

The lectin domains of polypeptide GalNAc-transferases exhibit carbohydrate-binding specificity for GalNAc: lectin binding to GalNAc-glycopeptide substrates is required for high density GalNAc-*O*-glycosylation

Hans H. Wandall^{2,3}, Fernando Irazoqui^{3,4}, Mads Agervig Tarp³, Eric P. Bennett³, Ulla Mandel³, Hideyuki Takeuchi⁵, Kentaro Kato³, Tatsuro Irimura⁵, Ganesh Suryanarayanan⁶, Michael A. Hollingsworth⁶, and Henrik Clausen^{1,3}

³Department of Medical Biochemistry and Genetics and the Department of Oral Diagnostics, Dental School, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark;

⁴CIQUIBIC-CONICET/Department of Biological Chemistry, Faculty of Chemical Sciences, National University of Cordoba, 5000 Cordoba, Argentina; ⁵Laboratory of Cancer Biology and Molecular Immunology, Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan; and ⁶Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, Omaha, NE 68198

Received on August 9, 2006; revised on November 21, 2006; accepted on December 19, 2006

Initiation of mucin-type *O*-glycosylation is controlled by a large family of UDP GalNAc:polypeptide *N*-acetylgalactosaminyltransferases (GalNAc-transferases). Most GalNAc-transferases contain a ricin-like lectin domain in the C-terminal end, which may confer GalNAc-glycopeptide substrate specificity to the enzyme. We have previously shown that the lectin domain of GalNAc-T4 modulates its substrate specificity to enable unique GalNAc-glycopeptide specificities and that this effect is selectively inhibitable by GalNAc; however, direct evidence of carbohydrate binding of GalNAc-transferase lectins has not been previously presented. Here, we report the direct carbohydrate binding of two GalNAc-transferase lectin domains, GalNAc-T4 and GalNAc-T2, representing isoforms reported to have distinct glycopeptide activity (GalNAc-T4) and isoforms without apparent distinct GalNAc-glycopeptide specificity (GalNAc-T2). Both lectins exhibited specificity for binding of free GalNAc. Kinetic and time-course analysis of GalNAc-T2 demonstrated that the lectin domain did not affect transfer to initial glycosylation sites, but selectively modulated velocity of transfer to subsequent sites and affected the number of acceptor sites utilized. The results suggest that GalNAc-transferase lectins serve to modulate the kinetic properties of the enzymes in the late stages of the initiation process of *O*-glycosylation to accomplish dense or complete *O*-glycan occupancy.

Key words: GalNAc/transferases/lectins/glycans/mucins

Introduction

A large homologous family of uridine diphosphate (UDP)-*N*-acetyl- α -*D*-galactosamine (GalNAc):polypeptide *N*-acetylgalactosaminyltransferases (GalNAc-transferases, EC 2.4.1.41) initiate mucin-type *O*-glycosylation by transferring GalNAc to the hydroxyl group of serine and threonine residues (GalNAc α 1-*O*-Ser/Thr). Human and rodent GalNAc-transferase families are predicted to include up to 20 distinct isoforms, and to date 16 of these have been cloned and characterized (Clausen and Bennett 1996; Hassan, Bennett, et al. 2000; Ten Hagen et al. 2003). The GalNAc-transferase family is conserved in evolution, and distinct subfamilies of orthologous isoforms with conserved kinetic properties have been identified among vertebrates and invertebrates (Bennett, Hassan, Mandel, et al. 1999; Schwientek et al. 2002; Stwora-Wojczyk et al. 2004). The GalNAc-transferase isoforms have different kinetic properties and cell and tissue expression patterns, suggesting that they serve different and nonredundant functions (Clausen and Bennett 1996; Hassan, Bennett, et al. 2000; Ten Hagen et al. 2003). Although ablation of several GalNAc-transferase isoforms in mice have not demonstrated a phenotype (Hennet et al. 1995; Ten Hagen et al. 2003), the finding that impairment of a single isoform in *Drosophila melanogaster* disrupts development (Ten Hagen and Tran 2002; Schwientek et al. 2002), and that the human GalNAc-transferase isoform, GalNAc-T3, is implicated in the disease familial tumoral calcinosis (Topaz et al. 2004) demonstrate the nonredundant function of some GalNAc-transferase isoforms.

Our understanding of the functions of these enzymes is largely based on *in vitro* analysis of their acceptor substrate specificities with short peptides and glycopeptides. Studies with GalNAc-glycopeptides (GalNAc residues attached at some but not all potential acceptor sites in peptide sequence) have demonstrated that a number of GalNAc-transferase isoforms exhibit selective substrate specificities for these glycopeptides, which may indicate that *O*-glycosylation involves several successive biosynthetic steps catalyzed by different GalNAc-transferase isoforms (Bennett et al. 1998; Bennett, Hassan, Hollingsworth, et al. 1999; Ten Hagen et al. 1999, Pratt et al. 2004). The GalNAc-glycopeptide substrate specificity exhibited by some isoforms has been associated with a ricin-like lectin domain found in the C-terminal region of most GalNAc-transferase isoforms (Hassan, Reis, et al. 2000). This distinct lectin-like domain was originally identified by Hazes et al. (1996) and Imberty et al. (1997), and recent studies have demonstrated that the catalytic and lectin domains of

¹To whom correspondence should be addressed; e-mail: hc@imbg.ku.dk

²Present address: Division of Hematology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115.

GalNAc-transferases fold into separate domains (Fritz et al. 2004, 2006; Kubota et al. 2006). In agreement with predictions from studies of the role of the lectin domain in directing GalNAc-glycopeptide substrate specificities (Hassan, Bennett et al. 2000; Hassan, Reis et al. 2000) Fritz and colleagues were able to model a glycopeptide into the catalytic pocket and show potential interactions of the GalNAc residues with the carbohydrate-binding sites of the lectin domain. To date, however, direct experimental evidence in support of actual carbohydrate-binding activity and specificity of the GalNAc-transferase lectin domains have not been demonstrated.

Hagen et al. (1999) initially investigated the function of the lectin domain of GalNAc-T1, and found that selective mutational disruption of the lectin domain of GalNAc-T1 did not severely affect the catalytic function of the enzyme with peptide substrates. Subsequently, we found that the GalNAc-T4 isoform exhibited a unique GalNAc-glycopeptide substrate specificity (Bennett et al. 1998). Mutational analysis of GalNAc-T4 demonstrated that a mutation in the α -repeat of the C-terminal lectin domain of GalNAc-T4 selectively inactivated the GalNAc-glycopeptide catalytic function of the enzyme, whereas activity with unglycosylated peptides was unaffected, in accordance with the original study of GalNAc-T1 (Hassan, Reis, et al. 2000). Furthermore, inhibition studies with monosaccharides identified free GalNAc as a selective inhibitor of the GalNAc-glycopeptide specificity of GalNAc-T4. These results suggest that the GalNAc-T4 lectin domain modulates the function of the enzyme through interaction with the GalNAc residues of the glycopeptide substrate. Similar studies subsequently performed with GalNAc-T1 and -T2 have suggested that the lectin domains of these may also be important for the glycosylation of partially GalNAc-glycosylated peptide substrates (Tenno, Saeki, et al. 2002; Fritz et al. 2006). Some GalNAc-transferases, including GalNAc-T7 and -T10, have been found to display apparent strict GalNAc-glycopeptide specificities (Ten Hagen et al. 1999; Bennett, Hassan, Hollingsworth, et al. 1999; Ten Hagen et al. 2001), and it is likely that the lectin domains of these enzymes play similar roles.

Nonetheless, the studies to date have not demonstrated direct carbohydrate-binding properties of the GalNAc-transferase lectin domain. To determine the potential carbohydrate-binding properties of GalNAc-transferase lectin domains, we have now developed a carbohydrate-binding assay and investigated the binding specificity of the lectin domain of GalNAc-T4, which has a selective GalNAc-glycopeptide substrate specificity, and GalNAc-T2, which represents enzymes without apparent unique glycopeptide specificity. Finding that both isolated lectin domains bind GalNAc, we investigated the importance of a functional lectin domain in GalNAc-T2 representing enzymes without apparent glycopeptide specificity. The results demonstrate that the catalytic function of GalNAc-T2, similar to GalNAc-T4, is affected by mutations in the lectin domain and the mechanism involves selective GalNAc-glycopeptide substrate specificity. The results suggest that GalNAc-transferase lectins serve a general function on all isoforms: to improve kinetic properties with partially GalNAc-glycosylated substrates, presumably by allowing both the catalytic and the lectin domain to interact with the substrate for improved kinetic properties.

Results

Demonstration of direct binding of truncated GalNAc-T2 and -T4 lectins to GalNAc-glycosylated mucin peptides

Several different truncated lectin constructs were analyzed for expression in insect cells, and constructs yielding significant protein expression were found to include residues 405–578 and 432–571 for GalNAc-T4 and -T2, respectively, with the N-terminal sequence containing His- and T7-tags (see Figure 1A for design of expression constructs). The expression levels were significantly lower than that of the corresponding soluble secreted enzyme constructs (not shown). An enzyme-linked lectin assay (ELLA) was developed to investigate the carbohydrate-binding properties of truncated GalNAc-transferase lectin domains. Biotinylated purified recombinant lectin domains of GalNAc-T2 and -T4 selectively bound GalNAc-glycosylated MUC1 peptides (Figure 1B and C), whereas no binding was observed to the corresponding nonglycosylated peptide. With the same assay design, we were able to show binding of the unlabeled lectin domains using antibodies to the enzymes or the N-terminal His- and T7-tags. Biotinylated anti-Tn lectin from *Helix pomatia* used as control showed the same binding pattern, whereas an anti-MUC1 antibody HMFG2 reacted with both the unglycosylated and the GalNAc-glycosylated peptide (Fig. 1F).

Subsequently, biotinylated soluble constructs of GalNAc-T2 and GalNAc-T4 were tested. The biotinylated enzymes exhibited the same binding specificity as the truncated biotinylated and nonbiotinylated lectin domains, albeit with stronger binding to GalNAc-glycosylated MUC1 peptides (Figure 1B and C). In order to confirm that the carbohydrate-binding activity demonstrated with the GalNAc-T2 and -T4 enzymes were derived from the lectin domains, several mutant constructs were tested. Previously, we demonstrated that the lectin-mediated GalNAc-glycopeptide substrate specificity of GalNAc-T4 was impaired by mutation of a critical residue (D459) in the CLD motif of the α -repeat. A single amino acid substitution was therefore introduced in the α -repeat changing the aspartic acid to a histidine in the lectin domain of GalNAc-T2 and -T4. As expected, mutant GalNAc-T2^{458H} and GalNAc-T4^{459H} did not exhibit binding to GalNAc-MUC1. In contrast, mutation of the CLD motif in the γ -repeat of GalNAc-T2 (GalNAc-T2^{541A}) did not impair binding to the GalNAc-MUC1 glycopeptide, suggesting that the γ -repeat is not involved in carbohydrate binding. To further exclude the possibility that the GalNAc-binding property of the biotinylated enzymes was partly due to the catalytic domain, we also tested the effect of mutations in the catalytic domain. The catalytically inactive biotinylated GalNAc-T2^{224H}, in which the DXH nucleotide-binding motif was mutated, showed the same binding to glycosylated MUC1 glycopeptides as the biotinylated wild type enzyme and the truncated lectins. The binding of the biotinylated enzymes was found to be insensitive to addition of EDTA, UDP, and MnCl₂. Surface plasmon resonance experiments confirmed the binding of the isolated nonbiotinylated lectin domain of GalNAc-T2 to GalNAc-MUC1. The binding affinity was concentration dependent, with a dissociation constant measured to 3.75×10^{-5} M ($K_{\text{on}} = 0.43 \text{ M}^{-1} \text{ s}^{-1}$, $K_{\text{off}} = 1.62 \times 10^{-5} \text{ s}^{-1}$) (Fig. 2). No binding was observed

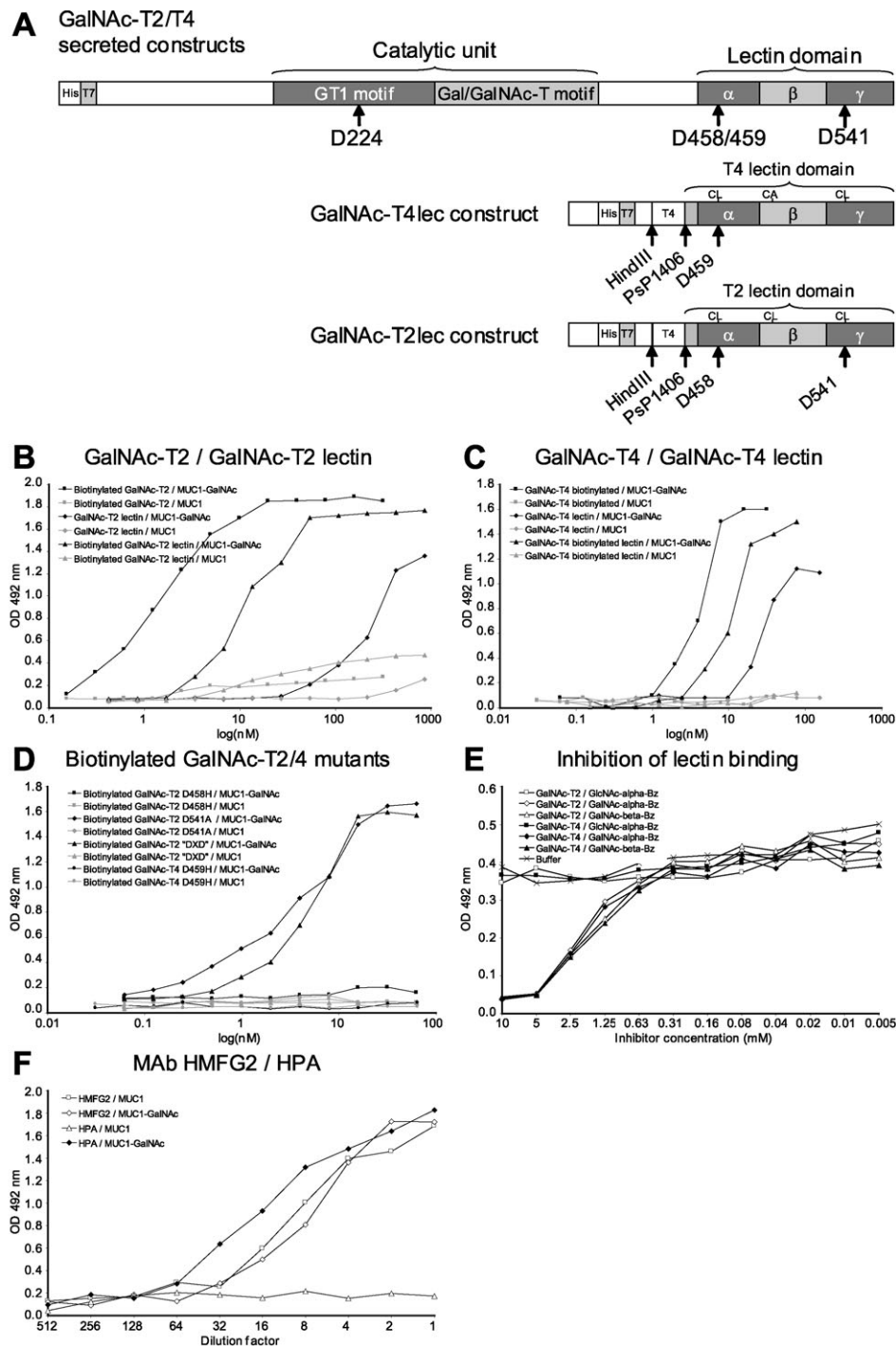


Fig. 1. ELLA analysis of GalNAc-transferase lectin binding to GalNAc-glycopeptides. (A) The localization of mutations in GalNAc-T2^{458H}, GalNAc-T2^{541A}, GalNAc-T2^{224H}, GalNAc-T4^{459H}, and GalNAc-T4lec^{459H} is shown (arrows). (B) Nonbiotinylated GalNAc-T2lec, biotinylated GalNAc-T2lec, and biotinylated GalNAc-T2 bind GalNAc glycosylated MUC1 peptide. Polystyrene microtiter plates coated with 60-merMUC1 and GalNAc-60merMUC1 were incubated with decreasing concentrations of non-biotinylated GalNAc-T2lec, biotinylated GalNAc-T2lec, and biotinylated GalNAc-T2. (C), Nonbiotinylated GalNAc-T4lec, biotinylated GalNAc-T4lec, and biotinylated GalNAc-T4 bind GalNAc glycosylated MUC1 peptide. (D) The α -repeat of the lectin domain is essential for binding of the GalNAc-T2 and -T4 lectin to GalNAc-60merMUC1. Microtiter plates coated as described above were incubated with biotinylated transferases in which either the catalytic function or the putative lectin function was disrupted by site directed mutagenesis: GalNAc-T2^{224H} (DXE mutant), GalNAc-T2^{458H} (α -repeat mutant), GalNAc-T2^{541A} (γ -repeat mutant), GalNAc-T4^{459H} (α -repeat mutant) and GalNAc-T4 wildtype. (E) Inhibition of lectin binding to GalNAc-60merMUC1 was performed by coincubated of biotinylated GalNAc-T2 and -T4 with increasing concentrations of Bn- α -GalNAc and Bn- α -GlcNAc, demonstrating selective inhibition of the interaction between GalNAc-T2 and GalNAc-60merMUC1. Designations as indicated. (F) Detection of 60merMUC1 and GalNAc-60merMUC1 with anti-MUC1 MAb HMFG2 and biotinylated lectin HPA.

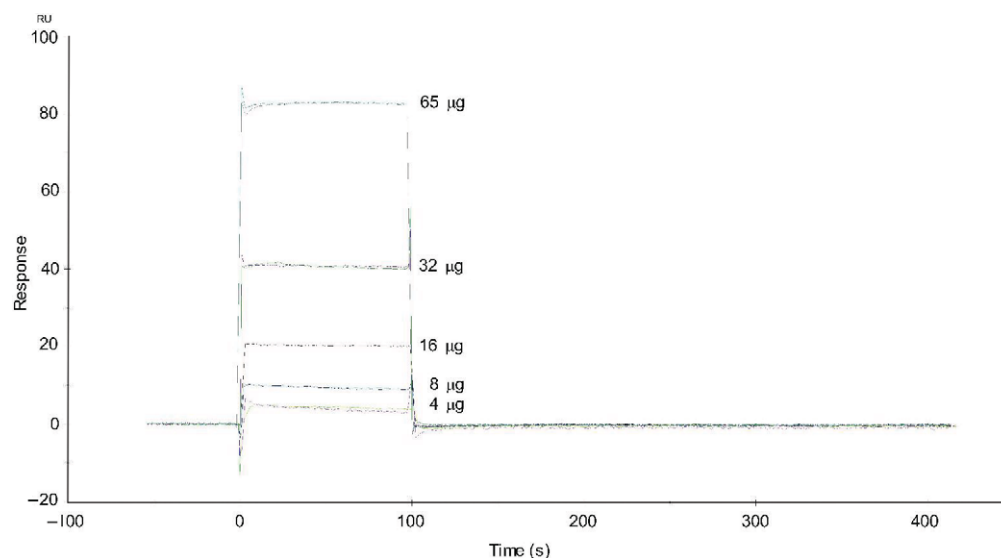


Fig. 2. Evaluation of GalNAc-MUC1 interaction with the isolated lectin domain of GalNAc-T2lec by BIAcore analysis. GalNAc-MUC1 peptide was covalently immobilized onto a CM3 sensor flow cell. Various concentrations of non-biotinylated GalNAc-T2lec (as indicated) were injected over the GalNAc-MUC1 flow cell. Results from parallel injections of each concentration over a control flow cell with immobilized Representative sensograms for GalNAc-T2lec are shown. The binding affinity was concentration dependent with a dissociation constant measured to 3.75×10^{-5} M ($K_{on} = 0.43 \text{ M}^{-1} \text{ s}^{-1}$, $K_{off} = 1.62 \times 10^{-5} \text{ s}^{-1}$). No binding was observed to the non-glycosylated MUC1 peptide.

Table I. Inhibition of biotinylated GalNAc-transferase lectins by saccharides and related structures

Inhibitor	Protein				
	T2lec	T2	T4lec	T4	HPA
	<i>mM (IC50)^a</i>				
Glc	≥100	>100	>100	>100	>100
GlcNAc	>100	>100	>100	>100	8
Bn-α-GlcNAc	>100	>100	>100	>100	2
Bn-β-GlcNAc	>100	>100	>100	>100	>20
Gal	>100	>100	>100	50	>100
Me-α-Gal	>100	>100	>100	50	>100
Me-β-Gal	>100	>100	>100	40	>100
GalNAc	37	15	5	1	2
Bn-α-GalNAc	20	15	5	1	0.5
Ph-α-GalNAc	15	5	5	1	1
oNP-α-GalNAc	12	7	5	0.5	1.5
oNP-β-GalNAc	12	10	8	1	>20
pNP-α-GalNAc	10	10	10	1	2
pNP-β-GalNAc	10	10	8	0.8	>20
Lactose	>100	>100	>100	>100	Not determined
UDP-α-GalNAc	>100	65	50	30	2
UDP	>100	>100	>100	>100	>100
EDTA	>10	>10	>10	>10	>10

^aConcentration or higher (>) required for 50% inhibition.

^bA different experimental setup with fixed concentrations of lectins resulted in a significant difference in IC50 when compared to Fig. 1, panel E, where limited dilutions of lectins were used.

to the nonglycosylated peptide. No binding was observed with the native GalNAc-T2 enzyme to GalNAc-MUC1 in binding buffer at pH 7.4 (not shown).

Binding specificity of GalNAc-transferase lectins

To define the specificity of the carbohydrate-binding properties of biotinylated GalNAc-T2 and -T4, a panel of monosaccharides and aglycon derivatives were tested as inhibitors. Binding of both biotinylated GalNAc-T2 and -T4 lectin domains to GalNAc-MUC1 were inhibited by 100 mM free GalNAc with IC50 values of 15–37 and 1–5 mM, respectively. Weak inhibition of biotinylated GalNAc-T4 was observed with high concentrations of galactose, whereas other monosaccharides Glc, GlcNAc, and Man showed no inhibition for biotinylated GalNAc-T2, GalNAc-T4 or their lectins (Table I). Benzyl or phenyl α/β-GalNAc compounds showed lower IC50 values for biotinylated GalNAc-T4 compared with biotinylated GalNAc-T2. In Figure 1E, the binding of biotinylated GalNAc-T2 and -T4 were inhibited with benzyl inhibitors using limiting dilutions of enzyme lectins. Both GalNAc-α-benzyl and GalNAc-β-benzyl showed similar inhibitory effects, demonstrating that biotinylated GalNAc-T2 and -T4 lectins do not exhibit specificity for the anomeric configuration of GalNAc.

Analysis of a panel of GalNAc-glycopeptides derived from the tandem repeat sequence of MUC1, MUC2, and MUC7, which were glycosylated in vitro with different densities and patterns of GalNAc residues, did not reveal significantly different binding properties, indicating that the lectins do not have selective specificities for peptide sequences, or the density or pattern of GalNAc glycosylation (not shown).

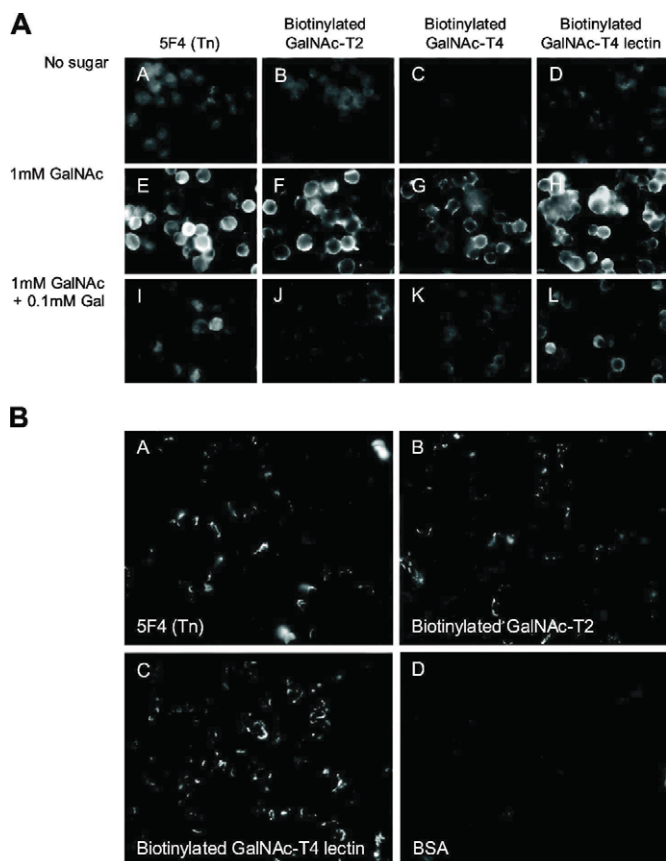


Fig. 3. Biotinylated GalNAc-transferase lectins exhibit similar immunocytochemistry staining pattern as anti-Tn monoclonal antibodies. **(A)** CHO-IdID cells stably transfected with the full coding human MUC1 gene were grown with and without GalNAc and Gal/GalNAc fixed and stained with biotinylated GalNAc-T2, GalNAc-T4, and GalNAc-T4lec, and α -Tn monoclonal antibody 5F4 recognizing GalNAc α S/T. **(B)** Salivary glands were fixed and sections labeled with biotinylated GalNAc-T4lec, GalNAc-T2, and 5F4.

Binding of biotinylated GalNAc-transferases and GalNAc-T4 lectin to cells and tissue

To further examine the specificity of the GalNAc-transferase lectins, we analyzed the lectin staining patterns of biotinylated GalNAc-T2 and -T4 enzymes and the isolated biotinylated lectins with cultured cells and tissues (Figure 3). Chinese hamster ovary (CHO) IdID cells lack the UDP-Gal/UDP-GalNAc 4-epimerase and are deficient in GalNAc O-glycosylation and galactosylation in the absence of exogenous addition of GalNAc and Gal, respectively (Kingsley et al. 1986). CHO IdID cells stably transfected with the full-length human MUC1 cDNA (CHO IdID-MUC1) were grown with and without GalNAc and Gal/GalNAc. CHO IdID-MUC1 cells grown in the presence of GalNAc stain strongly with anti-Tn (GalNAc- α -Ser/Thr) monoclonal antibody 5F4 (Figure 3A). Without exogenous addition of GalNAc, the cells do not stain with anti-Tn antibody (5F4). When cells are grown with both Gal and GalNAc no or very weak anti-Tn staining is observed because GalNAc residues are extended to form the sialyl-T structures (NeuAc α 2-3Gal β 1-3[NeuAc α 2-6]GalNAc- α -Ser/Thr), but instead cells stain strongly with anti-T antibodies after pretreatment with neuraminidase (not shown) (Sorensen et al. 2005). As shown in Figure 3A, the biotinylated GalNAc-transferases, the biotinylated GalNAc-T4 lectin, and the anti-Tn monoclonal antibody showed the same staining pattern on CHO IdID-MUC1 cells grown in the presence of GalNAc. No anti-Tn reactivity was observed on cells grown without GalNAc, and only minimal staining was observed when cells were grown with both GalNAc and Gal. Further immunohistological studies of human salivary glands showed the same correlation of staining of enzymes, lectin, and the anti-Tn antibody (Figure 3B). Human salivary glands stain with anti-Tn antibodies in Golgi-like supranuclear locations, very similar to the staining pattern found with antibodies to glycosyltransferases (Mandel et al. 1991). The combined results indicate that the specificity of the lectin domains of GalNAc-T2 and -T4 is correlated with the expression of

Table II. Substrate specificities measured by short time essays of GalNAc-T1, GalNAc-T2^{458H}, GalNAc-T2^{541A}, and GalNAc-T11

Peptide	Peptide sequence	GalNAc-transferase				
		T1 (mU/mL)	T2 (mU/mL)	T2 ^{458H} (mU/mL)	T2 ^{541A} (mU/mL)	T11 (mU/mL)
MUC1	AHGVT <u>S</u> APDTRPAPGSTAPP	122 ± 7	158 ± 10	121 ± 13	151 ± 12	97 ± 2
MUC1-2GalNAc	AHGVT <u>S</u> APDTRPAPGSTAPP ^a	25 ± 8	46 ± 6	5 ± 7	55 ± 8	0 ± 0
MUC1a'	AHGVT <u>S</u> APDTR	86 ± 7	31 ± 3	29 ± 7	28 ± 11	85 ± 3
MUC1b'	RPAPGSTAPPA	0 ± 0	87 ± 6	94 ± 8	128 ± 7	18 ± 4
MUC2	PTTTPITTTTTVTPPTPTGTQTPTTTPISTTC	283 ± 10	152 ± 18	164 ± 7	165 ± 13	125 ± 3
MUC2-6GalNAc	PTTTPITTTTTVTPPTPTGTQTPTTTPISTTC ^b	Not determined	62 ± 10	0 ± 0	46 ± 9	0 ± 0
MUC7	CPPTPSATTPAPPSSAPPETAA	36 ± 3	29 ± 8	53 ± 10	46 ± 5	18 ± 3
IgA hinge	VPSTPPTSPSTPPTSPSK	281 ± 13	161 ± 5	185 ± 6	142 ± 5	224 ± 2
OSM fragment	LSESTTQLPGGGPGCA	108 ± 2	0 ± 1	0 ± 0	0 ± 1	0 ± 0
hCG- β	PRFQDSSSSKAPPLPSPSRLPG	0 ± 0	51 ± 6	54 ± 1	65 ± 7	0 ± 1
Erythropoietin T	PPDAATAAPLR	74 ± 5	39 ± 10	47 ± 11	66 ± 8	76 ± 3

^aGalNAc attachment sites underlined.

^bGalNAc attachment sites not characterized.

GalNAc residues on proteins irrespective of density and pattern of glycosylation.

Demonstration of lectin-mediated catalytic functions of GalNAc-T2

Although it is now clear that GalNAc-T4 uses the lectin domain to mediate GalNAc-glycopeptide substrate specificity, the function has not been elucidated for the lectin domains of GalNAc-T2 and other GalNAc-transferases that efficiently utilize peptide substrates without a particular specificity for GalNAc-glycopeptides. We used GalNAc-T2 as a representative example and tested the effect of lectin mutations on catalytic activity by performing a comparative study of the catalytic activity of soluble wild type and mutant GalNAc-T2. The majority of the studies were performed with the GalNAc-T2^{458H} α -repeat mutant because our lectin binding studies demonstrated that GalNAc-T2^{458H} lost capacity for carbohydrate binding, while a similar mutation in the γ -repeat had no effect (Fig. 1, Panel D). Initially the specific activities of purified wild type and mutant GalNAc-T2 were found to be similar with several peptide substrates (Table II). Only minor loss of activity was found following shaking incubation at 37°C for 12 h (GalNAc-T1: 25%, GalNAc-T2: 23%, GalNAc-T2^{458H}: 15%, GalNAc-T2^{541A}: 19%, and GalNAc-T11: 31%). Thus, the mutation introduced in the lectin domain of GalNAc-T2 did not significantly affect the stability of the enzyme. Comparison of K_m for the donor substrates UDP-GalNAc and UDP-Gal revealed only minor differences (Table III). Furthermore, inhibitory effects of UDP with GalNAc-T2 and GalNAc-T2^{458H} were similar (not shown). The minor differences in K_m are not considered significant. In conclusion, the wild type and mutant GalNAc-T2 enzymes appear indistinguishable, except that GalNAc-T2^{458H} does not exhibit lectin-mediated glycopeptide binding activity.

Using a panel of both single-site and multiple-site acceptor substrates, the potential influence of a lectin domain mutation on the substrate specificity of GalNAc-T2 was analyzed. Previously, we have shown that GalNAc-T2 catalyzes glycosylation of three of the five potential *O*-glycosylation sites of the MUC1 tandem repeat sequence (AHGVT⁵SAPDTRPAPGS¹⁶T¹⁷APPA, sites utilized indicated by numbers), and that this follows an ordered sequence with Thr¹⁷ first, followed rapidly by Thr⁵ and more slowly by Ser¹⁶ (Wandall et al. 1997). As shown in Figure 4, the GalNAc-T2^{458H} lectin mutant efficiently incorporates GalNAc into two sites of the MUC1 derived peptide, whereas the wild type incorporates into three sites. From previous studies, it is known that GalNAc-T11 incorporates two GalNAc residues into the MUC1 tandem repeat at Thr⁵ and Thr¹⁷ (AHGVT⁵SAPDTRPAPGST¹⁷APPA) (Schwientek et al. 2002). The importance of the lectin domain in the glycosylation of the Ser¹⁶ in the MUC1 tandem repeat could therefore be tested using MUC1 tandem repeats preglycosylated by GalNAc-T11. For this MUC1 GalNAc-glycopeptide (AHGVT⁵SAPDTRPAPGST¹⁷APPA), only wild-type GalNAc-T2 glycosylated the third site (Ser¹⁶), whereas GalNAc-T2^{458H} did not. The GalNAc-T2^{541A} mutant with retained lectin carbohydrate-binding properties functioned like the wild type (not shown). Further analysis of other peptide substrates derived from mucin tandem repeats of

Table III. Substrate specificities measured by short time assays of GalNAc-T2, GalNAc-T2^{458H}, and GalNAc-T2^{541A}

Substrate	Sequence	GalNAc-T2		GalNAc-T2 ^{458H}		GalNAc-T2 ^{541A}	
		K_m (mM)	V_{max} (pmol/min/ μ g)	K_m (mM)	V_{max} (pmol/min/ μ g)	K_m (mM)	V_{max} (pmol/min/ μ g)
<i>Acceptor</i>							
MUC1	AHGVT ⁵ SAPDTRPAPGSTAPP	0.26 \pm 0.16	97 \pm 14	0.36 \pm 0.19	94 \pm 5	0.15 \pm 0.03	143 \pm 15
MUC1-2GalNAc	AHGVT ⁵ SAPDTRPAPGSTAPP	0.87 \pm 0.23	72 \pm 21	Not Applicable	Not Applicable	0.86 \pm 0.21	118 \pm 39
MUC2	PTTPTTTT ¹⁷ VTPTPTGTQPTTTP ¹⁶ ISTTC	0.010 \pm 0.001	81 \pm 1	0.025 \pm 0.006	150 \pm 6	0.018 \pm 0.003	273 \pm 12
MUC2-6GalNAc	PTTPTTTT ¹⁷ VTPTPTGTQPTTTP ¹⁶ ISTTC	0.062 \pm 0.007	302 \pm 17	Not Applicable	Not Applicable	0.066 \pm 0.010	292 \pm 18
IgA hinge	VPSTPTPSPSTPTPTPSPSK	0.015 \pm 0.002	321 \pm 18	0.020 \pm 0.005	236 \pm 1	0.010 \pm 0.01	145 \pm 21
IgA hinge-4GalNAc	VPSTPTPSPSTPTPTPSPSK	0.12 \pm 0.01	83 \pm 13	0.81 \pm 0.20	119 \pm 24	0.32 \pm 0.09	49 \pm 5
<i>Donor</i>							
UDP-GalNAc		0.011 \pm 0.002	69 \pm 12	0.030 \pm 0.001	45 \pm 13	0.013 \pm 0.001	46 \pm 12
UDP-Gal		0.030 \pm 0.002	56 \pm 6	0.036 \pm 0.001	65 \pm 5	0.041 \pm 0.005	85 \pm 5

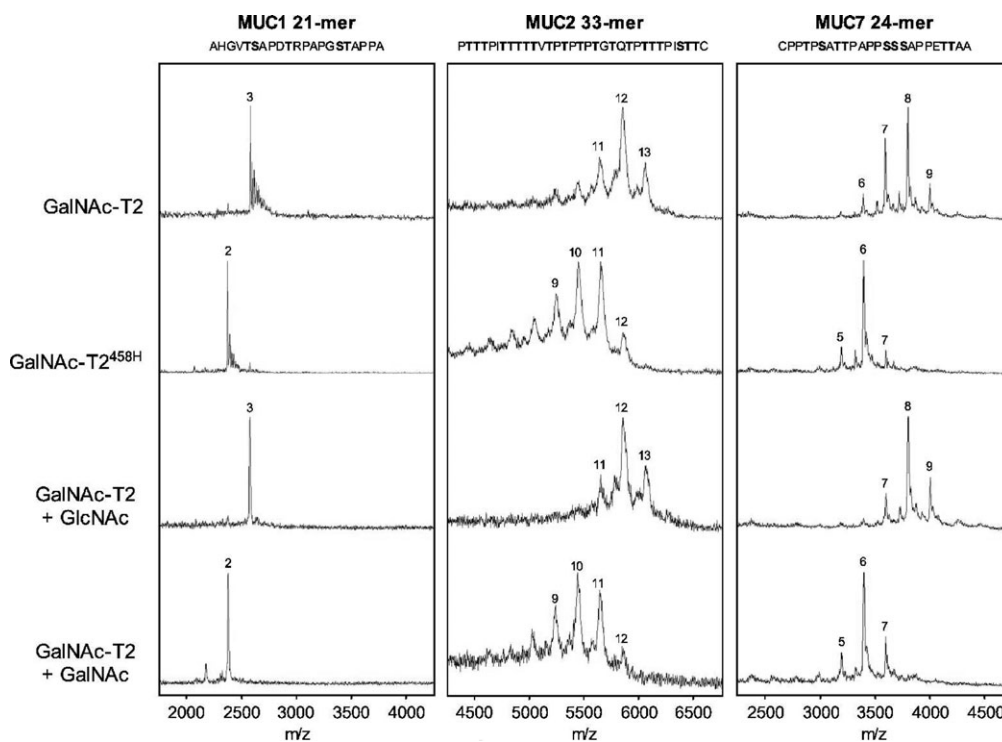


Fig. 4. MALDI-TOF analysis of GalNAc-T2 lectin mutant reaction end products. Purified recombinant GalNAc-T2 and -T2^{458H} were used in excessive glycosylation of the mucin tandem repeat peptide substrates from MUC1, MUC2, and MUC7, and end product analyzed by MALDI TOF. The GalNAc-T2 reaction was tested in the presence of 250 mM GalNAc (and 10 mM benzyl- α -GalNAc, not shown) and 250 mM GlcNAc (and 10 mM benzyl- α -GlcNAc, not shown) to inhibit the lectin mediated function (Hassan, Reis, et al. 2000).

MUC2 and MUC7 with higher densities of potential *O*-glycosylation sites revealed the similar patterns. Thus, GalNAc-T2^{458H} incorporated 1–2 fewer (than wild-type GalNAc-T2) GalNAc residues per substrate molecule of MUC2 and MUC7 peptide substrates with 20 and 9 potential glycosylation sites, respectively (Figure 4). Inhibition with free GalNAc (250 mM) selectively inhibited the wild-type glycosylation of substrates with multiple acceptor sites, abolishing the difference seen in glycosylation between wild-type GalNAc-T2 and lectin deficient GalNAc-T2. No inhibition was seen with GlcNAc. Similar inhibition was observed with GalNAc- α -benzyl (10 mM) (not shown). These results are comparable to those observed with inhibition of lectin-mediated functions of GalNAc-T4 (Hassan, Reis, et al. 2000) and support that GalNAc-specific lectins are important for the efficient glycosylation of substrates with multiple acceptor sites.

The above analysis includes evaluation of end products formed in exhaustive endpoint reactions. To provide further insight, the activity of GalNAc-T2 and its lectin mutants were tested with a panel of peptides and corresponding partially GalNAc-glycosylated glycopeptides (Table II). The specific activities of GalNAc-T2, GalNAc-T2^{458H}, and GalNAc-T2^{541A} measured in short-time reactions with minimum product development, were similar to activities with peptides encoding mucin tandem repeat sequences from MUC1 and MUC2, and the hinge region of immunoglobulin A. In striking contrast, only GalNAc-T2 and GalNAc-T2^{541A} readily utilized the partially GalNAc-glycosylated peptides (MUC1-2GalNAc and MUC2-6GalNAc), whereas

GalNAc-T2^{458H} did not. Furthermore, the K_m of the glycopeptide IgA hinge-4GalNAc was evidently higher with the GalNAc-T2^{458H} enzyme compared with the wildtype and GalNAc-T2^{541A} enzyme. The kinetic parameters of GalNAc-T2, GalNAc-T2^{541A}, and GalNAc-T2^{458H} for the nonglycosylated peptide substrates (MUC1, MUC2, and IgA) were very similar, with only minor differences in K_m values (Table III).

Finally, we performed time course studies using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) to monitor product developments of a panel of peptides with GalNAc-T2 and GalNAc-T2 lectin mutants (Figures 5 and 6). Wild-type GalNAc-T2 and GalNAc-T2^{458H} initially exhibited similar rates of GalNAc incorporation into the 33-mer MUC2 substrate with 20 potential acceptor sites, which is in accordance with initial-rate glycosylation assays that demonstrated similar activity of the wild type and the mutant GalNAc-T2 with the nonglycosylated MUC2 peptide (Table II). However, the reaction with GalNAc-T2^{458H} was reduced significantly after addition of the first 2–3 GalNAc residues, and the final product was slightly less glycosylated than that obtained with the wild-type GalNAc-T2 (Figure 5A). This is in agreement with the finding that wild type GalNAc-T2 showed good activity with the MUC2-6GalNAc peptide, whereas GalNAc-T2^{458H} showed essentially undetectable activity (Table II) in short-time assays. The reaction kinetics with the MUC2 tandem repeat sequence was further studied by high-performance liquid chromatography HPLC with fluorescent labeled short- and long-MUC2 peptides (PTTTPK and GTQTPTTPTTTTTVPTPTPTG) (Figure 5B and C), verifying the slow catalytic activity of

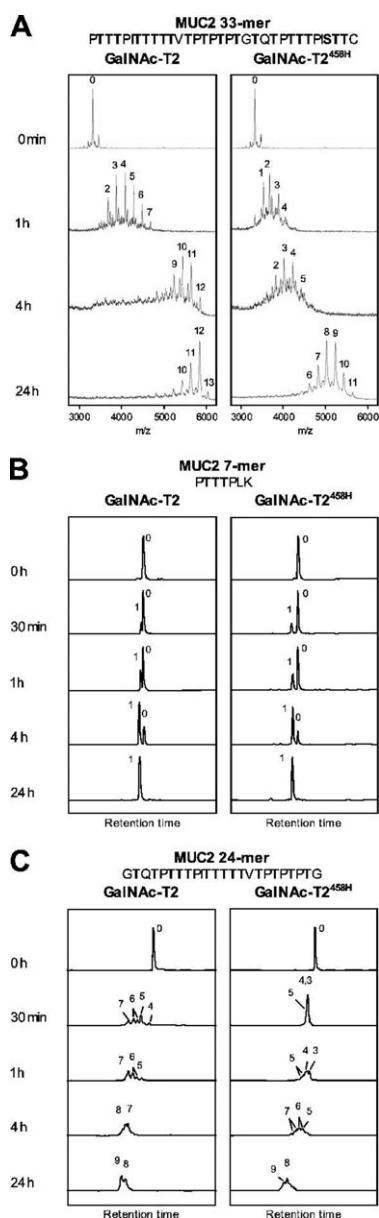


Fig. 5. Time-course analysis of GalNAc-T2 and -T2^{458H} reactions monitored by MALDI-TOF. (A) The GalNAc-T2 lectin mutant exhibits a significantly slower reaction velocity compared to the wildtype enzyme with the MUC2 tandem repeat peptide substrate. MUC2 synthetic peptide was glycosylated with GalNAc-T2 and -T2^{458H} and samples were analyzed by MALDI-TOF MS at the indicated timepoints. This result was further verified by HPLC analysis of different MUC2 peptide substrates; (B) MUC2 short (FITC-PTTTPLK) and (C) MUC2 long (FITC-GTQTPTTTTPTTTTITTTVTPITPTG) with GalNAc-T2 and GalNAc-T2^{458H}. No difference in glycosylation rate was seen with the short peptide, whereas the longer MUC2 substrate showed a slower reaction velocity with the mutant GalNAc-T2^{458H}. The number of GalNAc residues incorporated per MUC2 peptide is indicated by number above elution peaks.

GalNAc-T2^{458H} compared with wild-type GalNAc-T2. In this experiment, the short MUC2 peptide serves as an internal control for the initial GalNAc-glycosylation reaction with both wild-type GalNAc-T2 and GalNAc-T2^{458H} exhibiting the same reaction velocity.

IgA hinge peptide substrate

It has been reported (Iwasaki et al. 2003) that the mucin domain sequence of the IgA hinge region served as a substrate for the GalNAc-T2 isoform, and that several other GalNAc-transferase isoforms including GalNAc-T1 were inactive with peptides derived from this sequence. The IgA hinge peptide could therefore represent a relevant *in vivo* substrate requiring the concerted action of both the catalytic domain and a functional lectin domain of GalNAc-T2 in order to be fully glycosylated. However, as shown in Table II, we found that both GalNAc-T1, -T2, and -T11 efficiently catalyze glycosylation of the IgA peptide. The peptide design used was identical to the design used by Iwasaki et al. (2003), except that the C-terminus was labeled with 5-carboxyfluorescein. The glycosylation reaction of IgA hinge peptide with GalNAc-T1, GalNAc-T2, GalNAc-T2^{458H}, and GalNAc-T11 was further studied in a time-course assay monitored by MALDI-TOF. Similarly to the findings with other mucin peptide sequences, GalNAc-T2^{458H} exhibited slower reaction kinetics than wild-type GalNAc-T2 after the first initial incorporation of 1-2 GalNAc residues, and the end product had fewer GalNAc residues incorporated. In agreement with the results of the initial velocity assay (Table II), GalNAc-T1 had a higher reaction velocity but incorporated fewer GalNAc residues. GalNAc-T11 has been shown to incorporate fewer GalNAc residues into a number of mucin peptide substrates, including those derived from MUC1, MUC7, and MUC2 tandem repeats (Schwientek et al. 2002). As shown in Figure 6, GalNAc-T11 essentially only incorporates two GalNAc residues into the IgA hinge peptide, but the initial reaction velocity is similar to that of the other isoforms tested.

Discussion

This study addressed the potential carbohydrate-binding properties and function of the lectin-domains of two distinct GalNAc-transferase isoforms, GalNAc-T4 and -T2. These transferases represent examples of GalNAc-transferases with apparent lectin-mediated GalNAc-glycopeptide glycosylation functions (GalNAc-T4), and GalNAc-transferases for which such functions have not been determined or predicted (GalNAc-T2). Previous studies of catalytic functions of GalNAc-T4 and -T1 with lectin mutations, as well as inhibition studies of wild-type enzymes with saccharides, have provided indirect evidence that the lectins modulate the catalytic functions through interactions with GalNAc (Hassan, Bennett et al. 2000; Tenno, Kezdy, et al. 2002; Tenno, Saeki, et al. 2002). The present study now demonstrates direct carbohydrate binding to GalNAc of the non biotinylated and biotinylated isolated lectins and biotinylated soluble enzymes of both GalNAc-T2 and -T4, confirming that these lectins are functional.

We expressed and tested several different tagged lectin and GalNAc-transferase construct. The lectin constructs presented in Figure 1 were expressed and secreted in the insect cell system, albeit with relatively low yields compared with the secreted soluble enzyme constructs. Both nonbiotinylated and biotinylated GalNAc-T4 and -T2 lectins bound GalNAc-glycopeptides, and the binding was inhibited by

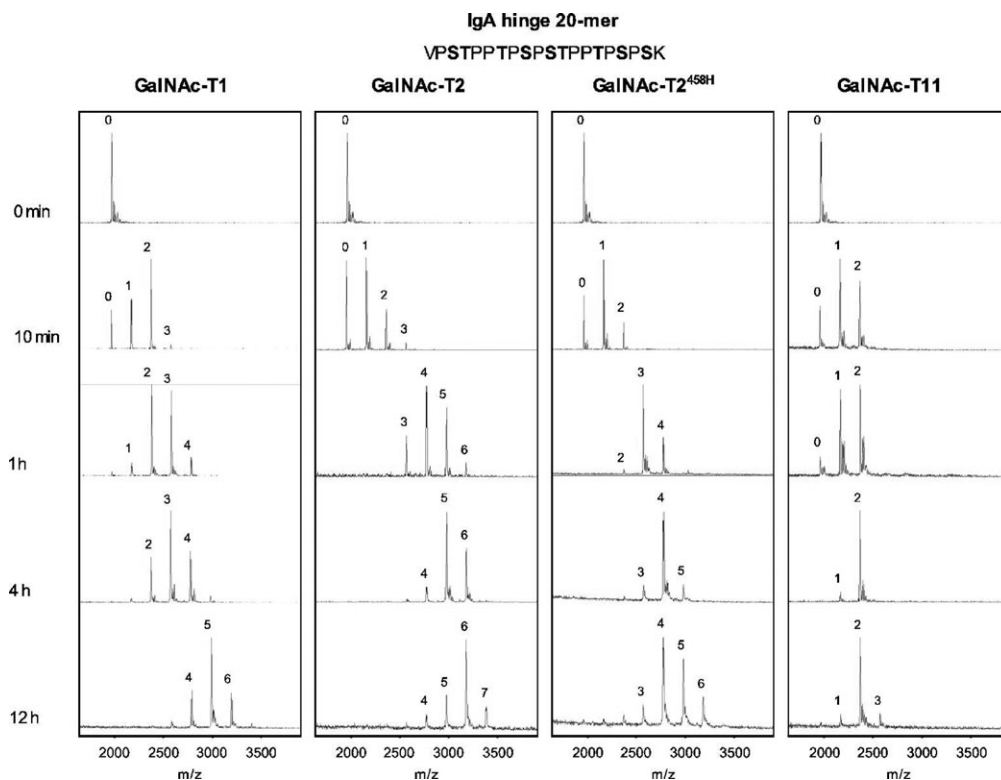


Fig. 6. Time-course analysis of GalNAc-glycosylation of IgA hinge peptide with a panel of GalNAc-transferase isoforms monitored by MALDI-TOF. IgA hinge peptide was glycosylated with GalNAc-T1, GalNAc-T2, -T2^{458H}, and -T11 and samples were analyzed by MALDI-TOF MS at the indicated timepoints. The mutant GalNAc-T2^{458H} exhibited slower reaction velocity than the wildtype GalNAc-T2 and incorporated fewer GalNAc residues.

GalNAc- α/β -derivatives (Table I). A minor inhibition of GalNAc-T4 to GalNAc-MUC1 was observed with high concentrations of galactose. Single mutations in the DXD motif disrupting the catalytic function did not interfere with GalNAc-transferase binding to glyco peptides. Further verification of the carbohydrate specificity was provided using the lectins in immunocytology of CHO/dID and tissue sections of mucin producing cells. The lectins showed binding patterns similar to anti-Tn antibodies specifically recognizing glyco-proteins presenting penultimate GalNAc α -S/T structures (Figure 3). The interaction between the GalNAc-T2lec (GalNAcTlec latin domain of UDP-*N*-acetyl- α -D-galactosamine: polypptide *N*-acetylgalacto, saminy transferase) and GalNAc-MUC1 was additionally characterized with surface plasmon resonance, demonstrating binding of the nonlabeled GalNAc-T2 lectin domain to GalNAc-MUC1 (Figure 2).

The ricin-like lectin domain consists of three putative carbohydrate-binding motifs (α -, β -, and γ -repeats). A single mutation in the GalNAc-T4 α -repeat abolishes the GalNAc-glycopeptide substrate specificity of the enzyme as well as carbohydrate binding (Hassan, Bennett, et al. 2000) (Figure 1D). Similarly, a single mutation in the CLD motif in the α -repeat of GalNAc-T2 enzyme and the GalNAc-T2 isolated lectin domain abolished binding to GalNAc-glycopeptides and influences the GalNAc-T2 glycosylation kinetics of GalNAc-glycopeptides. In contrast, a D541A substitution in the CLD motif in the γ repeat did not affect binding to GalNAc-glycopeptides (Figure 1D). Moreover, no obvious effect on enzymatic activity was observed by mutating the γ

repeat of GalNAc-T2. These data are consistent with the findings by Tenno, Sacki, et al. (2002), demonstrating that mutations in the α and β repeats, but not in the γ repeat, adversely affects the kinetic properties of GalNAc-T1 with an apomucin substrate. Interestingly, inhibition experiments of GalNAc-T1 activity with free GalNAc have shown a cooperative effect of α and β repeat mutations (Tenno, Saeki, et al. 2002).

Binding of nonbiotinylated GalNAc-T2 and -T4 isolated lectin domains to GalNAc-MUC1 in the ELLA assay was detected with antibodies to their tags (anti-HIS and -T7), as well as with monoclonal antibodies directed to the lectin domains. The binding of the isolated nonlabeled GalNAc-T2 lectin domain was confirmed with surface plasmon resonance analysis. Surprisingly, biotinylation of both lectins and enzymes resulted in increased lectin binding activity. Hence, in vitro lectin binding was only observed with the catalytic inactive biotinylated GalNAc-T2 and -T4 enzymes, and not with the catalytic active non labeled enzymes. Importantly, the unlabeled active GalNAc-T2 and -T4 enzymes did not inhibit binding of either lectins or the biotinylated enzymes to GalNAc-glycopeptides (not shown), excluding that differences in detection methods could explain the binding differences. The reason for the increased binding of the biotinylated lectins is not known. Recently, the crystal structure of murine GalNAc-T1 and GalNAc-T2 were elucidated (Fritz et al. 2004, 2006). The proposed structures indicate that the carbohydrate-binding sites of the α , β , and γ repeats of both GalNAc-T1 and T2 are accessible for glycopeptide

binding and provide novel insight into extensive interactions between the catalytic domain and the lectin domain. For GalNAc-T1, the two domains are separated by a short random coil and β -strand, and they form a deep cleft in the surface of the enzyme predicted to contain the catalytic site. The interaction between the catalytic domain and the lectin domain is substantially reduced in GalNAc-T2, and interestingly substrate binding induces lectin domain, mobility. It is possible that biotinylation of the intact enzyme influence the interaction of the catalytic domain and lectin domain, thereby exposing the carbohydrate-binding pocket of lectin domain, but this requires further investigation.

GalNAc-transferases with selective or unique GalNAc-glycopeptide substrate specificity include human GalNAc-T4, -T7, and -T10, and such isoforms are conserved during evolution as found in the fruit fly (Schwientek et al. 2002) and in parasites (Stwora-Wojczyk et al. 2004). The lectin domains of these three isoforms are predicted to enhance the kinetic properties of the GalNAc-transferases with GalNAc-glycopeptide substrates. However, what are the functions of the lectin domain of GalNAc-transferases, such as GalNAc-T1 and T2, which are believed to act primarily as enzymes that initiate glycosylation of previously unglycosylated protein substrates? In this context, Tenno et al. (Tenno, Ember, et al. 2002; Tenno, Saeky et al. 2002) demonstrated an apparently selective decrease in activity of lectin mutated GalNAc-T1 with multiacceptor peptide substrates sites. Furthermore, while this manuscript was under preparation Fritz and colleagues demonstrated that removal of the lectin domain of GalNAc-T2 affected the glycosylation of glycopeptides but not peptides (Fritz et al. 2006). The results collectively suggest that the role of the lectin domain is not in catalysis of the initial GalNAc, but in the “follow-up” reaction. In contrast to the studies by Tenno et al., we did not see differences in kinetic properties for single acceptor substrates and nonglycosylated multiacceptor peptide substrates in short-time assays when the lectin domain was mutated (Tables II and III). This discrepancy is most likely because we determined K_m under enzyme reaction conditions limited to ensure that the GalNAc incorporation was <25% of peptide substrate (mol/mol). When we tested the corresponding partially glycosylated multiacceptor peptides, we found a significant difference between wild-type GalNAc-T2 and the α repeat lectin mutated enzyme, where the latter was almost inactive. Analysis of the final glycosylation products formed with the multiacceptor substrates confirmed that the lectin mutated enzyme did not transfer as many GalNAc residues to the acceptor substrates as the wild-type enzyme. In particular, the reaction with the MUC1 peptide showed that the last-and third-site in the MUC1 repeat (Wandall et al. 1997) (serine in -GSTA-) catalyzed by GalNAc-T2 was the site requiring a functional lectin domain (Figure 4). Time-course assays monitoring product development with several multiacceptor peptide substrates showed that the initial glycosylation velocity for the addition of the first GalNAc were similar, while subsequent site glycosylation velocities were considerably (three-to-six- fold) slower with the lectin mutant (Figures 5 and 6). Thus, similar to what we previously demonstrated for GalNAc-T4, GalNAc-T2 requires the lectin domain for transfer to unique GalNAc-glycopeptide substrate sites.

GalNAc-transferases catalyze glycosylation of unique substrate sites in peptides with multiple acceptor sites in an ordered fashion, in which there is preference for some sites and less favorable sites with the highest K_m are catalyzed last (Wandall et al. 1997). The GalNAc-transferase lectins may serve as an effective mechanism to enhance kinetic properties for these last sites through substrate approximation. The two GalNAc-transferase lectins studied in this report exhibit specificity for GalNAc and hence did not function with glycopeptide substrates where GalNAc residues were elongated by e.g. the core 1 structure (Gal β 1-3GalNAc α 1-O-Ser/Thr) (Hanisch et al. 2001). Although, it is conceivable that not all GalNAc-transferase lectins have the same specificity, it is interesting in this context that a recent study has demonstrated that the large secreted mucin MUC5AC appears to be processed through an intermediate stage with complete GalNAc-O-glycosylation before O-glycan extension (Sheehan et al. 2004). Several bacterial enzymes have lectin domains with carbohydrate-binding properties. In bacterial sialidases the lectin domain increase the affinity between the hydrolases and the acceptor substrates (Thobhani et al. 2003). In the same way, bacterial xylanases have a ricin-like lectin domain, which like the polypeptide GalNAc-transferases contain α -, β -, and γ repeats that are important for efficient catalysis of insoluble xylan by enhancing affinity of the enzyme to the multiunit insoluble xylan substrate (Dupont et al. 1998; Fujimoto et al. 2002). A similar mechanism is envisioned for GalNAc-transferases, where the lectin domain combined with the catalytic unit is proposed to provide a bivalent binding mode to partially completed substrates.

A recent report suggests that GalNAc-T2 is the sole GalNAc-transferase responsible for O-glycosylation of the IgA hinge region (Iwasaki et al. 2003). Glycosylation of the hinge region of IgA has attracted interest because of the potential involvement of decreased O-glycosylation in the pathogenesis of IgA glomerulonephritis. Iwasaki et al. (2003) identified six GalNAc-transferases (GalNAc-T1, -T2, -T3, -T4, -6, and -T9) in B cells and showed that only GalNAc-T2 produced significant glycosylation incorporating multiple GalNAc-residues in the IgA peptide design listed in Tables II and III and shown in Figure 6. The IgA hinge peptide could therefore represent a relevant *in vivo* substrate, requiring the concerted action of both the catalytic domain and a functional lectin domain of GalNAc-T2 in order to be fully glycosylated. Hence, we assessed the importance of the GalNAc-T2 lectin domain in glycosylating the IgA peptide (Figure 6). However, although the lectin domain influenced the rate of glycosylation, only a minor difference was found between GalNAc-T2 wild type and GalNAc-T2 with a nonfunctional lectin domain. Furthermore, to our surprise, all tested GalNAc-transferases utilized the substrate efficiently, including GalNAc-T1 (Table II and Figure 6) that was shown to have 100-fold less activity with the IgA hinge peptide compared with GalNAc-T2 in the Iwasaki et al. (2003) study. The reason for this discrepancy is unclear, but it could be related to the use of a fluorescent tagged peptide substrate. Regardless, the present results clearly rule out that GalNAc-T2 is the sole isoform responsible for O-glycosylation of the IgA hinge region. This is in agreement with the general finding that GalNAc-transferases exhibit a high degree of redundancy in functions with mucin-type peptide sequences

containing high density of acceptor sites (Hassan, Bennett, et al. 2000).

In summary, this study has provided direct evidence for carbohydrate-binding properties of polypeptide GalNAc-T2 and -T4 lectins, and demonstrated that the lectin domains may have an important role to improve the kinetic properties with partially GalNAc-glycosylated substrates produced either by that same transferase or by other GalNAc-transferases, and thereby to accomplish and regulate high-density glycosylation of certain protein substrates.

Materials and methods

Expression and purification of GalNAc-transferases

Expression constructs of the soluble coding region of human GalNAc-T1, GalNAc-T2, GalNAc-T4 wild type and GalNAc-T4 with a mutation in the α -subdomain (GalNAc-T4^{459H}) were prepared and cloned into the baculovirus expression vector pAcGP67 (BD, Pharmingen, San Diego, CA) as previously described (White et al. 1995; Hassan, Bennett, et al. 2000). A soluble construct of human GalNAc-T11 (pAcGP67-His-T11 sol) was prepared by inserting a nontagged GalNAc-T11 construct (Schwientek et al. 2002) into pAcGP67 vector containing a 6xHis-T7-tag (-SSHHHHHHSSGLVPRGSHMASMTGGQQMD-, T7 tag shown in bold). GalNAc-T2 with a mutation in the γ -subdomain (GalNAc-T2^{458H}) and GalNAc-T2 with a mutation in the γ -subdomain (GalNAc-T2^{541A}) were generated using QuickChange site-directed mutagenesis Kit (Stratagene, La Jolla, CA) as recommended by the supplier. Primer pair T2MUT1 (5'-GGAACAACTGCCTCCACACTTTGGGACACTTTG) and T2MUT2 (5'-CAAAGTGTCCCAAAGTGTG GAGGCAGTTAGTTCC) was used for GalNAc-T2^{458H} and primer pair T2MUT3 (5'-CTGTGCCTGGCCAGTCGCACG) and T2MUT4 (5'-CGTGC GACTGGCCAGGCACAG) was used for GalNAc-T2^{541A}. Likewise, an enzymatic inactive GalNAc-T2 construct (GalNAcT2^{224H}) was generated by site-directed mutagenesis changing the amino acid position D224 to H224 in the DXH motif using QuickChange site-directed mutagenesis kit (Stratagene) as recommended by the supplier with oligos T2DXHFOR:5'-GGTCTGACCTTCCTG C ACAGTCACTGCGAGTGT-3', T2DXHREV:5'-ACACTCGC AGTGACTGT G ACAGGAAGGTCAGGACC-3', and pAcGP67-T2sol as template DNA.

A GalNAc-T2 lectin construct was made using the primer pair T2lec (5'-GCGAAGA**ACGTTT**ATCCAGAGTTAA GGGTTCC) and EBHC68 (5'-GCGA**ATTC**TACT GCTGCAGGTTGAGC) (*Psp*1406I and *Eco*RI overhangs are shown in bold) with Pfu Ultra polymerase (Stratagene) on a TC2400 thermocycler (Applied Biosystems, Tokyo, Japan) as recommended by the supplier. The generated GalNAc-T2 lectin fragment included coding nucleotides 1297–1716 and was cloned into the pAcGP67-His-vector.

A truncated GalNAc-T4 lectin expression construct was prepared using a *Hind*III restriction site at nucleotide position 1211 of the coding region for excision of nucleotides 1211–1737 with *Hind*III/*Bam*HI from the existing soluble GalNAc-T4 expression construct (Bennett et al. 1998). The lectin construct was inserted into pAcGP67-His downstream from a previously described His-tag generating pAcGP67-His-T4lec

expression construct. A GalNAc-T4 lectin construct with a mutation in the α -subdomain (GalNAc-T4lec^{459H}) was similarly made from the existing GalNAc-T4^{459H} expression construct (Hassan, Bennett, et al. 2000) generating pAcGP67-His-T4lec^{459H}.

Plasmids pAcGP67-T1, -T2, -T2^{458H}, -T2^{541A}, T2^{224H}, -T4, -T4^{459H}, -T11, -T2lec, -T4lec, and -T4lec^{459H} were co-transfected with BaculoGold™ DNA (Pharmingen), and recombinant baculovirus was obtained after two successive amplifications in Sf9 cells as described previously (Wandall et al. 1997). Amplified virus was used for infection of High Five™ cells grown in serum-free media (Invitrogen, Carlsbad, CA) in upright roller bottles shaken 140 rpm in water baths at 27 °C. Secreted, soluble recombinant proteins were harvested (centrifugation at 2000 × g) and supernatants were subjected to chromatography on Amberlite® IRA-95 followed by dilution with 25 mM Bis-Tris (pH 6.5) and further purified by ion-exchange chromatography on 10 mL SP Sepharose™ Fast Flow (Sigma-Aldrich, St. Louis, MO) and eluted with a gradient of NaCl from 10 to 500 mM. Fractions containing enzyme were pooled and concentrated using a Centrprep® YM-10 centrifugation filter unit with 10 000 Da cut-off (Millipore, Billerica, MA). The concentrated GalNAc-transferases were diluted five-fold in Bis-Tris buffer, applied to a Mini-S™ column (PC 3.2/3, GE Healthcare, Amersham Biosciences, PA), and eluted with a NaCl gradient from 0.01 to 1 M. Activity and purity were analyzed by standard GalNAc-transferase assay (reference substrate MUC1 21-mer peptide) and sodium (SDS-PAGE) stained for proteins with Coomassie Blue R-250 using bovine serum albumin as a standard.

Purification of the lectin domains

Truncated lectin proteins were purified on iminodiacetic acid metal affinity chromatography (IMAC) Ni²⁺-charged (QIAGEN) with elution using 250 mM imidazole in 50 mM sodium phosphate, pH 8.0, 500 mM NaCl. Eluted proteins were dialyzed three times against phosphate buffered saline PBS (10 mM sodium phosphate, pH 7.4, 150 mM NaCl) and concentrated using a centrifugation filter unit (Millipore; 10 000 kDa cut-off). Purity was analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) under reducing conditions, and stained for proteins with Coomassie Blue R 250.

GalNAc-transferase assay

Standard assays were performed in 25–50 μ L of total reaction mixtures containing 25 mM Cacodylate (pH 7.4), 10 mM MnCl₂, 0.25% Triton X-100, 50 μ M UDP-[¹⁴C]GalNAc (1–3,000 cpm/nmol) (GE Healthcare), 0.01–0.5 mU of GalNAc-transferase, and 20–100 μ M of acceptor peptide. Peptides were synthesized by Neosystem (Strasbourg, France), Cancer Research UK, or fluorescein isothiojanate (FITC-labeled MUC2 7- and 24-mers) as previously described (Takeuchi et al. 2002). Enzymatic activity towards the acceptor peptides was routinely determined by scintillation counting after Dowex-1 formic acid chromatography. Peptides and products produced by in vitro glycosylation were evaluated by mass spectrometry (MS). Assuming that the initial velocity reactions display apparent Michaelis–Menten kinetics K_m was determined. GalNAc-transferase

assays used for determination of K_m of acceptor substrates were modified to include 300 μM UDP-[^{14}C]GalNAc (2000 cpm/nmol), with peptides in varying concentrations from 0.005 to 2 mM. Assays to determine K_m for UDP-GalNAc and UDP-Gal were performed with saturating concentrations of acceptor substrates (GalNAc-T1, 500 μM IgA hinge; GalNAc-T2 and mutant GalNAc-T2, 100 μM IgA hinge; and GalNAc-T11, 500 μM IgA hinge). Reaction incubation times did not exceed 20 min. Assays were performed in duplicates or triplicates. Exhaustive glycosylation of peptides was performed with 3 to 10-fold excess of potential Ser/Thr acceptor sites, and 0.25–5 mU of GalNAc-transferase (specific activity determined using the relevant acceptor peptide to be glycosylated). Reactions were incubated for extended periods (>12 h) at 37 °C as indicated, and product development monitored by MALDI-TOF or in the case of FITC-labeled peptides by HPLC (JASCO, Tokyo, Japan equipped with a Cosmosil column C₁₈, 6 × 150 mm; Nacalai Tesque, Japan). HPLC used a linear gradient from 10 to 30% solvent B (0.05% trifluoroacetic acid/70% 2-propanol in acetonitrile) in solvent A (0.05% trifluoroacetic acid in water). Eluates were monitored by fluorescence intensity at 520 nm (ex: 492 nm). MUC1-2GalNAc was produced by in vitro glycosylation of MUC1 with recombinant GalNAc-T11. MUC2-6GalNAc and IgA-4GalNAc was produced by in vitro glycosylation with recombinant GalNAc-T2, using limited amount of donor acceptor substrate. Product development was monitored using MALDI-TOF analysis and glycosylation reaction stopped with ethylene EDTA and glycopeptides purified by C-18 HPLC. Stability of GalNAc-transferase activity was analyzed as previously described (Wandall et al. 1997).

MALDI-TOF mass spectrometry

Sampling of reactions (1 μl) was purified by Poros[®] 20 R2 nanoscale reversed-phase chromatography (Perceptive Biosystems, Applied Biosystems, Foster City, CA) and applied directly to the probe with matrix (Mirgorodskaya et al. 1999). The matrix used was 2,5-dihydroxybenzoic acid (25 mg/mL, Aldrich) dissolved in a 2:1 mixture of 0.1% trifluoroacetic acid in 30% aqueous acetonitrile (Rathburn Chemicals Ltd, Scotland, UK). Mass spectra were acquired on a Voyager-DE mass spectrometer equipped with delayed extraction (Perceptive Biosystems). Analysis of FITC-labeled MUC2 peptides was done with α -cyano-4-hydroxycinnamic acid (10 mg/mL) dissolved in 0.1% trifluoroacetic acid/50% ethanol in water, and analysis with a Voyager Elite instrument (Perceptive Biosystems) operating at an accelerating voltage of 20 kV (grid voltage 93.5%, guide wire voltage 0.05%) in the positive ion and linear mode with the delayed extraction setting.

Lectin binding assay

Polystyrene microtiter plates (Maxisorb, Nunc, Roskilde, Denmark) were coated with peptides or glycopeptides in carbonate buffer (pH 9.6) overnight at 4 °C. Plates were washed and blocked with 0.1% Tween-20 and 0.2% bovine serum albumin (BSA) in PBS for 1 h at room temperature, followed by incubation with nonbiotinylated or biotinylated proteins in PBS with 0.05% Tween 20 for 2 h at room temperature. Plates were washed with PBS and incubated with 1:2000 dilution Streptavidin-Horseradish peroxidase

(Sigma-Aldrich) in PBS for 30 min at room temperature, washed with PBS and developed. Development was performed with 0.5 mg/mL *o*-phenylenediamine and 0.02% H₂O₂ in sodium citrate (pH 5.0) at room temperature and the reaction stopped by adding 100 μL /well of 0.5 N H₂SO₄. Competitive inhibition assays were performed either using concentrations of GalNAc-transferase proteins and lectins corresponding to end-point titers in the ELLA binding assay, or using a concentration of GalNAc-transferase and lectins proteins yielding an optical density of 1 in the ELLA assay. The proteins were preincubated for 1 h with the inhibitors. Protein biotinylation was performed by the addition of 40 μL /mL N-hydroxy-succinimidobiotin (10 mg/mL DMF) (Sigma-Aldrich) of purified protein in PBS (0.3 mg/mL). The solution was mixed end-over-end for 2 h at room temperature and dialyzed extensively against PBS.

Immunocytology/-histology

CHO Id1D (Kingsley et al. 1986) cells stably transfected with full coding human MUC1 containing 32 tandem repeats were grown in Dulbecco's minimum Eagles medium and Ham's (1:1) with 3% FCS without GalNAc and Gal in medium, in the presence of 1 mM GalNAc, or in the presence of 1 mM GalNAc and 0.1 mM Gal for 48 h. Cells were trypsinized, washed, air-dried, and fixed in ice cold acetone on multiwell slides for 10 min and then kept at -70 °C until used. Sections from paraffin blocks of minor salivary glands ($n = 4$) were obtained and processed as previously described (Mandel et al. 1991). Cells and tissues were preincubated with 10% rabbit serum for 30 min, then incubated with either 1 μg /mL biotinylated GalNAc-T2, 1 μg /mL biotinylated GalNAc-T4, 3 μg /mL biotinylated His-tagged GalNAc-T4 lectin or anti-Tn monoclonal antibody 5F4, recognizing GalNAc α S/T, overnight at 4 °C. Bound lectins were detected with either mouse monoclonal anti-T7 tag antibody (IgG2b) (Novagen, Darmstadt, Germany), monoclonal antibody 6B7 directed to the GalNAc-T2 lectin domain or Mab 4G2 directed to GalNAc-T4 lectin (incubation for 1 h at room temperature), followed by incubation with FITC-conjugated rabbit antimouse immunoglobulin (F261, Dako, Glostrup, Denmark). Control reactions consisted of replacement of the lectins with 0.1% BSA. Slides were examined in a Zeiss fluorescence microscope using epi-illumination.

Surface plasmon resonance

Surface plasmon resonance experiments were performed to evaluate the interaction between GalNAc-T2 and GalNAc-T2lec and the glycosylated synthetic peptide GalNAc-MUC1 using a Biacore system (Biacore, Uppsala, Sweden). A GalNAc-MUC1 60mer glycopeptide (HGVT*SAPDT*RPAPGST*APPA)₃ and the corresponding nonglycosylated peptide were amine-coupled to CM5 sensor chips using the amine coupling protocol and kit purchased from Biacore. GalNAc-T2 lectin domain (4, 8, 16, 32, and 65 μM) and GalNAc-T2 in biacore binding buffer (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, and 0.005% Surfactant P20, pH 7.4) was passed over the sensor chips. All experiments were run at 30 μL /min, a 50 μL sample size, and a 300 second dissociation phase. After the dissociation, the sensor chip was regenerated by the injection of 5 μL of 100 mM GalNAc at 5 μL /min. The kinetic parameters were calculated using

the BIAevaluation 4.1 Software using the 1:1 Langmuir binding model.

Acknowledgments

This work was supported by: The Danish Cancer Society, the Velux Foundation, the Danish Medical Research Council, EU QLK3-CT 2002-02010 and 2002-01980.

Conflict of interest statement

None declared.

Abbreviations

BSA, bovine serum albumin; CHO, Chinese hamster ovary; ELLA, enzyme linked lectin assay; FITC, fluorescein isothiocyanate; Gal, galactose; GalNAc, *N*-acetyl- α -D-galactosamine; GalNAcTlec, lectin domain of UDP-*N*-acetyl- α -D-galactosamine: polypeptide *N*-acetylgalactosaminyltransferase; GalNAc-transferase, UDP-*N*-acetyl- α -D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase; HPLC, high performance liquid chromatography; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; MS, mass spectrometry; PBS, phosphate-buffered saline polypeptide *N*-acetylgalactosaminyltransferase; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; UDP, uridine diphosphate.

References

- Bennett EP, Hassan H, Hollingsworth MA, Clausen H. 1999. A novel human UDP-*N*-acetyl-Dgalactosamine: polypeptide *N*-acetylgalactosaminyltransferase, GalNAc-T7, with specificity for partial GalNAc-glycosylated acceptor substrates. *FEBS Lett*, 460:226–230.
- Bennett EP, Hassan H, Mandel U, Hollingsworth MA, Akisawa N, Ikematsu Y, Merckx G, van Kessel AG, Olofsson S, Clausen H. 1999. Cloning and characterization of a close homologue of human UDP-*N*-acetyl- α -D-galactosamine: polypeptide *N*-acetylgalactosaminyltransferase-T3, designated GalNAc-T6—evidence for genetic but not functional redundancy. *J Biol Chem*. 274:25362–25370.
- Bennett EP, Hassan H, Mandel U, Mirgorodskaya E, Roepstorff P, Burchell J, Taylor-Papadimitriou J, Hollingsworth MA, Merckx G, et al. 1998. Cloning of a human UDP-*N*-acetyl- α -D-galactosamine: polypeptide *N*-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. *J Biol Chem*. 273:30472–30481.
- Clausen H, Bennett EP. 1996. A family of UDP-GalNAc: polypeptide *N*-acetylgalactosaminyltransferases control the initiation of mucin-type O-linked glycosylation. *Glycobiology*. 6:635–646.
- Dupont C, Roberge M, Shareck F, Morosoli R, Kluepfel D. 1998. Substrate-binding domains of glycanases from *Streptomyces lividans*: characterization of a new family of xylan-binding domains. *Biochem J*. 330:(Pt 1), 41–45.
- Fritz TA, Hurley JH, Trinh LB, Shiloach J, Tabak LA. 2004. The beginnings of mucin biosynthesis: the crystal structure of UDP-GalNAc:polypeptide α -N-acetylgalactosaminyltransferase-T1. *Proc Natl Acad Sci USA*. 101:15307–15312.
- Fritz TA, Raman J, Tabak LA. 2006. Dynamic association between the catalytic and lectin domains of human UDP-GalNAc:polypeptide α -N-acetylgalactosaminyltransferase-2. *J Biol Chem*. 281:8613–8619.
- Fujimoto Z, Kuno A, Kaneko S, Kobayashi H, Kusakabe I, Mizuno H. 2002. Crystal structures of the sugar complexes of *Streptomyces olivaceoviridis* E-86 xylanase: sugar binding structure of the family 13 carbohydrate binding module. *J Mol Biol*. 316:65–78.
- Hagen FK, Hazes B, Raffo R, deSa D, Tabak LA. 1999. Structure-function analysis of the UDP-*N*-acetyl-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase. Essential residues lie in a predicted active site cleft resembling a lactose repressor fold. *J Biol Chem*. 274:6797–6803.
- Hanisch FG, Reis CA, Clausen H, Paulsen H. 2001. Evidence for glycosylation-dependent activities of polypeptide *N*-acetylgalactosaminyltransferases rGalNAc-T2 and-T4 on mucin glycopeptides. *Glycobiology*. 11: 731–740.
- Hassan H, Bennett EP, Mandel U, Hollingsworth MA, Clausen H. 2000. In: Carbohydrates in chemistry and biology—a comprehension handbook. VCH Wiley: p. 273–292.
- Hassan H, Reis CA, Bennett EP, Mirgorodskaya E, Roepstorff P, Hollingsworth MA, Burchell J, Taylor-Papadimitriou J, Clausen H. 2000. The lectin domain of UDP-*N*-acetyl-Dgalactosamine: polypeptide *N*-acetylgalactosaminyltransferase-T4 directs its glycopeptide specificities. *J Biol Chem*. 275:38197–38205.
- Hazes B. 1996. The (QxW)₃ domain: a flexible lectin scaffold. *Protein Sci*. 5:1490–1501.
- Hennet T, Hagen FK, Tabak LA, Marth JD. 1995. T-cell-specific deletion of a polypeptide *N*-acetylgalactosaminyltransferase gene by site-directed recombination. *Proc Natl Acad Sci USA*. 92:12070–12074.
- Imberty A, Piller V, Piller F, Breton C. 1997. Fold recognition and molecular modeling of a lectinlike domain in UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferases. *Protein Eng*. 10:1353–1356.
- Iwasaki H, Zhang Y, Tachibana K, Gotoh M, Kikuchi N, Kwon YD, Togayachi A, Kudo T, Kubota T, Narimatsu H. 2003. Initiation of O-glycan synthesis in IgA1 hinge region is determined by a single enzyme, UDP-*N*-acetyl- α -D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase 2. *J Biol Chem*. 278:5613–5621.
- Kingsley DM, Kozarsky KF, Segal M, Krieger M. 1986. Three types of low density lipoprotein receptor-deficient mutant have pleiotropic defects in the synthesis of N-linked, O-linked, and lipidlinked carbohydrate chains. *J Cell Biol*. 102:1576–1585.
- Kubota T, Shiba T, Sugioka S, Furukawa S, Sawaki H, Kato R, Wakatsuki S, Narimatsu H. 2006. Structural basis of carbohydrate transfer activity by human UDP-GalNAc: polypeptide α -N-Acetylgalactosaminyltransferase (pp-GalNAc-T10). *J Mol Biol*. 359:708–27.
- Mandel U, Petersen OW, Sorensen H, Vedtofte P, Hakomori SI, Clausen H, Dabelsteen E. 1991. Simple mucin-type carbohydrates in oral stratified squamous and salivary-gland epithelia. *J Invest Dermatol*. 97:713–721.
- Mirgorodskaya E, Hassan H, Wandall HH, Clausen H, Roepstorff P. 1999. Partial vapor-phase hydrolysis of peptide bonds: A method for mass spectrometric determination of O-glycosylated sites in glycopeptides. *Anal Biochem*. 269:54–65.
- Pratt MR, Hang HC, Ten Hagen KG, Rarick J, Gerken TA, Tabak LA, Bertozzi CR. 2004. Deconvoluting the functions of polypeptide *N*- α -acetylgalactosaminyltransferase family members by glycopeptide substrate profiling. *Chem Biol*. 11:1009–1016.
- Schwientek T, Bennett EP, Flores C, Thacker J, Hollmann M, Reis CA, Behrens J, Mandel U, Keck B, Shafer MA, et al. 2002. Functional conservation of subfamilies of putative UDP-*N*-acetylgalactosamine: polypeptide *N*-acetylgalactosaminyltransferases in *Drosophila*, *Caenorhabditis elegans*, and mammals—One subfamily composed of I(2)35Aa is essential in *Drosophila*. *J Biol Chem*. 277:22623–22638.
- Sheehan JK, Kirkham S, Howard M, Woodman P, Kutay S, Brazeau C, Buckley J, Thornton DJ. 2004. Identification of molecular intermediates in the assembly pathway of the MUC5AC mucin. *J Biol Chem*. 279: 15698–15705.
- Sorensen AL, Reis CA, Tarp MA, Mandel U, Ramachandran K, Sankaranarayanan V, Schwientek T, Graham R, Taylor-Papadimitriou J, Hollingsworth MA, et al. 2005. Chemoenzymatically synthesized multimeric Tn/STn MUC1 glycopeptides elicit cancer specific anti-MUC1 antibody responses and override tolerance. *Glycobiology*. 2:96–107.
- Stwora-Wojczyk MM, Dzierszynski F, Roos DS, Spitalnik SL, Wojczyk BS. 2004. Functional characterization of a novel *Toxoplasma gondii* glycosyltransferase: UDP-*N*-acetyl-D-galactosamine: polypeptide *N*-acetylgalactosaminyltransferase-T3. *Arch Biochem Biophys*. 426:231–240.
- Takeuchi H, Kato K, Hassan H, Clausen H, Irimura T. 2002. O-GalNAc incorporation into a cluster acceptor site of three consecutive threonines. Distinct specificity of GalNAc-transferase isoforms. *Eur J Biochem*. 269: 6173–6183.

- Ten Hagen KG, Bedi GS, Tetaert D, Kingsley PD, Hagen FK, Balys MM, Beres TM, Degand P, Tabak LA. 2001. Cloning and characterization of a ninth member of the UDPGalNAc: polypeptide N-acetylgalactosaminyltransferase family, ppGaNase-T9. *J Biol Chem.* 276:17395–17404.
- Ten Hagen KG, Fritz TA, Tabak LA. 2003. All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology.* 13:1R–16R.
- Ten Hagen KG, Tetaert D, Hagen FK, Richet C, Beres TM, Gagnon J, Balys MM, Van Wuyckhuysse B, Bedi GS, Degand P, et al. 1999. Characterization of a UDPGalNAc: polypeptide N-acetylgalactosaminyltransferase that displays glycopeptide N-acetylgalactosaminyltransferase activity. *J Biol Chem.* 274:27867–27874.
- Ten Hagen KG, Tran DT. 2002. A UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase is essential for viability in *Drosophila melanogaster*. *J. Biol. Chem.* 277:22616–22622.
- Tenno M, Kezdy FJ, Elhammer AP, Kurosaka A. 2002. Function of the lectin domain of polypeptide N-acetylgalactosaminyltransferase I. *Biochem Biophys Res Commun.* 298:755–759.
- Tenno M, Saeki A, Kezdy FJ, Elhammer AP, Kurosaka A. 2002. The lectin domain of UDPGalNAc: polypeptide N-acetylgalactosaminyltransferase I is involved in O-glycosylation of a polypeptide with multiple acceptor sites. *J Biol Chem.* 277:47088–47096.
- Thobhani S, Ember B, Siriwardena A, Boons GJ. 2003. Multivalency and the mode of action of bacterial sialidases. *J Am Chem Soc.* 125: 7154–7155.
- Topaz O, Shurman DI, Bergman R, Indelman M, Ratajczak P, Mizrahi M, Khamaysi Z, Behar D, Petronius D, Friedman V, et al. 2004. Mutations in GALNT3, encoding a protein involved in O-linked glycosylation, cause familial tumoral calcinosis. *Nat. Genet.* 36:579–581.
- Wandall HH, Hassan H, Mirgorodskaya E, Kristensen AK, Roepstorff P, Bennett EP, Nielsen PA, Hollingsworth MA, Burchell J, TaylorPapadimitriou J, et al. 1997. Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3. *J Biol Chem.* 272:23503–23514.
- White T, Bennett EP, Takio K, Sorensen T, Bonding N, Clausen H. 1995. Purification and cDNA cloning of a human Udp-N-Acetyl-Alpha-D-Galactosamine-Polypeptide N-Acetylgalactosaminyltransferase. *J Biol Chem.* 270:24156–24165.