To whom correspondence should be addressed: Center for Oral Biology, Rochester Institute of Biomedical Sciences, University of Rochester, 601 Elmwood Ave., Box 611, Rochester, NY 14642. Tel.: 716-275-0770; Fax: 716-473-2679; E-mail: Lawrence_Tabak@urmc.rochester.edu.
Characterization of ppGaNTase-T6

GaNTase-T3, 102 in ppGaNTase-T4, 454 in ppGaNTase-T5, and 524 in ppGaNTase-T6). The segment of amino acid sequence aligned, one pair at a time, using the pairwise ClustalW (1.4) algorithm. The following alignment modes and parameters were used: slow alignment, open gap penalty 1.2, extension gap penalty 0.1, similarity matrix blosum, delay 50, initation codon. The N-terminal transmembrane domain (at nucleotide position 107 of ppGaNTase-T6). GAT), which generates an excess of probe. All hybridizations were performed in 5% formamide at 42 °C with two final washes in 2× SSPE, 50% formamide at 42 °C with two final washes in 2× SSC, 0.1% SDS at 65 °C for 20 min.

**Northern Blot Analysis**—Total RNA from BALB/c mouse and Wistar rat tissues was extracted according to the single step isolation method described in Ausubel et al. (13). Following electrophoresis in a 1% formaldehyde-agarose gel, rat and mouse total RNA samples were transferred to Hybond-N membranes (Amersham Pharmacia Biotech) (12). A segment of the ppGaNTase-T6 cDNA sequence, from nucleotide position 1305 to 1473, was labeled by asymmetric PCR (14) using the antisense oligonucleotide d(GACTTC-CACAACACGA) as a probe. Levels of recombinant enzymes were analyzed by Tricine SDS-PAGE (16) after labeling with [γ-32P]ATP using heart muscle kinase (HMK) as described previously (data not shown) (10). Gels were dried under vacuum and exposed to film (XAR, Eastman Kodak Co.) or quantitated on a PhosphorImager (Molecular Dynamics).

**Functional Assays of Secreted Recombinant ppGaNTase-T6 from COS7 Cells**—Activities of ppGaNTase-T1 and -T6 were initially measured by assaying for the following panel of peptide substrates as described previously (9, 10): EA2 (PTTDSSTTPAPTUK from the tandem repeat of rat submandibular gland mucin (17); HIV (RGPGRAFVTIGKIGNMR) from the human immunodeficiency virus gp120 protein (7); MUC2 (PTTPITISSTTVPTTPPTFC) from human intestinal mucin (18); MUC1b (DPTRPAPGSTAPPAC) from human MUC1 mucin (19); MUC1 (PDTRPAPGSTAPPAC) from human gastric mucin (20); MUC5AC (SPTTSTPISSTPQPTS) from rat intestinal mucin (20); mG-MUC (EQPVTPTTPCSTE) from mouse gastric mucin (21); and MUC5AC (TPSPSTTPISSTPQPTS) from human intestinal mucin (20). Equivalent amounts (units) of each enzyme (as determined by SDS-PAGE gel quantitation) were used in each assay. No enzymatic activity for ppGaNTase-T6 was detected in any of these initial assays. Subsequent activity (15) was performed with cell culture media and the MUC5AC peptide (GTPSVPTTPTTSSAP) in a total volume of 40 μl at the following final concentrations: 1 μM peptide substrate; 125 μM MES buffer (pH 7.0) containing 0.2% (v/v) Triton X-100; 12.5 mM MnCl2; 1 mM 4-(2-amino-3-methylbutane; 1 mM phenylthioureasulfonanil fluoride; 1 mM trans-epoxy-succinyl-l-leucyl-arginine-6-methyl butanone; 1 mM phenylmethylsulfonyl fluoride,

**Amino Acid Similarity Determinations**—Amino acid sequences were aligned, one pair at a time, using the pairwise ClustalW (1.4) algorithm. The following alignment modes and parameters were used: slow alignment, open gap penalty 10, extended gap penalty 0.1, similarity matrix blosum, delay 50, initation codon. The N-terminal transmembrane domain (at nucleotide position 107 of ppGaNTase-T6). GAT), which generates an excess of probe. All hybridizations were performed in 5% formamide at 42 °C with two final washes in 2× SSPE, 50% formamide at 42 °C with two final washes in 2× SSC, 0.1% SDS at 65 °C for 20 min.

**Functional Assays of Secreted Recombinant ppGaNTase-T6 from COS7 Cells**—Activities of ppGaNTase-T1 and -T6 were initially measured by assaying for the following panel of peptide substrates as described previously (9, 10): EA2 (PTTDSSTTPAPTUK from the tandem repeat of rat submandibular gland mucin (17); HIV (RGPGRAFVTIGKIGNMR) from the human immunodeficiency virus gp120 protein (7); MUC2 (PTTPITISSTTVPTTPPTFC) from human intestinal mucin (18); MUC1b (DPTRPAPGSTAPPAC) from human MUC1 mucin (19); MUC1 (PDTRPAPGSTAPPAC) from human gastric mucin (20); MUC5AC (SPTTSTPISSTPQPTS) from rat intestinal mucin (20); mG-MUC (EQPVTPTTPCSTE) from mouse gastric mucin (21); and MUC5AC (TPSPSTTPISSTPQPTS) from human intestinal mucin (20). Equivalent amounts (units) of each enzyme (as determined by SDS-PAGE gel quantitation) were used in each assay. No enzymatic activity for ppGaNTase-T6 was detected in any of these initial assays. Subsequent activity (15) was performed with cell culture media and the MUC5AC peptide (GTPSVPTTPTTSSAP) in a total volume of 40 μl at the following final concentrations: 1 μM peptide substrate; 125 μM MES buffer (pH 7.0) containing 0.2% (v/v) Triton X-100; 12.5 mM MnCl2; 1 mM 4-(2-amino-3-methylbutane; 1 mM phenylthioureasulfonanil fluoride; 1 mM trans-epoxy-succinyl-l-leucyl-arginine-6-methyl butanone; 1 mM phenylmethylsulfonyl fluoride,
and 1.25 mM AMP. The enzyme samples were preincubated in this reaction mixture for 20 min at 37 °C, and then the reaction was initiated with the addition of UDP-\(\text{N-}\)acetylgalactosamine and UDP-[3H]GalNAc (7.8 Ci mmol\(^{-1}\); 288.6 GBq mmol\(^{-1}\); 0.1 mCi ml\(^{-1}\)) to final concentrations of 1 mM and 1.25 nM, respectively. Reactions proceeded for 24 h at 37 °C. To minimize the possibility of dipeptidylaminotransferase or other peptidase activities that could confound the MALDI-MS analysis, the thiol inhibitor \(\text{trans-epoxysuccinyl-L-leucylamido-3-methyl butane}\) and serine peptidase inhibitors were included in the reaction mixture as described previously (22). Reactions were stopped by the addition of 8 volumes of 20 mM sodium borate, 1 mM EDTA (pH 9.1). Reaction products were passed through AG1-X8 resin and eluted with 3 ml of water, and incorporation was determined by scintillation counting. Background values obtained from controls incubated without peptide substrate were subtracted from each experimental value. Products from the above mentioned incubations using ppGaNTase-T1 cell culture media were recovered by using Sep-Pak C18 reverse-phase cartridges (Waters Corp., Milford, MA) as described previously (23). The products of the reaction by ppGaNTase-T1 were used as substrates (in place of MUC5AC parent peptide) in subsequent incubations with ppGaNTase-T6 and mock (pIMKF1) media. For this second step of N-acetylgalactosaminylation, the conditions were identical to those described above, except that 1 nM UDP-[14C]GalNAc (54.7 mCi mmol\(^{-1}\); 2.02 GBq mmol\(^{-1}\); 0.02 mCi ml\(^{-1}\)) replaced 1.25 nM UDP-[3H]GalNAc. Reactions were performed for 24 h at 37 °C and were stopped as described above.

Reaction products from the aforementioned enzyme assays were analyzed by mass spectrometry and/or capillary electrophoresis. To desalt samples prior to capillary electrophoresis and/or mass spectrometry, Sep-Pak C18 cartridges were used as described previously (23). Matrix-assisted laser desorption ionization mass spectrometry was performed using a Vision 2000 time-of-flight instrument (Finnigan MAT, Bremen, Germany) equipped with a 337 nm UV laser. The mass spectra were acquired in reflectron mode under 6 kV acceleration voltage and positive detection. The samples were prepared by mixing directly onto the target 1 ml of analyzed solution (typically 50 pmol) and 1 ml of 2,5-dihydroxybenzoic acid matrix solution (12 mg ml\(^{-1}\) in CH\(_3\)OH/H\(_2\)O, 70:30, v/v) and then allowed to crystallize at room temperature. External calibration was performed using the MUC5AC peptide (\(M_r\) 1502.7). 10–30 shots were accumulated for the mass spectrum.

Capillary electrophoresis was performed on a P/ACE system model 5000 (Beckman, Fullerton, CA) under conditions previously described (23). For the separation of the hexadecapeptides, 2N formic acid buffer with 2.5% polyvinyl alcohol (\(M_r\) 15,000) (v/v) was used. The mass spectra were acquired in reflectron mode under 6 kV acceleration voltage and positive detection. The samples were prepared by mixing directly onto the target 1 ml of analyzed solution (typically 50 pmol) and 1 ml of a 2.5-dihydroxybenzoic acid matrix solution (12 mg ml\(^{-1}\) in CH\(_3\)OH/H\(_2\)O, 70:30, v/v) and then allowed to crystallize at room temperature. External calibration was performed using the MUC5AC peptide (\(M_r\) 1502.7). 10–30 shots were accumulated for the mass spectrum.

and 1.25 mM AMP. The enzyme samples were preincubated in this reaction mixture for 20 min at 37 °C, and then the reaction was initiated with the addition of UDP-N-acetylgalactosamine and UDP-[3H]GalNAc (7.8 Ci mmol\(^{-1}\); 288.6 GBq mmol\(^{-1}\); 0.1 mCi ml\(^{-1}\)) to final concentrations of 1 mM and 1.25 mM, respectively. Reactions proceeded for 24 h at 37 °C. To minimize the possibility of dipeptidylaminotransferase or other peptidase activities that could confound the MALDI-MS analysis, the thiol inhibitor \(\text{trans-epoxysuccinyl-L-leucylamido-3-methyl butane}\) and serine peptidase inhibitors were included in the reaction mixture as described previously (22). Reactions were stopped by the addition of 8 volumes of 20 mM sodium borate, 1 mM EDTA (pH 9.1). Reaction products were passed through AG1-X8 resin and eluted with 3 ml of water, and incorporation was determined by scintillation counting. Background values obtained from controls incubated without peptide substrate were subtracted from each experimental value. Products from the above mentioned incubations using ppGaNTase-T1 cell culture media were recovered by using Sep-Pak C18 reverse-phase cartridges (Waters Corp., Milford, MA) as described previously (23). The products of the reaction by ppGaNTase-T1 were used as substrates (in place of MUC5AC parent peptide) in subsequent incubations with ppGaNTase-T6 and mock (pIMKF1) media. For this second step of N-acetylgalactosaminylation, the conditions were identical to those described above, except that 1 nM UDP-[14C]GalNAc (54.7 mCi mmol\(^{-1}\); 2.02 GBq mmol\(^{-1}\); 0.02 mCi ml\(^{-1}\)) replaced 1.25 nM UDP-[3H]GalNAc. Reactions were performed for 24 h at 37 °C and were stopped as described above.

Reaction products from the aforementioned enzyme assays were analyzed by mass spectrometry and/or capillary electrophoresis. To desalt samples prior to capillary electrophoresis and/or mass spectrometry, Sep-Pak C18 cartridges were used as described previously (23). Matrix-assisted laser desorption ionization mass spectrometry was performed using a Vision 2000 time-of-flight instrument (Finnigan MAT, Bremen, Germany) equipped with a 337 nm UV laser. The mass spectra were acquired in reflectron mode under 6 kV acceleration voltage and positive detection. The samples were prepared by mixing directly onto the target 1 ml of analyzed solution (typically 50 pmol) and 1 ml of a 2.5-dihydroxybenzoic acid matrix solution (12 mg ml\(^{-1}\) in CH\(_3\)OH/H\(_2\)O, 70:30, v/v) and then allowed to crystallize at room temperature. External calibration was performed using the MUC5AC peptide (\(M_r\) 1502.7). 10–30 shots were accumulated for the mass spectrum.

Capillary electrophoresis was performed on a P/ACE system model 5000 (Beckman, Fullerton, CA) under conditions previously described (23). For the separation of the hexadecapeptides, 2 N formic acid buffer with 2.5% polyvinyl alcohol (\(M_r\) 15,000) (v/v) was used. To determine O-linkage sites, a preparative scale procedure was performed as described by Bielker and Schwartz (25), and the recovered fractions

Table I

|        | rT1 | hT2 | mT3 | mT4 | rT5 | rT6 |
|--------|-----|-----|-----|-----|-----|-----|
| rT1    | NA  | 70  | 71  | 70  | 68  | 61  |
| hT2    | 70  | NA  | 66  | 68  | 65  | 59  |
| mT3    | 71  | 66  | NA  | 69  | 64  | 57  |
| mT4    | 70  | 68  | 69  | NA  | 65  | 59  |
| rT5    | 68  | 65  | 64  | 65  | NA  | 57  |
| rT6    | 61  | 59  | 57  | 59  | 57  | NA  |

![FIG. 2. Amino acid sequence alignments of ppGaNTase-T1, -T2, -T3, -T4, -T5, and -T6 from human, murine, and rat clones.](image-url)
were then analyzed by Edman degradation using an Applied Biosys-
tems gas-phase sequencer, model 477A, as described previously (26).

**Peridate Oxidation, Sodium Borohydride Reduction, and Enzyme Assays**—Large quantities of glycosylated MUC5AC were prepared by incubation with *Pichia pastoris*-derived recombinant ppGaNTase-T1. 2 Briefly, the ppGaNTase-T1 coding segment from pPIC-mT1 (27) was inserted into the EcoRI site of a modified *Pichia* expression vector, pPIC (Invitrogen). *Pichia* containing this vector were grown and expression was induced according to the manufacturer’s instructions (Invitrogen). Recombinant ppGaNTase-T1 was purified as described previously (4). *Pichia*-derived recombinant ppGaNTase-T1 and ppGaNTase-T1 expressed from COST cells displayed similar substrate specificities and kinetic parameters.* 2 *Pichia*-derived ppGaNTase-T1 (0.028 μg) was in-
cubated with 1 mg of MUC5AC under the conditions described above, using UDP-[3H]GalNAc (7.8 Ci mmol–1; 288.6 GBqmmol–1; 0.1 mCrmol–1) at a final concentration of 128.2 μM and cold UDP-GalNAc at a final concentration of 30 mM. Reaction products were passed through an AG1-X8 column, and incorporation was determined by scintillation counting. The reaction products were isolated on a Waters 265 HPLC using a Vydac C-18 reverse phase column (0.46 × 25 cm) with a flow rate of 1 ml/min using a linear gradient of 5% acetonitrile, 0.1% triflu-
oroacetic acid to 20% acetonitrile, 0.1% trifluoroacetic acid for 20 min at 22 °C. Mass spectrometry was performed on purified products at the Louisiana State University Mass Spectrometry Facility using extraction in reflector mode on a Bruker (Billerica, MA) ProFLEX III MALDI-TOF mass spectrometer. The matrix used was 2-cyano-4-
hydroxydicyanamide. Two-point calibrations were performed using peptides that have masses above and below the range of the masses of our samples. The most abundant product recovered (the tri-glycopeptide) was subjected to Edman degradation as described previously (28) using a PE Applied Biosystems 473A protein sequencer (Foster City, CA).

This purified MUC5AC tri-glycopeptide (100 nmol) and the MUC5AC parent peptide (100 nmol) were oxidized with 20 μl of 0.08 M NaIO₄ in 0.05 M acetic acid buffer (PH 4.5) at 4 °C for 60 h in the dark (29) 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 N NaOH. Sodium borohydride was added to a final concentration of 0.2 M and reduction continued for 24 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 in the previous paragraph. Capillary electrophoresis was performed on an Applied Biosys-
tems 270A-HT capillary electrophoresis system using 2 N formic acid, 2.5% polyvinyl alcohol (w/v) and a fused silica capillary column (0.75 μm inner diameter) with 50 cm to the optical path and running voltage of 15 kV.

Periodate-treated MUC5AC and the MUC5AC tri-glycopeptide as well as untreated MUC5AC and MUC5AC tri-glycopeptide were then used as substrates in reactions with COST cell-derived ppGaNTase-T1, ppGaNTase-T6, or mock media. Reactions were carried out in duplicate as described above with the following modifications: 15 μg of each peptide/glycopeptide substrate were used in each reaction; the final concen-
tration of UDP-[14C]GalNAc (54.7 mCi mmol–1; 2.02 Gbqmmol–1; 0.02 mCrmol–1) was 21.78 mM and the final concentration of cold UDP-GalNAc was 0.96 mM; the final reaction volume was 50 μl. Reaction products were passed through AG1-X8 resin; incorporation was determined by scintillation counting.

**RESULTS**

cDNA Cloning and Sequence Analysis of ppGaNTase-T6—Primary sequence alignments of previously identified members of the ppGaNTase family revealed many conserved regions within an approximately 420-amino acid segment of the proteins (9). We designed degenerate PCR primers to short blocks of highly conserved sequences, EIWGGHEN and YWDMEYK. PCR was performed on cDNA from rat SLG, and clones were purified and sequenced to identify the nature of the insert as described previously (10). From this screening, a novel PCR product was identified that shared homology with previously described isoforms. The insert from this clone was used as a probe to screen a rat sublingual gland cDNA library. A cDNA containing a complete open reading frame was sequenced and given the designation, ppGaNTase-T6.

As shown in Fig. 1, the cDNA encoding ppGaNTase-T6 con-
tains a 2228-base pair insert encoding a unique 657-amino acid protein. No upstream termination codon or Kozak sequence was found. Conceptual translation of this cDNA revealed a type
II membrane protein architecture, typical of the ppGaNTase family. The enzyme consists of a potentially short N-terminal cytoplasmic region, a 27-amino acid hydrophobic region, a 147-amino acid stem region, and a 483-amino acid lumenal region. As shown in Fig. 2, ppGaNTase-T6 is distinct from previously identified mammalian isoforms yet shares many blocks of sequence similarity or identity between consensus amino acid 174 and 657. Table I summarizes the degree of amino acid similarity between each of the known isoforms within the conserved lumenal region; ppGaNTase-T6 has the lowest similarity when compared with the other isoforms.

**Northern Blot Analysis**—Northern blots of mouse and rat total RNA were probed with a ppGaNTase-T6-specific probe (Fig. 3) as well as probes specific for previously characterized isoforms. The highest levels of the 4.8-kilobase ppGaNTase-T6 message were found in the SLG, with lower levels seen in stomach, small intestine, and colon of both rat and mouse. Trace amounts were detectable in ovary, cervix, and uterus. As reported previously, ppGaNTase-T4 transcripts were found in the digestive and reproductive tracts as well as other tissues.

**Functional Expression**—The truncated coding region of ppGaNTase-T6 was cloned downstream of the insulin secretion signal, HMK site, and FLAG\textsuperscript{TM} epitope tag in the vector pMKF1 (9). The ppGaNTase-T6 truncation began at amino acid position 38. The ppGaNTase-T6 expression construct as well as a similar construct containing ppGaNTase-T1 were independently transfected into COS7 cells. The expressed products from these transfections were secreted into the culture media. Initially, equivalent amounts of each isoform (as judged by densitometric scanning of Tricine SDS-PAGE gels) (data not shown) were used for \textit{in vitro} glycosylation assays against a panel of peptides (10). Although ppGaNTase-T1 glycosylated a number of peptide substrates, \textit{in vitro} glycosylation activity was seen for ppGaNTase-T6 (data not shown). Further assays were then conducted using the MUC5AC peptide and cell culture media from cells transfected with either ppGaNTase-T1 or ppGaNTase-T6. When the MUC5AC peptide was incubated with media from ppGaNTase-T6-transfected cells, capillary electrophoresis revealed only a single peak that displayed a mass ($m/z = 1525.2 \ [M^+ + Na^+]^+$) corresponding to the parent peptide (Fig. 4A). Incubation of the same peptide with ppGaNTase-T1 resulted in incorporation of GalNAc into the peptide fraction (258.3 nmol of GalNAc/h/unit of recombinant ppGaNTase-T1, where a unit of ppGaNTase-T1 is defined as an arbitrary amount of ppGaNTase-T1 normalized to ppGaNTase-T6 after gel quantitation as described under “Experimental Procedures”). The capillary electrophoresis profile revealed the formation of one major peak (36% of initial peptide substrate) (peak 1, Fig. 4B) and two minor species (6.0 and 3.9% of initial peptide substrate, respectively) (peaks 2 and 3, Fig. 4B). Thus, the total of glycosylated products represented 45.9% of the initial peptide presented to the enzyme. MALDI-MS confirmed that peaks 1–3 consisted of mono- ($m/z = 1728.5, i.e. 203$ greater than the parent peptide), di- ($m/z = 1931.5$), and tri-substituted ($m/z = 2134.4$) glycopeptides, respectively. Direct sequence analysis revealed that threonines 3 and 13 were substituted with GalNAc in the major purified fraction, obtained by capillary electrophoresis at preparative scale (peak 2,
TABLE II

ppGaNTase-T6 activity on treated and untreated MUC5AC and MUC5AC tri-glycopeptide

| Enzyme source       | Amount of [14C]GalNAc incorporated into untreated MUC5AC tri-glycopeptide | Amount of [14C]GalNAc incorporated into treated MUC5AC tri-glycopeptide | Amount of [14C]GalNAc incorporated into untreated MUC5AC | Amount of [14C]GalNAc incorporated into treated MUC5AC |
|---------------------|---------------------------------------------------------------------------|-----------------------------------------------------------------------|---------------------------------------------------------|---------------------------------------------------------|
| pMKP1^b             | 565 ± 25                                                                   | 342 ± 40                                                              | 356 ± 6                                                 | 357 ± 1                                                  |
| ppGaNTase-T1        | 368 ± 2                                                                   | 604 ± 56                                                              | 5539 ± 11                                               | 5038 ± 152                                               |
| ppGaNTase-T6        | 3960 ± 120                                                                 | 648 ± 4                                                               | 350 ± 54                                                | 346 ± 11                                                 |

Values shown are the average and standard deviation generated from two reactions.

* Treatment refers to periodate oxidation and sodium borohydride reduction.

Parent vector representing the negative control.

As an initial step in defining the requirement of ppGaNTase-T1 and ppGaNTase-T6, we employed in combination, an increase in the level of GalNAc incorporation into the MUC5AC peptide was observed over that obtained with ppGaNTase-T1 alone (294.5 nmol of GalNAc/h/unit of recombinant ppGaNTase-T1 and -T6); 49% of the initial peptide presented was distributed in five discrete fractions resolved by capillary electrophoresis (peaks 1–5; 17.1, 10.0, 11.8, 7.4, and 2.7% of initial peptide presented, respectively) (Fig. 4C). Analysis of the products by MALDI-MS indicated that they corresponded to glycopeptides that were substituted with one to five residues of GalNAc (m/z = 1728.5, 1931.5, 2134.4, 2337.6, and 2540.5, respectively). Insufficient amounts of material were present to determine the positions of GalNAc residues in these species. Collectively, these results suggested that ppGaNTase-T6 catalyzes the transfer of GalNAc from UDP-GalNAc to a GalNAc-containing glycopeptide (i.e. a UDP-GalNAc glycopeptide-GaNTase (gpGaNTase)).

To confirm the presence of ppGaNTase activity, the substrate GTTPSPVPTTTSAP was first incubated for 24 h with recombinant ppGaNTase-T1 cell culture media using UDP-[14C]GalNAc as the sugar donor. The reaction products (containing the di-substituted glycopeptide and unmodified parent peptide as well as small amounts of mono- and tri-substituted peptide) were next incubated with ppGaNTase-T6 in the presence of UDP-[14C]GalNAc (40,000 dpm) as the sugar donor. As a control, an equivalent quantity of culture media from mock-transfected (pMKP1) COS7 cell cultures was also used as an "enzyme" source. As expected, little incorporation of [14C]GalNAc was observed when the mock-transfected media was used as the enzyme source (360 dpm; <1% of the initial truncated substrate was labeled with [14C]). In contrast, significant incorporation of [14C]GalNAc was obtained when ppGaNTase-T6 was used (5,960 dpm; 18.6% of initial substrate was [14C]-labeled, corresponding to 283.8 nmol of GalNAc/h/unit). Fig. 5 compares the products generated after the second incubation with mock-transfected supernatant and ppGaNTase-T6. In contrast to the products obtained after incubation with media from the mock-transfected control, ppGaNTase-T6 yielded 3 additional glycopeptide fractions with longer retention times; fractions 4–6 correspond to glycopeptides substituted with four to six residues of GalNAc (m/z = 2337.6, 2540.5, and 2743.6). The relative level of the di-substituted glycopeptide (peak 2) present after ppGaNTase-T6 incubation versus mock incubation was much reduced (2.7% of total profile versus 32.7%, respectively), suggesting that it had been converted to the more heavily glycosylated species (peaks 3–6), whereas the quantity of the parent peptide (4.6% of total profile) and the mono-substituted glycopeptide (peak 1) (2.7% of total profile) remained unchanged (Fig. 5).

As an initial step in defining the requirement of the ppGaNTase-T6 isoform for a GalNAc-containing substrate, we modified GalNAc residues by periodate oxidation and sodium borohydride reduction. To obtain sufficient amounts of glycosylated MUC5AC, we used the P. pastoris expression system to generate large quantities of recombinant ppGaNTase-T1. The ppGaNTase-T1 coding segment used in the COS7 cell expression system was cloned into a Pichia expression vector (pPIC, Invitrogen) and was expressed under methanol induction conditions. Approximately 500 µg/liter ppGaNTase-T1 was purified as described (4) and a portion of that was incubated with MUC5AC and UDP-[3H]GalNAc as outlined under "Experimental Procedures." This incubation resulted in the production of 4 glycopeptide fractions (1887.6 nmol of GalNAc/µg of ppGaNTase-T1 incorporated), corresponding to mono- (m/z = 1729.5) (21.7% of initial peptide presented), di- (m/z = 1930.2) (28.7% of initial peptide presented), tri- (m/z = 2133.1) (35.3% of initial peptide presented), and tetra-substituted (m/z = 2336.1) (1.3% of initial peptide presented) glycopeptides, respectively. The most abundant peak recovered after HPLC purification of all reaction products was that representing the tri-glycopeptide, as determined by mass spectrometry (m/z = 2133.1). Edman degradation of this species revealed that a GalNAc residue was present at serine 5 and, like the di-substituted species generated by COS7 cell-derived ppGaNTase-T1, at threonines 3 and 13.

This purified tri-glycopeptide along with the MUC5AC parent peptide were subjected to periodate oxidation followed by sodium borohydride reduction. Periodate-treated and untreated tri-glycopeptide and MUC5AC parent peptide were purified by HPLC, analyzed for integrity by capillary electrophoresis (data not shown), and subsequently incubated with COS7 cell-derived ppGaNTase-T1, ppGaNTase-T6, or mock-transfected (pMKP1) media in the presence of UDP-[14C]GalNAc. Table II compares the counts incorporated into each substrate by each enzyme. Treatment of the tri-glycopeptide with periodate and sodium borohydride clearly reduces the ability of ppGaNTase-T6 to use it as a substrate (compare 3960 cpm incorporated into untreated material to 648 cpm incorporated into treated material). This reduction in incorporation is not due to the peptide itself being compromised during periodate treatment because ppGaNTase-T1 works equally well on both treated and untreated MUC5AC (Table II). These data suggest that ppGaNTase-T6 requires the presence of intact GalNAc on the MUC5AC peptide for it to be used as a substrate.

DISCUSSION

Through the use of degenerate PCR, we have cloned a novel isoform of the ppGaNTase family. ppGaNTase-T6 is a type II membrane protein, consisting of a potentially short N-terminal cytoplasmic domain, a transmembrane domain, a stem region, and a luminal domain, characteristic of the other previously identified isoforms. This isoform displays the lowest level of amino acid similarity within the putative catalytic domain among the members of the ppGaNTase family and is the only isoform identified to date that lacks any potential N-glycosylation sites. Unlike previously identified isoforms, ppGaNTase-T6 fails to act on a panel of 8 peptide substrates but...
rather catalyzes the transfer of GalNAc from UDP-GalNAc to a GalNAc-containing peptide substrate. Furthermore, the modification of the GalNAc residues on the glycopeptide substrate by periodate oxidation and sodium borohydride reduction inhibits further incorporation of GalNAc by ppGaNTase-T6. Our data, therefore, indicate that at least two free GalNAc residues must be incorporated (by ppGaNTase-T1) into the MUC5AC peptide before it can be used as a substrate by ppGaNTase-T6; this requirement is specific to the GalNAc structure itself and is not simply satisfied by the presence of the chemical constituents that make up the GalNAc residue. Whether or not there exist strict positional requirements for these GalNAc residues as well as their effect on the site of transfer of subsequent GalNAc residues by another isoform, highlighting a potential hierarchical relationship between members of the ppGaNTase family.

Glycosylation of the MUC5AC peptide by ppGaNTase-T1 also appears to occur in a regulated manner. The di-glycopeptide contains GalNAc residues at threonines 3 and 13; the tri-glycopeptide has an additional GalNAc at serine 5. It will be of interest to determine and compare the $K_m$ values of the threonine positions versus the serine. These data demonstrate there exists a hierarchical addition of GalNAc within the MUC5AC substrate by ppGaNTase-T1.

The pattern of transcript expression for ppGaNTase-T6 is very restricted, being found predominantly in the SLG, with lower levels in the remainder of the digestive tract and female reproductive tract. This distinct expression pattern is conserved across species, between rat and mouse. Thus, ppGaNTase-T6 appears to require the prior addition of GalNAc-containing peptide substrate. Furthermore, the modification of the GalNAc residues by another isoform, highlighting a potential hierarchical relationship between members of the ppGaNTase family.

ACKNOWLEDGMENTS—We thank Meng Qian for help in preparing this manuscript. We are grateful to Dr. G. Ricart, D. Demeyer, and N. Parsy for their technical assistance.

REFERENCES

1. Dowbenko, D., Andalihi, A., Young, P. E., Lusis, A. J., and Lasky, L. A. (1995) J. Biol. Chem. 268, 4525–4529.
2. Kinloch, R. A., Sakai, Y., and Wasserman, P. M. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 283–287.
3. Jenoff, N. (1996) Trends Biochem. Sci. 15, 291–294.
4. Hagen, F. K., VanWyckhuyse, B. C., and Tabak, L. A. (1993) J. Biol. Chem. 268, 18960–18965.
5. Homa, F. L., Hollander, T., Lehman, D. J., Thomsen, D. R., and Elhammer, J. G. (1999) J. Biol. Chem. 274, 21269–21275.
6. Serensen, T., White, T., Wendall, H. H., Kristensen, A. K., Roepstorff, P., and Clausen, H. (1995) J. Biol. Chem. 270, 24166–24173.
7. Bennett, E. P., Hassan, H., and Clausen, H. (1996) J. Biol. Chem. 271, 17006–17012.
8. Zara, J., Hagen, F. K., Ten Hagen, K. G., VanWyckhuyse, B. C., and Tabak, L. A. (1996) Biochem. Biophys. Res. Commun. 228, 38–44.
9. Hagen, F. K., Ten Hagen, K. G., Beres, T. M., Balye, M. M., VanWyckhuyse, B. C., and Tabak, L. A. (1997) J. Biol. Chem. 272, 13843–13848.
10. Ten Hagen, K. G., Hagen, F. K., Balye, M. M., Beres, T. M., VanWyckhuyse, B., and Tabak, L. A. (1998) J. Biol. Chem. 273, 27749–27754.
11. Guyonnet-Duperat, V., Audie, J. P., Debaillie, V., Laine, A., Buisine, M. P., Galiegue-Zouitina, S., Figny, P., Degond, P., Aubert, J. P., and Porchet, N. (1995) Biochem. J. 305, 211–219.
12. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
13. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (eds) (1995) Current Protocols in Molecular Biology, pp. 4.2.4–4.2.8, John Wiley & Sons, Inc., New York.
14. Bednarczuk, T. A., Wiggins, R. C., and Konat, G. W. (1991) BioTechniques 11, 2049–2056.
15. Bennett, E. P., Hassan, H., Mandel, U., Hollingsworth, M. A., Akisawa, N., Ikematsu, Y., Merkx, G., van Kessel, A. G., Oldston, A., and Clausen, H. (1999) J. Biol. Chem. 274, 25692–25700.
16. Schagger, H., and von Jagow, G. (1987) Anal. Biochem. 166, 368–379.
17. Albone, E. F., Hagen, F. K., VanWyckhuyse, B. C., and Tabak, L. A. (1994) J. Biol. Chem. 269, 2440–2446.
18. Gum, J. R., Jr., Hickes, J. W., Toribara, N. W., Siddiki, B., and Kim, Y. S. (1994) J. Biol. Chem. 269, 2440–2446.
19. Gendler, S. J., Lancaster, C. A., Taylor-Papadimitriou, J., Dhugid, N., Peat, N., Burchell, J., Pemberton, L., Lalani, E.-N., and Wilson, D. (1990) J. Biol. Chem. 265, 15296–15297.
20. Ohmori, H., Doehman, A. F., Gallop, M., Tuda, T., Kaj, H., Gum, J. R., Jr., Kim, Y. S., and Bashaumb, C. B. (1994) J. Biol. Chem. 269, 17833–17840.
21. Shekels, L. L., Lytolf, C., Kieliszewski, M., Filipe, J. D., Kozak, C. A., and Ho, Y. S. (1995) Biochem. Biophys. Res. Commun. 215, 775–780.
22. Soudan, B., Tetaert, D., Hennebicq, S., Briand, G., Zerineck, F., Richet, C., Demeyer, D., Gagnon, J., Petilis, Y., and Degand, P. (1998) J. Peptide Res. 52, 346–354.
23. Hennebicq, S., Soudan, B., Tetaert, D., Briand, G., Richet, C., Demeyer, D., and Degand, P. (1996) Carbohydr. Res. 280, 121–126.
24. Bonn, T., Richet, C., Haupt, K., Hennebicq, S., Soudan, B., Tetaert, D., Degand, P., and Vijayalakshmi, M. A. (1998) J. Chromatogr. A. 805, 285–293.
25. Biehler, R., and Schwartz, H. E. (1995) Technical Bulletin TIBB 105, 1–9.
26. Hennebicq, S., Tetaert, D., Soudan, B., Briand, G., Richet, C., Gagnon, J., and Degand, P. (1998) Glycoconj. J. 15, 275–282.
27. Hagen, F. K., and Nehrke, K. (1998) J. Biol. Chem. 273, 8268–8277.
28. Gerken, T. A., Owen, C. L., and Pasumarthy, M. (1997) J. Biol. Chem. 272, 9709–9719.
29. Goldstein, J. I., Hong, G. W., Lewis, B. A., and Smith, F. (1965) Methods Carbohydr. Chem. 3, 361–367.
30. Nielsen, P. A., Bennett, E. P., Wandall, H. H., Therkildsen, M. H., Hanibal, K. G. Ten Hagen and L. A. Tabak, unpublished observations.
31. Troxler, R. F., Iontcheva, I., Oppenheim, F. G., Nunes, D. P., and Offner, G. D. (1997) *Glycobiology* 7, 965–973
32. Desseyn, J.-L., Guyonnet-Duprat, V., Perchet, N., Aubert, J.-P., and Laine, A. (1997) *J. Biol. Chem.* 272, 3168–3178
33. Piller, V., Piller, F., and Fukuda, M. (1996) *J. Biol. Chem.* 271, 9264–9271
34. Nehrke, K., and Tabak, L. A. (1997) *Biochem. J.* 323, 497–502
35. Bennett, E. P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M. A., Merkl, G., van Kessel, A. G., Elberg, H., Steffensen, R., and Clausen, H. (1998) *J. Biol. Chem.* 273, 30472–30481
36. Röttger, S., White, J., Wandall, H. H., Olivo, J.-C., Stark, A., Bennett, E. P., Whitsu, C., Berger, E. G., Clausen, H., and Nilsson, T. (1998) *J. Cell Sci.* 111, 45–60