JOURNAL OF MASS SPECTROMETRY
J. Mass Spectrom. **33**, 1124–1133 (1998)

Nanospray-ESI low-energy CID and MALDI Post-source Decay for determination of O-glycosylation sites in MUC4 peptides

Kim Alving[1], Roman Körner[2], Hans Paulsen[3] and Jasna Peter-Katalinic[1]*
[1] Institut für Medizinische Physik und Biophysik der Westfälischen Wilhelms-Universität Münster, Germany
[2] Department of Molecular Biology, Odense University, Denmark
[3] Organisch-Chemisches Institut der Universität Hamburg, Hamburg, Germany

Mucins are glycoproteins containing from one to several hundreds of carbohydrate chains attached to the peptide by *O*-glycosidic linkages between *N*-acetylgalactosamine (GalNAc) and either serine (Ser) or threonine (Thr). Apomucin proteins are encoded by different mucin genes (MUC1–8) and show in general tandem repeats in which the carbohydrate chains can be highly clustered. The glycosylation density of human mucins seems to be related to some type of cancer. Fast atom-bombardment mass spectrometry (FAB-MS) has previously been used for the determination of *O*-glycosylation sites in combination with β-elimination. In this contribution, a comparative study by electrospray (ESI)-MS/MS and post-source decay (PSD)-MALDI-MS is presented. Nanospray-ESI-MS/MS experiments were performed using two analyzers, a triple quadrupole and a quadrupole ion trap. A series of *O*-GalNAc substituted synthetic MUC4 fragments with blocked N- and C-termini, and their nonglycosylated counterparts were investigated. In the sequence analysis of differently glycosylated and nonglycosylated MUC4 decapeptides by MALDI-PSD-MS, the spectra were dominated by b- and *y*-type ions. Substantial and structure-unrelated loss of the sugar from molecular $[M + H]^+$ and fragment ions was generally observed. The single GalNAc attachment site out of three to five potential substitution Ser/Thr sites was identified by attribution of N- and C-terminal fragment ions still carrying the GalNAc moiety. Similar fragmentation patterns were produced by nanospray ESI–low-energy CID on a Q3 instrument using Ar as a collision gas or on ion trap CID using He as a collision gas. Less abundant sequence ions were obtained from CID experiments on sodiated molecular $[M + Na]^+$ ions, which required much higher energy. © **1998 John Wiley & Sons, Ltd.**

KEYWORDS: mucins; *O*-glycosylation; mass spectrometry; ESI-MS; MALDI-MS

INTRODUCTION

Mucins are *O*-glycoproteins of high molecular weight (100-1000 kDa) and a high degree of glycosylation (50–80%). They play a major role in defense of mucosae, covering and protecting epithelium against various types of aggression. In most cases, glycan chains are located in the densely clustered regions on the apomucin protein backbone, showing structural heterogeneity concerning the core type, the chain length and degree of sialyation. Apomucins are encoded by different mucin genes (MUC 1–8 genes) showing tandem repeats with a high proportion of hydroxylated amino acids. Biological significance and pathophysiology-related structural features of mucins have been the subjects of many investigations. Aberrant glycosylation in cancer results in expression of simpler carbohydrate structures, such as Thomsen–Friedenreich (TF) antigen Galβ1-3GalNAc-Ser(Thr) or its precursor, Tn epitope GalNAc-Ser(Thr) [1]. The reason for the glycosylation site-related microheterogeneity is still largely unknown, owing to the structural problems related to the complexity and the size of the mucin fragments investigated. In model studies, synthetic apomucin MUC1 fragments were glycosylated *in vitro* in order to determine factors involved in the substrate recognition [2].

Mass spectrometry was shown to represent the most direct approach for analysis of mucins. Using fast atom bombardment mass spectrometry (FAB-MS) it was possible to demonstrate molecular diversity of *O*-glycans in mucins, released from their parent apomucin by reductive elimination reaction and permethylation [3]. Their general architecture, like linearity and patterns of branching, and finally the expression of blood-group and/or tumor-associated antigens as determinants on the non-reducing termini were elucidated from fragmentation patterns, obtained from unimolecular decays in the positive ion mode. The structural determination of the core region, showing eight possible structures, became accessible using a combined strategy, including oxidative cleavage of the core GalNAc and derivatization of the cleavage fragments [4]. Their difference in

* *Correspondence to:* J. Peter-Katalinic, Institut für Medizinische Physik, Wilhelms-Universitaẗt, Robert-Koch-Str. 31 D-48149 Munster, Germany.
E-mail: jkp@uni-muenster.de

CCC 1076–5174/98/111124–10 $17.50

Received 5 April 1998
Accepted 30 June 1998

mass was shown to be an excellent parameter for FAB-MS analysis after separation on thin layer chromatography (TLC) and direct desorption from the plate [5]. By positive and negative ion mode ESI-MS the degree of sialyation and the glycosylation domains were demonstrated for clustered, multiple glycosylated short peptides [6]. Skimmer voltage sequencing in ESI-MS was suitable for the sequencing of the sugar portion in glycosylated amino acids and short peptides [7]. For identification of O-glycosylation sites in peptides with up to 10 amino acids a combination of reductive elimination and high energy CID FAB-MS was shown to be a successful approach [8]. High-energy CID was also used in sequencing native and *N*-derivatized fetuin-derived *O*-glycopeptides, obtained by trypsin and pronase proteolysis [9]. Post-source decay (PSD) MALDI-MS was applied for sequencing of MUC1 glycopeptides, rich on proline [10]. However, a clear-cut molecular ion map to document microheterogeneity and to sequence the peptide chain for identification of the glycosylation sites is still a challenging enterprise.

In this contribution, a systematic investigation of three instrumental strategies for sequencing synthetic model MUC4 peptides with blocked N- and C-termini containing only one GalNAc moiety linked to one of several potential *O*-glycosylation sites is presented.

EXPERIMENTAL

Instruments

Mass spectrometry experiments were performed using the mass spectrometric facilities at Department of Chemistry and Department of Molecular Biology, Odense University, except ESI Fourier-transform ion-cyclotron sustain off resonance irradiation CID (ESI-FTICR-SORI-CID) which was performed at Bruker Franzen, Bremen, Germany.

PSD-MALDI was performed on a Voyager Elite instrument (Perseptive Biosystems, Framingham, MA, USA). Target voltage was 20 kV, and for delayed extraction at the first grid 91% of target voltage, and 250 nsec delay was used.

The nanospray-ESI *x*, *y*, *z*-manipulator was home-built. The needle voltage applied was 700 V. For nanospray ESI a prototype Esquire Ion Trap (Bruker Franzen, Bremen, Germany) and the TSQ 700 triple quadrupole analyzer (Finnigan MAT, San Jose, CA, USA) were used. On the Bruker instrument the buffer and collision gas He was held at approximately 1 mtorr. A maximum of 800 mass units broad scans was allowed by the instrument set-up. On the Thermo Bioanalysis instrument Q^2 was used as a collision cell with Ar as a collision gas at the pressure of 2.0 mtorr or slightly higher. A collision energy of -15 to -40 V for protonated and of -65V in case of sodiated molecular ions was chosen. All spectra were run in the positive ion mode.

For the ESI-FTICR-SORI-CID experiment a Bruker bioAPEX II (Bruker Franzen, Bremen, Germany) equipped with a 3.0 Tesla superconducting magnet was used.

Peptides and Glycopeptides

All MUC4 peptides and glycopeptides (for sequences see Table 1) were synthesized at the University of Hamburg and used without further purification. An attempt to desalt the peptides P10, P11, and P12 was carried out on a high-performance liquid chromatography (HPLC) SMART system (Pharmacia Biotech, Freiburg, Germany) fitted with a reversed phase C18 column, run with 0.1% TFA (v/v) in water and eluted with an acetonitrile gradient. Another attempt to desalt was carried out on an ion exchange column (cation exchange polymer: Biorad, 50W-X8 mesh size 100–200 μm) activated with 0.1% aq. HCI. Ion exchange beads were used also to wash the target before adding the matrix and analyte.

For MALDI (DHB), 2,5 dihydroxybenzoic acid was used as a matrix. It was prepared as a 12 g/l in water : acetonitrile 2 : 1 (v/v) 0.1% trifluoroacetic acid (TFA) solution in water.

RESULTS AND DISCUSSION

All MUC4 substrates represent different decapeptide stretches from the MUC4 repeating domain, starting from different positions in the peptide chain and therefore containing different numbers of potential glycosylation sites. Hydroxylated amino acids, serine and threonine, are abundant and represent up to 50% of the total amount of amino acid building blocks. The decapeptide sequences are given in Table 1.

Table 1 MUC4 Sequences

Sample		*Mass* $[M+H]^+$ (*monoisotopic*)	*Mass* $[M+H]^+$ (*average*)
P1	Ac-TSSVSTGHAG–NH_2	944.43	944.94
P2	Ac-TSSVST(GalNAc)GHAG–NH_2	1147.43	1147.94
P8	Ac-GHATSLPVTG–NH_2	980.51	981.06
P9	Ac-GHAT(GalNAc)SLPVTG–NH_2	1183.51	1184.06
P10	Ac-ATSLPVDTG–NH_2	1002.5	1003.06
P11	Ac-AT(GalNAc)SLPVTDTG–NH_2	1205.5	1206.06
P12	Ac-ATSLPVT(GalNAc)DTG–NH_2	1205.5	1206.06

Ac-TSSVSTGHAG-NH_2 (P1)

Spectra are shown in Fig. 1, interpretations in Table 2. P1 peptide contains five potential glycosylation sites close to or at the N-terminus. Its MALDI-PSD spectrum ($[M + H]^+ = 944$) shows two uninterrupted series of b and y fragment ions, b_{2-9} and y_{3-8}, although of significantly differing abundance (Fig. 1a). The most abundant ions are y_6, at the serine cleavage site, and b_8 and b_9, at the histidine and alanine cleavage site, respectively. Symptomatic for the high T/S content at the N-terminus is the abundant loss of H_2O from b ions. A number of internal cleavage fragment ions of lower abundance was also present. A similar fragmentation tendency was found also in the ESI-ion trap spectrum of Pl, containing more abundant fragment ions around the C-terminal histidine, b_8 and b_9 (Fig. 1b). The relative abundance distribution of fragment ions appears in the ESI/Q3 spectrum in an opposite way for both series, b_{2-9} and y_{3-8}, y_3 at the histidine cleavage site showing the highest intensity (Fig. 1c). All three spectra give sufficient and relevant information for sequencing.

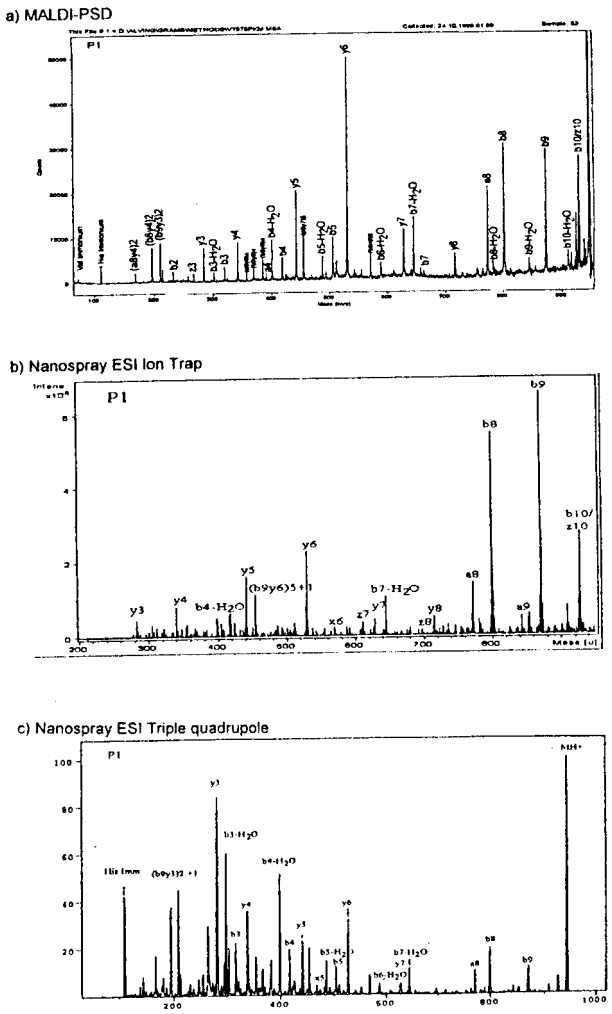

Figure 1. Product ion spectra of Ac-TSSVSTGHAG–NH_2 (P1). (a) PSD-MALDI, precursor ion selected: $[M + H]^+$. (b) Nano-ESI/ion trap, precursor ion selected: $[M + H]^+$. (c) Nano-ESI/Q3, precursor ion selected: $[M + H]^+$.

Table 2.

Calc. mass average (m/z)	*Obs. mass (m/z)*	*Δm(u)*	*Assignment*
P1 MALDI-PSD			
72	72.5	0.5	Val Immonium ion
110	110.4	0.46	His Immonium ion
167.2	167.6	−0.4	(a8y4)2 + 1
181.2	181.6	−0.38	(a9y3)2 + 1
187.2	187.6	−0.38	(b5y7)2 + 1
189.2	189.8	−0.61	(b6y6)2 + 1
195.2	195.6	−0.4	(b8y4)2 + 1
209.2	209.6	−3.7	(b9y3)2 + 1
213.2	213.6	−0.4	b2−H_2O
231.2	231.7	−0.5	b2
238.3	238.7	−0.4	(a9y4)3 + 1
248.3	248.7	−0.4	z3−H_2O
266.3	266.7	−0.4	z3
283.3	283.7	−0.4	y3
296.3	296.7	−0.4	(b8y5)3 + 1
300.3	300.7	−0.4	b3−H_2O
318.3	318.8	−0.5	b3
323.3	323.8	−0.5	z4
340.4	340.8	−0.4	y4
355.4	355.8	−0.4	(a8y6)4 + 1
367.4	367.9	−0.5	(b9y5)4 + 1
383.4	383.8	−0.4	(b8y6)4 + 1
389.4	389.9	−0.5	a4
399.4	399.8	−0.4	b4−H_2O
417.4	417.9	−0.5	b4
423.5	423.9	−0.4	y5−H_2O
426.5	426.9	−0.4	(a9y6)5 + 1
441.5	441.9	−0.4	y5
454.5	454.9	−0.4	(a8y7)5 + 1
486.5	486.9	−0.4	b5−H_2O
504.5	505.1	−0.6	b5
528.5	529.0	−0.5	y6
553.6	553.9	−0.3	(b9y7)6 + 1
569.6	569.9	−0.3	(b8y8)6 + 1
587.6	588.1	−0.5	b6−H_2O
610.6	610.2	+4	z7 ??
627.7	628.1	−0.4	y7
644.6	645.0	−0.4	b7−H_2O
662.2	662.9	−0.3	b7
714.8	715.3	−0.5	y8
771.8	772.2	−0.6	a8
781.8	782.1	−0.7	b8−H_2O
799.8	800.2	−0.4	b8
870.9	871.5	−0.4	b9
909.9	909.7	−0.2	b10−H_2O
927.9	928.5	−0.4	b10/z10
P1 ion trap			
283.4	283.2	0.1	y3
340.4	340.1	0.3	y4
398.4	398.8	−0.4	b4−H_2O
417.4	417.0	0.4	b4
424.4	424.3	0.1	z5
441.5	441.2	0.3	y5
454.5	454.2	0.3	(b9y6)5 + 1
510.5	510.4	0.1	y6−H_2O
528.5	528.3	0.2	y6
554.6	554.4	0.1	x6
610.6	610.5	0.1	z7
627.7	628.4	−0.7	y7
644.6	644.1	0.5	b7−H_2O
697.7	697.6	0.1	z8
714.8	714.5	0.3	y8
771.8	771.4	0.4	a8
799.8	799.5	0.3	b8

Table 2—*Continued.*

Calc. mass average (*m/z*)	*Obs. mass* (*m/z*)	*Δm(u)*	*Assignment*
842.8	843.4	−0.6	a9
852.9	853.5	0.6	b9−H_2O
870.9	870.5	0.4	b9
927.9	927.5	0.4	b10/z10
P1 QQQ			
110	110.0	0.0	His Immonium ion
138.2	138.1	0.1	(b8y3)1 +1
144.1	143.9	0.2	b1
167.2	167.0	0.2	(a8y4)2 +1
209.2	209.1	0.1	(b9y3)2 +1
231.2	231.1	0.1	b2
266.3	266.0	0.3	y3-NH3
283.3	283.0	0.3	y3
300.3	300.0	0.3	b3−H_2O
318.3	318.1	0.2	b3
340.4	339.8	0.6	y4
355.4	355.1	0.3	(a8y6)4 +1
367.4	367.1	0.3	(b9y5)4 +1
383.4	383.2	0.2	(b8y6)4 +1
399.4	399.0	0.4	b4−H_2O
417.4	417.2	0.2	b4
441.5	440.9	0.6	y5
454.5	454.2	0.3	(b9y6)5 +1
467.5	467.9	−0.4	x5
486.5	485.9	0.6	b5−H_2O
504.5	504.1	0.4	b5
528.5	528.2	0.3	y6
553.6	553.1	0.5	(b9y7)6 +1
569.6	569.0	0.6	(b8y8)6 +1
587.6	587.1	0.5	b6−H_2O
627.7	627.5	0.2	y7
644.6	644.1	0.5	b7−H_2O
771.8	771.3	0.5	a8
799.8	799.4	0.4	b8
842.8	842.0	0.8	a9
870.9	870.2	0.7	b9
926.9	926.3	0.6	MH−H_2O

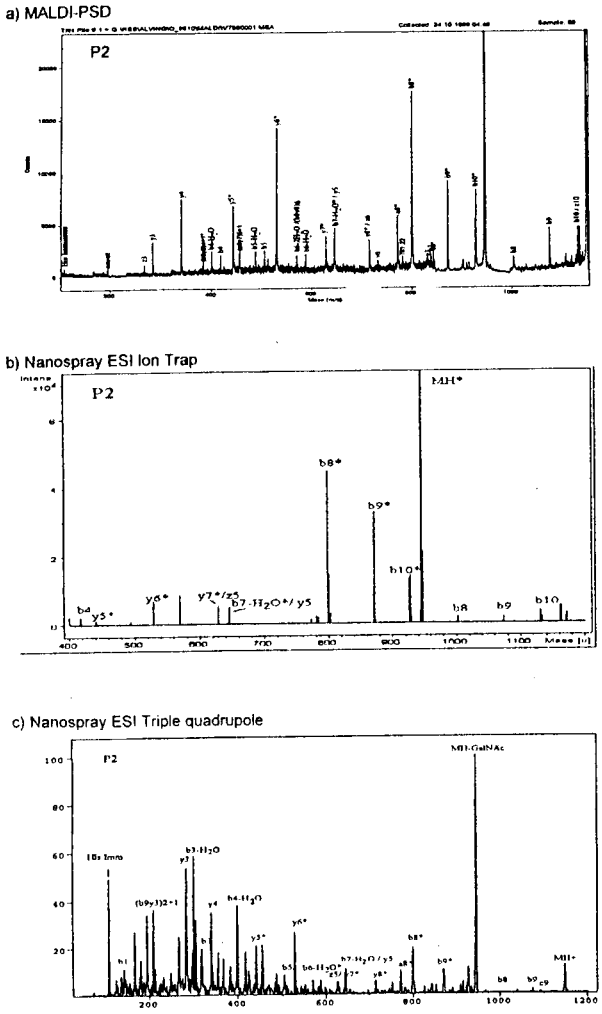

Figure 2. Product ion spectra of Ac-TSSVST(GalNAc) GHAG–NH_2 (P2). (a) PSD-MALDI, precursor ion selected: $[M+H]^+$ (b) Nano-ESI/ion trap, precursor ion selected: $[M+H]^+$. (c) Nano-ESI/Q3, precursor ion selected: $[M+H]^+$.

Ac-TSSVST(GalNAc)GHAG-NH_2 (P2)

Spectra are shown in Fig. 2, interpretations in Table 3. P2 is the glycosylated homolog of P1 carrying a single GalNAc unit on the Thr_6, and leaving other potential glycosylation sites unsubstituted. MALDI-PSD of the P2 glycopeptide ($[M+H]^+ = 1147$) is shown in Fig. 2(a). The most abundant fragment ion is the loss of GalNAc from the molecular ion, beside a number of fragment ions, identical to those from its non-glycosylated counterpart. An asterisk (*) is used to indicate that a fragment that should be glycosylated has lost its sugar. For determination of the glycosylation site, glycosylated fragments are to be used that either include no other potential glycosylation site, or by combination of N- and C- terminus fragments overlap to isolate the glycosylated site from other potential sites. Diagnostic for the identification of the glycosylation site in the position 6 would be the sequence ions x_5, y_5 or z_5, still carrying the attached GalNAc moiety. In the absence of y_5, y_6 or y_7 it would reduce the number of potential glycosylation sites from five to two. In the MALDI-PSD spectrum, the *m/z* value of the potential diagnostic ion y_5 is indistinguishable from that of the nondiagnostic b_7-H_2O* ion. They differ by 0.01 mass unit solely, y_5 with a theoretical value at $m/z = 644.3004$ and b_7-H_2O* ion at $m/z = 644.2891$ (both monoisotopic). This signal was also present in the ESI-ion trap spectrum and in the ESI/Q3 of the P2 glycopeptide (Fig. 2b and c). The presence of y_7 in MALDI-PSD reduces the number of potential glycosylation sites from five to two.

In order to assign the m/z 644 ion either to y_5 or to b_7-H_2O*, ESI Fourier-transform ion-cyclotron mass spectrometry (ESI-FT-ICR-MS), operated in SORI (sustained off resonance irradiation) CID mode was applied, according to its ability to determine the mass values with high precision in order to discriminate between these two fragment ions with different isotopic distribution.

FT-ICR-ESI-SORI-CAD-MS of P2

This experiment was carried out on a Bruker FT-ICR-MS instrument equipped with a 3 T hyperconducting

Table 3.

Calc. mass average (m/z)	*Obs. mass. (m/z)*	*Δm(u)*	*Assignment*
P2 MALDI-PSD			
110	110.4	−0.4	His Immonium ion
266.3	266.5	−0.2	z3
283.3	283.5	−0.2	y3
340.4	340.4	0.0	y4
383.4	383.5	−0.1	(b8y6)4 + 1*
399.4	399.4	0.0	b4−H_2O
417.4	417.4	0.0	b4
423.5	423.6	−0.1	y5−*
441.5	441.8	0.3	y5*
454.5	454.5	0.0	(a8y7)5 + 1
486.5	486.5	0.0	b5−H_2O
504.5	504.5	0.1	b5
510.5	510.6	−0.1	y6−H_2O*
528.5	528.5	0.0	y6*
553.6	553.6	0.0	b9y7)6 + 1*
569.6	569.6	0.0	(b8y8)6 + 1*
587.6	587.4	0.2	b6−H_2O*
627.7	627.6	0.1	y7*
644.6/644.5	655.5	0.1/0.0	b7−H_2O*/y5
662.6	662.4	0.2	b7*
714.8	714.5	0.3	y8*
731.5	731.5	0.0	y6
771.8	771.4	0.4	a8*
799.8	799.3	0.5	b8*
830.7	830.7	0.0	y7
837.6	837.2	0.4	a7
842.8	842.6	0.2	a9*
852.9	852.7	0.2	b9−H_2O
870.9	870.5	0.4	b9*
909.9	908.4	1.5	b10−H_2O*
927.9	927.3	0.6	b10*
944.9	944.4	0.5	MH*
1002.8	1002.8	0.0	b8
1073.9	1073.6	0.3	b9
1130.9	1130.6	0.3	b10/z10
P2 ion trap			
417.4	416.8	0.6	b4
441.5	441	0.5	y5*
528.5	528.5	0.0	y6*
627.7/627.4	628.3	0.9/0.6	y7*/z5
644.6/644.5	645.2	0.6/0.7	b7−H_2O*/y5
771.8	772.3	−0.5	a8*
799.8	799.4	0.4	b8*
870.9	871.7	−0.8	b9*
927.9	927.8	0.1	b10*
944.9	944.7	0.2	MH*
1002.8	1002.5	0.3	b8
1073.9	1073.5	0.4	b9
1130.9	1130.4	0.5	b10
P2 QQQ			
110	110.2	−0.2	His Immonium ion
144.1	144.2	−0.1	b1
167.2	167.0	0.2	(a8y4)2 + 1
209.2	209.0	0.2	(b9y3)2 + 1
213.2	213.0	0.2	b2−H_2O
266.3	266.1	0.2	y3-NH3
283.3	283.1	0.2	y3
300.3	299.9	0.4	b3−H_2O
318.3	318.0	0.3	b3
340.4	339.8	0.6	y4
355.4	355.1	0.3	(a8y6)4 + 1*
367.4	367.2	0.2	(b9y5)4 + 1*
383.4	383.2	0.2	(b8y6)4 + 1*
399.4	398.9	0.5	b4−H_2O
417.4	417.0	0.4	b4
441.5	441.5	0.0	y5*
454.5	454.0	0.5	(a8y7)5 + 1*
486.5	485.7	−0.2	b5−H_2O
504.5	504.4	0.1	b5
528.5	528.1	0.4	y6*
587.6	587.3	0.3	B6−H_2O*
627.4/627.7	626.9	0.5/0.8	z5/y7*
644.6/644.5	644.4	0.2/0.1	b7−H_2O/y5
714.8	715.7	−0.9	y8*
771.8	771.4		a8*
799.8	799.2	0.6	b8*
842.8	842.5	0.3	a9*
870.9	870.1	0.8	b9*
927.9	927.1	0.8	b10*
944.9	944.5	0.4	MH*
1002.8	1003.2	−0.4	b8
1073.9	1073.5	0.4	b9
1090.9	1090.4	0.5	c9

Table 3—*Continued*

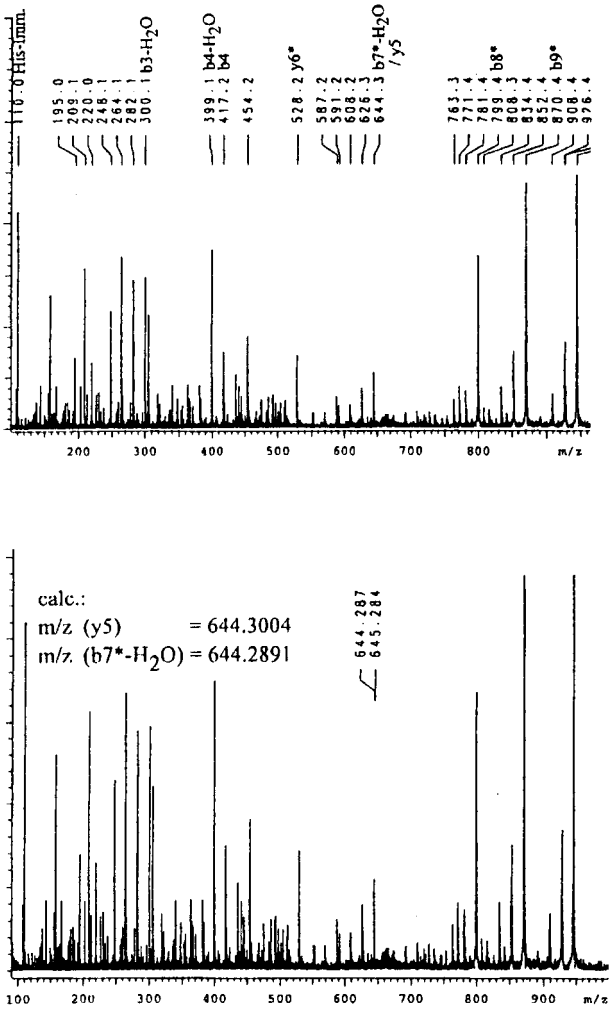

Figure 3. ESI-FT-ICR-SORI-CID mass spectrum of Ac-TSSVST(GalNAc)GHAG–NH_2 (P2). The ion at $m/z = 644$ was attributed to the (b7−H_2O)* according to the experimentally determined value of $m/z = 644.2891$. The alternative attribution of this ion to the y5 fragment at $m/z = 644.3004$ could be ruled out according to the internal calibration using dominant fragment ions.

magnet, using ESI desorption in a SORI-CID mode. The result is presented in Fig. 3. Using internal calibration with prominent b- and y-type ions it was possible to assign the signal at $m/z = 644$ as a $b_7 - H_2O^*$ ion.

Ac-GHAT(GalNAc)SLPVTG-NH_2 (P9)

Spectra are shown in Fig. 4, interpretations in Table 4. The actual glycosylation site in P9 is the first threonine from the N-terminus. MALDI-PSD of the P9 glycopeptide ($[M + H]^+ = 1183$) is shown in Fig. 4a. From three potential glycosylation sites only one, Thr_4, is glycosylated by GalNAc. The evidence for the identification of the glycosylation site should be documented from the fragment ion b_4 still carrying the sugar moiety. In the depicted spectrum, b_4 was present in its deglycosylated form only; however, the next sequence ions, b_{5-8}, could be used for assignment, leaving the possibility of glycosylation on Ser_5 open. In ESI ion trap and ESI Q3 spectra the relevant b_4 ions of low abundance were present, as well as the b_6 ions (Fig. 4b and c). The most abundant ion in all three spectra was the one presenting the deglycosylation of the molecular ion.

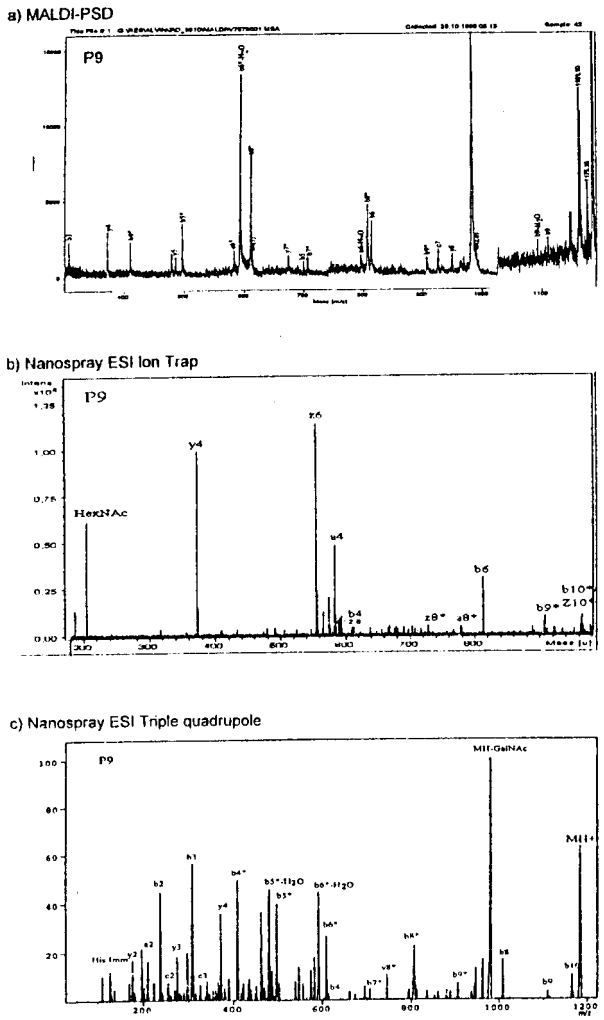

Figure 4. Product ion spectra of Ac-GHAT(GalNAc) SLPVTG–NH_2 (P9). (a) PSD-MALDI, precursor ion selected: $[M + H]^+$. (b) Nano-ESI/ion trap, precursor ion selected: $[M + H]^+$. (c) Nano-ESI/Q3, precursor ion selected: $[M + H]^+$.

Table 4.

Calc. mass average (m/z)	*Obs. mass (m/z)*	*Δm(u)*	*Assignment*
P9 MALDI-PSD			
308.3	308.4	−0.1	b3
372.4	372.5	−0.1	y4
409.4	409.5	−0.1	b4*
478.5	478.5	0.0	b5−H_2O*
496.5	496.4	0.1	b5*
485.6	485.5	0.1	y5
581.6	581.5	0.1	a6*
591.6	591.6	0.0	b6−H_2O*
609.6	609.5	0.1	b6*
612.4	612.5	−0.1	b4
673.8	673.4	0.4	y7*
699.5	699.4	0.1	b5!
706.7	706.5	0.2	b7*
794.6	794.4	0.2	b6−H_2O
805.9	805.5	0.4	b8*
812.6	812.3	0.3	b6
907	906.3	0.7	b9*
926.8	926.0	0.8	c7
947.9	948.8	−0.9	y8
981.1	980.5	0.6	MH*
990.9	990.4	0.5	b8−H_2O
1092	1091.9	0.1	b9−H_2O
1110	1109.5	0.5	b9
P9 ion trap			
204	204.3	−0.3	HexNAc + H
372.4	372.1	0.3	y4
555.7	555.3	0.4	z6
584.4	583.8	0.6	a4
612.4	612.8	−0.4	b4
727.8	728.2	−0.4	z8*
777.9	777.5	0.4	a8*
812.6	812.4	0.2	b6
907.0	907.7	−0.7	b9*
964.0	965.1	−1.1	b10*/z10*
P9 QQQ			
110	110.1	−0.1	His Immonium ion
159.2	159.3	−0.1	y2
176.2	175.9	0.3	y2
209.2	208.9	0.3	a2
237.2	236.8	0.4	b2
254.2	253.6	0.6	c2
275.3	274.0	1.3	y3
308.3	308.0	0.3	b3
325.3	325.9	0.6	c3
372.4	372.3	0.1	y4
409.4	409.1	0.3	b4*
460.5	460.3	0.2	b5−2H_2O*
478.5	478.1	0.4	b5−H_2O*
496.5	495.8	0.7	b5*
555.7	555.2	0.3	z6*
572.7	573.3	−0.6	y6*
581.6	581.6	0.0	a6*
591.6	591.1	0.5	b6−H_2O*
609.6	609.2	0.4	b6*
612.4	613.3	−0.9	b4
706.7	706.4	0.3	b7*
744.9	744.1	0.8	y8*
805.9	805.6	0.3	b8*
907.0	905.9	1.1	b9*
947.9	947.0	0.9	y8
963.1	962.4	0.7	MH−H_2O*
981.1	980.0	1.1	MH−H_2O*
1008.9	1008.1	0.8	b8
1110	1109.7	0.3	b9
1167	1165.3	1.7	b10

Ac-AT(GalNAc)SLPVTDTG-NH_2 (P11) and Ac-ATSLPVT(GalNAc)DTG-NH_2 (P12)

Spectra for P11 are shown in Fig. 5, those for P12; Fig. 6, and interpretations are given in Table 5 and 6. P11 and P12 are topological isomers of the same peptide ($[M + Na]^+ = 1227$), showing in both their glycosylated/non-glycosylated forms high affinity for cationization. The removal of sodium ions did not succeed, either by RP-HPLC using elution with aq. 0.1% TFA, or by use of ion exchange beads. High affinity for sodium ions was also observed in the case of another MUC4 peptide, Ac-SLPVTDTSSG–NH_2 (spectra not shown), containing an overlapping SLPVDT sequence with P11 and 12, a possible site of affinity for sodium ions. If the two latter amino acids are missing, as in P9, the affinity for sodium ions is decreased significantly. This aspect was not further pursued experimentally. Instead of $[M + H]^+$ ions, the sodiated molecular ions as precursor ions were used. They require significantly more energy to fragment, giving rise to highly noisy spectra, an observation already reported by others [11]. Retention of sodium on fragment ions was observed as well.

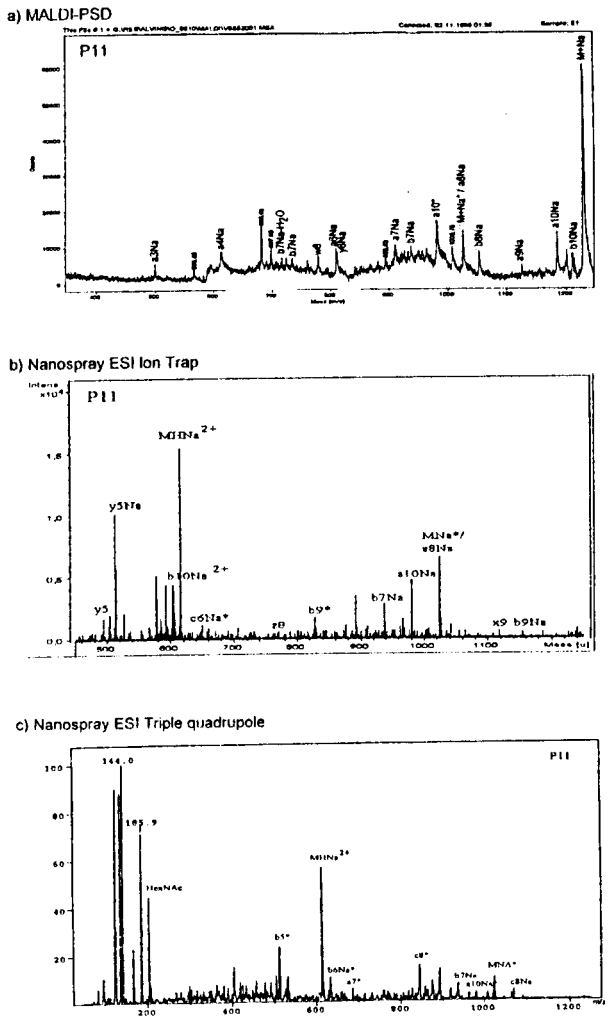

Figure 5. Product ion spectra of Ac-AT(GalNAc)SLPVTDG–NH_2 (P11). (a) PSD-MALDI, precursor ion selected: $[M + Na]^+$. (b) Nano-ESI/ion trap, precursor ion selected: $[M + H + Na]^{2+}$. (c) Nano-ESI/Q3, precursor ion selected: $[M + 2H]^{2+}$.

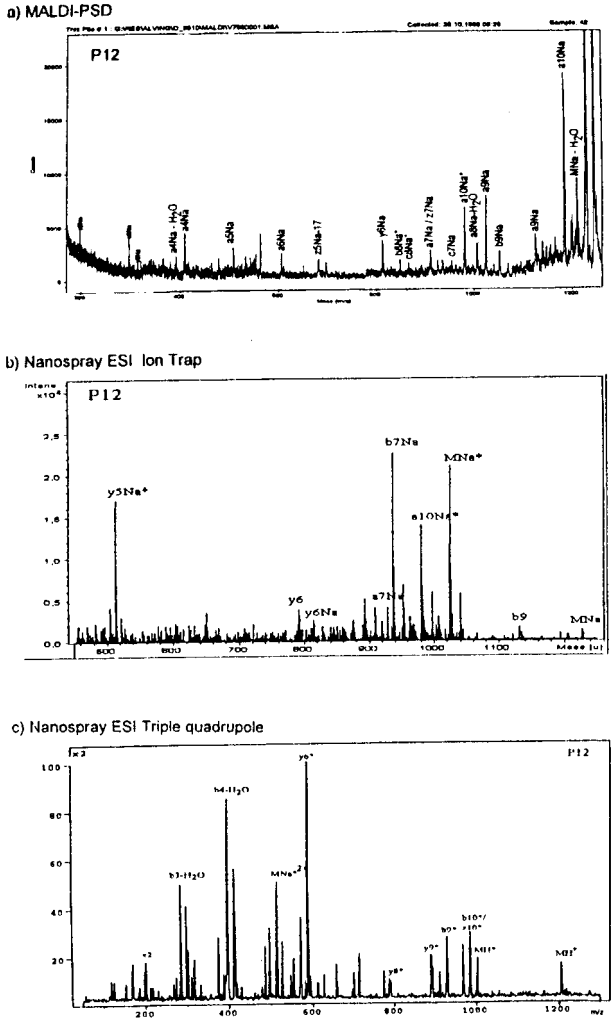

Figure 6. Product ion spectra of Ac-ATSLPVT(GalNAc) DTG–NH_2 (P12). (a) PSD-MALDI, precursor ion selected: $[M + Na]^+$. (b) Nano-ESI/ion trap, precursor ion selected: $[M + Na]^+$. (c) Nano-ESI/Q3, precursor ion selected: $[M + H]^+$.

For the MALDI-PSD sequencing of P11 already a short N-terminal sequence can give diagnostic ions for the identification of the single out of four potential glycosylation sites, namely at the Thr_2. In Fig. 5(a) the a and b series of fragment ions appear as sodiated and glycosylated species in order to demonstrate the glycosylation site near the N-terminus. For the distinction of glycosylation between Thr_2 or Ser_3 the crucial fragment b_2 ion was not present in spectra. The glycosylated ions a_3Na and b_4Na provide evidence for both alternatives. The abundant y_5Na ion in the ESI ion trap spectrum, is an isobaric structure with the doubly charged $[M + H + Na - GalNAc]^{2+}$ (Fig. 5b). An isotope pattern distribution with 0.5 Da mass differences between the peaks proves the later correct. Likewise presence of c_6Na (ion trap) and b_4 (triple quadrupole) reduces the number of potential glycosylation sites to Thr_2/Ser_3.

Diagnostic ions for the P12 sequence are those from the N-terminus at position 7 or 8, in addition to those from the C-terminus at position 4 or 5. The metastable

Table 5.

Calc. mass average (m/z)	*Obs. mass (m/z)*	*Dm(u)*	*Assignment*
P11 MALDI-PSD			
499.2	499.6	−0.4	a3Na
567.6	566.7	−0.1	w6Na
612.4	612.4	0.0	a4Na
734.8	734.0	0.4	b7Na*
777.8	778.2	−0.4	w8Na
808.7	808.8	−0.1	a6Na
810.9	810.6	0.3	y8Na
880.9	880.7	0.2	w9Na
893.8	893.8	0.0	d7Na
909.8	910.6	−0.8	a7Na
937.8	937.6	0.2	b7Na
980.8	981.3	−0.5	d8Na
1007.8	1008	−0.2	a8Na−NH_3
1025.1/1024.8	1025	0.1/−0.2	Mna*/a8Na
1052.9	1053.1	−0.2	b8Na
1126	1125.7	0.3	a9Na
1183	1184.1	0.9	a10Na
1211	1210.0	1.0	b10Na/z10Na
P11 ion trap			
491.5	490.8	0.7	y5
513.5	512.9	0.4	y5Na
606	605.6	0.4	B10HNa2 +
614.6	615.4	−0.8	MHNa2 +
650.7	649.5	1.2	c6Na*
771.8	770.5	1.3	z8
827.4	828.6	−1.2	b8*
937.8	936.7	1.1	b7Na
979.5	980.8	−1.3	a10Na*
1025.1/1024.8	1024.8	0.3/0.0	MNa*/a8Na
1118.4	1118.2	0.2	x9
1154	1154	0.0	b9Na
P11 QQQ			
204	203.9	0.1	HexNac
302.3	303.0	0.7	b3*
418.2	418.1	0.1/0.3	b4/x4
512.5	512.6	−0.1	$[M + H + Na]^{2+}$*
633.7	633.0	0.7	b6Na*
684.8	685.9	−1.1	a7*
844.9	844.0	0.9	c8*
919.8	918.4	1.4	b7Na−H_2O
937.8	936.6	1.2	b7Na
980	980.3	−0.3	a10Na*
1025.1	1024.7	0.4	MNa*
1069.9	1070.2	−0.3	C8Na

fragmentation of P12, presented in Fig. 5(a), gives rise to a series of sodiated a type ions in the MALDI-PSD. The diagnostic a_7Na ion is isobaric with the z_7Na and therefore only of limited value. The c_7 fragment ion in combination with the y_6Na at proline cleavage site, however, offers clear-cut evidence for the glycosylation of the Thr_7. In the ESI ion trap spectrum the evidence on GalNAc attachment site is presented according to the y_6Na and b_7Na ions (Fig. 6b). The ESI Q3 spectrum does not offer relevant evidence (Fig. 6c).

Sequential losses of carboyhydrate residues from the molecular ion were observed by all groups working on *O*-glycopeptide sequencing using different instrumentation. Although this aspect can be applied in an useful way for sequencing of carbohydrate chains, for determination of glycosylation sites; however, this feature represents an obstacle. Only by using a combination of assignments, even for ions of low abundance, it is in most cases possible to identify one of the several possible sites of glycosylation or to reduce their number.

Another major problem in mass spectrometric sequencing is the assignment of m/z values to isobaric structures. This aspect can be met using high-resolution mass spectrometry, as in our case, the FT-ICR-MS, feasible for precise mass determination of fragment ions containing different isotopic composition. Another option would be to apply MS^3 to the relevant fragment ion using either FT-ICR-MS or quadrupole ion trap. MS^3 of glycosylated fragment ions in the quadrupole ion trap is in most real cases not feasible owing to the rather low abundance of these fragment ions in MS^2. Chemical modification of the N- or C-terminal followed by a second MS/MS analysis represents a third choice.

Blocked termini in peptides with N-acylation and C-amidation promote in high-energy CID the fragmentation of the peptide chain due to the even charge distribution over the amide bonds, depending on the acidity and basicity of the amino acids involved [12]. A recent low-energy CID study showed reduction in fragmentation with increasing basicity of the N- and C-terminal amino-acids for doubly charged [13]. They show furthermore increased fragmentation of singly charged peptides when primary amine groups are acetylated.

Relative position of proline in the peptide chain regarding to the N- and C-terminus seems to be important owing to the preferential cleavage at its N-terminal side [14]. Furthermore, the presence of proline is thought to induce a general destabilization of the peptide chain in its vicinity, favorable for fragmentation. A high number of proline residues in the amino acid chain will therefore give rise to multiple carbohydrate-retaining cleavages and enhance the determination of glycosylation sites. Successful sequencing of glycopeptides with "unfavourable" amino acid sequences will–in practice–require the strategy in which a combination of methods and a general good knowledge in interactive use of mass spectrometric and other instruments are to be applied.

A higher degree of internal fragmentation in MALDI-PSD, also observed and discussed by others [11, 14], in comparison to a negligible amount hereof in ESI-ion trap MS/MS and ESI Q3 MS/MS, can be rationalized by taking the considerably shorter time for energy accumulation in the MALDI-TOF ion source into account. Activation in MALDI is thought to take place during plume expansion by collisions between ions and neutrals. The time window for acceleration in MALDI-PSD is in the nanosecond range, while the time for metastable post-source fragmentation is in the 30–300 msec range. The yield of PSD fragmentation has been shown to be dependent on the field strength of the extraction field [11].

The time needed to achieve an energy level sufficient to fragment is significantly longer in the ion trap (the Bruker Esquire ion trap operated with 10 msec time allowed–by the software–for excitation of the intact parent ion). Excitation in the ion trap is performed by

Table 6.

Calc. mass average (*m*/*z*)	*Obs. mass* (*m*/*z*)	Δ*m*(*u*)	*Assignment*
P12 MALDI-PSD			
198.2	198.4	−0.2	y2Na
296.2	296.4	−0.2	a3Na
313.3	313.5	−0.2	y3Na
391.4	391.5	−0.1	a4Na−H_2O
409.4	409.4	0.0	a4Na
506.5	506.3	0.2	a5Na
539.5	540.5	−1.0	x5 + Na*
551.6	552.6	−1.0	c6Na
605.7	605.3	0.4	a6Na
681.1	681.1	0.0	z5Na−H_2O
813.6	812.9	0.7	y6Na!
849.9	849.2	0.7	b8Na*
866.9	867.0	−0.1	c8Na*
909.8/909.8	910.6	−0.8/−0.8	z7Na/a7Na
926.8	925.9	0.9	y7Na
954.8	954.2	0.6	c7Na
980	980	0.0	a10Na*
996.8	996.7	0.1	z8Na
1007.1/1006.8	1006	1.1/0.8	MNa*−H_2O/a8Na−H_2O
1025.1/1024.8	1024.1	1.0/0.7	MNa*/a8Na
1052.9	1052.4	0.5	b8Na
1069.9	1070.5	0.6	c8Na
1126	1125.3	0.7	a9Na
1154	1154.8	−0.8	b9Na
1183	1183.1	0.1	a10Na
1210.1	1209.4	0.7	MNa−H_2O
P12 ion trap			
512.7	513.1	−0.4	$(M + H + Na)^{2+}$ − GalNAc
791.6	791.8	−0.2	y6
813.6	814.0	−0.4	y6Na
909.8	909.0	0.8	a7Na
937.8	937.0	0.8	b7Na
952.8	952.9	−0.1	x7Na
980	981	−1.0	a10Na*
1025.1	1025	0.1	MNa*
1131.5	1131.2	0.3	b9
P12 QQQ			
169.1	170.4	1.1	a2−H_2O
202.0	200.9	1.3	x2
284.3	283.8	0.5	b3−H_2O
317.1	317.6	−0.5	x3
397.4	396.8	0.6	b4−H_2O
512.3	513.9	−1.6	b5
513.1	513.9	−0.8	MNa*2 +
588.6	588.3	0.3	y6*
788.9	787.7	1.2	y8*
890.0	889.1	0.9	y9*
929.0	927.7	1.3	b9*
986.0	984.5	1.5	b10*/z10*
1003.1	1002.8	0.4	MH*

applying the low-amplitude dipolar field resonance frequency, specific for the precursor ion, to the endcaps. In this way product ions are not affected by the precursor ion resonance frequency. The probability for a molecule to accumulate sufficient energy for the double fragmentation needed for the formation of internal fragments prior to the first fragmentation even is negligible, which is also reflected in the fragment patterns obtained.

In the triple quadrupole instrument, the use of the heavier Ar as a collision gas and the option to vary the gas pressure and collision energy over a broader range. offers a more efficient energy accumulation per time in Q^2, used as a collision chamber, than in the ion trap. Contrary to the excitation in quadrupolar ion trap, not only the precursor ion, but also fragments hereof, are excited in the linear quadrupole.

The time available for fragmentation in Q^2 is quite limited (20–30 msec) [11], since it must be concluded before entering the second analyser Q_3. In MALDI-PSD the available time scale for PSD is limited by the

length of the flight-tube before reaching the reflector, and by the extraction potentials, while in an ion trap for a fragmentation event in principle an unlimited time scale is allowed.

In MALDI-PSD spectra of P1 and P2, relatively large amount of internal fragments are detectable, while the two other techniques hardly show any internal fragmentation. Contrary to observations by Rouse *et al.* [15] the amount of the C-terminal ions was not significantly lower in MALDI, compared with the ESI spectra from the two low-energy CID instruments presented here. The main difference between product ion spectra of the nonglycosylated P1, P8 and P10, and the glycosylated peptides was the presence of predominantly low-intensity glycosylated fragment ions. The glycosylation did not seem to have major influence on the presence and intensity of the non- or deglycosylated fragments.

Comparing the experimental handling of the triple quadrupole set-up obtained with Ar as a collision gas with that of the ion trap obtained with He buffer-collision gas, it can be noted that to adjust the intensity of high or low mass fragments–especially for the energy-demanding cationized P11 and P12–was easier using the triple quadrupole. Collision conditions representing the best compromise between presence of most or all peptide backbone fragments, and highest possible intensity of those, can be found [16]. CID conditions optimal for peptide backbone fragmentation yield are not always optimal for formation of glycosylated fragments though. In some cases for general sequencing MS^n would be necessary to obtain complete peptide chain sequence using the ion trap. In our case the determination of the glycosylation sites was slightly more successful using the ion trap, probably because of higher sensitivity of this instrument compared to the Q3 used (Table 7).

CONCLUSIONS

It has been shown that using the PSD-MALDI, ESI Q3 or ESI-IT techniques, it is possible to determine the glycosylation site in P9 and P12 and to reduce the number of potential glycosylation sites from four and five to two in P2 and P11, respectively.

The presence of proline is hypothesized to increase the probability for sucessful detection of the glycosylation sites by destabilization of the peptide chain.

For glycopeptide sequencing high-impact requirements on instrumentation are top sensitivity and wide range of collision energy settings.

When loss of sugar is a predominant peak in the fragment spectra, it follows that:

(1) The presence of *A nonglycosylated fragment cannot be used* to exclude a potential glycosylation site;
(2) The only way to demonstrate the actual glycosylation site in a monoglycosylated peptide is (a) to obtain *fragment ions reaching the actual glycosylation sites and* (b) to obtain *combinations of N- and C-terminal fragment ions overlapping exclusively at the actual glycosylation site.*

Acknowledgements

We thank Prof. Peter Roepstorff for generous support for use of instruments. The Finnigan TSQ 700 mass spectrometer was granted by Danish Natural Science Research Council. The Bruker Esquire Ion Trap mass spectrometer was on loan from Bruker Franzen (Bremen, Germany), and the PerSpetive Voyager Elite on loan from Perseptive Biosystems (Framingham, MA, USA). The financial support for running costs was provided by the Danish Biotechnology Programme.

The help with ESI-FT-ICR-SORI-CID-MS experiment by Dr Walter Jertz at Bruker Franzen, Bremen, Germany, is gratefully acknowledged.

Table 7 Efficiency of Determination of Glycosylation Sites

	MALDI-PSD	*Ion trap*	*Triple-quadrupole*
P2	Reduced to S_5 or T_6 (y7)	None	None
P9	Reduced to T_4 or S_5 (b5)	Identified (b4)	Identified (b4)
P11	Reduced to T_2 or S_3 (b_3)	Red to T_2 or S_3 (c6Na)	Red to T_2 or S_3 (b4/b7Na)
P12	Identified (c7 and y6Na)	Identified (y6Na and b7Na)	None

REFERENCES

1. F.-G. Hanisch, T. R. E. Stadie, F. Deutzmann and J. Peter-Katalinic, *Eur. J. Biochem;* **236**, 318 (1996).
2. T. R. E. Stadie, W. Chai, A. M. Lawson, P. G. H. Byfield and F.-G. Hanisch, *Eur. J. Biochem;* **229**, 140 (1995).
3. J. Peter-Katalinic, F.-G. Hanisch, G. Uhlenbruck and H. Egge, in *Protein Glycosylation: Cellular, Biotechnological and Analytical Aspects*, H. S. Conradt, ed., GBF Monographs no. **15**, p. 179 (1991).
4. M. S. Stoll, E. F. Hounsell, A. M. Lawson, W. Chai and T. Feizi, *Eur. J. Biochem.*, **189**, 499 (1990).
5. F.-G. Hanisch and J. Peter-Katalinic, *Eur. J. Biochem.*, **205**, 527 (1992).
6. J. Peter-Katalinic, A. Ashcroft, B. Green, F.-G. Hanisch, Y. Nakahara, H. Ilijima and T. Ogawa, *Org. Mass Spectrom.*, **29**, 747 (1994).
7. J. Peter-Katalinic, K. Williger, H. Egge, B. Green, F.-G. Hanisch and D. Schindler, *J. Carbohydr. Chem.*, **13**, 447 (1994).
8. G. J. Rademaker, J. Haverkamp and J. Thomas-Oates, *Org. Mass Spectrom.*, **28**, 1536 (1993).
9. K. F. Medzihradszky, B. L. Gillece-Castro, R. R. Townsend, A. L. Burlingame and M. R. Hardy, *J. Amer. Soc. Mass Spectrom.*, **7**, 313 (1996).
10. S. Goletz, B. Thiede, F.-G. Hanisch, M. Schulz, J. Peter-Katalinic, S. Müller, O. Seitz and U. Karsten, *Glycobiology*, **7**, 881 (1997).
11. R. Kaufmann, B. Spengler and F. Lützenkirchen, *Rapid Commun. Mass Spectrom.*, **7**, 902 (1993).
12. K. Biemann, *Methods Enzymol.*, **193**, 351 (1990).
13. S. G. Summerfield and S. J. Gaskell, *Int. J. Mass. Spectrom. Ion Proc.*, **165/166** 509 (1997).
14. I. A. Papayannopoulos, *Mass Spectrom. Rev.*, **14**, 49 (1995).
15. J. B. Rouse, W. Yu and S. A. Martin, *J. Amer. Soc. Mass Spectrom.*, **6**, 822 (1995).
16. I. Haller, U. A. Mirz and B. T. Chait, *J. Amer. Soc. Mass Spectrom.*, **7**, 677 (1996).