Glycobiology vol. 7 no. 7 pp. 881–896, 1997

A sequencing strategy for the localization of O-glycosylation sites of MUC1 tandem repeats by PSD-MALDI mass spectrometry

Steffen Goletz[5], Bernd Thiede, Franz-Georg Hanisch[1], Michael Schultz[2], Jasna Peter-Katalinic[3], Stefan Müller[1], Oliver Seitz[4] and Uwe Karsten

Max-Delbrück-Centrum für Molekulare Medizin, Robert-Rössle Strasse 10, D-13125 Berlin-Buch, Germany. [1]Institut für Biochemie, Universität Köln, D-50924 Köln, Germany, [2]Institut für Organische Chemie, Universität Mainz, D-55099 Mainz, Germany, [3]Institut für Medizinische Physik und Biophysik, Universität Münster, D-48149 Münster, Germany, and [4]Scripps Research Institute, La Jolla, CA 92037, USA

[5]To whom correspondence should be addressed

It is demonstrated with glycopeptides of the polymorphic epithelial mucin (MUC1) that post-source decay matrix-assisted laser desorption ionization (PSD-MALDI) is a fast, highly sensitive, and reproducible method for the localization of O-glycosylation sites by reflectron time-of-flight (TOF) mass spectrometry. We have analyzed GalNAc-carrying peptides of up to 25 amino acids, and could distinguish even neighboring glycosylation sites. This method was also able to localize and characterize disaccharides (e.g., the Thomsen-Friedenreich disaccharide) on MUC1 derived peptides. PSD-MALDI-MS fragment ion patterns were recorded in the positive ion mode from the synthetic peptide TAP25 [(T^{1a}APPAHGVT^{9}S^{10}APDT^{14}RPAPGS^{20}) T^{1b}APPA], an overlapping sequence of MUC1 tandem repeats, which was glycosylated with GalNAc *in vitro*. The glycosylation sites found were either Thr9 or Thr1b in the monoglycosylated, Thr9 and Thr1b in the diglycosylated, and Thr9, Thr1b, and Ser20 in the triglycosylated peptide. A single PSD-MALDI-MS spectrum of the underivatized and uncleaved di- or triglycosylated TAP25 peptide was sufficient to identify the glycosylation sites, thereby distinguishing six potential, partly adjacent, glycosylation sites. The monoglycosylated fraction was found to consist of a mixture of two glycosylated species with the same molecular weight. This was shown by the analysis of proteolytic digests. PSD-MALDI-MS of the resulting peptides right out of the digestion probe was sufficient to identify the GalNAc-glycosylation sites as either Thr9 or Thr1b, respectively. Beyond the methodical aspects the results revealed that *in vitro* glycosylation of the TAP25 peptide with a transferase system from human milk differs from that obtained with a breast cancer cell transferase system.

Key words: glycosylation sites/O-glycosylation/PSD-MALDI-MS/MUC1/mucins

Introduction

The muc1 gene product is a high molecular mass glycoprotein that is expressed by normal and malignant cells of many human glandular epithelia. About 50% of the molecular mass of this polymorphic epithelial mucin found in human milk is made up of O-linked oligosaccharides which are mainly attached to a large domain of tandem repeats of 20-amino acid peptides (Gendler *et al.*, 1990). Tumor-associated mucins carry often aberrantly synthesized shorter oligosaccharide chains which alter the immunoreactivity of the tandem repeat core peptide (Burchell *et al.*, 1987; Hanisch *et al.*, 1989, 1996; Hull *et al.*, 1989; Brockhausen *et al.*, 1995) probably by changing the peptide conformation and/ or the degree of steric hindrance by the carbohydrates. The enhanced accessibility of the peptide core for antibodies reactive to the PDTRP-motif may also be explained by changes of the site-specific glycosylation patterns. In contrast to N-glycosylation, there is no consensus sequence known for O-glycosylation by UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase (GalNAc-transferase), the first step in the synthesis of mucin-type oligosaccharides, which transfers N-acetylgalactosamine (GalNAc) to Thr or Ser residues. However, the selection of binding sites appears not to be random. Results from several groups suggest that in many cases the presence of Pro and certain other amino acids at certain positions in the neighborhood of Ser or Thr and the absence of charged residues may predispose these sites for O-glycosylation (Wilson *et al.*, 1991; O'Connell *et al.*, 1992; Pisano *et al.*, 1993, 1994). There may exist distinct threonine- and serine-specific UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases (O'Connell *et al.*, 1991; Eckhardt *et al.*, 1991) or enzymes which can form both GalNAc-O-Thr and GalNAc-O-Ser (Wang *et al.*, 1993).

In order to define the actual status of glycosylation, a highly sensitive method is required for the analysis of O-glycosylation sites of peptides derived from mucins *ex vivo*. Preferably this method should work even with peptide mixtures and a certain amount of solvent impurities in order to circumvent extensive purification steps. One problem in determining the glycosylated amino acids of mucins/MUC1 is their high number of potential glycosylation sites. MUC1 has two Ser and three Thr residues per repeating unit. Two serines are neighboring threonines, an analytical problem which can not be resolved by common endoproteinase digestions in combination with MALDI-MS or other MS techniques.

MALDI-MS is the most sensitive mass spectrometrical method for underivatized polymers, but standard instrumentation lacks the ability for structural analysis. The recently established analysis in the reflectron mode of post-source decay (PSD) fragments overcomes this problem (Spengler *et al.*, 1992), and it has been shown that PSD-analysis of MALDI-generated polymer ions give valuable structural information for peptides, cross-linked peptides, carbohydrates, and DNA (Spengler *et al.*, 1993, 1995; Kaufmann *et al.*, 1994; Machold *et al.*, 1995; Talbo and Mann, 1996). PSD-MALDI-MS analyzes product ions from metastable decay of parent ions which takes place in the first field-free drift path of the reflectron time-of-flight mass spectrometer. It has been shown that a

© Oxford University Press

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

complete sequence analysis for peptides up to 26 amino acids in the femtomolar range can be achieved by this technique (Kaufmann *et al.*, 1993).

Here we show that PSD-MALDI-MS is a fast, highly sensitive, reproducible, and comparably easy technique for the localization of glycosylation sites on glycopeptides, able to distinguish even neighboring sites which can not be resolved by conventional endoproteinase treatment in combination with mass spectrometry. Additionally, PSD-MALDI-MS can be used to localize and characterize disaccharide side chains according to their masses. The method is relatively insensitive to solvent impurities, and is applicable to glycopeptide mixtures by choosing one molecular parent ion by deflection of all other ions using a pulsed field (precursor selection).

Results

Localization of the glycosylation site of MUC1–9AA-GalNAc by PSD-MALDI-MS

The synthetic glycopeptide MUC1–9AA-GalNAc (Ac-APDT(αGalNAc)RPAPG-OH; theoretical $M+H^+$ = 1126.95; Seitz and Kunz, 1995) was analyzed to show that PSD-MALDI-MS is able to localize the glycosylation site and to gain some general information about the fragment ion pattern of MUC1-derived glycopeptides, e.g., concerning major and minor peaks and subfragmentation known to be caused by cleavage of Xaa-Pro (proline fragments) with subsequent unidirectional internal fragmentation. Standard conditions (see *Materials and methods*) and the matrix CCA (α-cyano-4-hydroxycinnamic acid) were used for the analyses. Repetitive analyses showed no obvious differences in the quality and quantity of fragment ions. Mass accuracy for most of the peaks was better than 0.4 *m/z*, which allowed, together with the occurring companions (e.g. a_m, a_m-17, b_m, b_m-17), an exact assignment of the fragment ions. More than 90% of the peaks could be reliably assigned, and the remaining nonassigned peaks were minor peaks. Table I lists the calculated fragment ion masses, the observed masses, and the PSD-MALDI-MS fragment ions assigned according to the Biemann modification (1988) of the Roepstorff and Fohlmann nomenclature (1984). The fragment ions were labeled as shown in Figure 1B with the N-terminal series ions (a_m, b_m, c_m) and the C-terminal series ions (x_n, y_n, z_n; n = 9-m) when un- or deglycosylated, and the corresponding fragment ions +S (for HexNAc; +203 u) when glycosylated. The only exceptions to the Biemann nomenclature are the fragment ions originating from the nonreducing end of the sugar which were labeled B^+ and C^+ according to the nomenclature of Domon and Costello (1988; Figure 1B). Figure 1A shows the PSD-MALDI-MS spectrum with the observed masses. The schematic presentation in Figure 1C shows the occurring fragment ions sorted into C- and N-terminal ions, glycosylated and nonglycosylated ions, and lined up with the glycopeptide (primary) structure in order to show the data in a condensed fashion, and to facilitate their interpretation by the reader. These schemes are sufficient to determine the localization of the glycosylation site. Internal fragment ions derived from the subfragmentation of proline fragments and the main fragment are not included in the schemes in Figure 1C since they are not necessary for the determination of the glycosylation site (however, C- and N-terminal ions which overlap with these internal fragment ions or with other C- and N-terminal ions are indicated in the scheme within brackets). The y-fragment ions of nonglycosylated and glycosylated species of the C-terminal series led to the allocation of GalNAc to the threonine (in this case the only possible glycosylation site), which was confirmed by the rather complete N-terminal series of b- and a-fragment ions. Y-fragment ions sometimes coincided with y-18 species, which occurred at a lower intensity and originated from the loss of water. Additionally, y-17 fragment ions were found which originated from the loss of NH_3 from arginine or from the z-ion cleavage of the peptide backbone (*m/z* of z_n = *m/z* of y_n-17), and some minor x-fragment ion peaks. B-ions were often accompanied by a-ions which were mainly less abundant. B- and a-ions had both frequently -17 companions which resulted from the loss of NH_3 but also some -18 companions due to the loss of water. To distinguish between a -17 or -18 loss is sometimes not possible. The -17 companions showed, compared to the b- and a-ions, smaller fragment ion peaks for peptide stretches without arginine but higher peaks for fragments including Arg (especially subfragment ions from the proline P8 fragment). This increase in -17 companions of b- and a-ions was due to the loss of NH_3 from arginine. The main fragment ion (main fragment) was the y_6 which resulted from the cleavage at the C-terminal side of aspartic acid, a known preferred site of backbone cleavage (Kaufmann *et al.*, 1993; Yu *et al.*, 1993). We found abundant N-terminally directed internal fragment ions of the proline fragments, especially from Pro2 as well as internal fragment ions originating from subfragmentation of the main fragment y_6 (Figure 1, Table I).

Glycosylated fragments appeared as glycosylated fragment ions but also as deglycosylated species. In the C-terminal series, the y-series, the intensity of the deglycosylated fragment ions was much lower (10–15%) than that of their respective glycosylated forms. The glycosylated species of the N-terminal series, however, showed fragment ion peaks of about the same size or even lower ones in comparison to their deglycosylated species. However, we could not identify glycosylated forms of internal fragment ions from proline fragments and only few from the main fragment (Figure 1A, Table I). Two additional fragment ions were detected which originated from the cleavage of the sugar HexNAc of the peptide backbone, one major peak at *m/z* 204.5, indicating the B^+-ion (= $[B]^+$) and resulting from the simple cleavage of the protonated glycosidic bond, and one minor peak at *m/z* 222.6, indicating the C^+-ion (= $[C+2H]^+$; Domon and Costello, 1988).

To determine the sensitivity of PSD-MALDI-MS for glycoprotein analysis, we recorded PSD-MALDI-MS spectra from various amounts of MUC1–9AA-GalNAc. Virtually identical spectra without decrease of the signal-to-noise ratio were obtained down to 100 fmol. Within these sensitivity limits, all sites on the target were equally well suited for obtaining good spectra, and less than 5% of the coated target area was used for the analysis (3000 shots per analysis). Coating the target with down to 15 fmol (CCA as matrix; 3000 shots per analysis), it was still possible to obtain the same quality of spectrum without decrease of the signal-to-noise ratio, however, this was somewhat more time-consuming since appropriate places on the target had to be selected.

Localization of glycosylation sites in mono-, di-, and triglycosylated TAP25

TAP25 [(T^{1a}APPAHGVT^{9}S^{10}APDT^{14}RPAPGS^{20})T^{1b}APPA; theoretical $M+H^+$ = 2325.54] is a synthetic peptide that corresponds to one repeat (T1a-S20) and five overlapping

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

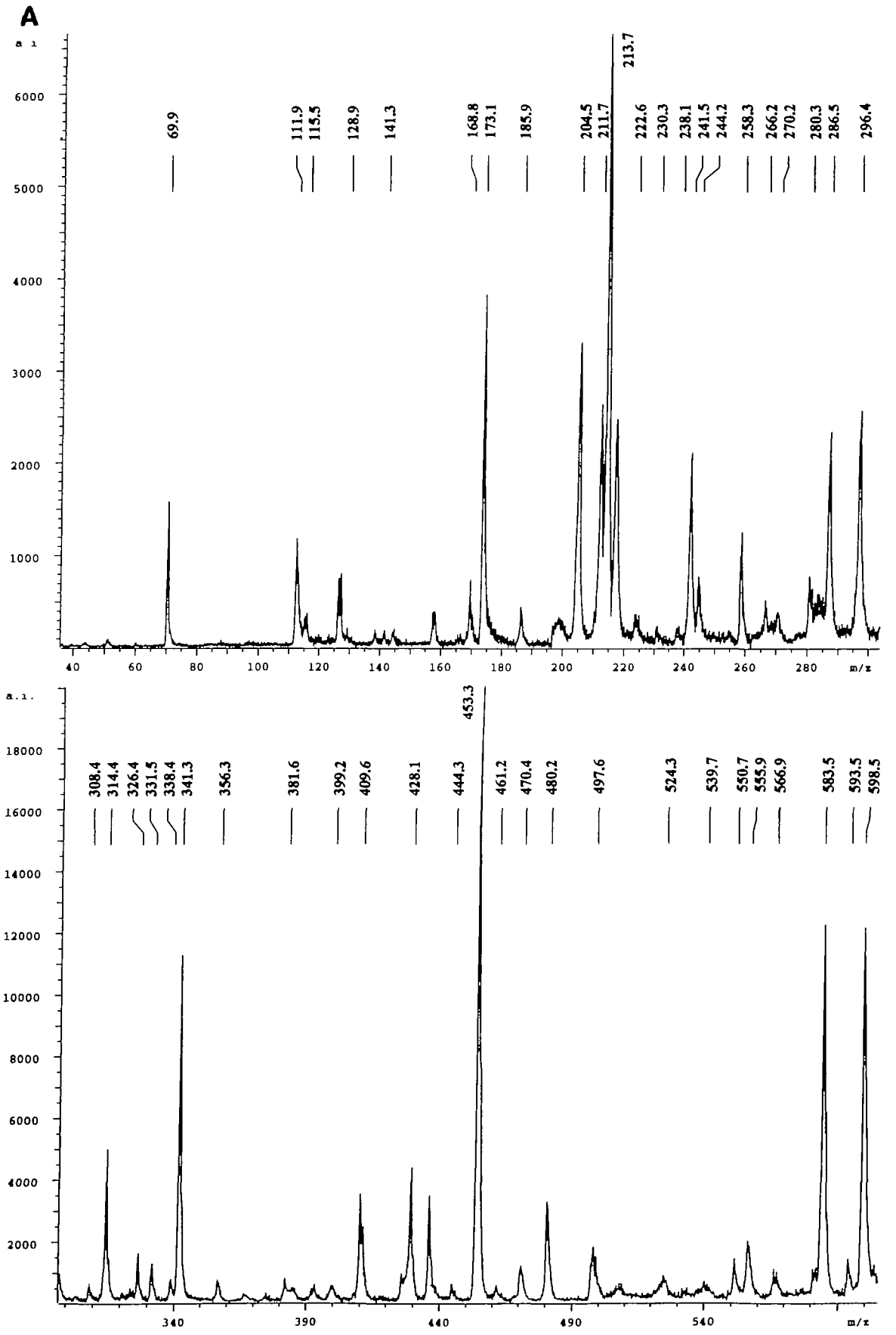


Fig. 1A. Continued on next page. Complete legend is on page 885.

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

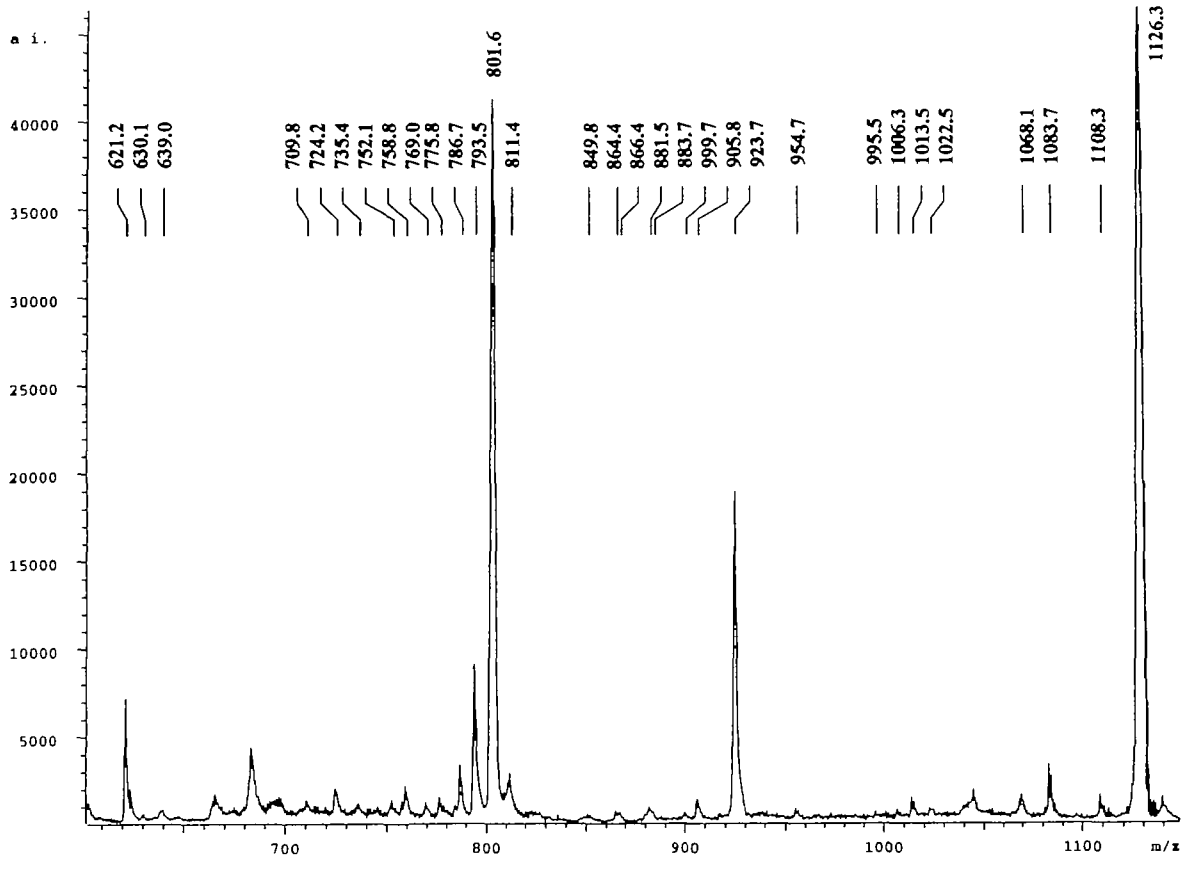


Fig. 1A. Continued. Complete legend is on next page.

amino acids (T1b-A5b) of the MUC1 core protein. It was glycosylated by detergent-solubilized UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases from human premature skim milk according to Stadie *et al.* (1995). The resulting glycopeptides were isolated and purified by ultrafiltration through a 10 kDa cut-off membrane and RP-HPLC. Four different molecules were identified by MALDI-MS: non-, mono-, di-, and triglycosylated TAP25 at $M+H^+$ = 2325.6, 2528.8, 2731.8, and 2935.3, respectively. PSD-MALDI-MS analysis of the nonglycosylated TAP25 (Figure 2A; table with fragment ions not shown) gave a complete series (except y_{22}) of mainly strong C-terminal y_n- and (y_n-18)-fragment ions with a main fragment y_{12}, resulting from cleavage at the C-terminal side of aspartic acid, which corresponds to the same cleavage site as for the main fragment y_6 in the PSD-MALDI-MS spectrum of MUC1–9AA-GalNAc. The main fragment of TAP25 gave, in relation to the other fragment ions, a much stronger signal than the main fragment from MUC1–9AA-GalNAc (Figures 1A, 2A). A_m, b_m, as well as a_m-17 and b_m-17 (m = 25-n) fragment ions were also rather strong. Most, but not a complete series, of the N-terminal fragments were displayed (except fragment ions according to the size of m = 12, 18, 23, and 24 amino acids). Additionally, there were a number of mainly minor a_m, b_m, a_m-17 and b_m-17 internal fragment ions detectable which originated from the unidirectional N-terminal subfragmentation of the various prolines and the main fragments mentioned above. These subfragments were mainly low mass fragment ions. These were partially superimposed and therefore complicated the spectra in the lower mass region. Nevertheless, nearly all peaks, especially the fragment ions of the C- and N-terminal series, but also minor peaks, could be clearly assigned due to the low signal-to-noise ratio which is typical for PSD-MALDI-MS and the mass accuracy which was in most cases better than 0.6 u. The accuracy was always better within an ion series (e.g., C-terminal nonglycosylated fragment ions) and among peaks in a given mass range showing approximately the same deviation. Knowing these tendencies of mass deviation in relation to the peak shape and the occurrence of the companion ions greatly facilitates the reliable assignment of the peaks.

Figure 3A shows the schematical drawing and Table II the calculated masses, observed masses, and the assignments of the fragment ions from the PSD-MALDI-MS spectrum of the diglycosylated TAP25 (TAP25–2). The most prominent fragment ion y_{12} (T14-A5b), the main fragment, was monoglycosylated (Figure 2C, 3A, Table II). It is obvious that the glycosylation was very stable under the used conditions in the PSD analysis. This was confirmed by the detection of all theoretically expected glycosylated y-ions. The few detectable deglycosylated fragment ions of the C-terminal ion series were mainly on a much lower level than the respective glycosylated form (e.g., y_{25} and y_{12}; about 3% of the glycosylated form). The fragment ion spectrum showed a complete series of C-terminal y_n- and some of (y_n-18)-cleavages. Y_2-y_4 yielded nonglycosylated, y_5-y_{16} monoglycosylated, and y_{17}-y_{25} diglycosylated fragment ions, thereby locating Thr9 and Thr1b as glycosylation sites (Figure 3A). The N-terminal b, b-17, a, and a-17 series was not complete, but confirmed the results of the y-series. Some -18 companions also occurred. B_2-b_8 are nonglycosylated; b_9-b_{20} correspond to monoglycosylated, and b_{21}-b_{25} to diglycosylated fragment species. Most of the N-terminal

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

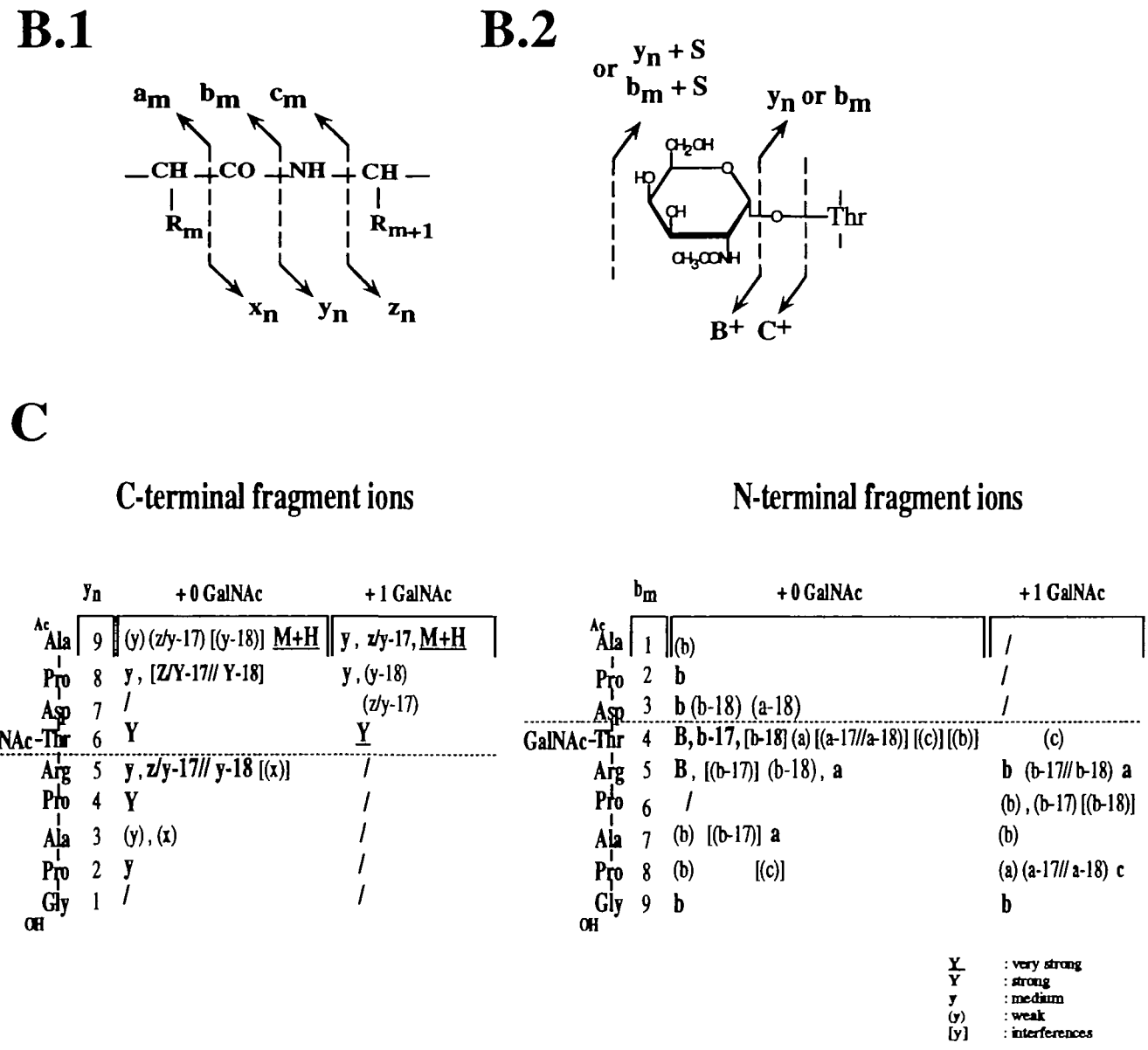


Fig. 1. PSD-MALDI spectrum of MUC1–9AA-GalNAc [theoretical $M+H^+$ = 1126.95]. (**A**) shows the raw spectrum and the observed masses (*m/z*). The calculated masses, observed masses, and mass assignments are listed in Table I. (**B.1**) illustrates the nomenclature of peptide fragmentation according to the Biemann modification of the Roepstorff & Fohlmann nomenclature, and (**B.2**) the nomenclature for glycopeptides, a combination of the peptide fragmentation nomenclature and the nomenclature for the fragmentation of carbohydrates (according to Domon and Costello). (C) shows a schematical presentation of the occurring C- and N-terminal fragment ions according to (B): x_n, y_n, z_n fragment ions of the C-terminal series (e.g., y_n = y fragment ion consisting of n amino acids counted from the C-terminus; z_n has the same *m/z* as y_n-17 and the respective peak is indicated as z_n/ y_n-17), and the a_m, b_m, c_m fragment ions of the N-terminal series (e.g., b_m = b fragment ion consisting of m amino acids counted from the N-terminus). Fragment ions which have lost NH_3 or water are indicated with –17 and –18, respectively. The fragment ions are separated into two columns of nonglycosylated (+0 GalNAc) and glycosylated fragment ion species (+1 GalNAc, which equals +S in Table I with an additional mass of 203 u), respectively. Peak intensities are indicated (Y, very strong, Y, strong; y, medium, (y), weak). The y_n and b_m column indicate, in parallel alignment with the glycopeptide sequence, the fragment number n or m of the horizontally following fragment ions, which equals the number of the included amino acids of the sequence, counted from the C-terminus and N-terminus, respectively (m = peptide length in amino acids – n). The horizontal dotted line divides fragment ions which occur as glycosylated species from those which do not. Internal subfragment ions of proline fragments and the main fragment are not shown. C- and N-terminal fragment ions indicated in brackets (e.g., [y_n]) are superimposed by other C- or N-terminal fragment ions or internal fragment ions. Besides the Biemann nomenclature for the peptide fragmentation, the carbohydrate fragment ions, which occur by fragmentation from the nonreducing end of GalNAc (**B.2**), are labeled with B^+ and C^+. The absolute mass accuracy is in most cases better than 0.4 u of the calculated values. Detailed fragment ion assignments with the respected calculated and observed masses are listed in Table I.

fragment ions were less prominent than the C-terminal ones, and a couple of minor deglycosylated species did also exist. The high number of prolines (7) within the TAP25 peptide sequence resulted in a series of internal proline fragments which fragmented unidirectionally from the N-terminal direction (Table II). They complicated the spectrum, as mentioned above, and no glycosylated species were identified within this series as had already been observed for the synthetic MUC1–9AA-GalNAc. Again, we did not find fragment ions stemming from glycosylated single amino acids or glycosylated dipeptides. The B^+-ion, originating from the cleavage of GalNAc from its nonreducing end, was only vaguely detectable as a

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

Table I. PSD-MALDI-MS fragment ions of MUC1-9AA-GalNAc

Calculated mass (*m/z*) (average)	Observed mass (*m/z*)	Δm(u)	Assignment
1126.9	1126.3	–0.6	M+H
1108.9	1108.3	–0.6	b_9+S
1085.0	1083.7	–1.3	y_9+S
1067.9	1068.1	+0.2	z_9/y_9-17+S
1068.9	1069.0	+0.1	c_8+S
1023.9	1022.5~	–1.4	a_8+S
1013.9	1013.5	–0.4	y_8+S
1006.9, 1005.9	1006.3	–0.6, +0.4	a_8-17+S, a_8-18+S
995.9	995.5	–0.4	y_8-18+S
954.8	954.7	–0.1	b_7+S
923.9	923.7	–0.2	M+H-S
905.9	905.8	–0.1	b_9
899.7	899.7		z_7/y_7-17+S
883.7	883.7~		b_6+S
882.0	881.5	–0.5	y_9
866.7	866.4	–0.3	b_6-17+S
865.9, 865.7	865.5	–0.4, –0.2	c_8, b_6-18+S
864.9, 863.9	864.4	–0.5, +0.5	z_9/y_9-17, y_9-18
848.9	849.8	+0.9	b_8
810.9	811.4	+0.5	y_8
801.7	801.6	–0.1	y_6+S
793.9, 792.9	793.5	–0.4, +0.6	z_8/y_8-17, y_8-18, PDTRPAPG
786.6	786.7	+0.1	b_5+S
775.8	775.8		PDTRPAPG-17
769.5, 768.6	769.0	–0.5, +0.4	b_5-17+S, b_5-18+S
758.6	758.8	+0.2	a_5+S
751.8	752.1	+0.2	b_7
735.8, 734.7	735.4~	–0.4, +0.7	PDTRPAP, b7-17
723.8	724.2	+0.4	a_7
709.9	709.8	–0.1	TRPAP-17+S
647.4	647.5~	+0.1	c_4+S
638.7	639.0	+0.3	PDTRPA
630.4, 629.8	630.1	–0.3, +0.3	b_4+S, TRPA+S
621.7, 620.7	621.2	–0.5, +0.5	PDTRPA-17, PDTRPA-18
598.7	598.5	–0.2	y_6
593.7	593.5	–0.2	PDTRPA-28-17
583.6	583.5	–0.1	b_5
567.6, 566.5	566.9	–0.7, +0.4	PDTRP, b_5-17
565.6	565.6		b_5-18
555.6	555.9	+0.3	a_5
550.6	550.7	+0.1	PDTRP-17
539.6	539.7	+0.1	PDTRP-28
523.6	524.3~	+0.7	TRPAP, x5
497.6	497.6		y_5
480.5, 479.6	480.2	–0.3, +0.7	z_5/y_5-17, y_5-18
470.5	470.4	–0.1	PDTR
461.6	461.2~	–0.4	TA+S
453.5	453.3	–0.2	PDTR-17
444.6, 444.4	444.3	–0.3, –0.1	TA-17+S, c_4
427.4	428.1	+0.7	b_4
425.5	425.4	–0.1	PDTR-28-17
410.4	410.6	+0.2	b_4-17
409.5, 409.6	409.6	+0.1, +0.2	TRPA-17, b_4-18
399.4	399.2	–0.2	a_4
382.4, 381.4	381.6	–0.8, +0.2	a_4-17, a_4-18, TRPA-28-17
355.4	356.3	+0.9	TRP
341.4	341.3	–0.1	y_4
338.4, 337.4	338.4	+/–0, +1.0	TRP-17, TRP-18
331.3	331.5	+0.2	PDT+17
326.3	326.4	+0.1	b_3
314.3	314.4	+0.1	PDT
308.3	308.4	+0.1	b_3-18
296.3	296.4	+0.1	PDT-18
286.3	286.5	+0.2	PDT-28
280.3	280.3~		a_3-18
270.6	270.2	–0.4	x_3
266.3	266.2	–0.1	PAP
258.3	258.3		TA
244.3	244.2	–0.1	y_3
241.3	241.5	+0.2	TA-17
238.3	238.1	–0.2	PAP-28
230.3,230.2	230.3	+/–0, +0.1	TR-28
222.2	222.6	+0.4	C^+
213.2	213.7	+0.5	PD
211.2	211.7	+0.4	b2
204.2	204.5	+0.3	B^+
185.2	185.9	+0.7	PD-28
173.2	173.1	–0.1	y_2
169.2,168.2	168.8	–0.4, +0.6	PA, PD-28-17
166.2	166.1	–0.1	a_2-17
141.2	141.3	+0.1	PA-28
129	128.9	–0.1	R (I)
115.2	115.5	+0.3	P+17
114.1	114.5	+0.4	b1
112	111.9	–0.1	R-17 (I)
70	69.9	–0.1	P-28, P(I)

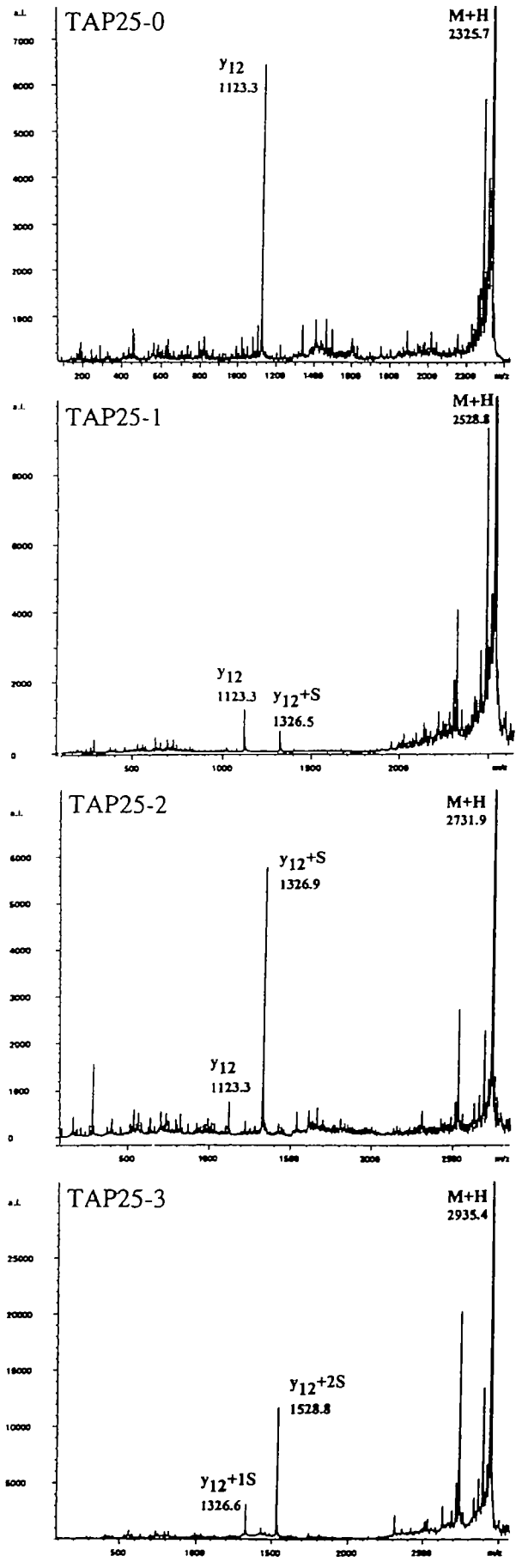
Fragment ions +S equal the indicated fragment ion + 203 u for the sugar GalNAc; (I) represents the immonium ion of the indicated amino acid; nomenclature according to Biemann *et al.* (1988) and Domon and Costello (1988).

very small peak, and no C^+-ion was detected. This was, in comparison to the B^+- and C^+-ions of the MUC1–9AA-GalNAc fragmentation, a very low signal, which correlates with the apparently lower amounts of deglycosylated fragments in the PSD-analysis of TAP25–2. A reason for this might be the poor transmission of low-energy ions. B^+-ions originating from TAP25–2 would be expected to retain a smaller percentage of the initial energy than B^+-ions from MUC1–9AA-GalNAc.

Figure 3B shows the results obtained with the triglycosylated TAP25 (TAP25–3) in the schematic presentation proposed by us. Again, the y-series, including the y-18 companions, was the most informative one and sufficient for localizing the glycosylated amino acids to Thr9, Ser20, and Thr1b. Y_2-y_4 corresponded to nonglycosylated, y_5 to monoglycosylated, y_6-y_{16} to diglycosylated, and y_{17}-y_{25} to triglycosylated fragment ions. B-ions, and a-ions, together with their -17 and -18 companions, confirmed these results but did not show a complete N-terminal series.

The localization of the GalNAc on the monoglycosylated peptide, TAP25–1, was more complex. The first hint for the possibility that it might represent a mixture of monoglycosylated TAP25 species was obtained from the PSD-MALDI-MS spectra. In TAP25–2 and TAP25–3 the glycosylated main fragment was by far more abundant than the fragment ion which had lost one sugar, showing, together with the other fragment ions, that the glycosylation is rather stable during the PSD-MALDI-MS analysis under the used conditions (Figure 2). The monoglycosylated main fragment of TAP25–1, which located the glycosylation site to the region T14-A5b, was less abundant than its nonglycosylated counterpart. Furthermore, the PSD-MALDI-MS spectrum could not clearly localize a single glycosylation site on T14-A5b, but did also show glycosylated fragment ions in T1-D13. Additionally, Edman degradation of TAP25–1 revealed two clear signals of similar intensity for Pth-Thr and the Pth-derivative of Thr-GalNAc at position Thr9, which is very close to the Pth-Ser peak (Stadie *et al.*, 1995), indicating a mixture of monoglycosylated species, and only nonglycosylated signals for Ser10 and Thr14. The

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022



glycosylation status of Ser20 and Thr1b could not be determined because the signals were too low in these late cycles to distinguish between carry-over products from the previous cycle and a potential mixture of glycosylated and nonglycosylated derivatives of the present cycle. Enzymatic digestions with Asp-N and trypsin (Figure 4A,B) were performed to prove that TAP25–1 is a mixture of monoglycosylated peptides. The Asp-N digest showed a mixture of T1a-P12 and D13-A5b, both glycosylated and nonglycosylated. T1a-R15, in the glycosylated and nonglycosylated form, was the only fragment which was detected by MALDI-MS from the tryptic digest. The other expected fragment P16-A5b was neither observed in the tryptic digest of TAP25–1 nor in a parallel digest of TAP25–0. The various glycosylated fragments were analyzed by PSD-MALDI-MS using the digested unpurified samples by applying precursor selection. The monoglycosylated AspN fragment D13-A5b +S showed a clear localization of the N-acetylhexosamine to Thr1b and the monoglycosylated trypsin fragment T1a-R15 +S to Thr9 of TAP25 (not shown). The monoglycosylated AspN fragment T1a-P12 was not used because of its low signal-to-noise ratio in the linear MALDI-MS spectrum.

Determination of glycosylation sites and saccharide structures on TF-MUC1 glycopeptides

PSD-MALDI-MS analysis of the synthetic glycopeptide MUC1–6AA-αTF ($_{Ac}$S(Galβ1–3GalNAcα1-)TAPPA; theoretical $M+H^+$ = 950.95), which is glycosylated with the Thomsen-Friedenreich disaccharide Galβ1–3GalNAcα1-O- at serine, revealed the localization of the glycan attachment site to Ser1 as well as a characterization of the disaccharide (Figure 5, Table III). The C-terminal fragment ion series was sufficient to localize and characterize the disaccharide. Only y_6 fragment ions displayed glycosylated species. The glycosylated fragment species of y_6 were those which were glycosylated with the complete disaccharide (Figure 5: +Gal-GalNAc; additional 365 u) and those which were partially deglycosylated with only the core GalNAc remaining (+GalNAc; additional 203 u), indicating an inner N-acetylhexosamine and a terminal hexose. The N-terminal fragment ion series confirmed the disaccharide structure (peptide-HexNAc-Hex) with the existence of GalNAc-Gal glycosylated fragment species as well as GalNAc glycosylated ones. The N-terminal fragment ion series localized the glycosylation site to S1-T2 since the glycosylated fragment species occurred for b_2-b_6. One peak at *m/z* 500.3 could be assigned to a_1-17-H+Na+GalNAc+Gal or to b_5-18-H+Na (this overlapping is indicated in Figure 5 with brackets). It is more likely that this peak represents the latter fragment ion because of the existing b_5-H+Na ion (*m/z* 518.3), and since the observed fragmentation patterns of MUC1–9AA-GalNAc, TAP25–1, TAP25–2, and TAP25–3 render it unlikely to find a fragment ion of the terminal glycosylated amino acid. However, it can not be excluded that fragment ions of one glycosylated amino acid exist and are part of the observed peak at *m/z* 500.3. Therefore, without the y-series, Ser1 and Thr2 were not distinguishable in this case.

Additional fragment ions were detected which originated

Fig. 2. Raw spectra of the PSD-MALDI-MS analyses of TAP25–0, TAP25–1, TAP25–2, and TAP25–3. The main fragment ion (y_{12}) and the parent ion (M+H) are indicated with their respective masses. Fragment ions which include one or two GalNAc are labeled + 1 S and + 2 S, respectively.

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

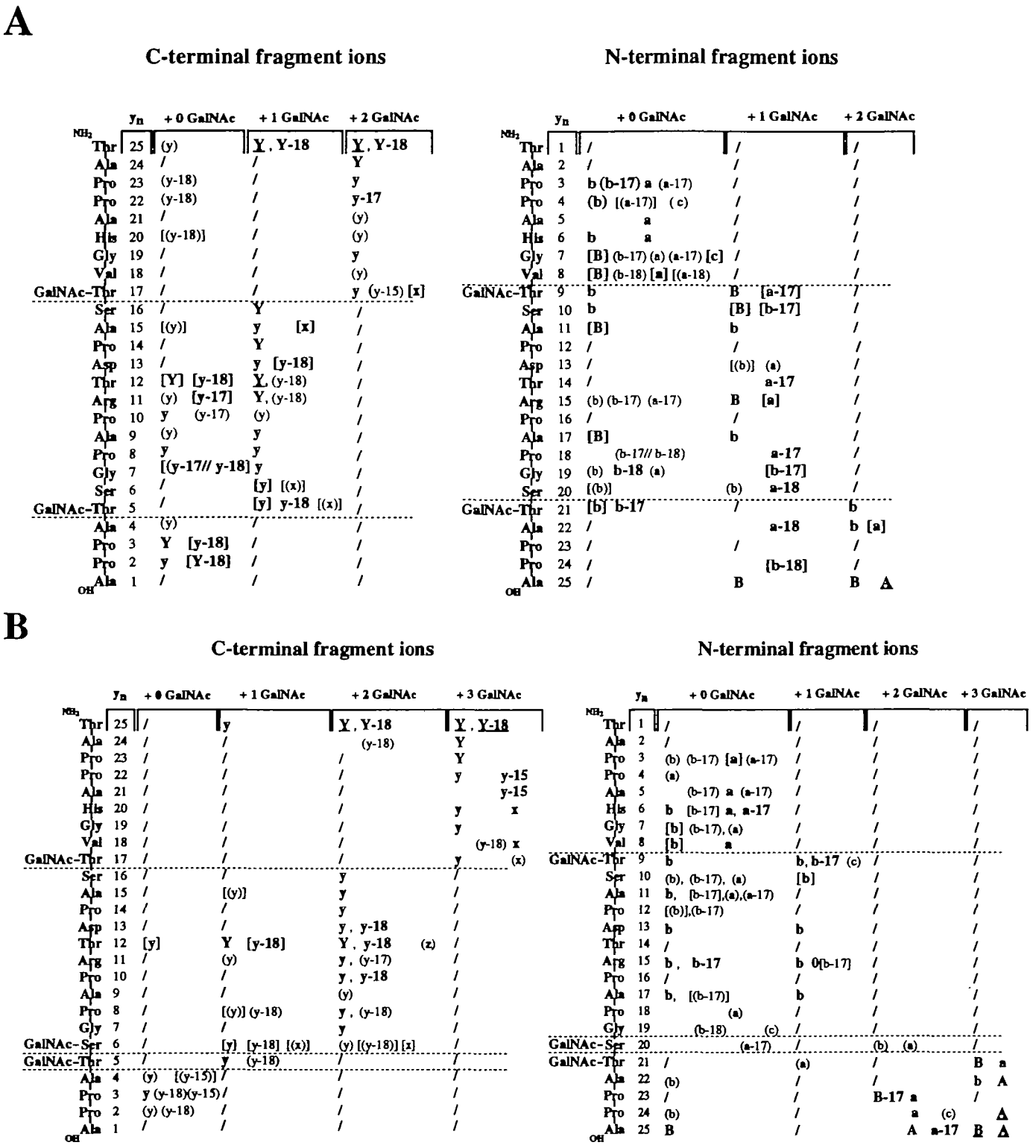


Fig. 3. Schematical presentation of the PSD-MALDI-MS fragmentation pattern of diglycosylated TAP25–2 (A; [theoretical M+H$^+$ = 2528.7]; Table II) and triglycosylated TAP25–3 (B; [theoretical M+H$^+$ = 2731.9]. TAP25 [(T^{1a}APPAHGVT^{9}S^{10}APDT^{14}RPAPGS^{20})T^{1b}APPA; theoretical M+H$^+$ = 2325.5] was glycosylated *in vitro* with detergent-solubilized UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases from human premature skim milk. Schematical presentation as in Figure 1C. Fragment ions which include no GalNAc, one GalNAc, two GalNAc, and three GalNAc are indicated as +0 GalNAc, + 1 GalNAc, + 2 GalNAc, and + 3 GalNAc, respectively.

from the cleavage of the disaccharide from the nonreducing end (Figure 5, top). The B_1^+ and C_1^+ ions (nomenclature according to Domon and Costello, 1988) resulted from the cleavage between Gal and GalNAc, and B_2^+ and C_2^+ ions from the cleavage between the GalNAc and the serine. The fragment ion $^{3,5}A_2$ originated from an internal sugar ring cleavage as indicated in Figure 5. All peaks were prominent except the B_1^+ ion which displayed a weak signal.

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

Table II. PSD-MALDI-MS fragment ions of TAP25-2

Calculated mass (*m/z*) (average)	Observed mass (*m/z*)	Δm(u)	Assignment
2731.9	2731.8	−0.1	y_{25}+2S
2713.9	2714.0	+0.1	y_{25}-18+2S, b_{25}+2S
2685.9	2686.7	+0.8	a_{25}+2S
2630.8	2630.3	−0.5	y_{24}+2S
2559.8	2559.4	−0.4	y_{23}+2S
2528.7	2528.5	−0.2	y_{25}+1S, (b_{23}-17+2S)
2510.7	2510.7		y_{25}-18+1S, b_{25}+1S
2448.6	2448.5	−0.1	b_{22}+2S
2445.6	2446.3~	+0.7	z_{22}/y_{22}-17+2S
2421.6, 2420.6	2421 2	−0.4, +0.6	b_{24}-18+1S, a_{22}+2S
2377.5	2377.1	−0.4	b_{21}+2S
2365.5	2366.3	+0.8	y_{21}+2S
2325.5	2326.2	+0.7	y_{25}, (b_{23}-17+1S)
2296.5	2296.7	+0.2	a_{22}-18+1S
2294.4	2294.7~	+0.3	y_{20}+2S
2135.3	2135.5~	+0.2	y_{23}-18
2157.3	2156.9	−0.4	y_{19}+2S, (b_{21}-17+1S)
2100.3	2100 6	+0.3	y_{18}+2S
2073.2	2072.7	+0.5	y_{20}+1S
2038.2	2038.6~	+0.4	y_{22}-18
2027.2, 2027.1	2026.1	−1.1, −1.0	a_{20}-18+1S, x_{17}+2S
2001.1	2001.6~	+0.5	y_{17}+2S
1986.1	1986.6	+0.5	y_{17}-15+2S
1971.1, 1969.1	1970.3	−0.8, −0.4	b_{21}, b_{19}-17+1S
1954.1	1954.4	+0.3	b_{21}-17
1870.0	1870.3~	+0.3	y_{20}-18, b_{20}
1884.1	1884.5	+0.4	a_{18}-17+1S
1832.0	1832.0		b_{17}+1S
1783.0	1783.5	+0.5	b_{19}
1764.9	1764 5	−0.4	b_{19}-18
1754.9	1755.4	+0.5	a_{19}
1708.9, 1707.9	1708.3	−0.6, +0.4	b_{18}-17, b_{18}-18
1696.8	1697.2	+0.4	y_{16}+1S
1663.8	1663.9	+0.1	b_{15}+1S
1635.8, 1635.7	1635.5	−0.3, −0.2	a_{15}+1S, x_{15}+1S
1611.8	1611.6	−0.2	b_{17}-17
1609.7	1609.8	+0.1	y_{15}+1S
1538.7	1537.9	−0.8	y_{14}+1S
1462.6	1462.4	−0.2	a_{14}-17+1S
1460.6	1461.1	+0.5	b_{15}
1443.6	1443.7	+0.1	b_{15}-17
1441.5	1442.0~	+0.5	y_{13}+1S
1423.5	1423.5		y_{13}-18+1S, (c_{13}+1S)
1415.6	1415.5	−0.1	a_{15}-17
1406.5	1406.3	−0.2	y_{15}, b_{13}+1S
1378.5	1378.1	−0.4	a_{13}+1S
1356.7	1356.4	−0.3	PPAHGVTSAPDTRP-28
1326.5	1326.9	+0.4	y_{12}+1S
1308.4	1308.4		y_{12}-18+1S, (c_{12}+1S)
1225.3	1225.1	−0.2	y_{11}+1S
1207.3	1207.6	+0.3	y_{11}-18+1S
1194.3	1193.7	−0.6	b_{11}+1S
1131.5	1131.6	+0.1	PPAHGVTSAPDTRT
1123.3	1123.3		y_{12}, (b_{10}+1S)
1105.2	1105.5	+0.3	y_{12}-18, (b_{10}-18+1S)
1069.2	1069.8	+0.6	y_{10}+1S
1036.1	1036.8	+0.7	b_9+1S
1034.5	1034.4	−0.1	PAHGVTSAPDT
1022.1	1022.4	+0.3	y_{11}
1005.2, 1005.1	1005.3	+0.1, +0.2	(c_{12}), z_{11}/y_{11}-17
991.1	991.5	+0.4	b_{11}, a_9-17+1S
972.0	972.3~	+0.3	y_9+1S
933.4	933.3	−0.1	PAHGVTSAPD
920.0	919.8	−0.2	b_{10}
901.0	900.7~	−0.3	y_8+1S
866.0	865.6	−0.4	y_{10}
848.9	849.2	+0.3	z_{10}/y_{10}-17
832.9	833.5	+0.6	b_9
818.4	819.2	+0.8	PPAHGVTSA, PAHGVTSAP

Table II. *Continued*

Calculated mass (*m/z*) (average)	Observed mass (*m/z*)	Δm(u)	Assignment
803.8	804.4	+0.6	Y_7+1S PPAHGVTSA-17, PAHGVTSAP-17, PPAHGVTSA-18,
801.4, 800.4	801.4	+/−0, +1.0	PAHGVTSAP-18 PPAHGVTSA-28,
790.4	791.2	+0.8	PAHGVTSAP-28 PPAHGVTSA-28-17,
773 4	773 2	−0.2	PAHGVTSAP-28-17,
772.8	773.2	+0.4	x6+1S, PPAHGVTSA-28-18,
772.4	773.2	+0.8	PAHGVTSAP-28-18
768.8	769.1	+0.3	y_9
764.4	765.4	+1.0	PPAHGVTS+17
747.4, 746.8	748.0	+0.6, +1.2	PPAHGVTS, y6+1S
731.8, 730.4	731.3	−0.5, +0.9	b_8, PPAHGVTS-17
721.4	721.3	−0.1	PAHGVTSA
713.8	713.2	−0.6	b_8-18
703.8, 703.4	703.2	−0.6, −0.2	a_8, PAHGVTSA-18,
702.4	703.2	+0.8	PPAHGVTS-28-17
697.8	697.3	−0.5	y_8
693.4	693.9	+0.5	PAHGVTSA-28
685.8, 685.7	685.5~	−0.3, −0.2	a_8-18, x_5+1S
660.3, 659.7	660.2~	−0.1, +0.5	PPAHGVT, y_5+1S
650.3, 649.7	650.5	−0.3, +0.3	PAHGVTS, c_7
641.7	641.9	+0.2	y_5-18+1S
633.3, 632.7	633.0	−0.3, +0.3	PAHGVTS-17, b_7,
632.4	633.0	+0.6	PPAHGVT-28
622.3	622.3		PAHGVTS-28
615.7	615.7		b_7-17
604.7	604.9	+0.2	a_7
587.7	587.8	+0.1	a_7-17
583.8, 583.6	583.4	−0.4, −0.2	(PAPGSTA), z_7/y_7-17,
582.6	583.4	+0.8	y_7-18
575.6	576.3	+0.7	b_6
565.3	565.2	−0.1	PAPGSTA-17
559.3	560.2	+0.9	PPAHGV
554.3	554.8	+0.5	PAPGSTA-28
547.6	548.3	+0.7	a_6
535.3	536.2	+0.9	PAHGVT-28
531.3, 530.6	532.0	+0.7, +1.4	PPAHGV-28, (a_6-17)
493.2	493.6	+0.4	PAPGST-18
462.2, 460.2	461.3	−0.9, +1.1	PAHGV, PPAHG
414.2	414.7	+0.5	PGSTA, (PPAHG-28-18)
410.5, 410.2	410.7	+0.2, +0.5	a_5, (PAPGS)
403.2	404.1	+0.9	PPAH
386.2	387.0	+0.8	PPAH-17
384.5	384.1	−0.4	c_4
375.2	375.7	+0.5	PPAH-28
367.4	367.5	+0.1	b_4
363.2	364.0	+0.8	PAHG
355.4	356.1	+0.7	y_4
343.2	343.7	+0.5	PGST
331.2	331.8	+0.6	PDT+17
323.7, 323.2	323.3	−0.4, +0.1	PAH+17, PAPG,
322.4	323.3	+1.1	a_4-17
314.1	314.3	+0.2	PDT
306.2, 306.1	306.5	+0.3, +0.4	PAH, PAPG-17
284.3	285.0	+0.7	y_3
270.3	269.7	−0.6	b_3
266.3, 266.2	266.7	+0.4, +0.5	y_3-18, PPA, PAP
259.1	259.1		PGS+17
253.3, 252.3	253.0	−0.3, +0.7	b_3-17, (b_3-18)
242.3, 242.1	242.0	−0.3, −0.1	a_3, PGS
225.3	225.0	−0.3	a_3-17
214.1, 213.2	213.9	−0.2, +0.7	PGS-28, x_2,
214.1	213.9	+0.8	PD
195.1	195.3	+0.2	PA, PD-18
187.2	187.3	+0.1	y_2

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

Table II. *Continued*

Calculated mass (*m/z*) (average)	Observed mass (*m/z*)	Δm(u)	Assignment
172.2	172.1~	–0.1	y_2-15
169.2, 169.1	169.1	–0.1, +/–0	y_2-18, PA
155.2, 155.1	155.5~	+0.3, +0.4	(b_2-18), PG
141.1	141.2	+0.1	PA-28
138.1	138.0	–0.1	PG-17
115.1	114.8	–0.3	P+17
70.1	70.1		P(I), P-28

Fragment ions +1S and +2S equal the indicated fragment ions with 1 GalNAc and 2 GalNAc sugar residues (203.2 u and +406.4 u, respectively; (I) represents the immonium ion of the indicated amino acid; nomenclature according to Biemann et al. and Domon and Costello; ( ), assumed minor contribution to peak intensity; ~, peak not well defined.

All fragment ions shown in the schematical presentation of Figure 5 were sodium adducts (additional 22 u). Sodium stayed with the glycosylated and nonglycosylated peptide fragment species as well as the with carbohydrate fragment ions from the nonreducing end. The parent ion occurred as $M+H^+$, $M+Na^+$, and $M-H^++2Na^+$ peaks, where $M+Na^+$ displayed by far the strongest signal. Only few fragment ions did also occur without the additional sodium and none were detected with two additional Na ions. Therefore, sodium ions did not hinder the glycopeptide analysis, which could have been the case if all fragments would have occurred with and without sodium ions thereby complicating the spectra.

This shows that PSD-MALDI-MS glycopeptide analysis can also localize and sequence a disaccharide, determine the glycosylation site even at the end of the peptide, and tolerate sodium contaminations.

Discussion

In this report we show that matrix-assisted laser desorption ionization time-of-flight mass spectrometry in combination with the reflectron mode analysis of post-source decay fragments (PSD-MALDI-MS) is a fast, highly sensitive and reproducible technique for the localization and sequencing of mono- and disaccharides on glycopeptides corresponding to the MUC1 core peptide of the tandem repeats. The method is applicable to glycopeptide mixtures and tolerates salts to a certain extent.

Most of the rules of occurring fragment ions which were established with the simple glycopeptide, MUC1–9AA-GalNAc, held also true for the TAP25 glycopeptides. The only difference seemed to be the observation that the glycosyl linkages of GalNAc on the *in vitro* glycosylated TAP25 were more stable in the PSD-MALDI-MS analyses compared to the chemically synthesized MUC1–9AA-GalNAc. The latter is glycosylated at a position which can not be glycosylated *in vitro* using enzyme preparations from milk. In general, the stability of the sugars in PSD-MALDI-MS analysis in all tested substances was sufficient to allow their unequivocal localization. However, glycosylated fragments occurred not only as glycosylated fragment ions but did, in addition, also occur occasionally as deglycosylated fragment ions. The intensities of the deglycosylated ions were usually considerably lower than their glycosylated counterparts. In case of the disaccharide side chains, this partial cleavage of sugars off the peptide backbone during the metastable decay of the parent ions in the first field free path was even of advantage, since potentially glycosylated fragments diversified into fully glycosylated, partially deglycosylated and completely deglycosylated ions which allowed a simultaneous sequencing of the small glycan. Additionally, fragment ions from the nonreducing end of the disaccharide showed characteristic daughter ions which aided the characterization of the sugar moiety. The stability of the sugars seemed to be higher on y-fragment ions from the C-terminal series in comparison to b- and a-ions from the N-terminal series. Therefore, and because of the more complete ion-series and their in general higher signal intensities, the y-fragment ion series was in all observed cases sufficient to localize and characterize the sugar part. The b- and a-ion series were also rather complete and backed the y-ion series. Y-Fragments resulting from the cleavage of the peptide backbone at the C-terminal side of aspartic acid, which is known to be a major site for backbone cleavage, represented the main fragments. This fragmentation is especially valuable when Asp divides potential glycosylation sites as it is the case for the MUC1 tandem repeat peptide TAP25. This allowed a quick localization of sugars to one or the other side of Asp on the glycopeptide. Internal subfragmentation of the main fragments and the proline fragments which yield new b- and a-ions unidirectionally, did complicate the spectra and showed few or no glycosylated species, respectively. In order to minimize the formation of these subfragments in future investigations, appropriate enzymatic digestions of the MUC1 tandem repeats should be chosen which will result in shorter proline fragments and less internal fragment ions (especially fragments which contain the preferred proline fragmentation PPA). This can be achieved, for example, by means of the Gly-C-specific protease IV from *Papaya,* which splits TAP25 into two peptides, V8-G19 and S20-G7 of TAP25. Similarly, the size of the main fragment can be controlled by appropriate enzymatic digestions in order to provide a quick localization of glycosylation sites to one or the other side of Asp.

The schematical presentation of the occurring C- and N-terminal fragment ions introduced by us (see Figures 1C, 3A,B, 5) summarizes the important data of the tables, and simplifies their interpretation and the localization of the glycosylation sites for the reader.

Our results show that Thr9, Thr1b, and Ser20 are glycosylation sites for the detergent-solubilized UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases from human premature skim milk on the overlapping tandem repeat TAP25 derived from MUC1. The monoglycosylated TAP25–1 was found to be a mixture of two monoglycosylated species with the same molecular mass but with distinct glycosylation sites at Thr9 or at Thr1b. Since MALDI-MS of TAP25–1 and its digestions does not give quantitative results, it was not possible to determine exactly the proportions of the two monoglycosylated species. However, by comparing the Edman sequencing data, which showed approximately equal amounts of glycosylated Thr9 and nonglycosylated Thr9 in the fraction of monoglycosylated species, with the MALDI-MS results from the enzymatic digests and PSD-MALDI-MS of TAP25–1, we have no indication for a preferred glycosylation site. This implies a statistical usage of either Thr9 or Thr1b of MUC1 as the first glycosylation sites for the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases of human premature skim milk. The homogeneous fraction of TAP25–2, with only Thr9 and Thr1b as glycosylation sites used by the detergent-

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

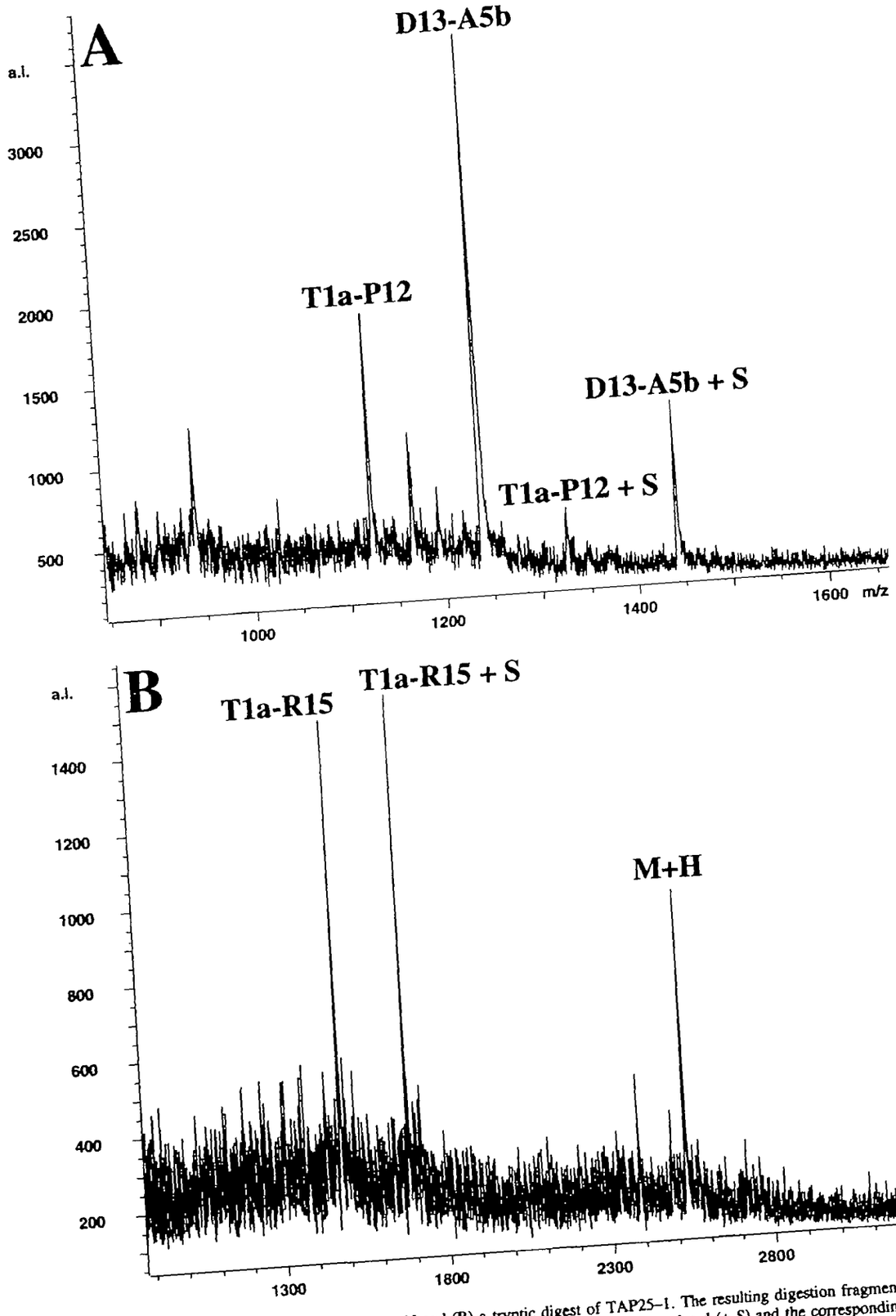


Fig. 4. MALDI-MS analysis in the linear mode of the (A) Asp-N and (B) a tryptic digest of TAP25–1. The resulting digestion fragments are indicated with their amino acid sequence in relation to the TAP25 sequence. The mass difference between the glycosylated (+ S) and the corresponding nonglycosylated fragment is 203 u.

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

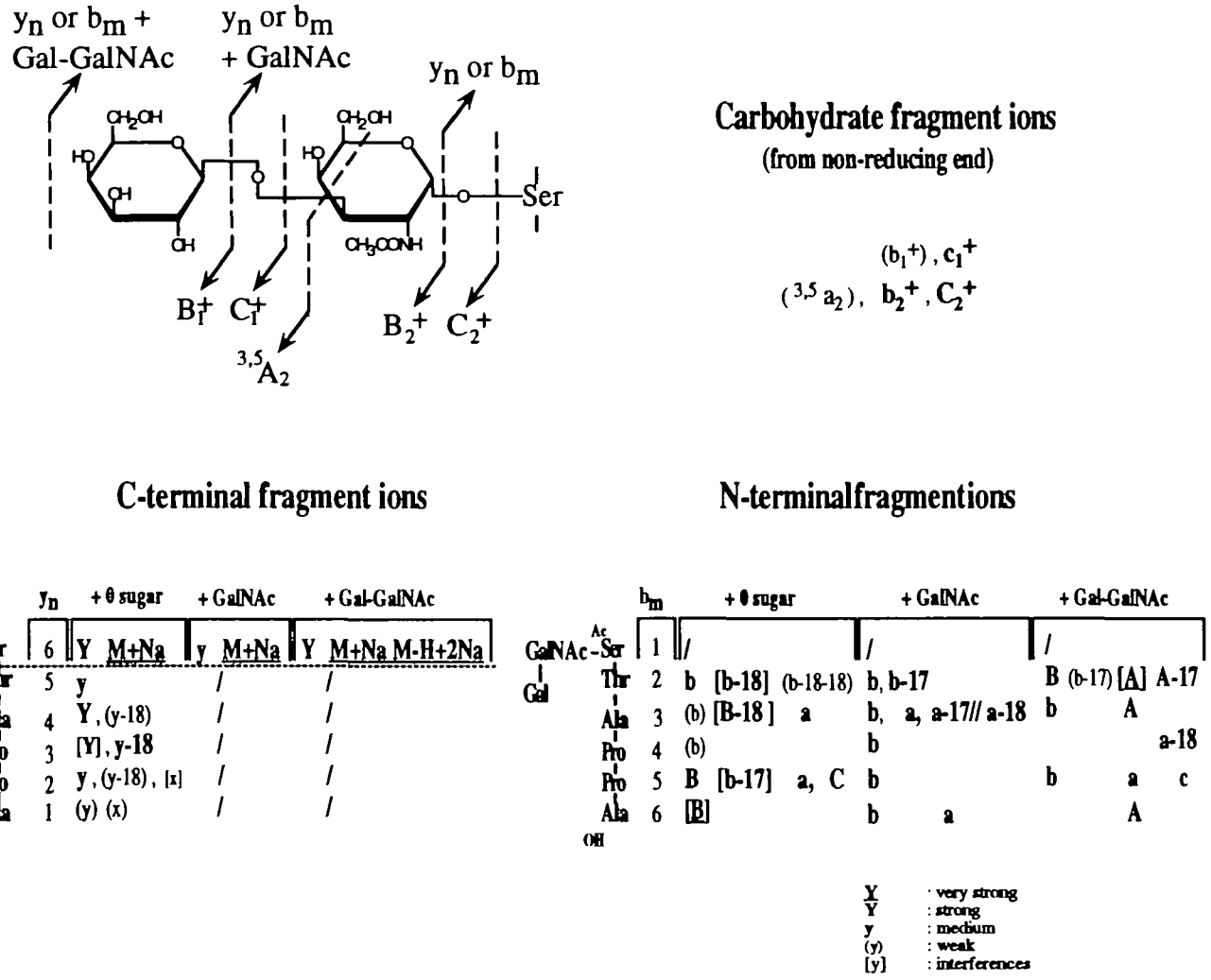


Fig. 5. Schematical presentation (as in Figure 1C) of the PSD-MALDI-MS fragmentation pattern of MUC1–6AA-αTF [$_{Ac}$S(αTF)TAPPA; theoretical M+H = *m/z* 950.95], which is glycosylated with the Thomsen-Friedenreich disaccharide Galβ1–3GalNAcα1-O at the serine. Fragment ions including the complete disaccharide are indicated with + Gal-GalNAc, an additional 365 mass units, and those which include only the core GalNAc with + GalNAc, additional 203 u. Fragment ions originating from cleavage from the non-reducing end of the disaccharide are listed separately using the nomenclature of Domon and Costello as indicated at the top of Figure 5. All indicated fragment ions are cationized with sodium. Detailed fragment ion assignments are listed together with their calculated and observed masses in Table III.

solubilized UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases, suggests that glycosylations at Thr9 and Thr1b are a prerequisite for glycosylation at Ser20. This is in accordance with the results obtained for TAP25–3 which was shown to be homogeneously glycosylated at Thr9, Thr1b, and Ser20. Thr14 and Ser10 were not found to be glycosylated in any of the glycosylated TAP25 peptides, which is in accordance with reports using other enzyme sources (Nishimori *et al.*, 1994a,b; Stadie *et al.*, 1995). Therefore, we propose that *in vitro* glycosylation with GalNAc, using the detergent-solubilized UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases from human skim milk, is a step-wise process. Thr9 or Thr1b are glycosylated independently at random. However, glycosylation of both Thr9 and Thr1b is apparently necessary for a subsequent glycosylation of Ser20. The glycosylation of TAP25 using the milk enzyme preparation could, for example, be achieved by the parallel action of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase 1 and/or 3, which prefer VT^{9}SA, together with GalNAc-transferase 2, which prefers ST^{1b}AP, leading to the diglycosylated TAP25–2 (H. Clausen *et al.*, 4th International Workshop on Carcinoma-Associated Mucins; Cambridge 1996). It appears that the enzyme(s) which glycosylates Thr1b is not dependent on glycosylations at Thr9 or other sites, but require some peptide stretch around this site, at least towards the N-terminus, since Thr1 was not glycosylated. This is in agreement with the finding that GalNAc-glycosylation on Thr1b is favored by Pro at positions -3, +2, and +3, and at -6 and +3 for Thr9 (Eckhardt *et al.*, 1987; Wilson *et al.*, 1991; O'Connell *et al.*, 1992). The glycosylation of Ser20 could then be achieved by one of these or another enzymes, which might be serine specific.

Comparing these results with TAP25 which had been analogously glycosylated *in vitro* with a microsomal UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase preparation from the human breast carcinoma cell line T47D (Stadie *et al.*, 1995), it becomes apparent that there is a difference between the two transferase systems (Figure 6). In both cases three different glycoforms were detected, the mono-, di-, and triglycosylated TAP25–1, TAP25–2, and TAP25–3, respectively. TAP25–2 and TAP25–3 were apparently identical when

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

Table III. PSD-MALDI-MS fragment ions of MUC1-6AA-αTF

Calculated mass (*m/z*) (average)	Observed mass (*m/z*)	Δm(u)	Assignment
994.6	993.2	−1.4	M-H+2Na+GalNAc+Gal
988.6	987.0, 988 0	−1.6, −0.6	M+K+GalNAc+Gal
972.6	971.9	−0.7	M+Na+GalNAc+Gal
950.6	950.1	−0.5	M+H+GalNAc+Gal
930.6	930.1	−0.5	y_6-H+Na+GalNAc+Gal
926.6	927.1	+0.5	a_6-H+Na+GalNAc+Gal
900.5	901.3	+0.8	c_5-H+Na+GalNAc+Gal
883.5	884.0	+0.5	b_5-H+Na+GalNAc+Gal
855.5	855.9	+0.4	a_5-H+Na+GalNAc+Gal
810.6	810.7	+0.1	M+Na+GalNAc
792.6	792.5	−0.1	b_6-H+Na+GalNAc
768.6	768.2	−0.4	y_6-H+Na+GalNAc
764.6	765.0~	+0.4	a_6-H+Na+GalNAc
740.4	739.9	−0.5	a_4-18-H+Na+GalNAc+Gal
721.5	721.4	−0.1	b_5-H+Na+GalNAc
689.3	689.6	+0.3	b_3-H+Na+GalNAc+Gal
661.3	661.8	+0.5	a_3-H+Na+GalNAc+Gal
624.4	624.5	+0.1	b_4-H+Na+GalNAc
618 2	618.4	+0.2	b_2-H+Na+GalNAc+Gal
607.6, 607 4	607.8	+0.2, +0.4	M+Na, (b_4-17-H+Na+ GalNAc)
601.2	601.5~	+0.3	b_2-17-H+Na+GalNAc+Gal
590.2, 589.6	589.9	−0.3, +0.3	a_2-H+Na+GalNAc+Gal, b_6-H+Na
585.4	585.5~	+0.1	b_4-17+GalNAc
573.2	573.1	−0.1	a_2-17-H+Na+GalNAc+Gal
568.2	568.6	+0 4	a_2+GalNAc+Gal
565.6	566.0	+0.4	y_6-H+Na
535.5	536.2	+0.7	c_5-H+Na
527.3	527.4	+0.1	b_3-H+Na+GalNAc
518.5	518.3	−0.2	b_5-H+Na
500.5, 500.0	500.3	−0.2, +0.3	b_5-18-H+Na, a_1-17-H+Na+ GalNAc+Gal
499.3	499.4~	+0.1	a_3-H+Na+GalNAc
490.5	490.5		a_5-H+Na
482.2, 481.2	481.6~	−0.6, +0.4	a_3-17-H+Na+GalNAc, a_3-18-H+Na+GalNAc
478.5	478.5		y_5-H+Na
456.5, 456.2	456.5~	+/−0, +0.3	(y_5), b_2-H+Na+GalNAc
439.5, 439.2	439.3	−0.2, +0 1	(z_5/y_5-17), b_2-17-H+Na+GalNAc
434.2	434.2		b_2+GalNAc
421.4	421.6~	+0.2	b_4-H+Na
406.3, 406.2	406.4	+0.1, +0.2	C_2-H+Na, (a_2+GalNAc)
399.4	399.1	−0.3	b_4
388.3, 388.2	388.2	−0.1, +/−0	B_2-H+Na, (a_2-18+GalNAc)
377.4	377.4		y_4-H+Na
324.3	324.0	−0.3	b_3-H+Na
316.0	315.9	−0.1	b_1-17+GalNAc
306.3, 306.2	306.2	−0.1, +/−0	y_3-H+Na, b_3-18-H+Na
296.3	296.3		a_3-H+Na
288.3	288.2	−0.1	y_3-18-H+NA
278.2	278.4	+0.2	a_3-18-H+Na
259.1	259.4	+0.3	$^{3,5}A_2$-H+Na
253.2	252.8	−0.4	b_2-H+Na
235.2	235.0	−0.2	x_2-H+Na, b_2-18-H+Na
226.2	226.1	−0.1	B_1/B_2-H+Na, C_1^+/C_2^+-H+Na
217.2	217.2		b_2-18-18-H+Na
213.2	213.1	−0.1	b_2-18, (x_2)
209.2	209.3	+0.1	y_2-H+Na
203.2	203.3	+0.1	C_1-H+Na
191.2	191.0	−0.2	y_2-18-H+Na
185.2	185.4~	+0.2	B_1-H+Na
138.1	138.3	+0.2	x_1-H+Na
112.1	111.8	−0.3	y_1-H+Na, (b_1-18)

Fragment ions +GalNAc and +Gal equal the indicated fragment ion + 203 u and + 162 u, respectively; nomenclature according to Biemann *et al.* (1988) and Domon and Costello (1988); fragment ions in brackets represent assignments whose contributions for the observed peaks are assumed to be minor or can be neglected compared to the nonframed fragment ion assignments.

using both transferase systems. However, TAP25–1 was exclusively glycosylated at Thr9 using the T47D transferase preparation, but at Thr9 or Thr1b when the transferase system from human premature skim milk was applied. This difference could be explained by the lack of an active peptide GalNAc-transferase 2 in the T47D preparation, which preferentially glycosylates ST^{1b}AP. If only core GalNAc-transferases 1 and/or 3 (which prefer VT^{9}SA over ST^{1b}AP) are active or present in T47D, Thr9 would be expected to become glycosylated first. A minor activity towards Thr1b could then allow a slower glycosylation of Thr1b, yielding the diglycosylated species which is necessary for the glycosylation of Ser20.

Our results show that PSD-MALDI-MS is superior to gas-phase Edman degradation analysis in the identification of glycosylation sites since the latter can not detect the glycosylated Thr1b in TAP25–1. Additionally, PSD-MALDI-MS is more sensitive (femtomolar range compared to low picomolar range for conventional Edman degradation). The gas-phase Edman degradation analysis has the further disadvantage that the Pth-derivative of Thr-GalNAc is nearly isographic with Pth-Ser, and the Pth-derivative of Ser-GalNAc is close to Pth-Asp in Edman degradation analysis. Therefore, the discrimination between glycosylated and nonglycosylated neighboring sites can be impaired by carry-over products during successive cycles. Solid-phase sequencing (Pisano *et al.*, 1993) overcomes this problem and is able to identify the sugar and its position on the glycopeptide. The advantage of this technique is its ability to quantify the glycosylation of a given site which might be important for the analysis of a peptide glycosylated at alternative positions in order to determine their relative amounts, as it would be the case with TAP25–1. In contrast, PSD-MALDI-MS might need an additional enzymatic cleavage in case of structural isomers of a glycopeptide which have the same mass but are glycosylated at alternative positions in order to separate the alternative glycosylation sites. In that case, PSD-MALDI-MS will be able to determine the glycosylation sites right out of the digestion mixture without further purification, making additional RP-HPLC unnecessary. In other words, the great advantage of PSD-MALDI-MS is that it can analyze mixtures of glycopeptides which can be separated by a pulsed field (precursor selection) and analyzed separately. This holds true not only for glycopeptide mixtures with various peptide sequences, but also for glycopeptide mixtures with the same peptide backbone. A mixture of TAP25–0/1/2/3, for example, which results from the above described *in vitro* glycosylation, can be resolved by PSD-MALDI-MS as shown in this manuscript. Solid-phase sequencing could identify the three glycosylation sites and their relative usage in the mixture, but would be unable to show the existence of the various TAP25-glycoforms (0/1/2/3). Solid-phase sequencing could be successful after separation of the various glycoforms. This can be difficult or impossible in many cases because of the similar physico-chemical properties of the glycoforms but has been partly achieved in the case of TAP25 glycoforms (Stadie *et al.*, 1995).

Apart from the mentioned advantage of PSD-MALDI-MS in case of mixtures, it is superior to other mass spectrometrical techniques due to its high sensitivity, ease of use, high signal-to-noise ratio, high reproducibility, and its relative insensitivity to solvent impurities. The applicability of PSD-MALDI-MS in glycopeptide analysis should not be restricted to *in vitro* glycosylated or chemically synthesized mucin peptides, but should also hold true for glycopeptides purified from cells or

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

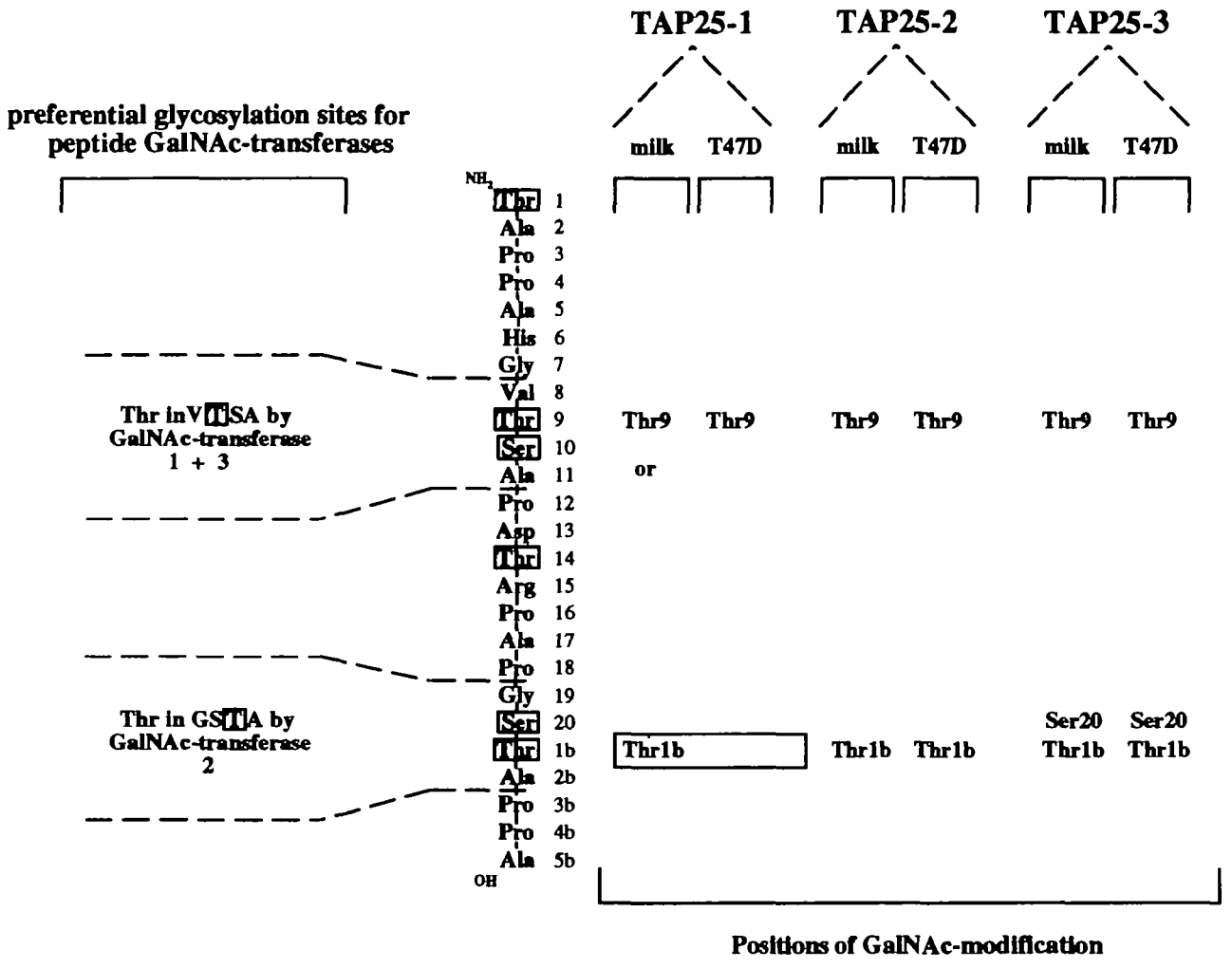


Fig. 6. Summary of the glycosylation site analysis of the overlapping MUC1 repetitive sequence TAP25 which was glycosylated *in vitro* with detergent-solubilized UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases (peptide GalNAc-transferases) from human premature skim milk or by a microsomal preparation from the human breast carcinoma cell line T47D.

body fluids. We are currently examining the use of PSD-MALDI-MS for the analysis of *ex vivo* prepared MUC1 tandem repeat units. Furthermore, there seems to be no principal restriction for the analysis of longer glycan chains linked to peptides. However, long glycans on peptides could complicate the spectra. We are currently investigating how glycan complexity influences the resolution of current PSD-MALDI mass spectrometers. We believe that the method is extremely useful and reliable for the localization and characterization of the sugar part as long as the peptide sequence is known, because the percentage of peaks which can be reliably assigned to certain fragment ions is very high (better than 90%). The determination of glycosylation sites of glycopeptides with unknown peptide sequence is in principle possible and will be feasible for shorter and less complex glycopeptides (e.g., MUC1–9AA-GalNAc). It might, however, be difficult or even impossible with longer and proline-rich sequences (e.g., TAP25–2 or –3). The critical limiting parameter to deal with long glycans or with longer glycopeptides of unknown peptide sequences is the mass accuracy and hence the resolution of neighboring fragment ion peaks. The mass accuracy of PSD-MALDI-MS is known to be lower than that of ESI-MS/MS instruments. In our experiments the mass accuracy of the vast majority of peaks was better than 0.4 u for lower and 0.6 u for higher mass ranges. It was always better within an ion series and among peaks in a certain mass range as opposed to the absolute overall accuracy. In the present study, the accuracy was sufficient to reliably identify most of the peaks even in the case of the glycosylated TAP25 peptide with its high amount of internal subfragmentation due to the number of proline residues. One can expect that this young technique will probably be improved considerably in its mass accuracy in the near future. One improvement can already be achieved by delayed extraction (Brown and Lennon, 1995), which is described to increase the mass accuracy to 10–100 ppm deviation of the theoretical mass by MALDI-MS experiments.

Materials and methods

PSD-MALDI-MS and MALDI-MS

Linear, reflectron, and PSD mass spectra analyses were performed with a MALDI-TOF instrument (Bruker-Reflex, Bruker-Franzen Analytic, Bremen, Germany). Desorption and ionization were performed by a pulsed UV laser beam (nitrogen laser, λ = 337 nm). Irradiance was used at a level slightly above the threshold of ion detection. Residual gas pressure was at $1–1.5 \times 10^{-7}$ mbar. Ion spectra were obtained in the positive ion mode. Acceleration and reflector voltages were set to 28.5 kV and 30 kV, respectively. The molecular

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

parent ion was, if necessary, isolated by deflection of all other ions with a pulsed field (precursor selection). Complete PSD-fragment ion spectra were obtained by stepwise reducing the reflector voltage to produce overlapping mass ranges, combination of the part-spectra and calibration according to the recommendations of the manufacturer. No smoothing or filtering was applied to the shown spectra. Mass accuracy was in most cases better than 0.4 u.

Sample preparation

Matrix substances CCA (α-cyano-4-hydroxycinnamic acid; saturated solution in TA (2/3 aqueous 0.1% trifluoroaceticacid, 1/3 acetonitrile)) or DHB (2,5-dihydroxybenzoic acid; 15 mg/ ml in TA) were mixed 1:1 (v/v) with the probe. The mixtures (0.7–0.8 μl) were placed onto a polished aluminum target and air-dried.

Glycopeptides

TAP25 [(T^{1a}APPAHGVT^{9}S^{10}APDT^{14}RPAPGS^{20})T^{1b}APPA; theoretical M+H$^+$ = 2325.54] was prepared and glycosylated *in vitro* as described in Stadie *et al.* (1995). TAP25 corresponds to one repeat (T1a-S20) and five overlapping amino acids (T1b-A5b) of the MUC1 core protein which served as an acceptor substrate for *in vitro* glycosylation with detergent-solubilized UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase prepared from human premature skim milk (taken from five individuals during the first week after parturition). This was chosen as a non-tumorous source for the transferase system in order to compare the results with those of a transferase system from the human breast cancer cell line T47D (Stadie *et al.*, 1995). The glycosylated peptides were isolated by ultrafiltration through a 10 kDa cut-off membrane (Millipore), purified by reverse-phase HPLC (PLRP-S, 4.6 × 250 mm; Polymer laboratories), freeze-dried, and dissolved in methanol/ water (1:1). Three different glycosylation species were separated and identified by MALDI-MS: mono-, di-, and triglycosylated peptides, TAP25-GalNAc$_1$ (TAP25–1), TAP25-GalNAc$_2$ (TAP25–2), and TAP25-GalNAc$_3$ (TAP25–3), respectively.

MUC1–9AA-GalNAc is a chemically synthesized and N-terminally acetylated part of the MUC1 tandem repeat peptide [$_{Ac}$-APDT(αGalNAc)RPAPG-OH; theoretical M+H$^+$ = 1126.95; Seitz and Kunz, 1995]. It is O-glycosylated with αGalNAc at the threonine which corresponds to Thr14 in TAP25.

MUC1–6AA-αTF corresponds to S20-A5b of TAP25 [$_{Ac}$S(αTF)TAPPA; theoretical M+H$^+$ = 950.95] which is acetylated and glycosylated with the Thomsen-Friedenreich disaccharide Galβ1–3GalNAcα1-O- at the serine, which equals the Ser20 of TAP25.

Proteolytic digestion of TAP25-GalNAc$_j$

Cleavages with Asp-N (sequencing grade, Boehringer Mannheim, Mannheim, Germany) and modified trypsin (sequencing grade, Promega, Madison, WI) were performed with an enzyme/substrate ratio of 1:50 in 10 mM Tris–HCl, pH 7.8, and in 10 mM Tris–HCl pH 8.0 with 1 mM $CaCl_2$, respectively, for 16 h at 37°C.

Automated Edman sequencing of glycopeptides

The peptide was loaded onto a biobrene-coated glass fiber filter and analyzed with a Procise protein sequencer (Applied Biosystems, Weiterstadt, Germany) using standard cycles of Edman degradation. The phenylthiohydantoin derivatives of the amino acids (Pth-Xaa) were identified by HPLC whereby the GalNAc-glycosylated Ser and Thr lack the corresponding peaks at the positions of unmodified Ser and Th. In the case of GalNAc-Thr the peak of its derivative is expected close to the position of Pth-Ser and for GalNAc-Ser just behind the Pth-Asp peak (Stadie *et al.*, 1995).

Acknowledgments

This study was partially supported by the Deutsche Forschungsgemeinschaft (grants Ka 921/1–1, Hu 146/19–1, Ha 2092/6–1, and Ha 2092/4–1), and an MDC fellowship to S.G.

Abbreviations

Ac-, acetyl-; αGalNAc, α-linked N-acetylgalactosamine; CCA, α-cyano-4-hydroxycinnamic acid; DHB, 2,5-dihydroxybenzoic acid; GalNAc, N-acetylgalactosamine; GalNAc-transferase, UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase; Hex, hexose; HexNAc, N-acetylhexosamine; MALDI, matrix-assisted laser desorption ionization; MS, mass spectrometry; *m/z*, mass per charge; PSD, post-source decay; Pth-Xaa, phenylthiohydantoin derivative of amino acids; RP-HPLC, reverse-phase high performance liquid chromatography; TFα, Thomsen-Friedenreich disaccharide Galβ1–3GalNAcα1-O-; TOF, time-of-flight.

References

Biemann,K. (1988) Contributions of mass spectrometry to peptide and protein structure. *Biomed. Environ. Mass Spectrom.*, **16**, 99–111.

Brockhausen,I., Yang,J.-M., Burchell,J., Whitehouse,C. and Taylor-Papadimitriou,J. (1995) Mechanisms underlying aberrant glycosylation of MUC1 mucin in breast cancer cells. *Eur. J. Biochem.*, **233**, 607–617.

Brown,R.S. and Lennon,J.J. (1995) Mass resolution improvement by incorporation of pulsed ion extraction in a matrix-assisted laser desorption/ ionization time-of-flight mass spectrometer. *Anal. Chem.*, **67**, 1998–2003.

Burchell,J., Gendler,S.J., Taylor-Papadimitriou,J., Girling,A., Lewis,A., Millis,R. and Lamport.D. (1987) Development and characterization of breast cancer reactive monoclonal antibodies directed to the core protein of the human milk mucin. *Cancer Res.*, **47**, 5476–5482.

Domon,B. and Costello,C. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J.*, **5**, 397–409.

Eckhardt,A.E., Timpte,C.S., Abernethy,J.L., Zhao,Y. and Hill,R.L. (1991) Porcine submaxillary mucin contains a cystine-rich, carboxyl-terminal domain in addition to a highly repetitive, glycosylated domain. *J. Biol. Chem.*, **266**, 9678–9686.

Gendler,S.J., Lancaster,C.A., Taylor-Papadimitriou,J., Duhig,T., Peat,N., Burchell,J., Pemberton,L., Lalani,E.-E. and Wilson,D. (1990) Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. *J. Biol. Chem.*, **265**, 15286–15293.

Hanisch,F.-G., Uhlenbruck,G., Peter-Katalinic,J., Egge,H., Dabrowski,J. and Dabrowski,U. (1989) Structures of neutral O-linked polylactosaminoglycans on human skim milk mucins. A novel type of linearly extended poly-N-acetyllactosamine backbones with Gal β(1–4)GlcNAc β(1–6) repeating units. *J. Biol. Chem.*, **264**, 872–883.

Hanisch,F.-G., Stadie,T.R., Deutzmann,F. and Peter-Katalinic,J. (1996) MUC1 glycoforms in breast cancer cell line T47D as a model for carcinoma-associated alterations of O-glycosylation. *Eur. J. Biochem.*, **236**, 318–327.

Huberty,M.C., Vath,J.E., Yu,W. and Martin,S.A. (1993) Site-specific carbohydrate identification in recombinant proteins using MALDI-TOF MS. *Anal. Chem.*, **65**, 2791–2800.

Hull,S.R., Bright,A., Carraway,K.L., Abe,M., Hayes,D.F. and Kufe,D.W. (1989) Oligosaccharide differences in the DF3 sialomucin antigen from normal human milk and the BT-20 human breast carcinoma cell line. *Cancer Commun.*, **1**, 261–267.

Kaufmann,R. (1995) Matrix-assisted laser desorption ionization (MALDI) mass spectrometry: a novel analytical tool in molecular biology and biotechnology. *J. Biotechnol.*, **41**, 155–175.

Kaufmann,R., Kirsch,D., and Spengler,B. (1994) Sequencing of peptides in a time-of-flight mass spectrometer: evaluation of postsource decay following matrix-assisted laser desorption ionisation (MALDI). *Int. J. Mass Spectrom. Ion Processes*, **131**, 355–385.

Kaufmann,R., Spengler,B. and Lützenkirchen,F. (1993) Mass spectrometric sequencing of linear peptides by product-ion analysis in a reflectron time-of-flight mass spectrometer using matrix-assisted laser desorption ionization. *Rapid Commun. Mass Spectrom.*, **7**, 902–910.

Machold,J., Utkin,Y., Kirsch,D., Kaufmann,R., Tsetlin,V. and Hucho,F. (1995) Photolabeling reveals the proximity of the α-neorotoxin binding site to the M2 helix of the ion channel in the nicotinic acetylcholine receptor. *Proc. Natl. Acad. Sci. USA*, **92**, 7282–7286.

Medzihradszky,K.F., Gillece-Castro,B.L., Settineri,C.A., Townsend,R.R., Masiarz,F.R. and Burlingame,A.-L. (1990) Structure determination of O-linked glycopeptides by tandem mass spectrometry. *Biomed. Environ. Mass Spectrom.*, **19**, 777–781.

Nishimori,I., Johnson,N.R., Sanderson,S.D., Perini,F., Mountjoy,K., Cerny,R.L., Gross,M.L. and Hollingsworth,M.A. (1994a) Influence of acceptor substrate primary amino acid sequence on the activity of human UDP-N-acetylgalactosaminyltransferase. Studies with the MUC1 tandem repeat. *J. Biol. Chem.*, **269**, 16123–16130.

Nishimori,I., Perini,F., Mountjoy,K.P., Sanderson,S.D., Johnson,N., Cernay,R.L., Gross,M.L., Fontenot,J.D. and Hollingworth,M.A. (1994b) N-Acetylgalactosamine glycosylation of MUC1 tandem repeat peptides by pancreatic tumor cell extracts. *Cancer Res.*, **54**, 3738–3744.

O'Connell,B.C., Tabak,L.A. and Ramasubbu,N. (1991) The influence of flanking sequences on O-glycosylation. *Biochem. Biophys. Res. Commun.*, **180**, 1024–1030.

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022

O'Connell,B.C., Hagen,F.K. and Tabak,L.A. (1992) The influence of flanking sequences on the O-glycosylation of threonine *in vitro*. *J. Biol. Chem.*, **267**, 25010–25018.

Pisano,A., Redmond,J.W., Williams,K.L., and Gooley,A.A. (1993) Glycosylation sites identified by solid-phase Edman degradation: O-linked glycosylation motifs on human glycophorin A. *Glycobiology*, **3**, 429–435.

Pisano,A., Packer,N.H., Redmond,J.W., Williams,K.L., and Gooley,A.A. (1994) Characterization of O-linked glycosylation motifs in the glycopeptide domain of bovine kappa-casein. *Glycobiology*, **4**, 837–844.

Roepstorff,P. and Fohlmann,J. (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.*, **11**, 601.

Seitz,O. and Kunz,H. (1995) A novel allylic anchor for solid-phase synthesis of protected and unprotected O-glycosylated mucin-type glycopeptides. *Angew. Chem. Int. Ed. Engl.*, **34**, 803–805.

Spengler,B., Lützenkirchen,F. and Kaufmann,R. (1993) On-target deuteration for peptide sequencing by laser mass spectrometry. *Org. Mass Spectrom.*, **28**, 1482–1490.

Spengler,B., Kirsch,D., Kaufmann,R. and Lemoine,J. (1995) Structure analysis of branched oligosaccharides using post-source-decay in matrix-assisted laser desorption ionization mass spectrometry. *J. Mass Spectrom.*, **30**, 782–787.

Spengler,B, Kirsch,D., Kaufmann,R. and Jaeger,E. (1992) Peptide sequencing by matrix-assisted laser-desorption mass spectrometry. *Rapid Commun. Mass Spectrom.*, **6**, 105–108.

Stadie,T.R.E., Chai,W., Lawson,A.M , Byfield,P.G.H. and Hanisch,F.-G. (1995) Studies on the order and site specificity of GalNAc transfer to MUC1 tandem repeats by UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase from milk or mammary carcinoma cells. *Eur. J. Biochem.*, **229**, 140–147.

Talbo,G. and Mann,M. (1996) Aspects of the sequencing of carbohydrates and oligonucleotides by matrix-assisted laser desorption/ ionization post-source decay. *Rapid Commun. Mass Spectrom.*, **10**, 100–103.

Wilson,I.B.H., Gavel,Y. and von Heijne,G. (1991) Amino acid distributions around O-linked glycosylation sites. *J. Biol. Chem.*, **275**, 529–534.

Yu,W., Vath,J.E., Huberty,M.C., Martin,S.A. and Scoble,H.A. (1993) Proceedings of the 41st Annual Conference on Mass Spectrometry and Allied Topics, San Francisco, CA. ASMS, Santa Fe.

Received on September 30, 1996, revised on February 7, 1997; accepted on February 27, 1997

Downloaded from https://academic.oup.com/glycob/article/7/7/881/725497 by guest on 20 August 2022