Expression and characterization of the first snail-derived UDP-*N*-acetyl-α-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase

**Christopher Taus · Chantal Lucini · Takeshi Sato ·
Kiyoshi Furukawa · Reingard Grabherr ·
Erika Staudacher**

Abstract UDP-GalNAc:polypeptide GalNAc transferase
(ppGalNAcT; EC 2.4.1.41) catalyzes the first step in mucin-
type O-glycosylation. To date, several members of this large
enzyme family have been analyzed in detail. In this study we
present cloning, expression and characterization of the first
representative of this type of glycosyltransferase from mollusk
origin, namely from *Biomphalaria glabrata*. The full length
sequence of the respective gene was obtained by screening of
a cDNA library using homology-based PCR. The entire gene
codes for a protein consisting of 600 amino acids comprising
the features of a typical type II membrane protein containing a
cytoplasmic tail at the N-terminus, a transmembrane and a
catalytic domain as well as a ricin-like motif at the C-terminus.
Sequence comparison with ppGalNAcTs from various species
revealed high similarities in terms of structural architecture.
The enzyme is O-glycosylated but does not have any putative
N-glycosylation sites. All four tested acceptor peptides were
functional substrates, with Muc2 being the best one. Further
biochemical parameters tested, confirmed a close relationship
to the family of yet known ppGalNAcTs.

Abbreviations

Bge cells	*Biomphalaria glabrata* embryonic cells
GalNAc	*N*-acetylgalactosamine
GlcNAc	*N*-acetylglucosamine
MES	2-(*N*-morpholino)ethanesulphonic acid
ppGalNAcT	polypeptide *N*-acetylgalactosaminyltransferase
Sf9-cells	*Spodoptera frugiperda* cells

C. Taus · C. Lucini · E. Staudacher
Department of Chemistry, University of Natural Resources and Life
Sciences, Vienna, Muthgasse 18, 1190 Vienna, Austria

T. Sato · K. Furukawa
Laboratory of Glycobiology, Department of Bioengineering,
Nagaoka University of Technology, Kamitomioka 1603-1,
Niigata 940-2188, Nagaoka, Japan

R. Grabherr
Department of Biotechnology, University of Natural Resources and
Life Sciences, Vienna, Muthgasse 18, 1190 Vienna, Austria

E. Staudacher (✉)
Department of Chemistry, Glycobiology, University of Natural
Resources and Life Sciences, Vienna, Muthgasse 18,
1190 Vienna, Austria
e-mail: erika.staudacher@boku.ac.at

Introduction

Mucin-type O-glycosylation is the most common type of O-
glycosylation. Besides mammals it has been found in insects,
snails, worms, and various parasites [1, 2]. The initiating en-
zymes of this type of glycosylation are members of an evolu-
tionarily conserved family of UDP-GalNAc:polypeptide
GalNAc transferases (ppGalNAcTs, [EC 2.4.1.41]), which
transfer a GalNAc residue from UDP-GalNAc to a serine
(Ser) or threonine (Thr) within a polypeptide chain. Depending
on the organism, tissue and developmental stage, the protein
bound GalNAc residue is elongated first by a GlcNAc and/or
galactose residue and subsequently is modified by more of
these or other monosaccharides (fucose or sialic acid).
Mucin-type O-glycosylation has been suggested to influence
conserved processes in development and changes in O-
glycosylation patterns have been connected with pathogenic

events. Therefore, understanding the biosynthesis of these structures, as well as their biological role is of high interest in basic and applied research.

In humans already 20 representatives of this enzyme family have been identified and several more in other organisms. They have been characterized and grouped into subfamilies according to conserved structural as well as functional aspects [3]. These sub-families differ in details of structure, tissue distribution and their acceptor peptide preferences. Yet, some common structural elements have been identified: ppGalNAcT is a type II membrane bound glycosyltransferase located in the Golgi apparatus with a highly conserved UDP-GalNAc-binding region, a manganese-binding site and quite often a lectin-like domain located at the carboxyl terminus, which seems to have a major influence on substrate specificity [4, 5].

So far, only few invertebrates have been characterized in terms of their ppGalNAcT activities. As a representative of insects *Drosophila melanogaster* has been found to contain 14 members of this enzyme family with some of them being essential for normal development [6, 7]. In the nematode *Caenorhabditis elegans* 11 homologues of vertebrate ppGalNAcTs were identified showing 60–80 % amino acid sequence similarity [8]. Also parasites (*Fasciola hepatica*, *Trypanosoma cruzi* and *Toxoplasma gondii*) contain ppGalNAcTs initiating their O-glycosylation [9–11]. The recent review by Bennett [3] gives a comprehensive coverage of the ppGalNAcTs that have been discovered so far in different organisms.

In the course of our previous work, we identified small mucin type O-glycans containing a protein bound GalNAc-residue elongated by galactose residues and, in some cases also fucose residues, in several snail species (*Arion lusitanicus*, *Achatina fulica*, *Biomphalaria glabrata*, *Cepaea hortensis*, *Clea helena*, *Helix pomatia*, *Limax maximus* and *Planorbarius corneus*) [2]. Many of the terminal hexoses were found to be methylated which is a typical structural element in snails [12, 13].

In the current study we present the identification, molecular cloning and characterization of the first ppGalNAcT from mollusk origin, in particular from *Biomphalaria glabrata*, a snail living in fresh water. The identified enzyme initiates mucin-type O-glycosylation and shares all important domains with the other members of the family. Our findings are another proof that O-glycosylation is a highly conserved process of protein modification throughout the animal kingdom.

Material and methods

Materials All DNA manipulations were carried out essentially as summarized by Sambrook *et al.* [14], restriction enzymes, T4 ligase, T4 polynucleotide kinase, DNA Polymerase I, Large (Klenow) Fragment and calf intestine alkaline phosphatase were purchased from New England Biolabs (Frankfurt, Germany). All enzymes were used according to manufacturers' recommendations. All primers were synthesized commercially (Sigma-Aldrich, Vienna, Austria). Embryonic cells from *Biomphalaria glabrata* (Bge cells, NR-40248, BEI Resources, NIAID, NIH) were grown as described at 26 °C [15]. Chemically transformed *Escherichia coli* (High Efficiency) cells (New England Biolabs, Frankfurt, Germany), were spread on Lysogeny broth (LB) agar plates containing 100 µg/ml Ampicillin and incubated overnight at 37 °C. *Spodoptera frugiperda* cells (Sf9, ATCC CRL-1711) were cultivated in IPL41 medium (SAFC Biosciences, St. Louis, USA) containing yeast extract, a lipid mixture supplemented with 10 % fetal calf serum, at 27 °C [16].

Synthesis of a cDNA library from Biomphalaria glabrata Total RNA was isolated from Bge cells by RNeasy Mini Kit (Qiagen, Hilden, Germany). Polyadenylated RNA was purified from 12 µg of total RNA by MicroPoly(A)Purist™ Kit (Ambion, Austin, USA). 100 ng of polyadenylated RNA were used as a template for first strand cDNA synthesis by In-Fusion® SMARTer™ Directional cDNA Library Construction kit (Takara, Saint-Germain-en-Laye, France). Double-stranded cDNA was synthesized according to the manufacturer's manual, ligated and transformed chemically into NEB 10-beta Competent *Escherichia coli* cells. Recombinant *E. coli* colonies were used as a template for control PCR using the primer set 5′TCACACAGGAAACAGCTATGA3′ and 5′CCTCTTCGCTATTACGCCAGC3′.

Homology-based PCR Highly conserved regions were identified by alignment of ppGalNAcT genes from *Homo sapiens* 2 (NP_004472), *Mus musculus* (NP_038842), *Caenorhabditis elegans* GLY-3 (NP_498722) and *Drosophila melanogaster* (AAQ56699). Based on the conserved amino acid sequences WGGEN and VWMDEY/F sense and antisense degenerated primers (5′TGGGGWGGWGAR AACYT3′ and 5′TGWAYTCRTCCATCCADAC3′) were designed. The resulting 187 base pair product was purified by NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel, Düren, Germany), ligated with the pGEM-T Easy vector (Promega, Fitchburg, USA) and sequenced to examine identity of the fragment.

Isolation of the full-length snail ppGalNAcT gene Recovery of the complete sequence of this novel ppGalNAcT gene was obtained by inverse PCR of self-ligated cDNA, as well as a 5′ and 3′ RACE PCR, respectively. Double-stranded cDNA (3.5 µg) was precipitated with ethanol, resuspended in

NEBuffer 4 and incubated with DNA Polymerase I, Large (Klenow) Fragment in a total volume of 25 μl for 15 min at 25 °C. After DNA purification and ethanol precipitation the cDNA was treated with 10 units of T4 Polynucleotide Kinase in a total volume of 30 μl for 30 min at 37 °C. The enzyme was inactivated at 65 °C for 20 min. T4 DNA Ligase was used for the self-ligation reaction (16 °C, over night). Correct self-ligated cDNAs coding for the ppGalNAcT gene were amplified by inverse PCR using the primer set 5′GAA ACACAAAGCGTGCAGCAGAAG3′ and 5′CACTGCCA AACACGGAAGGATATC3′. The amplified products were purified and ligated with the pGEM-T Easy vector for sequencing.

5′ and 3′ Ready RACE cDNA was synthesized from 1 μg of total RNA by using the SMARTer™ RACE cDNA Amplification Kit (Takara, Saint-Germain-en-Laye, France). The complete 5′ end of the novel ppGalNAcT gene was amplified by using the Phusion High-Fidelity DNA Polymerase (Thermo Scientific, Vienna, Austria), 5′ Ready RACE cDNA as template, the gene specific primer 5′CATCCAGACTTCTG CTGCACGCTTTGTG3′ and the 10X Universal Primer A Mix provided by the kit. The amplified 5′ end was purified and ligated with pUC19 vector for sequencing. A nested PCR resulted in the amplification of the complete 3′ end using ExTaq Polymerase (Takara, Saint-Germain-en-Laye, France) and 3′ Ready RACE cDNA as template. The amplified 3′ end was purified and ligated with pGEM-T Easy vector for sequencing.

The nucleotide sequence reported in this paper has been deposited in the GenBank database: *GenBank*: KC18251.

Recombinant expression, purification and Western blot analysis A cDNA fragment without cytoplasmic tail and transmembrane (amino acid 1–26) domain was amplified by using the forward primer 5′GGGGAGCTCGGAGATGAT CAAAGTGAGTTTG3′ and the reverse primer 5′GGGGGT ACCCTATCTGTTTTTACTTAAAGAAAATGTCCAC3′. The purified ppGalNAcT fragment was ligated with the pGEM-T Easy vector, cut, purified and ligated to pVT-Bac [17]. To verify the gene sequence and the correct insertion into the cloning site, this construct was sequenced. 2 μg of the recombinant construct (containing the honeybee melittin secretion signal, a 6 x His tag and the ppGalNAcT fragment) were co-transfected with 300 ng linearized BD BaculoGold™ Bright Baculovirus DNA (Becton Dickinson, Schwechat, Austria) into Sf9 insect cells. The secreted recombinant protein was purified on Ni-NTA agarose (Qiagen, Hilden, Germany) and analyzed by SDS-PAGE and Western blotting using Penta-His monoclonal antibody (Qiagen, Hilden, Germany; dilution 1:2000) followed by alkaline phosphatase conjugated anti-mouse IgG from goat (Sigma-Aldrich, Vienna, Austria; dilution 1:3000) [18].

Fetuin was used as a negative control for Penta-His monoclonal antibody.

Analysis of the recombinant ppGalNAcT activity The enzyme activity of the recombinant ppGalNAcT was determined in 20 μl reaction mixture containing 50 mM MES (2-(*N*-morpholino)ethanesulfonic acid), pH 7.0, 10 mM $MnCl_2$, 0.1 % Triton X-100, 40 nmol UDP-GalNAc (Sigma-Aldrich, Vienna, Austria), 20 μg acceptor peptide (Cellmano Biotech Co., Ltd., Shanghai, China) and 2 μl enzyme solution (protein concentration 5.75 μg/ml) at 37 °C for 45 min. The peptides were chosen randomly: Muc1a (APPAHGVTSAPDTRPA PGC) and 1a' (AHGVTSAPDTR) were taken from [19], Muc2 (PTTTPITTTTVTPTPTPTGTQTK) from [20] and Muc5Ac (GTTPSPVPTTSTTSAP) from [21]. The reaction was terminated by boiling at 96 °C for 5 min and analyzed by HPLC on a reversed phase C18 column (4.6×250 mm, 5 μm, Thermo Scientific, Vienna, Austria) in 0.1 % trifluoroacetic acid in water, applying a linear gradient from 10 to 30 % of eluent (0,1 % trifluoroacetic acid in acetonitrile) in 25 min at a flow rate of 1 ml/min. The long term incubation was carried out for 24 h analogous to the standard assay with additions of 10 nmol UDP-GalNAc and 2 ul of enzyme solution every 3 hours.

For the analysis of the biochemical parameters the standard assay conditions using Muc2 as acceptor peptide were modified as follows. For determination of the manganese optimum the concentration of $MnCl_2$ was varied from 0–300 mM, for the determination of cation requirement the standard assay was carried out without any cation addition or in the presence of 10 mM of EDTA, Mn^{2+}, Mg^{2+}, Ca^{2+}, Co^{2+}, Cu^{2+}, Ni^{2+}, or Ba^{2+}. The determination of enzyme stability and the pH-optimum were done according to Peyer *et al.* [22]. Kinetic data were acquired using 1:4 diluted enzyme solution and Muc2 at eleven different concentrations ranging from 0.01 to 0.82 mM, with UDP-GalNAc kept constant at a concentration of 2 mM. Similarly UDP-GalNAc was varied between concentrations from 0.01 to 2 mM while Muc2 was kept constant at 0.82 mM. Km-values were obtained from Lineweaver-Burk plots.

All quantitative values were calculated from the area of HPLC patterns. Each assay was carried out at least in duplicate with appropriate controls.

Determination of protein content Protein concentrations were determined by the Micro-BCA protein assay (Pierce, Bonn, Germany) with bovine serum albumin as the standard.

MALDI-TOF MS analysis MALDI-TOF MS analysis was carried out on an Autoflex Speed MALDI-TOF (Bruker Daltonics, Germany) equipped with a 1,000 Hz Smartbeam.II laser in positive mode using α-cyano-4-hydroxycinnamic acid

as matrix. Spectra were processed with the manufacturer's software (Bruker Flexanalysis 3.3.80).

Results

Isolation of the full-length ppGalNAcT gene A full-length cDNA library from Bge cells was created in order to isolate and characterize the first mollusk ppGalNAcT. The double-stranded cDNA synthesized from purified polyadenylated RNA was cloned and propagated in *E.coli*, resulting in a full length cDNA library. A control PCR of 50 randomly selected colonies showed high diversity with inserts ranging from about 500 up to 9,000 base pairs (data not shown).

Our PCR strategy was based on homology to the conserved regions (amino acid sequence WGGEN and VWMDEY/F) of four different ppGalNAcTs (*Homo sapiens* 2 NP_004472, *Mus musculus* NP_038842, *Caenorhabditis elegans* NP_498722 and *Drosophila melanogaster* AAQ56699). The worm and the insect enzyme were chosen due to their estimated closeness to mollusks, the two mammalian enzymes were chosen arbitrarily. By using degenerated primers an 187 base pair product was amplified. In order to receive the full-length ppGalNAcT gene, an inverse PCR of self-ligated cDNA was performed, which resulted in the amplification of a 1,659 base pair product coding for 553 amino acids (6 – 558). The remaining 5′ (amino acid 0 – 417) and 3′ end (amino acid 481 – 600) was amplified by RACE-PCR. Thereby, the full-length ppGalNAcT gene could be recovered, which turned out to be coding for a 600 amino acid type-II membrane protein containing a putative N-terminal cytoplasmic tail, a transmembrane domain (amino acid 7–25, predicted by TMHMM Server v. 2.0), a stem region, a luminal catalytic domain and a ricin-like motif at the C-terminus similar to almost all of the ppGalNAcTs described.

Purification and determination of specificity A truncated version of the Bge ppGalNAcT, without transmembrane domain and with the addition of an insect specific secretion signal, was expressed in Sf9 cells. The supernatant was tested for ppGalNAcT enzyme activity, as well as the supernatant of Sf9 cells expressing another glycosyltransferase. Whereas the supernatant of cells infected with ppGalNAcT expressing baculoviruses showed high enzyme activity, no transfer at all was detected in the control supernatant, confirming no interfering ppGalNAcT activity derived from the baculovirus insect cell expression system. Recombinant ppGalNAcT was further purified using metal chelate affinity chromatography. Quality of the purification was determined by Coomassie staining and immunoblotting using anti-His antibody. A band at approximately 65 kDa was detected, which correlated with the molecular weight as calculated

by the sum of the amino acids of the truncated form (Figs. 1a and b). For further characterization, the purified sample (specific activity 46 milliunits/ml; one unit is defined as the amount of enzyme that transfers 1 μmol of GalNAc in 1 min of the standard reaction mixture) was subjected to enzyme activity assays. Four different peptides, which had been successfully used in previous studies, were chosen as acceptors for enzyme activity. Qualitative analysis was done by MALDI-TOF MS; for quantification of the substrate conversion the relative amount of peak areas of the HPLC patterns were used for calculation.

All four arbitrarily selected peptides were functional acceptor substrates indicating that the snail derived ppGalNAcT displays a broad specificity. Similar to already described ppGalNAcTs, we observed the transfer of more than one GalNAc residue upon extended incubation times. Muc2 turned out to be the best acceptor where the addition of a second GalNAc residue was detectable already after 90 min and clearly visible after 120 min (Fig. 2a–d and 3). Using higher enzyme concentrations extension by a third GalNAc residue could be visualized within the standard incubation time of 2 h. By increasing the incubation time up to 24 h it was possible to detect a transfer of 8 GalNAc residues onto Muc2 acceptor peptide (Fig. 2e–h). On Muc5Ac, which was the second best acceptor peptide, traces of a second sugar residue could be seen after 2 h (Fig. 4), but less than 50 % of Muc1a or Muc1a' were converted into monoglycosylated peptides within 2 h of incubation time.
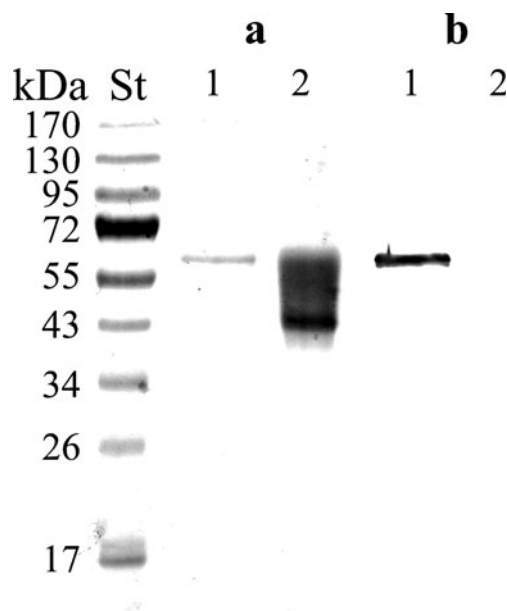

Fig. 1 Purification control of recombinant snail ppGalNAcT by SDS-PAGE and Western blot analysis (lane 1), **a** Coomassie staining **b** Immunoblotting of purified ppGalNAcT with Penta-His monoclonal antibody; Fetuin (lane 2) was used as a negative control for Penta-His monoclonal antibody
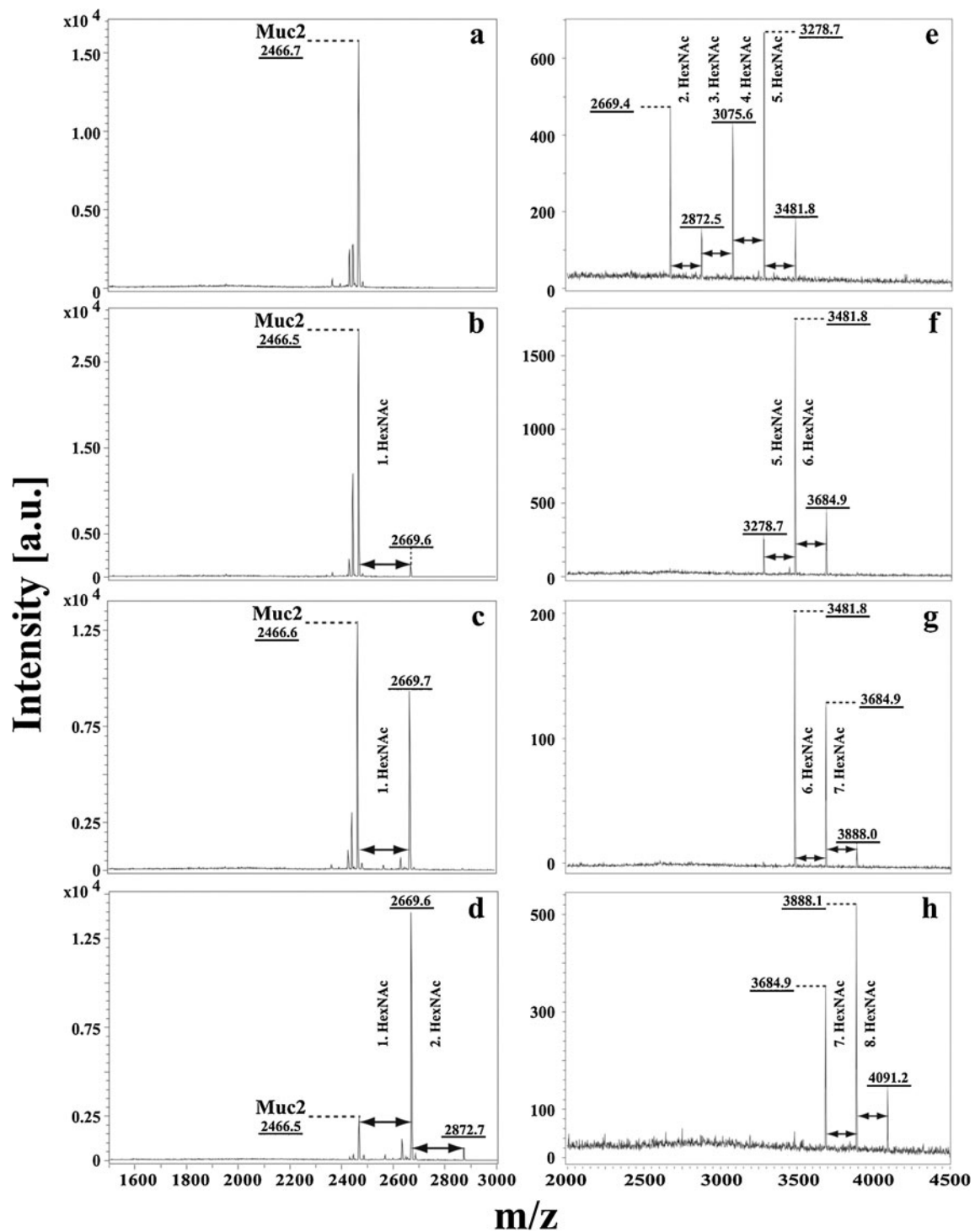
Fig. 2 Mass spectrometric ([M+Na]$^+$) analysis of the transfer of GalNAc using Muc2 as acceptor peptide. Muc2 peptide (**a**); Assays with incubation times of 10 min (**b**), 60 min (**c**), 2 h (**d**), 3 h (**e**), 6 h (**f**), 9 h (**g**) and 24 h (**h**)

Properties of the enzyme According to the amino acid sequence no N-glycosylation site was expected, however, O-glycans linked to the recombinant enzyme expressed in insect cells could be detected by lectin blot analysis using peanut agglutinin in combination with O-glycosidase treatment (data not shown).

The activity of the recombinant ppGalNAcT was not affected by storage for 72 h in a temperature range from −20 °C to room temperature, or by the addition of up to 10 % of methanol or glycerol. Addition of 5 % of acetonitrile had no negative effect on enzyme activity, whereas 10 % of acetonitrile or lyophilization reduced enzyme activity to
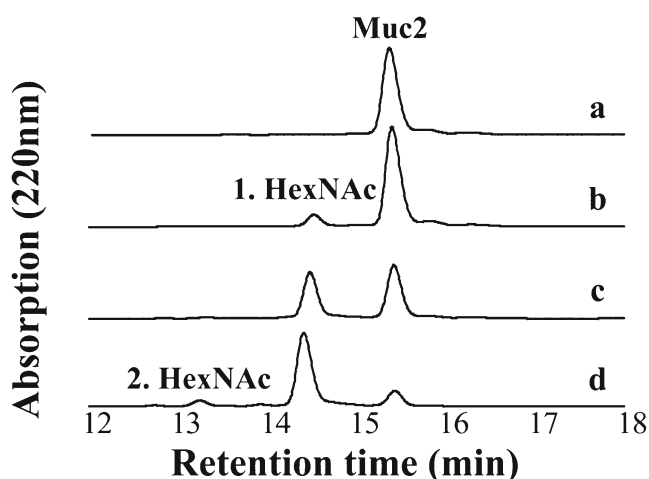
Fig. 3 HPLC analysis of the transfer of GalNAc onto Muc2 acceptor peptide. Muc2 peptide (**a**), incubation for 10 min (**b**), 60 min (**c**) and 2 h (**d**)

about 60 %. The activity was drastically reduced at storage temperatures above room temperature. After 72 h at 37 °C less than 10 % of activity was detectable. However, 37 °C was the optimal incubation temperature for short assays up to 2 h (data not shown). Further, ppGalNAcT showed a pH-optimum at 6.0–6.5 using MES as the appropriate buffer salt (Fig. 5a). The enzymatic transfer was dependent on divalent cations (no activity in the presence of EDTA) with the order of increasing support being $Mn^{2+}>Co^{2+}>Mg^{2+}>Ca^{2+}$. Cu^{2+}-ions abolished activity completely (Fig. 5b). Standard assays resulting in nearly 100 % of conversion into product were also carried out with half amount of enzyme to confirm the validity of the original assay. Maximal rates of transfer were achieved with Mn^{2+}-concentrations from 10–20 mM. Concentrations above 80 mM reduced the enzyme activity.

Analysis of enzyme activity revealed a K_M of 0.064 mM, with V_{max} at 0,16 nmol/min/μg for Muc2 acceptor peptide and a K_M of 0.046 mM and V_{max} at 0,054 nmol/min/μg for UDP-GalNAc in the presence of constant amounts of Muc2. Both values are within the range of published data (K_M 0.01–0.23 mM [19, 23]) for human ppGalNAc T2 using Muc2 acceptor, however the experiments have been performend under slightly different conditions. No kinetic data are available for the closely related enzymes from other invertebrates.

Discussion and conclusion

ppGalNAcTs are the largest family of glycosyltransferases, which catalyze a single linkage. They can be found throughout the animal kingdom, *e.g.* in mammals, insects and worms and are highly conserved regarding their amino acid sequence.

Mollusks are a large and diverse phylum of invertebrates, living in marine and freshwater as well as in terrestrial habitats.
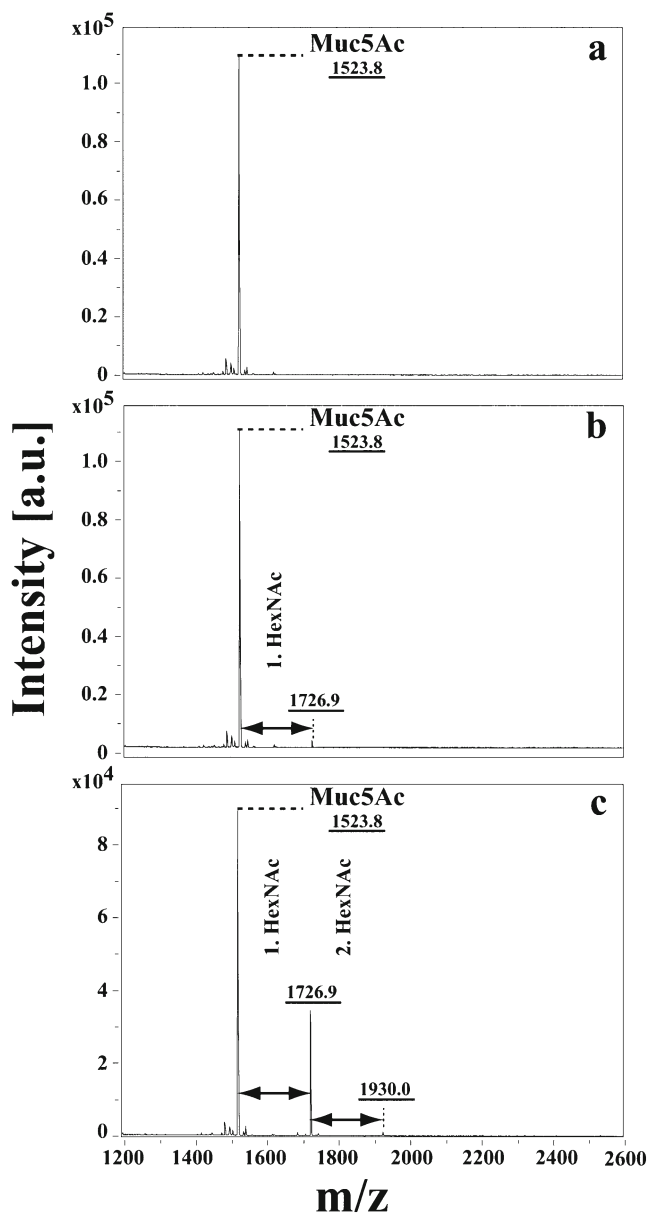
Fig. 4 Mass spectrometric ($[M+Na]^+$) analysis of the transfer of GalNAc onto Muc5Ac acceptor peptide. Muc5Ac peptide (**a**); Standard assay conditions using Muc5Ac as acceptor with incubation times of 20 min (**b**) and 120 min (**c**)

They span a wide diversity in terms of body size, body shape and the size of the brain. Thus, developmental processes are expected to be regulated quite differently. Yet, findings of conserved pathways sharing a common function with other organisms points towards a common origin in the biological past.

We previously found O-glycosylation with GalNAc being the protein bound sugar in several snail tissues [2] and thus, we expected to be able to identify a ppGalNAcT gene representative of this enzyme family from snail origin. Though, at the time we started our project, no complete mollusk genome was available. There were just parts of *Biomphalaria glabrata*
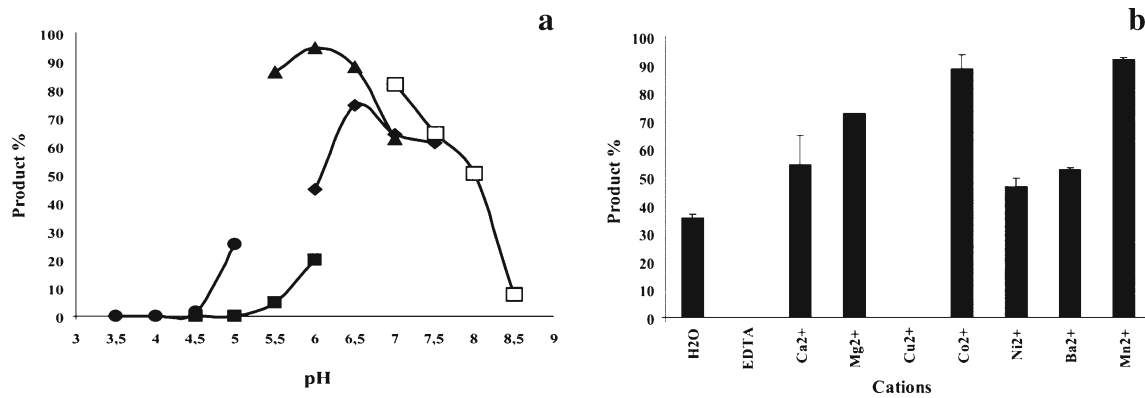
Fig. 5 Properties of ppGalNAcT from Bge cell line. **a** pH pattern using the following buffer systems: (●–●) acetate/NaOH, (■–■) citrate/NaOH, (▲–▲) MES, (♦–♦) phosphate and (□–□) Tris/HCl; **b** Influence of divalent cations on enzyme activity

(*Biomphalaria glabrata* genome initiative; http://biology. unm.edu/biomphalaria-genome/) and *Lottia gigantea* (http:// genome.jgi-psf.org/Lotgi1/Lotgi1.home.html) sequenced. Data existed from *Lymnea stagnalis*, which described a GalNAcT activity acting on terminal residues of N-glycans [24]. This enzyme had been characterized but not produced recombinantly. In the meantime, the complete sequence from Pacific oyster (*Crassostrea gigas*) has been published but no expression or characterization of glycosyltransferases has been performed [25].

In search for the snail ppGalNAcT gene we started by screening a full-length cDNA library, which we had generated from a *Biomphalaria glabrata* Bge cell line, with degenerated primers that were homologous to highly conserved parts of ppGalNAcTs that were published so far. As concluded from

sequence comparisons, the first piece of sequence obtained (coding for the amino acids 359–420) showed high homology to ppGalNAcTs from several other organisms (*e.g.* 92 % identity to *Crassostrea gigas* (EKC38600), 92 % to *Homo sapiens* 2 (NP_004472), 65 % to *Drosophila melanogaster* (AAQ56699) and 76 % to *Caenorhabditis elegans* (NP_ 498722)). After three further PCR steps we succeeded in the elucidation of the complete sequence of the full length snail derived ppGalNAcT, comprising a total of 600 amino acids. In the first part of the catalytic domain the glycosyltransferase 1 (GT1) structural motif (residues 166–272) described by [26], containing the conserved $D^{252}xH^{254}$ sequence was identified. This conserved motif has been shown to be involved in Mn^{2+} coordination [27]. C-terminal of the $D^{252}xH^{254}$ sequence there were two cysteine residues (255 and 257) located, which have
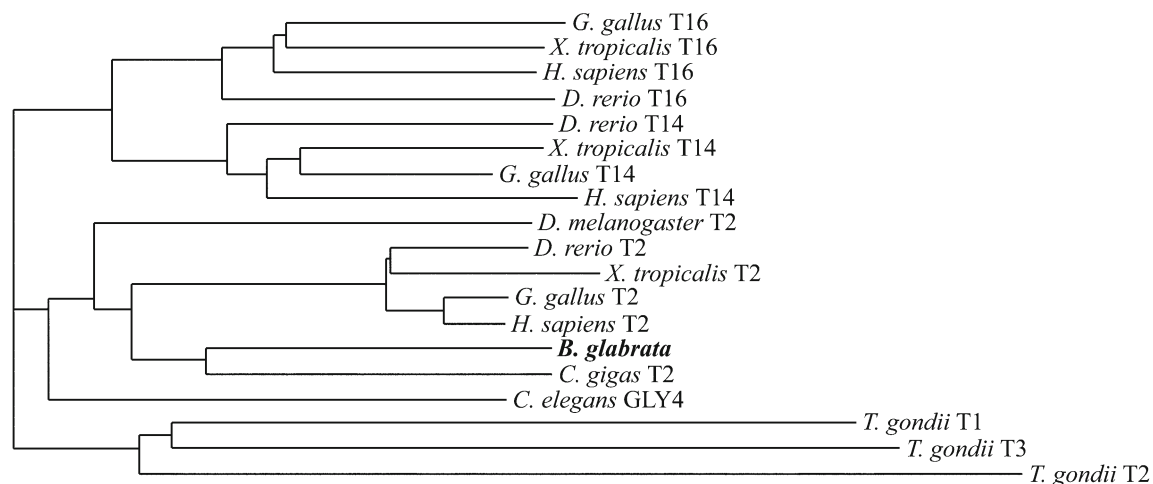
Fig. 6 Branch (group Ib according to [3]) of the phylogenetic tree of ppGalNAcTs. The phylogenetic tree is based on the neighbor-joining method based on amino acid sequences using ClustalW software. The branch length represents evolutionary distance between the members of the family ppGalNAcT enzymes. The following Genbank accession numbers were used: *Gallus gallus* T16 (XP_001231965.1), *Xenopus tropicalis* T16 (NP_001039091.1), *Homo sapiens* T16 (AJ505951), *Danio rerio* T16 (XP_001339749.3), *Danio rerio* T14 (NP_001038460.1), *Xenopus tropicalis* T14 (NP_001072369), *Gallus gallus* T14 (XM_419370.2), *Homo sapiens* T14 (Y09324), *Drosophila melanogaster* T2 (NP_608773.2), *Danio rerio* T2 (NP_001121823.1), *Xenopus tropicalis* T2 (XP_002931524.1), *Gallus gallus* T2 (XP_419581.2), *Homo sapiens* T2 (X85019), *Biomphalaria glabrata* (KC182513), *Crassostrea gigas* T2 (EKC38600), *Caenorhabditis elegans* GLY4 (NP_507850.2), *Toxoplasma gondii* T1 (XP_002365147.1), *Toxoplasma gondii* T3 (XP_002369811.1), *Toxoplasma gondii* T2 (XP_002365091.1)

been reported to be responsible for UDP-binding [28]. The second part of the catalytic domain contained the D^{353}xxxxxWGGENLE^{365} sequence within the Gal/GalNAc-transferase motif, which is shared among ppGalNAcTs and ß1,4-galactosyltransferases [26]. It has been shown that this domain interacts with the GalNAc moiety of the sugar donor [27]. Similar to most of the other ppGalNAcTs the C-terminus of the snail enzyme comprised a lectin domain, which consisted of three homologous repeats (α, ß and γ) belonging to the ricin-type lectin structural family. It has been hypothesized that this domain is involved in recognizing carbohydrate moieties on glycopeptide substrates [26, 29].

An extended phylogenetic tree with 102 ppGalNAcTs from vertebrate and invertebrate sources has been presented recently [3]. There, the new ppGalNAcT from Bge cells fits according to its amino acid homology into the subfamily I, group b (Fig. 6). As expected, the snail enzyme is closely related to the ppGalNAcT from *C. gigas*, another representative from the mollusk kingdom. Furthermore, the snail ppGalNAcT is highly homologous to the T2-enzymes from many other species (human, chicken, frog, fish, fly).

For activity and specificity tests the truncated version of the enzyme without the transmembrane domain was expressed in insect cells. To allow easy purification from the supernatant a honeybee melittin secretion signal and a His tag were fused to the N-terminus. The Bge ppGalNAcT was able to recognize all four chosen peptides as acceptor substrates, Muc2, Muc5Ac, Muc1a and Muc1a'. For Muc2 and Muc5Ac we observed a multiple GalNAc transfer within a reaction time of two hours of the standard assay, whereas Muc1 and Muc1a' were possible, but rather weak acceptors. Increasing the incubation time up to 24 hours in combination with several additions of enzyme yielded in an incorporation of 8 GalNAc-residues into Muc2. These data demonstrate that the enzyme has a broad specificity and is capable of utilizing non-glycosylated as well as mono- or multi-glycosylated substrates. Comparing the specificity of the snail enzyme with human enzymes, especially with human ppGalNAc T2 and ppGalNAc T14 [20], which are most close to our snail enzyme from the structural point of view (Fig. 6), the results indicate that Bge ppGalNAcT is a member of a ppGalNAc T2 sub-group. This is also supported by comparing the specificity with the insect ppGalNAc T2 from *Drosophila* [30]

Based on the amino acid sequence of the enzyme no potential N-glycosylation site was predicted. The recombinant enzyme expressed in Sf9-cells was O-glycosylated (detected by peanut agglutinin staining and O-glycosidase sensitivity; data not shown). We do not know the native glycosylation status of this protein in the snail, but in principle O-glycosylation in snails is possible [2].

In terms of biochemical parameters the Bge ppGalNAcT correlated with those determined for other members of the family. Its stability performance was comparable as well as the essential requirement of divalent cations with Mn^{2+} being optimal in a range of 10–20 mM. However, while most of the other enzymes have a pH-optimum at 7.0–7.5, the Bge ppGalNAcT shows its highest activity at pH 6.0–6.5.

In this study we could identify, clone, express and analyze a ppGalNAcT from gastropod origin. It is the first expressed and fully characterized enzyme of this family from the mollusk kingdom. According to its structure and its substrate specificity it is a close relative to ppGalNAcT2s from various sources.

References

1. Wilson, I.B.H., Paschinger, K., Rendić, D.: Glycosylation of model and "Lower" organisms. In: Gabius, H.-J. (ed.) The Sugar Code, pp. 139–154. Wiley-VCH, Weinheim (2009)
2. Stepan, H., Pabst, M., Altmann, F., Geyer, H., Geyer, R., Staudacher, E.: O-Glycosylation of snails. Glycoconjugate J **29**, 189–198 (2012)
3. Bennett, E.P., Mandel, U., Clausen, H., Gerken, T.A., Fritz, T.A., Tabak, L.A.: Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Glycobiology **22**, 736–756 (2012)
4. Brockhausen, I., Schachter, H., Stanley, P.: O-GalNAc glycans. In: Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W., Etzler, M.E. (eds.) Essentials of Glycobiology, pp. 115–127. Cold Spring Harbor, Laboratory Press, New York (2008)
5. Raman, J., Guan, Y., Perrine, C.L., Gerken, T.A., Tabak, L.A.: UDP-N-acetyl-α-D-galactosamine:polypeptide N-acetyl-galactosaminyltransferases: completion of the family tree. Glycobiology **22**, 768–777 (2012)
6. Ten Hagen, K.G., Tran, D.T.: A UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase isessential for viability in *Drosophila melanogaster*. J. Biol. Chem. **277**, 22616–22622 (2002)
7. Ten Hagen, K.G., Hagen, F.K., Balys, M.M., Beres, T.M., Van Wuyckhuyse, B., Tabak, L.A.: Cloning and expression of a novel, tissue specifically expressed member of the UDP-GalNAc:poly-peptide N-acetylgalactosaminyltransferase family. J. Biol. Chem. **273**, 27749–27754 (1998)
8. Hagen, F.K., Nehrke, K.: cDNA cloning and expression of a family of UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalac-tosaminyltransferase sequence homologs from *Caenorhabditis elegans*. J. Biol. Chem. **273**, 8268–8277 (1998)
9. Freire, T., Casaravilla, C., Carmona, C., Osinaga, E.: Mucin-type O-glycosylation in *Fasciola hepatica*: characterisation of carcinoma-associated Tn and sialyl-Tn antigens and evaluation of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase activity. Int. J. Parasitol. **33**, 47–56 (2003)
10. Freire, T., Robello, C., Soulé, S., Ferreira, R., Osinaga, E.: Sialyl-Tn antigen expression and O-linked GalNAc-Thr synthesis by *Trypanosoma cruzi*. Biochem. Biophys. Res. Commun. **312**, 1309–1316 (2003)

11. Stwora-Wojczyk, M.M., Kissinger, J.C., Spitalnik, S.L., Wojczyk, B.S.: O-glycosylation in *Toxoplasma gondii*: identification and analysis of a family of UDP-GalNAc:polypeptide N-acetylgalactosaminyl-transferases. Int. J. Parasitol. **34**, 309–322 (2004)

12. Stepan, H., Bleckmann, C., Geyer, H., Geyer, R., Staudacher, E.: Determination of 3-O- and 4-O-methylated monosaccharide constituents in snail glycans. Carbohydr. Res. **345**, 1504–1507 (2010)

13. Staudacher, E.: Methylation - an uncommon modification of glycans. Biol. Chem. **393**, 675–685 (2012)

14. Sambrook, J., Fritsch, E.F., Maniatis, T.: Molecular Cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)

15. Knight, M., Ittiprasert, W., Odoemelam, E.C., Adema, C.M., Miller, A., Raghavan, N., Bridger, J.M.: Non-random organization of the *Biomphalaria glabrata* genome in interphase Bge cells and the spatial repositioning of activated genes in cells co-cultured with *Schistosoma mansoni*. Int. J. Parasitol. **41**, 61–70 (2011)

16. Summers, M.D., Smith, G.E.: A manual of methods for baculovirus vectors and insect cell culture procedures. Tex. Agric. Exp. Station. Bull. 1555 (1987)

17. Tessier, D.C., Thomas, D.Y., Khouri, H.E., Laliberté, F., Vernet, T.: Enhanced secretion from insect cells of a foreign protein fused to the honeybee melittin signal peptide. Gene **98**, 177–183 (1991)

18. Gutternigg, M., Bürgmayr, S., Pöltl, G., Rudolf, J., Staudacher, E.: Neutral N-glycan patterns of the gastropods *Limax maximus*, *Cepaea hortensis*, *Planorbarius corneus*, *Arianta arbustorum* and *Achatina fulica*. Glycoconjugate J. **24**, 475–489 (2007)

19. Wandall, H.H., Mirgorodskaya, E., Kristensen, A.K., Roepstorff, P., Bennett, E.P., Nielsen, P.A., Hollingsworth, M.A., Burchell, J., Taylor-Papadimitriou, J., Clausen, H.: Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine: Polypeptide N-acetylgalactosaminyltransferase family. GalNAc-T1, -T2, and -T3. J. Biol. Chem. **272**, 23503–23514 (1997)

20. Wang, H., Tachibana, K., Zhang, Y., Iwasaki, H., Kameyama, A., Cheng, L., Guo, J., Hiruma, T., Togayachi, A., Kudo, T., Kikuchi, N., Narimatsu, H.: Cloning and characterization of a novel UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase pp-GalNAc-T14. Biochem. Biophys. Res. Commun. **300**, 738–744 (2003)

21. Ten Hagen, K.G., Tetaert, D., Hagen, F.K., Richet, C., Beres, T.M., Gagnon, J., Balys, M.M., VanWuyckhuyse, B., Bedi, G.S., Degand, P., Tabak, L.A.: Characterization of a UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase that displays glycopeptide N-acetylgalactosaminyltransferase activity. J. Biol. Chem. **274**, 27867–27874 (1999)

22. Peyer, C., Bonay, P., Staudacher, E.: Purification and characterization of a β-xylosidase from potatoes (*Solanum tuberosum*). Biochim. Biophys. Acta **1672**, 27–35 (2004)

23. Sørensen, T., White, T., Wandall, H.H., Kristensen, A.K., Roepstorff, P., Clausen, H.: UDP-N-acetyl-α-D-galactosamine: polypeptide N-Acetylgalactosaminyltransferase. Identification and separation of two distinct transferase activities. J. Biol. Chem. **270**, 24186–24173 (1995)

24. Mulder, H., Spronk, B.A., Schachter, H., Neeleman, A.P., van den Eijnden, D.H., De Jong-Brink, M., Kamerling, J.P., Vliegenthart, J.F.: Identification of a novel UDP-GalNAc:GlcNAc β -R β1-4 N-acetylgalactosaminyltransferase from the albumen gland and connective tissue of the snail *Lymnaea stagnalis*. Eur. J. Biochem. **227**, 175–185 (1995)

25. Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H., Xiong, Z., Que, H., Xie, Y., Holland, P.W., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang, J., Peng, C., Meng, J., Yang, L., Liu, J., Wen, B., Zhang, N., Huang, Z., Zhu, Q., Feng, Y., Mount, A., Hedgecock, D., Xu, Z., Liu, Y., Domazet-Lošo, T., Du, Y., Sun, X., Zhang, S., Liu, B., Cheng, P., Jiang, X., Li, J., Fan, D., Wang, W., Fu, W., Wang, T., Wang, B., Zhang, J., Peng, Z., Li, Y., Li, N., Wang, J., Chen, M., He, Y., Tan, F., Song, X., Zheng, Q., Huang, R., Yang, H., Du, X., Chen, L., Yang, M., Gaffney, P.M., Wang, S., Luo, L., She, Z., Ming, Y., Huang, W., Zhang, S., Huang, B., Zhang, Y., Qu, T., Ni, P., Miao, G., Wang, J., Wang, Q., Steinberg, C.E., Wang, H., Li, N., Qian, L., Zhang, G., Li, Y., Yang, H., Liu, X., Wang, J., Yin, Y., Wang, J.: The oyster genome reveals stress adaptation and complexity of shell formation. Nature **490**, 49–54 (2012)

26. Hagen, F.K., Hazes, B., Raffo, R., de Sa, D., Tabak, L.A.: Structure-function analysis of the UDP-N acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase. Essential residues lie in a predicted active site cleft resembling a lactose repressor fold. J. Biol. Chem. **274**, 6797–6803 (1999)

27. Fritz, T.A., Hurley, J.H., Trinh, L.B., Shiloach, J., Tabak, L.A.: The beginnings of mucin biosynthesis: the crystal structure of UDP-GalNAc:polypeptide alpha-N-acetylgalactosaminyltransferase-T1. Proc. Natl. Acad. Sci. U. S. A. **101**, 15307–15312 (2004)

28. Tenno, M., Toba, S., Kézdy, F.J., Elhammer, A.P., Kurosaka, A.: Identification of two cysteine residues involved in the binding of UDP-GalNAc to UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase 1 (GalNAc-T1). Eur. J. Biochem. **269**, 4308–4316 (2002)

29. Hassan, H., Reis, C.A., Bennett, E.P., Mirgorodskaya, E., Roepstorff, P., Hollingsworth, M.A., Burchell, J., Taylor-Papadimitriou, J., Clausen, H.: The lectin domain of UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-T4 directs its glycopeptide specificities. J. Biol. Chem. **275**, 38197–38205 (2000)

30. Ten Hagen, K.G., Tran, D.T., Gerken, T.A., Stein, D.S., Zhang, Z.: Functional characterization and expression analysis of members of the UDP-GalNAc:polypeptide N-Acetylgalactosaminyltransferase family from *Drosophila melanogaster*. J. Biol. Chem. **278**, 35039–35048 (2003)