Order and maximum incorporation of *N*-acetyl-D-galactosamine into threonine residues of MUC2 core peptide with microsome fraction of human-colon-carcinoma LS174T cells

Shin-ichiro IIDA, Hideyuki TAKEUCHI, Kentaro KATO, Kazuo YAMAMOTO, and Tatsuro IRIMURA¹

Laboratory 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

Mucin 2 (MUC2) is the major intestinal mucin. O-glycans are attached to MUC2 in a potentially diverse arrangement, which is crucial for their interaction with endogeneous and exogeneous lectins. In the present report, five oligopeptides [PTTTPITTTT(K), ITTTTIVTPT(K), TVTPTPTPTG(K), PTPTGTQTPT(K) and TQTPTTPTIT(K)] corresponding to portions of the MUC2 tandem repeat domain were synthesized, and incubated with UDP-*N*-acetyl-D-galactosamine (UDP-GalNAc) and detergent-soluble microsomes, prepared from the human colon carcinoma cell line LS174T. The products were fractionated by reverse-phase HPLC and characterized by matrix-assisted laser-desorption ionization–time-of-flight mass spectrometry. Oligopeptides with GalNAc residues derived from PTTTPITTTT(K), containing consecutive threonine residues, were found to be glycosylated with 1–7 GalNAc residues per single peptide. The sequences of all glycopeptides were de-

termined. The results indicated that the predominant sites of the first through to the sixth GalNAc incorporation were Thr³, Thr⁶, Thr⁵, Thr², Thr⁴ and Thr¹, respectively. An exception was the presence of a glycopeptide with three GalNAc residues at Thr¹, Thr⁴ and Thr⁵. Oligopeptides containing alternating threonine residues [TVTPTPTPTG(K) and PTPTGTQTPT(K)] were not fully glycosylated under the same conditions or even after prolonged incubations. Thus, the preferential order and maximum number of GalNAc incorporation into threonine residues of MUC2 core peptides depends on the peptide sequence, when the microsome fraction of LS174T cells is used as a source of *N*-acetyl-D-galactosaminyltransferases.

Key words: mucin, O-glycosylation, tumour antigen, UDP-GalNAc:polypeptide *N*-acetyl-D-galactosaminyltransferase.

INTRODUCTION

Epithelial mucins are characterized by their tandem repeat domains which contain many serine and threonine residues [1,2]. The tandem repeat portion of one of the predominant intestinal mucins, mucin 2 (MUC2), consists of 23 amino acids, including 14 threonine residues [3,4]. These threonine residues are reported to be glycosylated up to 78 % in the LS174T colon carcinoma cell line [5]. Although the tandem repeat domain of MUC2 in normal intestinal epithelia is also thought to be highly O-glycosylated, the glycosylation patterns in different cells and tissues are not well-understood. The sequence of the MUC2 tandem repeat has two characteristic features, consecutive and alternating threonine residues. It has been widely accepted that multivalent display of carbohydrate ligands is favoured by recognition molecules, in a variety of cell-to-cell recognition systems. These include C-type lectins [6], galectins [7] and lectins expressed on the surface of microorganisms [8,9]. Mucin architecture represents an ideal framework for a variety of multiple ligand arrangements that should serve as a recognition unit. Our previous study clearly indicated that multimeric arrangement of *N*-acetyl-D-galactosamine (GalNAc) residues on mucins served as a preferential ligand for a macrophage C-type lectin that recognizes Tn antigen [6]. Mucin O-glycosylation is initiated by the transfer of GalNAc to serine and threonine residues by UDP-GalNAc: polypeptide

GalNAc-transferases (pp-GalNAc-Ts). Five pp-GalNAc-Ts have been identified and characterized to date: T1, T2 [10], T3 [11], T4 [12–14] and T5 [15]. There are many possible pp-GalNAc-T genes encoding enzymes that potentially have different acceptor substrate-specificities. Using various mucin core peptides, Wandall and co-workers suggested that the acceptor substrate-specificities of human pp-GalNAc-T1, -T2 and -T3 were different in their kinetic properties [16]. Although glycosylation of all threonine and serine residues was not achieved by these enzymes, pp-GalNAc-T4 was found to possess a complementary specificity towards the MUC1 sequence. Brockhausen and co-workers reported that the amino acid sequence in oligopeptide substrates affected GalNAc transfer [17]. Gerken and co-workers determined the sequences of proteolytic fragments of porcine submaxillary mucin and concluded that there are preferred amino acid sequences having a high incidence of O-glycosylation [18]. However, little was known about the biological significance of variations in the arrangement of O-glycosylation of intestinal mucins, particularly MUC2 that has consecutive and alternating threonine residues in the tandem repeat domain. Towards this end, we prepared a mixture of glycopeptides with various numbers of attached GalNAc residues using the sequence of the different portions within the MUC2 tandem repeat domain. Microsome fractions from the human colon carcinoma LS174T cell line were used as a source of pp-GalNAc-Ts. Peptides with

Abbreviations used: GalNAc, *N*-acetyl-D-galactosamine; MALDI-TOF, matrix-assisted laser-desorption ionization–time-of-flight; MUC2, mucin 2; pp-GalNAc-T, UDP-GalNAc: polypeptide α -GalNAc-transferase (EC 2.4.1.41); PTH, phenylthiohydantoin; TFA, trifluoroacetic acid.

¹ To whom correspondence should be addressed (e-mail irimura@mol.f.u-tokyo.ac.jp).

various numbers of O-glycosylations were obtained. We made two interesting observations. First, the maximum number of GalNAc incorporation was limited, depending on the peptide sequence. Second, GalNAc incorporation into threonine residues seemed to occur with a strict order. Although a similar study by Hennebicq and co-workers with MUC5AC peptides and microsome fractions of gastric and colonic mucosa indicated that the peptide sequence and source of pp-GalNAc-T influenced the patterns of GalNAc incorporation, the order and maximum number of GalNAc incorporation were not addressed [19–21].

MATERIALS AND METHODS

Synthesis of acceptor substrates

Synthetic oligopeptides used as acceptor substrates were synthesized with a Model 9020 peptide synthesizer (Milligen, Burlington, MA, U.S.A.) according to the general principles of solid-phase methodology. Peptides composed of ten amino acids, corresponding to different portions within the tandem repeat domain of the MUC2 core polypeptide (PTTPTITTTTVP-TPTPTGTQT) [4], were synthesized with a lysine residue attached to each C-terminus. These were PTTPTITTTTK (FM2-1), ITTTTVPPTK (FM2-2), TVTPTPTPTGK (FM2-3), PTP-TGTQTPTK (FM2-4) and TQTPTTTPITK (FM2-5). All peptides were purified by reverse-phase HPLC on columns packed with C_{18} . The fidelity and molecular masses of all peptides were confirmed using an amino acid sequencer (PSQ-1, Shimadzu, Kyoto, Japan) and matrix-assisted laser-desorption ionization–time-of-flight (MALDI–TOF) MS (Voyager™ Elite, Nippon PerSeptive Biosystems, Tokyo, Japan), as described below. The peptides were N-terminally labelled with fluorescein isothiocyanate, in 100 mM Hepes buffer at pH 7.5.

Preparation of microsome fraction as a source of pp-GalNAc-Ts

Human colon carcinoma LS174T cells were cultured in a mixture of Dulbecco's modified Eagle's medium and Ham's F12 medium (1:1, v/v) supplemented with 10% fetal bovine serum. Cells were homogenized in 50 mM Tris/HCl buffer (pH 7.5) containing 250 mM sucrose, 1 µg/ml aprotinin (Sigma, St. Louis, MO, U.S.A.), 1 µg/ml leupeptin (Peptide Institute Inc., Osaka, Japan) and 0.5 µg/ml pepstatin A (Sigma). After centrifugation of this homogenate at 3000 g for 10 min at 4 °C, the decanted supernatant was centrifuged at 100 000 g for 1 h. The pellet was suspended in the buffer used during the homogenization step containing an additional 0.1% Triton X-100 (Sigma). Protein concentrations were determined using a bicinchonic acid ('BCA') protein assay kit (Pierce Chemical Co., Rockford, IL, U.S.A) with BSA as a standard. The solutions were aliquoted and stored at –80 °C until use.

Assays of GalNAc transferase activity

The standard enzyme reaction mixture consisted of 20 mM Hepes buffer (pH 7.5), 5 mM $MnCl_2$, 1 mM UDP-GalNAc (Sigma), 0.5 mM synthetic oligopeptide and 100 µg of microsomes, in a final volume of 100 µl. Reactions were performed at 37 °C for up to 18 h and were terminated by addition of 300 µl of 10 mM EDTA. Additional enzyme (100 µg of protein) and UDP-GalNAc were added after 6 h.

In time-course experiments UDP-[3H]GalNAc (0.1 mCi, Dupont-NEN, Boston, MA, U.S.A.) was used. Reaction mixtures were applied to columns packed with C_{18} -silica, washed with solvent A [0.05% trifluoroacetic acid (TFA) in water], and eluted with solvent B (0.05% TFA, 35% propanol, and 15%

acetonitrile in water). Radioactivity incorporated into the peptides was measured with a liquid-scintillation counter using Ready Cap (Beckman, Fullerton, CA, U.S.A.).

Characterization of the glycosylated peptides

The glycosylated peptides were separated by reverse-phase HPLC (JASCO, Tokyo, Japan). A COSMOSIL column (5C₁₈-AR, 10 mm × 250 mm, Nacalai Tesque, Kyoto, Japan) was used. The column was eluted with a linear gradient ranging from 0 to 50% solvent B in solvent A, at a flow rate of 2 ml/min, for 30 min. Eluates were excited at 492 nm and fluorescence emission was monitored at 520 nm.

MALDI–TOF MS analysis of glycosylated peptides

Concentrated glycosylated peptides (1–10 pmole) were applied to a stainless steel plate and mixed with a 10 mg/ml solution of α -cyano-4-hydroxycinnamic acid dissolved in 0.1% TFA/50% ethanol in water. All mass spectra were obtained using a Voyager™ Elite mass spectrometer in the linear mode with the delayed extraction setting. Acceleration voltage was set at 20 000, grid voltage at 93.5 and guide wire voltage at 0.05%. Laser intensity was chosen so that optimum resolution could be obtained. Relative molecular mass numbers were determined using angiotensin I as a standard.

Amino acid sequencing

Pulsed liquid Edman degradation amino acid sequencing of glycopeptides was performed on the Applied Biosystems 490 Procise protein sequencing system (Perkin Elmer, Norwalk, CT, U.S.A.). With this system, a phenylthiohydantoin (PTH) derivative of threonine_{GalNAc} was identified as a pair of peaks eluting near the positions of PTH-Ser and PTH-Thr [18]. Amino acid sequencing of a fully glycosylated peptide (PT_{GalNAc}T_{GalNAc}T_{GalNAc}P) confirmed these eluting positions (results not shown).

RESULTS

Glycosylation of acceptor substrates having consecutive threonine residues (FM2-1 peptide)

Five overlapping oligopeptides, FM2-1, FM2-2, FM2-3, FM2-4 and FM2-5, corresponding to the MUC2 tandem repeat domain, were synthesized in order to investigate the variations in the mode of attachment of GalNAc residues. These peptides contained either consecutive or alternating threonine residues. We investigated the maximum number and the preferential order of GalNAc incorporation. Representative HPLC chromatographs for glycosylated FM2-1 peptides, which contained three and four consecutive threonine residues, are shown in Figure 1. In these experiments, incubation times with the enzyme source were 3, 6 and 18 h. The unglycosylated FM2-1 peptide eluted at a retention time of 26.8 min after the sample injection (peak g). In cases of incubations without UDP-GalNAc, additional peaks were detected at retention times of 27.4 and 27.8 min. They probably represent degradation products of the peptide (molecular masses of 706 and 904 were estimated by MALDI–TOF MS, results not shown) derived from the unglycosylated peptides. Glycosylated peptides were eluted in fractions at retention times between 23 and 26 min (Figure 1, peaks a–f). Peak a was observed in the incubation mixture after 6 h but not after 3 h. The relative intensity of peaks a, b and c increased depending on the incubation time, whereas the relative intensity of peak g

Figure 1 Separation of glycosylated FM2-1 peptide corresponding to MUC2 by reverse-phase HPLC

FM2-1 peptides were incubated with detergent-soluble microsomes, prepared from the human colon carcinoma cell line LS174T. They were subsequently separated by reverse-phase HPLC on a C_{18} column. Elution profiles after 3, 6 and 18 h incubation periods are shown. Peaks a–g represent the major peaks.

decreased. Each peak was separately pooled and concentrated by evaporation before being subjected to MS analysis. A five-fold excess of heat-inactivated microsomes was added to the incubation mixture containing FM2-1 and the glycosylated products at the 16.5 h time point, and was subsequently analysed by HPLC. There was no difference detected in the elution profiles, indicating that the microsomes did not contain a significant amount of endogenous acceptor substrates.

MALDI-TOF MS analysis of glycosylated peptides

As shown in Figure 2, the molecular mass of the separated peptides was characterized by MALDI-TOF MS. The molecular mass of the unglycosylated FM2-1 peptide was identified to be 1551.8, the $[M+H]^+$ ion, and the molecular mass of the peptide eluted at a retention time of 25.8 min (peak f) was 1754.6, indicating that a single GalNAc residue (203 units) was attached to the FM2-1 peptide. In a similar manner, MALDI-TOF MS analysis revealed the molecular mass of the glycosylated peptide present in each fraction (peaks a–e), that clearly corresponded to the number of GalNAc residues attached to the peptide. For example, peak e gave a molecular mass of 1958.3, indicating that two GalNAc residues were attached to the peptide. Peaks d, c and b gave molecular masses of 2161.0, 2364.8 and 2567.6, respectively, indicating that three, four and five GalNAc residues were attached to the peptides. Peak a gave two peaks having molecular masses of 2771.3 and 2974.6, indicating that peak a

HPLC peak

Figure 2 Profiles of MALDI-TOF MS of the glycosylated FM2-1 peptide separated by reverse-phase HPLC

Mass indicates $[M+H]^+$ form.

Figure 3 For legend see opposite page.

contained two components, corresponding to two peptides with six and seven GalNAc residues, respectively. An increase in the number of GalNAc residues attached made the retention time shorter. When the FM2-1 peptide (PTTTPITTTTK) was used as an acceptor, the maximum number of GalNAc residues were transferred in the presence of the microsome fraction prepared from LS174T cells.

Identification of *N*-acetylgalactosaminylated threonine residues

Glycopeptides corresponding to peaks a–g after 6 h incubations were subjected to amino acid sequence analysis. Although the results of the Edman degradation study are difficult to interpret, due to the cycle-to-cycle carryover of PTH-amino acids, our results were interpreted based on an assumption that the

Figure 3 Profiles of amino acid sequencing chromatograms of the MUC2 peptide FM2-1 and its derivatives, with incorporated GalNAc residues

Materials corresponding to peaks from reverse-phase HPLC (6 h incubation) were analysed on the Applied Biosystems 490 Procise protein sequencing system. Untreated FM2-1 and FM2-1 with 1–6 GalNAc residues were analysed. These correspond to peaks g, f (G1), e (G2), d (G3 and G3'), c (G4) and b (G5). Materials in peak a were separated into two components by re-chromatography with an analytical C_{18} column and the second peak was sequenced (G6). Materials in peak d were further separated by additional HPLC fractionation and analysed separately (G3 and G3'). G4' was obtained, as a peak having a similar retention time as peak c, when the incubation of peptide with the enzyme was extended to 60 h. All other components remained as single peaks on C_{18} columns after separation under a variety of conditions. Asterisks indicate putative peaks of PTH derivatives of α -GalNAc-Threonine (T_{GalNAc}). Other amino acids are indicated by the one letter abbreviations.

Figure 4 Schematic representation of the putative pathways of GalNAc incorporation into a MUC2 peptide FM2-1, PTTPITTTTK, incubated with a microsomal fraction of LS174T human colon carcinoma cells

According to the results of peptide sequencing, the majority of the peptide was estimated to be processed through the pathway drawn on the left-hand side. Although G3' was detected in an almost equal quantity as G3, G1' and G4' it was identified only in a trace amount after prolonged incubation periods for 60 h. The sequence of G1' was not conclusive. Therefore, the sequences of G1' and G2' should be considered as putative sequences and are shown in parentheses.

materials obtained after the C_{18} reverse-phase separation were not contaminated with glycopeptides containing different numbers of GalNAc residues. As shown in Figure 3, the fractions eluted as peaks f and e on a C_{18} column were identified as $\text{PTTT}_{\text{GalNAc}}\text{PITTTTK}$ (G1) and $\text{PTTT}_{\text{GalNAc}}\text{PITTT}_{\text{GalNAc}}\text{TK}$ (G2), respectively. The fraction corresponding to peak d was separated into two components when applied to an analytical C_{18} instead of a preparative column. The analytical column was also eluted with a less steep gradient. The two fractions were separately processed for amino acid sequence analysis. The first peak was composed mainly of $\text{PTTT}_{\text{GalNAc}}\text{PITTT}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{TK}$ (G3) (Figure 3), whereas the second peak corresponded to $\text{PT}_{\text{GalNAc}}\text{TTPIT}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{TTK}$ (G3') (Figure 3). The relative quantity of these two components was approximately the same, according to the fluorescence detector responses. Peak c could not be separated into two components when applied to an analytical C_{18} column, as indicated for peak d. Sequence analysis indicated that peak c corresponded to $\text{PTT}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{PITTT}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{TK}$ (G4). However, a peak eluted at a similar position, when the incubation time was extended to 60 h and was shown to have a sequence $\text{PT}_{\text{GalNAc}}\text{TTPIT}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{TK}$, designated as G4' in Figure 3. Analysis of peak b indicated that it was composed mainly of $\text{PTT}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{PIT}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{TK}$ (G5). The second peak from re-chromatography of peak a was $\text{PT}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{PIT}_{\text{GalNAc}}\text{T}_{\text{GalNAc}}\text{TK}$ (G6). These

results strongly suggest that GalNAc transfer to the mucin core peptide, with multiple potential glycosylation sites, is not a random process but a process with preferential orders. The putative pathway is shown schematically in Figure 4.

O-glycosylation of other acceptor substrates

Results of the analysis of five acceptor peptides following incubation with the microsomal fraction from LS174T cells and UDP-GalNAc are shown in Table 1. After incubation for 6 h all glycopeptides were analysed using reverse-phase HPLC. Three peaks were obtained from FM2-2 (ITTTTITPTTK), three peaks from FM2-3 (TVTPTPTPTGK), only one peak from FM2-4 (PTPTGTQPTTK) and four peaks from FM2-5 (TQTPITTPITK), in addition to each acceptor peptide. Even after prolonged incubation for 18 h with additional enzyme sources, no additional peaks were observed.

We separately pooled each peak fraction and subjected them to MALDI-TOF MS analysis. A peak derived from FM2-2 that eluted at 26.2 min gave molecular masses of 2185.0 and 2388.1. This apparently corresponded to the peptide with three and four GalNAc residues attached in $[M+\text{Na}]^+$ forms. We also determined the molecular mass of the glycopeptide in the other fractions. The results indicated that a maximum of four of six threonine residues were substituted in FM2-2. In FM2-3, a maximum of three of five threonine residues were substituted. The apparent substitution with GalNAc in FM2-4 was two of five threonine residues, and in FM2-5, four of six threonine residues were apparently substituted with GalNAc. All threonine residues were glycosylated only when the FM2-1 acceptor peptide was used among the synthetic peptides we tested. For both FM2-2 and FM2-3 separated on C_{18} columns, two peaks having different retention times gave identical molecular mass numbers that corresponded to peptides with single GalNAc residues. Each pair of peaks probably represents peptides substituted by single GalNAc residues at different positions. Because of the relatively low recovery rate of Edman degradation of proline-rich peptides, we were not able to determine the sequence of glycosylated FM2-2 and FM2-3.

Comparison of the time-course of GalNAc-transfer to different peptides

The initial rate of GalNAc incorporation into the five peptides (FM2-1, -2, -3, -4 and -5) in the presence of microsomes was compared (Table 2). FM2-3 (TVTPTPTPTGK) showed the highest rate of GalNAc incorporation during the first hour, as shown by the low K_m value. The initial rate of GalNAc incorporation into FM2-1, the peptide on which all threonine residues are potentially substituted, was low. FM2-2 was shown to have an intermediate level of GalNAc incorporation, between FM2-3 and FM2-1. Although the FM2-4 peptide had alternating threonine residues, similar to FM2-3, the rate of GalNAc incorporation was not as high as for FM2-3.

DISCUSSION

MUC2 is the major insoluble mucin lining the luminal surface of the small and large intestines in humans [4]. MUC2 serves as a binding site for infectious microorganisms and lectins. Both of these interactions require clusters of GalNAc residues [6,8,22,23]. In this report, five different oligopeptides, representing different portions of MUC2, were synthesized and used to assess whether a variety of O-glycosylation could be achieved with pp-GalNAc-Ts expressed in human colon carcinoma cells. Although a

Table 1 Summary of the number of GalNAc residues attached to five different synthetic peptides corresponding to different portions of MUC2 tandem repeats

Acceptor	HPLC retention time (min)	MALDI-TOF MS molecular mass (<i>m/z</i>)	No. of GalNAc residues attached	No. of GalNAc residues attached/total no. of Thr residues
FM2-1 PTTTPITTTTK	23.3	2974.6	7	7/7
	23.3	2771.3	6	
	23.6	2567.6	5	
	24.4	2364.8	4	
	24.6	2161.0	3	
	25.0	1958.3	2	
	25.8	1754.6	1	
	26.8	1551.8	0	
FM2-2 ITTTTIVTPTK	26.2	2388.1*	4	4/7
	26.2	2185.0*	3	
	26.8	1960.4	2	
	27.3	1756.8	1	
	27.6	1756.6	1	
	28.1	1553.8	0	
FM2-3 TVTPTPTPTGK	24.5	2121.2*	3	3/5
	24.5	1918.0*	2	
	25.1	1692.0	1	
	25.5	1692.2	1	
	25.9	1489.4	0	
FM2-4 PTPTGTQTPTK	23.0	1925.4	2	2/5
	23.0	1722.3	1	
	23.3	1518.5	0	
FM2-5 TQPTTTTPTK	24.2	2392.0	4	4/6
	24.6	2189.0	3	
	25.2	1985.8	2	
	25.7	1782.3	1	
	26.3	1579.0	0	

* denotes $[M + Na]^+$ form**Table 2** Initial rate of incorporation of GalNAc into five different synthetic peptides corresponding to different portions of MUC2 tandem repeats

Peptide	Initial rate of GalNAc incorporation (nmol/h/mg)
FM2-1	0.37
FM2-2	1.02
FM2-3	2.54
FM2-4	1.04
FM2-5	0.36

previous study by Nishimori and co-workers, as well as Hanisch and co-workers, used the MUC1 core peptide, it does not contain consecutive or alternating O-glycosylation sites [24,25]. A preliminary attempt with the MUC2 core peptide has been published by Inoue and co-workers [26], but a systematic approach was not applied. Hennebicq and co-workers studied GalNAc incorporation into the MUC5AC motif, TTSAPPTS, to show the specificity was regulated both by enzyme sources (different organs) and peptide sequences [19–21]. We previously used a peptide, PTTTPLK, mimicking a tandem repeat portion of MUC2, to examine how GalNAc was incorporated into three consecutive threonine residues by pp-GalNAc-T1, -T2 and -T3 [27]. These enzymes showed different orders of GalNAc incorporation into this peptide. The maximum numbers of GalNAc incorporation were 2, 1 and 3, respectively [27]. The present

report is designed to investigate the order and the maximum incorporation of GalNAc into peptides corresponding to different portions of MUC2 tandem repeats using microsomes prepared from human colon carcinoma LS174T cells.

Interestingly, FM2-1 (PTTTPITTTTK), which included consecutive threonine residues, was fully glycosylated under the conditions used in this study. In contrast, peptides with alternating threonine residues (FM2-4 peptide, PTPTGTQTPTK and FM2-3 peptide, TVTPTPTPTGK) were not fully glycosylated even after prolonged incubation periods with additions of the enzyme source and UDP-GalNAc. Time-course studies showed that the fastest initial GalNAc incorporation was observed with the FM2-3 peptide (TVTPTPTPTGK), not with the FM2-1 peptide (PTTTPITTTTK) or the FM2-4 peptide (PTPTGTQTPTK). These results indicated that a sequence of alternating threonine and proline residues was the preferred substrate for rapid glycosylation when the microsome fraction of LS174T cells was used as the source of pp-GalNAc-T, which might be reminiscent of a preferred motif for bovine pp-GalNAc-T1 [28]. However, the initial site of GalNAc incorporation, indicated by sequence analysis of glycosylated FM2-1 peptides, was virtually identical to the site generated on a peptide PTTTPLK by the action of pp-GalNAc-T3 [27]. Therefore, pp-GalNAc-T3 appeared to play a predominant role when FM2-1 was used as a substrate. However, relative contributions of different pp-GalNAc-Ts possibly present in the microsome fraction of LS174T cells, in the formation of the various glycosylated peptides generated in the present study need to be evaluated further.

Additional products G1', G3' and G4' from FM2-1, shown as the minor pathway, could be explained by the contribution of

pp-GalNAc-T1, as judged from our previous study with PTT-TPLK and recombinant pp-GalNAc-T1 [27]. Using recombinant human pp-GalNAc-T1, -T2 and -T3 and MUC2 peptide (PT-TTPISTTTMTPTPTPTC), kinetic parameters were compared by Wandall and co-workers [16]. They found that the $V_{\max}$ of the peptide with both pp-GalNAc-T1 and -T3 was relatively high. Their results might be due to the fact that this peptide contained both consecutive and alternating threonine residues as substrates. It should also be noted that pp-GalNAc-T1, -T2 or -T3 alone was not able to achieve the formation of G6 or G7 (Kato, K., Takeuchi, H., Iida, S., Hassan, H., Clausen, H. and Irimura, T., unpublished work). Thus, LS174T cells may contain additional pp-GalNAc-Ts. According to preliminary observations [29], LS174T cells express pp-GalNAc-T1, -T2, -T3, -T4 and -T6. It would be interesting to test a mixture of these enzymes and observe whether it shows the same order and maximum number of GalNAc incorporation into the peptides used in this study.

Peptides with a single GalNAc residue possibly attached to different threonine residues were identified after HPLC separation when using FM2-2 or FM2-3 peptides as substrates (Table 1). Besides these glycopeptides and G1', G3' and G4' of FM2-1, the glycosylated peptide products with a given number of attached GalNAc residues obtained in this study gave single peaks on reverse-phase HPLC equipped with a C_{18} column under a variety of conditions as far as we tested. By the use of MUC2 tandem repeat domains and microsome fractions of LS174T cells, the order of GalNAc incorporation and the arrangement of GalNAc residues on the core polypeptides was very strict and did not generate a library with diverse components. Particularly, in the case of FM2-1 that contains seven threonine residues, only one predominant pathway was strongly suggested as shown in the left-hand side of Figure 4. An additional minor pathway seemed to be present as shown in the right-hand side of Figure 4. It is not clear whether there was further glycosylation of G4'.

In conclusion, we have shown in this paper that (a) the order of GalNAc incorporation into a portion of MUC2 tandem repeat containing three and four consecutive threonine residues (FM2-1) was strictly regulated when a microsome fraction of LS174T cells was used as the source of pp-GalNAc-T in cell-free systems. (b) Not all threonine residues in MUC2 peptides with alternating threonine residues (FM2-3 and FM2-4) incorporate GalNAc under the same conditions. The unique order and maximum number of GalNAc incorporated into each peptide could be due to the nature of the cells from which the enzyme source was prepared i.e. a combination of different pp-GalNAc-Ts may have been present.

This work was supported by grants-in-aid from the Ministry of Education, Science, Sports and Culture of Japan (05274101, 05557104, and 07407063), the Ministry of Health and Welfare, the Japan Health Science Foundation, the Research Association for Biotechnology, the New Energy and Industrial Technology Development Organization, and Program for Promotion of Basic Research Activities for Innovative Biosciences. We also thank Dr David M. Wildrick (The University of Texas MD

Anderson Cancer Center, Houston, TX, U.S.A.) for his editorial assistance and Ms Chizu Hiraiwa for her assistance in preparing this manuscript.

REFERENCES

- Gendler, S. J. and Spicer, A. P. (1996) *Annu. Rev. Physiol.* **57**, 607–634
- Kim, Y. S., Gum, Jr., J. R. and Brockhausen, I. (1996) *Glycoconjugate J.* **13**, 693–707
- Gum, Jr., J. R., Byrd, J. C., Hicks, J. W., Toribara, N. W., Lampert, D. T. and Kim, Y. S. (1989) *J. Biol. Chem.* **264**, 6480–6487
- Gum, Jr., J. R., Hicks, J. W., Toribara, N. W., Siddiki, B. and Kim, Y. S. (1994) *J. Biol. Chem.* **269**, 2440–2446
- Byrd, J. C., Nardelli, J., Siddiqui, B. and Kim, Y. S. (1988) *Cancer Res.* **48**, 6678–6685
- Iida, S., Yamamoto, K. and Irimura, T. (1999) *J. Biol. Chem.* **274**, 10697–10705
- Bresalier, R. S., Byrd, J. C., Wang, L. and Raz, A. (1996) *Cancer Res.* **56**, 4354–4357
- Adler, P., Wood, S. J., Lee, Y. C., Lee, R. T., W. A. Retri, J. and Schnaar, R. L. (1995) *J. Biol. Chem.* **270**, 5164–5171
- Belley, A., Keller, K., Grove, J. and Chadee, K. (1996) *Gastroenterology* **111**, 1484–1492
- White, T., Bennett, E. P., Takio, K., Sørensen, T., Bonding, N. and Clausen, H. (1995) *J. Biol. Chem.* **270**, 24156–24165
- Bennett, E. P., Hassan, H. and Clausen, H. (1996) *J. Biol. Chem.* **271**, 17006–17012
- Clausen, H. and Bennett, E. P. (1996) *Glycobiology* **6**, 635–646
- Hagen, F. K., Ten Hagen, K. G., Beres, T. M., Balys, M. M., Van Wuyckhuysse, B. C. and Tabak, L. A. (1997) *J. Biol. Chem.* **272**, 13843–13848
- Bennett, E. P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M. A., Merckx, G., van Kessel, A. G. et al. (1998) *J. Biol. Chem.* **273**, 30472–30481
- Ten Hagen, K. G., Hagen, F. K., Balys, M. M., Beres, T. M., Van Wuyckhuysse, B. and Tabak, L. A. (1998) *J. Biol. Chem.* **273**, 27749–27754
- Wandall, H. H., Hassan, H., Mirgorodskaya, E., Kristensen, A. K., Roepstorff, P., Bennett, E. P., Nielsen, P. A., Hollingsworth, M. A., Burchell, J., Taylor-Papadimitriou, J. and Clausen, H. (1997) *J. Biol. Chem.* **272**, 23503–23514
- Brockhausen, I., Toki, D., Brockhausen, J., Peters, S., Bielfeldt, T., Kleen, A., Paulsen, H., Meldal, M., Hagen, F. and Tabak, L. A. (1996) *Glycoconjugate J.* **13**, 849–856
- Gerken, T. A., Owens, C. L. and Pasumath, M. (1997) *J. Biol. Chem.* **272**, 9707–9719
- Hennebicq, S., Tetaert, D., Soudan, B., Briand, G., Richet, C., Demeyer, D., Gagnon, J., Petillot, Y. and Degand, P. (1998) *Biochimie* **80**, 69–73
- Hennebicq, S., Tetaert, D., Soudan, B., Boersma, A., Briand, G., Richet, C., Gagnon, J. and Degand, P. (1998) *Glycoconjugate J.* **15**, 275–282
- Soudan, B., Tetaert, D., Hennebicq, S., Briand, G., Zermach, F., Richet, C., Demeyer, D., Gagnon, J., Petillot, Y. and Degand, P. (1998) *J. Peptide Res.* **51**, 346–354
- Chadel, K., Johnson, M. L., Orozco, E., Petri, W. A. and Ravdin, J. I. (1988) *J. Infect. Dis.* **158**, 398–406
- Suzuki, N., Yamamoto, K., Toyoshima, S., Osawa, T. and Irimura, T. (1996) *J. Immunol.* **156**, 128–135
- Nishimori, I., Johnson, N. R., Sanderson, S. D., Perini, F., Mountjoy, K., Cerny, R. L., Gross, M. L. and Hollingsworth, M. A. (1994) *J. Biol. Chem.* **269**, 16123–16130
- Hanisch, F.-G., Müller, 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
- Inoue, M., Yamashina, I. and Nakada, H. (1998) *Biochem. Biophys. Res. Commun.* **245**, 23–27
- Iida, S., Takeuchi, H., Hassan, H., Clausen, H. and Irimura, T. (1999) *FEBS Lett.* **449**, 230–234
- Yoshida, A., Suzuki, M., Ikenaga, H. and Takeuchi, M. (1997) *J. Biol. Chem.* **272**, 16884–16888
- Mandel, U., Hassan, H., Bennett, E. and Clausen, H. (1999) *Glycoconjugate J.* **16**, S77

Received 6 September 1999/7 December 1999; accepted 11 February 2000