Glycoconjugate Journal 16, 787–791, 1999

Molecular cloning of cDNA encoding *N*-acetylglucosaminyltransferase II from *Arabidopsis thaliana*

R. Strasser[1], H. Steinkellner[1]*, M. Borén[1], F. Altmann[2], L. Mach[1], J. Glössl[1] and J. Mucha[1]

[1]*Zentrum für Angewandte Genetik, Universität für Bodenkultur Wien, Muthgasse 18, 1190 Wien, Austria,*
[2]*Institut für Chemie, Universität für Bodenkultur Wien, Muthgasse 18, 1190 Wien, Austria*

***N*-acetylglucosaminyltransferase II (GnTII, EC 2.4.1.143) is a Golgi enzyme involved in the biosynthesis of glycoprotein-bound N-linked oligosaccharides, catalysing an essential step in the conversion of oligomannose-type to complex N-glycans. GnTII activity has been detected in both animals and plants. However, while cDNAs encoding the enzyme have already been cloned from several mammalian sources no GnTII homologue has been cloned from plants so far. Here we report the molecular cloning of an *Arabidopsis thaliana* GnTII cDNA with striking homology to its animal counterparts. The predicted domain structure of *A. thaliana* GnTII indicates a type II transmembrane protein topology as it has been established for the mammalian variants of the enzyme. Upon expression of *A. thaliana* GnTII cDNA in the baculovirus/insect cell system, a recombinant protein was produced that exhibited GnTII activity.**

***Keywords:* *N*-acetylglucosaminyltransferase II, *Arabidopsis thaliana*, glycosyltransferase, plant**
EMBL Database Acc#: AJ24927

In both plant and animal cells, a highly conserved multistep biosynthetic pathway covalently links carbohydrate side-chains to asparagine residues of newly synthesised proteins. This process of N-linked protein glycosylation is initiated in the endoplasmic reticulum (ER) and requires transport to the most distal regions of the Golgi apparatus for its completion. Animal and plant glycoproteins contain both oligomannosidic and complex-type N-linked glycans. However, after the action of *N*-acetylglucosaminyltransferases I and II the pathway for complex N-glycan biosynthesis in plants and animals diverge significantly. Complex plant N-glycans are generally smaller than their animal counterparts, lack sialic acid but contain instead β1,2-xylose and/or core α1,3-fucose residues that are not found in mammalian glycoproteins [1,2].

Although first steps in the biosynthesis of N-glycans in plants and animals are closely related [for reviews see 3,4] and a variety of glycosyltransferases have been cloned from different animal species [for recent review see 4,5], only two enzymes involved in the plant N-glycosylation pathway, namely N-acetylglucosaminyltransferase I and core α1,3-fucosyltransferase, have been cloned and characterised at a molecular level [6,7]. Both sequences revealed significant homology to their animal counterparts. Although *N*-acetylglucosaminyltransferase II (GnTII, EC 2.4.1.143) cDNAs have been isolated from different mammals [8,9] and from *X. laevis* [10], the cloning of a plant GnTII cDNA has so far remained elusive, despite that the enzyme has been partially purified from mung beans [11].

We have now isolated a cDNA encoding *A. thaliana* GnTII. In the present paper we report its nucleotide and amino acid sequence and its homology to animal GnTII, demonstrate its authentic enzymatic activity and confirm the expression of the corresponding mRNA in *A. thaliana* leaves.

Materials and Methods

Plant material: *Arabidopsis thaliana* (var. Columbia) plants were cultivated in a controlled growth chamber with 22°C day and night temperature, a 16 h photoperiod and 50% humidity.

Reverse Transcription-PCR: Total RNA was isolated from *A. thaliana* leaves using the TRIzol reagent (Life Technologies). The RNA was treated with DNAse (Promega, RQ1 RNase Free DNase) to remove traces of genomic DNA. First-strand cDNA was synthesised from 2 µg of total RNA primed by the pagt2/5A primer (position 107892-107911 correspond-

*To whom correspondence should be addressed. Tel: +43-1-36006-6700; Fax: +43-1-36006-6392; E-mail: steink@mail.boku.ac.at

ing to the genomic sequence, Acc. Nr: AC007018, database search: May 1999), 5′-GCGCGAATTCTCATGGAGATG-CACTGCTAC-3′, using AMV Reverse Transcriptase (Promega). PCR amplification was carried out using primer pagt2/5A and primer pagt2/9S (position 106611-106631), 5′-GCGCGGATCCTGTGGTGATGGCAAATCTTT-3′, with an initial denaturation step at 94°C for 2 min, followed by 30 amplification cycles: 54°C for 1 min followed by 72°C for 2 min and 92°C for 1 min. After completion of 30 cycles, the reaction mixture was maintained at 72°C for 8 min. An EcoRI restriction site created in pagt2/5A and a BamHI restriction site was included into pagt2/9S to facilitate subsequent subcloning of the fragment. The resulting PCR product, extracted from an agarose gel using the QIAEX II kit (QIAGEN), was digested with EcoRI and BamHI and cloned into pUC 19 (Amersham Pharmacia Biotech) and pVL 1393 (PharMingen).

PCR using primers pagt2/9S and pagt2/8A (genomic clone: position 107073-107093), 5′-CGAATAAGGCGAG AAAATCTG-3′, was done with an initial denaturation step at 94°C for 2 min, followed by 30 amplification cycles: 53°C for 45 sec, 72°C for 1 min and 92°C for 1 min. After completion of 30 cycles, the reaction mixture was maintained at 72°C for 8 min.

DNA sequencing: Plasmid DNA was purified with the aid of the QIAGEN plasmid purification kit and subjected to cycle sequencing using the Big Dye™ Terminator Sequencing Ready Mix (PE Applied Biosystems) according to the recommendations of the manufacturer. DNA and protein sequence analysis was performed using the DNASTAR PC-software package.

Expression in insect cells: *A. thaliana* GnTII cDNA was cloned into the BamHI-EcoRI sites of the baculovirus transfer vector pVL 1393 (PharMingen). The sequence of the cloning product was confirmed by DNA sequencing. *Spodoptera frugiperda* Sf-9 cells were grown in serum-free IPL-41 medium (Sigma) at 27°C. 1 µg of recombinant baculovirus transfer vector was cotransfected with 200 ng linear Baculo-Gold DNA (PharMingen) into 1×10^6 Sf-9 cells using Lipofectin (Life Technologies) according to the manufacturer's protocol. Cells were incubated for 6 days at 27°C. The supernatant was used to infect 2×10^6 Sf-21 cells in IPL-41 medium supplemented with 5% heat-inactivated fetal calf serum. Cells were incubated for 4 days at 27°C and then harvested for enzyme assays.

GnTII activity assay: The activity of the recombinant enzyme was determined with UDP-GlcNAc and pyridylaminated GnTI and GnTII acceptor substrates: M5-PA ((Manα1-3)(Manα1-6)Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAc-pyridylamine), MM-PA (Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAc-pyridylamine, or MGn-PA (GlcNAcβ1-2Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAc-pyridylamine). As the enzyme source insect cell culture supernatants or homogenized cells were used as described previously [12]. The reaction products were analysed by reversed-phase HPLC. Partial *β-N*-acetylglucosaminidase digestion and methylation analysis of the product were performed according to published procedures [13].

Results and Discussion

Cloning of *Arabidopsis thaliana* GnTII cDNA

An *Arabidopsis thaliana* BAC sequence present in the EMBL/Genbank database (*A. thaliana* chromosome II BAC F5G3 genomic sequence, Acc. Nr: AC007018) contains sequences homologous to mammalian GnTII cDNAs [8,9]. Comparative analysis of these sequences using the Genefinder programme (Phil Green, University of Washington, http://dot.imgen.bcm.tmc.edu) revealed an open reading frame of 1290 nucleotides. To isolate the cDNA clone corresponding to this genomic sequence, the oligonucleotide primers pagt2/9S and pagt2/5A were designed which include the putative start and stop codons of the gene. Using these primers we obtained a PCR fragment of 1290 bp from *A. thaliana* leaf cDNA. This PCR fragment revealed upon sequence analysis 100% identity to the genomic sequence, thus indicating the absence of introns. The open reading frame encodes a 430 amino acid protein with a calculated molecular mass of 48.9 kD (Figure 1).

Homology between *A. thaliana* GnTII and its mammalian counterparts

Hydrophobicity analysis [14] of the deduced amino acid sequence of GnTII cDNA revealed a type II transmembrane protein topology as typical of all previously cloned Golgi located glycosyltransferases, exhibiting a putative cytoplasmic domain of 14 residues at the N-terminus followed by a transmembrane region of 23 amino acids, using SOSUI software package (http://azusa.proteome.bio.tuat.ac.jp/). Amino acid sequence alignment of all known GnTII sequences revealed little homology considering the putative cytoplasmic, transmembrane and stem regions. However, in the putative catalytic domain several single amino acids and peptide motifs were found to be conserved in all species. For example, 54 amino acids between positions 100 and 242 (*A. thaliana* numbering) were found to be conserved (Figure 1). One of the conserved regions contains an EED sequence flanked by hydrophobic residues. It has been suggested, for different classes of glycosyltransferases, that this motif is most probably involved in donor substrate binding [15–17]. Furthermore, seven cysteine residues are invariant throughout all GnTII sequences available so far.

GenBank and SwissProt database searches did not reveal significant sequence similarities to known plant DNA or protein sequences except to the genomic *A. thaliana* sequence AC007018.5 [18] and partially to a tomato EST clone (AI777297). Lin et al. [18] found a putative open reading frame in the BAC clone AC007018 that was by homology to mammalian sequences tentatively assigned to encode a GnTII-like protein of 419 residues (GenPept accession number:

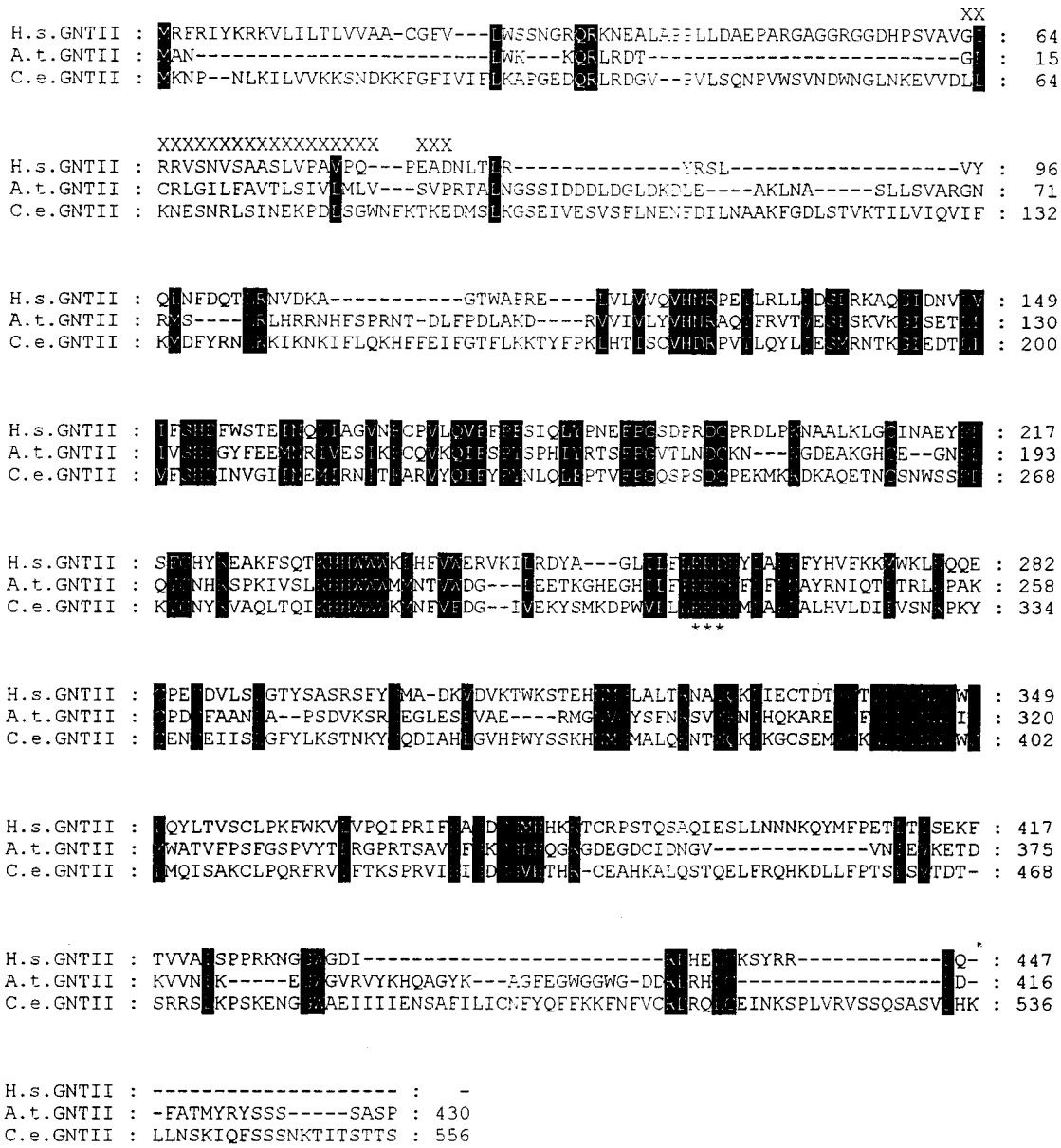

Figure 1. Multiple sequence alignment of GnTII from different species. The sequence alignments were generated using the Clustal software and refined manually. The invariant or closely related residues are given in white letters on black background. Dashes indicate gaps. Asterisks (*) indicate the invariant acidic motif suggested to be responsible for UDP-GlcNAc binding, xxx indicate the proposed *A. thaliana* GnTII membrane-spanning region. *H.s.* GNTII (human), *A.t.* GNTII (*A. thaliana*), *C.e.*GNTII (*C. elegans*, from Genbank Acc. Nr: Z81458, cosmid-sequence C03E10)).

AAD29068). The 371 C-terminal amino acids of this protein are identical to our GnTII sequence. However, the N-terminal region of the AAD29068 protein is significantly different from that in our GnTII sequence. To investigate this discrepancy, we designed specific primers for either transcript and performed RT-PCR experiments with total RNA isolated from *A. thaliana* leaves. While a band of the expected size was detected with the primer combination based on our GnTII sequence, no signal was obtained with AAD29068-specific primers (Figure 2 and data not shown). Close examination of the AC007018.5 and AAD29068 sequences revealed that Lin et al. [18] predicted one intron which is, according to our data, in fact not present in the *A. thaliana* GnTII gene.

Expression of *A. thaliana* GnTII cDNA in insect cells

A recombinant baculovirus was constructed to express full length *A. thaliana* GnTII cDNA in *Spodoptera frugiperda* (Sf-21) cells. The lysates and culture media of cells infected with the recombinant baculovirus were assayed for GnTII activity

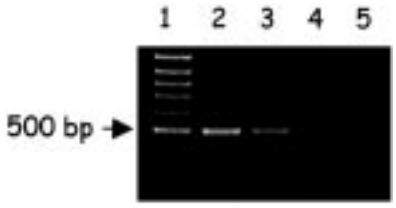

Figure 2. Detection of *A. thaliana* GnTII mRNA. Total RNA (2 μg) isolated from *A. thaliana* leaves served as template for reverse transcription using a GnTII specific primer (pagt2/8A). Subsequently a PCR using primers pagt2/8A and pagt2/9S (see Materials and Methods) was used (lane 2 and 3). Lane 1: marker, lane 4 and 5: corresponding negative controls, containing no RNA template. The size of the expected fragment is 500 bp.

with specific substrates for GnTII (Figure 3). Significant enzymatic activity was found in lysate and supernatant, whereas no activity was detectable in mock-infected cells (Figure 3).

The authenticity of the reaction product was verified by reversed-phase HPLC, partial β-*N*-acetylglucosaminidase digestion and methylation analysis. On reversed-phase HPLC, the product coeluted with authentic GnGn-PA bearing two β1,2-linked terminal GlcNAc-residues. Partial β-*N*-acetylglucosaminidase digestion of the product yielded two intermediate products, one being the substrate MGn-PA, the other an intermediate coeluting with GnM-PA where the GlcNAc residue is β1,2-linked to the 6-arm mannose. Methylation analysis of the product yielded permethylated alditol acetates indicative of 2- and 3,6-substituted mannose only. These data imply that the recombinant enzyme transferred GlcNAc in a β1,2-linkage to the terminal α1,6-mannosyl residue. No transfer of GlcNAc to Man5-PA and MM-PA, the substrates of *N*-acetylglucosaminyltransferase I (GnTI), was accomplished by recombinant *A. thaliana* GnTII. Apart from showing the absence of GnTI activity, this indicates the necessity of the presence of terminal GlcNAc on the α1,3-linked mannose also for GnTII from *A. thaliana*. Thus, from its substrate specificity and product structure, the cloned *A. thaliana* enzyme exhibits all features expected of GnTII.

Acknowledgements

We would like to thank Barbara Svoboda and Thomas Dalik for excellent technical support and Iain B.H. Wilson for critical reading of the manuscript.

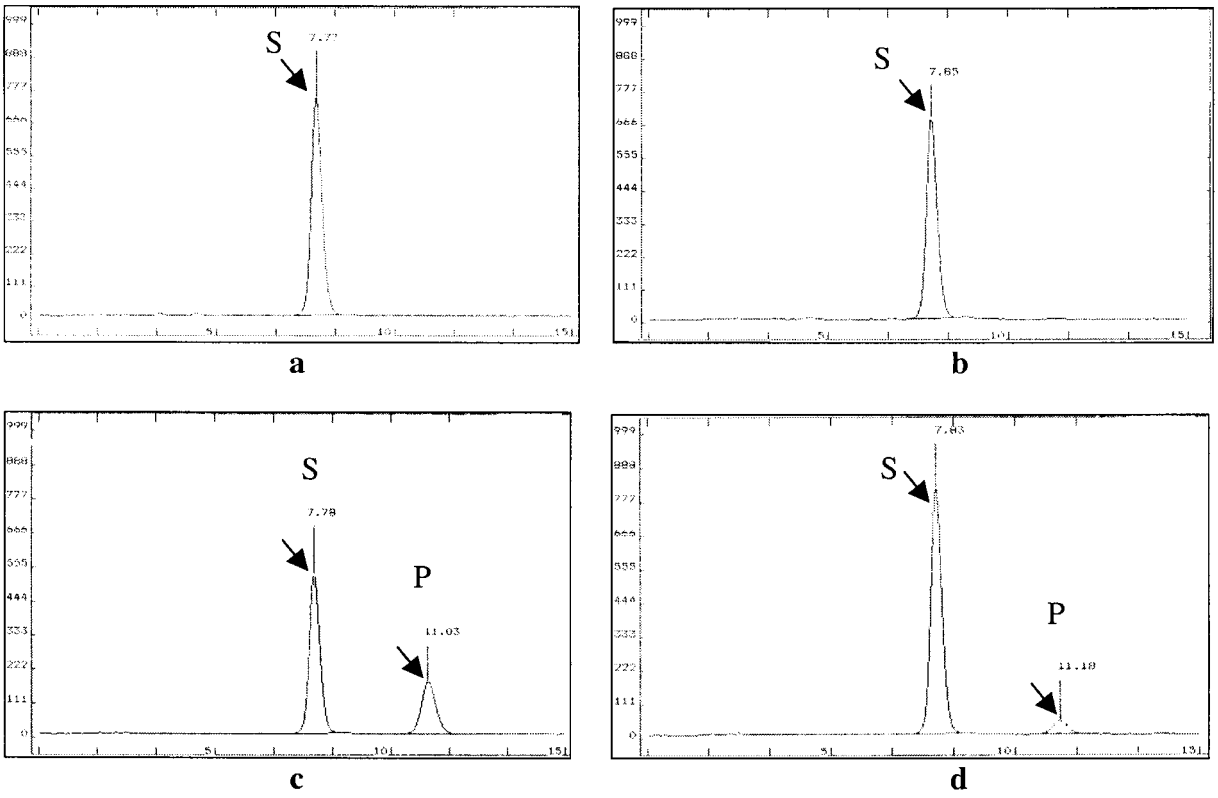

Figure 3. Demonstration of GnTII activity of the recombinant enzyme. Cell extract and supernatant of *A. thaliana* GnTII expressing insect cells were incubated with MGn-PA (2 nmol) as acceptor substrate and UDP-GlcNAc (40 nmol) as donor substrate for 18 h at 37°C. Acceptor substrate (S) and GnTII product (P) were subsequently separated by reversed-phase HPLC and detected by spectrofluorometry. Arrows indicate substrate and product peaks, the numbering indicates elution time. a and b: lysate and supernatant of approx. 4×10^4 mock infected Sf-21 cells. c and d: lysate and supernatant of approx. 4×10^4 Sf-21 cells expressing *A. thaliana* GnTII.

References

1 van Kuik JA, van Halbeek H, Kamerling JP, Vliegenthart JF, *J Biol Chem* **260**, 13984–8 (1985).
2 Kubelka V, Altmann F, Staudacher E, Tretter V, März L, Hard K, Kamerling JP, Vliegenthart JF, *Eur J Biochem* **213**, 1193–204 (1993).
3 Lerouge P, Cabanes-Macheteau M, Rayon C, Fichette-Lainé AC, Gomord, V, Faye L, *Plant Mol Biol* **38**, 31–48 (1998).
4 Bill RM, Revers L, Wilson IBH, *Protein Glycosylation*, (Kluwer Academic Publishers, Hingham, MA, USA, 1998).
5 Colley KJ, *Glycobiology* **7**, 1–13 (1997).
6 Strasser R, Mucha J, Schwihla H, Altmann F, Glössl J, Steinkellner H, *Glycobiology* **9**, 779–85 (1999).
7 Leiter H, Mucha J, Staudacher E, Grimm R, Glössl J, Altmann F, *J Biol Chem* **274**, 21830–9 (1999).
8 D'Agostaro GAF, Zingoni A, Moritz RL, Simpson RJ, Schachter H, Bendiak B, *J Biol Chem* **270**, 15211–21 (1995).
9 Tan J, D'Agostaro GAF, Bendiak B, Reck F, Sarkar M, Squire JA, Leong P, Schachter H, *Eur J Biochem* **231**, 317–28 (1995).
10 Mucha J, Kappel S, Schachter H, Hane W, Glössl J, *Glycoconjugate J* **12**, 473 (1995).
11 Szumilo T, Kaushal GP, Elbein AD, *Biochemistry* **26**, 5498–505 (1987).
12 Altmann F, Kornfeld G, Dalik T, Staudacher E, Glössl J, *Glycobiology* **3**, 619–25 (1993).
13 Tretter V, Altmann F, März L, *Eur J Biochem* **199**, 647–52 (1991).
14 Kyte J, Doolittle RF, *J Mol Biol* **157**, 105–32 (1982).
15 Breton C, Bettler E, Joziasse DH, Geremia RA, Imberty A, (1998) *J Biochem* **123**, 1000–9.
16 Breton C, Imberty A, *Curr Opin Struct Biol* **9**, 563–71 (1999).
17 Wiggins CAR, Munro S, *Proc Natl Acad Sci USA* **95**, 7945–50 (1998).
18 Lin X, Kaul S, Shea TP, Fujii CY, Shen M, VanAken SE, Barnstead ME, Mason TM, Bowman CL, Ronning CM, Benito M, Carrera AJ, Creasy TH, Buell CR, Town CD, Nierman WC, Fraser CM, Venter JC, Arabidopsis thaliana chromosome II BAC F5G3 genomic sequence, direct submission to NCBI, AAD29068 (1999).

Received 27 January 2000, revised 15 March 2000,
accepted 17 March 2000