Glycoconjugate Journal 19, 187–195, 2003
© 2003 Kluwer Academic Publishers. Manufactured in The Netherlands.

Two closely related forms of UDP-GlcNAc: α6-D-mannoside β1,2-*N*-acetylglucosaminyltransferase II occur in the clawed frog *Xenopus laevis**

Jan Mucha[1], Barbara Svoboda[1], Sonja Kappel[1], Richard Strasser[1], Peter Bencur[1], Ulrike Fröhwein[1], Harry Schachter[2], Lukas Mach[1†] and Josef Glössl[1]

[1]*Zentrum für Angewandte Genetik, Universität für Bodenkultur Wien, Muthgasse 18, A-1190 Wien, Austria,* [2]*The Hospital for Sick Children, Program in Structural Biology and Biochemistry, Toronto, M5G 1X8 Canada*

UDP-GlcNAc:α6-D-mannoside β1,2-*N*-acetylglucosaminyltransferase II (GnT II; EC 2.4.1.143) is a medial-Golgi resident enzyme that catalyses an essential step in the biosynthetic pathway leading from high mannose to complex N-linked oligosaccharides. Screening a cDNA library from *Xenopus laevis* ovary with a human GnT II DNA probe resulted in the isolation of two cDNA clones encoding two closely related GnT II isoenzymes, GnT II-A and GnT II-B. Analysis of the corresponding genomic DNAs revealed that the open reading frame of both *X. laevis* GnT II genes resides within a single exon. The GnT II-A gene was found to be transcriptionally active in all *X. laevis* tissues tested. In contrast, expression of the GnT II-B gene was detected only in a limited number of tissues. Both GnT II-A and GnT II-B exhibit a type II transmembrane protein topology with a putative N-terminal cytoplasmic tail of 9 amino acids followed by a transmembrane domain of 18 residues, and a C-terminal luminal domain of 405 residues. The two proteins differ at 28 amino acid positions within their luminal regions. Heterologous expression of soluble forms of the enzymes in insect cells showed that GnT II-A and GnT II-B are both catalytically active and exhibit similar specific activities. Both recombinant proteins are modified with N-linked oligosaccharides. N-terminal deletion studies demonstrated that the first 49 amino acid residues are not essential for proper folding and enzymatic activity of *X. laevis* GnT II.
Published in 2003.

Keywords: **GlcNAc-transferase II, amphibian, glycosyltransferase, N-glycan biosynthesis**

Abbreviations: **BSA, bovine serum albumin; CTS, cytoplasmic-transmembrane-stem; DTT, dithiothreitol; GalT, UDP-galactose:D-GlcNAc β1,4-galactosyltransferase; GGnM$_3$-octyl, Manα-1,6(Galβ-1,4-GlcNAcβ-1,2-Manα-1,3)Manβ-1-O-octyl; GlcNAcMan$_3$GlcNAc$_2$, Manα-1,6(GlcNAcβ-1,2-Manα-1,3)Manβ-1,4-GlcNAcβ-1,4-GlcNAc; GnM$_3$-octyl, Manα-1,6 (GlcNAcβ-1,2-Manα-1,3)Manβ-1-O-octyl; GnT I, β1,2-*N*-acetylglucosaminyltransferase I; GnT II, β1,2-*N*-acetylglucosaminyltransferase II; GST, glutathione S-transferase; HCA, hydrophobic cluster analysis; M$_3$-octyl, Manα-1,6(Manα-1,3)Manβ-1-O-octyl; ORF, open reading frame; PA, pyridylamine; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; PMSF, phenylmethylsulfonyl fluoride; MES, 2-(N-morpholino)ethane-sulfonic acid; PNGase F, peptide:N-glycosidase F; RT, reverse transcriptase; SDS, sodium dodecyl sulfate; SSC, sodium chloride-sodium citrate.**

Introduction

N-glycans, in particular those of the complex type, have an enormous diversity due to the number and composition of their outer chains and are often characteristic for a particular cell type or a specific stage of differentiation. The first step in the biosynthesis of the outer chains of N-glycans is the transfer of a GlcNAc residue to the Manα-1,3-Manβ- branch of {Manα-1,6(Manα-1,3)Manα-1,6}{Manα-1,3}Manβ-1,4-GlcNAcβ-

*The DNA sequences for *Xenopus laevis* GnT II-A and GnT II-B have been deposited in the GenBank database under the accession numbers *X89 002 and AJ517298.*

†To whom correspondence should be addressed: Lukas Mach, Zentrum für Angewandte Genetik, Universität für Bodenkultur Wien, Muthgasse 18, A-1190 Wien, Austria. Tel.: +43-1-36006-6360; Fax: +43-1-36006-6392; E-mail: lukas.mach@boku.ac.at

1,4-GlcNAc-Asn by β1,2-N-acetylglucosaminyltransferase I (GnT I), followed by the action of α-3,6-mannosidase II and the transfer of GlcNAc to the Manα-1,6-Manβ- arm by β1,2-N-acetylglucosaminyltransferase II (GnT II). The latter step is a key event in the biosynthesis of complex-type N-glycans containing two or more branches [1,2]. Congenital Disorders of Glycosylation (CDG) are a family of diseases characterized by incomplete N-glycosylation of proteins. Patients with CDG type IIa display psychomotor retardation, speech impairment and facial dysmorphy, caused by mutations within the catalytic domain of GnT II [3–7]. Over 60% of mouse embryos lacking *MGAT2,* the gene encoding GnT II, reach parturition but 99% of the newborns die during the first week of postnatal development. Survivors deficient in GnT II activity and complex N-glycans are runted, and display hematologic and osteogenic abnormalities [8,9].

The GnT II sequences available to date were obtained from human [10], rat [11], mouse [12], pig [13], *D. melanogaster* (*AY 055120*, *AE003772*; J. Tan, A.M. Spence, and H. Schachter, unpublished data), *C. elegans* [14] and *A. thaliana* cDNA [15]. Only one copy of the gene was found in the human genome [10]. While the open reading frames of the human, rat, pig and *A. thaliana MGAT2* loci are on a single exon, the *C. elegans* and *D. melanogaster* GnT II genes contain at least six and seven exons, respectively. Oocytes of the clawed frog *Xenopus laevis* are often used for the production of recombinant proteins. The system is particularly well suited for the biochemical and electrophysiological analysis of transporter proteins and ion channels [16]. Interestingly, N-glycosylation is frequently a prerequisite for stability and proper subcellular localisation of these membrane proteins [17]. However, information on the N-glycosylation potential of *X. laevis* cells is restricted to evidence that oocytes synthesize both oligomannosidic and complex-type N-glycans [18,19].

In a previous paper [20] we have demonstrated that the *X. laevis* genome encodes two closely related GnT I isoenzymes. We now report the molecular cloning and biochemical characterization of two homologous forms of *X. laevis* GnT II, GnT II-A and GnT II-B.

Materials and methods

Adult *X. laevis* frogs were provided by Dr. H. Wiener (Institut für klinische Pharmakologie, Universität Wien). *Escherichia coli* strain C600Hf and packaging extracts were purchased from Promega. The TOPO-cloning kit was purchased from Invitrogen. Oligonucleotide primers for PCR analysis and DNA sequencing were synthesized by Codon Genetics Systems, Weiden and VBC-Genomics Bioscience Research, Wien. Restriction enzymes were purchased from Roche, New England BioLabs and MBI Fermentas. Prestained protein standards were from New England BioLabs. Hybond-N membranes and [α-^{32}P]dCTP (3000 Ci/mmol) were purchased from Amersham Pharmacia Biotech. The DNA sequencing kit ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction was from Perkin-Elmer. Nylon membranes and PNGase F were from Roche. Baculovirus transfer vector pVTBacHis-1 [21] was kindly provided by David Joziasse (Department of Medical Chemistry, Vrije Universiteit Amsterdam, The Netherlands). BaculoGold DNA was purchased from PharMingen. *Spodoptera frugiperda* (Sf9 and Sf21) insect cells were obtained from ATCC. Lipofectin and RNA standards (0.24–7.4 kb ladder) were purchased from Life Technologies. UDP-[^{14}C]GlcNAc (288 mCi/mmol) was from New England Nuclear. M_3-octyl was purchased from Toronto Research Chemicals. GlcNAcMan$_3$GlcNAc$_2$-PA was kindly provided by Dr. F. Altmann (Institut für Chemie, Universität für Bodenkultur Wien). β-N-Acetylglucosaminidase (jack beans) was from Sigma.

cDNA library construction and isolation of cDNA clones

cDNA was synthesized from *X. laevis* ovary poly(A)$^+$-RNA (5 μg) using the cDNA Synthesis kit (Amersham Pharmacia Biotech) according to the instructions of the manufacturer. Recombinant phages were grown on *Escherichia coli* XL1-Blue cells at a density of approximately 25,000 plaques per 150 mm plate. Plaques were transferred to Hybond-N membranes (Amersham) and 4 $\times$ 10^5 plaques were hybridized with a 1.2 kb coding sequence of human GnT II generated by PCR from the clone pHG30 (3.0 kb genomic human GnT II DNA fragment cloned into pBluescript II KS$^+$ vector) [10]. The PCR product amplified by employing sense (5′-GCCTGCGGCTTCGTCCTCTG-3′) and antisense (5′-TCTCCCCACCCTCCATTTTT-3′) primers was labeled by the random primer method with [α-^{32}P]dCTP (3000 Ci/mmol) using the Oligolabelling kit (Amersham Pharmacia Biotech). Hybridization as well as plaque purification was performed as previously described [20].

DNA sequencing

Sequences of cDNAs and genomic DNA were determined using *Taq* DNA polymerase and the dideoxy dye-terminator technique and analyzed on an Applied Biosystems 373 A DNA Sequencer following the standard protocol as described by the manufacturer. Sequence data were obtained using gene-specific or vector-derived sense and antisense oligonucleotides as primers. DNA sequence manipulation was carried out using DNASTAR software.

Cloning of genomic DNA encoding *X. laevis* GnT II isoenzymes

PCR primers used for amplification of genomic DNA encoding GnT II isoenzymes were designed based on the coding sequences of the respective cDNAs. A genomic fragment encoding isoenzyme A was amplified employing sense (5′-GCCGCTGCAGATCTCCCCGGCAGCTAGA-3′) and antisense (5′-GCGGCGGTACCTCACTGTAACCGATGATAACT-3′) primers, and a fragment encoding isoenzyme B was amplified employing (5′-GCCGCTGCAGATATCCCCGGCCGCTAGA-3′) and

(5′-CGCGAATTCCAAGGTCCAGTTTCACTGTAA-3′) as sense and antisense primers, respectively. Genomic *X. laevis* DNA (∼100 ng) from liver was subjected to PCR with 0.5 U Turbo *Pfu* DNA polymerase (Stratagene) and 20 pmol of each primer in 50 μl buffer (Stratagene) and 125 μM of each dNTP. The samples were first incubated for 3 min at 94°C, then amplification was carried out for 35 cycles at 94°C/1 min, 51°C/2 min, and 72°C/1 min. The final elongation step was extended to 7 min. The PCR product was extracted from an agarose gel slice with the aid of QIAEX II (Qiagen), engineered with the restriction enzymes PstI and KpnI and subcloned into the pVTBacHis-1 transfer vector [21].

Isolation of mRNA from *X. laevis* tissues and Northern blot analysis

Adult *X. laevis* tissues were frozen in liquid N_2 immediately after excision and processed for isolation of total RNA according to a protocol based on the method of Chomczynski and Sacchi [22] as described previously [20]. Poly $(A)^+$-RNA was selected by two cycles of binding to oligo (dT)-cellulose as described previously [23]. Total RNA (5 μg) or poly $(A)^+$-RNA (1 μg) were separated by agarose electrophoresis and transferred onto a Nylon membrane [20]. The blot was probed with [α-^{32}P]dCTP-labeled *X. laevis* DNA encoding isoenzyme B essentially as previously described [20].

Southern blot analysis

Genomic *X. laevis* DNA (∼15 μg) was exhaustively digested overnight at 37°C with the restriction enzymes BamHI, EcoRI or HindIII, fractionated by agarose electrophoresis and transferred onto a Nylon membrane [20]. The blot was hybridized with [α-^{32}P]dCTP-labeled *X. laevis* DNA encoding isoenzyme B under the conditions described previously [20].

RT-PCR analysis

Total RNA (5 μg) isolated from various *X. laevis* tissues was treated with RNase-free DNase and reverse transcribed using the First-strand cDNA synthesis kit (Amersham Pharmacia Biotech). The cDNA (0.5 μg) was subjected to PCR for 35 cycles using sense (5′-GTATCAATGCAGAATATCCAG-3′) and antisense (5′-ACGTTTCCCTGTTCATTGCC-3′) primers with an annealing temperature of 50°C. PCR products were digested with the restriction enzyme NdeI prior to analysis by electrophoresis on a 1.5% agarose gel. To improve the detection of weak bands, the agarose gel was subsequently subjected to Southern blotting with [α-^{32}P]dCTP labeled *X. laevis* DNA encoding isoenzyme B.

Construction of baculovirus transfer vectors

DNA fragments encoding truncated forms of GnT II-B were obtained by PCR with sense (5′-CGCCGGATCCATATCCCC-GGCCGCTAGA-3′) and antisense (5′-CGGGATCCAAGGT-CCAGTTTCACTGTAA-3′) primers (deletion of 49 amino acid residues from the N-terminus), and with sense (5′-CGC-GGATCCGTAGTCCAGGTGCAGAACAG-3′) and antisense (5′-CGGGATCCAAGGTCCAGTTTCACTGTAA-3′) primers (deletion of 106 amino acid residues from the N-terminus). Fragments cleaved with the restriction enzyme BamHI at underlined positions were ligated into BamHI-digested pAcSecG2T (PharMingen) baculovirus transfer vector. In these constructs, the truncated *X. laevis* GnT II protein is placed downstream of the complete coding region of *S. japonicum* GST headed by the gp67 leader sequence. In all baculovirus constructs used, the heterologous sequences are placed under the control of the strong constitutive polyhedrin promoter.

A second set of DNA fragments encoding truncated forms of GnT II-A and GnT II-B (missing their first 49 amino acids) were obtained by PCR employing sense (5′-GCCG-CTGCAGATCTCGCCGGCAGCTAGA-3′) and antisense (5′-GCGGCGGTACCTCACTGTAACCGATGATAACT-3′) primers (isoenzyme A), and sense (5′-GCCGCTGCAGATAT-CCCCGGCCGCTAGA-3′) and antisense (5′-CGCGAATTC-CAAGGTCCAGTTTCACTGTAA-3′) primers (isoenzyme B). The PCR products were cleaved with PstI and KpnI at the underlined sites and ligated into pVTBacHis-1 [21] baculovirus transfer vector digested with the same enzymes. In these constructs, the truncated *X. laevis* GnT II proteins are placed downstream of the melittin signal peptide, a His-tag and an enterokinase cleavage site.

Heterologous expression of *X. laevis* GnT II isoenzymes in insect cells

Sf9 and Sf21 cells were grown in IPL-41 medium (Sigma) containing 5% heat-inactivated fetal bovine serum (Life Technologies). Recombinant baculoviruses were constructed to express the individual truncated forms of *X. laevis* GnT II in *Spodoptera frugiperda* cells. Each recombinant baculovirus transfer vector (1 μg) was cotransfected with 200 ng BaculoGold viral DNA (PharMingen) into Sf9 cells using lipofectin as recommended by the manufacturer. After 5 days at 27°C, supernatant containing recombinant virus was used for infection of Sf21 cells. Cells and conditioned media were harvested and subjected to enzymatic analysis and immunoblotting [20].

Purification of recombinant *X. laevis* GnT II isoenzymes

Culture supernatants (50 ml) of Sf21 cells infected with the respective baculoviruses were cleared by centrifugation and dialyzed two times against 2 L of 10 mM sodium phosphate buffer, pH 7.0, 40 mM NaCl, 0.02% NaN_3. Supernatant containing 20 mM imidazole and 10% (v/v) glycerol was loaded onto a 5 ml column of Chelating Sepharose (Pharmacia) charged with Ni^{2+} ions, equilibrated in the same buffer. After successive washes with 40 mM and 80 mM imidazole, the enzyme was eluted with 250 mM imidazole in dialysis buffer. After concentration by ultrafiltration, the concentrate was analysed by SDS-PAGE

and silver staining. Recombinant human GnT II lacking the N-terminal 82 amino acids (J. Tan and H. Schachter, unpublished results) and recombinant rabbit GnT I [24] were produced and purified by the same procedure. The GnT II protein yield was estimated by densitometric analysis of the stained gel, using bovine serum albumin as a standard.

Preparation of GnM$_3$-octyl

Large scale synthesis of the GnT II acceptor substrate GnM$_3$-octyl was done in a total volume of 500 μl of buffer (0.1 M MES, pH 6.1, 1 mg/ml BSA, 0.05 M $CoCl_2$) containing 3.0 mM M$_3$-octyl and 10 mM UDP-GlcNAc. Synthesis was performed using purified recombinant rabbit GnT I (1 mU) overnight at room temperature. The reaction product was purified by sequential ion-exchange chromatography with Dowex-1X8 and Dowex-50X8 (Sigma) eluted in water. As a final preparation step, the substrate was applied to a Sep-Pak C$_{18}$ reverse phase cartridge (Whatman) and eluted with methanol. Synthesis of GnM$_3$-octyl was monitored by thin layer chromatography using silica gel plates (Merck) and CH_2Cl_2/methanol/H_2O (65:35:6, by volume) as the chromatography solvent. The chromatogram was sprayed with 0.2% of orcinol (Fluka) in 20% H_2SO_4 and heated for 10 minutes at 110°C. The yield of purified product was determined by amino sugar analysis [25].

Preparation of GGnM$_3$-octyl

Synthesis of GGnM$_3$-octyl was carried out in a total volume of 200 μl of buffer (0.05 M MES, pH 7.0, 1 mg/ml BSA, 20 mM $MnCl_2$) containing 0.5 mM GnM$_3$-octyl and 3.0 mM UDP-galactose. The reaction was performed using bovine β1,4-galactosyltransferase (~30 mU, Sigma) overnight at 30°C. GlcNAc was removed from non-converted residual GnM$_3$-octyl by incubation with 300 μl 0.1 M sodium citrate, pH 5.0 and ~0.3 U of β-*N*-acetylglucosaminidase overnight at 37°C. Monitoring of product synthesis and purification were carried out as described above for the preparation of the GnT II substrate. The concentration of the purified product was determined by amino sugar analysis [25].

Assay of GnT II activity

Standard GnT II activity assays were performed in a total volume of 20 μl of buffer containing 0.1 mM GnM$_3$-octyl and 0.1 mM UDP-[^{14}C]GlcNAc (3000–4000 c.p.m./nmol) as substrates and 0.1 M MES (pH 6.3), 0.5% Triton X-100, 0.5 mg/ml BSA, 0.5 mM PMSF, 0.5 mM DTT, 20 mM $MnCl_2$, 2.5 μg/ml leupeptin, 2.5 μg/ml E-64 and 2 μl of enzyme solution. After incubation at 37°C for one hour reactions were stopped by addition of 0.5 ml of 20 mM sodium tetraborate containing 2 mM EDTA. The radioactive product was isolated by anion-exchange chromatography and quantified by liquid scintillation counting [20].

Alternatively, GnT II activity was determined with 0.1 mM pyridylaminated GlcNAcMan$_3$GlcNAc$_2$ and 2 mM UDP-GlcNAc as substrates under the same reaction conditions as above. The reaction products were analysed by reversed-phase HPLC as reported previously [26]. *X. laevis* GnT II was also assayed with M$_3$-octyl and GGnM$_3$-octyl under the conditions described above.

Western blotting analysis

Cell lysates and culture supernatants of infected Sf21 cells were subjected to 12% SDS-PAGE under reducing conditions. Proteins after separation and subsequent blotting on Hybond-C membranes were detected with affinity-purified rabbit antibodies to *S. japonicum* GST (Santa Cruz Biotechnology) or with a mouse monoclonal antibody to the enterokinase recognition sequence (Invitrogen). Detection of bound antibodies was done as described previously [20].

Enzymatic deglycosylation with PNGase F

Purified protein (2 μg) was denatured in the presence of 0.5% (w/v) SDS and 0.05 M β-mercaptoethanol in a final volume of 10 μl for 5 min at 95°C. After adding 10 μl of buffer (0.3 M NaH_2PO_4, pH 7.2, 15 mM EDTA), 5 μl 7.5% Triton X-100 and 2.0 μl PNGase F (200 U/ml), the sample was incubated overnight at 37°C. The reaction was stopped with 120 μl methanol, the sample was then incubated for 30 min at room temperature and centrifuged for 10 min at 15,000 × g. Pellets were washed with 1 ml acetone, dried and resuspended in 40 μl of SDS-PAGE sample buffer and analysed by SDS-PAGE and Western blotting as above.

Other methods

Total cellular protein content was determined by the Bradford method with the Bio-Rad Protein Assay kit (Bio-Rad), using bovine serum albumin as a standard. Densitometric analysis of silver stained SDS-PAGE gels and Western blots was done using ImageQuaNT v4.2 software (Molecular Dynamics). Lectin blotting with digoxigenin-labelled concanavalin A and anti-digoxigenin antibodies (both from Roche) was done according to the instructions of the manufacturer.

Results

Isolation and analysis of cDNAs encoding *X. laevis* GnT II isoenzymes

Hybridization of approximately 4 × 10^5 independent plaques from a *X. laevis* ovary lambda ZAP II cDNA library with a DNA probe derived from the region encoding the catalytic domain of human GnT II [10] yielded two cDNAs with high identity in the coding region indicating their origin from closely related but clearly distinct genes. Upon comparison with other GnT II cDNA sequences, one cDNA clone (1.7 kb) was found to encode a putative 432 amino acid protein with a calculated molecular size of 50.6 kDa (GnT II-A). The other cDNA (1.9 kb) encoded a highly homologous protein with the same

number of amino acids (432) and a calculated molecular size of 50.0 kDa (GnT II-B). This second cDNA clone includes 315 bp of 5′- and 271 bp of 3′-flanking noncoding sequences. No additional ATG triplets were observed within the 5′-untranslated region. The 3′-untranslated region contains one presumptive polyadenylation signal at positions 1758 to 1764, but the poly (A) tail itself is absent. Cloning and DNA sequencing of genomic *X. laevis* GnT II-A and GnT II-B DNAs showed that there are no introns in the coding regions of the two genes.

Hydropathy plot analysis [27] of the predicted proteins indicated a similar domain topology with a putative hydrophilic cytoplasmic tail (C) of 9 amino acids at the N-terminus followed by a hydrophobic transmembrane domain (T) (residues 10–27). The luminal part of each protein consists of a putative stem region (S) and a carboxy-terminal catalytic domain consisting of a total of 405 amino acids. Thus, both *X. laevis* GnT II isoforms exhibit a type II transmembrane protein topology typical for Golgi-bound glycosyltransferases. The two proteins differ at 28 amino acid positions within the luminal part. The predicted GnT II isoenzymes A and B exhibit two potential N-glycosylation sites at positions N-59 and N-73, respectively (Figure 1). As calculated for GnT II-B, the overall identities of the *X. laevis* GnT II isoforms are 34%, 39%, 46% and 64% relative to their counterparts from *A. thaliana*, *C. elegans*, *D. melanogaster* and humans, respectively.

Hydrophobic cluster analysis revealed that both *X. laevis* GnT II isoforms exhibit three regions previously observed to be conserved between GnT I and GnT II from other species [28]. Region I represents the start of the putative catalytic region, region II constitutes the EED motif found in all GnT II species characterized to date [28,29] and region III has the D-327 residue proposed to be a catalytic base [24,28]. A BLAST search of the GenBank database retrieved two short EST clones from unfertilized *X. laevis* eggs, AW636033 (457 bp) and BF 426411 (506 bp) corresponding to cDNAs encoding isoenzymes A and B respectively.

Heterologous expression of recombinant *X. laevis* GnT II isoenzymes in insect cells

In order to confirm the authenticity of the cDNA clones, the coding regions were engineered for expression in *Spodoptera frugiperda* Sf21 insect cells using a baculovirus transfer vector. In initial studies the cDNA encoding GnT II-B was engineered for expression of two different soluble proteins, deleted in their hydrophobic transmembrane region and, to a

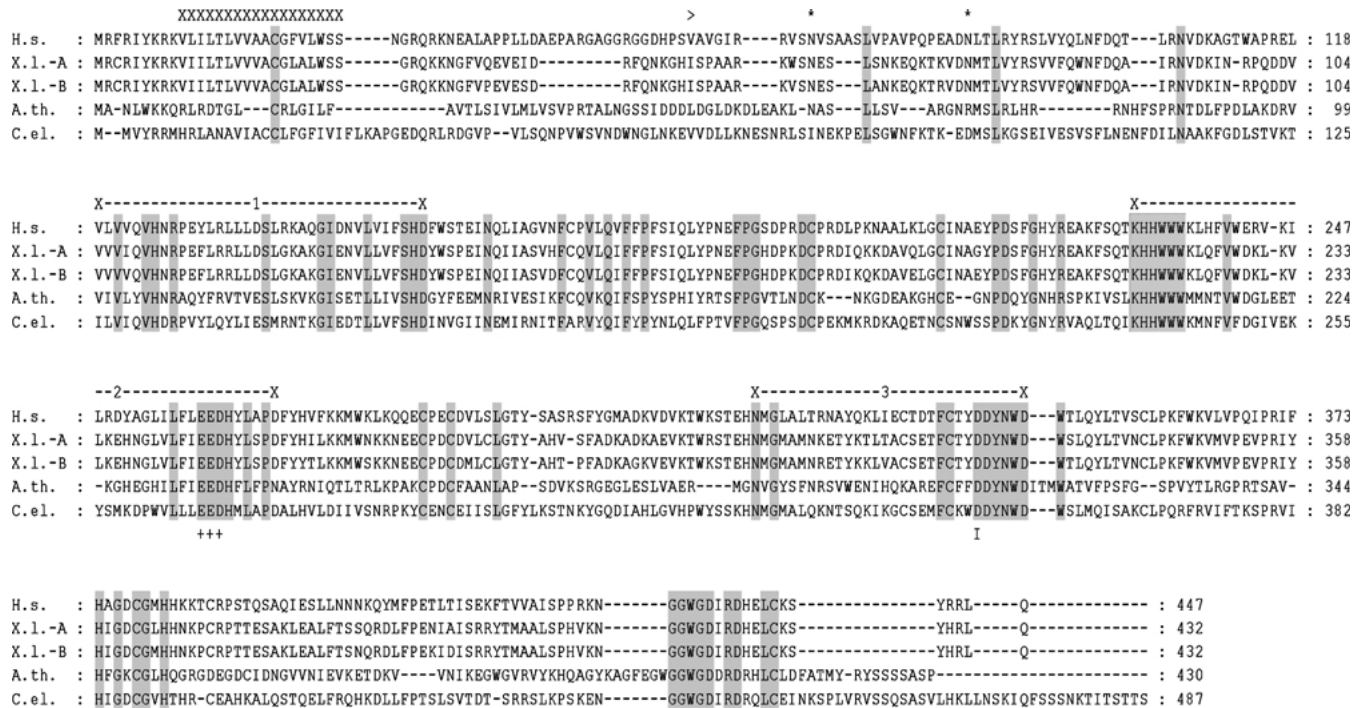

Figure 1. Multiple protein sequence alignment of *X. laevis* GnT II-A and GnT II-B with their counterparts from man, *A. thaliana* and *C. elegans*. Numbering refers to the amino acid position of the respective proteins: H.s. (H. sapiens), X.l.-A and X.l.-B (*X. laevis*), A. th. (*A. thaliana*), C.el. (*C. elegans*). The sequence alignments were generated using Clustal-W and refined manually. The strictly conserved residues are given in white letters on a black background. Dashes indicate gaps; (+ + +)- indicates the invariant acidic motif (EED, involved in donor substrate binding); xxx- proposed GnT II membrane-spanning region (http://azusa.proteome.bio.tuat.ac.jp); (>)- indicates the start of the truncated *X. laevis* GnT II forms used in this study. The proposed catalytic base (D-327) is designated (I), (X--------X) are highly conserved regions as observed by hydrophobic cluster analysis [28]; (*)-potential site for N-linked glycosylation of *X. laevis* GnT II.

variable extent, of the putative stem region. cDNA sequences encoding *X. laevis* GnT II-B devoid of its N-terminal 49 and 106 amino acids, respectively, were fused to the 3′ end of the gene encoding *Schistosoma japonicum* glutathione S-transferase (GST) within the baculovirus transfer vector. The resulting plasmids were inserted into the viral genome by homologous recombination, and the recombinant viruses thereby generated were used to infect Sf21 cells. Supernatants and lysates of cells infected with recombinant baculovirus were analyzed for GnT II activity *in vitro* using pyridylaminated Manα-1,6(GlcNAcβ-1,2-Manα-1,3)Manβ-1,4-GlcNAcβ-1,4-GlcNAc ($GlcNAcMan_3GlcNAc_2$-PA) as acceptor substrate. With this substrate no activity was detected in cell lysates and in conditioned media of uninfected cells. However, cell lysates and conditioned medium of the Sf21 cells expressing the GnT II-B fusion protein lacking the first 49 amino acid residues displayed a significant amount of GnT II activity. Production of GST-tagged GnT II was also monitored by immunoblotting with anti-GST antibodies showing presence of the fusion protein in the cell lysates and conditioned media. The apparent molecular mass of the GnT II-B fusion protein was in agreement with the theoretical molecular mass of 70 kDa (data not shown). When cell lysates and the corresponding conditioned media of Sf21 cells expressing GST fused to GnT II-B lacking the N-terminal 106 amino acids were assayed, no GnT II activity was detected although the fusion protein was present within the cells (data not shown). These results demonstrate that GnT II-B lacking its 49 (but not 106) N-terminal amino acid residues is properly folded and thus catalytically functional.

A truncated form of GnT II-A with a deletion of its first 49 amino acid residues was fused to a leader sequence containing a cleavable signal peptide, a hexahistidine tag (as a means to permit rapid purification of the enzyme), and an enterokinase cleavage site. Concomitantly, GnT II-B truncated as above, as well as human GnT II lacking its 82 N-terminal amino acids were engineered into the same vector. These constructs were inserted into the baculovirus genome, and the recombinant viruses thus generated were used to infect Sf21 cells. Radiometric GnT II assays were carried out on insect cell lysates and conditioned media using Manα-1,6(GlcNAcβ-1,2-Manα-1,3)Manβ-1-O-octyl (GnM_3-octyl) as acceptor substrate. Lysates of uninfected Sf21 cells showed a low level of endogenous GnT II activity (0.48 nmol of product formed per h per mg of total cellular protein). The corresponding culture medium was devoid of detectable GnT II activity. In contrast, culture medium conditioned by Sf21 cells expressing truncated *X. laevis* GnT II-A contained significant amounts of GnT II activity (32.9 nmol per h per mg of total cellular protein). The corresponding cell lysate had slightly elevated activity compared to the control lysate (1.09 nmol of product formed per h per mg of total cellular protein) indicating that most of the functional enzyme was secreted into the culture medium. Similar results were observed when the truncated forms of GnT II-B and human GnT II were expressed in insect cells (data not shown).

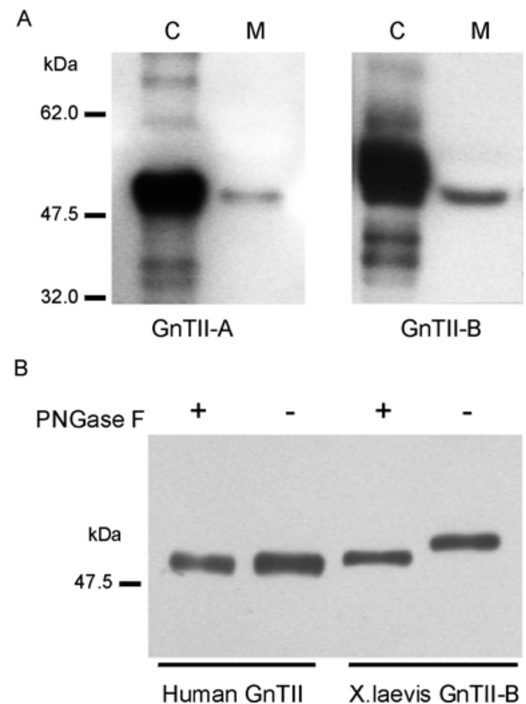

Figure 2. Characterisation of recombinant *X. laevis* GnT II produced in insect cells. (A) Heterologous expression of *X. laevis* GnT II isoenzymes in insect cells. Conditioned media (M) and protein extracts from Sf21 insect cells (C) infected with recombinant baculoviruses encoding the truncated versions of isoenzyme A and isoenzyme B were separated under reducing conditions by SDS-PAGE, transferred to a nitrocellulose membrane and probed with rabbit anti-enterokinase recognition site antibodies as outlined in *Materials and Methods.* The migration position of prestained bovine glutamic dehydrogenase (62.0 kDa), rabbit aldolase (47.5 kDa) and rabbit triosephosphate isomerase (32.0 kDa) are indicated. The detected size of the recombinant polypeptides was in close agreement with their theoretical molecular masses of 49.0 kDa (not accounting for any N-glycans). No specific signals were obtained with lysates and conditioned media of uninfected Sf21 cells. Please note that the medium samples correspond to ten times fewer cells than the cell lysates. (B) Analysis of *X. laevis* GnT II-B and human GnT II after treatment with PNGase F. Purified proteins ($\sim$2.0 μg) were incubated overnight in the presence (+) and absence (−) of PNGase F as described in *Materials and Methods*. Proteins separated by SDS-PAGE were transferred to a nitrocellulose membrane and probed with rabbit anti-enterokinase recognition site antibodies as above.

The results from heterologous protein expression demonstrate that both isoenzymes display significant catalytic activity and can be produced in insect cells as secreted soluble proteins lacking their 49 N-terminal amino acids. Recombinant protein production was also determined by immunoblotting with antibodies to the enterokinase cleavage site showing that the two *X. laevis* GnT II isoenzymes exhibit similar expression levels in insect cells (Figure 2A). Secreted His-tagged GnT II-A, GnT II-B and human GnT II were purified from Sf21 media

by nickel-chelate affinity chromatography. Radiometric GnT II assays using GnM_3-octyl as an acceptor substrate showed that isoenzymes A and B had similar specific activities (548 mU/mg and 433 mU/mg GnT II protein, respectively). Interestingly, both isoenzymes exhibited specific activities similar to that of purified recombinant human GnTII (488 mU/mg GnT II protein), assayed under the same reaction conditions.

Transfer of GlcNAc to the GnT I-substrate Manα-1,6(Manα-1,3)Manβ-1-O-octyl (M_3-octyl), was not detected, neither for GnT II-A nor for GnT II-B. The enzymes were also assayed with Manα-1,6(Galβ-1,4-GlcNAcβ-1,2-Manα-1,3)Manβ-1-O-octyl ($GGnM_3$-octyl), obtained by the action of UDP-galactose:GlcNAc β1,4-galactosyltransferase (GalT) on GnM_3-octyl. Less than 2% activity was observed with this substrate relative to GnM_3-octyl, indicating that substitution of the GlcNAcβ-1,2-Manα-1,3 branch with galactose prevents GnT II action.

X. laevis GnT II isoenzymes are N-glycosylated

Both *X. laevis* isoenzymes A and B have two potential N-glycosylation sites within the putative stem region at positions N-59 and N-73 (Figure 1). To determine whether these sites are indeed glycosylated *in vivo*, purified recombinant GnT II-A and GnT II-B were treated with peptide: N-glycosidase F (PNGase F). Western blot analysis of the proteins after treatment with the glycosidase showed a mobility shift of *X. laevis* GnT II-B (Figure 2B) and GnT II-A (data not shown). These results clearly indicate that both *X. laevis* isoenzymes are N-glycosylated. This conclusion was also confirmed by lectin blotting with concanavalin A (data not shown). In contrast, however, the single potential N-glycosylation site (N-86) of recombinant human GnT II is not used (Figure 2B). This is consistent with the results of Tan et al. [10], who suggested the absence of N-glycans on full-length recombinant human GnT II expressed in insect cells.

Expression of GnT II in adult *X. laevis* tissues

Northern blot analysis and RT-PCR were carried out to determine the expression of GnT II mRNA in adult frog tissues. Poly $(A)^+$-RNA from different tissues was hybridized with a cDNA encoding GnT II-B. A band of 2.7 kb was detected in ovary, liver, muscle, skin and lung (Figure 3A). The strong signal observed for liver is due to artifactual overloading of this lane (data not shown). RT-PCR analysis of total RNA was used to distinguish mRNA species coding for the two GnT II isoenzymes, by exploiting a diagnostic restriction site within the coding region of GnT II-B. While the transcript encoding GnT II-B was detected only in ovary and liver, GnT II-A mRNA was detected in all tested tissues (Figure 3B).

The genome of *X. laevis* harbors two related GnTII genes

In order to determine the GnT II copy number in the *X. laevis* genome, a Southern blot analysis was performed (Figure 4).

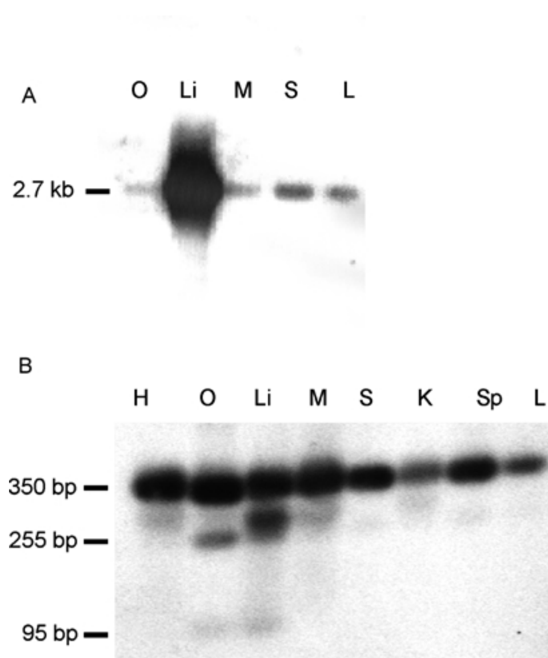

Figure 3. Expression of GnT II mRNA species in *X. laevis* tissues. (A) Northern blot analysis of poly $(A)^+$-RNA from various *X. laevis* tissues. Poly $(A)^+$-RNA ($\sim$1.0 μg) isolated from *X. laevis* ovary (O), liver (Li), muscle (M), skin (S) and lung (L), was electrophoresed on a 1.2% agarose/2.2 M formaldehyde gel. The mRNA was blotted onto a nylon membrane, probed with [α-^{32}P]dCTP-labeled isoenzyme B cDNA and finally washed in 0.5 $\times$ SSC/0.1% SDS at 60°C. The migration position of GnT II mRNA is indicated. Please note that the liver sample was artifactually overloaded. (B) RT-PCR products from different *X. laevis* tissues. RT-PCR products of total RNA from *X. laevis* heart (H), ovary (O), liver (Li), muscle (M), skin (S), kidney (K), spleen (Sp) and lung (L) were exhaustively treated with the restriction enzyme NdeI prior to analysis by Southern blotting with a ^{32}P-labeled *X. laevis* isoenzyme B cDNA. The amplified DNA sequence encoding isoenzyme B contains a diagnostic NdeI site that is not present in the respective cDNA sequence of isoenzyme A. The bands corresponding to isoenzyme B (255 bp and 95 bp) and isoenzyme A (350 bp) are indicated.

Total DNA from *X. laevis* liver was exhaustively digested with HindIII, EcoRI or BamHI. None of these enzymes cleaves within the coding sequences of either GnT II gene. The blot was hybridized with a DNA probe representing the coding region of isoenzyme B. Hybridization revealed two intense bands and several weak bands in the DNA samples. These results confirm that two GnT II genes occur in the *X. laevis* genome, with the possible additional presence of one or more pseudogenes.

Discussion

We have isolated and characterized cDNAs and the corresponding genomic DNAs from *X. laevis* encoding two closely related forms of GnT II, GnT II-A and GnT II-B and have used an insect

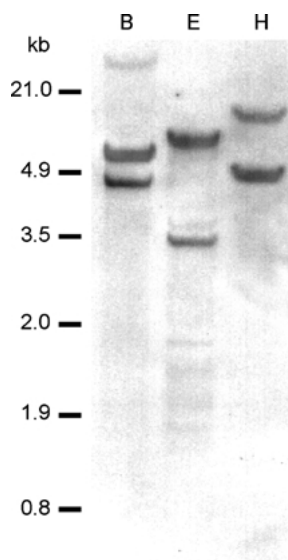

Figure 4. Southern blot analysis of genomic DNA from *Xenopus laevis*. *X. laevis* genomic DNA (15 μg) was digested to completion with BamHI (B), EcoRI (E) or HindIII (H). After DNA electrophoresis on a 0.8% agarose gel, denaturation and blotting onto a nylon membrane, the filter was hybridized with ^{32}P-labeled cDNA containing the entire coding region of isoenzyme B. The blot was finally washed in 0.5 × SSC/0.1% SDS at 65°C. The migration positions of DNA standards are indicated.

cell expression system to verify their enzymatic activities which are essentially identical to that of human GnT II.

The protein sequences of both isoforms share pronounced sequence similarity with their orthologues of mammalian, worm, fly and plant origin [10–13]. Analysis of the deduced amino acid sequences of both GnT II isoenzymes predicts the functional domain architecture typical for type II Golgi-bound glycosyltransferases that have been cloned to date. Despite a low sequence similarity with putative cytoplasmic-transmembrane-stem (CTS) regions of GnT II from man, worm and plants, hydrophobic cluster analysis (HCA) shows that both *X. laevis* GnT II isoforms share the conserved regions in the putative catalytic domain observed in all GnT I and GnT II enzymes cloned to date [28].

We have used the insect cell expression system to prove that the isolated genes encode enzymes with the expected catalytic properties of GnT II. Upon expression of the engineered DNAs, the isoenzymes were found to exhibit activity with the acceptor substrates GnM_3-octyl and $GlcNAcMan_3GlcNAc_2$-PA. When the terminal β1,2-linked GlcNAc residue coupled to α1,3-linked mannose is either absent or modified, no transfer of GlcNAc in β1,2-linkage to the α1,6 mannose branch is observed.

Soluble forms of recombinant GnT II-A and GnT II-B (both lacking their N-terminal 49 amino acid residues) were secreted into the culture medium in comparable amounts. The purified isoenzymes exhibit a similar specific activity even though the luminal segments differ in 28 amino acids. Interestingly, their specific activities are comparable with the specific activity of recombinant human GnT II expressed under the same conditions. This study shows that the N-terminal 49 amino acid residues of *X. laevis* GnT II are not required for GnT II activity. However, removal of a total of 106 amino acid residues from the N-terminus led to a complete loss of activity.

The three-dimensional structure of GnT II is as yet not known and there is little information on the minimal catalytic domain essential for GnT II activity. The crystal structure of rabbit GnT I has revealed that the catalytic domain of GnT I starts with a β strand at residue 107 [28]. Indeed, removal of 106 amino acid residues from the N-terminus of either rabbit or *X. laevis* GnT I has been shown to have no effect on enzyme activity [20,21]. However, in the case of GnT II equivalent N-terminal truncations of *X. laevis* GnT II-B (106 amino acids) and human GnT II (either 115 or 124 amino acids; J. Tan and H. Schachter, unpublished data) led to complete loss of activity. Nevertheless, recombinant human GnT II lacking its N-terminal 82 amino acids was found to be enzymatically active. Since amino acid position 82 of human GnT II corresponds to amino acid position 69 of the *X. laevis* enzymes, we presume that the catalytic region of *X. laevis* GnT II encompasses residues 70 to 432. The segment between residues 70 and 106 includes a stretch with pronounced homology between human and *X. laevis* GnT II (amino acids 72–96) which probably represents an important structural element for GnT II function. The catalytic domain of GnT II from *A. thaliana* requires an even more extended N-terminal region for proper folding of the protein since a 46 amino acid deletion from the N-terminus led to complete loss of activity (J. Mucha et al., unpublished data).

Absence of protein N-glycosylation may prevent the proper folding of nascent polypeptide chains. *X. laevis* and human GnT II contain two highly conserved potential N-glycosylation sites located in the putative stem region and close to the N-terminus of the catalytic domain, respectively. However, while GnT II-A and GnT II-B carry each at least one N-glycan, enzymatically active human GnT II lacking 82 amino acids from the N-terminus is not glycosylated at its remaining sequon (N-86). This suggests that improper folding of *X. laevis* GnT II missing its first 106 amino acids is not due to the absence of protein N-glycosylation.

This is the first report of multiple forms of GnT II found in a single organism. Previous studies have shown the existence of only one copy of the GnT II gene in various mammalian genomes [10]. Southern blot analysis indicates the presence of at least two gene copies in the *X. laevis* genome, both of which were identified to be transcriptionally active. Additional weak bands on Southern blots suggest the presence of transcriptionally inactive pseudogenes. Interpretation of gene copy numbers within the *X. laevis* genome is rather complicated, since this organism is a pseudotetraploid as a result of genome duplication 30 million years ago [30,31]. This event presumably resulted in the multiplication of the GnT II progenitor gene.

Our results show that the entire coding regions of both GnT II-A and GnT II-B are monoexonic as previously shown for

human, pig and rat GnT II [10,11,13]. A single band of 2.7 kb was detected in Northern blots of all tested tissues. The size of the mRNA indicates that the cloned cDNAs may not include the complete 5′- and 3′-ends of the untranslated regions. This is consistent with the observation that the 3′-end of the cDNAs did not contain a poly-A tail. While the mRNA encoding GnT II-A is expressed in all tested *X. laevis* tissues, the mRNA encoding GnT II-B was only detected in ovary and liver. Hence, the two GnT II forms present in *X. laevis* are only partially functionally redundant, with GnT II-A being of higher physiological relevance.

Acknowledgments

We thank Johannes Michel for purification and characterization of recombinant *X. laevis* isoenzyme A, Christelle Breton (Centre de Recherches sur les Macromolecules Vegetales, Grenoble, France) for assistance with hydrophobic cluster analysis, and Fritz Altmann and Thomas Dalik (Institut für Chemie, Universität für Bodenkultur Wien) for reagents and amino sugar analyses.

This project was supported in part by the Jubiläumsfonds der Österreichischen Nationalbank (project 5448) and the Austrian Science Foundation (project 14343).

References

1 Schachter H, *Biochem Cell Biol* **64**, 163–81 (1986).
2 Kornfeld R, Kornfeld S, *Annu Rev Biochem* **54**, 631–44 (1985).
3 Charuk JHM, Tan J, Bernardini M, Haddad S, Reithmeier RAF, Jaeken J, Schachter H, *Eur J Biochem* **230**, 797–805 (1995).
4 Jaeken J, Schachter H, Carchon H, De Cock P, Coddeville B, Spik G, *Arch Dis Child* **71**, 123–7 (1994).
5 Jaeken J, Spik G, Schachter H, *Glycoproteins and Disease* (Elsevier, Amsterdam, The Netherlands, 1996), pp. 457–67.
6 Tan J, Dunn J, Jaeken J, Schachter H, *Am J Hum Genet* **59**, 810–7 (1996).
7 Schachter H, Jaeken J, *Biochim Biophys Acta* **1455**, 179–92 (1999).
8 Wang Y, Tan J, Campbell RM, Ditto D, Le D, Schachter H, Marth JD, *Glycobiology* **10**, 1131–2 (2000).
9 Wang Y, Tan J, Sutton-Smith M, Ditto D, Panico M, Campbell RM, Varki NM, Long JM, Jaeken J, Levison SR, Wynshaw BA, Morris HR, Le D, Dell A, Schachter H, Mart JD, *Glycobiology* **11**, 1051–70 (2001).
10 Tan J, D'Agostaro GAF, Bendiak B, Reck F, Sarkar M, Squire JA, Leong P, Schachter H, *Eur J Biochem* **231**, 317–28 (1995).
11 D'Agostaro GAF, Zingoni A, Moritz RL, Simpson RJ, Schachter H, Bendiak BJ, *J Biol Chem* **270**, 15211–21 (1995).
12 Campbell R, Tan J, Schachter H, Bendiak B, Marth J, *Glycobiology* **7**, 1050 (1997).
13 Leeb T, Kriegesmann B, Baumgartner BG, Klett C, Yerle M, Hameister H, Brenig B, *Biochim Biophys Acta* **1336**, 361–6 (1997).
14 Chen S, Tan J, Reinhold VN, Spence AM, Schachter H, *Biochim Biophys Acta* **1573**, 271–9 (2002).
15 Strasser R, Steinkellner H, Boren M, Altmann F, Mach L, Glössl J, Mucha J, *Glycoconjugate J* **16**, 787–91 (1999).
16 Santacruz-Toloza L, Huang Y, John SA, Papazian, DM, *Biochemistry* **33**, 5607–13 (1994).
17 Khanna R, Myers MP, Laine M, Papazian DM, *J Biol Chem* **276**, 34028–34 (2001).
18 Roitsch T, Lehle L, *Eur J Biochem* **181**, 733–9 (1989).
19 Cantor AB, Kornfeld S, *Anal Biochem* **205**, 220–6 (1992).
20 Mucha J, Svoboda B, Fröhwein U, Strasser R, Mischinger M, Schwihla H, Altmann F, Hane W, Schachter H, Glössl J, Mach L, *Glycobiology* **11**, 769–78 (2001).
21 Sarkar M, Pagny S, Unligil UM, Joziasse D, Mucha J, Glössl J, Schachter H, *Glycoconjugate J* **15**, 193–7 (1998).
22 Chomczynski P, Sacchi N, *Anal Biochem* **162**, 156–9 (1987).
23 Ausubel FM et al. (ed.), *Current Protocols in Molecular Biology* (John Wiley & Sons, Inc., 1994–2002), Vol. 1, Ch. 2.
24 Unligil UM, Zhou S, Yuwaraj S, Sarkar M, Schachter H, Rini JM, *EMBO J* **19**, 5269–80 (2000).
25 Altmann F, *Anal Biochem* **204**, 215–9 (1992).
26 Altmann F, Kornfeld G, Dalik T, Staudacher E, Glössl J, *Glycobiology* **3**, 619–25 (1993).
27 Kyte J, Doolittle RF, *J Mol Biol* **157**, 105–32 (1982).
28 Breton C, Mucha J, Jeanneau C, *Biochimie* **83**, 713–8 (2001).
29 Wiggins CAR, Munro S, *Proc Natl Acad Sci USA* **95**, 7945–50 (1998).
30 Bisbee CA, Baker MA, Wilson AC, Haji-Azimi I, Fischberger M, *Science* **195,** 785–7 (1977).
31 Amaya E, Offield MF, Grainger RM, *Trends Genet* **14**, 253–5 (1998).

Received 7 January 2003;
accepted 11 February 2003