Expression and characterization of silkworm Bombyx mori β-1,2-N-acetylglucosaminyltransferase II, a key enzyme for complex-type N-glycan biosynthesis

Expression and characterization of silkworm *Bombyx mori* β-1,2-*N*-acetylglucosaminyltransferase II, a key enzyme for complex-type *N*-glycan biosynthesis

Takatsugu Miyazaki[1,2], Ryunosuke Miyashita[2], Sota Mori[2], Tatsuya Kato[1,2], and Enoch Y. Park[1,2,*]

[1]Laboratory of Biotechnology, Research Institute of Green Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka, 422-8529, Japan.
[2]Laboratory of Biotechnology, Department of Agriculture, Graduate School of Integrated Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529, Japan.

*__Correspondence:__ E.Y. Park, Laboratory of Biotechnology, Research Institute of Green Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka, 422-8529, Japan; Tel & Fax: +81-54-238-4887; E-mail: park.enoch@shizuoka.ac.jp

Short title: Characterization of silkworm GnTII

Abstract

N-glycans are involved in various physiological functions and their structures diverge among different phyla and kingdoms. Insect cells mainly produce high mannose-type and paucimannose-type glycans but very few mammalian-like complex-type glycans. However, many insects possess genes for proteins homologous to the enzymes involved in complex-type *N*-glycan synthesis in mammalian cells, and their *N*-glycosylation pathway is incompletely understood compared with that of mammals. Here, we cloned a candidate gene for β-1,2-*N*-acetylglucosaminyltransferase II (GnTII), which is a Golgi-localized enzyme involved in a key step in the conversion to complex-type *N*-glycans, from silkworm *Bombyx mori*, and the gene was found to be expressed ubiquitously in the larval and pupal stages. In addition, recombinant *B. mori* GnTII was expressed as a soluble form using a silkworm-*B. mori* nucleopolyhedrovirus bacmid expression system. The recombinant enzyme exhibited similar pH and temperature dependency and the same substrate specificity as human GnTII, but deglycosylation with peptide:*N*-glycanase F did not affect its enzymatic activity. Compared with the structure of human GnTII, the amino acid residues involved in catalytic activity and substrate recognition are almost fully conserved in *B. mori* GnTII, which is consistent with its enzymatic properties. These results raised the possibility of mammalian-like complex-type *N*-glycan synthesis using the GnTII ortholog in silkworm.

Keywords: β-1,2-*N*-acetylglucosaminyltransferase II, *Bombyx mori*, *N*-glycan, *N*-acetylglucosamine, complex type, glycosyltransferase family 16

N-glycosylation is one of the most common post-translational modifications of proteins and is involved in the folding, function, stability and trafficking of proteins in eukaryotes (1, 2). This modification is initiated in the endoplasmic reticulum (ER), where a tetradecasaccharide precursor $Glc_3Man_9GlcNAc_2$ (Glc, glucose; Man, mannose; GlcNAc, *N*-acetylglucosamine) is transferred to an asparagine residue in the Asn-X-Ser/Thr sequon (X is any amino acid residue except Pro) of a nascent polypeptide. Sequential modifications catalyzed by glycoside hydrolases and glycosyltransferases produce a wide variety of *N*-glycans, such as high mannose-type, hybrid-type, and complex-type glycans. The early steps of glucose- and mannose-trimming from the glycan precursor by α-glucosidase I, α-glucosidase II and α-mannosidase in the ER produce high mannosidic *N*-glycans, which are highly conserved among almost all eukaryotes (3, 4). In contrast, in the later steps, several Golgi-resident glycosidases and glycosyltransferases modify the structure of *N*-glycans, which diverge significantly among different phyla and kingdoms (5, 6).

Baculovirus-insect cell expression systems are widely used for recombinant protein production because they have the capacity for high-level production of many recombinant proteins with post-translational modifications that are similar to those found in mammalian cells, including the ability to glycosylate (7, 8). However, differences in *N*-glycan structures influence the biological activity of glycoproteins. Insect cells mainly produce paucimannose-type *N*-glycans, whose structure is quite different from the structure of complex-type glycans produced in mammals (Fig. S1). Both insect and mammalian cells produce Man(α1-6)[GlcNAc(β1-2)Man(α1-3)]Man(β1-4)GlcNAc(β1-4)GlcNAc (MGn) glycan as an intermediate during *N*-glycan processing. In insect cells, β-*N*-acetylglucosaminidase (fused lobes, FDL) removes a GlcNAc residue of the α1-3

arm of MGn glycan to produce Man(α1-6)[Man(α1-3)]ManGlcNAc$_2$ (MM), a core structure from paucimannose-type *N*-glycans (9, 10). In recent years, paucimannosidic glycans have been found in azurophilic granule-specific human neutrophil proteins from pathogen-infected sputum and are suggested to be generated by azurophilic granule-resident β-hexosaminidase A (Fig. S1) (11). On the other hand, β-1,2-*N*-acetylglucosaminyltransferase II (GnTII) catalyzes the transfer of GlcNAc from a uridine diphosphate (UDP)-GlcNAc donor to the α1-6 arm of MGn glycan to produce biantennary complex-type glycans in mammalian cells (12, 13). Additionally, galactosyltransferase and sialyltransferase sequentially elongate glycan with galactose and sialic acid, respectively.

Several approaches have successfully produced glycoproteins modified with mammal-like complex *N*-glycans in insect cells using baculovirus vector and transgenic technology. Insect cell lines can produce terminally galactosylated and sialylated *N*-glycoproteins via coexpression with glycosyltransferases, including β-1,2-*N*-acetylglucosaminyltransferase I (GnTI), GnTII, β-1,4-galactosyltransferase, and sialyltransferases (14–16). Silkworm *Bombyx mori* has been used for the production of recombinant proteins, and the transgenic silkworm that produces terminally galactosylated *N*-glyoproteins in silk glands was developed by using the *piggyBac* transposon system and a posterior silk gland-specific promoter (17). We also succeeded in converting the glycan structure produced in *B. mori* pupae by the co-expression of human GnTII (hGnTII) and β-1,4-galactosyltransferase using a silkworm-based *B. mori* nucleopolyedrovirus (BmNPV) bacmid system (18). Moreover, suppression of FDL function using RNA interference and inhibitor 2-acetamido-1,2-dideoxynojirimycin was attempted to improve a rate of GlcNAcylated *N*-glycoproteins in insect cells and

silkworm *Bombyx mori* (19, 20). In *Drosophila melanogaster* S2 cells a CRISPR-Cas9-mediated knockout of *fdl* gene elongated the structures of *N*-glycans (21).

Recently, many insects have been found to possess genes encoding proteins homologous to the mammalian glycosyltransferases needed to produce hybrid and complex-type *N*-glycans (22–26). In addition, complex-type *N*-glycans have been found in some insect tissues and cells (5, 27–29). However, small amounts of the complex-type, especially biantennary, *N*-glycans were observed in recombinant glycoproteins expressed in native insect cells (7), and the activity of GnTII was reported to be much less in insect cells than in mammalian cells (30). In the case of silkworm, the *N*-glycans attached to recombinant glycoproteins expressed in larval hemolymph using a baculovirus vector were not GlcNAcylated (20, 31) but GlcNAc-terminated biantennary *N*-glycans were expressed in silk glands (17, 32). Thus, the relationship between the function of glycosyltransferases, especially GnTII, and the *N*-glycan structure in silkworm is incompletely understood. Here, we cloned *B. mori* GnTII (BmGnTII) and expressed the enzyme in silkworm larvae using the BmNPV bacmid, a shuttle vector for *Escherichia coli* and *B. mori*, and the recombinant BmGnTII was biochemically characterized.

MATERIALS AND METHODS

Transcriptional analysis

First- to fifth-instar larvae and pupae (Ehime Sanshu, Ehime, Japan) were flash-frozen in liquid nitrogen and ground into powder. Total RNA was extracted using TRIzol reagent (Life Technologies Japan, Tokyo, Japan) from the resulting powder. First-strand cDNA was synthesized with a PrimeScript RT reagent kit (Takara, Shiga, Japan)

according to the manufacturer's protocol. Expression of the *BmGnTII* gene was estimated by PCR using a gene-specific primer set, BmGnTII_F and BmGnTII_R (Table 1), and the synthesized cDNA as a template. The gene for glyceraldehyde-3-phosphate dehydrogenase (*BmGAPDH*) was used as an internal control and was amplified using the BmGAPDH_F and BmGAPDH_R primers (Table 1). PCR products were analyzed by electrophoresis on a 2.0% agarose gel.

Construction of BmNPV bacmid of recombinant BmGnTII

The transmembrane region of BmGnTII was predicted using the TMHMM server (http://www.cbs.dtu.dk/services/TMHMM/) (33). DNA fragments encoding the N-terminally FLAG (DYKDDDDK)-tagged lumenal region (Leu27–Ala498) including the stem domain and the catalytic domain of BmGnTII were amplified by PCR using the synthesized cDNA from a fifth-instar larva as a template and a pair of primers, FLAG-BmGnTII_F and BmGnTII_KpnI_R (Table 1). A secretion signal peptide derived from bombyxin was added by PCR using the resultant DNA fragment as a template and a pair of primers, EcoRI-Bx-FLAG_F and BmGnTII_KpnI_R (Table 1), and then, the products were ligated into the pFastBac1 vector (Thermo Fisher Scientific K.K., Yokohama, Japan) using EcoRI and KpnI restriction sites. The constructs were verified by DNA sequencing. *Escherichia coli* BmDH10Bac-*CP*$^{-}$-*Chi*$^{-}$ competent cells, which contain the cysteine protease- and chitinase-deficient BmNPV bacmid (34), were transformed with the resultant plasmid and cultured on LB agar plate medium containing 50 μg/mL kanamycin, 7 μg/mL gentamycin, 10 μg/mL tetracycline, 40 μg/mL isopropyl β-D-1-thiogalactopyranoside, and 100 μg/mL 5-bromo-4-chloro-3-indolyl-4-galactopyranoside at 37°C for 18 h. The recombinant BmNPV bacmid containing the *BmGnTII* gene was

extracted from a white positive colony after blue-white selection and colony PCR identification with pUC/M13 forward and pUC/M13 reverse primers (Table 1).

Expression and purification of recombinant BmGnTII

To produce the soluble form of BmGnTII, chitosan/BmNPV bacmid nanocomplexes were prepared as described previously (35) and then injected into fifth-instar silkworm larvae. The bacmid-injected larvae were reared on an artificial diet (Silkmate S2, Nohsan Corporation, Yokohama, Japan) at 26ºC for 6 days. The hemolymph was collected by cutting a caudal leg in a tube containing 1 mM 1-phenyl-2-thiourea and the fat body was collected by cutting and dissection. The hemolymph and fat body samples were stored at −80ºC until further analysis. The hemolymph was mixed with 9 volumes of 50 mM Tris-HCl buffer (pH 7.4) containing 300 mM NaCl and 16% polyethylene glycol 4,000 (Wako Pure Chemical Co., Osaka, Japan) and incubated at 4ºC overnight. The precipitant was removed by centrifugation at 12,000×*g* for 10 min, and the supernatant was applied onto a DDDDK-tagged Protein Purification Gel (Medical and Biological Laboratories, Nagoya, Japan). The column was washed with 50 mM Tris-HCl buffer (pH 7.4) containing 300 mM NaCl and 0.1% Triton X-100, and then, the FLAG-tagged protein was eluted with 100 μg/mL FLAG-tag peptide (Medical and Biological Laboratories). The fraction containing recombinant BmGnTII was buffer-exchanged into 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)-NaOH buffer (pH 6.0) using an Amicon Ultra centrifugal device (Merck, Darmstadt, Germany). Protein expression and purity were confirmed by SDS-PAGE with Coomassie brilliant blue (CBB) staining and western blotting using an anti-DDDDK-tag monoclonal antibody (Medical and Biological Laboratories) as the primary antibody and an anti-mouse IgG

antibody labeled with horseradish peroxidase (Medical and Biological Laboratories) as the secondary antibody. Protein concentration was determined by bicinchoninic acid methods using a Pierce BCA Protein Assay Kit (Thermo Fisher Scientific K.K.) and bovine serum albumin (BSA) as a standard.

Enzyme assay

Fluorescent pyridylaminated (PA) glycans used were purchased from Masuda Chemical Industries (Takamatsu, Japan). MGn-PA was also prepared using non-labeled MGn glycan (N030; Chemily Glycoscience, Atlanta, GA, USA) and a Pyridylamination Manual Kit (Takara Bio, Shiga, Japan) according to the manufacturer's protocol. To measure the glycosyltransferase activity of the recombinant protein, a reaction mixture containing 50 μg/mL recombinant BmGnTII, 1 μM MGn-PA, 1 mM UDP-GlcNAc, 10 mM $MnCl_2$, 1 mg/mL BSA, and 50 mM HEPES-NaOH (pH 6.5) was prepared and incubated at 37ºC. To examine the pH effect on glycosyltransferase activity, the same reaction mixtures, with the exception that 50 mM sodium acetate (pH 4.0–6.0), HEPES-NaOH (pH 6.0–8.0), Tris-HCl (pH 8.0–9.0), or glycine-NaOH (pH 9.0–11.0) buffers were used, were incubated at 37ºC for 1 min. The effect of temperature was assayed at 20–70ºC using 50 mM HEPES-NaOH buffer (pH 6.5). Metal dependency was investigated at 37ºC in 50 mM HEPES-NaOH buffer (pH 6.0) with 10 mM $MgCl_2$, $CaCl_2$, or $CoCl_2$, instead of $MnCl_2$. After incubation, the reaction mixtures were boiled for 5 min to stop the enzymatic reaction. The amounts of PA-glycans were measured by reverse phase high-performance liquid chromatography (HPLC) using TSKgel ODS-80T_M (4.6 mm × 250 mm, Tosoh, Tokyo, Japan) and an HPLC system equipped with LC-10AD VP pumps and an RF-10A XL fluorescence detector (Shimadzu, Kyoto, Japan) at 320 nm excitation

and 400 nm emission. The samples were eluted under isocratic conditions at 1.0 mL/min using 100 mM ammonium acetate buffer (pH 4.0) with a column temperature of 40ºC. All activity measurements were performed in triplicate.

Glycan digestion

The recombinant protein was denatured and treated with peptide:*N*-glycanase F (PNGase F, Takara Bio) according to the manufacturer's protocol. For the activity measurement, recombinant BmGnTII was treated with PNGase F in native conditions.

Homology modeling

The homology model of BmGnTII was produced using the SWISS-MODEL server (https://swissmodel.expasy.org) (36) using the amino acid sequence of BmGnTII and the coordinate of hGnTII (PDB 5VCM) as a template. Figures were prepared using PyMOL (http://www.pymol.org/).

RESULTS AND DISCUSSION

Identification and expression of *BmGnTII* gene

The BLASTP search using the amino acid sequence of human GnTII (GenBank AAH06390.1) as a template found three putative transcriptional isoforms of *B. mori* GnTII ortholog (BmGnTII) (GenBank XP_012548039.1, XP_012548044.1, and XP_012548046.1). The sequence search in KAIKOBase (http://sgp.dna.affrc.go.jp/KAIKO/jp/index.html) (37) indicated that the *BmGnTII* gene was encoded on chromosome 11. The expression profile of *BmGnTII* mRNA was

investigated by RT-PCR using the *BmGnTII*-specific primers and cDNA derived from *B. mori* larvae at different stages and pupa. The specific bands were observed at all stages tested, suggesting that the *BmGnTII* gene was ubiquitously expressed in the larval and pupal stages (Fig. 1A). Moreover, the *BmGnTII* transcripts were detected in all organs tested (Fig. 1B). The DNA encoding full-length BmGnTII was amplified, and the cDNA from a fifth-instar larva was then sequenced (DDBJ/EMBL/GenBank accession number is LC381425).

Sequence analysis of BmGnTII protein

The open reading frame (1371 bp) of *BmGnTII* encodes 456 amino acid residues with a predicted molecular mass of the polypeptide chain of 52.8 kDa and 40.8, 56.4, and 82.2% amino acid sequence identities with human (hGnTII), *Drosophila melanogaster*, and *Spodoptera frugiperda* GnTII (SfGnTII) enzymes, respectively (Fig. 2A), all of which belong to glycosyltransferase family 16 in the CAZy database (http://www.cazy.org/) (38, 39). The cloned BmGnTII did not contain the insertion observed in SfGnTII (24). Golgi-resident glycosyltransferases, including GnTII, are type II transmembrane proteins (40), and the TMHMM server predicted that BmGnTII had a transmembrane region (residues Ala7–Met26) at the N-terminus (Fig. 2B). A dibasic [RK](X)[RK] motif (Arg3-Leu4-Lys5) was found at the N-terminal side of the predicted transmembrane part, which functions as an endoplasmic reticulum export signal in Golgi-localized glycosyltransferases (41).

Before the three-dimensional structure of GnTII was available, hydrophobic cluster analysis suggested that the catalytic domain structure of GnTII resembles that of β-1,2-*N*-acetylglucosaminyltransferase I (GnTI), which belongs to glycosyltransferase family

13 and catalyzes the transfer of GlcNAc from UDP-GlcNAc to the α1-3 arm of the trimannosyl core of *N*-glycan (42, 43). The sequence alignment revealed that BmGnTII had the EED motif, one of the DxD motifs that was proposed to interact with the substrate and metal cofactor in various glycosyltransferases, and it was similar to the EDD motif in GnTI (43) and has been found in all GnTII enzymes identified to date (Fig. 2B). Asp337 of BmGnTII is a fully conserved residue among all known GnTII species, and the corresponding residues were initially predicted to act as a catalytic base (43). In addition, residues involved in UDP-GlcNAc binding that were predicted by comparison with the crystal structure of GnTI (42, 43) are mostly conserved between BmGnTII and the other GnTIIs. These findings suggested that BmGnTII was catalytically active.

Expression and purification of recombinant BmGnTII

To investigate whether BmGnTII had glycosyltransferase activity, the recombinant enzyme (rBmGnTII) with deletion of the predicted cytosolic and transmembrane regions (Met1–Met26) and addition of a bombyxin signal peptide and a FLAG-tag at the N-terminus was constructed and expressed using the silkworm-BmNPV bacmid expression system (Fig. 3A and 3B). rBmGnTII was successfully expressed in silkworm larval hemolymph as a soluble form and had a molecular weight of 60 kDa as estimated by SDS-PAGE (Fig. 3C). Polyethylene glycol precipitation and FLAG-tag affinity chromatography resulted in a single band for rBmGnTII observed in SDS-PAGE analysis (Fig. 3D and 3E) and a yield of approximately 3 μg per larva. The molecular weight of rBmGnTII estimated based on the SDS-PAGE analysis (60 kDa) was higher than its theoretical molecular mass calculated from its amino acid sequence (51 kDa), suggesting that rBmGnTII was glycosylated. NetNGlyc server

(http://www.cbs.dtu.dk/services/NetNGlyc/) analysis showed that BmGnTII has nine potential *N*-glycosylation sites (Asn45, Asn69, Asn73, Asn82, Asn110, Asn245, Asn319, Asn384, and Asn441), which were fully conserved with SfGnTII, whereas hGnTII has only two sites (Fig. 2B) (13). Deglycosylation of rBmGnTII by PNGase F resulted in a decrease of its molecular size to approximately 52 kDa (Fig. 3D and 3E), which was similar to the theoretical molecular weight (51 kDa). These results indicated that rBmGnTII was *N*-glycosylated as recombinant SfGnTII expressed in Sf9 cells (25).

Activity and substrate specificity of rBmGnTII

To test whether rBmGnTII catalyzed the glycosyltransferase reaction, the enzyme was incubated with UDP-GlcNAc and pyridylaminated MGn glycan (MGn-PA), which were the substrates of the known GnTII, and then, the products were analyzed using reverse phase HPLC. rBmGnTII catalyzed the transfer of GlcNAc from UDP-GlcNAc to MGn-PA glycan to produce GnGn-PA glycan in 50 mM HEPES-NaOH buffer (pH 6.5) containing 10 mM $MnCl_2$ at 37ºC (Fig. 4). The specific activity of rBmGnTII in this condition was 12.6±0.5 nmol/min/mg, which was comparable to that of recombinant hGnTII expressed in silkworm larvae (13). In contrast, no product was observed in the reaction employing GnM-PA, GnGn-PA, and MM-PA as acceptor substrates (Fig. 4), although recombinant SfGnTII transferred a second GlcNAc residue to GnGn-PA with a long incubation period (25). Further investigation of its substrate specificity was performed using *p*-nitrophenyl monosaccharides as acceptor substrates, but rBmGnTII did not transfer the GlcNAc residue from UDP-GlcNAc to the acceptors (data not shown). These results suggested that rBmGnTII had strict substrate specificity, which is similar to characterized mammalian GnTIIs (13).

The effect of pH on the glycosyltransferase activity was examined using UDP-GlcNAc and MGn-PA as substrates over a pH range of 4.0–11.0. The enzyme had the optimum pH of 6.5 (Fig. 5A), which was similar to recombinant SfGnTII (25). The effect of temperature was tested at 20–70ºC, and the enzyme exhibited the highest activity at 37°C (Fig. 5B). Because glycosyltransferases possessing a DxD motif usually use a divalent metal ion as a cofactor to bind the sugar nucleotide substrate (40), the activity was investigated in the presence of several divalent metal chlorides. rBmGnTII had the highest activity under the condition with Mn^{2+}, whereas the enzyme exhibited 30% activity in the presence of Co^{2+} (Fig. 5C). These properties were almost identical to those of SfGnTII and hGnTII (13, 25). Furthermore, because BmGnTII has several *N*-glycosylation sites in its predicted catalytic domain (unlike hGnTII), we tested whether *N*-glycans are involved in the activity of deglycosylated rBmGnTII. Unexpectedly, PNGase F-treated rBmGnTII showed activity (99.5±1.6%) comparable to that of intact rBmGnTII (Fig. 5D) although rBmGnTII was deglycosylated in native condition (Fig. 3E), suggesting that *N*-glycosylation did not affect the catalytic reaction.

Very recently, the crystal structure of hGnTII was determined (44), allowing us to discuss the structure-function relationship of GnTII in more detail. The structure of hGnTII adopts a Rossmann-like fold resembling the catalytic domain of GnTI. Homology modeling of BmGnTII was performed using the SWISS-MODEL server and the coordinate of hGnTII (PDB 5VCM) as a template, and it indicated that BmGnTII has the same fold and active site architecture as hGnTII (Fig. S2). Compared with the structure of hGnTII, the active site residues, including the catalytic base (Asp342) and Mn^{2+}-binding sites (Asp256 and His370), are highly conserved in BmGnTII (Fig. 2B). In particular, the exosite pocket that binds the GlcNAc residue on the α1-3 arm of the

substrate is completely conserved in BmGnTII, agreeing with their activity and strict substrate specificity. Based on the homology model of BmGnTII, the catalytic domain has six potential *N*-glycosylation sites, four of which are located at the opposite side of the active site (Fig. 6). Further study is necessary to elucidate which sites are *N*-glycosylated and what the glycans contribute to.

In this study, we investigated the expression profile and enzymatic activity of BmGnTII to evaluate its physiological function and possibility to facilitate complex-type *N*-glycoprotein biosynthesis in silkworm. Although the GnTII activity in cultured Bm-N cells was low (30) and biantennary GlcNAcylated glycans were rarely observed in recombinant glycoproteins expressed in silkworm larvae and pupae (18), the transcriptional analysis in this study suggested that the *BmGnTII* gene is ubiquitously expressed at the larval and pupal stages and in several organs of larva including fat body and silk gland. Moreover, rBmGnTII had enzymatic properties comparable to those of hGnTII. These results suggested that BmGnTII could not efficiently GlcNAcylate glycan substrates because FDL may remove the GlcNAc residue of the α1-3 arm that is essential for BmGnTII activity. These findings also raised the possibility that the *BmGnTII* and *fdl* expression level may vary among different cells and tissues and/or that the translation of BmGnTII protein may be regulated. The previous study demonstrated that the ortholog SfGnTII expressed under a non-native *ie1* promoter control was reportedly located in the Golgi apparatus together with GnTI and FDL in Sf9 cells (25). Thus, further functional analyses using RNA silencing and cellular biological experiments are required to investigate the expression and localization of native BmGnTII enzyme in *B. mori* cells and tissues. Moreover, transgenic technologies will perhaps enhance the endogenous expression of BmGnTII, facilitating the production of biantennary complex-type *N*-

glycoproteins in silkworm.

ACKNOWLEDGMENTS

This work was supported by the Japan Society for the Promotion of Science KAKENHI (TM: grant no. 16H06847 and EYP: grant no. 16H02544).

REFERENCES

1. **Helenius, A. and Aebi, M.:** Intracellular functions of N-linked glycans, Science, **291**, 2364–2369 (2001).

2. **Moremen, K.W. and Molinari, M.:** *N*-linked glycan recognition and processing: the molecular basis of endoplasmic reticulum quality control, Curr. Opin. Struct. Biol., **16**, 592–599 (2006).

3. **Moremen, K.W., Tiemeyer, M., and Nairn, A.V.:** Vertebrate protein glycosylation: diversity, synthesis and function, Nat. Rev. Mol. Cell. Biol., **13**, 448–462 (2012).

4. **Miyazaki, T., Matsumoto, Y., Matsuda, K., Kurakata, Y., Matsuo, I., Ito, Y., Nishikawa, A., and Tonozuka, T.:** Heterologous expression and characterization of processing α-glucosidase I from *Aspergillus brasiliensis* ATCC 9642, Glycoconj. J., **28**, 563–571 (2011).

5. **Walski, T., De Schutter, K., Van Damme, E.J.M., and Smagghe, G.:** Diversity and functions of protein glycosylation in insects, Insect Biochem. Mol. Biol., **83**, 21–34 (2017).

6. **Schoberer, J., and Strasser, R.:** Plant glyco-biotechnology, Semin. Cell Dev. Biol., **80**, 133–141 (2018).

7. **Jarvis, D.L.:** Baculovirus-insect cell expression systems, Methods Enzymol., **463**, 191–222, (2009).

8. **Geisler, C., Mabashi-Asazuma, H., and Jarvis, D.L.:** An overview and history of glyco-engineering in insect expression systems, Methods Mol. Biol., **1321**, 131–152 (2015).

9. **Léonard, R., Rendic, D., Rabouille, C., Wilson, I.B., Préat, T., and Altmann, F.:** The *Drosophila fused lobes* gene encodes an *N*-acetylglucosaminidase involved in *N*-glycan processing, J. Biol. Chem., **281**, 4867–4875 (2006).

10. **Geisler, C., Aumiller, J.J., and Jarvis, D.L.:** A fused lobes gene encodes the processing β-*N*-acetylglucosaminidase in Sf9 cells, J. Biol. Chem., **283**, 11330–11339 (2008).

11. **Thaysen-Andersen, M., Venkatakrishnan, V., Loke, I., Laurini, C., Diestel, S., Parker, B.L., and Packer, N.H.:** Human neutrophils secrete bioactive paucimannosidic proteins from azurophilic granules into pathogen-infected sputum, J. Biol. Chem., **290**, 8789–8802 (2015).

12. **Bendiak, B., and Schachter, H.:** Control of glycoprotein synthesis. Purification of UDP-*N*-acetylglucosamine:α-D-mannoside β1-2 *N*-acetylglucosaminyltransferase II from rat liver, J. Biol. Chem., **262**, 5775–5783 (1987).

13. **Miyazaki, T., Kato, T., and Park, E.Y.:** Heterologous expression, purification and characterization of human β-1,2-*N*-acetylglucosaminyltransferase II using a silkworm-based *Bombyx mori* nucleopolyhedrovirus bacmid expression system, J. Biosci. Bioeng., **126**, 15–22 (2018).

14. **Palmberger, D., Wilson, I.B., Berger, I., Grabherr, R., and Rendic, D.:** SweetBac: a new approach for the production of mammalianised glycoproteins in insect cells, PLoS

One, **7**, e34226 (2012).

15. **Hollister, J., Grabenhorst, E., Nimtz, M., Conradt, H., and Jarvis, D.L.:** Engineering the protein *N*-glycosylation pathway in insect cells for production of biantennary, complex *N*-glycans, Biochemistry, **41**, 15093–15104 (2002).

16. **Mabashi-Asazuma, H., Shi, X., Geisler, C., Kuo, C.W., Khoo, K.H., and Jarvis, D.L.:** Impact of a human CMP-sialic acid transporter on recombinant glycoprotein sialylation in glycoengineered insect cells, Glycobiology, **23**, 199–210 (2013).

17. **Mabashi-Asazuma, H., Sohn, B.H., Kim, Y.S., Kuo, C.W., Khoo, K.H., Kucharski, C.A., Fraser, M.J. Jr., and Jarvis, D.L.:** Targeted glycoengineering extends the protein N-glycosylation pathway in the silkworm silk gland, Insect Biochem. Mol. Biol., **65**, 20–27 (2015).

18. **Kato, T., Kako, N., Kikuta, K., Miyazaki, T., Kondo, S., Yagi, H., Kato, K., and Park, E.Y.:** *N*-Glycan modification of a recombinant protein via coexpression of human glycosyltransferases in silkworm pupae, Sci. Rep., **7**, 1409 (2017).

19. **Kim, Y.K., Kim, K.R., Kang, D.G., Jang, S.Y., Kim, Y.H., and Cha, H.J.:** Suppression of β-*N*-acetylglucosaminidase in the *N*-glycosylation pathway for complex glycoprotein formation in *Drosophila* S2 cells, Glycobiology, **19**, 301–308 (2009).

20. **Nomura, T., Suganuma, M., Higa, Y., Kataoka, Y., Funaguma, S., Okazaki, H., Suzuki, T., Kobayashi, I., Sezutsu, H., and Fujiyama, K.:** Improvement of glycosylation structure by suppression of β-*N*-acetylglucosaminidases in silkworm, J. Biosci. Bioeng., **119**, 131–136 (2015).

21. **Mabashi-Asazuma, H., Kuo, C.W., Khoo, K.H., and Jarvis, D.L.:** Modifying an insect cell *N*-glycan processing pathway using CRISPRCas technology, ACS Chem. Biol., **10**, 2199–2208 (2015).

22. **Haines, N., and Irvine KD:** Functional analysis of *Drosophila* β1,4-N-acetlygalactosaminyltransferases, Glycobiology, **15**, 335–346 (2005).

23. **Vadaie, N., and Jarvis, D.L.:** Molecular cloning and functional characterization of a Lepidopteran insect β4-*N*-acetylgalactosaminyltransferase with broad substrate specificity, a functional role in glycoprotein biosynthesis, and a potential functional role in glycolipid biosynthesis, J. Biol. Chem., **279**, 33501–33518 (2004).

24. **Koles, K., Irvine, K.D., and Panin, V.M.:** Functional characterization of *Drosophila* sialyltransferase, J. Biol. Chem., **279**, 4346–4357 (2004).

25. **Geisler, C. and Jarvis, D.L.:** Substrate specificities and intracellular distributions of three *N*-glycan processing enzymes functioning at a key branch point in the insect *N*-glycosylation pathway, J. Biol. Chem., **283**, 11330–11339 (2008).

26. **Kajiura, H., Hamaguchi, Y., Mizushima, H., Misaki, R., and Fujiyama, K.:** Sialylation potentials of the silkworm, *Bombyx mori*; *B. mori* possesses an active α2,6-sialyltransferase, Glycobiology, **25**, 1441–1453 (2015).

27. **Koles, K., Lim, J.M., Aoki, K., Porterfield, M., Tiemeyer, M., Wells, L., and Panin, V.:** Identification of *N*-glycosylated proteins from the central nervous system of *Drosophila melanogaster*, Glycobiology, **17**, 1388–1403 (2007).

28. **Aoki, K., Perlman, M., Lim, J.M., Cantu, R., Wells, L., and Tiemeyer, M.:** Dynamic developmental elaboration of *N*-linked glycan complexity in the *Drosophila melanogaster* embryo, J. Biol. Chem., **282**, 9127–9142 (2007).

29. **Stanton, R., Hykollari, A., Eckmair, B., Malzl, D., Dragosits, M., Palmberger, D., Wang, P., Wilson, I.B., and Paschinger, K.:** The underestimated N-glycomes of lepidopteran species, Biochim. Biophys. Acta, **1861**, 699–714 (2017).

30. **Altmann, F., Kornfeld, G., Dalik, T., Staudacher, E., and Glössl, J.:** Processing

of asparagine-linked oligosaccharides in insect cells. *N*-Acetylglucosaminyltransferase I and II activities in cultured lepidopteran cells, Glycobiology, **3**, 619–625 (1993).

31. **Misaki, R., Nagaya, H., Fujiyama, K., Yanagihara, I., Honda, T., and Seki, T.:** N-linked glycan structures of mouse interferon-beta produced by *Bombyx mori* larvae, Biochem. Biophys. Res. Commun., **311**, 979–986 (2003).

32. **Iizuka, M., Ogawa, S., Takeuchi, A., Nakakita, S., Kubo, Y., Miyawaki, Y., Hirabayashi, J., and Tomita M.:** Production of a recombinant mouse monoclonal antibody in transgenic silkworm cocoons, FEBS J., **276**, 5806–5820 (2009).

33. **Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L.:** Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol., **305**, 567–580 (2001).

34. **Park, E.Y., Abe, T., and Kato, T.:** Improved expression of fusion protein using a cysteine-protease- and chitinase-deficient *Bombyx mori* (silkworm) multiple nucleopolyhedrovirus bacmid in silkworm larvae, Biotechnol. Appl. Biochem., **49**, 135–140 (2008).

35. **Kato, T., Arai, S., Ichikawa, H., and Park E.Y.:** Versatility of chitosan/BmNPV bacmid DNA nanocomplex as transfection reagent of recombinant protein expression in silkworm larvae, Biotechnol. Lett., **38**, 1449–1457 (2016).

36. **Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L., and Schwede, T.:** SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucleic Acids Res., **42**, W252–W258 (2014).

37. **Shimomura, M., Minami, H., Suetsugu, Y., Ohyanagi, H., Satoh, C., Antonio, B., Nagamura, Y., Kadono-Okuda, K., Kajiwara, H., Sezutsu, H., Nagaraju, J.,**

Goldsmith, M.R., Xia, Q., Yamamoto, K., and Mita, K.: KAIKObase: an integrated silkworm genome database and data mining tool, BMC Genomics, **10**, 486 (2009).

38. **Coutinho, P.M., Deleury, E., Davies, G.J., and Henrissat, B.:** An evolving hierarchical family classification for glycosyltransferases, J. Mol. Biol., **328**, 307–317 (2003).

39. **Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B.:** The carbohydrate-active enzymes database (CAZy) in 2013, Nucleic Acids Res., **42**, D490–495 (2014).

40. **Breton, C., and Imberty, A.:** Structure/function studies of glycosyltransferases. Curr. Opin. Struct. Biol., **9**, 563–571 (1999).

41. **Giraudo, C.G., and Maccioni, H.J.:** Endoplasmic reticulum export of glycosyltransferases depends on interaction of a cytoplasmic dibasic motif with Sar1. Mol. Biol. Cell., **14**, 3753–3766 (2003).

42. **Ünligil, U.M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., and Rini, J.M.:** X-ray crystal structure of rabbit *N*-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily, EMBO J., **19**, 5269–5280 (2000).

43. **Breton, C., Mucha, J., and Jeanneau, C.:** Structural and functional features of glycosyltransferases, Biochimie, **83**, 713–718 (2001).

44. **Kadirvelraj, R., Yang, J.-Y., Sanders, J.H., Liu, L., Ramiah, A., Prabhakar, P.K., Boons, G.-J., Wood, Z.A., and Moremen, K.W.:** Human *N*-acetylglucosaminyltransferase II substrate recognition uses a modular architecture that includes a convergent exosite, Proc. Natl. Acad. Sci. U. S. A., **115**, 4637–4642 (2018).

45. **Edgar, R.C.:** MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res., **32**, 1792–1797 (2004).

46. **Kumar, S., Stecher, G., and Tamura, K.:** MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets, Mol. Biol. Evol., **33**, 1870–1874 (2016).

47. **Robert, X., and Gouet, P.:** Deciphering key features in protein structures with the new ENDscript server, Nucleic Acids Res., **42**, W320–W324 (2014).

FIGURE LEGENDS

FIG. 1. RT-PCR analysis of *BmGnTII* mRNA expression in the larval and pupal stages. DNA fragments were amplified by PCR using gene-specific primers (Table 1) and the cDNA templates that were synthesized from total RNAs extracted from whole bodies of first- to fifth-instar larvae and pupa (A) and tissues of fifth-instar larvae (B). Tissues used are as follows: fat body (FB), midgut (Mg), silk gland (SG), epidermis (Ep), and Malpighian tubule (MT). *BmGAPDH* was used as a control.

FIG. 2. Phylogenetic tree (A) and multiple sequence alignment (B) of the identified GnTII enzymes. The amino acid sequences were aligned using the MUSCLE program (45), and the phylogenetic tree was generated using MEGA7 (46). The sequences used were from *Homo sapiens* (GenBank AAH06390.1), *Rattus norvegicus* (AAA86721.1), *Sus scrofa* (CAA70732.1), *Xenopus laevis* (CAD56909.1), *Drosophila melanogaster* (AAX53005.1), *Bombyx mori* (this study), *Spodoptera frugiperda* (AEX00083.1), *Caenorhabditis elegans* (AAF71273.1), and *Arabidopsis thaliana* (CAC08806.1). **(B)** The figure was generated using ESPript 3.0 (47). *N*-glycosylation sequons are highlighted in green, and predicted transmembrane regions are underlined with a dashed line. The residues involved in catalytic activity are indicated with triangles of the following colors: orange, UDP-GlcNAc donor-binding; magenta, acceptor substrate-binding; blue, manganese ion-binding; and cyan, catalytic base. The secondary structures and disulfide bonds (green numerals) are also indicated according to the crystal structure of hGnTII (PDB 5VCM).

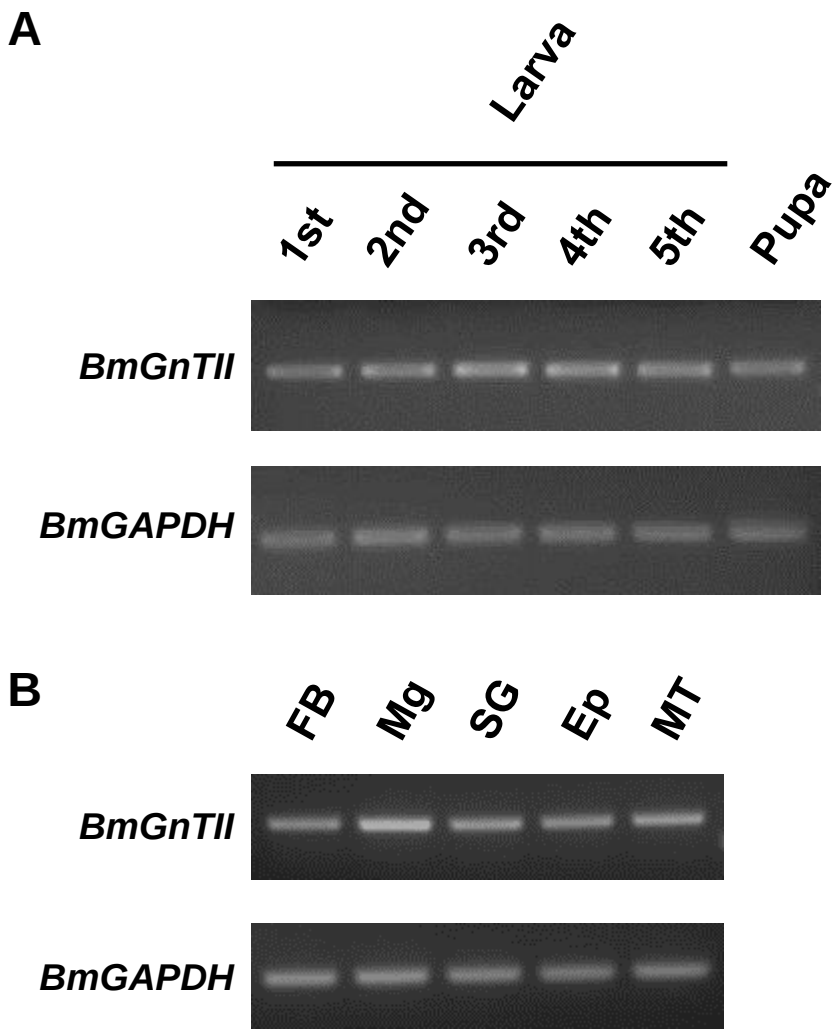
FIG. 3. Expression and purification of the soluble form of BmGnTII. **(A)** Predicted domain organization of BmGnTII and deletion of its N-terminal sequence containing the cytosolic (residues 1–7) and transmembrane (TM, residues 7–27) regions. **(B)** Construction of the expression vector for rBmGnTII containing the stem region and the catalytic domain. Abbreviations: *polh*, polyhedrin promoter; bx, bombyxin secretion signal; and F, FLAG-tag. EcoRI and KpnI were used for subcloning into the pFastBac1 vector. **(C)** Western blotting analysis of the expression of rBmGnTII in silkworm larvae using an anti-FLAG-tag antibody. Lane 1, molecular weight marker; lane 2, hemolymph from mock silkworm; lane 3, fat body from mock silkworm; lane 4, hemolymph from bacmid-injected silkworm; and lane 5, fat body from bacmid-injected silkworm. **(D and E)** SDS-PAGE analysis of the purified rBmGnTII with PNGase F treatment after denaturation **(D)** and under the native condition **(E)**. Lane 1, molecular weight marker; lane 2, PNGase F-treated rBmGnTII; and lane 3, intact rBmGnTII. Intact and deglycosylated rBmGnTIIs are indicated with black and white arrows, respectively.

FIG. 4. HPLC analysis of the reaction products generated by rBmGnTII. Purified rBmGnTII was incubated with MGn-PA, GnGn-PA, GnM-PA, or MM-PA in the presence of Mn^{2+} and UDP-GlcNAc at pH 6.5 for 1 min (MGn-PA) or 12 h (GnGn-PA, GnM-PA, and MM-PA), and then, the reaction mixtures were analyzed by reverse-phase HPLC.

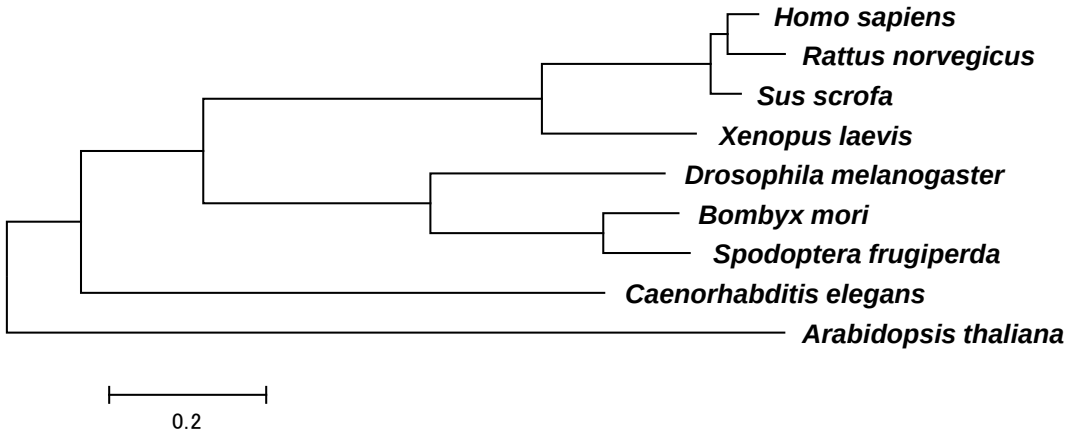
FIG. 5. General enzymatic properties of rBmGnTII. **(A)** pH dependence, **(B)** temperature dependence, **(C)** metal ion dependence and **(D)** effect of deglycosylation. The pH dependence was measured at 37ºC using 50 mM sodium acetate (pH 4.0–6.0, *circle*), HEPES-NaOH (pH 6.0–8.0, *triangle*), Tris-HCl (pH 8.0–9.0, *square*), or glycine-NaOH

(pH 9.0–11.0, *diamond*) buffers. The temperature dependence was measured at 20–70ºC using 50 mM HEPES-NaOH (pH 6.5). The effect of divalent metal ions was investigated at 37ºC using 50 mM HEPES-NaOH buffer (pH 6.0) containing 10 mM of each metal chloride. **(D)** Activity of the enzymes pre-incubated with (+) and without (−) PNGase F at 37ºC for 1 h. The activity of the recombinant BmGnTII without pre-incubation was taken to be 100%. Error bars represent the standard deviation in triplicate experiments.

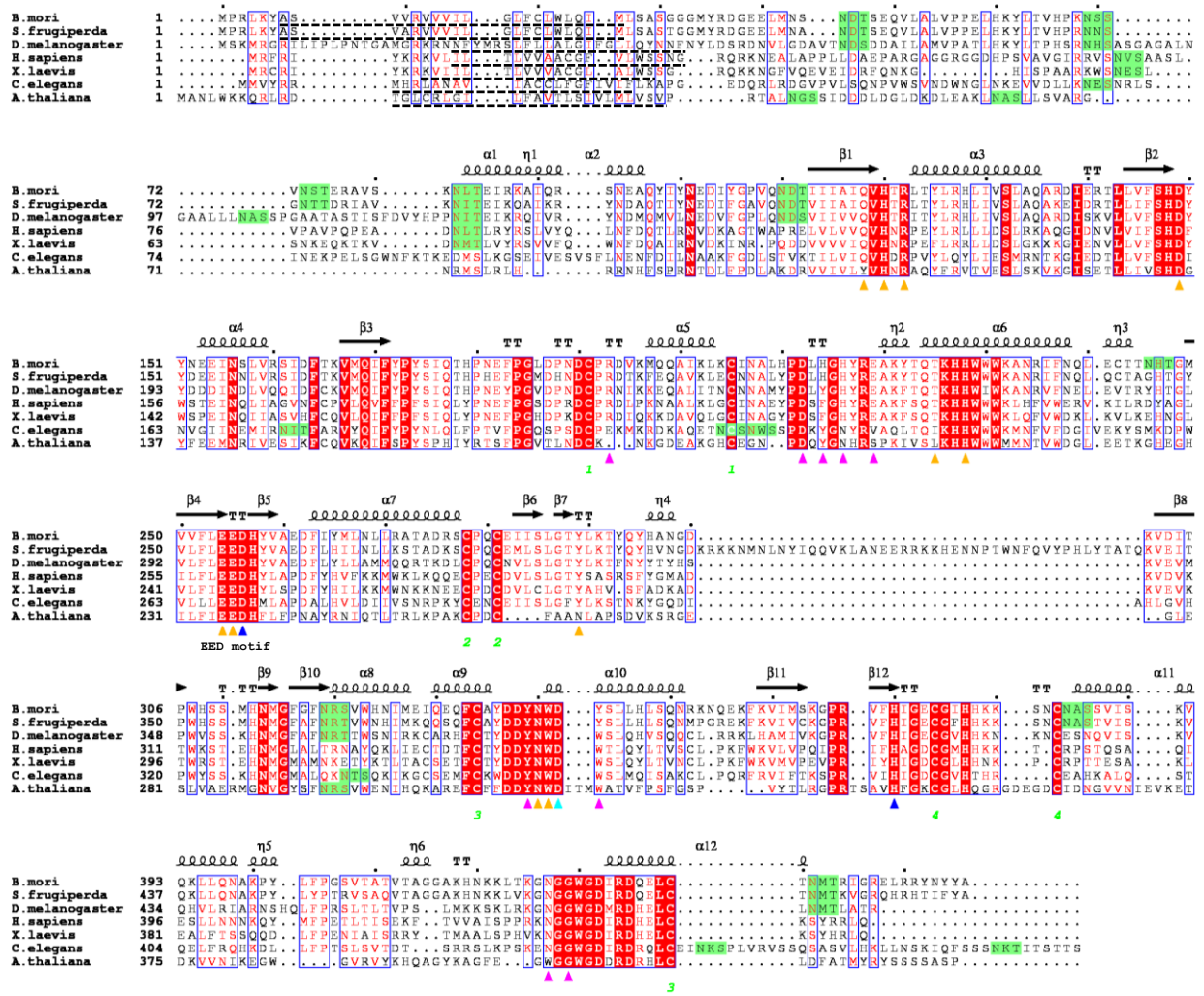
FIG. 6. *N*-glycosylation sites in BmGnTII. The homology model of BmGnTII catalytic domain was generated using the SWISS-MODEL server with the coordinate of hGnTII (PDB 5VCM). Putative *N*-glycosylation sites are highlighted in green. The models of UDP (yellow stick), an acceptor substrate (magenta stick) and Mn^{2+} ion (purple sphere) in the crystal structures of hGnTII complexed with UDP and Mn^{2+} (PDB 5VCM) or the acceptor (PDB 5VCS) are superposed into BmGnTII model.

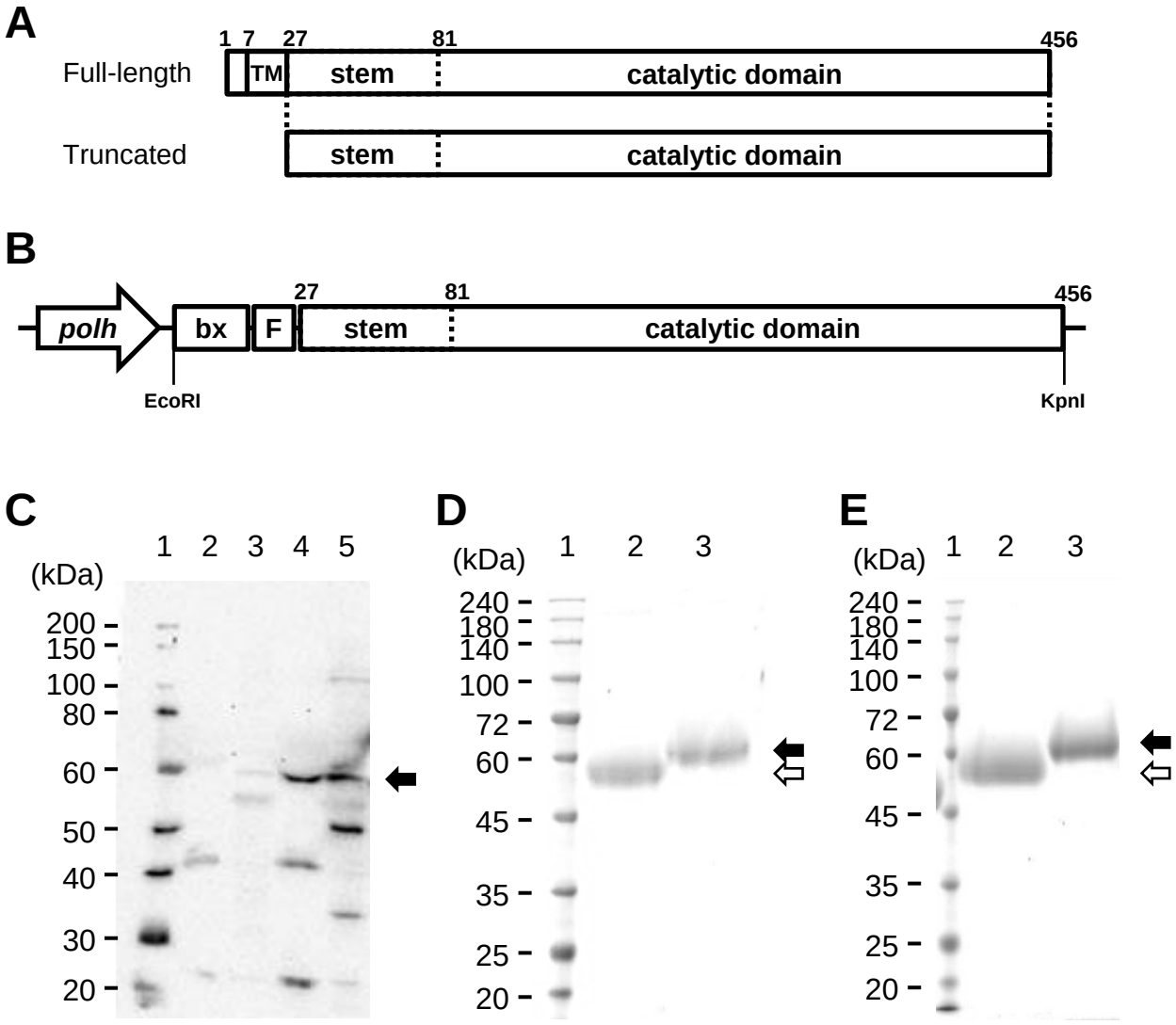

A
Larva
1st
2nd
3rd
4th
5th
Pupa
BmGnTII
BmGAPDH
B
FB
Mg
SG
Ep
MT
BmGnTII
BmGAPDH

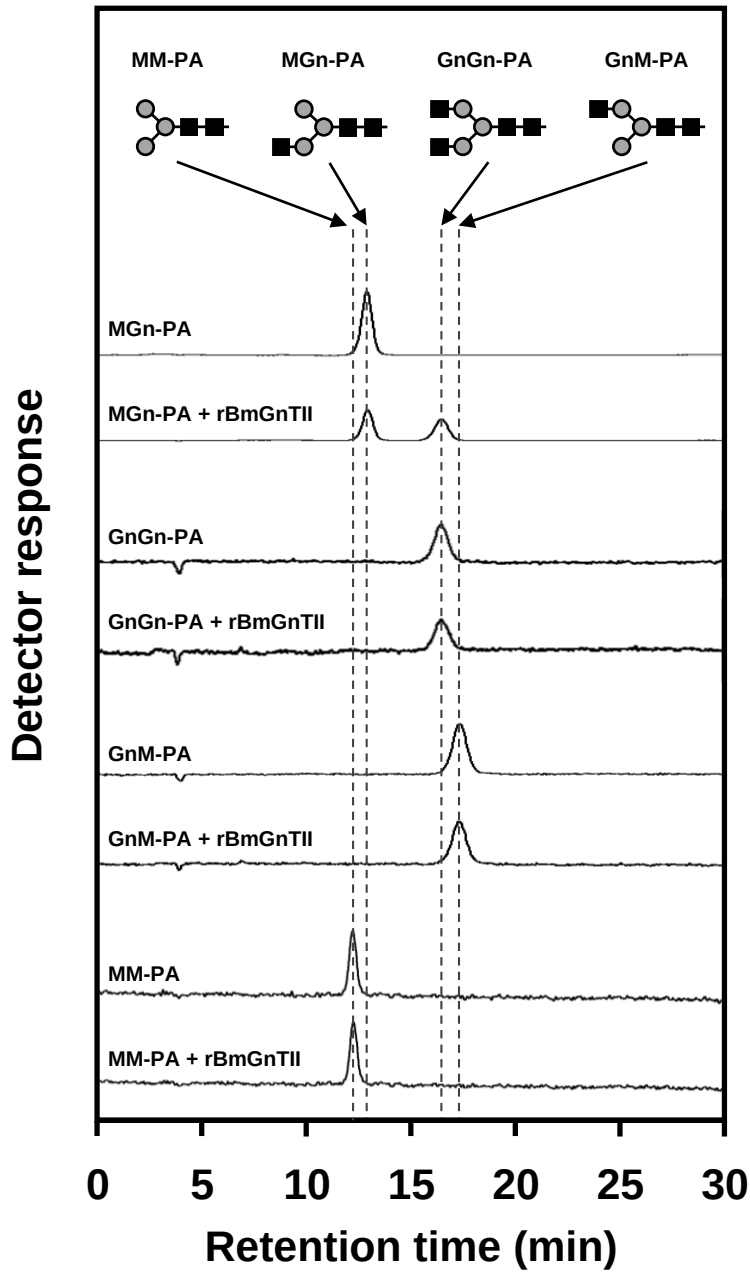
A

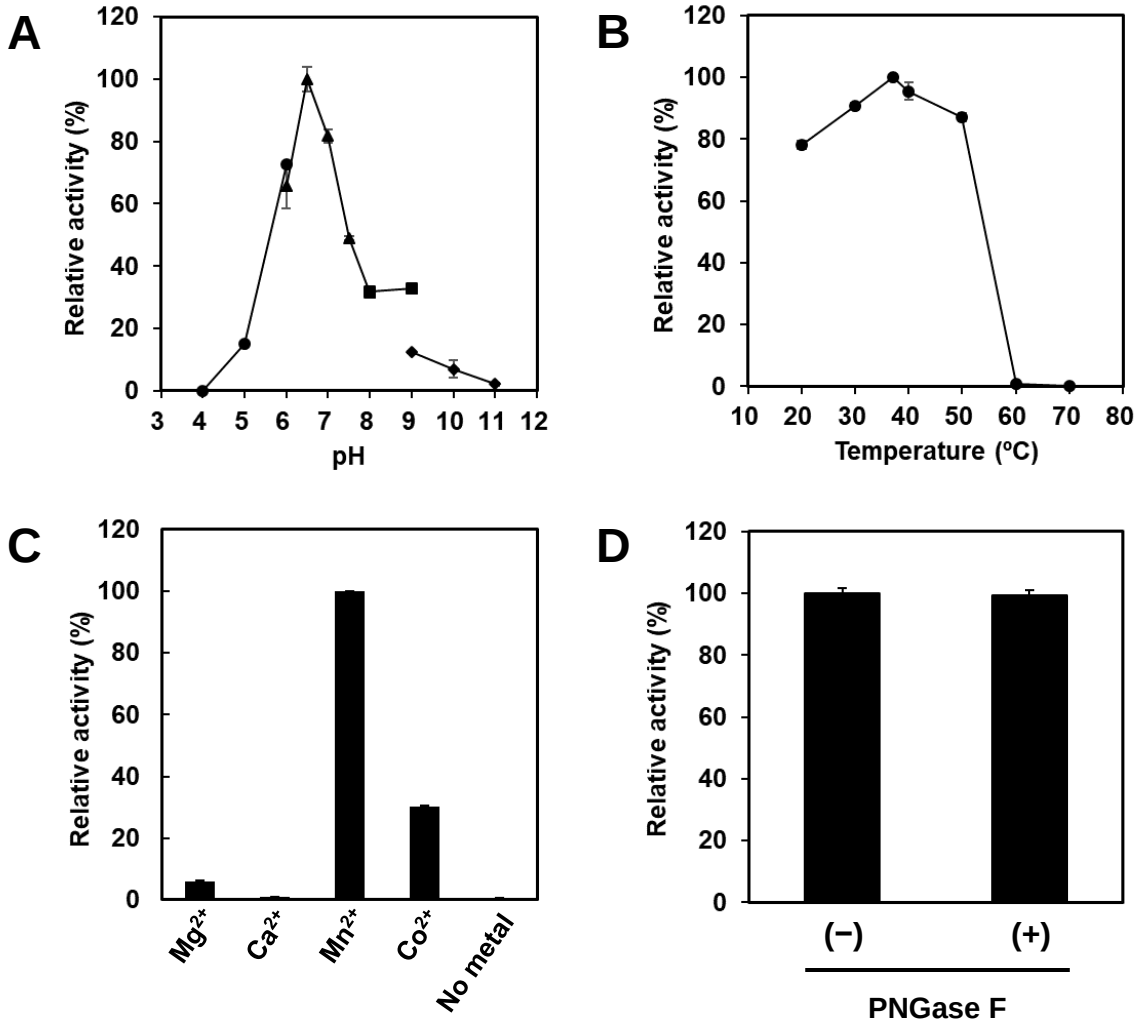

Homo sapiens
Rattus norvegicus
Sus scrofa
Xenopus laevis
Drosophila melanogaster
Bombyx mori
Spodoptera frugiperda
Caenorhabditis elegans
Arabidopsis thaliana
0.2

B


B.mori
S.frugiperda
D.melanogaster
H.sapiens
X.laevis
C.elegans
A.thaliana
EED motif


A
1 7 27 81 456
Full-length
TM
stem
catalytic domain
Truncated
stem
catalytic domain
B
27 81 456
polh
bx
F
stem
catalytic domain
EcoRI
KpnI
C
1 2 3 4 5
(kDa)
200
150
100
80
60
50
40
30
20
D
1 2 3
(kDa)
240
180
140
100
72
60
45
35
25
20
E
1 2 3
(kDa)
240
180
140
100
72
60
45
35
25
20


MM-PA
MGn-PA
GnGn-PA
GnM-PA
MGn-PA
MGn-PA + rBmGnTII
GnGn-PA
GnGn-PA + rBmGnTII
GnM-PA
GnM-PA + rBmGnTII
MM-PA
MM-PA + rBmGnTII
Detector response
0
5
10
15
20
25
30
Retention time (min)


A
Relative activity (%)
120
100
80
60
40
20
0
3 4 5 6 7 8 9 10 11 12
pH
B
Relative activity (%)
120
100
80
60
40
20
0
10 20 30 40 50 60 70 80
Temperature (°C)
C
Relative activity (%)
120
100
80
60
40
20
0
Mg2+
Ca2+
Mn2+
Co2+
No metal
D
Relative activity (%)
120
100
80
60
40
20
0
(−)
(+)
PNGase F

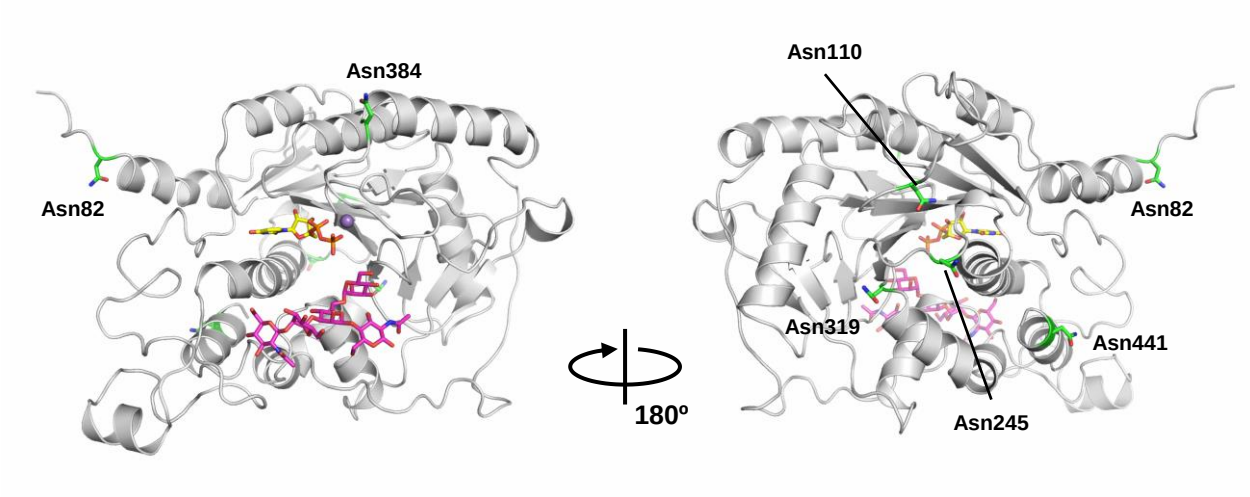

Asn384
Asn82
180º
Asn110
Asn82
Asn319
Asn441
Asn245

Table 1. Primers used in this study.

Primer	Sequence
BmGAPDH_F	5’-GCTGGAATTTCTTTGAATGAC-3’
BmGAPDH_R	5’-CAATGACTCTGCTGGAATAACC-3’
BmGnTII_F	5’-GCTCGAGACATCGAGAGGAC-3’
BmGnTII_R	5’-ATCTAAGCCCGGGAACTCAT-3’
FLAG-BmGnTII_F	5’-GATTACAAGGATGACGACGATAAGCTGTCTGCTTCTGGCGGG-3’
BmGnTII_KpnI_ R	5’-GCGGTACCTCAAGCGTAGTAATTGTATCTTC-3’
EcoRI-Bx-FLAG_F	5’-GCGAATTCATGAAGATACTCCTTGCTATTGCATTAATGTTGTCAACAGTAATGTGGGTGTCAACAGATTACAAGGATGACGACGATAAG-3’
pUC/M13 forward	5’-CCCAGTCACGACGTTGTAAAACG-3’
pUC/M13 reverse	5’-AGCGGATAACAATTTCACACAGG-3’

Restriction enzyme sites are underlined.

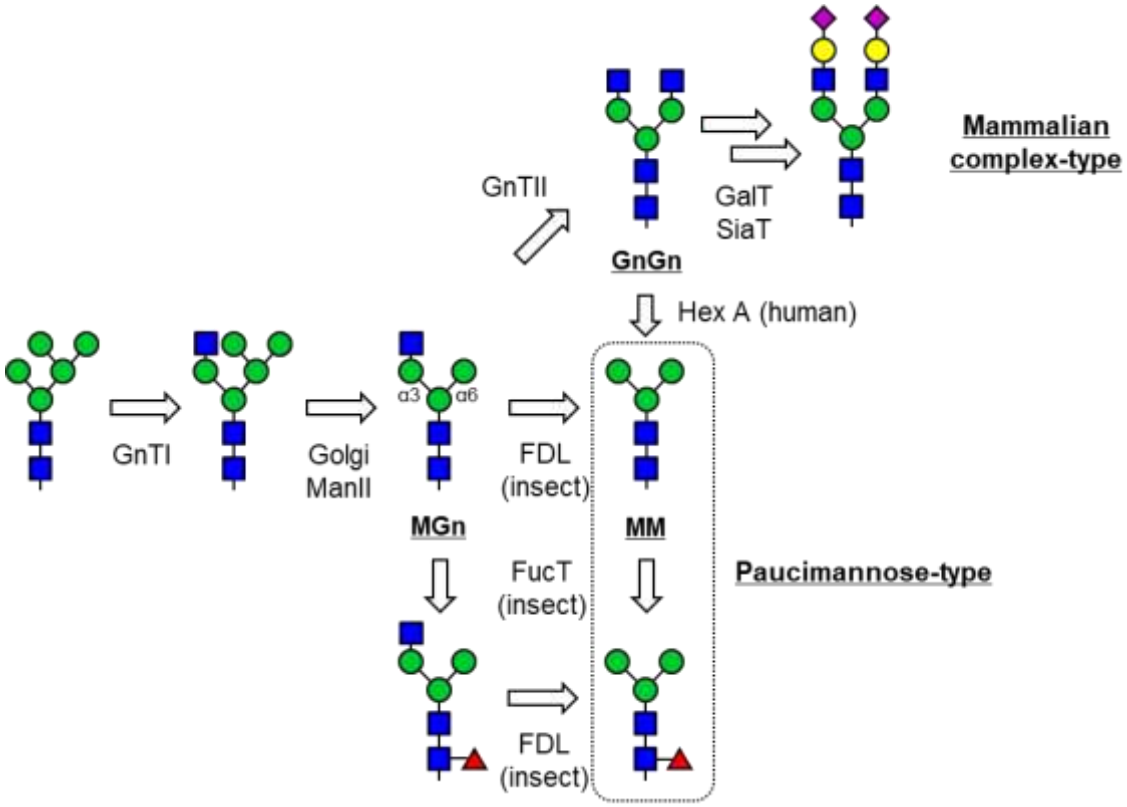
Supplementary Information

Expression and characterization of silkworm *Bombyx mori* β-1,2-*N*-acetylglucosaminyltransferase II, a key enzyme for complex-type *N*-glycan biosynthesis.

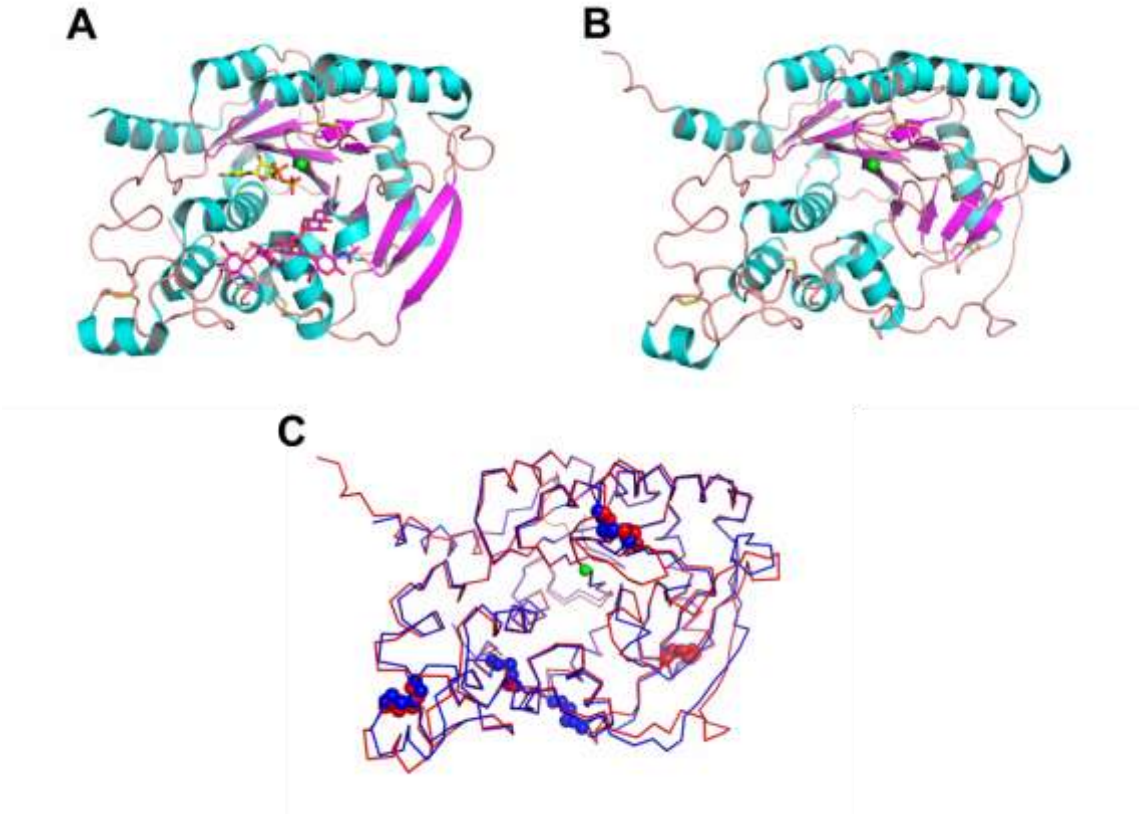
Takatsugu Miyazaki[1,2], Ryunosuke Miyashita[2], Sota Mori[2], Tatsuya Kato[1,2], and Enoch Y. Park[1,2,*]

[1]Laboratory of Biotechnology, Research Institute of Green Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka, 422-8529, Japan.
[2]Laboratory of Biotechnology, Department of Agriculture, Graduate School of Integrated Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529, Japan.

*__Correspondence:__ E.Y. Park, Laboratory of Biotechnology, Research Institute of Green Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka, 422-8529, Japan; Tel & Fax: +81-54-238-4887; E-mail: park.enoch@shizuoka.ac.jp



Supplementary Fig. S1. Biosynthesis pathway of complex-type and paucimannose-type *N*-glycans. Both insect and mammalian cells produce the same glycan intermediate MGn in the early steps of *N*-glycan processing. In mammalian cells, β-1,2-*N*-acetylglucosaminyltransferase (GnTII) transfers *N*-acetylglucosamine (GlcNAc) to α1-6 arm of MGn glycan to produce biantennary GnGn glycan, followed by β-1,4-galactosyltransferase (GalT)- and sialyltransferase (SiaT)-catalyzed modifications to generate complex-type glycans. In the case of insect cells, *N*-acetylglucosaminidase (FDL) removes the terminal GlcNAc from MGn glycan to produce paucimannose-type glycans and fucosyltransferases (FucT) add fucose residue(s) on an innermost GlcNAc residue. The core-fucosylation also occurs in mammalian cells, which is omitted for clarity. Recent studies have revealed that paucimannosidic *N*-glycans are produced in human neutrophils by hydrolysis of hexosaminidase A (Hex A) (1). Symbols: green circle, mannose; blue square, GlcNAc; yellow circle, galactose; magenta diamond, sialic acid; and red triangle, fucose.



Supplementary Fig. S2. Structures of GnTII. (A) Structure of hGnTII in complex with UDP (yellow and orange stick), manganese ion (green sphere), and an acceptor substrate Man(α1-6)[GlcNAc(β1-2)Man(α1-3)]ManGlcNAc (magenta stick). The coordinates of hGnTII complexed with UDP and Mn^{2+} (PDB 5VCM) and the model of the glycan derived from the coordinate of hGnTII complexed with the acceptor (PDB 5VCS) are superimposed. (B) Homology model of BmGnTII generated by the SWISS-MODEL server (https://swissmodel.expasy.org) using the coordinate 5VCM as a template. (C) Superimposition of the main chains of hGnTII (blue) and BmGnTII (red) model. Disulfide bonds are indicated as sphere model.

REFERENCES

1. **Thaysen-Andersen, M., Venkatakrishnan, V., Loke, I., Laurini, C., Diestel, S., Parker, B.L., and Packer, N.H.:** Human neutrophils secrete bioactive paucimannosidic

proteins from azurophilic granules into pathogen-infected sputum, J. Biol. Chem., **290**, 8789–8802 (2015).