ELSEVIER

International Journal of Mass Spectrometry 214 (2002) 37–51

International Journal of Mass Spectrometry

www.elsevier.com/locate/ijms

Review

The use of multiple ion chromatograms in on-line HPLC-MS for the characterization of post-translational and chemical modifications of proteins

Olaf Jahn, Bernhard Hofmann, Olaf Brauns, Joachim Spiess, Klaus Eckart*

Department of Molecular Neuroendocrinology, Max Planck Institute for Experimental Medicine, Hermann-Rein-Str. 3, 37075 Goettingen, Germany

Received 26 October 2001; accepted 13 December 2001

Abstract

In the post-genomic era, the characterization of the post-translational modifications and the three-dimensional structure of proteins will be of increasing interest. The post-translational modifications of proteins such as N-terminal processing, disulfide-bond formation, and glycosylation can be advantageously characterized by peptide mapping monitored with HPLC-MS. Cross-linking between a protein and a ligand can be used to identify contact points and thereby generate constraints for molecular modeling of the ligand–protein interaction. Here we demonstrate the use of multiple ion chromatograms, which represent an extension of selected ion extraction, for the selective detection of low abundant components in peptide mixtures obtained by enzymatic digestion of proteins. The power of this technique will be demonstrated for the characterization of multiple N-terminal processing sites, the usage of putative glycosylation sites, the determination of low abundant disulfide-bond scrambled forms of proteins, and the characterization of photoadducts produced with photoreactive peptide analogs. (Int J Mass Spectrom 214 (2002) 37–51) © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Peptide mapping; *N*-Glycosylation; Photoaffinity labeling; Selected ion extraction; Disulfide-bond structure

1. Introduction

The characterization of the post-translational modifications of a protein is of great importance for an understanding of its function. The N-terminal signal peptide cleavage site and the disulfide bonds are important structural features of proteins. Furthermore, the knowledge on the type and extent of glycosylation of potential glycosylation sites may facilitate the understanding of the biological function of proteins [1]. The characterization of these structural features will become increasingly important with the completion of the sequencing of the human genome [2,3]. The availability of all human DNA sequences in a genomic database generates additional interest in the structural characterization of the corresponding proteins of unknown function [4]. In view of this question, the use of cross-linking agents for the determination of distances of functional groups in native proteins [5] are of great value for the characterization of the tertiary and quarternary structure. Furthermore, photoaffinity labeling is commonly used to identify

* Corresponding author. E-mail: eckart@em.mpg.de

1387-3806/02/$20.00
PII S1387-3806(01)00593-0

contact points between a receptor and its ligand [6].

Peptide maps obtained by enzymatic digestion of proteins are the method of choice to analyze the modifications mentioned above. Reversed-phase high performance liquid chromatography (RP-HPLC) is the most versatile and most commonly used method for the separation of such peptide mixtures [7]. On-line coupling with electrospray (ES) mass spectrometry (ESMS) provides a powerful analytical tool for fast separation and identification of peptides [8]. A protein which is known on the basis of its cDNA sequence can be often unambiguously identified on the amino acid level by a single enzymatic digest analyzed with MS [9,10]. However, structural heterogeneity introduced by multiple N-terminal signal peptide cleavage sites, disulfide-bond scrambling, or incomplete glycosylation may be difficult to be determined and resolved.

The objective of the present study was to evaluate the use of ion chromatograms extracted from full scan HPLC-MS recordings for the identification of structural heterogeneity introduced into proteins by either post-translational or chemical modifications. Thereby, we introduced multiple ion chromatograms (MICs) as an extension of the previously reported method of selected ion extraction of single molecular ions [11,12]. In our laboratory, the use of MICs was established during the protein-chemical characterization of receptor proteins belonging to the corticotropin-releasing factor (CRF) system [13] such as the soluble N-terminal domain of the rat CRF receptor 1 (rCRFR1-NT) [14] and the rat CRF binding protein (rCRFBP) [15]. CRF is a C-terminally amidated 41 amino acid peptide which processes the stress signal within the hypothalamus–pituitary–adrenal (HPA) axis [16,17] and modulates a variety of brain functions such as anxiety, learning, food intake, and locomotor activity [13]. The heterogeneity of N-terminal processing, the extent of usage of the putative *N*-glycosylation sites, and the existence of disulfide-bond scrambling were characterized by generating MICs from HPLC-MS recordings of tryptic digests of rCRFR1-NT. The tryptic digest of a photoadduct formed between rCRFR1-NT produced in bacteria (brCRFR1-NT) and the photoreactive ovine CRF (oCRF) analog 4-(1-Azi-2,2,2-trifluoroethyl)benzoyl-[Tyr^0]oCRF (Diaz-oCRF) [18] was used to demonstrate the capabilities of MICs for the identification of photoreaction products and the resolution of heterogeneity of covalently modified proteins. Similarly, a photoadduct formed by rCRFBP and the photoreactive analog of the CRF-like peptide rat urocortin (rUcn), *para*-benzoylbenzoyl-rUcn (pBB-rUcn), was used to demonstrate the advantage of MICs for the detection of the site of photoincorporation. The advantages of MICs over total ion current (TIC) chromatograms will be discussed in view of detection limits, possibilities of semiquantitative estimations, and separation of co-eluting compounds.

2. Methods

2.1. Protein isolation, derivatization, and digestion

The isolation and characterization of recombinant rCRFR1-NT was recently described [14]. Because of the low production rates of rCRFR1-NT in human embryonic kidney (HEK) 293 cells, rCRFR1-NT was produced in the presence of the mannosidase I inhibitor kifunensine [19]. This treatment changed the type of glycosylation from complex to high mannose type and increased the production rate by two orders of magnitude [14]. All data were obtained using high mannose glycosylated rCRFR1-NT.

The methods for reduction of disulfide bonds and alkylation of Cys residues by *S*-carboxamidomethylation have been already reported [15]. Enzymatic deglycosylation of approximately 30 μg rCRFR1-NT was carried out by incubation with 1000 units endo-β-acetylglucosaminidase H ($EndoH_f$) or 2000 units peptide-*N*-glycosidase F (PNGaseF) for 6 h at 37 °C in 50 mM phosphate buffer pH 7.4 or 5.5, respectively, containing 2 M urea and 2 mM PMSF [14]. $EndoH_f$ hydrolyzes the *O*-glycosidic bond between the two *N*-acetylglucosamine residues linked to Asn, whereas PNGaseF hydrolyzes the *N*-glycosidic bond of Asn linked to *N*-acetylglucosamine and thereby

converts glycosylated Asn residues to Asp [20]. Digestion with TPCK-treated trypsin (Sigma) using an enzyme to substrate ratio of 1:20 was performed at 37 °C for 2–8 h in HEPES buffer (pH 8.0) containing 2 M urea, 5 mM $CaCl_2$, and 5% acetonitrile [14,15].

The production of recombinant rCRFR1-NT in bacteria (brCRFR1-NT) and its purification was recently described [21]. Due to the procaryotic production system, brCRFR1-NT was not post-translationally modified.

The production of recombinant rCRFBP in HEK 293 cells, its purification, and structural as well as pharmacological characterization was subject of a recent report [15].

2.2. Formation of photoadducts

The synthesis of the diazirine photophore 4-(1-Azi-2,2,2-trifluoroethyl)benzoic acid and its coupling to [Tyr^0]oCRF to generate Diaz-oCRF has been already described [18]. The diazirine photophore turned out to be stable under the conditions of peptide synthesis using HBTU activation, peptide cleavage from the resin and deprotection by standard trifluoroacetic acid (TFA) protocols, and RP-HPLC purification using TFA-containing elution systems. Photoadducts were formed by an 1 h incubation of 65 μM brCRFR1-NT and 65 μM Diaz-oCRF in 50 mM Tris buffer containing 5 mM $MgCl_2$ and 2 mM EGTA followed by 30 min irradiation with UV ($\lambda = 365$ nm; Stratalinker, Stratgene) at 4 °C. Unreacted brCRFR1-NT and its photoadduct were separated by HPLC on a Vydac C4 column of 15 cm length and 1 mm diameter (5 μm particles of 300 Å pore size). The photoadduct was digested with trypsin as described above.

pBB-rUcn was synthesized by coupling para-benzoylbenzoyl-*N*-hydroxysuccinimide ester to the de-protected N-terminal amino group of the resin-linked rUcn. The reaction product was cleaved from the resin under standard conditions and purified by preparative RP-HPLC on a Vydac C4 column of 25 cm length and 22 mm diameter (10 μm particles of 300 Å pore size). Approximately 1–10 nM rCRFBP was incubated for 4 h with 50 nM pBB-rUcn under standard binding conditions [15] in a total volume of 400 mL. The mixture was irradiated for 30 min at 0 °C with UV light using a 400 W metal vapor lamp (Osram Ultratech). Wavelength below 300 nm were excluded by a Schott B270 filter screen. Unlabeled rCRFBP and its photoadduct were co-purified by Ni-chelate chromatography utilizing the His-tag fused to the C-terminus of rCRFBP [15]. Unlabeled rCRFBP was separated from its photoadduct on a Vydac C4 column of 15 cm length and 1 mm diameter (5 μm particles of 300 Å pore size). Both protein fractions were digested with trypsin as described above.

2.3. On-line HPLC-mass spectrometry

The micro-HPLC system was composed of an ABI 140B syringe pump and an ABI 759A single wavelength detector (Applied Biosystems) [15]. Gradients formed by 0.07% TFA and 0.05% TFA containing 80% acetonitrile were applied for elution. The flow of the eluent was splitted before the column by using an IC CAP 100 gradient splitter (LC Packings). The samples were injected into the column at 4% acetonitrile and linear gradients of 0.5% acetonitrile per minute up to 60% acetonitrile were performed for elution. The micro-HPLC system was directly connected to the ES interface by using a fused silica capillary of 1 m length and 75 μm inner diameter. Mass spectra were recorded using an AutoSpec-T tandem mass spectrometer (Micromass, UK) equipped with a Micromass ES interface and an array detector in the second mass spectrometer. Exponential scan functions with mass ranges of 2000–400 or 2500–500 and a scan speed of 1 s per decade were used. The delay time was adjusted to obtain a cycle time of 1 s. The collector slit was adjusted to 500 resolution (10% valley definition) and five data points were recorded per mass signal by the OPUS data system (V 3.6). These parameters were chosen to reduce the size of the acquired data files. If necessary, unit mass resolution up to 2000 could be obtained. The acquired data were not centroided. MICs were generated by the OPUS data system using a mass window of ±1000 ppm. Combinations of masses were used as input for the chromatogram

option of the data system. Suitable masses for MICs were obtained by creating single ion chromatograms (SICs) representing the relevant charge states of the molecular ion of interest typically starting with the lowest charge state covered by the mass range of the scan function up to the highest possible charge state, which is usually determined by the number of basic sites in the peptide. By comparison of these SICs the compound of interest can be detected and the relevant molecular ion charge states for the MIC can be selected. In the case of semiquantitative estimations the potential difference between the sampling cone and the skimmer was adjusted to minimize fragmentation. The molecular ions showing a signal to noise ratio >10:1 in the SICs were chosen for the respective MIC.

The Maximum Entropy algorithm [22] included in the OPUS data system was used to deconvolute the ES mass spectra of proteins. The mass difference between two molecular masses was named ΔM.

3. Results

3.1. Characterization of the N-terminal cleavage sites of rCRFR1-NT

Heterogenous signal peptide cleavage C-terminal to Thr^{23} and Ser^{24} of rCRFR1-NT (Fig. 1) was recently demonstrated by Edman sequencing [14]. For further characterization, rCRFR1-NT was deglycosylated with PNGaseF and digested with trypsin. The co-eluting tryptic fragments rCRFR1-NT(24–57) and rCRFR1-NT(25–57) each linked to rCRFR1-NT(77–96) by a disulfide-bond were detected in the TIC chromatogram of the HPLC-MS analysis (Fig. 2A). By MICs, the fragments rCRFR1-NT(24–57)-SS-(77–96) (Fig. 2B, signal 4) and rCRFR1-NT(25–57)-SS-(77–96) (Fig. 2C, signal 5) were detected with absolute intensities of 1.4×10^4 and 9.5×10^3, respectively. On this basis, relative abundances of 60 and 40% for the cleavage C-terminal to Thr^{23} and Ser^{24}, respectively, were estimated. This result was in agreement with the ratio determined by Edman sequencing [14].

3.2. Detection of minor abundant disulfide-bond scrambled forms of rCRFR1-NT

The disulfide arrangement of rCRFR1-NT was characterized by comparison of the peptide maps of reduced and non-reduced rCRFR1-NT [14]. Thereby, three disulfide bonds have been assigned (Fig. 1). In the present study, we used MICs to screen the HPLC-MS analysis of the tryptic peptide map of non-reduced rCRFR1-NT for the existence of disulfide-bond scrambled forms of the fragment rCRFR1-NT(30–57). In addition to the predominant disulfide-linked fragment rCRFR1-NT(30–57)-SS-(77–96) (Fig. 3A and B, signal 9), the fragments rCRFR1-NT(30–57)-SS-(97–127), rCRFR1-NT(30–57)-SS-(97–111), and rCRFR1-NT(30–57)-SS-(97–110) (Fig. 3C, signals 10–12) were detected with an abundance of less than 10% relative to that of rCRFR1-NT(30–57)-SS-(77–96). These three peptides represented disulfide bonds formed between one of the three Cys residues

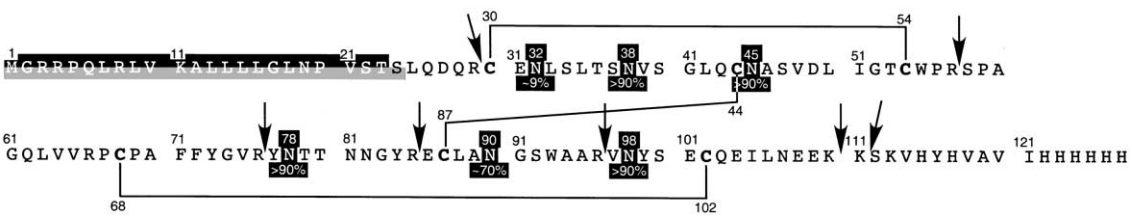


Fig. 1. Amino acid sequence of recombinant rCRFR1-NT. The two different signal peptides are marked by a black and a gray background. The connecting lines between the Cys residues indicate the disulfide bridges. The putative *N*-glycosylation sites are annotated with the position number of the Asn residue above and the extent of glycosylation below the sequence. The arrows mark the predominant tryptic cleavage sites.

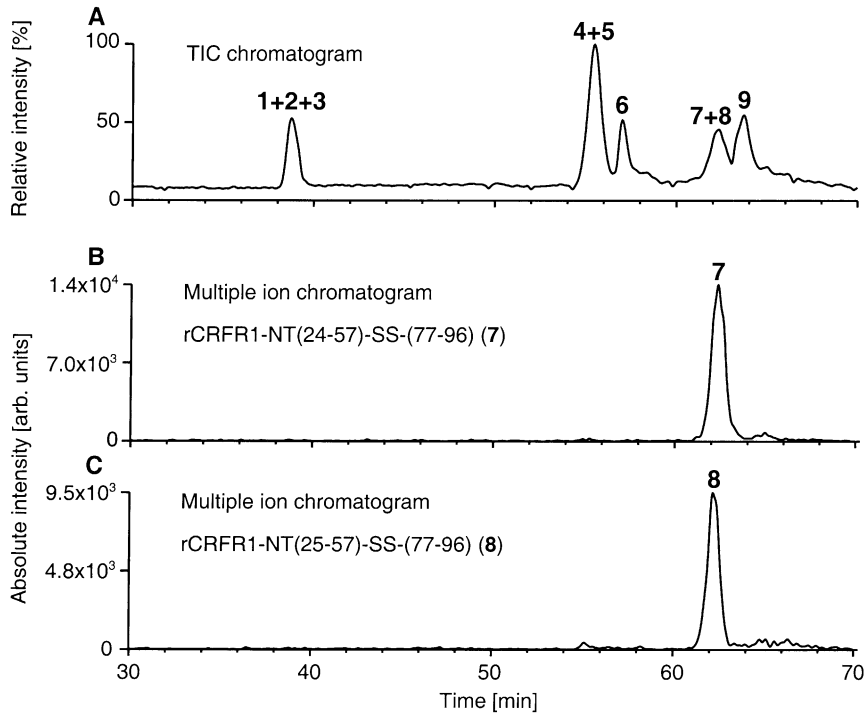


Fig. 2. TIC chromatogram (A) and MICs representing the tryptic fragments rCRFR1-NT(24–57)-SS-(77–96) (B) and rCRFR1-NT(25–57)-SS-(77–96) (C). The separation was carried out using a Vydac C18 column with 0.3 mm diameter. B: The MIC representing rCRFR1-NT(24–57)-SS-(77–96) was created using the masses of the triply (*m*/*z* 1991.5) and quadruply (*m*/*z* 1493.4) protonated molecular ions. C: The MIC representing rCRFR1-NT(25–57)-SS-(77–96) was created using the masses of the triply (*m*/*z* 1962.1) and quadruply (*m*/*z* 1471.7) protonated molecular ions. On the basis of the observed molecular masses, the following proteolytic fragments of rCRFR1-NT were assigned: 1: rCRFR1-NT(111–127); 2: rCRFR1-NT(112–127); 3: rCRFR1-NT(114–127); 4: rCRFR1-NT(58–76)-SS-(97–111); 5: rCRFR1-NT(58–76)-SS-(97–127); 6: rCRFR1-NT(58–76)-SS-(97–110); 7: rCRFR1-NT(24–57)-SS-(77–96); 8: rCRFR1-NT(25–57)-SS-(77–96); 9: rCRFR1-NT(30–57)-SS-(77–96).

of the fragment rCRFR1-NT(30–57) and Cys^{102} (Fig. 1). The presence of three different fragments may be explained by incomplete usage of the tryptic cleavage sites. Within the stretch of residues 97–127, non-reduced rCRFR1-NT was preferably cleaved C-terminal to Lys^{111} as indicated by the highest abundance observed for the fragment rCRFR1-NT(30–57)-SS-(97–111). Furthermore, the putative fragment rCRFR1-NT(58–76)-SS-(77–96) was detected in the MIC in low abundance (data not shown). In contrast, the putative fragment rCRFR1-NT(30–57)-SS-(58–76) was not detected using this approach. The remaining putative fragment rCRFR1-NT(76–96)-SS-(97–111) was observed close to the detection limit. These data suggested that only low amounts of disulfide-bond scrambled forms of rCRFR1-NT were present.

3.3. Characterization of the putative N-glycosylation sites of rCRFR1-NT

The extent of usage of the six putative *N*-glycosylation sites of rCRFR1-NT was analyzed using deglycosylation with $EndoH_f$ followed by digestion with trypsin [14]. Due to the specificity of $EndoH_f$, the *N*-acetylglucosamine (GlcNAc) residue linked to Asn remained on the protein after deglycosylation and was therefore used as a label for *N*-glycosylation.

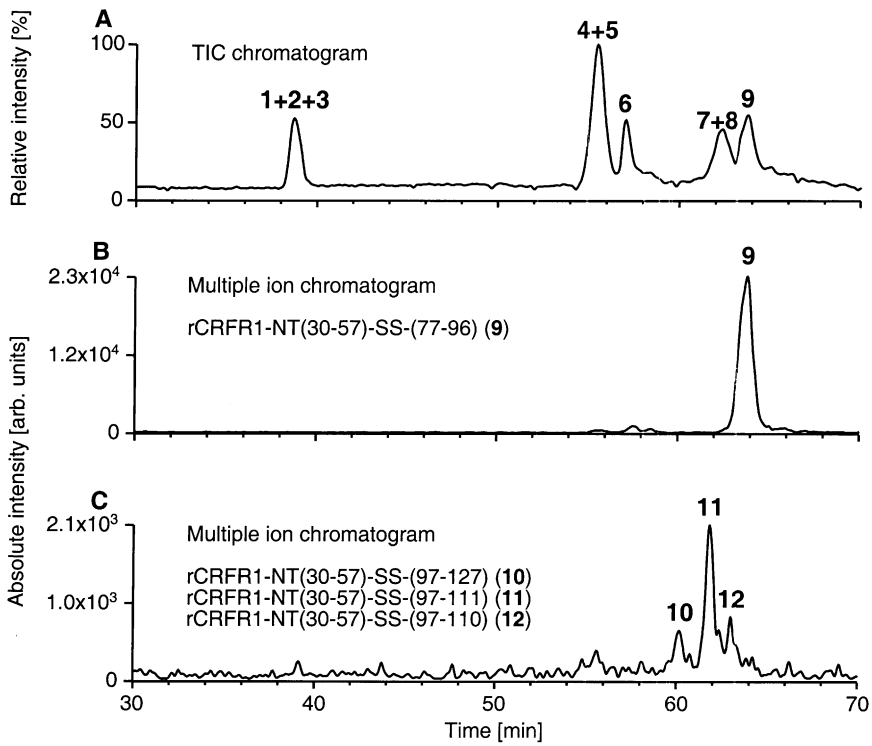


Fig. 3. TIC chromatogram (A) and MICs representing the disulfide-linked peptides containing the fragment rCRFR1-NT(30–57) (B, C). The separation was carried out on a Vydac C18 column with 0.3 mm diameter. B: The MIC representing rCRFR1-NT(30–57)-SS-(77–96) was created using the masses of the triply (*m*/*z* 1748.7) and quadruply (*m*/*z* 1311.4) protonated molecular ions. C: The MIC was created using the masses of the triply (*m*/*z* 1560.4) protonated molecular ion of rCRFR1-NT(30–57)-SS-(97–110), the triply (*m*/*z* 1603.1) and quadruply (*m*/*z* 1202.6) protonated molecular ions of rCRFR1-NT(30–57)-SS-(97–111), and the quadruply (*m*/*z* 1692.0) and quintuply (*m*/*z* 1358.8) protonated molecular ions of rCRFR1-NT(30–57)-SS-(97–127). On the basis of the observed molecular masses, the following proteolytic fragments of rCRFR1-NT were assigned: 1–9 see Fig. 2; 10: rCRFR1-NT(30–57)-SS-(97–127); 11: rCRFR1-NT(30–57)-SS-(97–111); 12: rCRFR1-NT(30–57)-SS-(97–110).

A large number of poorly resolved signals was observed in the TIC chromatogram of the tryptic digest of reduced and *S*-carboxamidomethylated rCRFR1-NT (Fig. 4A). The analysis of the extent of glycosylation was demonstrated by the characterization of the fragments rCRFR1-NT(86–96) and rCRFR1-NT(97–111). Both fragments contained a consensus sequence Asn-Xxx-Ser for *N*-glycosylation and were therefore expected to appear at least in part with one GlcNAc residue. By MICs, glycosylated rCRFR1-NT(86–96) and non-glycosylated rCRFR1-NT(86–96) were detected with absolute intensities of 3.7×10^3 and 1.6×10^3, respectively (Fig. 4B). Glycosylated rCRFR1-NT(97–111) and non-glycosylated rCRFR1-NT(97–111) were detected with absolute intensities of 1.5×10^4 and 5.6×10^2, respectively (Fig. 4C). These data indicated approximately 70 and 96% glycosylation of Asn^{90} and Asn^{98}, respectively. Even the low abundant non-glycosylated fragment rCRFR1-NT(97–111) was detected with a signal to noise ratio of approximately 10:1.

3.4. Characterization of the photoadduct formed by Diaz-oCRF and brCRFR1-NT

brCRFR1-NT (Fig. 5A) and Diaz-oCRF were irradiated with UV light to generate the photoadduct [brCRFR1-NT] × [Diaz-oCRF]. The proteins were

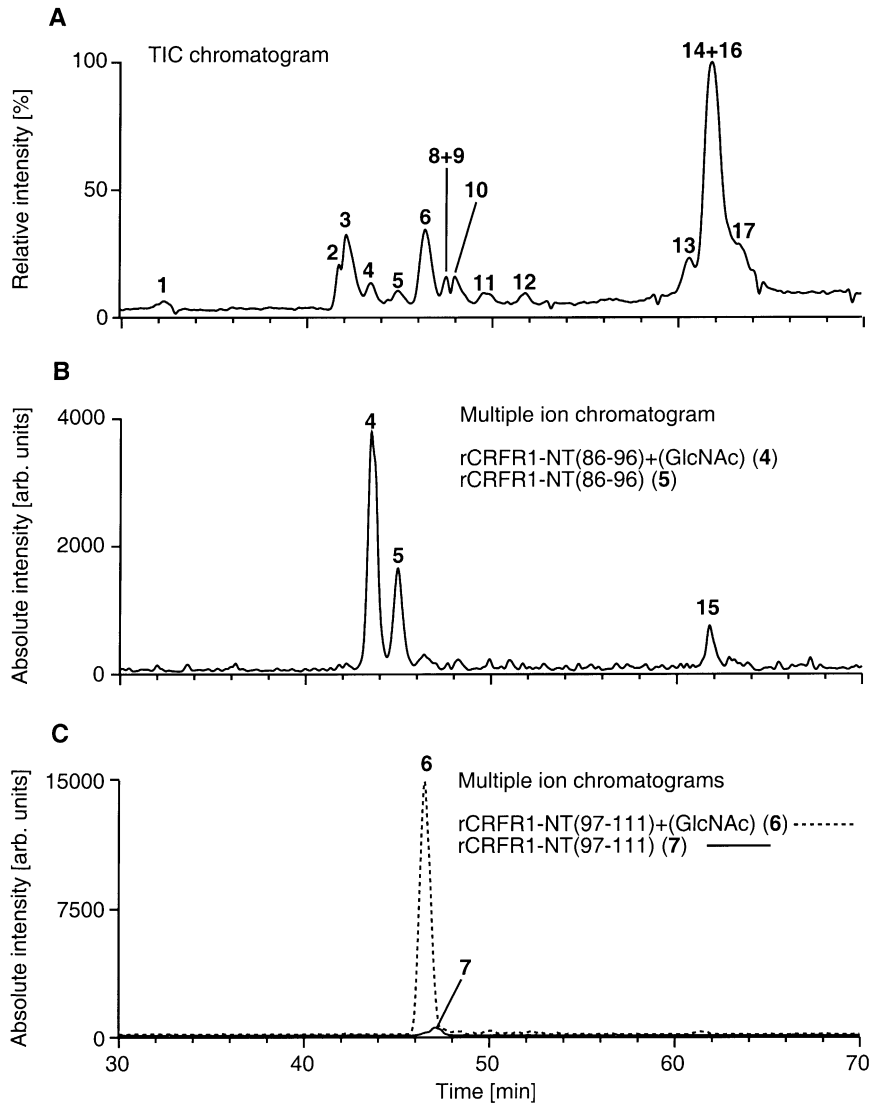


Fig. 4. Total ion current chromatogram (A) and MICs representing the differently glycosylated fragments rCRFR1-NT(86–96) (B) and rCRFR1-NT(97–111) (C). The separation was carried out on a Vydac C18 column with 0.3 mm diameter. B: The MIC representing the fragments rCRFR1-NT(86–96) and rCRFR1-NT(86–96) + (GlcNAc) was created using the masses of the molecular ions of rCRFR1-NT(86–96) (singly (*m*/*z* 1235.4) and doubly (*m*/*z* 618.2) protonated) and rCRFR1-NT(86–96) + (GlcNAc) (singly (*m*/*z* 1438.5) and doubly (*m*/*z* 719.8) protonated). C: The MIC representing the fragment rCRFR1-NT(97–111) carrying one GlcNAc residue was created using the masses of the doubly (*m*/*z* 1044.0) and triply protonated (*m*/*z* 696.3) molecular ions. The SIC representing the fragment rCRFR1-NT(97–111) carrying no GlcNAc residue was created using the mass of the doubly protonated molecular ion (*m*/*z* 943.0). On the basis of the observed molecular masses, the following proteolytic fragments of rCRFR1-NT were assigned: 1: rCRFR1-NT(77–85) + (GlcNAc); 2: rCRFR1-NT(114–127); 3: rCRFR1-NT(112–127); 4: rCRFR1-NT(86–96) + (GlcNAc); 5: rCRFR1-NT(86–96); 6: rCRFR1-NT(97–111) + (GlcNAc); 7: rCRFR1-NT(97–111); 8: rCRFR1-NT(97–110); 9: rCRFR1-NT(97–110) + (GlcNAc); 10: rCRFR1-NT(77–96); 11: rCRFR1-NT(77 − −96) + (GlcNAc); 12: rCRFR1-NT(97–127); 13: rCRFR1-NT(58–85); 14: rCRFR1-NT(30 − −57) + $(GlcNAc)_2$; 15: mass overlap of *m*/*z* 1235.4 of an unidentified minor component co-eluting with the fragments of signal 14; 16: rCRFR1-NT(58–76); 17: rCRFR1-NT(30–57) + (GlcNAc).

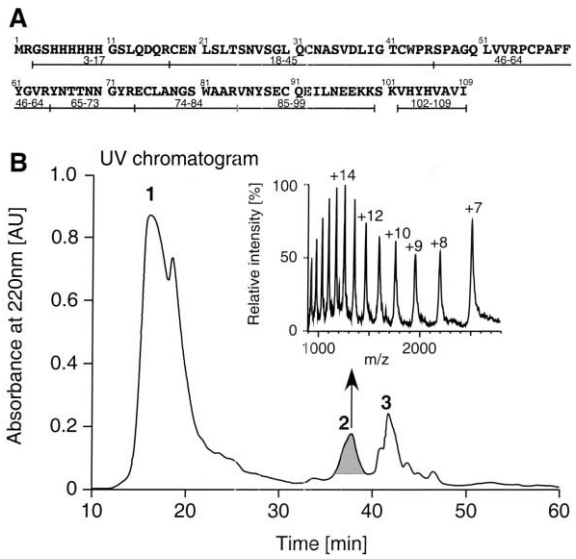
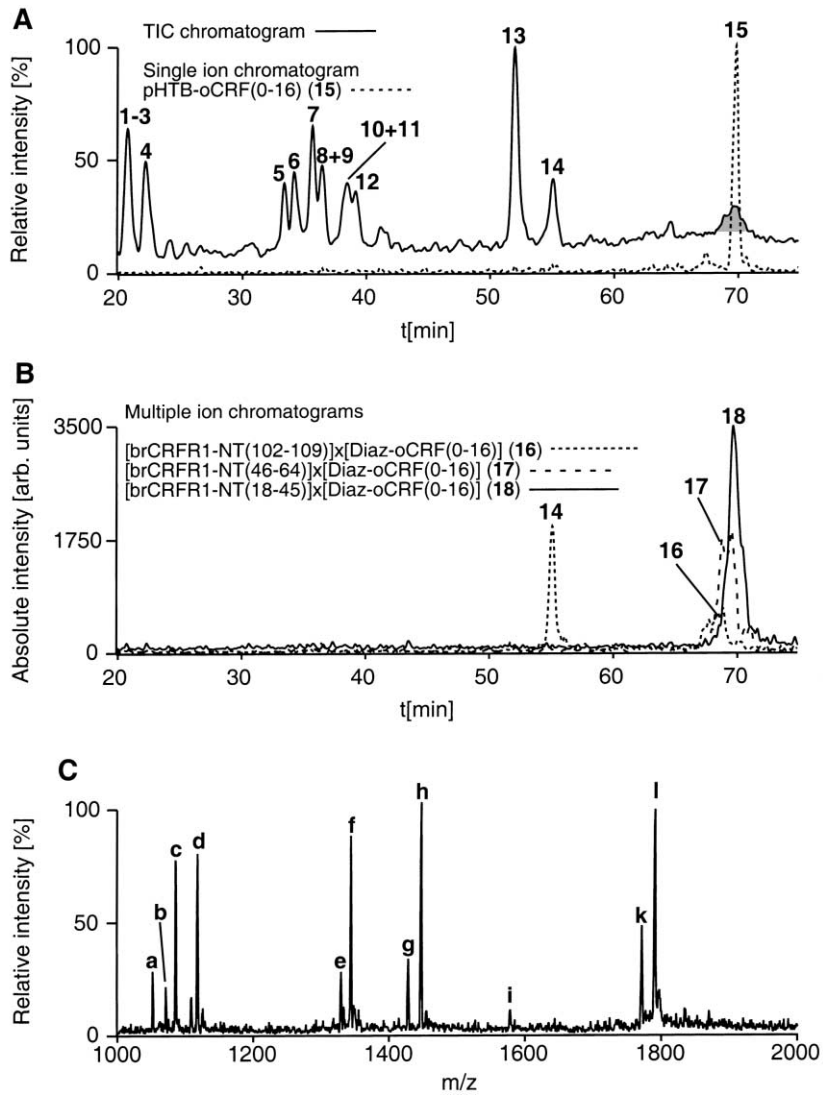


Fig. 5. Amino acid sequence of brCRFR1-NT (A) and UV chromatogram of the HPLC purification of the photoadduct [brCRFR1-NT] × [Diaz-oCRF] (B). A: The line and the range numbers below the sequence indicate the predominant tryptic fragments. B: The separation was carried out on a Vydac C4 column with 1 mm diameter. Signal 2 representing the photoadduct is shaded in gray. Signal 1 and 3 represented brCRFR1-NT and pHTB-oCRF, respectively. The HPLC-fraction containing the photoadduct was analyzed by HPLC-MS using a Vydac C4 column with 0.3 mm diameter. The ES mass spectrum obtained is shown in the inset. The observed molecular mass of the photoadduct was $M_{obs.} = 17{,}592$ Da ($M_{calc.} = 17{,}593$ Da).

reduced and *S*-carboxamidomethylated before separation of the photoreaction mixture by RP-HPLC. The compound eluting at a retention time of 38 min (Fig. 5B, signal 2) was collected and 5 μl of the isolated fraction were injected to the HPLC-MS system. A single component was eluting that was identified on the basis of its ES mass spectrum (Fig. 5B, inset) as the photoadduct of one molecule of brCRFR1-NT linked to one molecule of Diaz-oCRF. No unreacted brCRFR1-NT or Diaz-oCRF could be detected in this fraction. The major signal (Fig. 5B, signal 1) consisted of brCRFR1-NT. A third component (Fig. 5B, signal 3) was identified as photolyzed Diaz-oCRF formed by loss of N_2 and addition of H_2O (4-(1-hydroxy-2,2,2-trifluoroethyl)benzoyl-[Tyr^0]oCRF; pHTB-oCRF).

For peptide mapping of the photoadduct, HPLC purified [brCRFR1-NT] × [Diaz-oCRF] was digested with trypsin. brCRFR1-NT was digested equally and used as a control. The TIC chromatogram of the photoadduct showed the combined patterns of the tryptic fragments of brCRFR1-NT and Diaz-oCRF (Fig. 6A). Interestingly, none of the signals representing the tryptic fragments of brCRFR1-NT was absent in the TIC chromatogram of the peptide map of [brCRFR1-NT] × [Diaz-oCRF]. A new broad signal

Fig. 6. Tryptic peptide map of [brCRFR1-NT] × [Diaz-oCRF]. The TIC chromatogram (A), the SICs and MICs (A, B), and the ES mass spectrum (C) corresponding the gray shaded signal of the TIC chromatogram are shown. The separation was carried out on a Vydac C18 column with 0.3 mm diameter. A: The SIC representing pHTB-oCRF(0–16) was created using the mass of the doubly protonated molecular ion (*m*/*z* 1116.7). On the basis of the observed molecular masses, the following proteolytic fragments were assigned: 1: brCRFR1-NT(65–73); 2: Diaz-oCRF(24–35); 3: Diaz-oCRF(24–36); 4: brCRFR1-NT(3–17); 5: brCRFR1-NT(74–84); 6: Diaz-oCRF(17–23); 7: brCRFR1-NT(85–99); 8: brCRFR1-NT(85–98); 9: Diaz-oCRF(37–41); 10: brCRFR1-NT(100–109); 11: Diaz-oCRF(36–41); 12: brCRFR1-NT(102–109); 13: brCRFR1-NT(46–64); 14: brCRFR1-NT(18–45); 15: pHB-oCRF(0–16). B: The MIC representing the covalently-linked tryptic fragment [brCRFR1-NT(18–45)] × [Diaz-oCRF(0–16)] was created using the masses of the triply (*m*/*z* 1789.7) and quadruply (*m*/*z* 1342.4) protonated molecular ions. The MIC representing the covalently-linked tryptic fragment [brCRFR1-NT(46–64)] × [Diaz-oCRF(0–16)] was created using the masses of the triply (*m*/*z* 1446.2) and quadruply (*m*/*z* 1084.9) protonated molecular ions. The SIC representing the covalently-linked tryptic fragment [brCRFR1-NT(102–109)] × [Diaz-oCRF(0–16)] was created using the triply protonated molecular ion (*m*/*z* 1050.9). On the basis of the observed molecular masses, the fragments were assigned to the following signals: 16: [brCRFR1-NT(102–109)] × [Diaz-oCRF(0–16)]; 17: [brCRFR1-NT(46–64)] × [Diaz-oCRF(0–16)]; 18: [brCRFR1-NT(18–45)] × [Diaz-oCRF(0–16)]. C: On the basis of the observed molecular masses, differently protonated molecular ions of several fragments were assigned to the following signals. a: $[M + 3H]^{3+}$ of [brCRFR1-NT(102–109)] × [Diaz-oCRF(0–16)] (*m*/*z* 1050.9); b: $[M + 4H]^{4+}$ of [brCRFR1-NT(46–64)] × [Diaz-oCRF(0–16)] (*m*/*z* 1070.6); c: $[M + 4H]^{4+}$ of [brCRFR1-NT(46–64)] × [Diaz-oCRF(0–16)] + (CAM) (*m*/*z* 1084.9); d: $[M + 2H]^{2+}$ of pHB-oCRF(0–16) (*m*/*z* 1116.7); e: $[M + 4H]^{4+}$ of [brCRFR1-NT(18–45)] × [Diaz-oCRF(0–16)] + $(CAM)_2$ (*m*/*z* 1328.1); f: $[M + 4H]^{4+}$ of [brCRFR1-NT(18–45)] × [Diaz-oCRF(0–16)] + $(CAM)_3$ (*m*/*z* 1342.4); g: $[M + 3H]^{3+}$ of [brCRFR1-NT(46–64)] × [Diaz-oCRF(0–16)] (*m*/*z* 1427.3); h: $[M + 3H]^{3+}$ of [brCRFR1-NT(46–64)] × [Diaz-oCRF(0–16)] + (CAM) (*m*/*z* 1446.2); i: $[M + 2H]^{2+}$ of [brCRFR1-NT(102–109)] × [Diaz-oCRF(0–16)] (*m*/*z* 1578.0); k: $[M + 3H]^{3+}$ of [brCRFR1-NT(18–45)] × [Diaz-oCRF(0–16)] + $(CAM)_2$ (*m*/*z* 1770.7); l: $[M + 3H]^{3+}$ of [brCRFR1-NT(18–45)] × [Diaz-oCRF(0–16)] + $(CAM)_3$ (*m*/*z* 1789.7).



was observed in the TIC chromatogram at a retention time of approximately 70 min (Fig. 6A, shaded in gray). The corresponding ES mass spectrum showed the molecular ion signals of the photoadduct fragments [brCRFR1-NT(18–45)] × [Diaz-oCRF(0–16)] (Fig. 6C, signals f and l), [brCRFR1-NT(46–64)] × [Diaz-oCRF(0–16)] (Fig. 6C, signals c and h), and [brCRFR1-NT(102–109)] × [Diaz-oCRF(0–16)] (Fig. 6C, signals a and i). By MICs, these peptides were detected with decreasing intensities correlating with the decreasing size of the brCRFR1-NT fragment (Fig. 6B). In addition, the photoadduct fragments [brCRFR1-NT(3–17)] × [Diaz-oCRF(0–16)], [brCRFR1-NT(74–84)] × [Diaz-oCRF(0–16)], and [brCRFR1-NT(85–99)] × [Diaz-oCRF(0–16)] were detected with intensities close to the detection limit (data not shown). No adduct [brCRFR1-NT(65–73)] × [Diaz-oCRF(0–16)] was detected. The double signals of the fragments [brCRFR1-NT(18–45)] × [Diaz-oCRF(0–16)] (Fig. 6C, signals e, f and k, l)

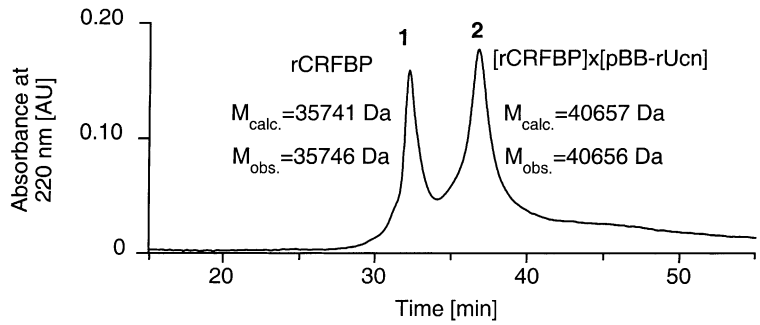


Fig. 7. UV chromatogram of the HPLC-MS analysis of the affinity-purified photoreaction mixture of pBB-rUcn and rCRFBP. The separation was carried out on a Vydac C4 column with 0.3 mm diameter. The difference of the molecular masses between rCRFBP and [rCRFBP] × [pBB-rUcn] was determined as $\Delta M_{\text{obs.}} = 4910$ Da ($\Delta M_{\text{calc.}} = 4916$ Da). The masses were determined for the most abundant glycosylated protein species.

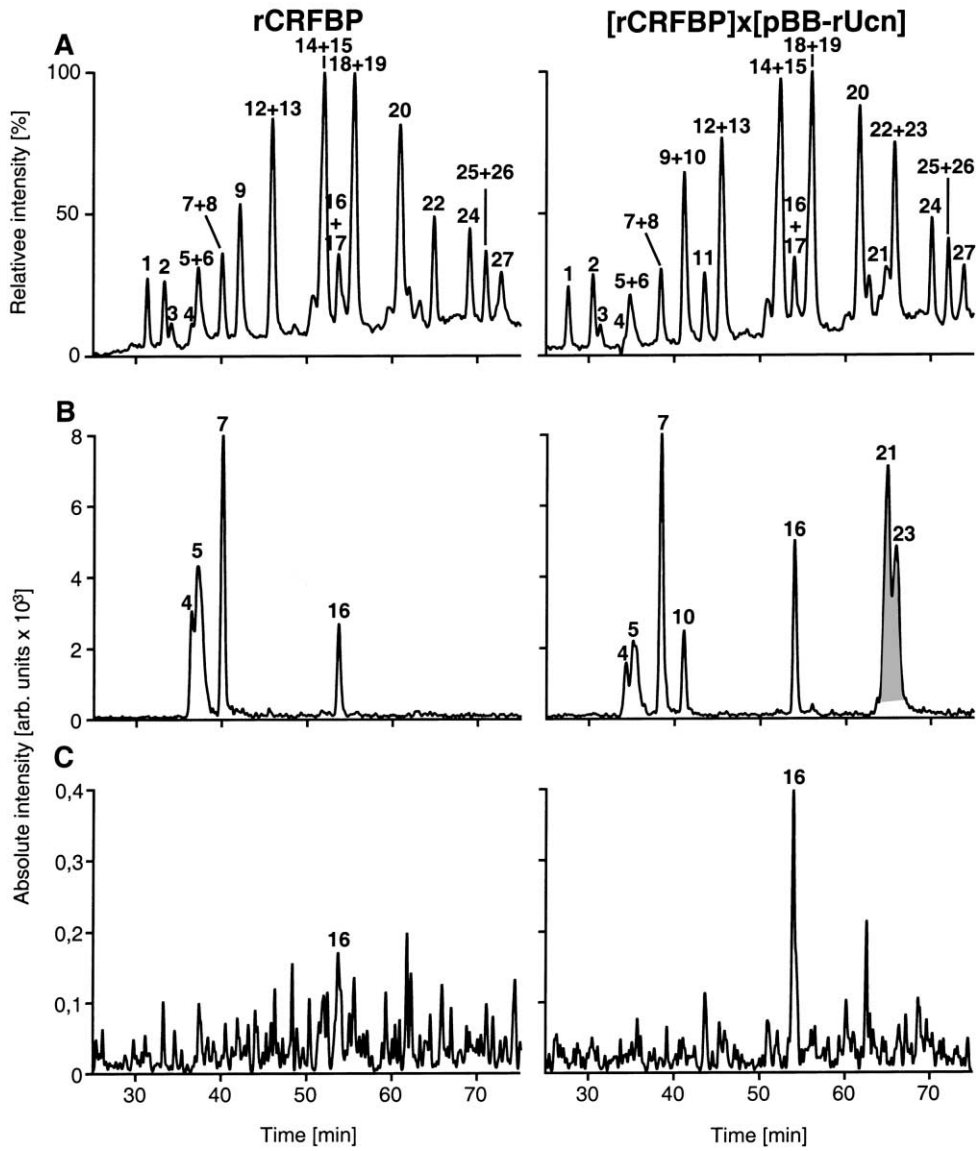
and [brCRFR1-NT(46–64)] × [Diaz-oCRF(0–16)] (Fig. 6C, signals b, c and g, h) may be explained by incomplete alkylation of the Cys residues with a *S*-carboxamidomethyl (CAM) moiety due to photoadduct formation.

In the SIC generated by using the mass of signal d (Fig. 6C), the ligand fragment pHTB-oCRF(0–16) (Fig. 6A, signal 15) was detected. Tryptic digestion of Diaz-oCRF did not yield detectable amounts of pHTB-oCRF(0–16) (data not shown). Furthermore, the incubation of [rCRFR1-NT] × [Diaz-oCRF] in digestion buffer without trypsin or with endoprotease AspN did not yield pHTB-oCRF or its fragments (data not shown). It was therefore concluded that the covalent linkage between brCRFR1-NT and Diaz-oCRF was not stable under the conditions of tryptic digestion. pHTB-oCRF(0–16) (Fig. 6A, signal 15) displayed a signal shape similar to that of brCRFR1-NT(46–64) (Fig. 6A, signal 13) indicating the presence of a homogenous peptide in contrast to the broad signals representing the fragments containing a covalent linkage of brCRFR1-NT and Diaz-oCRF (Fig. 6B, signals 16–18). This may be explained by micro-heterogeneity in the respective fragments.

3.5. Characterization of the photoadduct formed by pBB-rUcn and rCRFBP

rCRFBP and pBB-rUcn were irradiated with UV light to generate the photoadduct [rCRFBP] × [pBB-rUcn]. rCRFBP and [rCRFBP] × [pBB-rUcn]

Fig. 8. Tryptic peptide maps of rCRFBP (left column) and [rCRFBP] × [pBB-rUcn] (right column). The TIC chromatograms (A), the MICs representing the fragments of amino acids 89–124 of rCRFBP and the corresponding photoadducts with pBB-rUcn(1–15) (B), and the MICs of the fragment pBB-rUcn(1–15) (C) are shown. The separation was carried out on a Vydac C4 column with 0.3 mm diameter. B: The MICs were created using the masses of the singly protonated molecular ion of rCRFBP(89–96) (*m*/*z* 978.2), the doubly protonated molecular ions of rCRFBP(97–111) (*m*/*z* 848.9), rCRFBP(100–111) (*m*/*z* 691.8), and rCRFBP(114–124) (*m*/*z* 675.2), and the triply protonated molecular ions of [rCRFBP(97–113)] × [pBB-rUcn(1–15)] (*m*/*z* 1314.8) and [rCRFBP(100–113)] × [pBB-rUcn(1–15)] (*m*/*z* 1210.0). The signals representing the photoadduct fragments are shaded in gray. C: The MICs representing the peptide pBB-rUcn(1–15) were created using the masses of the singly (*m*/*z* 1961.2), doubly (*m*/*z* 980.1), and triply (*m*/*z* 653.7) protonated molecular ions. On the basis of the observed molecular masses, the following proteolytic fragments were assigned: 1: rCRFBP(252–260); 2: rCRFBP(24–33); 3: rCRFBP(24–32); 4: rCRFBP(100–111); 5: rCRFBP(97–111); 6: rCRFBP(271–276); 7: rCRFBP(114–124); 8: rCRFBP(114–125); 9: rCRFBP(289–305); 10: pBB-rUcn(35–40); 11: pBB-rUcn(16–22); 12: rCRFBP(140–152); 13: rCRFBP(277–288); 14: rCRFBP(153–170); 15: rCRFBP(194–208); 16: rCRFBP(89–96); 17: rCRFBP(125–139); 18: rCRFBP(126–139); 19: rCRFBP(209–233); 20: rCRFBP(171–193); 21: [rCRFBP(97–113)] × [pBB-rUcn(1–15)]; 22: rCRFBP(234–251); 23: [rCRFBP(100–113)] × [pBB-rUcn(1–15)]; 24: rCRFBP(1–23); 25: rCRFBP(2–23); 26: rCRFBP(1–22); 27: rCRFBP(2–22). Signal 10 representing pBB-rUcn(35–40) in panel B of [rCRFBP] × [pBB-rUcn] was due to a mass overlap with *m*/*z* 691.8 of rCRFBP(100–111).



were co-purified by Ni-chelate chromatography utilizing the C-terminal His-tag fused to rCRFBP. When the obtained protein mixture was analyzed by HPLC-MS, one component (Fig. 7, signal 1) was identified as rCRFBP on the basis of its observed molecular mass $M_{\text{obs.}} = 35{,}746$ Da ($M_{\text{calc.}} = 35{,}741$ Da). For the second component (Fig. 7, signal 2) a molecular mass higher than that of rCRFBP was observed ($\Delta M_{\text{obs.}} = +4910$ Da). The calculated mass difference between rCRFBP and its photoadduct was 4916 Da and thus the latter component was identified as the expected photoreaction product [rCRFBP] × [pBB-rUcn]. On the basis of UV absorption (Fig. 7) and SDS-PAGE combined with protein detection by silver staining (data not shown), a photoadduct yield of 60% was estimated.

rCRFBP and [rCRFBP] × [pBB-rUcn] were reduced with DTT and *S*-carboxamidomethylated with iodoacetamide. After separation by RP-HPLC on a Vydac

C4 column of 1 mm diameter, both proteins were digested with trypsin and the resulting peptide mixtures were analyzed by HPLC-MS. When the peptide maps of rCRFBP and [rCRFBP] × [pBB-rUcn] were compared, a decreased intensity of the signal representing the peptides rCRFBP(97–111) and rCRFBP(271–276) (Fig. 8A, signal 5 and 6) was observed in the TIC chromatogram of [rCRFBP] × [pBB-rUcn] (Fig. 8A). The corresponding mass spectrum indicated a significantly decreased abundance of rCRFBP(97–111) but not of rCRFBP(271–276) (data not shown). In addition, a decreased signal intensity was observed for the minor abundant fragment rCRFBP(100–111) (Fig. 8A, signal 4). Therefore, MICs were used to screen the HPLC-MS analysis for peptides composed of pBB-rUcn(1–15) and rCRFBP(97–111) or rCRFBP(100–111). However, no such peptides were found. By screening for tryptic fragments generated by incomplete cleavage, two new peptides matching the molecular mass of rCRFBP(97–113) and rCRFBP(100–113) each linked to pBB-rUcn(1–15) were identified (Fig. 8B). It was therefore concluded that the tryptic cleavage site at Arg^{111} was blocked due to photoadduct formation. The MICs representing the fragments rCRFBP(100–111) (Fig. 8B, signal 4) and rCRFBP(97–111) (Fig. 8B, signal 5), the two photoadduct fragments [rCRFBP(97–113)] × [pBB-rUcn(1–15)] (Fig. 8B, signal 21) and [rCRFBP (100–113)] × [pBB-rUcn(1–15)] (Fig. 8B, signal 23), and the N-terminal adjacent fragment rCRFBP(89–96) (Fig. 8B, signal 16) and the C-terminal adjacent fragment rCRFBP(114–124) (Fig. 8B, signal 7) were extracted from the HPLC-MS recordings mentioned above. A significant reduction of the abundance of rCRFBP(97–111) and rCRFBP(100–111) relative to the adjacent fragments rCRFBP(89–96) and rCRFBP(114–124) was observed in the peptide map of the photoadduct. Consistently, the signals representing [rCRFBP(97–113)] × [pBB-rUcn(1–15)] and [rCRFBP(100–113)] × [pBB-rUcn(1–15)] were detected as a result of the photoreaction. The presence of the unlabeled fragments rCRFBP(97–111) and rCRFBP(100–111) in the tryptic peptide map of [rCRFBP] × [pBB-rUcn] may be explained by an incomplete separation of [rCRFBP] × [pBB-rUcn] from rCRFBP by RP-HPLC (see Fig. 7).

Furthermore, MICs were used to screen the HPLC-MS recordings of the tryptic digests of rCRFBP and [rCRFBP] × [pBB-rUcn] for the fragment pBB-rUcn(1–15). Thereby, no signal representing a peptide of this molecular mass was detected indicating that the covalent linkage between rCRFBP and pBB-rUcn was stable under the conditions of tryptic digestion. Only a mass overlap with *m*/*z* 980.1 of rCRFBP(89–96) was observed (Fig. 8C, signal 16).

4. Discussion

MICs represent a chromatographic display simulating the selected ion monitoring (SIM) method, which is widely used for the detection of trace components in complex matrices by GC-MS such as dioxins in environmental samples [23]. In contrast to SIM, which is a distinct type of mass spectrometric recording for the detection of selected masses, MICs are generated from full scan HPLC-MS recordings. MICs combine the flexibility to choose the mass selection criteria after running the analysis with increased sensitivity. This was demonstrated by the characterization of the signal peptide cleavage sites and the minor abundant disulfide-bond scrambled forms of rCRFR1-NT on the basis of a single HPLC-MS recording. Compared with MICs, SIM is characterized by a higher sensitivity and more reliable quantitation capabilities, but requires preselection of the masses to be detected. Therefore, multiple recordings of the same sample are required when the mass selection parameters have to be changed.

As demonstrated in the present study, the use of MICs significantly decreased the detection limits for minor abundant components such as the fragment rCRFR1-NT(30–57)-SS-(97–110) (*m*/*z* 1560.4, Fig. 3C). This fragment was not detectable in the TIC chromatogram (Fig. 3A). From our experience with MICs we estimated an improvement of the sensitivity by approximately one order of magnitude. A similar enhancement of the sensitivity was described

when the signal suppression effect of the ion pairing agent TFA was prevented by post-column addition of propionic acid [24] or replacement of TFA by formic acid [Jahn et al., unpublished data]. However, MICs will provide the same enhancement in sensitivity to HPLC separations carried out under TFA-free conditions. The advantage of MICs is the additional enhancement of the selectivity which enables the detection of multiple components in a single HPLC peak. This enhancement was especially demonstrated by the characterization of the glycosylation sites of rCRFR1-NT (Fig. 4). The non-glycosylated fragment rCRFR1-NT(97–111) (Fig. 4C) was only detectable by the selective detection in the MICs. The sensitivity of this method was in addition demonstrated by the detection of this low abundant component (∼4%) with a signal to noise ratio of 10:1.

Semiquantitative estimations are of great difficulty when different peptides, which for example are generated by enzymatic cleavage of proteins, are compared. The abundances in positive ion ESMS may vary markedly depending on the presence of basic amino acid side chains which are preferably protonated. For instance, tryptic peptides containing a C-terminal Arg tend to higher abundances than tryptic peptides containing a C-terminal Lys. Therefore, the comparison of the abundances as shown for the different disulfide-linked fragments of rCRFR1-NT can only be used to distinguish between minor or major components. A different situation was present when the glycosylation of rCRFR1-NT was analyzed. Identical peptide chains differing only by a GlcNAc residue were compared. This carbohydrate moiety is not protonated or deprotonated to a significant extent under the conditions of ESMS. Therefore, a reliable semiquantitative estimation could be achieved. An alternative method for deglycosylation employing PNGaseF, which converts *N*-glycosylated Asn into Asp, may be less suited because an Asp residue is negatively charged in contrast to the neutral Asn residue. The fragments rCRFR1-NT(24–57)-SS-(77–96) and rCRFR1-NT(25–57)-SS-(77–96) were used to estimate the degree of cleavage on the respective processing site. Ser in position 24 adds a neutral amino acid side chain to a 53 amino acid peptide. Thereby, it is very likely that the protonation efficiency of both peptides is almost identical and the semiquantitative estimation is reliable as demonstrated by comparison with the results from Edman degradation. The characterization of N-terminal processing was carried out by using the disulfide-linked fragments rCRFR1-NT(24–57)-SS-(77–96) and rCRFR1-NT(25–57)-SS-(77–96), because the hydrophilic fragments rCRFR1-NT(24–29) and rCRFR1-NT(25–29) were not detected under the chromatographic conditions of the HPLC-MS analysis. The overlapping tryptic fragments rCRFR1-NT(24–57)-SS-(77–96) and rCRFR1-NT(25–57)-SS-(77–96) were only found in the tryptic digest of non-reduced rCRFR1-NT indicating partial inhibition of the tryptic cleavage site N-terminal to Cys^{30} probably due to its involvement in a disulfide-bond.

Heterogeneity introduced into peptides by modifications is difficult to resolve by HPLC. Therefore, a method named selected ion extraction was introduced [11] for the characterization of the different carbohydrate forms of tissue plasminogen activator [11,12,25,26]. This principle was also used to selectively detect low fmol amounts of co-eluting peptides in artificial peptide mixtures [27] and phosphopeptides in enzymatic digests of phosphorylated proteins [28]. In the present study, we extended the selected ion extraction to MICs. Thereby, the detection of low abundant peptides in protein digests and the semiquantitative estimation of protein modifications was facilitated. Disulfide-linked peptides representing minor abundant disulfide-bond scrambled forms of rCRFR1-NT (Fig. 3C) were displayed in a single chromatogram by using this method. Similarly, several covalently-linked fragments of [brCRFR1-NT] × [Diaz-oCRF], which co-eluted as a broad signal in the TIC chromatogram, were resolved by MICs (Fig. 6A and B). The semiquatitative estimations were demonstrated for the extent of glycosylation of the fragment rCRFR1-NT(86–96).

By using MICs it was shown that photoincorporation of Diaz-oCRF into brCRFR1-NT occurred in several tryptic fragments of brCRFR1-NT. Taking

into account the different physicochemical properties of the brCRFR1-NT fragments, the abundance of the brCRFR1-NT fragments did not provide a reliable estimate on the relative yield of photoincorporation in each fragment. However, the MIC data clearly demonstrated photoincorporation at multiple sites of brCRFR1-NT indicating the lack of regio-specificity for this reaction. This finding may be explained by a non-specific interaction between Diaz-oCRF and brCRFR1-NT. In contrast, regio-specific labeling in the stretch of residues 100–113 of rCRFBP was found for the photoadduct formed by pBB-rUcn and rCRFBP.

A minor abundant disulfide-bond scrambled form of rCRFR1-NT was detected with MICs. From incompletely resolved minor signals in the HPLC chromatogram of deglycosylated rCRFR1-NT (data not shown) it was concluded, that a significant extent of disulfide-bond scrambling appeared before tryptic digestion. This disulfide-bond scrambling was probably not due to the deglycosylation conditions since a buffer of acidic pH was used for PNGaseF cleavage. It could not be ruled out that the slightly alkaline conditions (pH 8.0) of the tryptic digestion resulted in additional disulfide-bond scrambling. However, the additional scrambling was not significant and did not interfere with the successful assignment of the disulfide bonds in rCRFBP [15] and rCRFR1-NT [14,29].

A significant amount of hydrolyzed Diaz-oCRF fragment pHTB-oCRF(0–16) was detected in the tryptic digest of the photoadduct formed by brCRFR1-NT and Diaz-oCRF. It was assumed that the linkage formed by the photoreaction was unstable under the conditions of tryptic digestion. It is most likely that the carbene generated by irradiation of the diazirine moiety is preferably inserting into the bond of an acidic hydrogen atom [30] as present in the side chain of Asp, Asn, Gln, Glu, and Tyr and in the backbone. This linkage may be hydrolyzed during tryptic digestion resulting in the release of pHTB-oCRF(0–16). In contrast, the covalent linkage formed by the benzophenone photoprobe [31] was not hydrolyzed during tryptic digestion.

Acknowledgements

Lars van Werven, Andrea Flaccus, and Thomas Liepold are gratefully acknowledged for their excellent technical help. We thank Dr. Sabine Sydow and Dr. Andreas Rühmann for helpful discussions.

References

[1] A. Varki, Glycobiology 3 (1993) 97.
[2] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. Fitz-Hugh, et al., Nature 409 (2001) 860.
[3] J.C. Venter, M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, et al., Science 291 (2001) 1304.
[4] A. Pandey, M. Mann, Nature 405 (2000) 837.
[5] M.M. Young, N. Tang, J.C. Hempel, C.M. Oshiro, E.W. Taylor, I.D. Kuntz, B.W. Gibson, G. Dollinger, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 5802.
[6] G. Dormán, G.D. Prestwich, Trends Biotechnol. 18 (2000) 64.
[7] J. Rivier, J. Spiess, M. Perrin, W. Vale, in: G. Hawk (Ed.), Biological/Biomedical Application of Liquid Chromatography II, Marcel Dekker, New York, 1979, p. 223.
[8] E. Gelpi, J. Chromatogr. A 703 (1995) 59.
[9] M. Mann, R.C. Hendrickson, A. Pandey, Ann. Rev. Biochem. 70 (2001) 437.
[10] M. Mann, A. Pandey, Trends Biochem. Sci. 26 (2001) 54.
[11] A.W. Guzzetta, L.J. Basa, W.S. Hancock, B.A. Keyt, W.F. Bennett, Anal. Chem. 65 (1993) 2953.
[12] W.S. Hancock, A. Apffel, J. Chakel, C. Souders, T. M'Timkulu, E. Pungor, A.W. Guzzetta, Meth. Enzymol. 271 (1996) 403.
[13] K. Eckart, J. Radulovic, M. Radulovic, O. Jahn, T. Blank, O. Stiedl, J. Spiess, Curr. Med. Chem. 6 (1999) 1035.
[14] B.A. Hofmann, S. Sydow, O. Jahn, L. van Werven, T. Liepold, K. Eckart, J. Spiess, Protein Sci. 10 (2001) 2050.
[15] O. Jahn, K. Eckart, S. Sydow, B.A. Hofmann, J. Spiess, Peptides 22 (2001) 47.
[16] J. Spiess, J. Rivier, C. Rivier, W. Vale, Proc. Natl. Acad. Sci. U.S.A. 78 (1981) 6517.
[17] W. Vale, J. Spiess, C. Rivier, J. Rivier, Science 213 (1981) 1394.
[18] A. Rühmann, A.K.E. Köpke, F.M. Dautzenberg, J. Spiess, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 10609.
[19] A.D. Elbein, J.E. Tropea, M. Mitchell, G.P. Kaushal, J. Biol. Chem. 265 (1990) 15599.
[20] F. Maley, R.B. Trimble, A.L. Tarentino, T.H. Plummer Jr., Anal. Biochem. 180 (1989) 195.
[21] S. Sydow, J. Radulovic, F.M. Dautzenberg, J. Spiess, Mol. Brain Res. 52 (1997) 182.
[22] A.G. Ferrige, M.J. Seddon, B.N. Green, S.A. Jarvis, J. Skilling, Rapid Commun. Mass Spectrom. 6 (1992) 707.

[23] J.B. Plomley, M. Lausevic, R.E. March, Mass Spectrom. Rev. 19 (2000) 305.
[24] A. Apffel, S. Fischer, G. Goldberg, P.C. Goodley, F.E. Kuhlmann, J. Chromatogr. A 712 (1995) 177.
[25] A. Apffel, J. Chakel, S. Udiavar, W.S. Hancock, C. Souders, E. Pungor, J. Chromatogr. A 717 (1995) 41.
[26] A. Apffel, J.A. Chakel, W.S. Hancock, C. Souders, T. Mtimkulu, E. Pungor, J. Chromatogr. A 732 (1996) 27.
[27] T.L. Quenzer, M.R. Emmett, C.L. Hendrickson, P.H. Kelly, A.G. Marshall, Anal. Chem. 73 (2001) 1721.
[28] R.S. Annan, M.J. Huddleston, R. Verma, R.J. Deshaies, S.A. Carr, Anal. Chem. 73 (2001) 393.
[29] M.H. Perrin, W.H. Fischer, K.S. Kunitake, A.G. Craig, S.C. Koerber, L.A. Cervini, J.E. Rivier, J.C. Groppe, J. Greenwald, S.M. Nielsen, et al., J. Biol. Chem. 276 (2001) 31528.
[30] J. Brunner, Ann. Rev. Biochem. 62 (1993) 483.
[31] G. Dormán, G.D. Prestwich, Biochemistry 33 (1994) 5661.