Characterization of a UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase from *Drosophila*

Naosuke NAKAMURA

(*Received September 12, 2005*)

Abstract

UDP-N-acetylgalactosamine: polypeptide N-acetylgalactosaminyltransferases (GalNAc-transferases) catalyze the initial reaction of mucin-type O-glycosylation. Here, I report biochemical characterization of one of the *Drosophila* GalNAc-transferases, pgant3. This enzyme retains conserved motifs essential for the catalytic activity, but is a novel isozyme in that it has several inserted sequences in its lectin-like domain. Northern hybridization analysis of this isozyme identified a 2.5-kb mRNA in *Drosophila* larva. Biochemical characterization was carried out using the recombinant soluble pgant3 expressed in COS7 cells. pgant3, which required Mn^{2+} for the activity, had a pH optimum ranging from pH 7.5 to 8.5, and glycosylated most effectively at 29–33°C. Its K_m for UDP-GalNAc was 10.7 μM, which is as low as that of mammalian isozymes. pgant3 glycosylated the peptides containing a sequence of XTPXP or TTAAP most efficiently. The enzyme was irreversibly inhibited by *p*-chloromercuriphenylsulphonic acid, indicating the presence of essential Cys residues for the activity, such as that of mammalian GalNAc-T1. Besides pgant3, I found several *Drosophila* cDNA clones homologous to the cloned GalNAc-transferases through the database search. *Drosophila* contains at least twelve putative GalNAc-transferase genes, suggesting that the GalNAc-transferase family comprises a large gene family as is found in mammals. These findings demonstrated that the GalNAc-transferases play an important role in controling the biosynthesis of mucin-type carbohydrates in insects as well as in mammals.

Keywords; GalNAc-transferase, mucin, *Drosophila*, O-glycosylation, insect, peptide

Contents

1. Introduction

Post-translational modifications are important process for maturation of protein in the cell, and glycosylation is one of the most common modifications of proteins. Mucin-type O-glycosylation is widely distributed in many secretory and membrane glycoproteins, and is involved in important biological functions such as cell-cell recognition, host-pathogen interaction, and protection of proteins from proteolytic degradation [1,2]. A UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase (GalNAc-transferase) (EC. 2.4.1.41) catalyzes the initial step of biosynthetic reaction of mucin-type carbohydrate by transferring GalNAc from UDP-GalNAc to a hydroxyl amino acid on a polypeptide acceptor [3]. It is one of the most important enzymes for the biosynthesis of mucin-type carbohydrates, since it determines the number and the positions of O-linked sugar chains in a protein. Recent progress in the molecular cloning of GalNAc-transferases has revealed a large gene family in mammals, nematodes, and insects [4-23]. The evolutionary conservation of GalNAc-transferases stresses the biological importance of the mucin carbohydrates.

Database search of the *Drosophila* genome indicated the presence of at least 14 GalNAc-transferases, and nine of these isozymes cloned have been shown to encode functional enzymes [17,24]. Moreover, the developmental importance of the GalNAc-transferases in *Drosophila* is demonstrated by the lethal mutations in the l(2)35Aa allele coding for dGalNAc-T1 /pgant35 [17,23]. However, *Drosophila* GalNAc-transferases have not been biochemically characterized in detail so far. Thus, their activity has been determined under the same conditions as the mammalian enzymes. Here, the author reports the detailed biochemical properties of one of the *Drosophila* GalNAc-transferases,

which is referred to as pgant3 by Kelly TenHagen et al [24]. They demonstrated that pgant3 has a relatively narrow substrate specificity, exclusively glycosylating a peptide derived from EA2 among the seven peptides tested. It was also reported that this isozyme has very poor glycopeptide glycosylation activity. In this thesis, I demonstrate a more extensive analysis of the substrate specificity of pgant3, in which synthetic peptides from insect mucins as well as mammalian ones were used as acceptors to assess the possibility of pgant3 glycosylating mucin polypeptides in insects.

2. Materials and Methods

2.1 Materials

An EST clone GH09147 (GadFly annotation CG4445), encoding for pgant3, was purchased from Research Genetics, Inc. A panel of peptides was synthesized using a Multipin Peptide Synthesis Kit (Chiron Mimotopes).

2.2 Northern Blot Analysis

Drosophila larva poly A^+ RNA (Clontech) was electrophoresed in the denaturing condition in a 1.0% agarose-formaldehyde gel with 20 mM 4-morpholinepropanesulfonic acid/5 mM sodium acetate buffer, pH 7.0. Separated RNA was then transferred from a gel to a positively charged nylon membrane (Roche) with 0.05M NaOH. The filter membrane was briefly washed with 2 X standard saline citrate (SSC), and then baked for 20 min at 50°C to immobilize RNA on the membrane. Following prehybridization in the hybridization buffer (0.2M sodium phosphate buffer, pH 7.5, containing 0.1% sodium dihydrogen pyrophosphate, 2 mM ethylenediamine tetraacetic acid, and 7% sodium dodecylsulfate (SDS)) for 8 h at 68°C, the membrane was hybridized overnight at 68°C with the ^{32}P-labeled pgant3-specific probe, which was prepared by polymerase chain reaction (PCR) with *pgant3* cDNA as template and a primer pair, DmG-PS (5'-ATGGGTCTCCGGTTCCAGCA-3') and DmG-PA (5'-CCTTGAGGGTGCGGTTCAGT-3'), designed for amplification of nucleic acid residue 1-600 bp, in the presence of [α-^{32}P] TTP, 3000Ci/mmol. The membrane was then washed in 2 X SSC, containing 0.1% SDS, and subsequently in 0.1 X SSC, containing 0.1% SDS, 3 times for 10 min at 68°C, respectively.

2.3 Expression of recombinant soluble pgant3 in COS7 cells

For the construction of soluble pgant3, a cDNA fragment, deleted with the sequences for putative cytoplasmic tail, the transmembrane domain, and a portion of the stem region, was amplified by PCR using *pgant3* cDNA as a template and a primer set of 5'-GAGATGCGATATCCGAAAC

CG-3' and 5'-AATGTCCGGGCCCCGGCATGA-3'. The nucleotides in the primers represented in boldfaces are mutated to introduce *Eco*RV and *Apa*I restriction sites (underlined) in the amplified cDNA, respectively. Following the amplification, the PCR product was digested with *Eco*RV and *Apa*I, and purified by a QIAquick Gel Extraction Kit (Qiagen). The digest was ligated into the *Bam*HI (blunt-ended) and the *Xho*I sites of a mammalian expression vector, *pInsProA*, using Ligation High (Toyobo). *pInsProA* contains a cDNA fragment encoding for the insulin signal sequence and the Protein A-IgG binding domain inserted into *pcDNA3.1* (Invitrogen) [25]. The generated construct coding for the truncated pgant3 fused with the signal sequence and the IgG-binding domain of Protein A was introduced into *E. coli* strain DH5α (Toyobo). Colonies were screened by PCR, and the plasmid (*pIP-ΔdGNT3*) containing the truncated *pgant3* was obtained using the QIAprep Miniprep (Qiagen). *pIP-ΔdGNT3* was then transfected into COS7 cells using FuGENE6 (Roche). Three days after transfection, the secreted recombinant enzyme in the conditioned medium was purified as described below.

2.4 Purification and Western blot analysis of recombinant pgant3

The secreted enzyme in the culture medium was purified on IgG-Sepharose (Amersham Pharmacia Biotech) as follows: Five hundred microliters of the conditioned medium were mixed with 10 μl of IgG-Sepharose and incubated under constant rocking at 4°C for 2 h. The IgG-Sepharose adsorbed with the recombinant pgant3 was washed three times with 25 mM imidazole buffer (pH 7.2) containing 300 mM NaCl and 1% Triton X-100, and three times with 25 mM imidazole buffer (pH 7.2) containing 100 mM NaCl. The adsorbed pgant3 on the resins was used for the enzyme assay as described below. For quantitation of the recombinant enzyme, the washed resins were suspended in the SDS-polyacrylamide gel electrophoresis loading buffer, boiled at 100°C for 3 min, and centrifuged at 12,000 rpm for 10 min. The resulting supernatant was loaded on the gel. The electrophoresed proteins in the gel were transferred to a polyvinilidene difluoride membrane (Bio-Rad), and visualized by incubating the blot with an alkaline phosphatase (AP)-conjugated rabbit antibody to mouse IgG (Cappel), followed by staining with nitrobluetetrazolium and 5-bromo-4-chloro-3-indolylphosphate. Intensity of the bands on the immunoblots was determined using the LAS-1000PLUS (Fuji) to evaluate the concentration of recombinant pgant3 in the medium.

2.5 Enzyme Assay

The activity of pgant3 was determined in a reaction mixture composed of 50 mM imidazole buffer (pH 7.2), 10 mM $MnCl_2$, 0.5 nmol UDP-[1-^{3}H] GalNAc (approximately 40,000 dpm), 5 nmol synthetic peptides, and an appropriate amount of enzyme, in a total volume of 40 μl. The mixture

was incubated at 25°C for 16 h under constant stirring to maintain a uniform suspension. It was confirmed that the progress of the reaction was linear with respect to the incubation time. Following the incubation, the reaction was stopped by adding 10 μl of 0.25 M EDTA. The reaction mixture was then applied to an AG1 X-8 anion exchange column (0.5 ml, Cl^- form). The pass-through fraction was recovered and the radioactivity was determined. K_m for UDP-GalNAc was obtained by increasing concentrations of UDP-GalNAc from 5 to 400 μM using 5 nmol of MUC2-1 as acceptor. Kinetic parameters were obtained by the double reciprocal plot (1/v versus 1/[S]), using standard procedures.

2.6 Modification of pgant3 with *p*-chloromercuriphenylsulphonic acid (PCMPS)

Recombinant pgant3 expressed in COS7 cells and adsorbed to the IgG-Sepharose was treated with increasing concentrations (0-0.5 mM) of PCMPS (Sigma) in 0.1 M imidazole buffer (pH 7.2), for 90 min at 25°C. After incubation with PCMPS, the enzyme adsorbed to the IgG-Sepharose was washed and its activity was determined as described above.

2.7 Phylogenetic reconstruction of the GalNAc-transferase family

Multiple alignment of the GalNAc-transferases from *Drosophila* and mammals was carried out by comparing the homology of the putative catalytic domain, which corresponds to 84-418 amino acid residues in human GalNAc-T1, using the Clustal X algorithm of MacVector (version 6.5). Distance-based phylogenic trees were constructed by an NJ image method of PAUP* version 4.0 [26].

3. Results and Discussion

3.1 Characterization of primary structure of pgant3

Comparison of the amino acid sequences of pgant3 with mammalian GalNAc-transferase cloned to date indicated that pgant3 is a putative type II membrane protein with seven potential N-glycosylation sites, and consists of a 10-amino acid N-terminal cytoplasmic region, a 23-amino acid transmembrane region, an 84-amino acid stem region, a 345-amino acid putative catalytic region, and a 154-amino acid lectin domain (Fig. 1). It contains a glycosyltransferase1 motif, which is a conserved sequence commonly found in a wide variety of glycosyltransferases including the GalNAc-transferases, and a DXH sequence, which is a putative binding site for a sugar donor and/or a metal ion [27-30]. A Gal/GalNAc-transferase motif is also found in the putative catalytic domain of this isozyme. It is a homologous sequence found in Gal- and GalNAc-transferases, which may be a

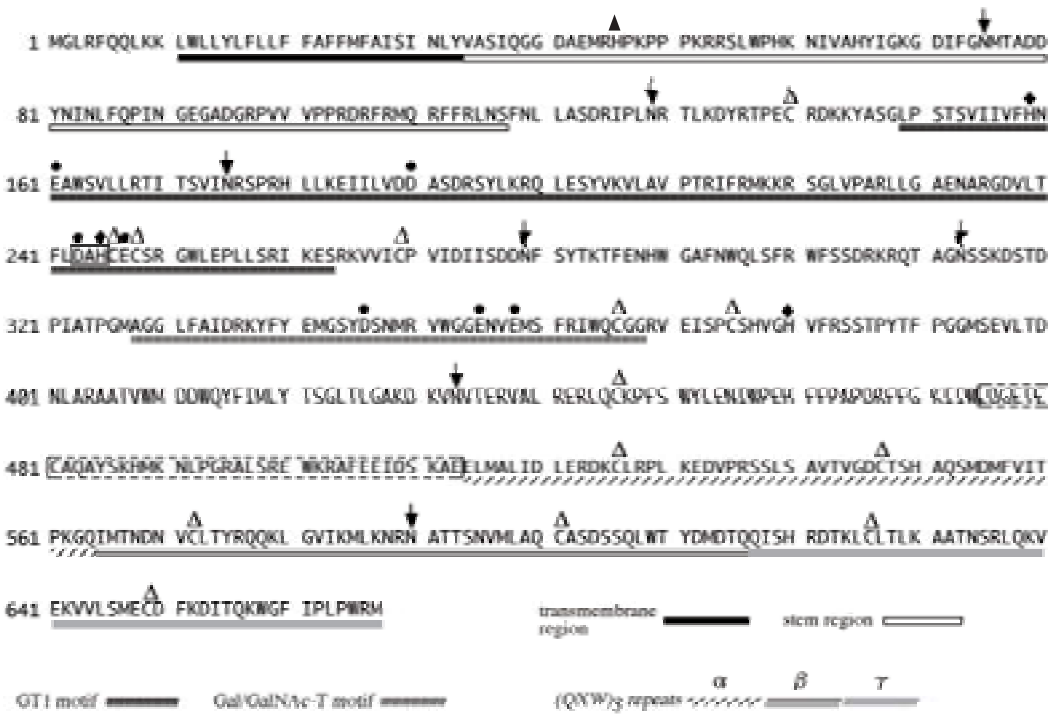

Fig. 1 Amino acid sequence of pgant3

Potential N-glycosylation sites are indicated by arrows. The DXH sequence in the glycosyltransferase 1 motif and the insert sequences are outlined with solid and dashed lines, respectively. Conserved acidic, histidine, and cysteine residues are indicated by •, ◆, and Δ, respectively. ▲ represents the N-terminus of the truncated pgant3 generated for the enzyme assay.

binding site for both a UDP-sugar and an acceptor substrate [31]. At the C-terminus, there is a lectin-like domain called the $(QXW)_3$ repeats, which consists of three tandemly repeated sequences (α, β, and γ repeats), and occurs only in the GalNAc-transferases among the glycosyltransferase family [32]. Furthermore, this isozyme contains several amino acid residues essential for catalytic activity as demonstrated by mutational studies of mammalian GalNAc-transferases. It has seven conserved acidic amino acids that are involved in the enzyme activity of human GalNAc-T1 [31]. It also contains four essential His residues [33]. In addition, it has several Cys residues conserved in the family, most of which are likely involved in forming disulfide bonds necessary for the proper folding of the enzymes. Two Cys residues among them, which are located at the C-terminus of the DXH motif, are also present in this isozyme. It is reported that these are reduced Cys residues, and are involved in binding with UDP-GalNAc [25]. Interestingly, pgant3 contains a 39 amino acid residue insertion at the N-terminus of the $(QXW)_3$ α repeat, which is not found in the other GalNAc-transferases. Moreover, it also has a short insertion (approximately 5-10 amino acid residues) in each

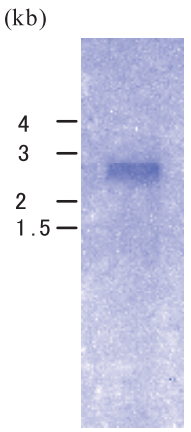

Fig. 2 Northern hybridization analysis

One microgram of *Drosophila* larva poly (A)$^+$RNA was separated by agarose gel electrophoresis and blotted onto a nylon membrane. The blot was probed with ^{32}P-labeled pgant3 cDNA.

repeat of the lectin-like domain. This raises the possibility that this isozyme has altered lectin activity. It is reported that the GalNAc-transferases have distinct manners of substrate recognition [34-36] ; one is based on protein-protein interaction, so-called the initial activity, and is involved in the glycosylation of non-glycosylated peptides, and the other requires the prior addition of GalNAc to peptide substrates to transfer more GalNAc residues to the substrates. The latter activity, called follow-up activity, is based on the recognition of GalNAc residues on the acceptor substrate by the lectin domain of the enzyme. Mutational studies on the rat GalNAc-T1 lectin domain identified an asparagine residue (Asp444) in the α repeat predominantly involved in the lectin activity [35]. In pgant3, this residue is replaced with arginine. Thus, together with the aberrant insertions, the lectin domain of this isozyme may not be fully functional. In fact, the glycopeptide glycosylation activity of this isozyme is reported to be very low [24].

3.2 Northern blot analysis

I examined the expression of pgant3 in *Drosophila* by Northern blot hybridization. A blot containing 1 μg of *Drosophila* larva poly (A)$^{2+}$ RNA was probed with a ^{32}P-labeled 600-bp cDNA fragment under stringent conditions. A single 2.5-kb mRNA was observed, confirming the expression of the pgant3 transcripts in the larva (Fig. 2).

3.3 Expression and purification of recombinant pgant3

For the characterization of pgant3, soluble recombinant pgant3 was expressed in COS7 cells. *pgant3* cDNA was deleted with the coding region for the cytoplasmic tail and the transmembrane

region, and fused with cDNA for an insulin signal sequence and a Protein A-IgG binding domain to the resulting 5'-end of the truncated *pgant3* cDNA. Following the subcloning of the resulting cDNA into the mammalian expression vector, the truncated *pgant3* cDNA was transfected into COS7 cells, and the recombinant enzyme was expressed as the secreted fusion protein, which consisted of Protein A-IgG binding domain fused to N-terminus of the truncated pgant3. The recombinant secreted protein was purified from culture medium with IgG-Sepharose. Western blot analysis of the purified protein demonstrated the expression of a polypeptide with an expected molecular weight of ~90 kDa. This recombinant soluble pgant3 was used for the following experiments.

3.4 Biochemical characterization of pgant3

I first examined the optimal temperature of pgant3. Reaction was performed at pH 7.2 at various temperature ranging from 9°C to 45°C with synthetic undecpeptide MUC2-1, which contains O-glycosylation sites of human MUC2 and is one of the most efficient acceptors of pgant3 as shown below. pgant3 exhibited a relatively broad temperature specificity, although it worked most efficiently at temperatures between 29°C and 33°C (Fig. 3A). The optimal temperature of the insect isozyme was significantly lower than that of mammalian GalNAc-transferases (40-45°C) [21]. Interestingly, the *C. elegans* GalNAc-transferases have an even lower optimal temperature (20-25°C) [21]. This indicates that the optimal temperature of enzymes is closely related to the environmental temperature where they inhabit. For the following experiments, the assay was performed at 25°C. It is a suboptimal temperature, but employed because the flies were kept at this temperature. I then investigated the optimal pH. Reaction was performed in the pH range from 5.4 to 9.0 (Fig. 3B). The enzyme preferred the mild alkaline conditions for the efficient catalysis, with an optimal pH of 7.5-8.5. Bovine GalNAc-T1 is similarly reported to have a mild alkaline optimal pH of 7.2-8.6 [37]. The following enzyme assay was, however, carried out at pH 7.2, since the GalNAc-transferases are supposed to reside in the ER and/or the cis-Golgi. pgant3 requires divalent cations for the activity, since its activity was lost upon the addition of EDTA in the reaction mixture. Among the divalent cations tested, Mn^{2+} most effectively enhanced the enzyme activity (Fig. 4). The activity in the presence of either Co^{2+} or Ni^{2+} was 60% and 20% of that of Mn^{2+}, respectively. Neither Ca^{2+}, Cu^{2+}, Mg^{2+}, nor Zn^{2+} was effective. pgant3 exhibited a similar divalent cation requirement as that of the mammalian GalNAc-transferases [37]. Taking these observations together, the activity of pgant3 was determined in the following experiments at 25°C with the pH 7.2 buffer containing 10 mM $MnCl_2$. I then investigated the affinity of UDP-GalNAc with the enzyme (Fig. 5). The K_m value (10.7 μM) of pgant3 for UDP-GalNAc is as low as those of the mammalian GalNAc-transferases [38].

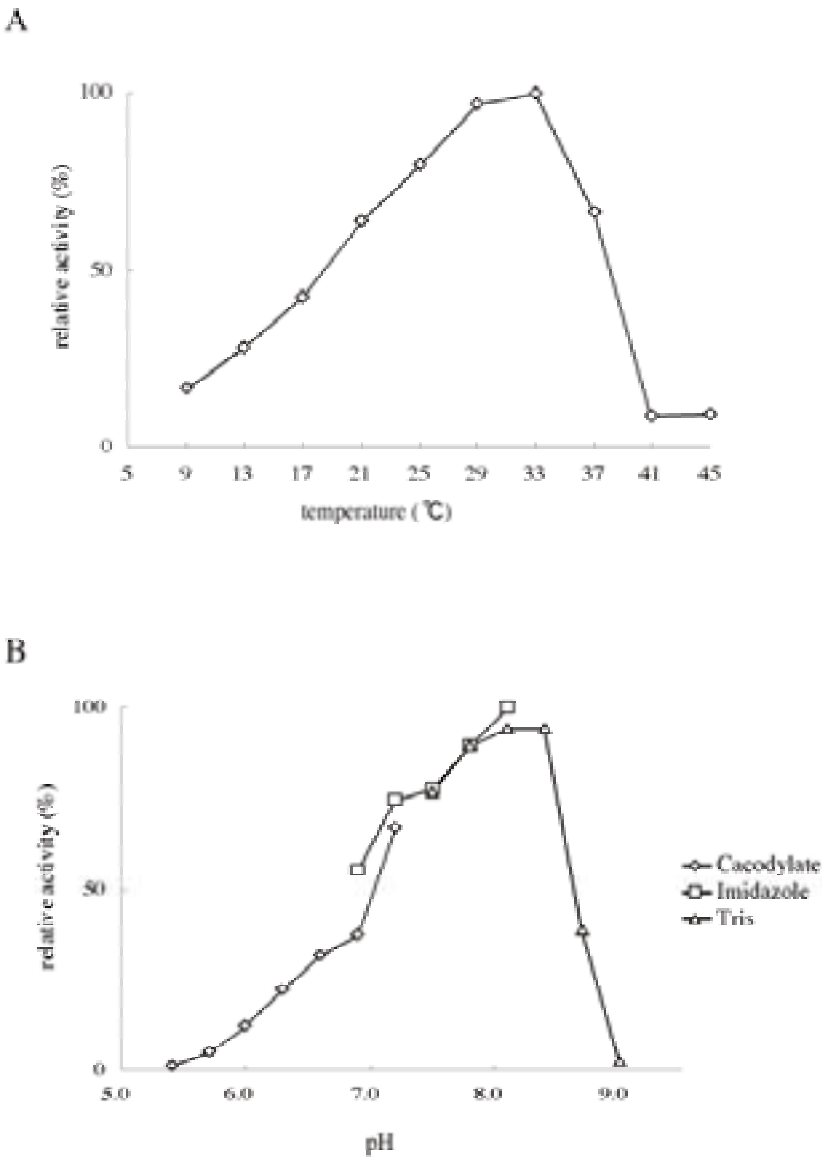

Fig. 3 (A) Optimal temperature

Reaction was carried out at the temperature between 9°C and 45°C.

(B) Optimal pH

Reaction was performed in 20 mM of the three different buffers. pH 5.4-7.2, cacodylate buffer; pH 6.9-8.1, imidazole buffer; pH 7.5-9.0, Tris buffer.

3.5 Modification of pgant3 with PCMPS

The effect of Cysteine modification of pgant3 was investigated. Tenno *et al.* reported that the modification of mammalian GalNAc-T1 with a Cys-specific reagent, PCMPS, irreversibly inhibited the enzyme with a K_i of 0.03 mM and that Cys212 and Cys214, which are C-terminal to the DXH

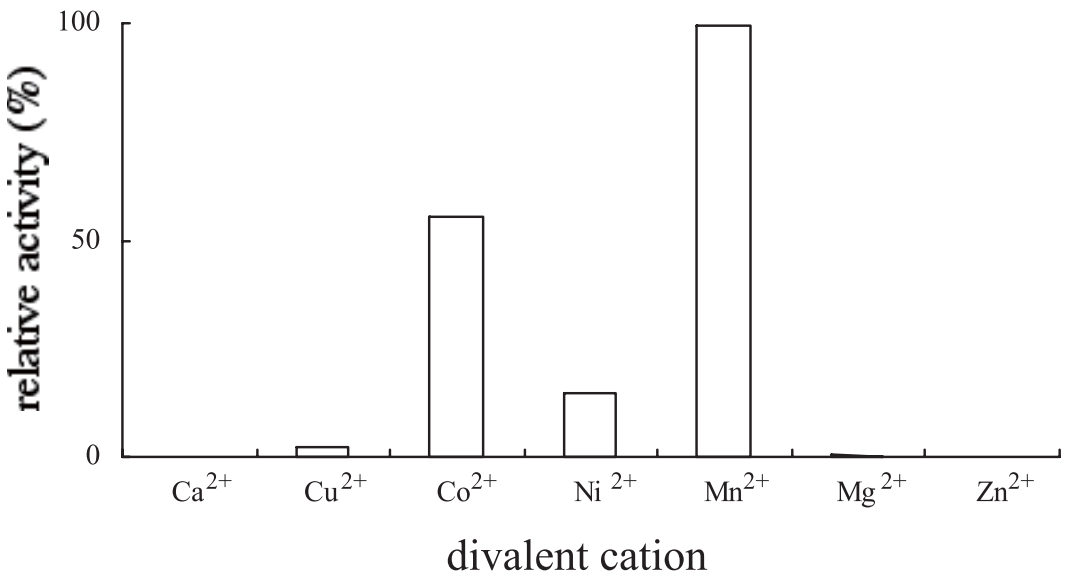

Fig. 4 Divalent cation requirement

Mixture containing 10 mM of each divalent cation was used. The activity was represented as a relative activity to that obtained in the presence of Mn^{2+}.

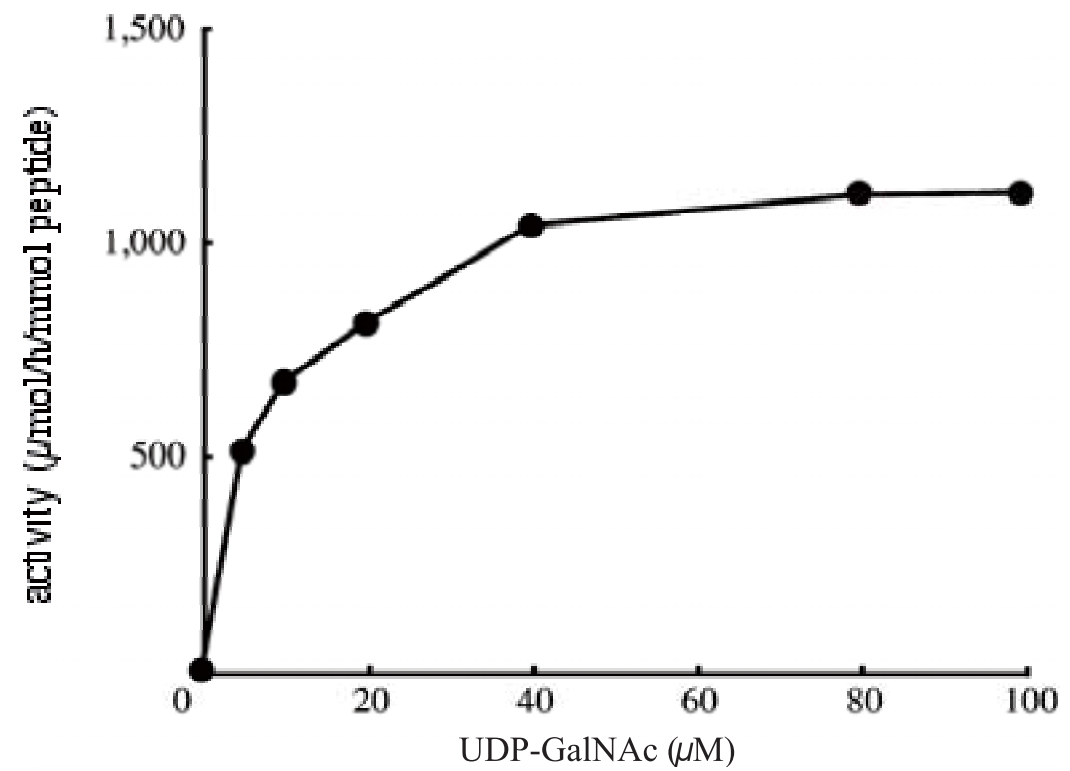

Fig. 5 Apparent Km determination for UDP-GalNAc

Reaction conditions are described under Materials and Methods.

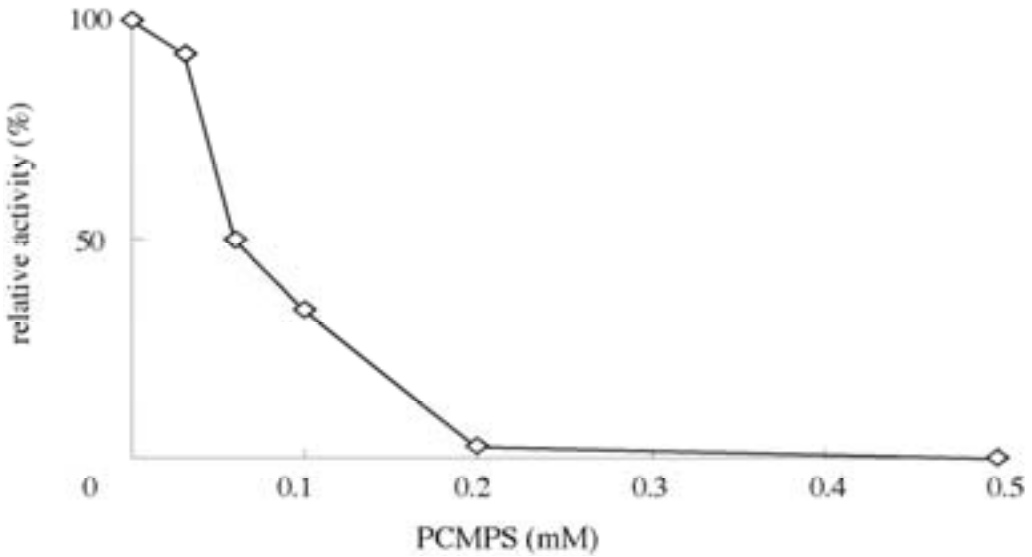

Fig. 6 Inhibition of pgant3 with PCMPS

Purified pgant3 was incubated with increasing concentrations of PCMPS, and assayed for activity as described under Materials and Methods.

motif, are the sites of the modification and are involved in the binding with UDP-GalNAc [25]. pgant3, containing both Cys residues corresponding to mammalian Cys212 and Cys214, was also incubated with increasing concentrations of PCMPS. PCMPS effectively inhibited pgant3 with an apparent Ki of 0.05 mM (Fig. 6). This raised the possibility that two Cys residues (Cys246 and Cys248) C-terminal to DXH in pgant3 are also involved in the recognition of UDP-GalNAc.

3.6 Assay for substrate specificity

Finally, a number of peptides were synthesized and used for the assay (Table1) in order to examine the acceptor substrate specificity of pgant3. The peptide sequences were obtained from typical mucin-like molecules of *Drosophila* and human origin that are characteristically rich in Ser, Thr, and Pro residues. As for the putative insect mucins, I searched for genes coding for a putative mucin-like molecule in the *Drosophila* database. I then examined the O-glycosylation probability of the candidate sequences according to the method described by Elhammer *et al.* [39], and the peptide sequences with a high probability were chosen for the peptide synthesis (mucin like-1 to -14, Table1). I also prepared the peptides containing putative O-glycosylation sites found in insect hemomucin, which is a putative self-defense protein found in hemocytes, in the gut [40]. In addition to insect peptides, those with sequences from mammalian mucins and glycophorin were included. All the peptides were synthesized onto the Gly residue covalently bound to the multipins. The synthe-

Table 1 Substrate specificity of pgant3

peptide	sequence	*activity	peptide	sequence	*activity
Drosophila			Mammals		
mucin like-1	PPVVPFTPLPG	107	MUC2-9	TPTGTQTPTTG	25
mucin like-2	MPTAGTPPPIG	299	MUC2-10	TQTPTTTPITG	39
mucin like-3	PGGDSSSYQPG	20	MUC3-1	HSTPSFTSSIG	20
mucin like-4	LPPVTTTPAPG	904	MUC3-2	SFTSSITTTEG	19
mucin like-5	SEVTTTLVPAG	26	MUC3-3	SITTTETTSHG	49
mucin like-6	PADVTTAAPGG	544	MUC3-4	TETTSHSTPSG	18
mucin like-7	SDAESSGPVPG	20	MUC4-1	TSSVSTGHATG	27
mucin like-8	AWIPTPPPIPG	1172	MUC4-2	STGHATSLPVG	21
mucin like-9	QWDPTPISTPG	20	MUC4-3	ATSLPVTDTSG	23
mucin like-10	PTPISTPACEG	14	MUC4-4	PVTDTSSVSTG	17
mucin like-11	RCESDTTPPPG	80	MUC5AC-1	GTTPSPVPTTG	148
mucin like-12	EEEIPATPRPG	53	MUC5AC-2	SPVPTTSTTSG	19
mucin like-13	GSPCTTDNPSG	19	MUC5AC-3	TTSTTSAPGTG	192
mucin like-14	EPPITTPHCSG	6	MUC5AC-4	TSAPGTTPSPG	520
hemomucin-1	PTTTTTTTTPG	90	MUC5B-1	SSTPGTAHTLG	19
hemomucin-2	TTTTTTPKPTG	605	MUC5B-2	GTAHTLTVLTG	18
hemomucin-3	TTTPKPTTKTG	32	MUC5B-3	TLTVLTTTATG	20
hemomucin-4	PKPTTKTTTTG	20	MUC5B-4	LTTTATTPTAG	37
hemomucin-5	PEPSKPKVKRG	17	MUC5B-5	ATTPTATGSTG	34
hemomucin-6	TTPTTPTTPTG	20	MUC5B-6	TATGSTATPSG	32
Mammals			MUC5B-7	STATPSSTPGG	40
MUC1-1	PDTRPDAPGSG	19	MUC7-1	TTAAPPTPSAG	516
MUC1-2	PDAPGSTAPPG	88	MUC7-2	PPTPSATTPAG	44
MUC1-3	GSTAPPAHGVG	34	MUC7-3	SATTPAPPSSG	2311
MUC1-4	PPAHGVTSAPG	267	MUC7-4	PAPPSSSAPPG	29
MUC1-5	GVTSAPDTRPG	35	MUC7-5	SSSAPPETTAG	25
MUC2-1	PTTTPSPPITG	533	MUC7-6	PPETTAAPPTG	484
MUC2-2	PSPPITTTTTG	109	glycophorin-1	VLLLSAIVSIG	21
MUC2-3	ITTTTTPLPTG	268	glycophorin-2	SAIVSISASSG	20
MUC2-4	TTPLPTTTPSG	128	glycophorin-3	SISASSTTGVG	25
MUC2-5	PTTTPITTTTG	46	glycophorin-4	SSTTGVAMHTG	27
MUC2-6	PITTTTTVTPG	139	glycophorin-5	GVAMHTSTSSG	23
MUC2-7	TTTVTPTPTPG	429	glycophorin-6	HTSTSSSVTKG	33
MUC2-8	TPTPTPTGTQG	58			

*The activity is indicated as μmol GalNAc/h/mmol peptide.
Origin of *Drosophila* mucin like peptides is as follows, mucin like-1 and -2; CG11570, mucin like-3; CG12480, mucin like-4; CG13722, mucin like-5; CG13817, mucin like-6 and -7; CG2779, mucin like-8; CG8697 (larval cuticle protein 2), mucin like-9 and -10; CG17058 (peritrophin A), mucin like-11 and-12; CG17014, mucin like-13; CG6004, mucin like-14; CG10287 (Gasp).

sized peptides, therefore, have a C-terminal Gly residue.

Assay with these peptides indicated that pgant3 has broad substrate specificity capable of glycosylating almost all of the peptides from both insect and mammalian sequences, but with a con-

siderably distinct reaction rate depending on the peptide sequences (Table1). Concerning the peptides derived from insects, mucin like-2 (MPTAGTPPPIG), -4 (LPPVTTTPAPG), -6 (PADV TTAAP GG), -8 (AWIPTPPPIPG), and hemomucin-2 (TTTTTTPKPTG) were excellent acceptors. Comparison of their sequence revealed that all the peptides except mucin like-6 contain XTPXP (underlined above), a sequence that is defined as "a mucin box" and is reported to be an efficient O-glycosylation site by Takeuchi *et al.* [41] A sequence TTAAP contained in mucin like-6 also seemed to be a good acceptor site, since mammalian MUC7-1 (TTAAP PTPSAG) and MUC7-6 (PPE TTAAP PTG), which similarly contain TTAAP, were glycosylated as efficiently as mucin like-6. It should, however, be noted that some peptides containing XTPXP were poor acceptors. For example, mucin like-11 (RCESDTPPPG), -12 (EEEIPATPRPG), and hemomucin-3 (TTTPKPTTKTG) were not glycosylated efficiently. These peptides, however, contain several charged amino acids (represented in the boldface in the above sequences), which are unfavorable residues for O-glycosylation [39,42]. in and around the XTPXP sequences. As to the mammalian synthetic peptides, eight peptides were glycosylated at rate higher than 200 μmol GalNAc/h/mmol peptide. Seven of these contain XTPXP or TTAAP (MUC2-1, 2-3, 2-7, 5AC-4, 7-1, 7-3, and 7-6), indicating again that pgant3 preferentially glycosylates the peptides containing these acceptor sites. Among the poor acceptors, there were peptides, such as MUC2-8 and MUC5AC-1, which contain XTPXP. This clearly demonstrates that the presence of a favorable primary sequence (e. g. XTPXP and TTAAP) is necessary but not a prerequisite for the efficient O-glycosylation by pgant3. Efficiency of the O-glycosylation would be influenced by several other factors than the primary sequence of the acceptor site as well, such as the amino acid sequence flanking the acceptor site and the secondary structure of the peptide.

As described above, pgant3 glycosylated, some of mucin like insect peptides efficiently. Among insect peptides, mucin like-8 was the most efficient substrate. It has a putative O-glycosylation site of larval cuticle protein II that is homologous to human MUC2A [43,44]. It appears that O-glycosylation of this protein is involved in a protective role in the larvae. Another efficient acceptor, mucin like-4, is derived from a putative uncharacterized polypeptide, but it is homologous to the insect salivary glue protein and human MUC5BC. Mucin like-6 is also correlated with the mammalian mucins, such as human MUC2 and zonadhesin. In addition, it is homologous to a putative insect intestinal mucin, insect intestinal mucin (IIM). IIM was identified in the larvae of a lepidopterous insect, *Trichoplusia ni* [45]. It is proposed that the O-glycosylation of IIM is involved in the protection of digestive tracts from the protease attack and/or the microbial infection. The substrate specificity of pgant3 suggests that pgant3 is involved in O-glycosylation of some of the insect mucin-like molecules. pgant3, on the other hand, did not glycosylate the peptides (mucin like-9 to -14) from peritrophin efficiently. Peritrophin is a component of the peritrophic membrane of insects that cov-

ers the digestive tract [46]. Some of the peritrophin family are characterized by the presence of a mucin like domain, which is expected to be O-glycosylated.

Importance of putative insect mucins, especially in host-pathogen interactions, has been reported. For example, the disruption of IIM by a baculovirus enhancin resulted in the enhanced baculovirus infection [47]. A *Drosophila* class C macrophage-specific scavenger receptor at the surface of macrophages (dSR-CI) has mucin-like domains highly homologous to hemomucin [48]. *Trypanosoma cruzi*, which causes American trypanosomiasis (Chagas' disease), has a mucin-like molecule (TcMUC) that is involved in the invasion to host cells [49]. Venturia, a parasitoid of wasp, has a hemomucin-like molecule that protects egg and larvae from host immunity [50]. The fact that pgant3 efficiently glycosylated one of the synthetic peptides, mucin like-6, related to IIM, suggests that these molecules are actually O-glycosylated in insects and involved in the recognition of the immune system. This also suggests that the other putative insect GalNAc-transferases found in the database may be related to the O-glycosylation of other mucin-like molecules such as hemomucin, and peritrophin, since these were not good acceptors for pgant3. pgant3 characterized in this study glycosylated typical mucin-like sequences without prior O-glycosylation of the peptides. In mammals, some isozymes such as GalNAc-T4, -T7, and -T10 [8,10,15,16] are regarded as the follow-up isozymes, which have either no or very low initial O-glycosylation activity, but recognize partially O-glycosylated peptides and transfer more GalNAc to them. It would reasonable to predict the presence of the so-called follow-up isozymes in insects as well.

3.7 GalNAc-transferase family in Drosophila

A survey of the *Drosophila* genome sequence identified 12 putative GalNAc-transferases in *Drosophila*, including pgant3 characterized in this study (Table2). Some of them were highly homologous to the mammalian isozymes. For example, pgant5 is 82% homologous to human GalNAc-T1, when the amino acid sequences of the whole coding region were compared (Table3). pgant2 is homologous to human GalNAc-T2 with 86% similarity as well. Most of the putative *Drosophila* GalNAc-transferase genes seem to code for the full-length isozymes. However, *CG10000*, *CG7579*, *and CG7304* apparently code for the truncated isozymes. The putative polypeptides encoded by *CG10000* and *CG7579* lack N-terminal sequences. The C-terminal $(QXW)_3$ repeats are deleted in *CG7304* as well. These truncated sequences may be regarded as pseudogenes. Pseudogenes similarly processed have been identified in human genome [51]. Interestingly, the *Drosophila* genome contains three genes (*pgant8*, *CG7304*, and *CG7579*) that are mapped at the same chromosomal location (3L: 71E5-71E5), with *CG7304* in the inverted orientation (Table2). Although they are located at the close position in the genome, they have diverged sequences. *pgant8* seems to code for the full-

Table 2 Putative GalNAc-transferase genes in *Drosophila*

Entry	FlyBase No.	Nucleotides (bp)	Amino Acids (residues)	Chromosome	Cytogenic Map
pgant6	FBan0002103	2658	666	3L	63A1--2
pgant2	FBan0003254/FBgn0031530	1773	590	2L	23F1-23F1
pgant3	FBan0004445/FBgn0027558	2858	667	2R	42E2-42E2
pgant7	FBan0006394	2340	591	X	17B3--5
pgant8	FBan0007297/FBgn0036529	1773	590	3L	71E5-71E5
CG7579	FBan0007579/FBgn0036528	1497	498	3L	71E5-71E5
CG7304	FBan0007304/FBgn0036527	1424	474	3L	71E5-71E5
CG7480	FBan0007480/FBgn0001970	2090	632	2L	35A1-35A1
pgant1	FBan0008182/FBgn0034025	2125	601	2R	52A6-52A6
pgant4	FBan0008845/FBgn0031541	3759	1252	2L	23F3-23F3
pgant5	FBan0009152/FBgn0031681	1893	630	2L	25C4-25C4
CG10000	FBgn0010000/FBgn0039596	1392	434	3R	98E2-98E2

Table 3 Amino acid homology of the putative catalytic domain among the GalNAc-transferases Amino aced similarity was calcurated by Clastal W method in Mac Vector. The paramiters for alignment were : slow alignment, open gap penalty=10, extend gap penalty=0.1, matrix=30, delay divergence=40%, and no hydrophilic gap penalty. h ; huma.

	hGalNAc-T1	hGalNAc-T2	pgant6	pgant2	pgant3	pgant7	pgant8	CG7304	CG7579	pgant35	pgant1	pgant4	pgant5
hGalNAc-T1	—	—	—	—	—	—	—	—	—	—	—	—	—
hGalNAc-T2	70	—	—	—	—	—	—	—	—	—	—	—	—
pgant6	64	58	—	—	—	—	—	—	—	—	—	—	—
pgant2	70	86	58	—	—	—	—	—	—	—	—	—	—
pgant3	74	62	60	64	—	—	—	—	—	—	—	—	—
pgant7	65	60	63	58	59	—	—	—	—	—	—	—	—
pgant8	59	58	62	56	58	55	—	—	—	—	—	—	—
CG7304	48	45	48	44	45	47	51	—	—	—	—	—	—
CG7579	51	50	53	49	51	47	57	51	—	—	—	—	—
pgant35	64	66	57	66	56	57	55	47	51	—	—	—	—
pgant1	64	60	58	59	59	59	52	44	50	61	—	—	—
pgant4	64	56	74	55	55	60	60	48	55	57	56	—	—
pgant5	82	68	66	69	75	62	59	47	53	62	61	63	—
CG10000	48	49	46	48	42	46	46	39	43	46	45	47	46

length isozyme, while the others (*CG7304*, and *CG7579*) for the pseudogenes as described above. Hence, they exhibit the relatively low homology (Table3). There are two possibilities that account for the occurrence of these genes in the same genomic region and their sequence divergence. One possibility is that they are triplicated genes derived from the same ancestral gene, and that the two pseudogenes accumulated mutations in their entire coding sequence. Alternatively, it may be possible that these genes are independent ones that happened to locate at the same genomic position.

A gene *pgant4* encodes for the largest protein among the insect GalNAc-transferases (Table2). This isozyme is characterized by a long stem region, which is similar to mammalian GalNAc-T5 [9]. Consequently, overall similarity is low (15~36%, data not shown) due to the presence of the long non-homologous stem region. However, when the sequence of the putative catalytic region is compared, its homology is as high as that of other genes (Table3).

To depict the phylogenetic relationship among the GalNAc-transferases from mammals and *Drosophila*, a neighbor-joining dendrogram was constructed (Fig. 7). The dendrogram shows seven separated clusters in the GalNAc-transferase family except for human GalNAc-T15 (Groups A to G in Fig. 7). Among these groups, Groups A, B, D, and E consist of the GalNAc-transferases from both mammals and *Drosophila*. By contrast, Groups C and F contain only mammalian isozymes, and Group G *Drosophila* ones. In Group G, only pgant8 codes for the full-length isozyme, while the other three clones (CG10000, CG7579, and CG7304) are regarded as pseudogenes. The sequence homology among them is significantly low (Table3), compared with the similarity of the other groups. Concerning the mammalian specific groups, the insect genes in these groups might be deleted after speciation. Otherwise, mammals might have acquired novel isozymes with new catalytic properties after speciation. Interestingly, the brain-specific isozymes, GalNAc-T9 [14] and pt-GalNAc-T, belong to Group F. This implies that these mammalian isozymes emerged after speciation and acquired specific catalytic properties essential for the O-glycosylation in the mammalian brain during the evolution. In general, the similarity of the GalNAc-transferases from *Drosophila* and from mammals is low (usually less than 60% when the sequence of the putative catalytic domain was compared). However, it should be noted that some GalNAc-transferases between these species are highly homologous. For example, pgant5 and hGalNAc-T1 (and -T13) in Group B exhibited 82% homology. Also, pgant2 and hGalNAc-T2 in Group D show high similarity (86%). Judging from the observation that the highly homologous human clones, GalNAc-T3 and -T6 (6,10), exhibit the similar substrate specificity, the homologous clones from the insect and the mammalian origin may have similar catalytic properties. Interestingly, hGalNAc-T1 and -T2 are ubiquitously expressed and regarded as essential isozymes for the general O-glycosylation in the cell.

In summary, I have characterized one of the *Drosophila* GalNAc-transferases (pgant3). Recombinant pgant3 expressed in COS7 cells was catalytically active and characterized by the low optimal temperature, compared with the mammalian enzymes. pgant3 glycosylated most of the mucin-like peptides, with strong preference to the sequences containing XTPXP and TTAAP, and did not require prior O-glycosylation of the peptides. This suggests that this isozyme may be responsible for the initiation of insect mucin O-glycosylation. pgant3 is expected to have low follow-up activity, as deduced by the primary structure of the enzyme, and, in fact, Ten Hagen *et al.* detected very low

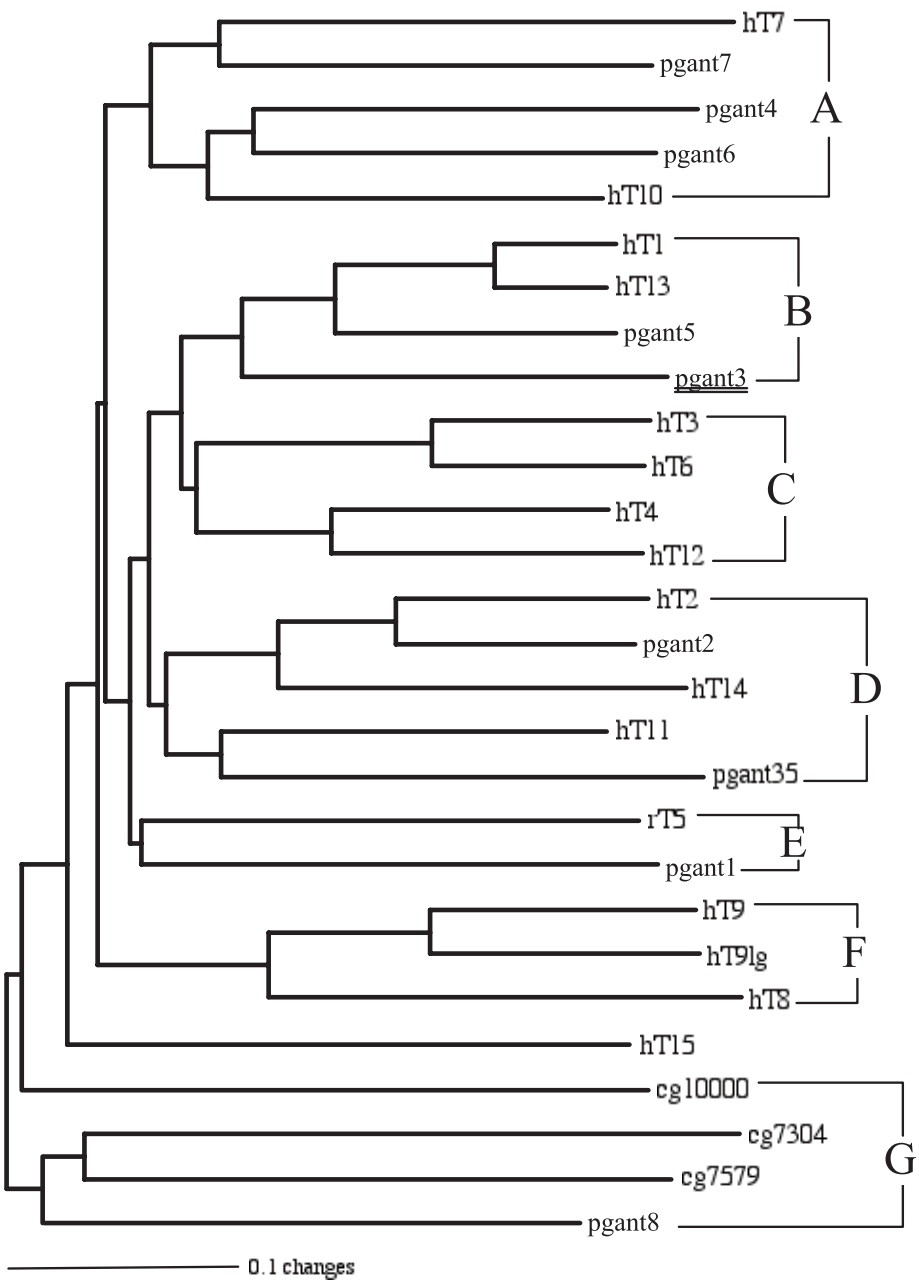

Fig. 7 Phylogenetic analysis of the GalNAc-transferases

A phylogenetic tree was constructed by comparing the sequences of the putative catalytic regions of the GalNAc-transferases. An NJ image method of the PAUP*version 4.0 (Swofford, 2002) was used for analysis. CG numbers are shown for putative *Drosophila* isozymes. h; human, r; rat, pgant; *Drosophila melanogaster*.

follow-up activity [24]. It would be reasonable to predict the presence of the so-called follow-up isozymes in insects, and a recent study reports the presence of insect isozymes with glycopeptide glycosylation activity [24]. Together with the various reports on the putative mucin-like molecules, it is expected that O-glycosylation in insect cells plays an important role in the cellular and/or molecular recognition, as observed in mammalian cells. Characterization of the GalNAc-transferases and elucidation of mucin-carbohydrate structures will help to understand the physiological importance of the O-glycosylation in insects.

Acknowledgments

I am extremely grateful to Dr. Akira Kurosaka, Department of Biotechnology, Faculty of Engineering, Kyoto Sangyo University, for great help on the entire study. I acknowledge Dr. Toru Terachi, Department of Biotechnology, Faculty of Engineering, Kyoto Sangyo University, for advise on molecular and genetic techniques. I am thankful to Dr. Shinya Toba, Department of Biotechnology, Faculty of Engineering, Kyoto Sangyo University, for advice and help on this study. I really would like to thank Dr. Naohiko Kodo, Department of Biotechnology, Faculty of Engineering, Kyoto Sangyo University, for experimental advices on molecular biological techniques. This work was in part supported by Grants-in-Aid for Scientific Research from Japan Society for the Promotion of Science, and the Foundation for Bioventure Research Center from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

Footnotes

Abbreviations: AP, alkaline phosphatase; EDTA, ethylenediamine tetraacetic acid; EST, expressed sequence tags; GalNAc, N-acetylgalactosamine; IgG, immunoglobulin G; IIM, insect intestinal mucin; NJ, neighbor-joining; PCMPS, *p*-chloromercuriphenylsulphonic acid; PCR, polymerase chain reaction; SDS, sodium dodecylsulfate; SSC, standard saline citrate; TTP, thymidine triphosphate;

References

[1] Paulson, J. C., and Colley, K. J. (1989) *J Biol Chem* 264, 17615-17618.

[2] Jentoft, N. (1990) *Trends Biochem Sci* 15, 291-294.

[3] Clausen, H., and Bennett, E. P. (1996) *Glycobiology* 6, 635-646.

[4] Homa, F. L., Hollander, T., Lehman, D. J., Thomsen, D. R., and Elhammer, A. P. (1993) *J Biol Chem* 268, 12609-12616.

[5] White, T., Bennett, E. P., Takio, K., Sorensen, T., Bonding, N., and Clausen, H. (1995) *J Biol Chem*

270, 24156-24165.

[6] Bennett, E. P., Hassan, H., and Clausen, H. (1996) *J Biol Chem* 271, 17006-17012.

[7] Hagen, F. K., Ten Hagen, K. G., Beres, T. M., Balys, M. M., VanWuyckhuyse, B. C., and Tabak, L. A. (1997) *J Biol Chem* 272, 13843-13848.

[8] Bennett, E. P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M. A., Merkx, G., van Kessel, A. G., Eiberg, H., Steffensen, R., and Clausen, H. (1998) *J Biol Chem* 273, 30472-30481.

[9] Ten Hagen, K. G., Hagen, F. K., Balys, M. M., Beres, T. M., Van Wuyckhuyse, B., and Tabak, L. A. (1998) *J Biol Chem* 273, 27749-27754.

[10] Ten Hagen, K. G., Tetaert, D., Hagen, F. K., Richet, C., Beres, T. M., Gagnon, J., Balys, M. M., VanWuyckhuyse, B., Bedi, G. S., Degand, P., and Tabak, L. A. (1999) *J Biol Chem* 274, 27867-27874.

[11] Bennett, E. P., Hassan, H., Mandel, U., Hollingsworth, M. A., Akisawa, N., Ikematsu, Y., Merkx, G., van Kessel, A. G., Olofsson, S., and Clausen, H. (1999) *J Biol Chem* 274, 25362-25370.

[12] Bennett, E. P., Hassan, H., Hollingsworth, M. A., and Clausen, H. (1999) *FEBS Lett* 460, 226-230.

[13] White, K. E., Lorenz, B., Evans, W. E., Meitinger, T., Strom, T. M., and Econs, M. J. (2000) *Gene* 246, 347-356.

[14] Toba, S., Tenno, M., Konishi, M., Mikami, T., Itoh, N., and Kurosaka, A. (2000) *Biochim Biophys Acta* 1493, 264-268.

[15] Hagen, K. G., Bedi, G. S., Tetaert, D., Kingsley, P. D., Hagen, F. K., Balys, M. M., Beres, T. M., Degand, P., and Tabak, L. A. (2001) *J Biol Chem* 276, 17395-17404.

[16] Cheng, L., Tachibana, K., Zhang, Y., Guo, J., Kahori Tachibana, K., Kameyama, A., Wang, H., Hiruma, T., Iwasaki, H., Togayachi, A., Kudo, T., and Narimatsu, H. (2002) *FEBS Lett* 531, 115-121

[17] Schwientek, T., Bennett, E. P., Flores, C., Thacker, J., Hollmann, M., Reis, C. A., Behrens, J., Mandel, U., Keck, B., Schafer, M. A., Haselmann, K., Zubarev, R., Roepstorff, P., Burchell, J. M., Taylor-Papadimitriou, J., Hollingsworth, M. A., and Clausen, H. (2002) *J Biol Chem* 277, 22623-22638

[18] Guo, J. M., Zhang, Y., Cheng, L., Iwasaki, H., Wang, H., Kubota, T., Tachibana, K., and Narimatsu, H. (2002) *FEBS Lett* 524, 211-218

[19] Zhang, Y., Iwasaki, H., Wang, H., Kudo, T., Kalka, T. B., Hennet, T., Kubota, T., Cheng, L., Inaba, N., Gotoh, M., Togayachi, A., Guo, J., Hisatomi, H., Nakajima, K., Nishihara, S., Nakamura, M., Marth, J. D., and Narimatsu, H. (2003) *J Biol Chem* 278, 573-584

[20] Wang, H., Tachibana, K., Zhang, Y., Iwasaki, H., Kameyama, A., Cheng, L., Guo, J., Hiruma, T., Togayachi, A., Kudo, T., Kikuchi, N., and Narimatsu, H. (2003) *Biochem Biophys Res Commun* 300, 738-744

[21] Hagen, F. K., and Nehrke, K. (1998) *J Biol Chem* 273, 8268-8277.

[22] Hagen, F. L., Michael, N. K., Gentile, K., Berbach, K., Tsao, C. C., and Forsythe, M. (2001) *Trends in Glycosci. Glycotech.* 13, 463-479

[23] Ten Hagen, K. G., and Tran, D. T. (2002) *J Biol Chem* 277, 22616-22622

[24] Ten Hagen, K. G., Tran, D. T., Gerken, T. A., Stein, D. S., and Zhang, Z. (2003) *J Biol Chem* 278, 35039-35048

[25] Tenno, M., Toba, S., Kezdy, F. J., Elhammer, A. P., and Kurosaka, A. (2002) *Eur J Biochem* 269, 4308-4316

[26] Swofford, D. L. (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods) version 4.* Sinauer Associates,, Sunderland, Massachusetts

[27] Chen, W., Unligil, U. M., Rini, J. M., and Stanley, P. (2001) *Biochemistry* 40, 8765-8772.

[28] Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A., and Negishi, M. (2000) *J Biol Chem* 275, 34580-34585.

[29] Gastinel, L. N., Bignon, C., Misra, A. K., Hindsgaul, O., Shaper, J. H., and Joziasse, D. H. (2001) *EMBO J* 20, 638-649.

[30] Unligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., and Rini, J. M. (2000) *EMBO J* 19, 5269-5280.

[31] Hagen, F. K., Hazes, B., Raffo, R., deSa, D., and Tabak, L. A. (1999) *J Biol Chem* 274, 6797-6803

[32] Hazes, B. (1996) *Protein Sci* 5, 1490-1501

[33] Wragg, S., Hagen, F. K., and Tabak, L. A. (1997) *Biochem J* 328, 193-197.

[34] Hassan, H., Reis, C. A., Bennett, E. P., Mirgorodskaya, E., Roepstorff, P., Hollingsworth, M. A., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (2000) *J Biol Chem* 275, 38197-38205

[35] Tenno, M., Saeki, A., Kezdy, F. J., Elhammer, A. P., and Kurosaka, A. (2002) *J Biol Chem* 277, 47088-47096

[36] Tenno, M., Kezdy, F. J., Elhammer, A. P., and Kurosaka, A. (2002) *Biochem Biophys Res Commun* 298, 755-759

[37] Elhammer, A., and Kornfeld, S. (1986) *J Biol Chem* 261, 5249-5255.

[38] Elhammer, A. P., Kezdy, F. J., and Kurosaka, A. (1999) *Glycoconj J* 16, 171-180.

[39] Elhammer, A. P., Poorman, R. A., Brown, E., Maggiora, L. L., Hoogerheide, J. G., and Kezdy, F. J. (1993) *J Biol Chem* 268, 10029-10038.

[40] Theopold, U., Samakovlis, C., Erdjument-Bromage, H., Dillon, N., Axelsson, B., Schmidt, O., Tempst, P., and Hultmark, D. (1996) *J Biol Chem* 271, 12708-12715.

[41] Yoshida, A., Suzuki, M., Ikenaga, H., and Takeuchi, M. (1997) *J Biol Chem* 272, 16884-16888.

[42] O'Connell, B. C., Hagen, F. K., and Tabak, L. A. (1992) *J Biol Chem* 267, 25010-25018.

[43] Snyder, M., Hirsh, J., and Davidson, N. (1981) *Cell* 25, 165-177.

[44] Steinemann, M., Steinemann, S., and Pinsker, W. (1996) *J Mol Evol* 43, 413-417.

[45] Wang, P., and Granados, R. R. (1997) *J Biol Chem* 272, 16663-16669.

[46] Tellam, R. L., Wijffels, G., and Willadsen, P. (1999) *Insect Biochem Mol Biol* 29, 87-101.

[47] Wang, P., and Granados, R. R. (1997) *Proc Natl Acad Sci U S A* 94, 6977-6982.

[48] Pearson, A., Lux, A., and Krieger, M. (1995) *Proc Natl Acad Sci U S A* 92, 4056-4060.

[49] Acosta-Serrano, A., Almeida, I. C., Freitas-Junior, L. H., Yoshida, N., and Schenkman, S. (2001) *Mol Biochem Parasitol* 114, 143-150.

[50] Kinuthia, W., Li, D., Schmidt, O., and Theopold, U. (1999) *Journal of Insect Physiology* 45, 501-506

[51] Meurer, J. A., Drong, R. F., Homa, F. L., Slightom, J. L., and Elhammer, A. P. (1996) *Glycobiology* 6, 231-241.

ショウジョウバエ UDP-GalNAc : polypeptide N-acetylgalactosaminyltransferase の性質について

中　村　直　介

要　旨

ムチン型糖鎖の生合成はタンパク質への GalNAc の付加により始まり，この反応は UDP-GalNAc : polypeptide N-acetylgalactosaminyltransferase（以降 GalNAc 転移酵素）が触媒する．本論文にて，筆者はショウジョウバエの GalNAc 転移酵素遺伝子のうちの一つ，pgant3 の最初の生化学的特徴について報告する．このアイソザイムは，触媒活性に必須な，よく保存されたモチーフ構造を有している一方で，レクチン様ドメイン内に何カ所かの挿入配列を持つという特徴がある．筆者はまず，ノーザンハイブリダイゼーション解析より本酵素 mRNA 発現を調べたところ，ショウジョウバエの幼虫にて，2.5kb の mRNA を認めた．次に，ほ乳類培養細胞で本酵素を組換えタンパク質として発現させ，生化学的特徴を解析した．その結果，pgant3 は酵素活性に Mn^{2+} を必要とし，pH 7.5～8.5の至適 pH を持ち，29-33℃ で最も効率よく糖を転移する事が分かった．UDP-GalNAc に対する Km は，10.7 μM と，ほ乳類のアイソザイムと同様に高い親和性を示した．また，数多くの合成ペプチドを用いて基質特異性を調べた結果，pgant3 は XTPXP や TTAAP の配列を含む典型的なムチン様配列を持つペプチドに効率よく GalNAc を転移することが分かった．さらに，Cys 残基特異的な不可逆的阻害剤である *p*-chloromercuriphenylsulfonic acid で強く阻害を受けたことから，本酵素はほ乳類 GalNAc-T1 と同様に酵素活性に必要な Cys 残基を持つ事が示された．筆者はまた，データベース検索により，pgant3 の他に，既知の GalNAc 転移酵素と相同性を持つショウジョウバエ由来の cDNA クローンをいくつか見いだした．ショウジョウバエには少なくとも12種類の GalNAc 転移酵素遺伝子が存在し，ほ乳類と同様に GalNAc 転移酵素は大きな遺伝子ファミリーを形成している事が示された．これらの知見より，ほ乳類と同様に，GalNAc 転移酵素は昆虫においても多くのアイソザイムの協奏的な働きによりムチン型糖鎖の生合成を調整している事が示唆された．

キーワード：GalNAc 転移酵素，ムチン，ショウジョウバエ，O-グリコシレーション，昆虫，ペプチド