Deletion of Two Exons from the *Lymnaea stagnalis* β1→4-N-Acetylglucosaminyltransferase Gene Elevates the Kinetic Efficiency of the Encoded Enzyme for Both UDP-sugar Donor and Acceptor Substrates*

(Received for publication, November 13, 1996, and in revised form, March 27, 1997)

Hans Bakker‡, Angelique Van Tetering‡, Marja Agterberg‡, August B. Smit§, Dirk H. Van den Eijnden‡, and Irma Van Die¶

From the Departments of ‡Medical Chemistry and §Experimental Zoology, Vrije Universiteit, 1081 BT Amsterdam, The Netherlands

**Lymnaea stagnalis** UDP-GlcNAc:GlcNAcβ-R β1→4-N-acetylglucosaminyltransferase (β4-GlcNAcT) is an enzyme with structural similarity to mammalian UDP-Gal:GlcNAcβ-R β1→4-galactosyltransferase (β4-GaIT). Here, we report that also the exon organization of the genes encoding these enzymes is very similar. The β4-GlcNAcT gene (12.5 kilobase pairs, spanning 10 exons) contains four exons, encompassing sequences that are absent in the β4-GaIT gene. Two of these exons (exons 7 and 8) show a high sequence similarity to part of the preceding exon (exon 6), suggesting that they have originated by exon duplication. The exon in the β4-GaIT gene, corresponding to β4-GlcNAcT exon 6, encodes a region that has been proposed to be involved in the binding of UDP-Gal. The question therefore arose, whether the repeating sequences encoded by exon 7 and 8 of the β4-GlcNAcT gene would determine the specificity of the enzyme for UDP-GlcNAc, or for the less preferred UDP-GalNAC. It was found that deletion of only the sequence encoded by exon 8 resulted in a completely inactive enzyme. By contrast, deletion of the amino acid residues encoded by exons 7 and 8 resulted in an enzyme with an elevated kinetic efficiency for both UDP-sugar donors, as well as for its acceptor substrates. These results suggest that at least part of the donor and acceptor binding domains of the β4-GlcNAcT are structurally linked and that the region encompassing the insertion contributes to acceptor recognition as well as to UDP-sugar binding and specificity.

Glycosyltransferases form a large family of functionally related, membrane-bound enzymes that are involved in the biosynthesis of the carbohydrate moieties of glycoproteins and glycolipids (1, 2). Recently we have identified a novel glycosyltransferase by the isolation of a UDP-GlcNAc:GlcNAcβ-R β1→4-N-acetylglucosaminyltransferase (β4-GlcNAcT) cDNA from the prostate gland of the snail *Lymnaea stagnalis* (3). In vitro, the recombinant β4-GlcNAcT catalyzes the transfer of GlcNAc from UDP-GlcNAc in β1→4 linkage to various β-N-acetylgalactosaminides (3, 4). The β4-GlcNAcT cDNA appeared to show a significant sequence similarity to the mammalian UDP-Gal:GlcNAcβ-R β1→4-galactosyltransferase (β4-GaIT) cDNAs, with an overall resemblance between the predicted amino acid sequence of about 30% (3, 5–7). Based on the genetic and enzymatic relationship of the β4-GlcNAcT and the β4-GaITs, we have suggested that these enzymes constitute a separate glycosyltransferase gene family, the members of which are capable of catalyzing the transfer of a specific sugar from their respective UDP-sugar donors in a β1→4 linkage toward a terminal β-linked GlcNAc residue in the acceptor (3, 8). Based on enzymatic properties, we have proposed that also UDP-GalNAc:GlcNAcβ-R β1→4-N-acetylglucosaminyltransferase (β4-GalNAcT), detected in several non-vertebrate species (9–13), belongs to this family (14). The primary structure of this enzyme, however, is still unknown.

The reaction catalyzed by glycosyltransferases typically involves two substrates and often a divalent cation cofactor. This suggests that the enzymes consist of several functional domains involved in substrate and cofactor binding, respectively. Comparison of the conserved and variable regions of genetically related glycosyltransferases with different properties would open possibilities to address structure-function relationships. As snails and mammals are evolutionary distant species, comparison of the genomic organization of the genes that encode these enzymes might give insight in their way of divergence, that resulted in genes encoding enzymes with a different UDP-sugar specificity. The genomic organization of the murine and human β4-GaIT genes have been described previously (15, 16). Here we report the organization of the *L. stagnalis* gene that codes for the β4-GlcNAcT. The intron-exon distribution was determined and compared with that of the β4-GaIT gene. Mutant β4-GlcNAcT cDNAs were constructed by deletion of sequences that do not have a counterpart in the β4-GaIT gene, and expressed in COS cells. Comparison of the kinetic parameters of the resulting mutant and parental enzymes showed that the insertion in the β4-GlcNAcT and its surrounding regions contribute to acceptor recognition as well as to UDP-sugar binding and specificity.

**EXPERIMENTAL PROCEDURES**

Materials, Bacterial Strains, and Cells—Acceptor substrates were obtained from Sigma (compounds 1, 2, and 3) and from Toronto Research Chemicals (compounds 5 and 6). Compound 4 was a gift of Dr. O. W. floribunda lectin.

*The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom reprint requests should be addressed: Dept. of Medical Chemistry, Vrije Universiteit, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands. Tel.: 31-20-4448157; Fax: 31-20-4448144; E-mail: im.van_die.medchem@med.vu.nl.

‡ The abbreviations used are: β4-GlcNAcT, UDP-GlcNAc:GlcNAcβ-R β1→4-N-acetylglucosaminyltransferase; β4-GalNAcT, UDP-GalNAc:GlcNAcβ-R β1→4-N-acetylgalactosaminyltransferase; β4-GaIT, UDP-Gal:GlcNAcβ-R β1→4-galactosyltransferase; kb, kilobase pair(s); PCR, polymerase chain reaction; HPAEC, high pH anion-exchange chromatography; PAD, pulsed amperometric detection; pNP, para-nitrophenyl; WFA, W. floribunda lectin.
Deletion of Exons of the Snail β1→4-GlcNAc Transferase Gene

Hindsaul, University of Alberta, Alberta, Canada. UDP-[3H]GlcNAc, UDP-[3H]GalNAc and UDP-[3H]Gal were purchased from NEN Life Science Products, and were diluted with unlabeled UDP-sugars (Sigma) to the desired specific radioactivity.

Recombinant plasmids were propagated in the Escherichia coli K12 strain XLI-Blue (Stratagene). COS-7M6 cells (ATCC 1651) were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, 2 mM L-glutamine, and 1:50 penicillin/streptomycin solution (all from Life Technologies, Inc.). Synthetic oligonucleotides were obtained from Isogen Bioscience BV (Maarsen, the Netherlands).

DNA Techniques—The genomic clone 45 has been isolated previously (3) from a genomic EMBL3A library of L. stagnalis (17). Isolation of plasmid DNA was carried out by a modification of the minilysate method, as described in Ref. 18. Plasmids used for transfection of COS cells were isolated by the Qiagen plasmid protocol, using a QIAGEN-tip 100 microlum. Restriction enzymes and other DNA-modifying enzymes were used according to the manufacturer. Dideoxyxunucleotide chain-terminating sequencing reactions (19) were performed on double-stranded plasmid DNA, with the T7 DNA sequencing kit (Pharmacia Biotech Inc.). [35S]dATP (Amersham), using M13 universal primer, the KS and SK primers, and several sequence-specific synthetic oligonucleotide primers. Southern blotting was performed as described previously (3). PCR with sequence-specific primers was performed using Ultima-polynucleotides (Perkin-Elmer) by 25 cycles (1 min at 95 °C, 1 min at 65 °C, 1 min at 72 °C). For cloning purposes, amplified fragments were subsequently purified according to the QiAquick PCR purification protocol (Qiagen Inc.).

Construction of pPROTA Hybrid Plasmids—The plasmid pCM135, containing a fusion between part of the protein sequence and β1→4-GlcNAcT cDNA, was constructed as follows. A BamHI-EcoRI adapter was ligated into the BamHI site of pVTBac-P11.4, carrying 5′-truncated β4-GlcNAcT cDNA (3). The resulting 1.4-kb EcoRI fragment from this construct was ligated into an EcoRI-digested pPROTA vector (20). Plasmids carrying mutant β4-GlcNAcT genes were derived from pCM135 by exchange of the 0.52-kb Xhol-BglII fragment for a PCR fragment carrying the desired deletion. For PCR of the deletion fragments, a sense primer (bases 439–456 of the β4-GlcNAcT cDNA; Ref. 3) and the antisense primer ID6 (ATACATGTTAAGATGCTGCTGAGTGACAT-TCGA) or ID16 (ATACATGTTAAGATGCTGCTGAGTGACATTCG) was used. The antisense primer ID8, used for construction of pCM124, consists of a part complementary to the 3′ end of exon 6, and a part (11 base pairs), complementary to the 5′ end of exon 9 (BglII restriction site underlined). The antisense primer ID16, used for construction of pCM166, consists of a part complementary to the 3′ end of exon 7, and the same 5′ part of exon 9 as ID8. PCR fragments obtained with these primers were digested with XhoI and BglII, and ligated into XhoI-BglII-digested pCM135. After transformation and plasmid isolation of several transformants, the desired plasmids (pMC142 encoding protA-β4-GlcNAcT57–8, GlcNAc-S-pNP, and UDP-[3H]GalNAc, was analyzed by lectin affinity chromatography with Wisteria floribunda lectin (WFA), immobilized on azlactone/bisacrylamide polymeric beads (33) (from Emphas™, Pierce) (24). The product was also analyzed by HPAEC-PAD (3). As a control, both WFA lectin chromatography and HPAEC-PAD analysis were performed with reference GalNAcβ1→4-GlcNAc-S-pNP produced with L. stagnalis albumen gland b4-GlcNAcT (9).

RESULTS

Isolation and Characterization of the L. stagnalis β4-GlcNAcT Gene—A genomic clone, denoted 45, was isolated previously from an EMBL3A library of L. stagnalis, and was shown to contain two short DNA sequences identical to β4-GlcNAcT cDNA sequences (3). A rough genomic map of 45 was constructed by PCR and Southern blot hybridization using specific β4-GlcNAcT cDNA fragments as probes (Fig. 1). 45 was found to encompass the complete coding sequence of the β4-GlcNAcT gene, spanning 12.5 kb of DNA, that was divided into 10 exons (Fig. 1, Table I). As probably part of the 5′-noncoding sequence is lacking from the cDNA (3), we cannot exclude the presence of the gene of one or more exons upstream of the denoted exon 1. Exon 10 was found to encompass the complete 3′-noncoding region. All exon sequences in the genomic clone were identical to those of the cDNA (3). Donor and acceptor splice junction sequences (Table II) are in agreement with consensus sequences reported (25).

Comparison of the L. stagnalis β4-GlcNAcT Gene with the Murine and Human β4-GaIT Genes—The cDNA encoding β4-GlcNAcT shows sequence identity with the mammalian β4-GaIT cDNAs identified (3, 5–7). A comparison of the protein-coding exons of the L. stagnalis β4-GlcNAcT gene with those of the murine and human β4-GaIT genes (15, 16) is shown in Fig. 2. The β4-GaIT gene was found to be divided into six exons, whereas the β4-GlcNAcT gene appeared to contain 10 exons. Exons 3, 4, 5, 6, and 9 of the β4-GlcNAcT gene were found to show similarity to exons 2–6 of the β4-GaIT gene, corresponding to the catalytic domain of the enzyme (Fig. 2). This similarity was not only confined to a high degree of sequence similarity.
Deletion of Exons of the Snail β1→4-GlcNAc Transferase Gene

### Table I

| Exon no. | Exon length | Corresponding cDNA residues | Amino acids encoded |
|----------|-------------|-----------------------------|--------------------|
| 1        | >52         | ~3–52                       | 1–17               |
| 2        | 399         | 53–451                      | 18–150             |
| 3        | 233         | 452–684                     | 151–328            |
| 4        | 191         | 685–875                     | 229–392            |
| 5        | 123         | 876–998                     | 293–333            |
| 6        | 102         | 999–1100                    | 334–367            |
| 7        | 78          | 1101–1178                   | 368–393            |
| 8        | 78          | 1179–1256                   | 394–419            |
| 9        | 139         | 1257–1395                   | 420–465            |
| 10       | 225         | 1396–1620                   | 466–490            |

### Table II

| Intron no. | Intron length | Splice junction sequence | Donor | Acceptor |
|------------|---------------|--------------------------|-------|----------|
| h1         | 51            | TCTCGAGAATTCCTGCAAATTCGCAAGCTTTT | GACCAAGCTTTCTTCACTTCTTCTTCTT | |
| h2         | 52            | GACCAAGCTTTCTTCACTTCTTCTTCTTCTT | GACCAAGCTTTCTTCACTTCTTCTTCTT | |
| h3         | 53            | GACCAAGCTTTCTTCACTTCTTCTTCTTCTT | GACCAAGCTTTCTTCACTTCTTCTTCTT | |
| h4         | 54            | GACCAAGCTTTCTTCACTTCTTCTTCTTCTT | GACCAAGCTTTCTTCACTTCTTCTTCTT | |
| h5         | 55            | GACCAAGCTTTCTTCACTTCTTCTTCTTCTT | GACCAAGCTTTCTTCACTTCTTCTTCTT | |
| h6         | 56            | GACCAAGCTTTCTTCACTTCTTCTTCTTCTT | GACCAAGCTTTCTTCACTTCTTCTTCTT | |
| h7         | 57            | GACCAAGCTTTCTTCACTTCTTCTTCTTCTT | GACCAAGCTTTCTTCACTTCTTCTTCTT | |
| h8         | 58            | GACCAAGCTTTCTTCACTTCTTCTTCTTCTT | GACCAAGCTTTCTTCACTTCTTCTTCTT | |
| h9         | 59            | GACCAAGCTTTCTTCACTTCTTCTTCTTCTT | GACCAAGCTTTCTTCACTTCTTCTTCTT | |

Fig. 2. Schematic representation of organization of the protein-coding exons of the genes encoding L. stagnalis β4-GlcNAcT and human β4-GalT (16). The shaded boxes indicate exons that show considerably amino acid identity between both genes. The open boxes indicate exons that show no significant sequence similarity between the genes, or are present only in the β4-GlcNAcT gene. Homologous splice sites are connected by dotted lines.

### Table III

| Exon | Amino acid sequence encoded |
|------|-----------------------------|
| 6    | AVHMKPLRIKLRTLHGLDMVH---VEAGWNVNPPS |
| 7    | KGAHLYDMKNAGVCAAGWNVPNS |
| 8    | KWPPLRF5D5SNVAFAEAGWNVPDR |

### Figure 3

Physical maps of the protein A-β4-GlcNAcT chimeric plasmids. The open box represents pBPTA vector, and the thin line inserted GlcNAcT cDNA. The positions of exons 6, 7, and 8 are indicated in the cDNA. E, X, S, and B indicate the position of the restriction sites for EcoRI, XhoI, SalI, and BglII, respectively, in the cDNA. pMC135 encodes protA-β4-GlcNAcT, pMC142 encodes protA-β4-GlcNAcTΔ7–8, and pMC166 encodes protA-β4-GlcNAcTΔ8.

### Figure 4

Enzyme Activity of the Chimeric Proteins—The enzyme activities of the mutant hybrid proteins and the native protA-β4-GlcNAcT (Fig. 4) were assayed using similar amounts of enzyme (Fig. 4). As the native membrane-bound β4-GlcNAcT shows a low GalNAcT activity (about 7% of the β4-GlcNAcT activity), the activity of the enzymes was measured with both UDP-GlcNAc and UDP-GalNAc. ProtA-β4-GlcNAcTΔ7–8 showed reproducibly an almost 2 times higher GlcNAcT activity and a 4 times higher GalNAcT activity than the parental protA-β4-GlcNAcT (Table IV). In contrast to the other chimeric proteins, protA-β4-GlcNAcTΔ8 appeared to be enzymatically inactive. To determine if the enzymes would show GalT activity, the hybrid proteins were purified using IgG-agarose beads. In this way, the recombinant enzymes were completely disposed of COS cell-derived β4-GalT. The bead-associated chimeras did not show detectable GalT activity, whereas they showed GlcNAcT and GalNAcT activities similar to those for the concentrated media (results not shown).

Acceptor specificity studies, using UDP-GlcNAc as sugar donor, showed no significant differences in acceptor preference between the two active hybrid enzymes at an acceptor concentration of 1 mM (Table V; Ref. 4). The acceptor specificity of the mutant chimeric protein using UDP-GalNAc as a sugar donor appeared very similar to the preference of the enzyme when using UDP-GlcNAc. By contrast, the prostate gland β4-GalNAcT is less specific for the linkage type of the terminal β-linked GlcNAc, and its acceptor substrate requirement resembles that of the albumen gland β4-GalNAcT (9).

### Kinetic Properties of the Chimeric β4-GlcNAcTs—To explain the differences in donor specificity in more detail, the K<sub>m</sub> and V values for UDP-GlcNAc and UDP-GalNAc were determined for both protA-β4-GlcNAcT and protA-β4-GlcNAcTΔ7–8. It appears that the increase in GlcNAcT activity of protA-β4-GlcNAcTΔ7–8 is due to a 3-fold reduction in K<sub>m</sub> for UDP-GlcNAc (Table VI), whereas the increased GalNAcT activity of the mutant enzyme is mainly due to an enhanced maximum velocity. These effects result in an elevated kinetic efficiency for both types of transfer. The data suggest an involvement of the region encoded by exons 7 and 8 in UDP-sugar donor binding. However, it was found that both active chimeras were...
inhibited to almost the same extent by UDP (50% at 4–5 mM UDP, results not shown).

The $K_m$ and $V$ values of two acceptor substrates were estimated (Table VII). While the $V$ values of both enzymes appeared to be of the same order for both compounds, the $K_m$ values were decreased in the mutant enzyme. Deletion of the two exons clearly elevates the kinetic efficiency with both acceptor substrates.

![image](https://example.com/fig4.png)

**FIG. 4.** Western blot analysis of protA-fusion proteins. COS-7M6 cells were transfected with the pPROTA-fusion plasmids, carrying the β4-GlcNAcT cDNA, or deletion derivatives. Total cellular proteins, secreted into the medium, were separated by SDS-polyacrylamide gel electrophoresis (10%). The proteins were transferred to nitrocellulose. Antiserum reactions were directed against the protein A part of the proteins. Lane 1, protA-β4-GlcNAcTΔ7–8; lane 2, protA-β4-GlcNAcTΔ5; lane 3, protA-β4-GlcNacT; lane 4, pPROTA; lane 5, molecular mass standards. Molecular masses on the right are indicated in kDa.

**TABLE IV**

UDP-sugar donor promiscuity of protein A-β4-GlcNAcT chimeras and deletion mutants derived thereof

| Enzyme            | Glycosyltransferase activity with UDP-GlcNAc | UDP-GalNAc |
|-------------------|---------------------------------------------|------------|
| protA-β4-GlcNAcT  | 210 pmol · min$^{-1}$ · ml$^{-1}$           | 15         |
| protA-β4-GlcNAcTΔ8| <1 pmol · min$^{-1}$ · ml$^{-1}$            | <1         |
| protA-β4-GlcNAcTΔ7,8| 361 pmol · min$^{-1}$ · ml$^{-1}$         | 67         |

**Characterization of the N-Acetylgalactosaminylated Product Obtained with β4-GlcNAcTΔ7–8**

The N-acetylgalactosaminylated product formed by incubating protA-β4-GlcNAcTΔ7–8, UDP-GalNAc, and GlcNAc-S-pNP was anticipated to be GalNAcβ1→4GlcNAc-S-pNP. To confirm this, the product was compared with authentic GalNAcβ1→4GlcNAc-S-pNP obtained by the action of *L. stagnalis* albumen gland β4-GlcNAcT (9). Both product and reference were subjected to lectin affinity chromatography with immobilized WFA, known to bind with high affinity to glycans containing terminal GalNAc in a β1→4 linkage (24, 28, 29). More than 90% of both compounds bound to the immobilized WFA, and could be eluted with 10 mM GalNAc. In addition, reference and product eluted on HPAEC with the same retention time (data not shown). These results indicate that β4-GlcNAcTΔ7–8 shows a β4-GlcNAcT activity.

**DISCUSSION**

Based on the sequence identity of the mammalian β4-GalTs and *L. stagnalis* β4-GlcNAcT, these enzymes have been proposed to be members of one gene family (3). The observed conservation of the exon-intron organization of the genes encoding these enzymes strengthens this view. All exons of the β4-GalT gene show sequence identity to parts of the β4-GlcNAcT gene, which suggests that both genes originated from an common ancestor, and diverged to genes encoding enzymes with a different sugar donor specificity.

The most remarkable difference between the genes is found in an insertion of two exons (exons 7 and 8) in the β4-GlcNAcT gene. These exons were found to encode partial repeats of exon 6. The sequences of these exons and those of the bordering introns strongly suggest that exons 7 and 8 were generated by internal exon duplications, and originate from exon 6. This is the first glycosyltransferase described in which exon duplication seems to have taken place during evolution. Generally, exon duplication is thought to play an important role in the evolution of genes, and many complex genes have been described that have evolved by internal exon duplication and subsequent modification of the primordial genes (30, 31).

**TABLE V**

Acceptor specificity of recombinant protA-β4-GlcNAcT and protA-β4-GlcNAcTΔ7,8

For the GlcNAcT assays 100% activity is 128 and 338 pmol · min$^{-1}$ · ml medium$^{-1}$, respectively. For the GalNAcT assays 100% activity is 110 pmol · min$^{-1}$ · ml medium$^{-1}$ for the mutant enzyme, and 5 pmol · min$^{-1}$ · ml medium$^{-1}$ for the prostate-derived enzyme.

| Acceptor substrate | Activity with UDP-GlcNAc | Activity with UDP-GalNAc |
|--------------------|--------------------------|--------------------------|
|                    | protA-β4-GlcNAcT         | protA-β4-GlcNAcTΔ7,8     | rec-β4-GlcNAcTα         | protA-β4-GlcNAcTΔ7,8 | Prostate gland membranes |
| 1 GlcNAcβ-S-pNP    | 100%                     | 100%                     | 100%                     | 100%                     | 100%                     |
| 2 GlcNAcβ-O-pNP    | 55%                      | 72%                      | 71%                      | 68%                      | 28%                      |
| 3 GlcNAcβ1→4GlcNAcβ-O-pNP | 5% | 7% | 1% | 1% | 11% |
| 4 GlcNAcβ1→2Manα1,3 Manβ-O-Rα | <1% | <1% | <1% | 48% |
| 5 GlcNAcβ1→3GalNAcO-O-pNP | 3% | 4% | 5% | 2% | 47% |
| 6 GlcNAcβ1→6GalNAcO-O-pNP | 96% | 86% | 104% | 136% | 67% |

a Recombinant soluble β4-GlcNAcT from baculovirus infected insect cells (4).

b $R = -(CH_2)_nCOOCH_3$.

**TABLE VI**

Kinetic parameters of protein A-β4-GlcNAcT chimeras and a deletion mutant derived thereof for two UDP-sugar donors

The recombinant enzymes were produced as protein A fusion proteins in COS cells.

| Enzyme             | UDP-GlcNAc | UDP-GalNAc |
|--------------------|------------|------------|
|                    | $K_m$      | $V$        | $K_m$      | $V$        |
|                    | mM         | pmol · min$^{-1}$ · ml$^{-1}$ | $V · K_m^{-1}$ | mM         | pmol · min$^{-1}$ · ml$^{-1}$ | $V · K_m^{-1}$ |
| protA-β4-GlcNAcT   | 0.23 ± 0.02 | 300 ± 11 | 1304 | 2.0 ± 0.6 | 92 ± 13 | 46 |
| protA-β4-GlcNAcTΔ7,8 | 0.08 ± 0.01 | 380 ± 12 | 4750 | 1.6 ± 0.2 | 358 ± 19 | 224 |
To study the effect of the exon duplications on enzyme catalysis, we deleted the sequence (52 amino acids) corresponding to the two additional exons from the cDNA sequence. The catalytic domain of the mutant protein thus obtained, might resemble that of a putative ancestor of the *L. stagnalis* β-4-GlcNAcT. The deletion enhanced the kinetic efficiency of the enzyme for both the transfer of GlcNAc and GalNAc, as well as for its acceptor substrates. The acceptor substrate specificity, i.e. the preference for a terminal β6-linked GlcNAc, was not affected. So the mutant enzyme has an enhanced catalytic potential, but also an increased sugar-donor promiscuity. Surprisingly, deletion of only the sequence encoded by exon 8 resulted in a completely inactive enzyme. An explanation could be that the enzyme is not properly folded with only one additional repeat, whereas the presence of two additional copies of the repeated sequence allows a correct folding.

From the results obtained here, it is difficult to deduce the selective advantage that was obtained by the exon duplications for the biological function of the β4-GlcNAcT. As can be inferred from the sequence, changes have occurred in the repeats after the exon duplications, and most likely also in other regions of the protein. It is possible, however, that reduction of the capacity of the primordial β-4-GlcNAcT to transfer GalNAc might have been advantageous for the snail. The increase in the specificity of the enzyme for UDP-GlcNAc would then have been of more importance for the snail than the loss of catalytic potency that coincided with the introduction of the additional exons.

In β4-GalT the corresponding region (encoded by exon 5) has been proposed to be involved in UDP-Gal binding (32–34). In the same region a tetrapeptide (KKKN) has been found that is conserved between β4-GalT and α3-GalT, which is in support of this proposition (35). In the human β4-GaIT a second region (in exon 4) has been proposed to be involved in UDP-Gal binding (36, 37). Interestingly, this latter region shows a high degree of similarity between the β4-GalTs and the β4-GlcNAcT, whereas the more downstream region of β4-GalT (encoded by exon 5) shows much less sequence identity with the corresponding region in the β4-GlcNAcT (encoded by exons 6, 7, and 8). As the mammalian β4-GalTs and the *L. stagnalis* β4-GlcNAcT use UDP-Gal and UDP-GlcNAc, respectively, it is tempting to assume that the domain that is most conserved between these enzymes, is involved in the interaction with the UDP part of the nucleotide-sugar, while the more downstream region is responsible for the specificity for donor Gal and GlcNAc, respectively. The observation that both native β4-GlcNAcT and the deletion mutant were inhibited to a similar extent by UDP, suggesting that the UDP-binding domain is not affected by the deletion, is in support of this supposition. The lower $K_m$ for UDP-GlcNAc that was found for the mutant enzyme could be explained by a higher affinity of this enzyme for the GlcNAc part of the UDP-sugar.

The change in kinetic properties observed with the mutant enzyme suggests that the region around the insertion in β4-GlcNAcT is not only involved in interaction with the sugar-donor, but also in binding of the acceptor substrate. Additionally, in β4-GalT a region (in exon 4) has been identified, that is involved in interaction of both donor and acceptor substrates (36, 37). So in both enzymes the donor and acceptor substrate binding domains seem to be structurally linked. This is conceivable, as both substrates have to be close together for the transfer reaction. In β4-GalT another binding domain for GlcNAc, however, has been localized in exon 2/3 (34). This region shows a high sequence identity with the corresponding region in β4-GlcNAcT, which supports the suggestion that this sequence is involved in acceptor binding.

We have shown here that the *L. stagnalis* β4-GlcNAcT can be transformed to an enzyme that is capable of catalyzing the transfer of GlcNAc and GalNAc with a similar maximum velocity. This suggests that the β4-GalNAcT, which has been observed in *L. stagnalis* tissues (Ref. 9 and this study), might be encoded by a structurally related enzyme belonging to the β4-GalT gene family. This is further supported by several studies that have shown a similarity between invertebrate β4-GalNAcTs and the mammalian β4-GalTs in acceptor specificity (9–12), and sometimes in responsiveness to α-lactalbumin (13). Furthermore, β4-GaIT shows sugar donor promiscuity at high concentrations of UDP-GalNAc (38), or in the presence of α-lactalbumin (24), resulting in the transfer of GalNAc. Interestingly, β4-GlcNAcT is not sensitive to α-lactalbumin (4), and this enzyme can not utilize UDP-Gal, but shows a relatively high β4-GlcNAcT activity (7%). By mutagenesis, we have constructed an enzyme with an even higher GalNAcT activity. A similar experiment has been documented for the blood group A (α2-GalNAcT) and B (α3-GalT) enzymes (39). These enzymes also show some nucleotide sugar donor promiscuity (40), but by construction of hybrids an enzyme with both activities was obtained (41). Our observations suggest that, in addition to the facilitation of the transfer of the desired sugar, the prevention of sugar donor promiscuity might have been a driving force in the evolution of enzymes of the β4-GalT gene family.

**Acknowledgments**—We thank J. Aino Andriessen, Casper Kee, and Carolien Koeleman for technical assistance in part of the experiments; Dr. Timo K. van den Berg for the gift of monoclonal antibody ED3; Dr. Bruce A. Macher for the gift of pPROTA vector; Dr. Ole Hindsgaul for the gift of pPROTA vector; Dr. Ole Hindsgaul for technical assistance in part of the experiments; Dr. Bruce A. Macher for the gift of pPROTA vector; and Dr. Richard D. Cummings for advice on lectin chromatography and the gift of WFA-Emphaze™.

**REFERENCES**

1. Schachter, H. (1991) *Curr. Opin. Struct. Biol.* 1, 755–765
2. Van den Eijnden, D. H., and Joziasse, C. A. M. (1993) *Curr. Opin. Struct. Biol.* 3, 711–721
3. Bakker, H., Agterberg, M., Van Tetering, A., Koeleman, C. A. M., Van den Eijnden, D. H., and Van Die, I. (1994) *J. Biol. Chem.* 269, 30326–30333
4. Bakker, H., Schoonmakers, P. S., Koeleman, C. A. M., Joziasse, D. H., Van Die, I., and Van den Eijnden, D. H. (1997) *Glycobiology* 7, 539–548
5. Shaper, N. L., Hollis, G. F., Douglas, J. G., Kirsch, I. R., and Shaper, J. H. (1986) *J. Biol. Chem.* 261, 10420–10428
6. Shaper, N. L., Shaper, J. H., Meuth, J. L., Fox, J. L., Chang, H., Kirsch, I. R., and Hollis, G. F. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 83, 1573–1577
7. Masri, K. A., Appert, H. E., and Fukuda, M. N. (1988) *Biochem. Biophys. Res. Commun.* 157, 657–663
8. Van den Eijnden, D. H., Neeleman, A. P., Van den Knaap, W. P. W., Bakker, H., Agterberg, M., and Van Die, I. (1995) *Adv. Exp. Med. Biol.* 376, 47–51
9. Mulder, H., Spronk, B. A., Schachter, H., Neeleman, A. P., Van den Eijnden, D. H., De Jong-Brink, M., Kamerling, J. P., and Vliegenthart, J. F. G. (1995) *Eur. J. Biochem.* 227, 175–185
10. Neeleman, A. P., Van den Knaap, W. P. W., and Van den Eijnden, D. H. (1994) *Biochem. Biophys. Res. Commun.* 207, 458–464
Deletion of Exons of the Snail β1→4-GlcNAc Transferase Gene