Structure-function Relationship of UDP-GalNAc: Polypeptide N-acetylgalactosaminyltransferase 1 (GalNAc-T1)

——Function of the Lectin Domain of GalNAc-T1——

Mari TENNO

(*Received September 12, 2003*)

Abstract

The initial step of mucin-type O-glycosylation is catalyzed by UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases (GalNAc-transferases). GalNAc-transferases contain a lectin domain at the C-terminus, consisting of three tandem repeats (α, β, and γ). I have reevaluated the role of the GalNAc-T1 lectin domain in this study, because the putative lectin domain of this enzyme is reportedly not functional. Deletion of the lectin domain resulted in a complete loss of enzymatic activity. I found that GalNAc-T1 has two activities distinguished by their sensitivities to inhibition with free GalNAc; one activity is sensitive, the other is resistant. In my experiments, the former activity is represented by the O-glycosylation of apomucin, an acceptor that contains multiple glycosylation sites, and the latter by synthetic peptides that contain a single glycosylation site. Site-directed mutagenesis of the lectin domain selectively reduced the follow-up activity, and identified Asp^{444} in the α repeat as the most important site for GalNAc recognition. I also found that the β repeat recognizes GalNAc and is involved in glycosylation of acceptors with multiple glycosylation sites. These results indicate that the lectin domain of GalNAc-T1 has at least two functional repeats, allowing the possibility of multivalent interactions with GalNAc residues on the acceptor peptide during glycosylation.

Keywords: mucin-type O-glycosylation, GalNAc-transferase, lectin domain, structure-function relationship, site-directed mutagenesis

Contents

1. Introduction
2. Materials and Methods
2.1 Site-directed mutagenesis of soluble rat recombinant GalNAc-T1
2.2 Expression of P-ΔN42 and mutant P-ΔN42 in COS7 cells
2.3 GalNAc-transferase assay
2.4 Preparation and expression of H-ΔN32 and H-ΔN32ΔC130
2.5 Binding assay with UDP-hexanolamine-agarose and apomucin-Sepharose
3. Results
3.1 Effect of deletion of the $(QXW)_3$ repeats on GalNAc-T1 activity
3.2 The disulfide bonds in the $(QXW)_3$ repeats are essential for the activity

3.3 Binding of the C-terminal deletion mutant to UDP-hexanolamine-agarose and apomucin-Sepharose
3.4 Characterization of the $(QXW)_3$ mutants
3.5 Each repeat in the lectin domain cooperatively interacts with GalNAc on the acceptor
4. Discussion

1. Introduction

The initial step in the biosynthesis of mucin-type O-glycosylation is catalyzed by a group of enzymes known as the UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases (EC 2.4.1.41) (GalNAc-transferases). These enzymes transfer GalNAc from the nucleotide-sugar donor, UDP-GalNAc, to certain serine and threonine residue on acceptor proteins [1–3]. Thus, GalNAc transferases are regarded as key enzymes in O-glycan biosynthesis because they define the number and position(s) of mucin carbohydrates on glycoproteins. Recent progress in the molecular cloning of GalNAc-transferases has revealed a large gene family [4–19]. All cloned GalNAc-transferases are type II membrane proteins [20] with a common domain structure that includes a short N-terminal cytoplasmic tail, a transmembrane domain, a stem domain, a putative catalytic domain, and C-terminal $(QXW)_3$ repeats. The central catalytic domain can be further subdivided into two regions. The N-terminal half is represented by a glycosyltransferase 1 (GT1) motif that is conserved among a wide range of glycosyltransferases [21]. The extreme C-terminal end of the GT1 motif contains a so-called DXH motif, which corresponds to the DXD sequence common to many glycosyltransferases [22]. The C-terminal half of the catalytic domain contains a so-called Gal/GalNAc-T motif, a sequence segment where significant homology can be seen between β1,4-galactosyltransferases and GalNAc-transferases [21,23].

The C-terminal $(QXW)_3$ repeats occur exclusively in the GalNAc-transferases [24]. This domain, which consists of three tandem repeats (α, β, and γ) with approximately 50 amino acid residues in each repeat unit, are homologous to the B-chain in the plant lectin ricin [24]. Hence, it is assumed that the $(QXW)_3$ repeats of the GalNAc-transferases can function as a lectin, although no direct sugar binding activity has been demonstrated. X-Ray crystallography studies have revealed distinct features of the ricin molecular structure [25–27]. Each repeat unit in the molecule has a unique, globular structure with several conserved hydrophobic amino acids at the core. Moreover, the cysteine residues in the repeats are essential for proper folding, and several amino acid residues in the molecule has been shown to interact with the haptenic monosaccharide, galactose. There is no information available about the three-dimensional structures of the $(QXW)_3$ repeats in the GalNAc-transferase lectin domains. However, computer-based analysis predicts conformational similarity with the ricin lectin domain [28,29]. Until recently, the function of the GalNAc-transferase lectin domain remained largely unknown. However, recent work has demonstrated the importance of this domain for GalNAc-T4, an isozyme that requires prior O-glycosylation of the acceptor peptides for catalytic activity [30,31]. In contrast, glycosylation by GalNAc-T1, one of the most ubiquitous isozymes, is considered independent of the $(QXW)_3$ repeats, since the mutations in this domain resulted in no, or only modest effects on

the activity [21]. In this study, I have reevaluated the functional role of the GalNAc-T1 lectin domain, and I demonstrate that this domain is involved in glycosylation of polypeptides with multiple acceptor sites. Detailed information on the structure-function relationship of the lectin domain is also presented.

2. Materials and Methods

2.1 Site-directed mutagenesis of soluble rat recombinant GalNAc-T1

Rat GalNAc-T1 cDNA was obtained as outlined by Hagen *et al.* [32]. The plasmid, *pInsProA*Δ*N*42, containing cDNA for soluble, tagged, rat GalNAc-T1 was prepared by deleting the coding sequence for the cytoplasmic tail and the transmembrane domain from full-length GalNAc-T1 cDNA, and by fusing cDNA for an insulin signal sequence and a Protein A-IgG binding domain to the resulting 5'-end of the truncated GalNAc-T1 cDNA as described previously [33]. Site-directed mutagenesis was performed on *pInsProA*Δ*N*42 using the LA PCR$^{\text{TM}}$ *in vitro* mutagenesis kit using the primers listed below [33]. Nucleotides expressed in bold face are altered nucleotides to introduce mutations to GalNAc-T1. DC130, AAT**C**AGAAATAGTGACGAGG; DC85, TTCTAATTT**A**TTTGTTGGCA; DC45, ACATGCTGC**TA**GGTTAATTT; DC1, TCTCA**TT**ATATTTCTGGAAGGG; DC3, TCTCAGAATATTT**A**TGGAAGGG; DC9, ATTTC**A**AAGAAGCCACTGCT; DC12, AAG**T**CACTGCTGGGACC; C442A, CTAGA**GC**TTGATTTGTTTCC; C459A, ACCATGA**GC**ATTAAAAATTC; C482A, GTCCAAA**GC**AAGGTCATCTG; C497A, GTGGTGG**GC**TTTGAGCATGG; C523A, TTATCCAGG**GC**CTGGTTACT; C540A, TTCCAGTG**GC**GTCTCTGATG; D444A, TTCTAGCCATGTTA**G**CTAGAC; G455Q, AAATT**TG**AACCTTTTCATTCTCT; F457A, ATACCATGACAATTA**GC**AATTCC; N465A, AGAGAAAACCTGA**GC**GCCTC; Q466A, AGAGAAAACC**GC**ATTGCCTC; F468A, TGGCAGTGTAAGAG**GC**AACC; F468Y, TGGCAGTGTAAGAG**T**AAACC; F468W, TGGCAGTGTAAGA**CC**AAACC; D484A, ACG**G**CCAAACAAAGGTCATC; D525A, GCTTTA**G**CCAGGCACTGGTT; MEcoT22I GNT1, GAATGCATCAAATGTGAAGA; 3.1R, TAGAAGGCACAGTCGAGG.

2.2 Expression of P-ΔN42 and mutant P-ΔN42 in COS7 cells

The fusion protein (P-ΔN42), containing a Protein A-IgG binding domain at the N-terminus, and mutant P-ΔN42 were expressed in COS7 cells and purified from the conditioned medium as described previously [33]. SDS-PAGE and Western blotting of the recombinant molecules were also carried out as described [33]. The protein bands on the immunoblots were scanned and quantitated with a Luminoimage Analyzer LAS-1000 PLUS (Fujifilm). The activities of P-ΔN42 and mutant P-ΔN42 were determined as described below and the activity levels were corrected for the enzyme protein concentration in the medium.

2.3 GalNAc-transferase assay

GalNAc-transferase assays using apomucin and peptide acceptors as acceptors were carried

out as described in [33] and [34], respectively. Kinetic analysis of P-ΔN42 and its mutants were carried out as described previously [33].

2.4 Preparation and expression of H-ΔN32 and H-ΔN32ΔC130

A truncated form of rat GalNAc-T1, lacking the 32 N-terminal amino acid residues, including the cytoplasmic tail and the transmembrane domain, was amplified by PCR, using full-length GalNAc-T1 cDNA, cloned in *pcDNA3.1*, as a template. For this reaction, the primer set, MEcoT22I-GNT1 and 3.1R, was used. A truncated form of rat GalNAc-T1, lacking the 32 N-terminal (including the cytoplasmic and transmembrane domains) and the 130 C-terminal amino acid residues (including the $(QXW)_3$ repeats), was amplified by PCR using the primer set, MEcoT22I-GNT1 and DC130. PCR was carried out with KOD dash DNA polymerase for 30 cycles of 96℃ for 30 sec, 50℃ for 30 sec, and 72℃ for 2 min. The PCR products were purified and ligated into the pGEM-T easy (Promega), to generate plasmids containing the truncated *GalNAc-T1 cDNAs* (*p*Δ*N*32 and *p*Δ*N*32Δ*C*130). A synthetic linker fragment, encoding His and HA tags, (HHHHHHGGYPYDVPDYAGG) with BamHI and EcoRI restriction sites at the 5'- and 3'-ends, respectively, was inserted into the BamHI-EcoRI sites of *pAcGP67A* (a baculovirus transfer vector) to produce the vector, *pAcGP67A-His/HA*. *pAcGP67A-His/HA* contains His and HA tag sequences, downstream of a secretion signal sequence and a signal peptide cleavage site, such that the in-frame ligation of a cDNA encoding a protein of interest leads to the production of a secreted protein with a His and an HA tag at the N-terminus. *p*Δ*N*32 and *p*Δ*N*32Δ*C*130 were digested with EcoT22I, and the coding sequences for Δ*N*32 and Δ*N*32Δ*C*130 were subcloned into the PstI sites of *pAcGP67A-His/HA*, generating the plasmids *pAcGP67A-His/HA/ΔN32* and *pAcGP67A-His/HA/ΔN32ΔC130*, for expression of secreted, truncated rat GalNAc-T1 fused to His and HA tags (H-ΔN32 and H-ΔN32ΔC130). Sf9 cells were grown at 27℃ in TMN-FH medium (Pharmingen) containing 10 % fetal calf serum. Co-transfection of Sf9 cells, with the plasmid (*pAcGP67A-His/HA/ΔN32* or *pAcGP67A-His/HA/ΔN32ΔC130*) and Baculo-Gold^{TM} DNA (Pharmingen), was performed according to the manufacture's specifications. Briefly, 0.4 μg of *pAcGP67A-His/HA/ΔN32* or *pAcGP67A-His/HA/ΔN32ΔC130* was mixed with 0.1 μg of Baculo-Gold^{TM} DNA and then co-transfected into Sf9 cells in 24-well plates. Five days after transfection, recombinant viruses were collected from the culture medium. To prepare large virus stocks, the recombinant viruses were inoculated in Sf9 cells with multiplicity of infection (MOI) of 0.01, and the culture media were harvested after 3–5 days. For the expression of H-ΔN32 and H-ΔN32ΔC130, recombinant viruses were inoculated into Sf9 cells with a MOI of 5. After 3–5 days, the culture medium containing the recombinant proteins was collected.

2.5 Binding assay with UDP-hexanolamine-agarose and apomucin-Sepharose

To assess binding activity with UDP, conditioned medium containing either H-ΔN32 or H-ΔN32ΔC130 was mixed with 20 μl of UDP-hexanolamine-agarose, in 600 μl of Buffer A (25 mM imidazole-HCl buffer, pH 7.2, containing 20 mM $MnCl_2$ and 100 mM NaCl), and incubated by rocking, at 4℃, for 2 hours. Following centrifugation at 2,500 rpm for 5 min, the

supernatant was recovered and the resin was washed with Buffer A. The supernatants were combined and referred to as the unbound fraction. The resin was washed twice with Buffer B (25 mM imidazole-HCl buffer, pH 7.2, containing 20 mM $MnCl_2$, 1 M NaCl, and 0.1% Triton X-100). Adsorbed proteins were eluted by washing the resin twice with Buffer C (25 mM imidazole-HCl buffer, pH 7.2, containing 1 M NaCl and 50 mM EDTA). Finally, strongly bound proteins that were not eluted with EDTA, were released by boiling the resin in SDS-PAGE sample buffer. Mutant GalNAc-T1 proteins in each fraction were immunoprecipitated by incubating with an anti-His_6 tag antibody (Invitrogen) and Protein-G Sepharose. The precipitated proteins were analyzed by SDS-PAGE, followed by Western blotting using PVDF membranes. The recombinant proteins on the membrane were probed with anti-HA 3F10 antibodies conjugated with peroxidase (Roche) and visualized using the ECL Western blotting analysis system from Amersham Pharmacia Biotech. In the binding experiments with apomucin-Sepharose, samples containing either H-ΔN32 or H-ΔN32ΔC130 were fractionated in a similar fashion with the exception that Buffers A and B contained 0.25 mM UDP. Proteins recovered from each fraction were analyzed by SDS-PAGE followed by Western blotting and detection by ECL.

3. Results

3.1 Effect of deletion of the $(QXW)_3$ repeats on GalNAc-T1 activity

To investigate the role of the GalNAc-T1 $(QXW)_3$ repeats in catalysis, I prepared three deletion mutants, P-ΔN42ΔC45, P-ΔN42ΔC85, and P-ΔN42ΔC130, in which the γ repeat unit, the γ and β repeat units, and the entire $(QXW)_3$ repeat domain were deleted, respectively (Fig. 1). The mutant P-ΔN42 molecules, as well as intact P-ΔN42 were subsequently expressed in COS7 cells, the secreted fusion proteins were isolated from the culture medium, and their activities were determined, using deglycosylated bovine submaxillary mucin (apomucin) as acceptor. Previous work has demonstrated that bovine submaxillary apomucin is an efficient substrate of GalNAc-T1 [35]. The molecule is characterized by a high content of serine, threonine, and proline residues and as a consequence contains a large number of putative sites for O-glycosylation [36–39]. Chemical analysis of the apomucin preparation used in this investigation failed to detect any remaining GalNAc, indicating that virtually all the GalNAc residues were released from the mucin molecule. The secreted fusion proteins were quantified by Western blotting, as described under Materials and Methods. Their enzymatic activities were corrected for the levels of enzyme protein, and expressed as relative to that of P-ΔN42 such that the effect of each mutation, on the enzymatic activity, could be evaluated. Western blotting of the COS7 cells produced proteins showed that their molecular weights correlated well with the degree of the truncation, and also that the expression levels of all the mutant proteins were significantly decreased, as compared to P-ΔN42 (Fig. 1). Moreover, the enzymatic activities of the mutant proteins were reduced to almost background level. This suggests that the $(QXW)_3$ repeats are important for stable expression and/or secretion, as well as for the enzymatic activity of GalNAc-T1.

To assess the importance of the C-terminal sequence for the activity of GalNAc-T1, the

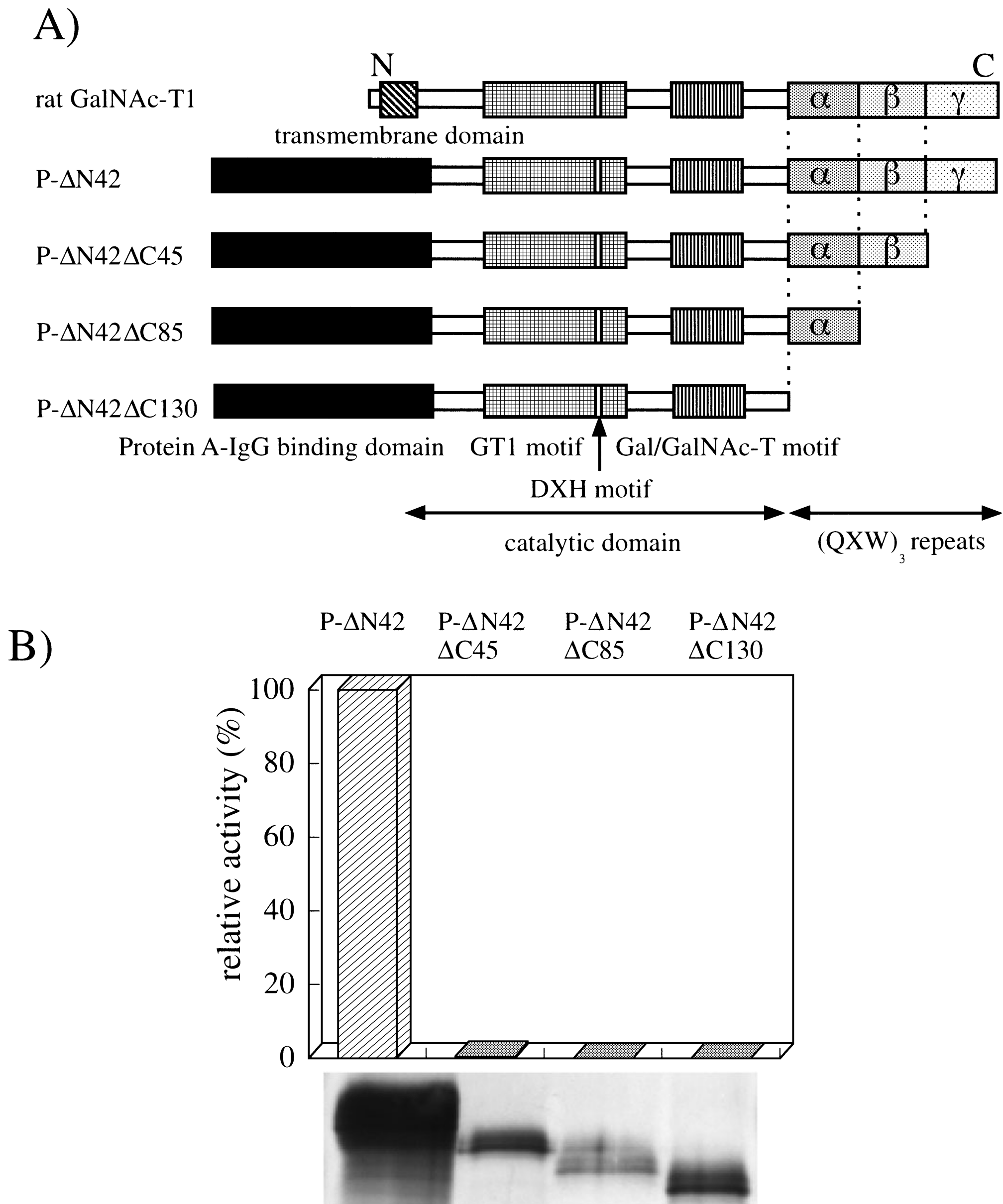


Fig. 1 Expression and enzymatic activity of the $(QXW)_3$ deletion mutants
A) Schematic representation of wild type and mutant GalNAc-T1. B) Enzymatic activity of mutant GalNAc-T1 enzymes. The mutant enzymes were expressed in COS7 cells and the secreted recombinant proteins were recovered from the culture medium. The amount of the secreted protein was determined by Western blotting followed by densitometric scanning of the blot (lower panel). The enzymatic activity secreted in the medium was determined using apomucin as acceptor, corrected for the amount of the enzyme protein in the medium, and expressed as activity relative to that of P-ΔN42. Mutant enzymes, solid bars; P-ΔN42, hatched bar

deletion mutants P-ΔN42ΔC1, P-ΔN42ΔC3, P-ΔN42ΔC9, and P-ΔN42ΔC12 were generated (Fig. 2A). As shown in Fig. 2B, P-ΔN42ΔC1 and P-ΔN42ΔC3 were fully active, while the activity of P-ΔN42ΔC9 was about 50% of that of P-ΔN42. Interestingly, the deletion of 12 amino acid residues from the C-terminus (P-ΔN42ΔC12) resulted in complete inactivation of P-ΔN42. The reason for the inactivation of this mutant enzyme is not clear, but it may be related to the loss of Trp^{548}. This tryptophan residue is conserved in ricin and in all GalNAc-transferases cloned to date. Moreover, it is reported to be essential for the formation of the hydrophobic core in the globular lectin domain [25–27]. The results from the C-terminal deletion experiments suggest that a properly folded $(QXW)_3$ repeat domain is essential for the enzymatic activity of

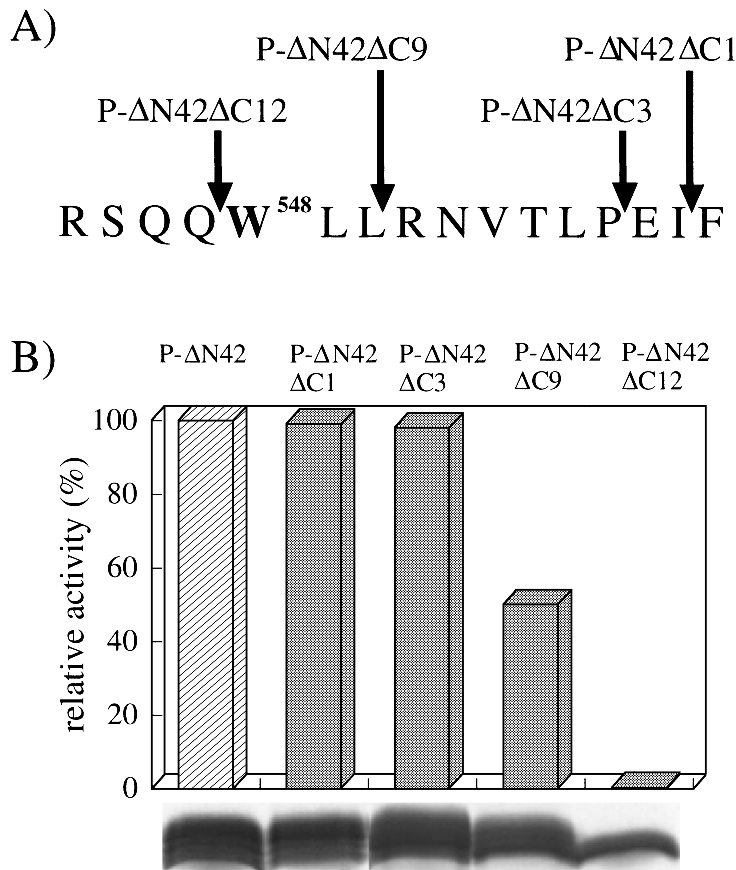


Fig. 2 Enzymatic activity of C-terminal deletion mutant enzymes
A) C-Terminal amino acid sequence of GalNAc-T1. The C-terminus of each deletion mutant is indicated by an arrow. B) Enzymatic activity of the C-terminal deletion mutants. Enzyme assay and Western blotting of the expressed mutants were carried out as described in the legend to Fig. 1. Mutant enzymes, solid bars; P-ΔN42, hatched bar

GalNAc-T1.

3.2 The disulfide bonds in the $(QXW)_3$ repeats are essential for the activity

The $(QXW)_3$ repeats are predicted to contain three disulfide bonds (Cys^{442}-Cys^{459}, Cys^{482}-Cys^{497}, and Cys^{523}-Cys^{540}), *i.e.*, one disulfide bridge in each repeating unit (Fig. 3A) [28]. It is likely that these disulfide bonds stabilize the globular conformation of each unit. To examine whether disruption of the disulfide bonds in the repeats affects the enzymatic activity, I prepared mutant constructs for each cysteine residue in the $(QXW)_3$ repeat domain. Fig. 3B shows that all the mutant proteins generated were completely inactive. Moreover, their expression levels were significantly decreased. Consistent with the results from my constructs with deleted $(QXW)_3$ repeats (see above), this again indicates that the integrity of this domain is important

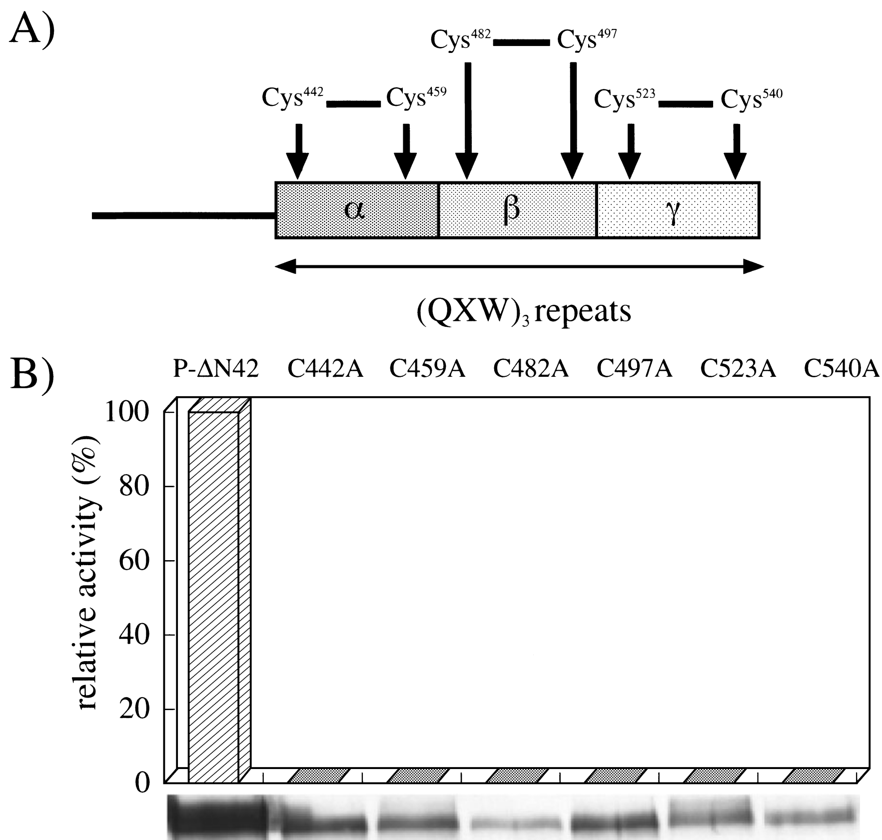


Fig. 3 Enzymatic activity of enzymes with mutated cysteine residues in the $(QXW)_3$ repeats
A) Positions of the disulfide bonds in the GalNAc-T1 $(QXW)_3$ repeats. B) Enzymatic activity of the mutant enzymes. Enzyme assays and Western blotting of mutant proteins were carried out as described in the legend to Fig. 1. Mutant enzymes, solid bars; P-ΔN42, hatched bar

for GalNAc-T1 activity.

3.3 Binding of the C-terminal deletion mutant to UDP-hexanolamine-agarose and apomucin-Sepharose

The inactivation of P-ΔN42, caused by truncation of the $(QXW)_3$ repeats or disruption of the disulfide bonds in the $(QXW)_3$ repeats, may result either from a conformational change in the entire enzyme molecule, or from malfolding of the C-terminal repeat domain. To evaluate the effects of deletion of the C-terminal domain, I compared the UDP-binding activities of the GalNAc-T1 mutants. As reported previously, GalNAc-T1 primarily recognizes the UDP portion of UDP-GalNAc [3,40]. In fact, UDP-affinity chromatography has been used successfully for purification of GalNAc-transferases [41,42]. To investigate the capacity of the mutant enzymes for nucleotide sugar binding, I generated the GalNAc-T1 mutant enzymes, H-ΔN32 and H-ΔN32ΔC130 that, instead of a large Protein A-IgG binding domain, have a His tag and a haemagglutinin (HA) epitope tag (19 amino acid residues in total) fused to N-terminus (Fig. 4A). This was done since my preliminary binding experiments demonstrated that enzymatically active P-ΔN42 failed to bind to UDP-hexanolamine-agarose, while P-ΔN42ΔC130 retained capacity for binding (data not shown). I considered it possible that the large Protein A-IgG binding domain ($\sim$ 30 kDa) in P-ΔN42 (Fig. 1A) could sterically interfere with the binding to UDP immobilized on agarose. Deletion of the entire C-terminal lectin domain might mitigate this steric hindrance, thereby making retention of P-ΔN42ΔC130 on the UDP-hexanolamine-agarose possible. The His/HA-tag fusion constructs were expressed as secreted proteins in insect cells, using the baculovirus expression system. The secreted molecules were purified with Ni^{+}-chelating chromatography and assayed for enzymatic activity. While H-ΔN32 was enzymatically active at levels similar to P-ΔN42, H-ΔN32ΔC130 was inactive. This was also the case for P-ΔN42ΔC130. Culture media containing either H-ΔN32 or H-ΔN32ΔC130 were mixed with UDP-hexanolamine-agarose. Following incubation, the suspension was centrifuged to obtain fractions containing UDP-hexanolamine-unbound (the supernatant) and UDP-hexanolamine-bound (the resin) material. The UDP-hexanolamine-bound fraction was washed with the buffer containing 1 M NaCl and 0.1% Triton X-100. The adsorbed proteins on the resin were then eluted by chelating Mn^{2+}, a metal ion essential for donor substrate binding, with EDTA. H-ΔN32 or H-ΔN32ΔC130 recovered in the bound and unbound fractions was immunoprecipitated with an anti-His_6 tag antibody, subjected to SDS-PAGE, and analyzed by Western blotting, using a peroxidase conjugated anti-HA-antibody. Adsorbed proteins not eluted with EDTA were released by boiling the resin in the SDS-PAGE sample buffer. In contrast to the results from the constructs containing a Protein A-IgG binding domain, both H-ΔN32 and H-ΔN32ΔC130 were capable of binding to UDP-hexanolamine-agarose (Fig. 4B). H-ΔN32 appeared to bind stronger than H-ΔN32ΔC130. Still, small amounts of the recombinant proteins were also found in the unbound fractions. This may represent denatured proteins, since no enzymatic activity was detected in these fractions (data not shown). The specificity of the interaction(s) of H-ΔN32 and H-ΔN32ΔC130 with UDP-hexanolamine-agarose was investigated by preincubating both molecules with either 2 mM UDP or 50 mM GDP, before

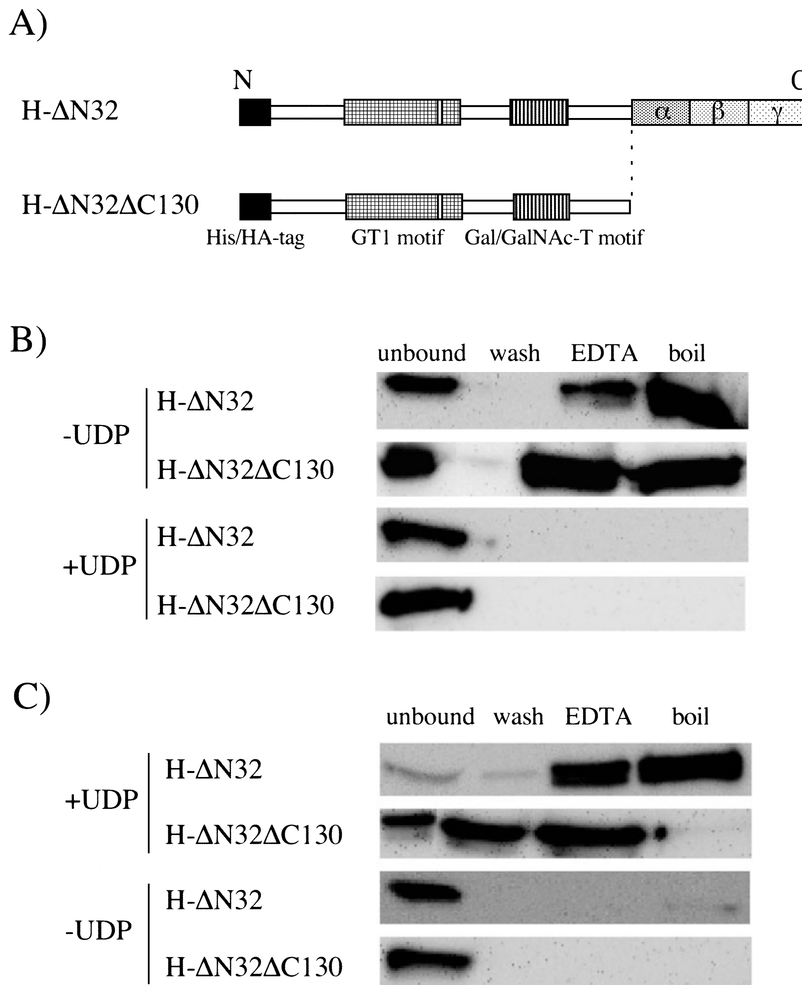


Fig. 4 Binding of H-ΔN32 and H-ΔN32ΔC130 to UDP-hexanolamine-agarose and apomucin-Sepharose A) Schematic representation of the His/HA tagged recombinant molecules used for the binding experiments. B) Binding of the His/HA tagged proteins to UDP-hexanolamine-agarose. H-ΔN32 and H-ΔN32ΔC130 were expressed in Sf9 cells, and the recombinant proteins in the conditioned medium were fractionated as described under Materials and Methods. In the control experiments, H-ΔN32 and H-ΔN32ΔC130 were incubated with either 2 mM UDP or 50 mM GDP before mixing with UDP-hexanolamine-agarose. The His/HA tagged proteins recovered in the bound and unbound fractions were immunoprecipitated with an anti-His_6 tag antibody. The immunoprecipitated proteins were separated on SDS-PAGE and blotted to a PVDF membrane. The His/HA tagged proteins on the membrane were probed with an anti-HA-antibody (3F10) conjugated to peroxidase and detected with an ECL Western blotting analysis system. C) Binding of the His/HA tagged proteins to apomucin-Sepharose. Binding assays were carried out as described in B) except that apomucin-Sepharose was used instead of UDP-hexanolamine-agarose and that the buffers contained 0.25 mM UDP. Control experiments were carried out in the absence of UDP.

adding the resin. Figure 4B shows that while GDP had no effect (data not shown), UDP completely blocked the binding of both molecules to UDP-hexanolamine-agarose, suggesting that the binding of both H-ΔN32 and H-ΔN32ΔC130 is specific for the UDP moiety on the resin. Taken together, these results suggest that the C-terminal deletion mutant proteins retain the capacity for UDP-binding, although they lack enzymatic activity.

Next, I compared the acceptor-binding activity of H-ΔN32 and H-ΔN32ΔC130. Medium containing either recombinant protein was mixed with apomucin-Sepharose. The apomucin-resins were treated as described above for UDP-hexanolamine-agarose, except that the buffers contained 0.25 mM UDP. H-ΔN32 bound quite avidly to apomucin (Fig. 4C). Most of H-ΔN32 added to the apomucin matrix was recovered in the EDTA and SDS-PAGE sample buffer eluates. H-ΔN32 recovered in the unbound fraction was enzymatically inactive, suggesting that all enzymatically active molecules bound to the matrix. H-ΔN32ΔC130 also bound to apomucin-Sepharose, although with the significantly lower affinity. Bound H-ΔN32ΔC130 was for the most part recovered in the EDTA eluate and in the wash fraction. None of the protein was recovered in the SDS-PAGE sample buffer eluate. It should be noted that both H-ΔN32 and H-ΔN32ΔC130 required UDP for binding to apomucin. Without UDP, both failed to bind and were recovered exclusively in the unbound fractions (Fig. 4C). It appears likely that the binding of UDP produces a change in the conformation of the catalytic domain that allows binding to apomucin. The lectin domain appears not to be involved in this conformational change.

Together with the experiments described above, these results suggest that H-ΔN32ΔC130, though enzymatically inactive, retains the capacity to bind to both UDP and apomucin. Consequently, deletion of the lectin domain from GalNAc-T1 apparently does not cause any change in the overall conformation of the catalytic domain of the enzyme. The data also indicate that the catalytic domain in itself may be sufficient for binding to the substrates and that the interaction of the enzyme with the acceptor substrate (apomucin) was based on the protein-protein interaction, since the mutant, H-ΔN32ΔC130, in which the entire $(QXW)_3$ domain is deleted, was capable of binding with apomucin. The reason for the weaker binding of H-ΔN32ΔC130 to apomucin and UDP, as compared to H-ΔN32, is not clear, but it may be related to the specific interactions between the catalytic and lectin domains of the molecule, as described in Discussion.

3.4 Characterization of the $(QXW)_3$ mutants

Figure 5 shows a comparison of the amino acid sequences of ricin and GalNAc-T1 $(QXW)_3$ repeats. X-Ray crystallography studies of the ricin $(QXW)_3$ repeats have demonstrated that they contain two independent active domains (1α and 2γ), both of which are capable of binding lactose [25,26]. In these domains, the haptenic monosaccharide galactose, is located in a cleft, in which conserved hydrophobic residues contribute to form a compact hydrophobic core. In the ricin 1α domain (Fig. 5A), Asp^{22}, Asn^{46}, Gln^{47}, Gln^{35}, and Trp^{37} (Asp^{234}, Asn^{255}, Gln^{256}, Ile^{246}, and Tyr^{248} in the 2γ domain) are involved in binding with galactose. Trp^{49} (Trp^{258} in the 2γ domain) is an essential aromatic residue for the hydrophobic core of the globular domain [25,26]. Although the sequence identities are modest between the $(QXW)_3$ repeats of

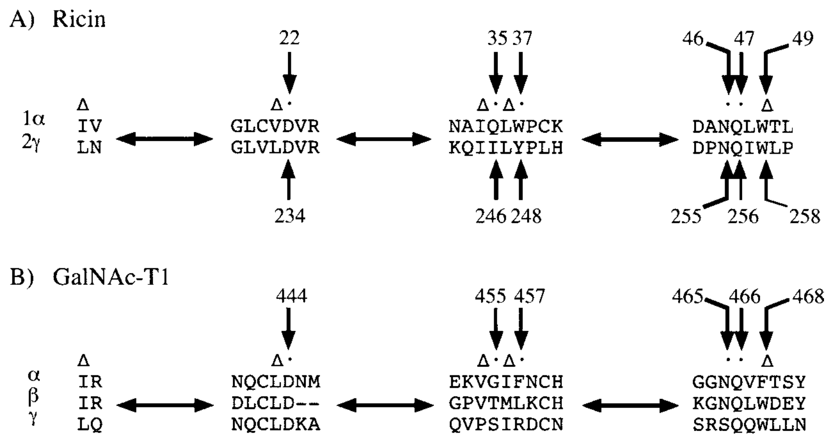
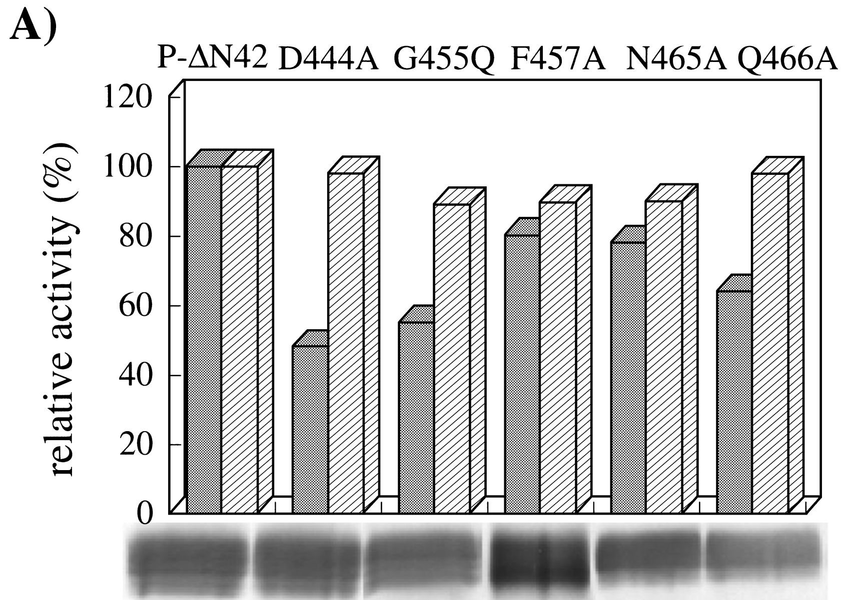
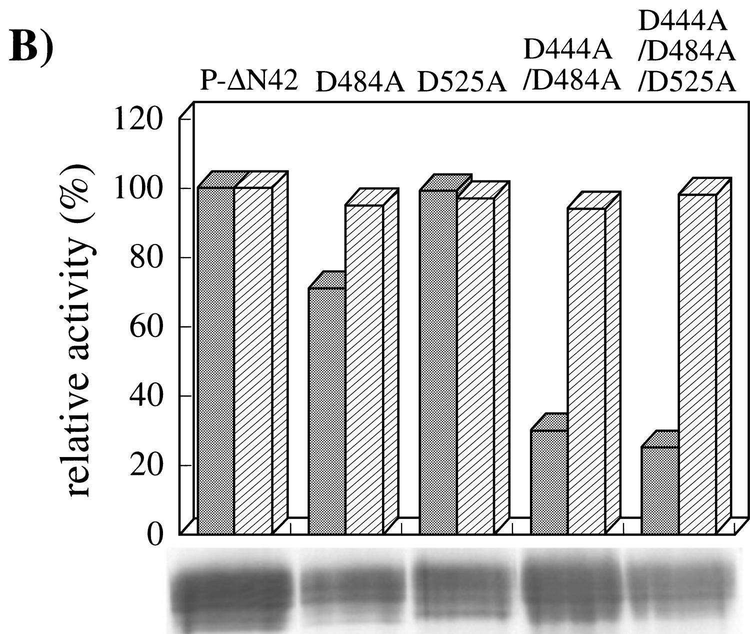


Fig. 5 Amino acid alignment of the $(QXW)_3$ repeats in ricin and GalNAc-T1
Putative amino acids involved in carbohydrate binding are marked by dots and amino acids required for formation of the hydrophobic core are denoted by triangles. Amino acid residues discussed in the text are indicated by arrows and numbered from the N-terminus of the proteins. Segments with low similarity are omitted. A) Amino acid sequence of the 1α and the 2γ repeats in the ricin B-chain. B) Amino acid sequence of the GalNAc-T1 $(QXW)_3$ repeat domain.

GalNAc-T1 and ricin, threading methods predict both molecules to have similar globular conformations [24,28]. Judging from the structure of ricin, the α repeat, of the three repeats in GalNAc-T1, is most likely to have sugar-binding activity, since this unit contains all amino acid residues required for the sugar binding and the hydrophobic core formation, with the exception of Gly^{455} (Fig. 5B). I investigated the role of the GalNAc-T1 α repeat unit using site-directed mutagenesis of amino acids corresponding to residues identified as being involved in ligand binding in ricin. I mutated the possible sugar-binding sites, Asp^{444}, Phe^{457}, Asn^{465}, and Gln^{466} of P-ΔN42, into alanine. I also mutated Gly^{455} into glutamine such that the mutated α repeat unit would have sugar-binding sites identical to those in ricin. Phe^{468}, a residue putatively essential for the hydrophobic core structure, was also mutated. Three different mutations (F468W, F468Y, and F468A) were generated at this site. F468W and F468Y represent conservative amino acid replacements, while F468A is expected to result in disruption of the hydrophobic core due to loss of the aromatic ring and also to cause a conformational change in the repeat domain. All mutant sequences were expressed in COS7 cells and the activities of the mutant enzymes were determined using apomucin or the synthetic peptide, PPDAATAAPL, as acceptors. PPDAATAAPL, which represents a modification of an O-glycosylated sequence in erythropoietin, contains a single glycosylation site. The peptide is a very efficient acceptor for GalNAc-T1, exhibiting K_m value of approximately 0.20 mM [3].

The enzymes with single point mutations at the putative sugar binding sites (Fig. 6A) were secreted into the medium, to the same extent as P-ΔN42, suggesting that the mutated sites had little impact on the secretion and stability of the molecules.


A)
P-ΔN42
D444A
G455Q
F457A
N465A
Q466A
relative activity (%)
120
100
80
60
40
20
0


B)
P-ΔN42
D484A
D525A
D444A
/D484A
D444A
/D484A
/D525A
relative activity (%)
120
100
80
60
40
20
0

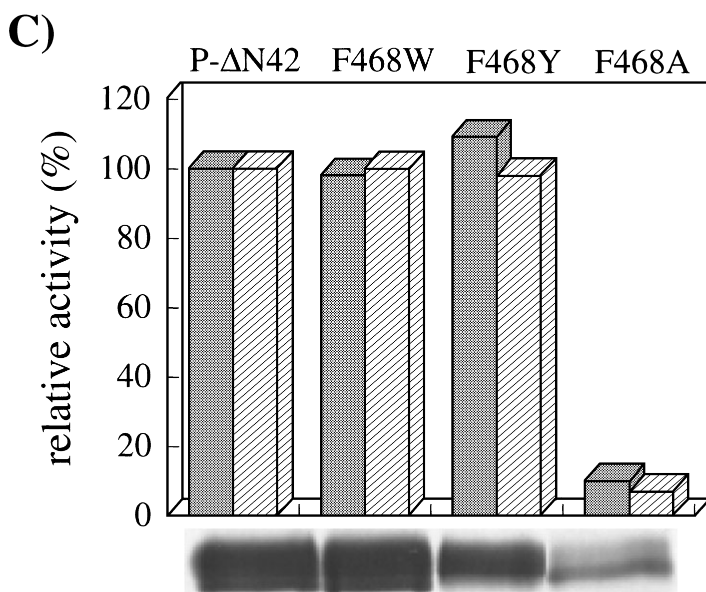


Fig. 6 Activity of enzymes containing mutations in the $(QXW)_3$ repeats
Enzyme activity measurements and Western blotting of mutant proteins were carried out as described in the legend to Fig. 1. Solid and hatched bars indicate enzymatic activity determined using apomucin and PPDAATAAPL as acceptor, respectively. A) Enzymes containing mutations at the putative sugar binding sites. B) Enzymes containing mutations at the Asp residues. C) Enzymes containing mutations of the putative amino acid required for formation the hydrophobic core.

Moreover, the activity when using the peptide PPDAATAAPL, was essentially unaffected. On the other hand, the relative activities of the mutant enzymes when using apomucin as acceptor, were significantly decreased, compared to the parent molecules. Asp^{444} appears most important among the putative sugar binding sites in the α repeat domain, since apomucin glycosylation by D444A was most severely impaired, among the mutant enzymes. The aspartate residues in the β and γ repeats, Asp^{484} and Asp^{525}, corresponding to Asp^{444} in the α repeat, are also conserved (Fig. 5) suggesting a possible role for these residues in the binding to apomucin. To investigate the functional importance of the β and γ repeats, I generated two additional mutants, D484A and D525A. The measured relative activity of D484A toward apomucin was significantly decreased, although not to the level of D444A (Fig. 6B). On the other hand, similar to D444A, the activity of D484A using the acceptor peptide, PPDAATAAPL, was essentially unaffected. In contrast to D444A and D484A, the activity of D525A was unchanged with both substrates (Fig. 6B). This suggests that the β repeat, as well as the α repeat, is important for glycosylation of substrates with multiple acceptor sites, but the γ repeat may not be. To confirm the role of individual repeat units in apomucin glycosylation, I generated double and triple mutant enzyme, D444A/D484A and D444A/D484A/D525A. These molecules contain mutations at the putative sugar-binding aspartic acid residues in the α and β repeats and in all three repeats, respectively. As shown in Fig. 6B, the activity of D444A/D484A with apomucin was

low as that of D444A/D484A/D525A. The activity of the both mutants, at 20% the activity of P-ΔN42, were significantly lower than that of D444A. Moreover, the mutant enzymes glycosylated the peptide acceptor as efficiently as the wild type enzyme. These results suggest that, at least, two of the $(QXW)_3$ repeats interact cooperatively with acceptors with multiple acceptor sites.

Taken together, the data indicate that the observed partial decrease in apomucin O-glycosylation by the mutant enzymes may be attributed to biased O-glycosylation of the acceptor sites in apomucin, rather to an overall decrease in glycosylation efficiency, since the peptide acceptor PPDAATAAPL was glycosylated at similar efficiency by both the wild type and mutant enzymes; *i.e.*, it appears that the mutant enzymes are capable of glycosylating some sites in apomucin as efficiently as the wild type, while other sites are glycosylated at a slower rate or not at all.

Mutation at Phe^{468} strongly influenced both the expression level and activity of the mutant enzymes (Fig. 6C). In contrast to the enzymes with conservative amino acid replacements (F468W and F468Y), the expression of F468A was significantly decreased, and its enzymatic activity, using either acceptor was reduced to almost background level. The presence of an aromatic group appears to be essential at position 468, as predicted by its role in the hydrophobic core formation of the lectin domain [28]. The inactivation of GalNAc-T1 resulting from site-directed mutagenesis as well as the C-terminal deletions is likely caused by conformational changes in the $(QXW)_3$ repeats.

To evaluate the potential involvement of the $(QXW)_3$ repeats in catalysis, kinetic analysis was carried out on the mutant enzymes (Table 1). All the mutant enzymes generated K_m values for UDP-GalNAc that were almost identical to that of P-ΔN42, in spite of the fact that the mutations were located at potential sugar-binding sites in the repeats. The sugar-binding activity of the $(QXW)_3$ repeats appears, therefore, to be involved in acceptor recognition, rather

Table 1 Kinetic analysis of the $(QXW)_3$ mutants

	UDP-GalNAc		apomucin	
	K_m (mM)	-fold	K_m (mg/ml)	-fold
P-ΔN42	5.1 ± 0.8	1.00	4.7 ± 0.01	1.00
D444A	5.1 ± 0.6	1.00	39.2 ± 7.3	8.34
G455Q	4.0 ± 0.1	0.78	16.2 ± 3.8	3.45
F457A	5.4 ± 0.2	1.06	8.2 ± 1.5	1.74
N465A	3.6 ± 0.5	0.71	6.5 ± 0.4	1.38
Q466A	6.1 ± 0.1	1.20	15.5 ± 0.1	3.30
D484A	4.1 ± 0.4	0.8	17.1 ± 0.4	3.6
D525A	4.3 ± 0.2	0.9	4.6 ± 0.4	1.0
D444A/D484A	5.9 ± 0.5	1.2	93.0 ± 9.7	19.8
D444A/D484A/D525A	4.4 ± 0.9	0.9	106.1 ± 2.8	22.6

K_m for UDP-GalNAc and apomucin was determined as described under Materials and Methods. Values are the average of three distinct reactions.

than in the binding to the sugar donor. Consistent with this, the affinity of the mutant enzymes for apomucin, was significantly decreased (Table 1). There was a clear correlation between the affinity for apomucin and the acceptor activity of this protein (Table 1 and Fig. 6A). Of the mutant enzymes tested, D444A, D444A/D484A, and D444A/D484A/D525A showed the marked increase (approximately 8- and 23-fold, respectively) in their K_m values for apomucin. The $(QXW)_3$ repeats, therefore, appear to be involved in the binding of acceptors with multiple glycosylation sites. It appears possible that the lectin domain promotes the glycosylation of polypeptides with multiple acceptor sites by recognizing GalNAc residues at adjacent sites on the acceptor. Since apomucin does not contain any detectable GalNAc residues, the lectin domain of the enzyme must interact with GalNAc residues transferred to apomucin in the course of the reaction.

3.5 Each repeat in the lectin domain cooperatively interacts with GalNAc on the acceptor

In order to further investigate the role of the $(QXW)_3$ repeats in the recognition of apomucin, the influence of monosaccharides on the GalNAc-T1 activity was examined. In a previous report, I demonstrated that micromolar concentrations of free GalNAc do not inhibit GalNAc-T1 catalyzed glycosylation [3]. However, a high concentration of GalNAc was recently shown to inhibit GalNAc-T4, a follow-up type isozyme that requires the prior O-glycosylation of the acceptor polypeptides [30]. To investigate if GalNAc-T1 was similarly affected by high sugar concentrations, I studied the effect of 300 mM monosaccharides on the activity of P-ΔN42. In reactions using apomucin as acceptor (Fig. 7A), GalNAc was a quite potent inhibitor. Approximately 80% of the GalNAc incorporation into apomucin was inhibited. By contrast, Gal, Glc, GlcNAc, Man, and Fuc had no significant inhibitory activity. The inhibition of P-ΔN42 by GalNAc was concentration-dependent, generating an apparent K_i of 68 mM (Fig. 8). This is almost as efficient as the reported inhibition of GalNAc-T4 by this sugar (K_i = 50 mM)[30]. Interestingly, the inhibitory effect of free GalNAc was entirely dependent on the type of acceptor used. When PPDAATAAPL was used, there was no measurable inhibition by GalNAc (or any of the other monsaccharides tested) (Fig. 7B). This indicates that GalNAc-T1 has two activities that can be distinguished by the inhibition with free GalNAc, *i.e.* one activity is sensitive to the inhibition, and the other is resistant. In my experiments, the former activity is represented by the O-glycosylation of apomucin that contains multiple glycosylation sites, and the latter by the single-acceptor site peptide. To confirm the involvement of the lectin domain in the GalNAc sensitivity of GalNAc-T1, analogous experiments were carried out with the $(QXW)_3$ mutants. As discussed above, these mutant enzymes display only moderate activity toward apomucin (Fig. 6A, 6B, and Table 1). The glycosylation of apomucin was affected significantly (Fig. 7A). Again, only GalNAc, of the monosaccharides tested, was effective. And, as expected, the GalNAc inhibition was correlated with the extent of mutations in the lectin domain (Fig. 8). Of the single-point mutant enzymes, D444A exhibited the highest degree of resistance to GalNAc inhibition. In analogy with P-ΔN42, the inhibition of D444A by GalNAc was concentration-dependent (Fig. 8), with K_i of 173 mM. This is 2.5-fold higher than that (68 mM)

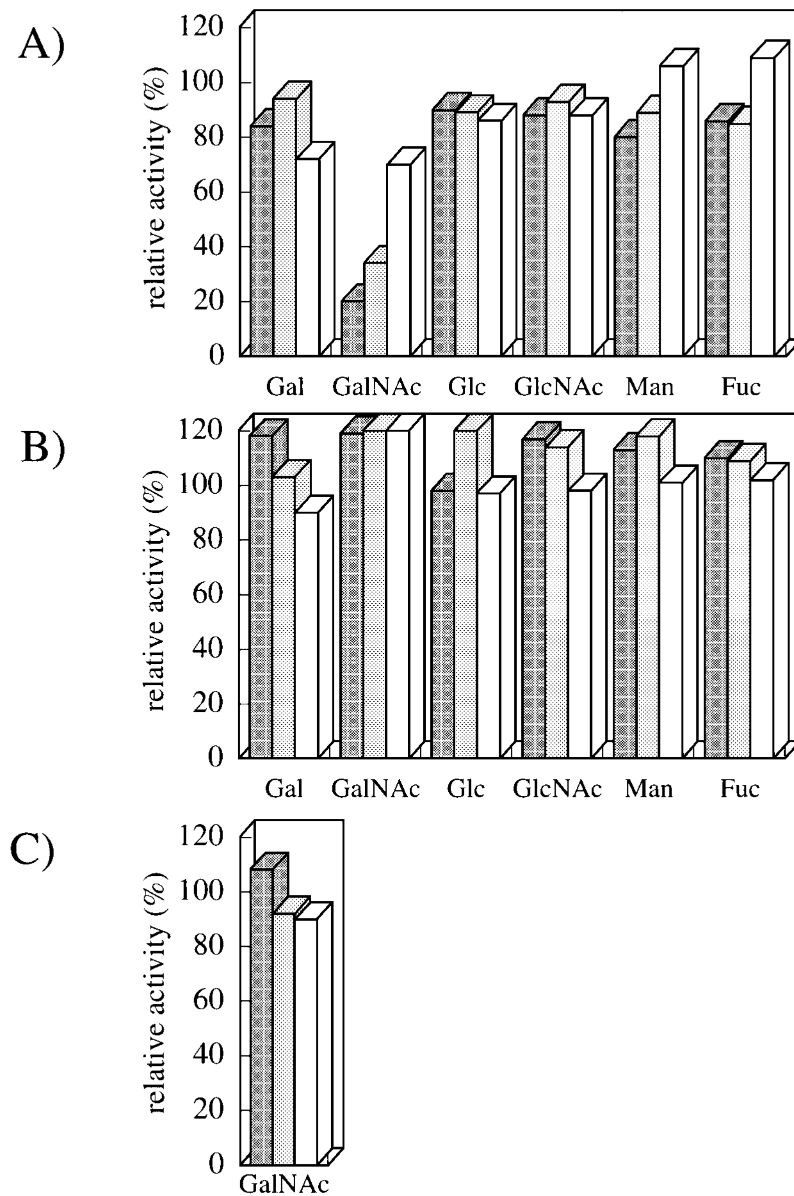


Fig. 7 Inhibition of enzymatic activity by free monosaccharides

The inhibitory activity of monosaccharides was determined using A) apomucin, B) PPDAATAAPL, and C) GVVPTVVPG, as acceptors. The enzymatic activity of the recombinant proteins is expressed as relative to the activity determined in the absence of added monosaccharides. Solid bars denote the activity of P-ΔN42, hatched and white bars that of the mutant enzymes D444A, and D444A/D484A/D525A, respectively.

of P-ΔN42. D484A was more resistant than D525A, while D525A was almost as sensitive to the GalNAc inhibition as the parent enzymes. I also observed an almost identical decrease in affinity for GalNAc with D444A/D484A and D444A/D484A/D525A, which were more resistant to inhibition than D444A. The very low affinity of D444A/D484A and D444A/D484A/D525A for GalNAc made determination of an inhibition constant difficult. These mutant enzymes retained 70% of its activity towards apomucin even in the presence of 300 mM GalNAc (Figs. 7A and 8). By contrast, when assayed with the peptide PPDAATAAPL (Fig. 7B), no inhibition by monosaccharides was observed for these mutant (Fig. 7B). I also investigated the inhibition by GalNAc, using another single acceptor site peptide, GVVPTVVPG. This peptide is also an efficient acceptor with a K_m value of approximately 0.80 mM. Again, the peptide was glycosylated at the similar efficiency by both the parent and the mutant enzymes, and no significant inhibition by GalNAc was observed (Fig. 7C).

Taken together, the data clearly suggest that the lectin domain of GalNAc-T1 is involved in glycosylation of substrates with multiple acceptor sites such as apomucin. GalNAc-T1,

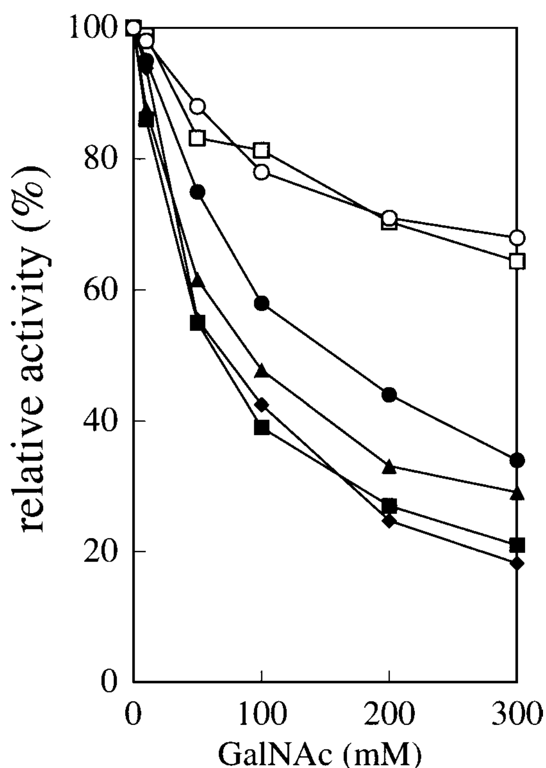


Fig. 8 Inhibition of enzymes containing mutations in the $(QXW)_3$ repeat domain, with free GalNAc
The recombinant enzymes were expressed in COS7 cells and purified from the conditioned medium on IgG-Sepharose. The enzymatic activity of the mutant proteins was assayed in the presence of increasing concentrations of free GalNAc. Apomucin was used as acceptor. ■: P-ΔN42,●: D444A, ▲: D484A, ◆: D525A, □: D444A/D484A, and ○: D444A/D484A/D525A.

appears to recognize GalNAc residues on the acceptor polypeptides that *in situ* may have been transferred by GalNAc-T1 itself or another GalNAc-transferase. Moreover, the α and β repeats in the lectin doman of GalNAc-T1 are capable of recognizing GalNAc, with the α repeat having the higher affinity. These two repeat units in the lectin domain functions in a cooperative manner in the recognition of GalNAc residues on acceptor polypeptides.

4. Discussion

This report demonstrates a measurable function for the (putative) lectin domain on GalNAc-T1. I demonstrate that the $(QXW)_3$ repeats of GalNAc-T1 are functional and that they have a defined function in the glycosylation of substrates with multiple acceptor sites, such as apomucin. Glycosylation of the synthetic peptides, PPDAATAAPL and GVVPTVVPG, was independent of the $(QXW)_3$ repeats, as demonstrated by the efficient glycosylation of these peptides by both the wild type and the mutant enzymes (Figs. 6 and 7). I have also demonstrated that the function of the $(QXW)_3$ repeats is related to carbohydrate recognition (Figs. 7 and 8). Based on these observations, I hypothesize that GalNAc-T1 catalyzes two distinct reactions. One is the initial O-glycosylation of unglycosylated polypeptide acceptors. This reaction is independent of the $(QXW)_3$ repeats and is not inhibitable by free GalNAc. Consequently, the substrate specificity of this reaction would be determined predominantly by information in the acceptor polypeptides, such as the primary and secondary structures [3,43]. The other reaction is a follow-up glycosylation. This would be dependent on the lectin activity of the $(QXW)_3$ repeats and would be inhibitable by GalNAc. The follow-up glycosylation would be significantly influenced by the structure of the acceptor polypeptides and their degree of glycosylation [30,31,44–46]. The addition of β1,3Gal to GalNAc would inhibit the follow-up reaction, since Gal is not recognized by the $(QXW)_3$ repeats (Fig. 7). A similar binding specificity has been observed for GalNAc-T4 [30,31]. Thus, it appears plausible that the presence of a disaccharide, Galβ1,3GalNAc, at one potential O-glycosylation site may reduce O-glycosylation at neighboring potential sites, on the same peptide [31,44].

Considering the possible role of the $(QXW)_3$ repeats, a model for O-glycosylation of polypeptides with multiple acceptor sites by GalNAc-T1 can be postulated (Fig 9). a) As reported in a previous paper, the reaction is initiated by the binding of the enzyme to UDP-GalNAc [3]. b) This binding brings about conformational changes in GalNAc-T1, enabling it to bind to acceptor polypeptides. Requirement of UDP for the binding of H-ΔN32 and H-ΔN32ΔC130 with apomucin-Sepharose, as well as results from previous work is consistent with this idea [3,35]. c) GalNAc-T1 then recognizes the acceptor polypeptide *via* the protein-protein interaction, which was demonstrated by the binding of the lectin domain deletion mutant, H-ΔN32ΔC130, to apomucin-Sepharose (Fig. 4C). d) The enzyme, then, transfers GalNAc residues to the acceptor by the initial O-glycosylation activity. Following the initial O-glycosylation of the acceptor, the glycosylated acceptor would be dissociated from the enzyme. This is essential for the recycling of the enzyme and for the subsequent glycosylation of substrates. e) The glycosylated acceptor may, then, be recognized by the lectin domain of

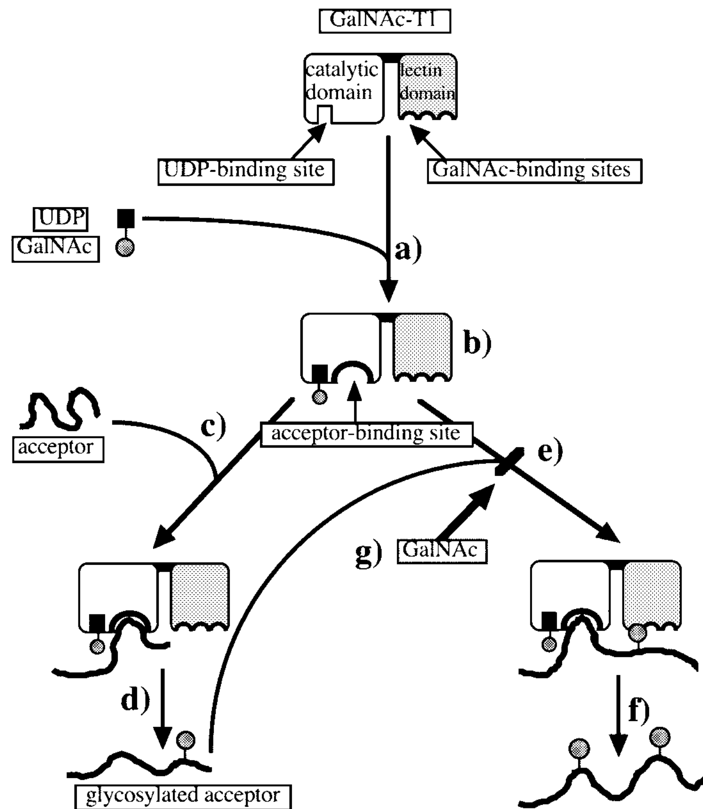


Fig. 9 A model for O-glycosylation by GalNAc-T1
a) Binding of the enzyme to UDP, b) conformational change in the catalytic domain, c) binding of acceptors, d) initial glycosylation, e) binding with GalNAc on the acceptor, f) followup glycosylation, and g) inhibition by free GalNAc. A detailed explanation of this model is included in the Discussion.

GalNAc-T1 through protein-carbohydrate interaction. Concomitant protein-protein interaction would make the binding between glycosylated acceptors and the enzyme more stable. The importance of the carbohydrate-protein interaction is demonstrated by the reduced glycosylation of apomucin by the mutants with alterations at the putative sugar-binding sites (Fig. 6A and Table 1). f) The enzyme would transfer GalNAc to the glycosylated acceptor, *i.e.* function as a follow-up enzyme. The addition of free GalNAc to the reaction mixture would interfere with binding of the enzyme to glycosylated acceptors, causing inhibition of the follow-up O-glycosylation. Consequently, it appears possible that the enzymatic activity remaining in the presence of high concentration of GalNAc represents the initial O-glycosylation activity of GalNAc-T1 (Figs. 7 and 8).

It is noteworthy that mutations causing conformational changes in the lectin domain completely ablated the enzymatic activity of GalNAc-T1, in spite of data suggesting that the lectin domain is involved in follow-up glycosylation only. Moreover, the deleterious effects on the

activity were observed by truncation of either of the three $(QXW)_3$ repeats (Fig. 1) or by mutation of the amino acid at the putative hydrophobic core of the lectin domain (Fig. 6). Follow-up glycosylation is expected to occur at potential acceptor sites located at close proximity to the initially glycosylated sites in the acceptor, since follow-up O-glycosylation can be assayed with relatively short synthetic peptides [30,31,44,46,47]. Hence, it appears reasonable to assume interaction between the catalytic and lectin domains of GalNAc-T1. This may be required to bring the two domains close together to efficiently carry out follow-up glycosylation. Consequently, mutations, which disrupt the lectin domain structure, could influence the conformation of the catalytic domain, thereby inactivating both the initial and the follow-up activities of the enzyme. The weaker binding of the C-terminal deletion mutant to UDP-hexanolamine-agarose and apomucin-Sepharose is consistent with conformational changes in the catalytic domain of the molecule (Fig. 4). By contrast, the enzymes with mutations at the putative sugar binding sites of the lectin domain could retain the native conformation of both domains. They would therefore, lose the follow-up activity, while their initial activity would be essentially unchanged. Interactions between the catalytic domain and the $(QXW)_3$ repeats have been demonstrated in the *streptomyces olvaceovirids* E-86 β-xylanase [48,49]. The interaction keeps both domains close together, thereby making it possible to hydrolyze insoluble xylan with the help of the lectin domain, in spite of the fact that this domain is not directly involved in catalysis [48–51]. Similarly, the $(QXW)_3$ repeats in GalNAc-T1 may play an auxiliary role in follow-up glycosylation in coordination with the catalytic domain. The lectin domain of GalNAc-T1 is not directly involved in the catalysis but it is essential for presenting the glycosylated acceptor polypeptides to the catalytic domain.

This study produced new information on the structure-function relationship of the GalNAc-T1 $(QXW)_3$ repeats. I have demonstrated that the amino acid residues, at potential sugar binding sites (Asp^{444}, Gly^{455}, Phe^{457}, Asn^{465}, and Gln^{466}) and at the hydrophobic core (Phe^{468}), in the α repeat, are important for the lectin function of this domain (Fig. 6). In particular, I found that Asp^{444} was an important site for sugar recognition, as evidenced by the low affinity of D444A for apomucin (Table 1) and by the decreased sensitivity of an enzyme carrying this mutation, to GalNAc inhibition (Fig. 7). Interestingly, G455Q, in which glutamine is substituted for Gly^{455}, so as to introduce the ricin-like Gal-binding site into the α repeat of GalNAc-T1, did not confer detectable Gal-binding activity to the mutant enzyme (unpublished data). Instead, this enzyme exhibited a reduced reactivity towards apomucin similar to the other mutant enzymes. This suggests that the $(QXW)_3$ repeats of GalNAc-T1 are conformationally different from those of ricin, in spite of the fact that both lectin domains are supposed to originate from a common ancestral Gal-binding protein [24,52]. The exact role of Gly^{455} in GalNAc-T1 is not clear, but the reduced activity of the mutant enzyme could be the result of a local conformational change in the α repeat. The bulky, polar side chain of glutamine, as compared to glycine, may restrict the flexibility of the polypeptide chain at position 455.

My results indicated that a predominant involvement of the α and β repeats in GalNAc recognition during glycosylation of apomucin. The α repeat is shown to have higher affinity for GalNAc than the β repeat, while no measurable GalNAc-binding activity was observed for

the γ repeat. The sugar binding activity of all three repeat units correlated well with the location of the putative sugar-binding residues, as deduced from sequence comparisons of the GalNAc-T1 lectin domains to the ricin B-chain. Of the three repeat units in GalNAc-T1, the putative sugar-binding amino acid residues in the α repeat, Asp444, Phe457, Asn465, and Gln466, show the highest degree of conservation (with the ricin B-chain) [24,28]. The β repeat unit is second in terms of conservation of putative sugar-binding sites. This repeat unit contains Asp484, Asn503, and Gln504, corresponding to Asp444, Asn465, and Gln466 in the α repeat, but there is a substitution (in the β repeat) at the site corresponding to Phe457 in the a repeat. Phe457 corresponding to Trp37 in ricin B-chain, a residue that makes a stacking interaction with galactose [25,26]. The γ repeat also does not contain an aromatic residue at the site corresponding to Phe457 in the α repeat. In addition, the γ repeat has an alteration at the site corresponding to Asn465 in the α repeat. In ricin, the corresponding asparagines is involved in direct hydrogen binding with galactose [25,26]. These considerations suggest a ranking of the sugar-binding activities of the three repeat units in the following order: $\alpha > \beta > \gamma$. Although no measurable sugar-binding activity was detected for the γ repeat in this study, more sensitive methods may yet reveal some capacity for GalNAc-binding in this unit. It is also possible that the γ repeat is involved in binding of sugars other than GalNAc.

I found that two of the repeat units in the lectin domain of GalNAc-T1 are functional in recognizing GalNAc. This indicates that the lectin domain of GalNAc-T1 is capable of at least divalent interactions with glycosylated ligands. Lectins often contain contiguous modules with multimeric complexes of two or more non-covalently linked modules [53,54]. Although the binding affinity of an individual module may be low, strong binding can still be achieved, through the interaction of multiple modules with ligand. Mucins are putative *in vivo* ligands for the GalNAc transferases that are characterized by the presence of numerous carbohydrate chains. Hence, it is likely that the multitude of GalNAc residues transferred to the core polypeptide during the glycosylation of mucins make multivalent interactions with the GalNAc transferase lectin domains possible. Although the binding of the individual repeat units in the lectin domain may be weak, as represented by the requirement of a high concentration of GalNAc to inhibit the apomucin glycosylation, the overall binding affinity between the GalNAc transferase enzyme and the mucin acceptor is expected to be high enough to tether partially glycosylated mucin molecules and carry out follow-up glycosylation.

The results presented in this report demonstrate the functional importance of the $(QXW)_3$ repeats. The data show that this domain in fact is a functioning lectin on GalNAc-T1, at least under the *in vitro* conditions tested. This led to the conclusion that GalNAc-T1, in addition to glycosylating unmodified polypeptides *in vivo*, may also specifically recognize and glycosylate partially glycosylated acceptors, *i.e.* it may also function as a "follow-up" enzyme. The results generated in this study also provide new information about the substrate specificity of GalNAc-T1. Although acceptor specificities of the GalNAc-transferases, has been studied quite extensively, a specific consensus sequence has not been identified for any enzyme [3,55]. Instead, GalNAc-T1 is regarded as a multi-substrate enzyme with broad acceptor specificity. Still, the sequence context surrounding glycosylated serine and threonine residues is known to be impor-

tant. Certain residues, at certain positions, in the vicinity of the acceptor amino acid are known to promote glycosylation, while other residues are known to be refractory. Based on this information, computer-based algorithms have been generated capable of predicting O-glycosylation, with an accuracy of approximately 80% [55,56]. The difficulty in exactly defining the substrate specificity may in part be ascribed to the multiple functions of GalNAc-transferases, as demonstrated by the two activities of GalNAc-T1. Discrimination between these two activities could lead to a more in-depth understanding of the molecular mechanism of this enzyme. By introducing multiple mutations in the $(QXW)_3$ repeats, I are currently preparing novel recombinant enzymes that retain only the initial O-glycosylation activity. More detailed structural information on the lectin domain would open the possibility of designing novel enzymes with the altered (follow-up) substrate specificity.

Acknowledgement

The author wishes to thank Dr. Akira Kurosaka, Department of Biotechnology, Faculty of Engineering, Kyoto Sangyo University, for helpful advice on the study and critical reading of the manuscript. I would like to thank Dr. Åke P Elhammer, Pharmacia Co. for kind advice on the study. This work was in part supported by Sasakawa Scientific Research Grant.

Footnotes

Abbrevations: EDTA, ethylenediaminetetraacetic acid; GalNAc, N-acetylgalactosamine; GalNAc-transferase, UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase; GT1, glycosyltransferase 1; HA, haemagglutinin; MOI, multiplicity of infection; PCR, polymerase chain reaction; PVDF, polyvinilidene difluoride; UDP, uridine diphosphate; SDS-PAGE, SDS-polyacrylamide gel electrophoresis.

References

[1] Clausen, H., and Bennett, E. P. (1996) *Glycobiology* **6**, 635–646

[2] Marth, J. D. (1996) *Glycobiology* **6**, 701–705

[3] Elhammer, Å. P., Kézdy, F. J., and Kurosaka, A. (1999) *Glycoconj. J.* **16**, 171–180

[4] Homa, F. L., Hollander, T., Lehman, D. J., Thomsen, D. R., and Elhammer, A. P. (1993) *J. Biol. Chem.* **268**, 12609–12616

[5] White, T., Bennett, E. P., Takio, K., Sorensen, T., Bonding, N., and Clausen, H. (1995) *J. Biol. Chem.* **270**, 24156–24165

[6] Bennett, E. P., Hassan, H., and Clausen, H. (1996) *J. Biol. Chem.* **271**, 17006–17012

[7] Hagen, F. K., TenHagen, K. G., Beres, T. M., Balys, M. M., VanWuyckhuyse, B. C., and Tabak, L. A. (1997) *J. Biol. Chem.* **272**, 13843–13848

[8] Bennett, E. P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., TaylorPapadimitriou, J., Hollingsworth, M. A., Merkx, G., vanKessel, A. G., Eiberg, H., Steffensen, R., and Clausen, H. (1998) *J. Biol. Chem.* **273**, 30472–30481

[9] Ten Hagen, K. G., Hagen, F. K., Balys, M. M., Beres, T. M., Van Wuyckhuyse, B., and Tabak, L.

A. (1998) *J. Biol. Chem.* **273**, 27749–27754
[10] 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., and Tabak, L. A. (1999) *J. Biol. Chem.* **274**, 27867–27874
[11] Bennett, E. P., Hassan, H., Hollingsworth, M. A., and Clausen, H. (1999) *FEBS Lett.* **460**, 226–230
[12] White, K. E., Lorenz, B., Evans, W. E., Meitinger, T., Strom, T. M., and Econs, M. J. (2000) *Gene* **246**, 347–356
[13] Toba, S., Tenno, M., Konishi, M., Mikami, T., Itoh, N., and Kurosaka, A. (2000) *Biochim. Biophys. Acta* **1493**, 264–268
[14] Hagen, K. G., Bedi, G. S., Tetaert, D., Kingsley, P. D., Hagen, F. K., Balys, M. M., Beres, T. M., Degand, P., and Tabak, L. A. (2001) *J. Biol. Chem.* **276**, 17395–17404
[15] Merla, G., Ucla, C., Guipponi, M., and Reymond, A. (2002) *Hum. Genet.* **110**, 429–438
[16] Guo, J. M., Zhang, Y., Cheng, L., Iwasaki, H., Wang, H., Kubota, T., Tachibana, K., and Narimatsu, H. (2002) *FEBS Lett.* **524**, 211–218
[17] Hagen, F. K., and Nehrke, K. (1998) *J. Biol. Chem.* **273**, 8268–8277
[18] Ten Hagen, K. G., and Tran, D. T. (2002) *J. Biol. Chem.* **277**, 22616–22622
[19] Schwientek, T., Bennett, E. P., Flores, C., Thacker, J., Hollmann, M., Reis, C. A., Behrens, J., Mandel, U., Keck, B., Schafer, M. A., Haselmann, K., Zubarev, R., Roepstorff, P., Burchell, J. M., Taylor-Papadimitriou, J., Hollingsworth, M. A., and Clausen, H. (2002) *J. Biol. Chem.* **277**, 22623–22638
[20] Paulson, J. C., and Colley, K. J. (1989) *J. Biol. Chem.* **264**, 17615–17618
[21] Hagen, F. K., Hazes, B., Raffo, R., deSa, D., and Tabak, L. A. (1999) *J. Biol. Chem.* **274**, 6797–6803
[22] Wiggins, C. A., and Munro, S. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 7945–7950
[23] Wragg, S., Hagen, F. K., and Tabak, L. A. (1997) *Biochem. J.* **328**, 193–197
[24] Hazes, B. (1996) *Protein Sci.* **5**, 1490–1501
[25] Rutenber, E., Ready, M., and Robertus, J. D. (1987) *Nature* **326**, 624–626
[26] Rutenber, E., and Robertus, J. D. (1991) *Proteins* **10**, 260–269
[27] Tahirov, T. H., Lu, T. H., Liaw, Y. C., Chen, Y. L., and Lin, J. Y. (1995) *J. Mol. Biol.* **250**, 354–367
[28] Imberty, A., Piller, V., Piller, F., and Breton, C. (1997) *Protein Eng.* **10**, 1353–1356
[29] Breton, C., and Imberty, A. (1999) *Curr. Opin. Struct. Biol.* **9**, 563–571
[30] Hassan, H., Reis, C. A., Bennett, E. P., Mirgorodskaya, E., Roepstorff, P., Hollingsworth, M. A., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (2000) *J. Biol. Chem.* **275**, 38197–38205
[31] Hanisch, F. G., Reis, C. A., Clausen, H., and Paulsen, H. (2001) *Glycobiology* **11**, 731–740
[32] Hagen, F. K., Gregoire, C. A., and Tabak, L. A. (1995) *Glycoconj. J.* **12**, 901–909
[33] Tenno, M., Toba, S., Kézdy, F. J., Elhammer, Å. P., and Kurosaka, A. (2002) *Eur. J. Biochem.* **269**, 4308–4316
[34] Tenno, M., Saeki, A., Kézdy, F. J., Elhammer, Å. P., and Kurosaka, A. (2002) *J. Biol. Chem.* **277**, 47088–47096
[35] Elhammer, Å. P., and Kornfeld, S. (1986) *J. Biol. Chem.* **261**, 5249–5255
[36] Bhavanandan, V. P., and Hegarty, J. D. (1987) *J. Biol. Chem.* **262**, 5913–5917
[37] Bhargava, A. K., Woitach, J. T., Davidson, E. A., and Bhavanandan, V. P. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6798–6802
[38] Jiang, W., Woitach, J. T., Gupta, D., and Bhavanandan, V. P. (1998) *Biochem. Biophys. Res. Commun.* **251**, 550–556
[39] Jiang, W., Gupta, D., Gallagher, D., Davis, S., and Bhavanandan, V. P. (2000) *Eur. J. Biochem.*

267, 2208–2217

[40] Tenno, M., Kézdy, F. J., Elhammer, Å. P., and Kurosaka, A. (2002) *Biochem. Biophys. Res. Commun.* **298**, 755–759

[41] Wang, Y., Abernethy, J. L., Eckhardt, A. E., and Hill, R. L. (1992) *J. Biol. Chem.* **267**, 12709–12716

[42] Hagen, F. K., Van Wuyckhuyse, B., and Tabak, L. A. (1993) *J. Biol. Chem.* **268**, 18960–18965

[43] Kirnarsky, L., Nomoto, M., Ikematsu, Y., Hassan, H., Bennett, E. P., Cerny, R. L., Clausen, H., Hollingsworth, M. A., and Sherman, S. (1998) *Biochemistry* **37**, 12811–12817

[44] Hanisch, F. G., Muller, S., Hassan, H., Clausen, H., Zachara, N., Gooley, A. A., Paulsen, H., Alving, K., and Peter-Katalinic, J. (1999) *J. Biol. Chem.* **274**, 9946–9954

[45] Hanisch, F. G., and Muller, S. (2000) *Glycobiology* **10**, 439–449

[46] Kato, K., Takeuchi, H., Kanoh, A., Mandel, U., Hassan, H., Clausen, H., and Irimura, T. (2001) *Glycobiology* **11**, 821–829

[47] Kato, K., Takeuchi, H., Miyahara, N., Kanoh, A., Hassan, H., Clausen, H., and Irimura, T. (2001) *Biochem. Biophys. Res. Commun.* **287**, 110–115

[48] Fujimoto, Z., Kuno, A., Kaneko, S., Kobayashi, H., Kusakabe, I., and Mizuno, H. (2002) *J. Mol. Biol.* **316**, 65–78

[49] Fujimoto, Z., Kuno, A., Kaneko, S., Yoshida, S., Kobayashi, H., Kusakabe, I., and Mizuno, H. (2000) *J. Mol. Biol.* **300**, 575–585

[50] Kuno, A., Shimizu, D., Kaneko, S., Koyama, Y., Yoshida, S., Kobayashi, H., Hayashi, K., Taira, K., and Kusakabe, I. (1998) *J. Ferment. Bioeng.* **86**, 434–439

[51] Black, G. W., Hazlewood, G. P., Millward-Sadler, S. J., Laurie, J. I., and Gilbert, H. J. (1995) *Biochem. J.* **307**, 191–195

[52] Hirabayashi, J., Dutta, S. K., and Kasai, K. (1998) *J. Biol. Chem.* **273**, 14450–14460

[53] Lis, H., and Sharon, N. (1998) *Chem. Rev.* **98**, 637–674

[54] Weis, W. I., and Drickamer, K. (1996) *Annu. Rev. Biochem.* **65**, 441–473

[55] Elhammer, Å. P., Poorman, R. A., Brown, E., Maggiora, L. L., Hoogerheide, J. G., and Kézdy, F. J. (1993) *J. Biol. Chem.* **268**, 10029–10038

[56] Hansen, J. E., Lund, O., Tolstrup, N., Gooley, A. A., Williams, K. L., and Brunak, S. (1998) *Glycoconj. J.* **15**, 115–130

UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase 1 (GalNAc-T1) の構造活性相関
――GalNA-T1 のレクチン様ドメインの機能解析――

天　野　麻　理

要　旨

糖タンパク質のムチン型糖鎖生合成開始反応を触媒する UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase (GalNAc-T) は，タンパク質中のセリンおよびトレオニン残基のヒドロキシル基に GalNAc を転移する．GalNAc-T の C 末端部には，植物レクチンのリシンと相同性が高く，α, β, γ の 3 つのタンデムリピートからなるレクチン様ドメインが存在する．GalNAc-T の主要なアイソザイムの 1 つである GalNAc-T1 のレクチン様ドメインが酵素活性に関与しないという報告がなされているが，筆者はこのドメインに種々の変異を導入することで，GalNAc-T1 のレクチン様ドメインが酵素活性に重要な役割を果たしていることを見いだしたので報告する．筆者は，まずレクチン様ドメインを欠失した変異体を作製した．この変異体が完全に失活したことから，このドメインの酵素活性への関与が示唆された．さらに筆者は，GalNAc-T1 が高濃度の GalNAc による阻害反応で区別できる 2 つの活性を持つことを見いだした．1 つの糖受容部位しか持たないアクセプターへの転移反応によって測定される酵素の initial 活性は，GalNAc の阻害を受けなかった．それに対して，アポムチンの様に，多くの糖受容部位を持ち，完全な糖付加には酵素の initial 活性と follow-up 活性の両方が必要であるアクセプターへの糖転移は，高濃度の GalNAc の添加により部分的に阻害された．この結果は，高濃度の GalNAc が，follow-up 反応を選択的に阻害し，かつ follow-up 反応では，initial 反応により転移された GalNAc が酵素により認識される可能性を示している．このことは，レクチン様ドメインに変異を導入したときに，アポムチンに対する活性が選択的に低下したことからも推測された．筆者はさらに，ドメイン中の3つのリピートが α ＞ β ＞ γ の順の強さで糖結合活性を持つこと，および α リピート中の Asp444 がレクチン様活性に最も重要な残基であることも見いだした．以上の結果は，GalNAc-T1 のレクチン様ドメインの GalNAc 認識機構が follow-up 活性によるアクセプターへの GalNAc の付加反応に必須であることを示している．

キーワード： ムチン型 O-グリコシレーション，GalNAc 転移酵素，レクチン様ドメイン，構造活性相関，部位特異的変異