Glycoconjugate Journal (1998) **15:** 193–197

EGC2
Removal of 106 amino acids from the N-terminus of UDP-GlcNAc : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyltransferase I does not inactivate the enzyme*

M. Sarkar[1], S. Pagny[2], U. Ünligil[3], D. Joziasse[4], J. Mucha[5], J. Glössl[5] and H. Schachter[†1,6]

[1]*Department of Biochemistry, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada*
[2]*Université de Rouen, CNRS URA203, France*
[3]*Department of Medical Genetics and Microbiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada*
[4]*Department of Medical Chemistry, Vrije University, Amsterdam, The Netherlands*
[5]*Universität für Bodenkultur, A-1180 Vienna, Austria*
[6]*Department of Biochem., University of Toronto, Toronto, Ontario M5S 1A8 Canada*

UDP-GlcNAc : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyltransferase I (GnT I, EC 2.4.1.101) plays an essential role in the conversion of oligomannose to complex and hybrid N-glycans. Rabbit GnT I is 447 residues long and has a short four-residue N-terminal cytoplasmic tail, a 25-residue putative signal–anchor hydrophobic domain, a stem region of undetermined length and a large C-terminal catalytic domain, a structure typical of all glycosyltransferases cloned to date. Comparison of the amino acid sequences for human, rabbit, mouse, rat, chicken, frog and *Caenorhabditis elegans* GnT I was used to obtain a secondary structure prediction for the enzyme which suggested that the location of the junction between the stem and the catalytic domain was at about residue 106. To test this hypothesis, several hybrid constructs containing GnT I with N- and C-terminal truncations fused to a mellitin signal sequence were inserted into the genome of *Autographa californica* nuclear polyhedrosis virus (AcMNPV), Sf 9 insect cells were infected with the recombinant baculovirus and supernatants were assayed for GnT I activity. Removal of 29, 84 and 106 N-terminal amino acids had no effect on GnT I activity; however, removal of a further 14 amino acids resulted in complete loss of activity. Western blot analysis showed strong protein bands for all truncated enzymes except for the construct lacking 120 N-terminal residues indicating proteolysis or defective expression or secretion of this protein. The data indicate that the stem is at least 77 residues long.

***Keywords*: N-glycan synthesis; GlcNAc-transferase; Sf 9 insect cells; baculovirus; catalytic domain**

Abbreviations*: AcMNPV, *Autographa californica* nuclear polyhedrosis virus; FCS, foetal calf serum; GnT, *N*-acetylglucosaminyltransferase; PCR, polymerase chain reaction; SDS–PAGE, sodium dodecyl sulphate-polyacrylamide gel electrophoresis; Sf 9, *Spodoptera frugiperda

***Enzymes*: UDP-GlcNAc : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyltransferase I, GnT I (EC 2.4.1.101)**

1. Introduction

UDP-GlcNAc : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyltransferase I (GnT I, EC 2.4.1.101) plays an essential role in the conversion of oligomannose to complex and hybrid N-glycans since the action of α-mannosidase II and the initiation of antennae by GnT II and IV all require prior GnT I action and GnT V in turn requires prior GnT II action [1]. It is therefore of interest to determine the mechanism of GnT I catalytic activity. GnT I genes have been cloned and expressed from man [2, 3], mouse [4, 5], rat [6] and rabbit [7] and GnT I-like sequences have been cloned from frog (unpublished work), hen [8] and *Caenorhabditis elegans* [9]. Rabbit GnT I is 447 residues long and has

*This paper was posted to the internet as part of the Second Electronic Glycosciences Conference held during September, 1996.
[†]To whom correspondence should be addressed.

0282–0080

```
        51                                                  93
Human   TREVIRLAQDAEVELERQRGLLQQIGD..ALSSQRGRVPTAAP.........
Mouse   -----H--E---A------------KEHY--WR--W----V--.........
Rat     -----H--E---A------------KEHYS-WR--W----V--.........
Rabbit  -------------------------REHH--W---WK------.........
Frog    -AH--Q--EE--A---K-K------HYYSG-LNHQQAPSHVRAEPITTNMSF
Cel B   LESAN--ERLLKF-AKKIAA-AEDVHK..IRANRK-KHVIMEE.........
Cel M   DEAGS--EKLLK--K-HMER-SETMNR..-TRLRIT................

        94                                                 141
Human   PAQPRVPVTPAPAVIPILVIACDR.STVRRCLDKLLHYRPSAELFPIIV
Mouse   --W-------S-VQ----------.-----------------R-----
Rat     --W----G--S------------.-----------------H-----
Rabbit  ----H-----P------------.----------------------
Frog    -SPSP-GSG-L-L------V----.PS------S--K------N-----
Hen                    -V--L----.-S------S--R-----QR---L-
Cel B   MVSQDLKQWKD-..--V--FS-N-AMA--DHVE--IR----Q-K-----
Cel M   -NSNNPNNW-H-..--VI-FS-N-PDS--AHVE--IM-----QQ---T-
```

Figure 1. Alignment of GnT I amino acid sequences at the 'stem-catalytic domain' junction. (–––), same amino acid as in human GnT I; (···), gaps introduced to optimize alignment. Accession numbers: human, M61829; mouse, M73491; rat, D16302; rabbit, M57301; Cel B, *C. elegans* clone B, U23516; Cel M, *C. elegans* clone M, Z46381. Neither the hen [8] nor the frog (JM and JG) sequences have been published. The hen N-terminal sequence has not as yet been determined. Numbering refers to the mouse, rat and rabbit enzymes with the initiation Met residue at 1.

a short four-residue N-terminal cytoplasmic tail, a 25-residue putative signal–anchor hydrophobic domain, a stem region of undetermined length and a large C-terminal catalytic domain, a structure typical of all glycosyltransferases cloned to date [10, 11]. Alignment of GnT I sequences from different species (Figure 1) shows a dramatic increase in sequence similarity commencing at about residue 105. An automatic mail server (PHD) for protein secondary structure prediction based on sequence profiles and neural networks [12–15] was applied to the GnT I sequences from human, rabbit, mouse and rat; the analysis predicted helix and loop regions up to amino acid residue 106 with the first β strand beginning at residue 107.

In the present report, we have attempted to delineate the catalytic domain of GnT I by expressing in the baculovirus/Sf9 insect cell system a series of GnT I constructs containing N- and C-terminal truncations. We show that removal of 29, 84 and 106 N-terminal amino acids had no effect on GnT I activity; however, removal of a further 14 amino acids resulted in complete loss of activity. Truncation of seven amino acids at the C-terminus resulted in over 60% loss of activity. Western blot analysis showed strong protein bands for all truncated enzymes except for the construct lacking 120 N-terminal residues and therefore loss of enzyme activity may be due to proteolysis or a defect in either transcription–translation or secretion. The data indicate that the stem is at least 77 residues long in agreement with the secondary structure prediction and sequence alignment data (Figure 1).

Materials and methods

Materials

The following materials were purchased from the indicated sources: Grace's insect medium, foetal calf serum (FCS), TC Yeastolate, TC lactalbumin hydrolysate, gentamycin, amphotericin (Fungizone), competent *E. coli* DH5α (Gibco laboratories, Grand Island, NY); Vent DNA polymerase (New England Biolabs); Sep-Pak C18 reverse phase cartridges (Waters); T7 Sequencing Kit, restriction endonucleases, T4 DNA ligase (Pharmacia); BaculoGold baculovirus linearized DNA, transfection buffers, cationic liposomes, wild-type AcMNPV (PharMingen); *Spodoptera frugipèrda* (Sf9) insect cells (ATCC); Triton X-100, goat anti-mouse IgG (whole molecule) coupled to alkaline phosphatase (Sigma). UDP-[^{3}H]GlcNAc (NEN) was diluted with non-radioactive UDP-GlcNAc (Sigma) to a specific activity of 3000 dpm/nmole. Man-α-1-6[Man-α-1-3]Man-β-octyl (M3-octyl) was kindly provided by Dr. Hans Paulsen, University of Hamburg, Hamburg, Germany. Oligonucleotides were synthesized on a Pharmacia DNA synthesizer and purified by the cartridge method (Hospital for Sick Children–Pharmacia Biotechnology Center, Toronto, Canada). The baculovirus transfer vector pVT-Bac-His was constructed by one of us (DJ) in collaboration with Drs G. Smit and R. van Elk (Faculty of Biology, Vrije Universiteit, Amsterdam, The Netherlands). A DNA fragment encoding six His residues followed by an enterokinase-sensitive site was ligated into the *Bam*H1 and *Pst*I sites of

pVT-Bac (kindly donated by Dr T. Vernet, Biotechnology Research Institute, Montreal, Quebec, Canada [16]). Recombinant proteins encoded by this vector will contain a melittin cleavable signal peptide at their N-terminus and be secreted from baculovirus-infected insect cells.

Expression of truncated GnT I

All molecular biology procedures were carried out by standard methods [17, 18]. Plasmids were amplified in competent *E. coli* DH5α. Rabbit GnT I cDNA [7] was used as a template for PCR amplification using Vent DNA polymerase with several sets of gene-specific primers (Table 1) designed to produce GnT I molecules truncated at either the N- or C-terminal ends. PCR conditions were one cycle at 96 °C × 3 min, 50 °C × 1 min, 72 °C × 1 min followed by 39 cycles at 95 °C × 1 min, 50 °C × 1 min, 72 °C × 1 min and a final 5 min extension at 72 °C. The PCR products were subcloned into pVT-Bac-His downstream from and in frame with the ATG start site of the plasmid using *Sac*I and *Kpn*I sites introduced by the primers to yield pVT-Bac-His-GnT I plasmids [16]. Plasmid DNA sequences were verified by the double-strand dideoxy method [19].

Sf9 cells were grown at 28 °C in Grace's insect medium supplemented with 10% FCS, TC Yeastolate, TC lactalbumin hydrolysate, 50 μg ml^{-1} gentamycin sulfate and 2.5 μg ml^{-1} amphotericin (Fungizone) in Falconware T flasks. The pVT-Bac-His-GnT I plasmids (60 ng) were co-transfected with BaculoGold baculovirus linearized DNA (14 ng) into 5×10^4 Sf9 cells at 80% confluency in 96-well microtitre plates using cationic liposomes according to the PharMingen protocol [20–24]. The cells were incubated for 4 days at 28 °C to allow homologous recombination and release of virus into the culture medium. The recombinant baculovirus was amplified twice and used to infect Sf9 cells at a multiplicity of infection of 2 p.f.u. cell^{-1}. At 5–6 days post infection, cells were sedimented and lysed in 0.5 ml 25 mM MES, pH 6.5, 0.1% Triton X-100, 0.02% sodium azide. Both cell lysates and supernatants were assayed for GnT I using M3-octyl and UDP-[^{3}H]GlcNAc as substrates and SepPak C18 cartridges to obtain radioactive product [23, 25]. Control GnT I assays were carried out with uninfected Sf9 cells and cells transfected with wild-type baculovirus.

Table 1. Primers used for PCR truncation of *N*-acetylglucosaminyltransferase I

PCR primer	*Amino acid at 5′-end of PCR product*	*Nucleotide sequence*
Upstream (forward)[a]	30	ACACGTCCAGTGCCTAGC
	85	CGGTGGAAGGTGCCTACT
	107	GTGATCCCCATCCTGGTA
	121	CGCCGCTGTTTGGACAAG
Down stream (reverse)[b]	447	CAGGCAGGAGCTGTTAAG
	440[c]	ATCCCAAGTCTGAGGGGG

[a]Upstream primers had an additional sequence (not shown) at the 5′-end which included the *Sac*I restriction site.
[b]Downstream primers had an additional sequence (not shown) at the 5′-end which included the *Kpn*I restriction site.
[c]This primer had an additional sequence (not shown) at the 5′-end which included both the *Kpn*I restriction site and a termination codon to permit premature termination seven residues prior to normal termination.

Western blot analysis

Aliquots of crude Sf9 supernatant (containing about 15–30 ng recombinant enzyme) were analysed by SDS–12.5% PAGE [26] and transferred by electrophoresis to an Immobilon-PVDF membrane (0.45 microns, Millipore) in the presence of 0.1% SDS. The membranes were probed with 1 : 5000 diluted mouse monoclonal antibody raised against the enterokinase cleavage site (Anti-Xpress antibody kit, Invitrogen). This enterokinase site is encoded by pVT-Bac-His and is expressed near the N-terminus of recombinant GnT I. Proteins were visualized by treatment with 1 : 2500 diluted alkaline phosphatase-conjugated goat anti-mouse IgG followed by staining with bromochloroindolyl phosphate and nitro blue tetrazolium (Promega) according to the manufacturer's instructions. The stained membranes were scanned (UVP scanner) and signal intensities were semi-quantitated in square pixels using the computer program provided by the scanner manufacturer (UVP-Grab IT).

Results

Expression of recombinant GnT I

The catalytic activity of recombinant rabbit GnT I was previously studied by expression in Sf9 cells of cDNA corresponding to amino acid residues 30–447 containing the full stem and catalytic domains but lacking the N-terminal and transmembrane domains [23]. In the present work, a series of GnT I cDNA constructs encoding various N- and C-terminal truncations were expressed in Sf9 cells to produce soluble secreted enzyme. Expression of GnT I containing amino acid residues 30–447 was repeated with the present expression system as a positive control. As shown in Table 2, constructs containing amino acid residues 30–447, 85–447 and 107–447 had similar enzyme activities. Deletion of an additional 14 amino acids (construct containing residues 121–447) resulted in almost complete loss of enzyme activity indicating that the stem region extends to at least amino acid residue 106. Evidence has been obtained that the large C terminal region of the enzyme is essential for catalytic activity [10, 27]. The effect of truncation at the carboxyl end on catalytic activity was therefore determined. Removal of the last seven amino acids (Table 2) resulted in

Table 2. Catalytic activity of recombinant rabbit *N*-acetylglucosaminyltransferase I in the Sf 9 supernatants and pellets. The numbers designating truncated enzyme indicate the amino-and carboxy-terminal ends with the N-terminal methionine residue of the native enzyme designated as 1 [7].

Recombinant truncated enzyme	*Enzyme activity (dpm × 10^{-5}/10^{6} cells/h)*		*Pellet/ supernatant*	*Theoretical mass (kDa)*	*Experimental mass (kDa)*[a]
	Supernatant	*Pellet*			
30–447	18	0.85	0.047	52.6	48
85–447	19	0.92	0.048	46.4	43
107–447	14.5	0.66	0.046	44.1	42
121–447	0.4	0.05	0.1	42.6	ND
85–440	6.0	ND		45.6	41, 43

[a]Figure 2.
ND, not determined.

a loss of over 60% of the catalytic activity. No GnT I activity was detected in the supernatants of uninfected Sf 9 cells nor of Sf 9 cells infected with wild-type baculovirus. The supernatants contained 96% of the total enzyme activity (Table 2).

Western blot analysis of recombinant GnT I

Western blot analysis for the presence of recombinant protein in Sf 9 supernatants was carried out to determine whether absence of enzyme activity was due to secretion of an inactive protein or to lack of protein secretion. The fully active recombinant enzymes all showed a band at the expected molecular weights (Table 2) but there was no signal for the inactive preparation (residues 121–447) (Figure 2). We have not yet determined whether this is due to a defect in either secretion or transcription–translation or to proteolysis of a destabilized protein during the expression and secretion processes.

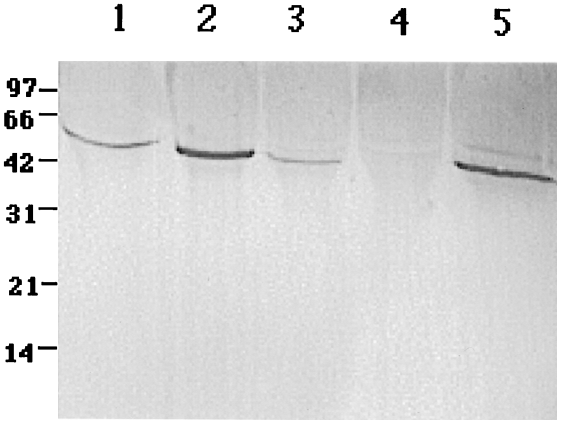

Figure 2. Western blot analysis of recombinant proteins. Crude Sf 9 supernatants containing recombinant protein (about 15–30 ng/lane) were analysed by SDS–12.5% PAGE and transferred to an Immobilon-PVDF membrane. The membranes were probed with mouse monoclonal antibody raised against the enterokinase cleavage site and proteins were visualized with alkaline phosphatase-conjugated goat anti-mouse IgG. Recombinant proteins are designated as described in Table 2. Lane 1, 30–447; lane 2, 85–447; lane 3, 107–447; lane 4, 121–447; lane 5, 85–440. Control Sf 9 supernatants from cells infected with wild-type virus did not show a protein band (data not shown).

Discussion

The studies described above indicate that removal of the first 106 amino acids at the N-terminus of rabbit GnT I has no effect on enzyme activity. Removal of the next 14 amino acids led to complete loss of activity due to the absence of protein secretion. The most likely explanation is that the truncated enzyme is unstable and degraded during the expression and secretion processes. However, further studies are required to rule out a defect in either secretion or transcription–translation. We can conclude that the stem is at least 77 amino acid residues long. This finding correlates precisely with the stem length estimate based on secondary structure prediction and on amino acid sequence alignment (Figure 1). The high similarity in amino acid sequences commencing at amino acid residue 105 strongly suggests that the catalytic domain begins at this residue.

Previous studies by Henion *et al.* [28] on recombinant marmoset UDP-Gal : Galβ1-4GlcNAc-R α-1,3-galactosyltransferase and by Xu *et al.* [29] on human α-1,3(4)-fucosyltransferases III and V have shown that removal, respectively, of 89, 51 and 75 N-terminal amino acids from these three enzymes had no effect on enzyme activity. The lengths of the native enzymes were 376, 361 and 374 amino acids respectively. Shorter forms of these enzymes had either reduced or no detectable enzyme activity. These studies show that the stem regions of the three enzymes respectively comprise approximately 67, 18 and 42 amino acid residues. Truncation of three amino acids from the C-terminal end of the α-1,3-galactosyltransferase and of two amino acids from the C-terminal end of fucosyltransferase V led to the production of inactive recombinant enzymes. The data illustrate the remarkable similarity in domain structure between

GnT I and the above three glycosyltransferases although there is no primary amino acid sequence similarity between GnT I and these enzymes. Structural characterization of the catalytic domains can therefore be carried out on truncated recombinant forms of these glycosyltransferases lacking the cytoplasmic, transmembrane and stem regions. Such truncated enzymes should lend themselves more readily to X-ray crystallography and to analysis by high resolution nuclear magnetic resonance spectroscopy.

Acknowledgements

This work was supported by grants from the Medical Research Council of Canada and PENCE, The Protein Engineering Network of Centres of Excellence of Canada.

References

1 Schachter H (1986) *Biochem Cell Biol* **64**: 163–81.
2 Hull E, Sarkar M, Spruijt MPN, Höppener JWM, Dunn R, Schachter H (1991) *Biochem Biophys Res Commun* **176**: 608–15.
3 Kumar R, Yang J, Larsen RD, Stanley P (1990) *Proc Natl Acad Sci USA* **87**: 9948–52.
4 Kumar R, Yang J, Eddy RL, Byers MG, Shows TB, Stanley P (1992) *Glycobiology* **2**: 383–93.
5 Pownall S, Kozak CA, Schappert K, Sarkar M, Hull E, Schachter H, Marth JD (1992) *Genomics* **12**: 699–704.
6 Fukada T, Iida K, Kioka N, Sakai H, Komano T (1994) *Biosci Biotechnol Biochem* **58**: 200–1.
7 Sarkar M, Hull E, Nishikawa Y, Simpson RJ, Moritz RL, Dunn R, Schachter H (1991) *Proc Natl Acad Sci USA* **88**: 234–38.
8 Narasimhan S, Yuen R, Fode CJ, Ali J, Rajalakshmi S, Schappert K, Dennis JW, Schachter H (1993) *Glycobiology* **3**: 531.
9 Wilson R, Ainscough R, Anderson K, Baynes C, Berks M, Bonfield J, Burton J, Connell M, Copsey T, Cooper J (1994) *Nature* **368**: 32–8.
10 Paulson JC, Colley KJ (1989) *J Biol Chem* **264**: 17615–18.
11 Schachter H (1994) In *Molecular Glycobiology*, (Fukuda M, Hindsgaul O, eds) pp 88–162, Oxford: Oxford University Press.
12 Rost B, Sander C (1993) *J Mol Biol* **232**: 584–99.
13 Rost B, Sander C (1993) *Proc Natl Acad Sci USA* **90**: 7558–62.
14 Rost B, Sander C, Schneider R (1994) *Computer Applications in the Biosciences* (*CABIOS*) **10**: 53–60.
15 Rost B, Sander C (1994) *Proteins* **19**: 55–72.
16 Tessier DC, Thomas DY, Khouri HE, Laliberté F, Vernet T (1991) *Gene* **98**: 177–83.
17 Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1993) *Current Protocols in Molecular Biology*. New York: John Wiley and Sons.
18 Sambrook J. Fritsch EF, Maniatis T (1989) *Molecular Cloning, A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
19 Sanger F, Nicklen S, Coulson AR (1977) *Proc Natl Acad Sci USA* **74**: 5463–7.
20 Smith GE, Summers MD, Fraser MJ (1983) *Mol Cell Biol* **3**: 2156–65.
21 Smith GE, Ju G, Ericson BL, Moschera J, Lahm HW, Chizzonite R, Summers MD (1985) *Proc Natl Acad Sci USA* **82**: 8404–8.
22 Summers MD, Smith GE (1987) *Texas Agricultural Experiment Station Bulletin No. 1555*.
23 Sarkar M (1994) *Glycoconjugate J* **11**: 204–9.
24 Tan J, D'Agostaro GAF, Bendiak B, Reck F, Sarkar M, Squire JA, Leong P, Schachter H (1995) *Eur J Biochem* **231**: 317–28.
25 Palcic MM, Heerze LD, Pierce M, Hindsgaul O (1988) *Glycoconjugate J*. **5**: 49–63.
26 Laemmli UK (1970) *Nature* **227**: 680–5.
27 Joziasse DH (1992) *Glycobiology* **2**: 271–7.
28 Henion TR, Macher BA, Anaraki F, Galili U (1994) *Glycobiology* **4**: 193–201.
29 Xu XH, Vo L, Macher BA (1996) *J Biol Chem* **271**: 8818–23.

Accepted in revised form 20 June 1997