Analytical Biochemistry **290,** 47–59 (2001)
doi:10.1006/abio.2000.4955, available online at http://www.idealibrary.com on IDEAL®

Glycoprotein Identification and Localization of O-Glycosylation Sites by Mass Spectrometric Analysis of Deglycosylated/Alkylaminylated Peptide Fragments

Franz-Georg Hanisch,*[,1] Marko Jovanovic,† and Jasna Peter-Katalinic†
*Institute of Biochemistry, Medical Faculty of the University, Joseph-Stelzmann-Strasse 52, 50931 Köln, Germany; and
†Institute of Medical Physics and Biophysics, University of Münster, Robert-Koch-Strasse 31, 48149 Münster, Germany

Received May 25, 2000

In-gel digestion of densely O-glycosylated proteins, an essential step in proteome analysis, is often hampered by steric hindrance of the proteases. To overcome this technical problem a simple and convenient method has been developed, which combines several advantages: (1) Approximately 70% of the oligosaccharides are cleaved without significant protein hydrolysis at the optimal reaction conditions of 70% ethylamine, and quantitative cleavage is achieved with 40% methylamine, at 50°C. (2) To the unsaturated derivatives of Ser and Thr the alkylamine is added as a label of previous O-glycosylation sites. (3) The alkylaminylated protein is effectively cleaved by proteolysis. (4) The modified peptides are identified by MALDI mass spectrometry under consideration of incremental mass increases. (5) The alkylamine label is stable under MALDI post-source-decay analysis as well as in collision-induced dissociation experiments allowing sequencing and peptide localization of O-glycosylation sites. Applicability of the method is evaluated with a series of synthetic glycopeptides, the densely O-glycosylated human glycophorin A, and with the mucin MUC1 from human milk fat globule membranes.

Key Words: **glycoproteins; O-glycosylation; mass spectrometry; proteome research.**

Analysis of proteomes, the total protein equivalent of a cellular genome, has been developed in recent years to a sophisticated, largely automated, technological level. Using subtractive approaches the actual patterns of protein expression in a cell can be defined under different metabolic conditions or in a differentiation-dependent manner. The protocols generally start with two-dimensional gel electrophoresis, followed by in-gel proteolysis of the protein bands and the mass spectrometric identification of the eluted fragments. An inherent technical problem in proteome analysis arises from the fact that most proteins are posttranslationally modified, in particular by N, O-glycosylation. While only about 10% of the data bank entries were defined as glycoproteins, it has been estimated recently that more than half of all proteins are glycoproteins (1). Proteins with a high carbohydrate content, in particular those with clustered O-linked glycans, are difficult to detect in gels. They do not stain well with Coomassie and often appear as diffuse bands due to a heterogeneity of the glycans. Accordingly, heavily glycosylated proteins may be lost or may present difficulties in cutting the gel pieces for in-gel digestion. A further problem, which is caused by dense O-glycosylation, is the resistance of the protein core to proteolytic cleavage, which is a prerequisite for the elution of fragments out of polyacrylamide gels and their identification by mass spectrometric analysis. Also in cases, where limited proteolytic cleavage occurs, the peptides are difficult to identify due to their sizes and to the heterogeneity introduced by complex O-glycosylation. To overcome this problem complete or partial chemical deglycosylation can be performed by several alternative approaches. However, none of these has so far been demonstrated to be applicable in protocols for proteome analysis.

Reductive alkali-catalyzed β-elimination (2) and hydrazinolysis (3) represent well-established methods for the effective liberation of O-linked oligosaccharides; however, both suffer from the drawback that the structural integrity of the protein core is not retained.

[1] To whom correspondence should be addressed. Fax: +49 221 478 6977. E-mail: franz.hanisch@uni-koeln.de.

0003-2697/01 $35.00

By contrast, trifluoromethanesulfonic acid has been reported to deglycosylate proteins without severe affection of their primary structures (4), thereby allowing their ready proteolytic cleavage. If optimized for a particular glycoprotein and carefully controlled, this approach leads to the formation of partially deglycosylated products with only minimal losses of core-GalNAc, which can serve for the identification of O-glycosylation sites. The resulting glycopeptides exhibit, however, poor qualities with respect to their mass spectrometric analysis, i.e., large molecular sizes and strong polarities. In particular, a further disadvantage lies in substantial losses of O-linked sugars during CID[2] or post-source-decay (PSD) experiments for the purpose of site determination. Moreover, partial de-O-glycosylation is difficult to control leading either to glycopeptide products with residual complex glycosylation and the resulting impairment of mass spectrometric or Edman sequencing or to cleavage of core-GalNAc residues with the consequence of underestimation of O-glycosylation sites. A further point, which restricts applicability of this method, is the necessity to exclude moisture from the sample.

A precursor of the present approach has been reported by Rademaker *et al.* (5), which combines base-catalyzed β-elimination with the addition of ammonia to the unsaturated amino acid derivatives. The aminylated products principally allow proteolytic cleavage, but also the identification of previous O-glycosylation sites. However, the mass difference of only one mass unit between -OH and -NH_2 groups renders mass spectrometric sequencing at low resolution, such as CID with triple quadrupoles or MALDI/PSD, difficult. However, for high-resolution instruments like Qtof a mass difference of 1 Da would be sufficient. The method suffers also from the drawback that a limited exchange of free hydroxyl groups occurs during prolonged incubation times resulting in artificial overestimation of O-glycosylation sites (unpublished observations). Ethylamine has been previously used for the de-O-glycosylation of proteins, however, with the aim of obtaining reducing oligosaccharides and, hence, the published data exclusively refer to the glycan portion (6).

The alkylamine label introduced by addition of methylamine or ethylamine into the elimination products combines several advantages: (1) It makes proteolytic hydrolysis easier or even possible for glycoproteins otherwise resistant to proteolysis. (2) The label is chemically stable under the conditions of MALDI- and ESI mass spectrometry and also under CID or PSD conditions when data for the sequencing are acquired. This compares to the partial gas-phase elimination of O-linked sugars via a rearrangement reaction. The alkylamine label allows for the ready site identification by mass spectrometric sequencing based on incremental mass increase. Compared to aminylation the mass shift is more easily detectable, but the total mass of the alkylaminylated peptides is still considerably smaller than that of the parent glycopeptides. This again favors mass spectrometric analysis with respect to mass accuracy and sensitivity. (3) The alkylaminylated peptides are less polar than the corresponding glycopeptides. The label, moreover, introduces a positive charge into the peptide. Both modifications enhance the yields of primary ions detectable in the positive-ion mode of MALDI-MS and ESI-MS and improves detection by reducing the glycosylation-dependent heterogeneity.

MATERIALS AND METHODS

Materials

Glycoproteins and glycopeptides. Human glycophorin A (hGA) from blood group MM donors was purchased from Sigma (Deisenhoven, Germany). hGA from blood group NN donors was isolated from human erythrocyte membranes by phenol-saline extraction followed by gel-exclusion chromatography on Sephadex G200 (7). MUC1 mucin from human milk fat globule membranes was extracted with 4 M guanidinium-HCl and isolated by combined gel-exclusion chromatography on Sephacryl S300 and affinity chromatography on immobilized BC3 antibody (8, 9). The purified mucin was partially deglycosylated by treatment with trifluoromethanesulfonic acid (TFMSA) as described previously (8). Synthesis of glycopeptides A2, A9, H1, and A13, corresponding to one repeat unit of the MUC1 repeat domain and carrying Galβ1-3GalNAc (A2, A9) or GalNAc (A13, H1) in position Thr10 (A2, A13), or Thr5 (H1) or in all five putative positions (A9) of the sequence AHGVTSAPDTRPAPGSTAPPA (AHG21) has been described previously (10, 11).

Amines. Aqueous solutions of amines were purchased from Sigma (ethylamine, 70%; ethanolamine, 99%); Merck, Darmstadt, Germany (ammonia, 25%); Aldrich, Steinheim, Germany (methylamine, 40%); and Fluka, Seelze, Germany (hydroxyl amine, 50%).

Polyacrylamide Gel Electrophoresis and High-Performance Liquid Chromatography

Glycoproteins were separated by polyacrylamide slab gel electrophoresis in the presence of SDS. The samples (1–10 μg) were analyzed in 0.75 mm 10%

[2] Abbreviations used: CID, collision-induced dissociation; PSD, post-source-decay; MALDI, Matrix-assisted laser desorption ionization; ESI-MS, electrospray ionization-mass spectrometry; hGA, human glycophorin A; TFMSA, trifluoromethanesulfonic acid; SDS, sodium dodecyl sulfate; EA, ethylamine; MA, methylamine; DTT, dithiothreitol; HCCA, α-cyano-4-hydroxycinnamic acid; QTOF, quadrupole/time-of-flight.

TABLE 1

Comparative Study of Base-Catalyzed β-Elimination/Aminylation Using Different Amines

Base	Structure	% aq. solution	Molarity (pH)	Reaction parameters	Yield of deglycosylated product (%)	Peptide degradation
Ammonia	NH_3	25	13.36	45°C, 18 h	100	Weak
Methylamine	CH_3NH_2	40	11.62	50°C, 18 h	100	Strong
				50°C, 6 h	100	Weak to moderate
Ethylamine	$CH_3CH_2NH_2$	70	12.36	50°C, 18 h	70–80	Weak
Ethanolamine	$HOCH_2CH_2NH_2$	70	11.52 (13.5)	50°C, 18 h	50	Weak
Hydroxylamine	$HONH_2$	50	16.95 (11.0)	50°C, 18 h	0	Very strong

polyacrylamide gels (hGA) or in gradient gels with 3.5 to 15% polyacrylamide (hMUC1) buffered with Tris-glycine, pH 8.9 (0.1% SDS). The proteins were dissolved in a sample buffer system containing 1% SDS in Tris-glycine and were loaded on top of a 3% stacking gel. Electrophoresis was performed in a mini protean cell (BioRad) for 45–60 min at 200 V constant voltage. Proteins were fixed and stained with Serva Blue G (Serva, Heidelberg, Germany).

The glycopeptides or their alkylaminylated derivatives were chromatographed on a narrow-bore ODS Ultrasphere column (150 × 2 mm, Beckman Instruments, Munich, Germany) by gradient elution in a mixture of acetonitrile in water (0.1% trifluoroacetic acid) from 2% (solvent A) to 80% (solvent B) during 80 min at a flow rate of 0.3 ml/min. The eluting peptides were registered by photometrical detection at 214 nm.

β-Elimination/Alkylaminylation and Proteolytic Digestion of Glycoproteins

β-Elimination/alkylaminylation. Gel pieces corresponding to coomassie blue positive protein bands in polyacrylamide slab gels were washed in 50% aqueous methanol (0.5 ml) for 1 h followed by drying in a speed vac. The dried gel was taken up in 100–300 μl of 70% ethylamine (EA) or 40% methylamine (MA) and incubated at 50°C for 18 or 6 h, respectively. The liquid

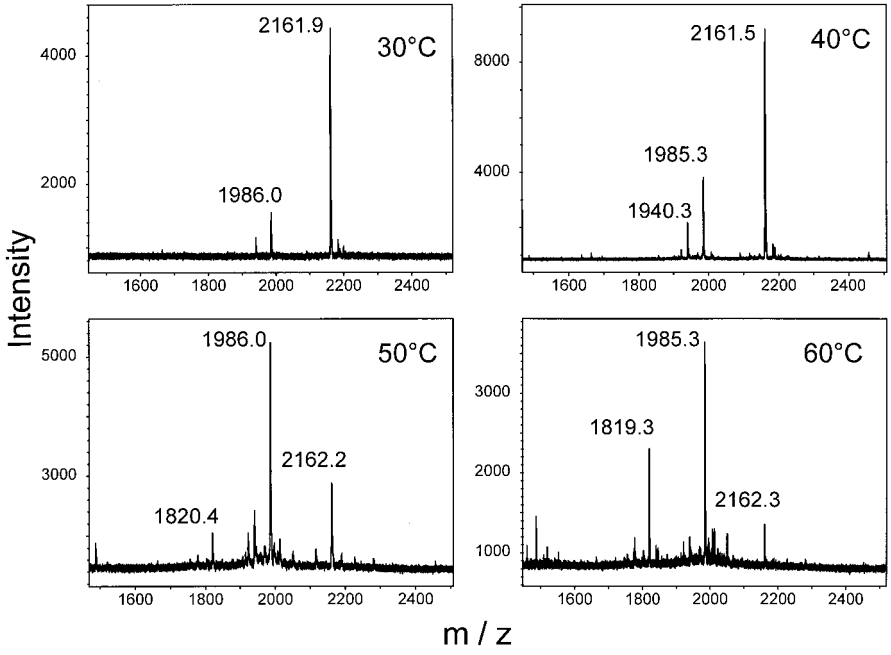


FIG. 1. Temperature dependency of β-elimination/ethylaminylation. Synthetic MUC1 glycopeptide A13 (carrying GalNAc at Thr10 of the peptide sequence AHGVTSAPDTRPAPGSTAPPA, average mass of MH 2162.3) was treated with 70% ethylamine at varying temperatures for 24 h. The dried samples were taken up in 20 μl 0.1% TFA and analyzed by reflectron MALDI mass spectrometry. The ion signals with average masses of *m*/*z* 2162 and 1986 correspond to the molecular ions MH^+ of the glycopeptide and of the derived alkylaminylated product, respectively.

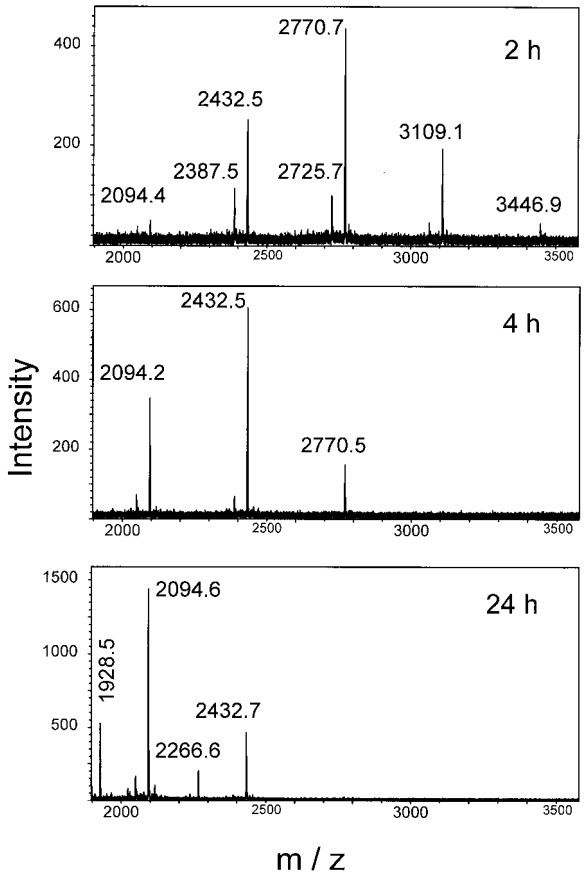


FIG. 2. Time dependency of β-elimination/ethylaminylation. Synthetic MUC1 glycopeptide A9 (carrying Gal-GalNAc disaccharides at Thr5, Ser6, Thr10, Ser16, and Thr17 of the peptide sequence AHGVTSAPDTRPAPGSTAPPA, average mass of MH 3786.1) was treated with 70% ethylamine at 50°C for varying time intervals. The molecular ions MH^+ measured after 2 h correspond to ethylaminylated products of A9 with four (*m*/*z* 3446.9), three (*m*/*z* 3109.1), two (*m*/*z* 2770.7), one (*m*/*z* 2432.5) and no residual glycan chains (*m*/*z* 2094.4). Signals at MH-45 (*m*/*z* 2387.5, 2725.7) correspond to partially deglycosylated peptide lacking addition of one EA residue.

phase containing the liberated glycans was removed and both phases were separately dried in a speed vac.

β-Elimination/alkylaminylation of glycopeptide A2 (10 μg), human glycophorin A (NN) (1–10 μg), native MUC1 from human milk fat globules (25 μg), and of partially deglycosylated MUC1 (10 μg) was also performed in solution under the same conditions as those described above. The alkylaminylated proteins were digested with trypsin (hGA) or with clostripain (hMUC1).

Trypsin digestion. The dried gel was repeatedly washed with water, dried again, and finally reswollen in 20 μl 0.1 M NH_4HCO_3, pH 7.8, containing 2 μg of sequencing grade trypsin (Promega). After fill-up of the reaction volume to 50 μl with the same buffer, the gel was incubated for 2 h at 37°C. The proteolytic fragments were eluted after addition of water (0.5 ml) or mixtures of water/acetonitril by overnight incubation at room temperature. The eluted peptide fragments were dried and finally solubilized in 0.1% trifluoroacetic acid/water. Occasionally, the peptides were desalted prior to mass spectrometric analysis by ZipTip clean-up according to the manufacturer's instructions (Millipore).

Clostripain digestion. Partially deglycosylated MUC1 or the alkylaminylated MUC1 derivative were digested with clostripain (Promega, Madison, WI) to obtain PAPGSTAPPAHGVTSAPDTR (PAP20) fragments from the tandem repeat region. Prior to proteolytic digestion, clostripain was activated for 3 h at room temperature in the presence of 5 mM DTT and 2 mM Ca-acetate. Deglycosylated ethylaminylated MUC1 was resuspended in 100 μl of 20 mM sodium-phosphate buffer, pH 7.6, containing 5 mM DTT and 0.2 mM Ca-acetate, and mixed with activated clostripain. The

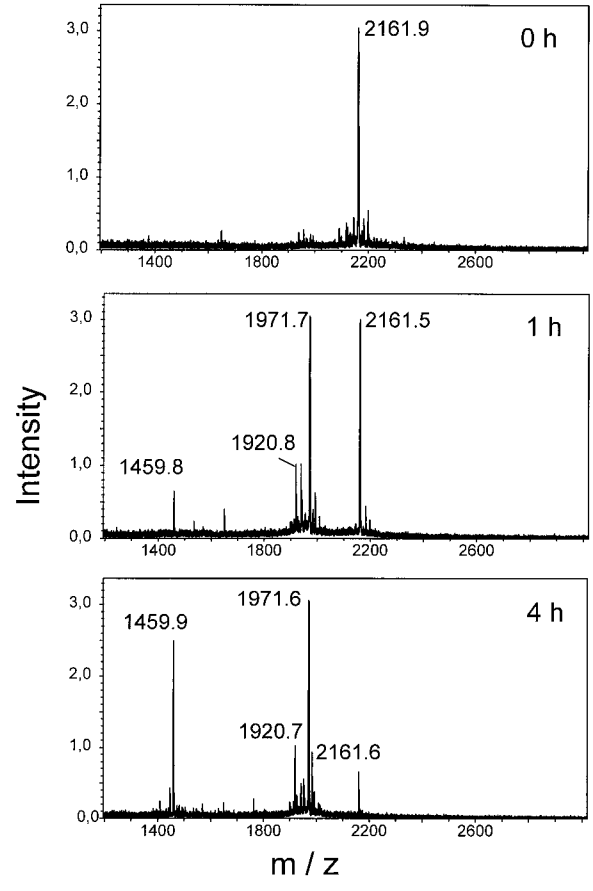


FIG. 3. Time dependency of β-elimination/methylaminylation. Synthetic MUC1 glycopeptide H1 (carrying GalNAc at Thr5 of the peptide sequence AHGVTSAPDTRPAPGSTAPPA, average mass of MH 2162.3) was treated with 40% methylamine at 50°C for varying time intervals. The molecular ions MH^+ correspond to the glycopeptide (*m*/*z* 2162) and to the methylaminylated product (*m*/*z* 1971). The ion signal at *m*/*z* 1460 represents a fragment derived by cleavage of the Gly15–Ser16 bond.

TABLE 2

MALDI Mass Spectrometric Analysis of Proteolytic Fragments from De-*O*-Glycosylated/Ethylaminylated Glycoproteins

Glycoprotein	Proteolytic treatment	Fragment	Number of EA groups	Detected as monoisotopic MH^+ at m/z (calculated monoisotopic mass)	
h Glycophorin A	Trypsin	19–30	2	1433.6 (1434.7)	
		32–39	1,2	921.3 (921.4)	948.5 (948.5)
		31–49	2	2212.0 (2212.2)	
		40–49	1,2	1153.6 (1153.6)	1179.6 (1180.6)
		50–61	1	1433.6 (1433.6)	
		40–61	2,3	2568.9 (2568.3)	2595.9 (2595.3)
		101–131	0	3308.9 (3307.6)	
		102–131	0	3180.8 (3179.5)	
h MUC1	Clostripain	PAP20-0	0	1887.5 (1886.9)	
		PAP20-1	1	1914.3 (1913.9)	
		PAP20-2	2	1941.3 (1940.9)	
		PAP20-3	3	1968.4 (1968.0)	
		PAP20-4	4	1995.5 (1995.0)	
		PAP20-5	5	2022.5 (2022.1)	

reaction mixture was incubated for 18 h at 37°C. Digestion was terminated by the addition of 10 mM EDTA and the digest was subjected to MALDI mass spectrometry after ZipTip clean-up.

Mass Spectrometric Analyses

MALDI-MS. The matrix HCCA (α-cyano-4-hydroxycinnamic acid; saturated solution in acetonitril: 0.1% trifluoroacetic acid, 2:1) was mixed on the target 1:1 by volume with 0.5-μl sample in 0.1% TFA and air-dried. Reflectron and PSD mass spectrometric analyses were performed on a Bruker Reflex III instrument (Bruker Daltonik, Bremen, Germany) using a pulsed laser beam (nitrogen laser, $\lambda = 337$ nm). Residual gas pressure was at $<1.0 \times 10^{-7}$ mbar. Ion spectra were recorded in the positive-ion mode. Acceleration and reflector voltages were set to 28.5 and 30 kV, respectively. The molecular parent ion was isolated with a pulsed field by deflection of all other ions. Complete PSD fragment ion spectra were obtained by stepwise reducing the reflector voltage to produce overlapping mass ranges and combination of the spectral sections (8).

The relative masses of major ion signals in the presented spectra correspond to monoisotopic or average masses depending on the type of experiment. The relative masses of ion signals registered during optimization of reaction parameters for β-elimination/alkylaminylation (refer to Figs. 1 to 3) and of post-source-decay fragments in the sequencing experiments (refer to Tables 3A and 3B to 5) correspond to average masses. Relative masses used for the identification of peptides after proteolytic digestion correspond to monoisotopic masses (refer to Table 2 and to Fig. 4).

QTOF-ESI MS. Positive ion electrospray mass spectrometry was conducted on a QTOF instrument (Micromass, Manchester, UK) (12). QTOF is a quadrupole/time-of-flight hybrid mass spectrometer fitted with a Z-spray source. A Kopf vertical pipette puller, Model 720 (David Kopf Instruments, CA), was used for production of nanospray glass capillaries. Capillary voltage applied on an internal wire electrode was 1100 V, while 40 V was applied to the cone. Peptide sequencing was conducted on doubly charged cation precursors, using argon as the collision gas, under the pressure of 4×10^{-3} mbar. The collision energy used was 25–40 V. The samples were dissolved in water, giving a concentration of 20 pmol/μl. Five microliters was used for the entire experiment. Relative masses of major ion signals presented in Figs. 5 and 6 correspond to monoisotopic masses.

RESULTS

Base-Catalyzed β-Elimination by Primary Amines

A series of primary amines was tested initially for comparison of the potential to liberate O-linked glycans from peptides and to introduce a label into the previously glycosylated position. The reaction products were analyzed by MALDI mass spectrometry with respect to the yields of aminylated products, but also with respect to the identification of by-products from side reactions, including uncontrolled peptide cleavage. The results are summarized for each base in Table 1. The commercially available aqueous base solutions were used at molar concentrations ranging from 11.5 to 13.4 M except for hydroxylamine, which was 17 M. De-O-glycosylation and formation of the respective aminylated products were quantitative in solutions of ammonia and methylamine after overnight incubation. While ammonia caused limited side reactions and resulted in a small series of by-products, methylamine-catalyzed de-O-glycosylation was accompanied by

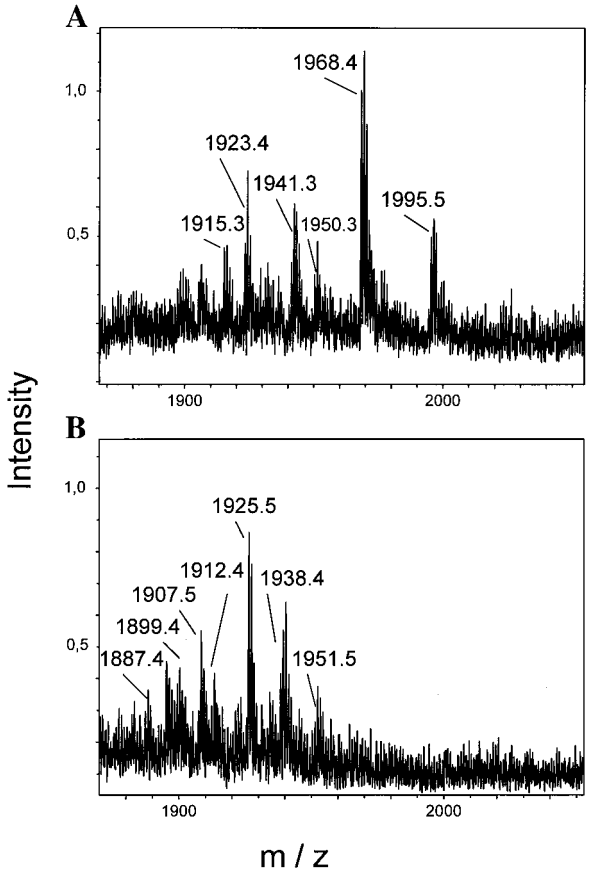


FIG. 4. Alkylaminylation and clostripain digestion of human MUC1 from milk fat membranes. Clostripain cleaves the MUC1 repeat domain into icosapeptides of the sequence PAPGSTAPPAHGVTSAPDTR (calculated monoisotopic mass of MH 1886.9). (A) Ethylaminylated products were identified via their average masses to be substituted with one (*m*/*z* 1915.3), two (*m*/*z* 1941.3), three (*m*/*z* 1968.4), and four EA residues (*m*/*z* 1995.5); (B) Methylaminylated products were identified via their average masses to be substituted with no (*m*/*z* 1887.4), one (*m*/*z* 1899.4), two (*m*/*z* 1912.4), three (*m*/*z* 1925.5), four (*m*/*z* 1938.4), and five MA residues (*m*/*z* 1951.5). A series of signals at MH-45 or -31 indicates incomplete alkylamine addition.

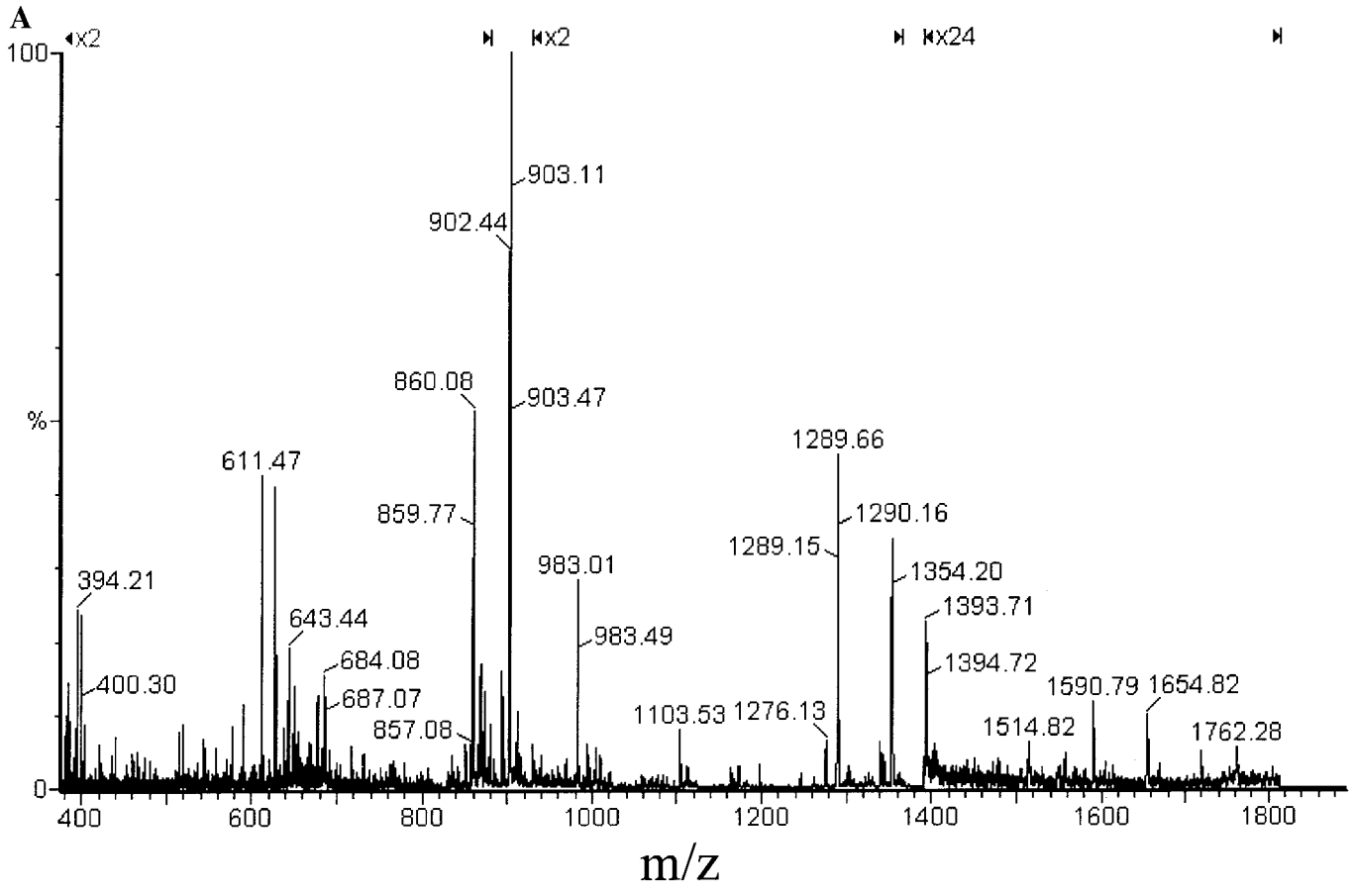
heavy peptide fragmentation during prolonged reaction times. Ethylamine did not yield complete conversion into the respective products, but under optimum conditions peptide fragmentation was weak. The results obtained with ethanolamine were comparable to those for ethylamine; however, the yields were significantly lower. Hydroxylamine cannot be used in this context, since the weak base is unable to liberate *O*-linked glycans to a measurable extent. Moreover, besides its known qualities in selective peptide fragmentation hydroxylamine causes extensive nonspecific cleavage of peptidyl bonds.

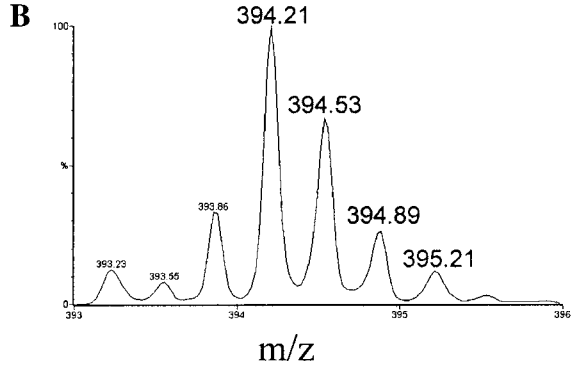
Optimization of Reaction Parameters for the Alkylamine-Catalyzed De-O-Glycosylation

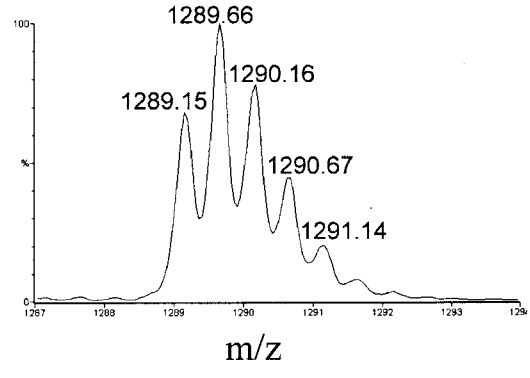
De-O-glycosylation of synthetic MUC1 glycopeptide A13 (GalNAc substituted at Thr10 of the sequence AHGVTSAPDTRPAPGSTAPPA, calculated monoisotopic/average mass of MH 2161.2/2162.3) was used for the optimization of parameters in ethylamine-catalyzed β-elimination and addition reactions. The base concentration was varied from 50 to 70% corresponding to 8.8–12.4 mol/L. Decreasing the ethylamine concentration to 50% resulted in complete cleavage of GalNAc from the glycopeptide, however, a series of unidentified by-products was formed by successive elimination reactions (not shown). In a 70% solution of base the conversion of glycopeptide into its ethylaminylated product was incomplete (approx 70%); however, the only considerable by-product formed during overnight incubation was the unsaturated amino acid derivative indicating incomplete ethylamine addition. The temperature of the reaction was varied over a range from 30 to 60°C using 70% ethylamine in overnight incubations (Fig. 1). Under these conditions product formation showed a strong temperature dependency with only 10 and 30% yields at 30 and 40°C, respectively. At higher temperatures conversion into the ethylaminylated product was 70% (50°C) and 90% (60°C), respectively. However, above 50°C a considerable by-product formation occurred.

Time-dependency of the reaction was followed by using the penta-O-glycosylated MUC1 peptide A9 (Gal-GalNAc disaccharides substituted at all five Ser/Thr positions within the peptide AHGVTSAPDTRPAPGSTAPPA, calculated monoisiotopic/average mass of MH 3784.9/3786.1). A mixture of partially de-O-glycosylated, ethylaminylated products was obtained at 50°C, which was dominated by peptide with one (1 h), three (2 h), four (4 h, 8 h), and five EA groups (24 h). De-O-glycosylation of four chains occurred rapidly within a few hours, while liberation of the fifth chain took considerably longer and did not reach completion after 24 h (Fig. 2). Methylamine-catalyzed β-elimination exhibited a similar temperature profile with optimal yields at about 50°C. Kinetic studies using H1 glycopeptide (the same peptide backbone as in A13 with GalNAc substitution at Thr5) revealed, however, that methylamine- vs ethylamine-catalyzed de-O-glycosylation/alkylaminylation was faster and complete

FIG. 5. QTOF ESI mass spectrum of hGA peptides after in-gel ethylaminylation and trypsin digestion. Deconvolution of doubly and triply charged ions revealed a pattern of relative monoisotopic masses for MH^+, which was largely identical to that measured by reflectron MALDI-MS. For fragment identification of MH^+ ion masses refer to Table 2.


A
x2
x2
x24
100
%
0
903.11
902.44
860.08
903.47
611.47
859.77
1289.66
1290.16
1289.15
983.01
1354.20
394.21
1393.71
643.44
983.49
1394.72
684.08
400.30
687.07
1590.79
1654.82
857.08
1103.53
1276.13
1514.82
1762.28
400
600
800
1000
1200
1400
1600
1800
m/z


B
394.21
394.53
394.89
395.21
m/z


1289.66
1290.16
1289.15
1290.67
1291.14
m/z

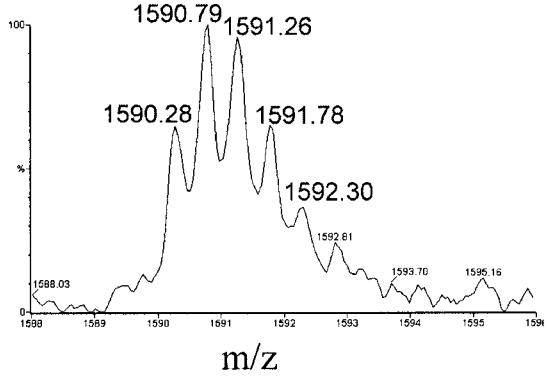

1590.79
1591.26
1590.28
1591.78
1592.30
m/z

TABLE 3A

Fragment Masses Measured in the Post-Source-Decay Mode at *m/z* 1985 for EA-Substituted A13 Peptide

		Masses of C-terminal fragments			Masses of N-terminal fragments		
y_n	Amino acid sequence	EA groups 0	1		EA groups 0	1	b_m
1	A						21
2	P	(187.1)(y)					20
3	P	284.9(y)					19
4	A	(355.1)(y) z					18
5	T						17
6	S						16
7	G						15
8	P	698.8(y)					14
9	A	(769.2)(y)					13
10	P						12
11	R	1022.7(y) z				1120.5(b)	11
12	T		1132.4(z)				10
13	D				836.2(b)		9
14	P		1363.1(y)		722.6(b)		8
15	A				624.8(b), 606.8(b-18)		7
16	S				554.4(b), 536.8(b-18)		6
17	T				466.4(b), 448.6(b-18)		5
18	V				365.3(b) a		4
19	G				266.2(b)		3
20	H				209.9(b)		2
21	A						1

Note. The relative masses *m/z* measured for the fragment ions from the C-terminal (x_n, y_n, and z_n ions and their companions) or N-terminal of A13-EA (a_m, b_m, and c_m ions and their companions) were listed in accordance to the peptide sequence. Masses given in parentheses were of low intensity, while for ions indicated by + accurate mass assignment was impossible. The masses are accompanied by assignments of the fragment species given in parentheses, followed by supporting ions which provide additional evidence for the cleavage at distinct sites.

within 6 h. The high degree of nonselective peptide degradation observable during overnight incubation in methylamine was strongly decreased, if reaction times between 4 and 6 h were used (Fig. 3). The prominent signal at *m/z* 1459.8 (average mass), which was registered already after a 1-h reaction time, corresponds to the AHG15 fragment of H1 (Ala1-Gly15). This fragment arises from a selective cleavage of the peptidyl bond between Gly and Ser in the PGST motif, which is highly susceptible to base-catalyzed hydrolysis. Similar selective cleavage was not observed with other peptide and glycopeptide substrates (data not shown).

Proteolytic Digestion of Ethylaminylated Proteins in Solution

Two heavily O-glycosylated proteins, human glycophorin A and MUC1 from human milk, were used as model glycoproteins for the study of enzymatic hydrolysis after alkylamine-catalyzed de-O-glycosylation. While hGA is a substrate for trypsin even in its native, fully glycosylated state, hMUC1 cannot be digested by any protease unless the glycan chains were reduced to the core-type level or removed completely. De-O-glycosylation/alkylaminylation of the glycoproteins was performed under standard conditions prior to proteolytic cleavage with trypsin (hGA) or clostripain (hMUC1). The mixtures of fragments were analyzed by MALDI mass spectrometry and the results summarized in Table 2. In case of hGA no glycosylated fragments were registered in the mass range up to 6000 mass units. Eight, partially overlapping fragments were identified as ethylaminylated peptides carrying one to three EA groups. The heavily glycosylated N-terminal peptides [1–30] and [1–39], which contain one N-glycosylation site, were not detectable in this mass range.

Native MUC1 from human milk fat globule membranes was treated with ethylamine or methylamine and the alkylaminated derivatives were digested with clostripain overnight. The degree of proteolytic digestion was checked by reversed-phase HPLC of the EA- or MA-modified PAP20 peptides (data not shown). MALDI mass spectrometry of the two preparations revealed that the alkylaminylated protein core of MUC1 was effectively digested by the protease into PAP20 peptides (Table 2, Fig. 4). A series of MH^+ ions at *m/z* 1914.3, 1941.3, 1968.4, 1995.5, and 2022.5 (monoisotopic masses, Table 2) indicated the formation of PAP20 peptides modified by addition of one, two, three, four, and five EA groups, respectively. The corresponding series of MH^+ ions in the methylamine-catalyzed reaction was found at *m/z* 1899.4, 1912.4, 1925.5, 1938.4, and 1951.5 (monoisotopic masses, Fig. 4B). In

TABLE 3B

Fragment Masses Measured in the Post-Source-Decay Mode at *m/z* 2162 for GalNAc-Substituted A13 Peptide

		Masses of C-terminal fragments		Masses of N-terminal fragments		
y_n	Amino acid sequence	GalNAc 0	1	GalNAc 0	1	b_m
1	A					21
2	P	187.6(y)				20
3	P	284.4(y)				19
4	A	338.1(z)				18
5	T					17
6	S					16
7	G					15
8	P	697.9(y)				14
9	A					13
10	P	866.3(y)				12
11	R					11
12	T	1124.3(y) y-18		937.4(b) b-18		10
13	D			837.9(b)		9
14	P		1539.1(y)			8
15	A		1609.2(y)	625.3(b), a, b-18		7
16	S			535.4(b-18)		6
17	T		+(y)	467.2(b)		5
18	V			365.7(b) a		4
19	G			266.8(b)		3
20	H		2092.8(y)	209.3(b)		2
21	A					1

Note. The relative masses *m/z* measured for the fragment ions from the C-terminal (x_n, y_n, and z_n ions and their companions) or N-terminal of A13-GalNAc (a_m, b_m, and c_m ions and their companions) were listed in accordance to the peptide sequence. Masses given in parentheses were of low intensity, while for ions indicated by + accurate mass assignment was impossible. The masses are accompanied by assignments of the fragment species given in parentheses, followed by supporting ions which provide additional evidence for the cleavage at distinct sites.

the spectra of both preparations the PAP20 species with three modified groups represented the most prominent MH^+ ion, which is in accord with the known average density of O-glycosylation per repeat peptide (8). No signals that matched the calculated masses of PAP20 peptides with residual glycosylation were detectable in the higher mass range up to *m/z* 3500. In the lower mass range a series of intense signals dominated the spectrum, which exhibited incremental mass differences of hexoses, deoxyhexoses, hexosamines, and sialic acids and corresponded presumably to partially degraded "peeling" products of the neutral and acidic mucin oligosaccharides.

Deglycosylation of MUC1 repeat peptide in the presence of alkylamine is also indicated by a series of ions corresponding to the β-eliminated products lacking ethylamine (*m/z* 1950.3, 1923.4 in Fig. 4A) or methylamine addition (*m/z* 1907.5, 1894.5 in Fig. 4B).

In-Gel Digestion of Ethylaminylated hGA and Protein Identification by Mass Search

HGA was separated by slab gel SDS polyacrylamide gel electrophoresis and the dominating coomassie-positive band at the apparent molecular mass of 60-kDa (corresponding to the dimer of hGA) was cut for analysis. In-gel trypsin digestion of hGA (between 1 and 10 μg protein) was performed under standard conditions after *in situ* de-O-glycosylation/ethylaminylation. On analysis by QTOF-ESI mass spectrometry (Fig. 5) the mixture of tryptic peptides eluted from the gel revealed a similar pattern of glycopeptide derivatives as listed in Table 2. The masses of the monoisotopic doubly and triply charged ions served for the identification of the protein in a database (NCBInr.8.28.2000). The deconvoluted monoisotopic MH^+ ion masses 1153.60, 1180.60, 1433.65, 3179.56, and 3307.68 were used at 50 ppm mass accuracy together with the calculated masses derived by subtraction of 27 and 54 Da (corresponding to the incremental mass of one or two ethylamine substituents).

Of 63257 entries (proteins of 1000–100,000 Da, full p*I* range, human and mouse proteins) 5 entries were selected that matched with human glycophorin A or preglycophorin A. Four matching masses with *m/z* 1126.57, 1406.65, 3179.56, and 3307.68 were identified with the tryptic peptides [40–49], [50–61], [102–131], and [101–131], respectively.

Sequencing of Ethylaminylated Peptides by PSD MALDI Mass Spectrometry

A comparative analysis of A13 glycopeptide and its ethylaminylated derivative by post-source-decay

TABLE 4

Fragment Masses Measured in the Post-Source-Decay Mode at *m/z* 2570 for EA-Disubstituted HGA Peptide 40-61

		Masses of C-terminal fragments				Masses of N-terminal fragments				
y_n	Amino acid sequence	EA groups	0	1	2	EA groups	0	1	2	b_m
1	R	175.1(y)								22
2	E	286.8(z)							(2395.1)(b)	21
3	G	361.9(y)	x						2238.9(a)b-17	20
4	T									19
5	E	592.3(y)								18
6	E	722.4(y)							1961.5(b-17)	17
7	E								(1849.5)(b)	16
8	P									15
9	P	1044.0(y)	z						1604.2(b-18)	14
10	Y								1542.5(c)	13
11	V	(1306.6)(y)								12
12	T			1435.1(y)						11
13	R			1574.2(z)						10
14	V			1674.6(z)					1122.7(c)	9
15	S									8
16	I									7
17	E									6
18	S							552.7(b)a,c		5
19	V					437.0(b)				4
20	E					338.1(b)				3
21	H					209.6(b)				2
22	A									1

Note. The relative masses *m/z* measured for the fragment ions from the C-terminal (x_n, y_n, and z_n ions and their companions) or N-terminal of HGA(40–61)-EA_2 (a_m, b_m, and c_m ions and their companions) were listed in accordance to the peptide sequence. Masses given in parentheses were of low intensity, while for ions indicated by + accurate mass assignment was impossible. The masses are accompanied by assignments of the fragment species given in parentheses, followed by supporting ions which provide additional evidence for the cleavage at distinct sites.

MALDI mass spectrometry revealed that equal molar amounts of the alkylaminylated and of the glycosylated peptides (50 pmol each) yielded fragment spectra of comparable ion intensities. However, in the case of the EA-modified peptide a 3- to 5-fold increase in signal intensity was observed in the low mass range ($m/z <$ 1000). Fragments of higher masses dominated in the PSD spectrum of A13 glycopeptide. The PSD spectrum of A13-EA (*m/z* 1985.3) was characterized by a consistent series of b ions and b-18-ions from the N-terminal, which broke off in the middle of the peptide beyond the site of modification (Table 3). As observed previously (8) the glycosylated peptide yielded a prominent series of y and b ions, which indicate absence of GalNAc, and probably result from post-source cleavage of the sugar. In accordance with this, the MH^+ ion of the deglycosylated A13 peptide at *m/z* 1961.4 was very prominent. By contrast, the ethylaminylated peptide derivative was stable under the conditions of PSD-MALDI-MS. The position of EA modification could be unequivocally assigned to Thr10. (refer to the intense b_{11}-ion with the average mass at *m/z* 1120.5).

Analysis of ethylaminylated peptide fragments from hGA by reflectron mode MALDI mass spectrometry yielded a series of intense MH^+ signals, which were structurally assigned according to their masses (Table 2). The ion registered at *m/z* 2568.9 (monoisotopic mass), which corresponds to tryptic peptide [40–61] and carries two EA groups, was sequenced in the PSD mode to localize the sites of alkylaminylation (Table 4). Fragments from the N-terminal (a, b, and c ions) revealed evidence of the EA substitution of peptide position Ser5 and Ser8, in accordance with Edman degradation data (13, 14). The series of C-terminal fragments (x, y, and z ions) indicated EA modification of Thr11, which is also in accord with previous sequence data (13, 14). However, the number of EA-modified sites identified by mass spectrometric fragmentation does not agree with the molecular mass of the parent ion, which corresponds only to two EA substitutions. This discrepancy can be explained by site heterogeneity of glycosylation on individual hGA molecules.

Human MUC1 was partially deglycosylated down to the level of core-GalNAc prior to proteolytic digestion of the repeat domain into PAP20 glycopeptides (see above). De-O-glycosylation/ethylaminylation yielded a series of derivatives, which were identified by reflectron mode MALDI-MS. The intense MH^+ ion at *m/z* 1941.3 (monoisotopic mass) corresponding to PAP20-

TABLE 5

Fragment Masses Measured in the Post-Source-Decay Mode at *m/z* 1942 for EA-Disubstituted PAP20 Peptide

		Masses of C-terminal fragments		Masses of N-terminal fragments		
y_n	Amino acid sequence	EA groups 0	1	EA groups 0	1	b_m
1	R					20
2	T	277.2(y)				19
3	D					18
4	P					17
5	A	559.6(y)				16
6	S	648.5(y)				15
7	T		774.9(y)			14
8	V					13
9	G					12
10	H		1067.4(y)			11
11	A					10
12	P					9
13	P		1334.2(y)		707.8(b)	8
14	A		(1405.2)(y)	583.6(b)		7
15	T			511.1(b)		6
16	S					5
17	G					4
18	P			265.9(b)		3
19	A			169.3(b)		2
20	P					1

Note. The relative masses *m/z* measured for the fragment ions from the C-terminal (x_n, y_n, and z_n ions and their companions) or N-terminal of PAP20-EA_2 (a_m, b_m, and c_m ions and their companions) were listed in accordance to the peptide sequence. Masses given in parentheses were of low intensity, while for ions indicated by + accurate mass assignment was impossible. The masses are accompanied by assignments of the fragment species given in parentheses, followed by supporting ions which provide additional evidence for the cleavage at distinct sites.

EA_2 was sequenced in the PSD mode of MALDI-MS to localize the sites of modification (Table 5). Although the y and b ion series were not consistent and did not cover the entire peptide sequence, one of the two positions (Thr14) could be identified as a substitution site by the relative masses of y_7, y_{10}, y_{13}, and y_{14}. The second EA group should be localized to Ser5 or Thr6, but no definitive assignment was possible due to contradictive b_6 and b_8 ions.

Sequencing of Ethylaminylated Peptides by QTOF-ESI Mass Spectrometry

MS/MS analysis of peptide A13 (EA) by QTOF-ESI mass spectrometry revealed that the total sequence of the peptide was covered within a single experiment. Both series, the y and the b ions, were nearly complete. Identification of the substitution site was possible on the basis of the intense y_{12} ion at the average mass *m/z* 1149.63 (Fig. 6), which represented the first indication of EA substitution within the series by an incremental mass shift of +27 Da. Summarizing the results obtained for a series of glycopeptide-derived EA peptides (data not shown) the MS/MS spectra generally were dominated by the presence of fragment ions belonging to b and y series, about two-thirds of which were detected within a single experiment. Fragment ions belonging to other series were present to much less extent. QTOF-ESI mass spectrometry proved to be a suitable method for defining the site of modification within the peptides. The average mass accuracy of fragment ions was around 50 ppm.

DISCUSSION

A simple and convenient method is described, which allows for (1) the ready de-O-glycosylation of glycoproteins in gels as a prerequisit for the generation of proteolytic peptide “finger prints”, and (2) the localization of previous O-glycosylation sites via mass spectrometric sequencing of alkylaminylated peptide fragments. The method has been tested to fullfil part of the requirements needed in proteome analysis by applying it to the densely O-glycosylated proteins hGA and hMUC1. In the case of methylamine the whole procedure comprising in-gel de-O-glycosylation/alkylaminylation and trypsin digestion can be performed within a working day. The performance does not need experimental expertise or special technical equipment. Since de-O-glycosylation/alkylaminylation is largely quantitative and the introduced label does not hamper the action of the proteases (trypsin, clostripain), the sensitivity of the total analytical procedure will not be limited by the additional steps. In the context of mass

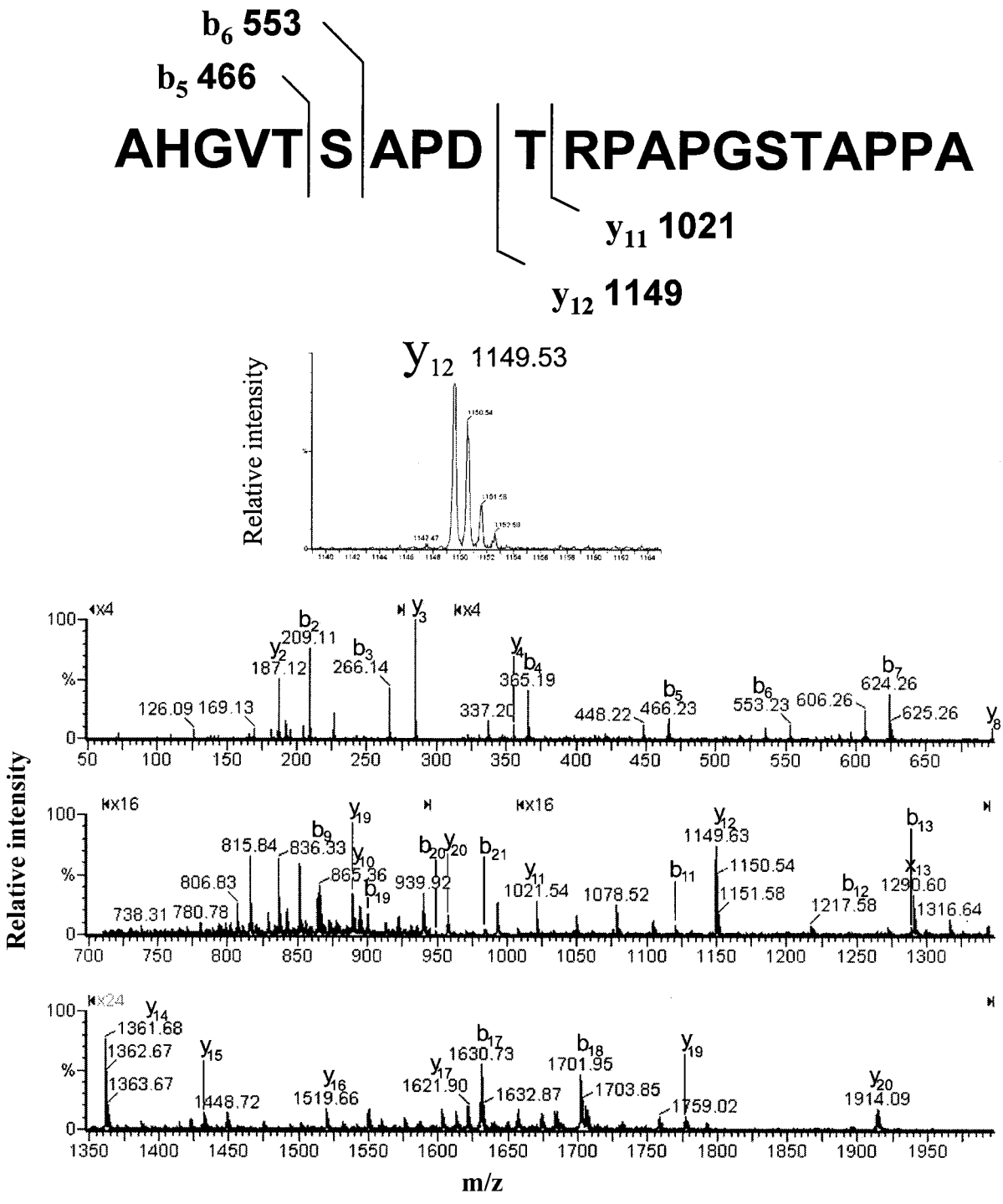


FIG. 6. MS/MS spectrum of peptide A13 (EA) measured by QTOF-ESI mass spectrometry. One hundred percent of the peptide sequence was covered within a single MS/MS experiment. The y series indicated the glycosylation site of the peptide at position Thr10. The substitution site was deduced from the y_{12} ion, the first in this series with an incremental mass shift of +27 Da due to the presence of the ethylamine group.

spectrometric analysis and the sequencing via post-source-decay fragments in MALDI-MS or by CID in ESI-MS there are several advantages resulting from the introduction of alkylamine groups into the peptides. First of all, the yields of MH^+ ions are increased due to the lower polarities of the analytes and to the

positive charge of the added label. The size reduction by liberation of the complex O-linked glycans makes even densely glycosylated peptides accessible to PSD fragment analysis. The small mass increments of +13 and +27 could, however, impair the selection of single parent ions due to the sizes of mass windows in MALDI mass spectrometry, which generally exceed 25 mass units. On the other hand, if a peptide species exists as a consistent series of alkylaminylated derivatives, a discrimination would not be necessary and one PSD spectrum would yield information on all modified positions within the peptide.

A further point of criticism refers to the degree of specific and nonspecific peptide fragmentation during EA and MA treatment. Specific cleavage of the peptide backbone in aqueous alkylamine solutions is not a general problem. Only distinct peptide bonds, which are known to be chemically more labile than others, show considerable cleavage under the conditions used in this study (refer to the preferential cleavage at Gly-Ser within the MUC1 repeat peptide). Nonspecific peptide fragmentation is more pronounced in methylamine solution, however, only during prolonged reaction times. Within a time frame of 4–6 h the degree of nonspecific peptide cleavage is negligible.

Applicability of the method was demonstrated by treatment of the densely O-glycosylated proteins hGA and hMUC1, which both are structurally defined with respect to the sites of glycan attachment (8, 13, 14). The alkylaminylated proteins were cleaved by the respective proteases and the fragment masses matched with the reported patterns and with the expected numbers of alkylaminylated sites.

In the case of hGA the heavily O-glycosylated fragments of the N-terminal, peptides [1–30] and [1–39], were not found in the mass spectra possibly due to incomplete de-O-glycosylation/alkylaminylation or to the presence of N-linked glycan. De-O-glycosylation of proteins in ethylamine (methylamine) suffers from the drawback that β-elimination of the glycans and also the addition of alkylamine are not quantitative. The extent of incomplete conversion into alkylaminylated products is possibly dependent on the peptide localization of glycans. Clustered O-linked chains with more than two adjacent glycans appear to be largely resistant to the treatment as revealed for the N-terminal tryptic peptides of hGA (this study) and for synthetic MUC2 glycopeptides (F.-G. Hanisch, unpublished observations). The inertness of glycopeptides with clustered *O*-glycans could be explained by mesomeric stabilization of the unsaturated intermediate product. Some peptides of hGA show a heterogeneity with respect to the numbers and sites of O-glycosylation. A similar statistical distribution of actual glycosylation sites has been previously revealed for MUC1 repeat peptides. The lactation-associated mucin has been shown to carry about 2.5 glycans per repeat with five potential sites (8). However, these glycans were demonstrated to be located at varying positions resulting in a partial glycosylation of each putative site (8).

The above described method could develop into a valuable tool in proteome analysis (identification of proteins) and in the characterization of site-specific O-glycosylation of proteins. In particular, in the latter context the method could aid comparative analyses of natural and recombinant glycoforms of proteins and allow the determination of host cell-specific substitution densities and patterns.

ACKNOWLEDGMENTS

This study was supported by a grant from the Deutsche Forschungsgemeinschaft Ha2092/4-2. The authors thank Prof. Dr. Hans Paulsen, University of Hamburg, for providing the synthetic glycopeptides and Dr. Stefan Müller for the samples of human MUC1 from milk fat membranes. We also acknowledge the kind support by Dr. Markus Macht, Center of Molecular Medicine, Cologne, during mass spectrometric analysis.

REFERENCES

1. Apweiler, R., Hermjakob, H., and Sharon, N. (1999) *Biochim. Biophys. Acta* **1473,** 4–8.
2. Iyer, R. N., and Carlson, D. M. (1971) *Arch. Biochem. Biophys.* **142,** 101–105.
3. Patel, T., Bruce, J., Merry, A., Bigge, C., Wormald, M., Jaques, A., and Parekh, R. (1993) *Biochemistry* **32,** 679–693.
4. Gerken, T. A., Gupta, R., and Jentoft, N. (1992) *Biochemistry* **31,** 639–648.
5. Rademaker, G. J., Pergantis, S. A., Blok-Tip, L., Langridge, J. I., Kleen, A., and Thomas-Oates, J. E. (1998) *Anal. Biochem.* **257,** 149–160.
6. Chai, W., Feizi, T., Yuen, C. T., and Lawson, A. M. (1997) *Glycobiology* **7,** 861–872.
7. Dahr, W., Uhlenbruck, G., Janssen, E., and Schmalisch, R. (1977) *Hum. Genet.* **35,** 335–343.
8. Müller, S., Goletz, S., Packer, N., Gooley, A. A., Lawson, A. M., and Hanisch, F.-G. (1997) *J. Biol. Chem.* **272,** 24780–24793.
9. Müller, S., Alving, K., Peter-Katalinic, J., Zachara, N., Gooley, A. A., and Hanisch, F.-G. (1999) *J. Biol. Chem.* **274,** 18165–18172.
10. Mathieux, N., Paulsen, H., Meldal, M., and Bock, K. (1997) *J. Chem. Soc. Perkin Trans. I,* 1, 2359–2368.
11. Karsten, U., Diotel, C., Klich, G., Paulsen, H., Goletz, S., Müller, S., and Hanisch, F.-G. (1998) *Cancer Res.* **58,** 2541–2549.
12. Hanisch, F.-G., Green, B. N., Bateman, R., and Peter-Katalinic, J. (1998) *J. Mass Spectrom.* **33,** 358–362.
13. Tomita, M., Furthmayr, H., and Marchesi, V. (1978) *Biochemistry* **17,** 4756–4770.
14. Pisano, A., Redmond, J. W., Williams, K. L., and Gooley, A. A. (1993) *Glycobiology* **3,** 429–435.