ELSEVIER

Journal of Chromatography B, 752 (2001) 247–261

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Characterization of a recombinant monoclonal antibody by mass spectrometry combined with liquid chromatography

Lobvi E. Matamoros Fernández[a,1], Dário E. Kalume[b], Loany Calvo[a], Minerva Fernández Mallo[a,1], Antonio Vallin[a], Peter Roepstorff[b,*]

[a] *Center of Molecular Immunology, P.O. Box 16040, Havana 11600, Cuba*

[b] *Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Campusvej 55, DK-5230 Odense M., Denmark*

Abstract

In this report, we present the characterization of a humanized monoclonal antibody specific for the human epidermal growth factor receptor (hEGFR). Direct analysis by matrix assisted laser desorption ionization mass spectrometry (MALDI-MS) of peptide mixtures and chromatographically isolated fractions allowed identification of 94.0% and 85.4% of the amino acid sequence of light and heavy chains, respectively. Microheterogeneity sources were identified in light and heavy chains and a previously unreported posttranslational modification for immunoglobulins was found. One N-glycosylation site was identified in the heavy chain with non-sialylated bianntenary fucosylated structures. This study is one of the first to assess the potential of MALDI-MS in combination with more conventional protein chemistry techniques for the characterization of monoclonal antibodies.

Keywords: Recombinant monoclonal antibody

1. Introduction

Overexpression of human epidermal growth factor receptor (hEGFR) in some human malignancies, such as gliomas [1] breast [2,3], bladder [4], colon [5] and lung [6,7] tumors, suggests that this receptor is important for the growth or regulation of these tumors. This has led to the development of diagnostic and therapeutic strategies, which use hEGFR as the target molecule. As part of this new experimental approach, monoclonal antibodies (MoAbs) have been successfully employed during recent years in radioimmunodiagnosis and radioimmunotherapy of these tumors. Thus, a murine IgG_{2a} antibody that binds the extracellular domain of the hEGFR has been developed [8] and its properties [9] and use for diagnosis and therapy of tumors of epithelial origin [10–13] has been reported earlier.

To reduce the response of the human immune system against murine antibodies we have recently developed a humanized version of this antibody [14]. The humanized antibody is an IgG_1, which was obtained by fusion of the complementarity determining regions (CDRs) of the murine antibody with a human framework region. As with all recombinant

*Corresponding author. Tel.: +45-6550-2404; fax: +45-6593-2661.

E-mail address: roe@pr-group.sdu.dk (P. Roepstorff).

[1] Present address: Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Campusvej 55, DK-5230 Odense M., Denmark.

0378-4347/01/$ – see front matter
PII: S0378-4347(00)00503-X

proteins intended for use in human therapy, it is necessary to rigorously characterize the structure of this antibody prior to use.

The combined use of matrix assisted laser desorption ionization (MALDI) and liquid chromatography–electrospray ionization (LC–ESI) mass spectrometry (MS) has been previously reported for the characterization of the amino acid sequence and the glycosylation of murine [15–18] and humanized [19–21] monoclonal antibodies. Simultaneous application of these two methods has provided high sequence coverage in some cases of both heavy and light chains.

Here we report the results of the structural characterization of a humanized monoclonal antibody specific for the hEGFR. Using a strategy which takes advantage of the analytical capabilities of MALDI-MS combined with reversed-phase chromatography, it was possible to verify the amino acid sequence and characterize the site specific N-glycosylation. The sequence of modified peptides and characterization of C- and N-terminal posttranslational modifications, including one not previously reported for antibodies were determined by ESI-MS–MS.

2. Experimental

2.1. Materials

Humanized monoclonal antibody R3 (MoAb R3Hu) specific for hEGFR expressed in NSO cells was produced at the Center of Molecular Immunology (Havana, Cuba) and was used as a 1 mg ml^{-1} sterile solution in sodium phosphate buffer, pH 7.2. Tris(hydroxymethyl)aminomethane, sodium carbonate, ammonium acetate, guanidine hydrochloride and hydrochloric acid were from Merck (Darmstadt, Germany). Dithiothreitol (DTT), iodoacetamide, ethylenediaminetetraacetic acid (EDTA, tetrasodium salt), and polyvinylpyrrolidone-40 (PVP-40) were obtained from Sigma (St. Louis, MO, USA). Recombinant peptide *N*-glycosidase F (EC 3.2.218, 3.5.1.52 from *Flavobacterium meningosepticum*), β-galactosidase (EC 3.2.1.23 from bovine testes), *N*-acetyl-β-D-glucosaminidase (EC 3.2.1.30 from *Diplococus pneumoniae*) and endoproteases Lys-C (EC 3.4.21.50 from *Lysobacter enzymogenes*) and Glu-C (from *Staphylococcus aureus* V8) were purchased from Boehringer Mannheim Biochemicals (Mannheim, Germany). Exoglycosidases α-manosidase (EC 3.2.1.24 from jack bean) and α-fucosidase (EC 3.2.1.51 from bovine kidney) were from Oxford GlycoSystem (Abingdon, UK). Trypsin endoprotease (EC 3.4.21.4) was from NovoNordisk (Denmark). MALDI matrices employed were 2,5-dihydroxybenzoic acid from Aldrich (Steinheim, Germany), α-cyano-4-hydroxycinamic acid from Sigma, and sinapinic acid from Fluka (Switzerland). An ion-exchange chromatographic support AG-50W-X8 (H^+ form) was purchased from Bio-Rad Labs. (Richmond, CA, USA). Trifluoroacetic acid (TFA) of sequence grade and acetonitrile (ACN) of HPLC grade were obtained from Pierce (Rockford, IL, USA). The reversed-phase chromatographic support POROS R2/50 was from Perseptive BioSystems (Framingham, MA, USA). Milli-Q-grade water (Millipore, Bedford, MA, USA) was used for all the solutions.

2.2. Monoclonal antibody purification

Humanized monoclonal antibody R3 was purified on a fast protein liquid chromatography (FPLC) system (Amersham Pharmacia Biotech, Uppsala, Sweden) by a combination of affinity chromatography with Protein A Sepharose Fast Flow (particle size 45–165 μm), ion-exchange chromatography with DEAE-Sepharose Fast Flow (particle size 45–165 μm) and gel filtration with Sephadex G-25 (particle size 172–516 μm). A volume of cell culture supernatant containing the MoAb is diluted with the same volume (1:1, v/v) of buffer 50 m*M* sodium phosphate, pH 7.0 and loaded into the DEAE Sepharose Fast Flow column (200×16 mm). The unbound fraction is collected and applied to the Protein A column (100×10 mm) and the column is washed with three volumes of 50 m*M* sodium phosphate, pH 7.0 until a baseline reading has been obtained. The Protein A column is then eluted with buffer 100 m*M* citric acid–NaOH, pH 3.0 and the eluted fraction containing the MoAb is applied to a Sephadex G-25 column (500×50 mm) equilibrated with a sterile solution of sodium phosphate buffer, pH 7.2. The obtained monoclonal antibody solution is filtered through a 0.22-μm sterile nitrocellulose

membrane. The purification process was performed at a temperature of 20°C and is controlled at a wavelength of 280 nm. Linear flow-rates of 0.5 ml min^{-1} were used in all the chromatographic steps.

The obtained MoAb is tested for physico-chemical characteristics by high-performance liquid chromatography–gel permeation chromatography (HPLC–GPC), sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and isoelectric focusing and for biological activity by immunohistochemistry and flow cytometry. Additionally DNA content, pyrogen level and murine virus content are also checked.

2.3. Sample reduction and S-carboxyamidomethylation

Lyophilized IgG (50 μg) was dissolved in 50 μl of 25 m*M* Tris–HCl buffer, pH 8.5 with 6 *M* guanidine. After addition of DTT to a final concentration of 50 m*M*, the mixture was flushed with argon and incubated for 3 h at a temperature of 37°C. Following this step the sample was cooled to room temperature, iodoacetamide added to a final concentration of 100 m*M* and the mixture incubated for 20 min in the dark at room temperature. The reduced and carboxyamidomethylated sample was purified by reversed-phase chromatography and the isolated light and heavy chain fractions dried in a Speedvac.

2.4. Reversed-phase HPLC

The HPLC system consisted of two LKB HPLC pumps 2150 (Pharmacia, Uppsala, Sweden), a solvent degasser, a column oven, and a variable-wavelength UV detector. Operation parameters were fixed and controlled through a personal computer. A 10-μm particle size 250×4 mm Vydac C_4 HPLC column (Hesperia, CA, USA) was used for peptide and protein purification. The solvent system consisted of a solution of 0.1% (v/v) TFA–water and a solution of TFA–water–ACN (0.08:20:80, v/v).

2.5. Digestion with trypsin

A 25-μg amount of protein of the light chain and 50 μg of the heavy chain were dissolved in 50 μl 50 m*M* ammonium hydrogencarbonate buffer, pH 7.8 with 1 *M* urea. After addition of 1 μg of trypsin the mixture was incubated for 18 h at a temperature of 37°C. The reaction was stopped by addition 50 μl of 2% (v/v) TFA–water. The peptide mixture was analyzed directly by MALDI-MS followed by fractionation of the mixture by reversed-phase chromatography and analysis of the isolated fractions by MALDI-MS.

2.6. Identification of the C-terminal

For identification of the C-termini of the two chains, 5 μg of light chain was digested with 0.5 μg of endoprotease Lys-C in 25 m*M* Tris–HCl buffer, pH 8.5 with 1 m*M* EDTA–50% (v/v) $H_2^{18}O$ for 18 h at a temperature of 37°C. For the heavy chain 7 μg of protein was digested with 1 μg of endoprotease Glu-C in 50 m*M* ammonium hydrogencarbonate buffer, pH 7.8–50% (v/v) $H_2^{18}O$ for 18 h at a temperature of 37°C. Samples were frozen at −20°C until they were analyzed by MALDI-MS.

2.7. Carbohydrate analysis

Oligosaccharides were obtained by the following procedure: a solution of MoAb R3Hu, 3 μg μl^{-1}, in reducing sample buffer was analyzed by SDS–PAGE employing the Tris–glycine buffer system [22]. To each sample well 30 μg of protein was applied. Concluding this step, the light and heavy chains were electrotransfered to a nitrocellulose membrane following the procedure described by Towbin et al. [23]. Proteins were located in the membrane by staining with Ponceau Red dye.

Bands containing the heavy chain were excised, washed with Milli-Q water and incubated for 60 min at room temperature with 200 μl of a solution of 1% (w/v) PVP-40 in Milli-Q water. After this step, the bands were washed twice with Milli-Q water (500 μl each time) and treated with 0.5 U of peptide *N*-glycosidase F (PGNase F) in 25 μl of 50 m*M* Tris–acetate buffer, pH 8.0 for 18 h at a temperature of 37°C. The supernatant (released carbohydrate pool) was recovered and the membrane pieces, washed twice with 50 μl of Milli-Q water and the washing solvent added to the supernatant. The total mixture was lyophilized in a SpeedVac, resuspended in Milli-Q water and analyzed by MALDI-MS.

Carbohydrate sequencing was performed by sequential digestions with β-galactosidase, *N*-acetyl-β-D-glucosaminidase, α-manosidase and α-fucosidase of tryptic glycopeptides isolated by reversed-phase chromatography. All the enzymatic digestions were performed in 5 μl of buffer 50 m*M* ammonium acetate buffer, pH 5.0 at a temperature of 37°C for 18–24 h. After each enzymatic step the resulting mixtures were analyzed by MALDI-MS.

The deglycosylated heavy chain on the nitrocellulose membrane pieces was treated with 50 μl of a trypsin solution, 12.5 ng μl^{-1}, in 50 m*M* ammonium hydrogencarbonate buffer, pH 7.8 for 18 h at a temperature of 37°C. The supernatant from the tryptic digestion was analyzed by MALDI-MS.

2.8. MALDI-MS

Analysis of peptides and proteins by MALDI-time-of-flight (TOF) MS were performed on a Voyager-Elite mass spectrometer (Perseptive Biosystems) equipped with a 337 nm nitrogen laser. Analyses were carried out in linear (M_r>2000) and reflector (M_r<2000) modes with delayed extraction. For peptide analysis in linear and reflector modes the instrument was calibrated externally with a mixture of known peptides and a low mass gate value of 500 was selected. For intact protein analysis, bovine serum albumin (BSA) was used to calibrate the instrument. Data analyses were performed using GPMAW software (Lighthouse data, Denmark), Perseptive GRAMS software version 3.02 (Galactic Industries, Salem, NH, USA) and with PeptideMass and PeptIdent programs available from the ExPASy WWW server [24].

2.9. Sample preparation for MALDI-MS

For peptide mixtures analysis a purification step using laboratory-made microcolumns filled with POROS R2/50 reversed-phase matrix was performed following a previously described procedure [25]. Peptide analyses in reflector and linear modes were performed by the sandwich sample preparation method [26] using two different matrices, sinapinic acid (SA), 10 μg μl^{-1}, in acetone and α-cyano-4-hydroxycinnamic acid (HCCA), 10 μg μl^{-1}, in TFA–water–ACN (0.1:30:70, v/v). First 0.5 μl of SA solution was deposited on the target and allowed to spread and dry. Then 0.5 μl of 2% (v/v) TFA in water and 0.5 μl of sample solution were placed on top of the sinapinic acid layer followed by 0.5 μl of HCCA solution. Protein analysis in linear mode was also performed with the sandwich method but for the second layer of matrix, a solution of SA in TFA–water–ACN (0.1:30:70, v/v) was employed. For glycopeptide analysis, a similar micropurification was performed and the eluted peptides mixed with a saturated matrix solution of 2,5-dihydroxybenzoic acid (DHB) in 50% (v/v) methanol in water onto the target plate.

The released carbohydrate structures were desalted on microcolumns containing ion-exchange resin AG-50W-X8 (H^+ form) prior to analysis by MALDI-MS. The mixture was passed over the column and the run-through applied onto the target plate, mixed with 0.5 μl of a saturated solution of DHB in 50% (v/v) methanol in water and allowed to dry. The sample-matrix layer was re-crystallized with absolute ethanol before analysis by MS.

2.10. ESI-MS

Nano-ESI-MS and MS–MS were performed in the positive ion mode on a Micromass Q-TOF mass spectrometer (Manchester, UK) equipped with a Z-spray sample introduction system. Peptide samples were sprayed from silver-coated glass capillaries from Protana (Odense, Denmark). The potential voltage applied to the nanoflow tip in the ion source and the nitrogen backpressure ($<0.68 \cdot 10^{-3}$ mbar) were adjusted to produce a flow-rate of the sample solution in the ~25 nl min^{-1} range to allow long analysis time. A potential of 800–1200 V was applied to the nanoflow tip in the ion source. The cone voltage was set to 30–55 V and the MCP (microchannel plate detector) voltage was set to 2600–2800 V. The source temperature was 40°C and nitrogen was used as a drying gas (flow-rate 25 l h^{-1}). Acquisition of data was performed on a Mass Lynx Windows NT PC data system (software version 3.2). For all the experiments, a sodium iodide solution (~1 μg μl^{-1} in 2-propanol–1% formic acid, 1:1, v/v) was used for calibration of the TOF analyzer. For collision-induced dissociation experiments, the precursor ion was selected in the first

Table 1
Masses for the MoAb R3Hu and the light and heavy chains

Sample	Theoretical value (M_r)	Observed value (M_r)	Mass difference
Light chain	24 415.11[a]	24 423.0	+8.29
Heavy chain	50 243.30[a]	51 484.4	+1241.0
Native MoAb	148 404.82[b]	150 863.0	+2458.18
Deglycosylated MoAb	148 404.82[b]	147 591.0	−813.0

[a] Masses have been corrected for the mass increase produced by the reduction and carboxyamidomethylation procedure (+57 per cysteine residue) considering five cysteine residues in the light chain and 11 in the heavy chain.
[b] Masses are calculated with all cysteines assumed disulfide bonded.

quadrupole (Q1) and fragmented in the hexapole collision cell with collision energy varying from 20 to 40 eV. Argon was used as the collision gas at a pressure of ~$6.0 \cdot 10^{-5}$ mbar. The fragment ions were analyzed using the TOF analyzer.

2.11. Sequences of light and heavy chains

Sequences of the constant regions of the light and heavy chain were obtained from Protein Information Resource (PIR) database release 61.00 [27].

3. Results

The M_r values of the intact glycoprotein, the deglycosylated protein and of the separated light and heavy chains following reduction and *S*-carboxyamidomethylation were determined by MALDI-MS (Table 1). The M_r values obtained for the light chain are in good agreement with the theoretical values calculated from the template sequence. The M_r values obtained for the heavy chain, the native monoclonal antibody and for the deglycosylated molecule indicate the presence of high mass modi-

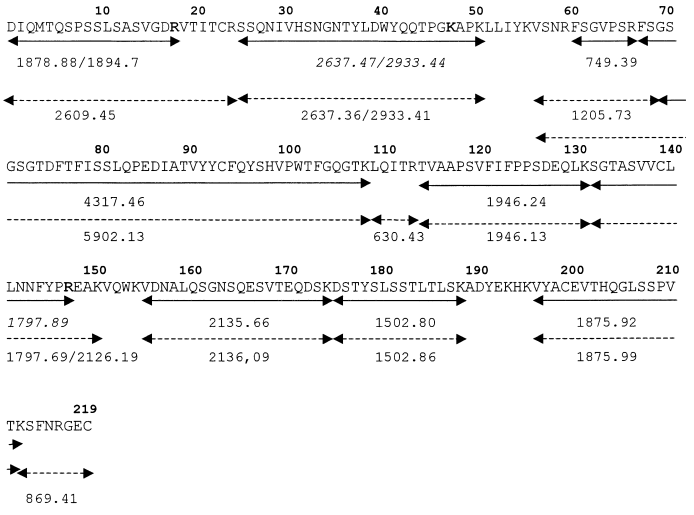

Fig. 1. Predicted amino acid sequence of the MoAb R3Hu light chain. The identified tryptic peptides are underlined with full lines. The masses for peptides identified by direct mixture analysis are given in italics and those identified after reversed-phase isolation in normal font. Peptides identified by direct mixture analysis after lysyl-endopeptidase digestion are indicated with dotted lines. The amino acid residues in bold letters denote missed or partial cleavage sites.

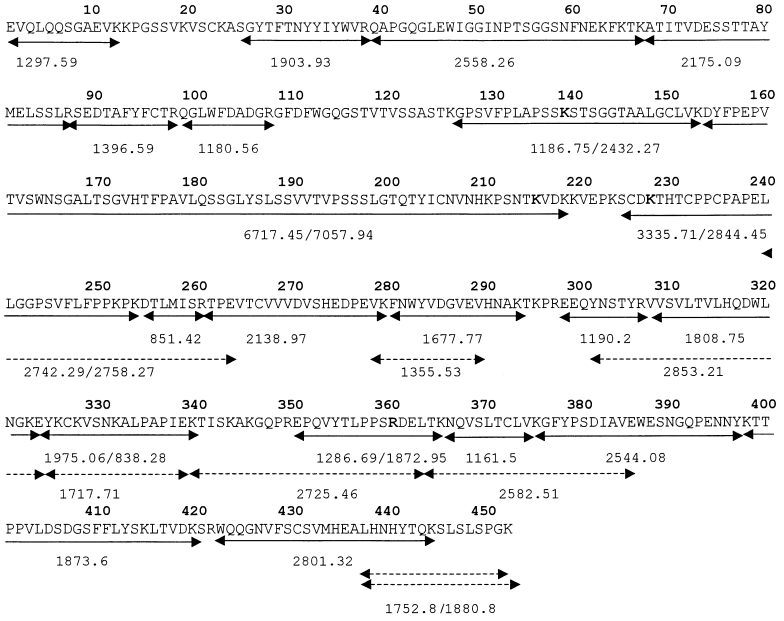

Fig. 2. Predicted amino acid sequence of the MoAb R3Hu heavy chain indicating the tryptic peptides identified by direct mixture analysis and after reversed-phase isolation (full lines), and Glu-C endopeptidase digestion with direct mixture analysis (dotted lines). The N-terminal amino acid is pyroglutamic acid. The amino acid residues in bold denote missed or partial cleavage sites.

fications in the heavy chain most likely due to the glycosylation.

Measurement of the M_r of the intact MoAb R3Hu and its component chains does not allow identification of the glycosylation site nor does it allow assignment of any additional modifications. On the basis of the predicted sequences of the light and heavy chains (Figs. 1 and 2) it was decided to perform a tryptic digestion of the light and heavy chains to generate a mixture of peptides for further

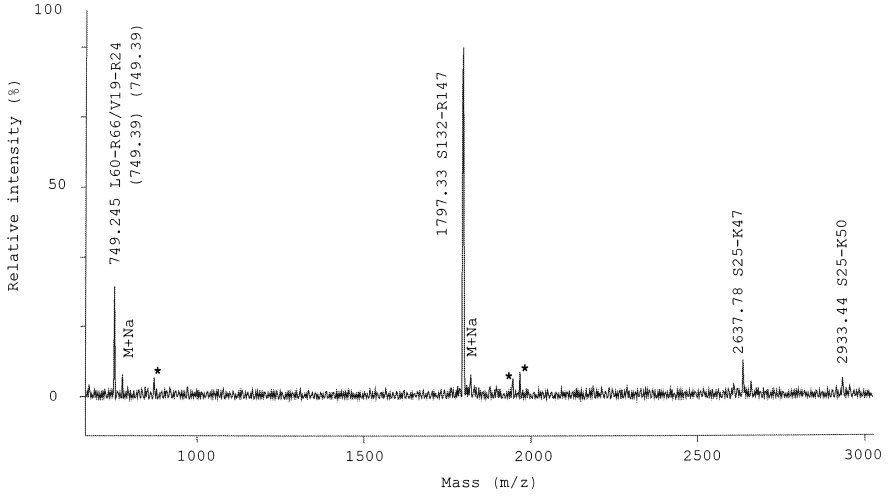

Fig. 3. MALDI spectrum of the tryptic digest of the reduced and carboxyamidomethylated light chain. Non identified signals are labeled with an asterisk.

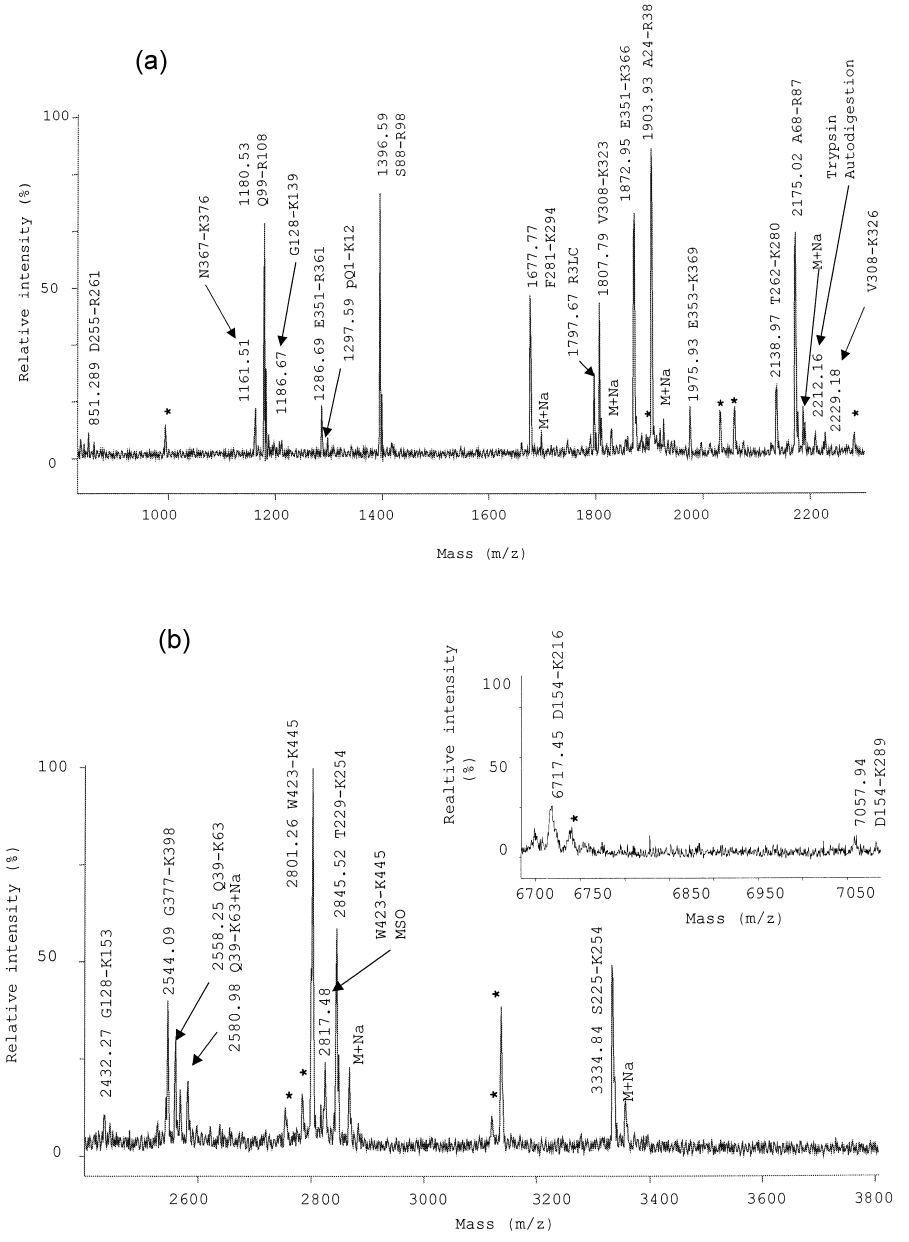

Fig. 4. MALDI spectrum of the tryptic digest of the reduced and carboxyamidomethylated heavy chain. (a) M_r range 800–2500, (b) M_r range 2300–3800/7060. R3LC: Contaminant peptides from the light chain. Non identified signals are labeled with an asterisk.

detailed analysis. The resulting peptide mixtures were analyzed directly by MALDI-MS and also submitted to reversed-phase chromatography and the isolated fractions analyzed by MALDI-MS. The MALDI spectrum in reflector mode of the tryptic peptide mixture derived from the light chain is shown in Fig. 3. The spectrum of the peptide mixture contains rather few peaks compared to the expected number. The sequence coverage is only 19.1% or 49 of the 219 amino acids of the light chain. This

sequence coverage assumes that the signal at m/z 749.24 represents just one of two peptides (theoretical m/z 749.39, L^{60}–R^{66} and V^{19}–R^{24}) that have the same monoisotopic mass.

The analysis by MALDI-MS of the tryptic peptide mixture from the heavy chain (Fig. 4) results in a sequence coverage of 77.7% or 352 of the 453 amino acids of the sequence and permitted identification of several modifications. The heavy chain N-terminal peptide (peak at m/z 1297.59) with the glutamine residue in position 1 converted to pyroglutamic acid was detected in the spectra (confirmed by MS–MS analysis, data not shown). The signal with m/z 1180.5 could not readily be assigned to any sequence and the mixture was therefore analyzed by ESI-MS–MS. The MS–MS spectrum of the corresponding doubly charged ion at m/z 590.27 showed a series of Y ions (Y_1–Y_7) which identified the peak to be derived from the peptide Q^{99}–R^{108} with an oxidized tryptophan residue resulting in a mass increment of 16.0 (Fig. 5). The signal with m/z 851.28 was identified, as derived from the peptide D^{255}–R^{261} with a methionine residue oxidized (not verified).

Analysis of the fractions obtained by reversed-phase chromatography of the tryptic peptides from the light chain (Fig. 6) results in sequence coverage of 80.3% corresponding to 176 of the 219 amino acids. In the fraction 9, containing the light chain N-terminal tryptic peptide, two signals were detected at m/z 1878.8 and m/z 1894.7. The first signal correspond to the unmodified peptide, and the other signal which differ by 16.0 from the expected m/z value, may indicate the presence of a modified form of the peptide. This form was investigated by ESI-MS–MS. The series of fragment ions obtained is similar to that predicted from the sequence except for the b_4 and Y_{15} fragment ions, which shows the presence of an oxidized methionine residue (Fig. 7). For the chromatographic fractions 1, 2 and 20 no signal was observed by MALDI-MS analysis.

The results obtained by chromatographic separation of the tryptic peptide mixture of the heavy chain by reversed-phase (RP) HPLC (Fig. 8) followed by MS analysis showed one new peptide with m/z 1873.9 (T^{399}–K^{415}) whereas several of the previously detected peptides were not found (m/z 1903.9, 2432.27, 3335.71 and 6717.4). The combination of the results from direct peptide mixture

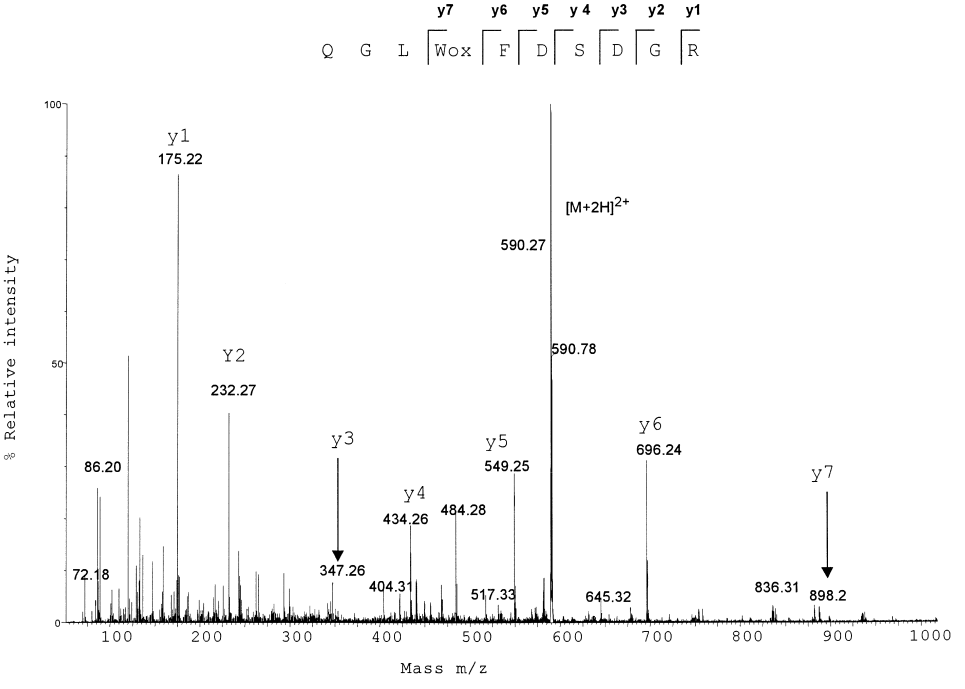

Fig. 5. ESI-MS–MS spectrum of $[M+2H]^{2+}$ for the peptide Q^{99}–R^{108} identifying an oxidized tryptophan residue.

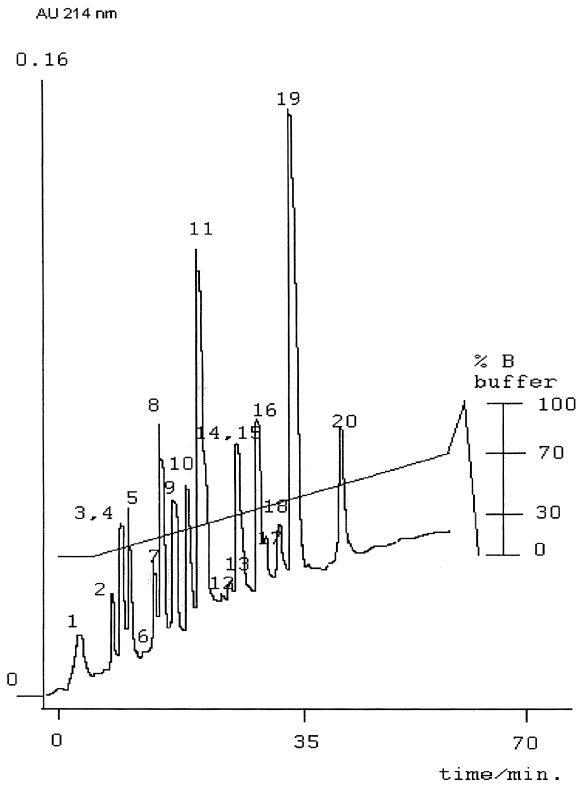

Fig. 6. Reversed-phase chromatogram of the tryptic digest of the reduced and carboxyamidomethylated light chain. Flow-rate=0.9 ml min^{-1}, wavelength=214 nm, oven temperature 45°C. Gradient: 0% B/10 min, 0–40% B/40 min, 40–60% B/20 min, 60–80% B/1 min. The identified peptides are shown in Table 2.

analysis with those obtained after reversed phase covered 369 residues out of 453 residues (81.4%) for the heavy chain.

To identify the C-terminal peptides in the two chains, digestion with specific proteases was carried out in buffer solutions containing 50% $H_2^{18}O$ resulting in mass spectra where all peptides, except the original C-terminal, will have ^{16}O and ^{18}O incorporated in a 1:1 proportion. This allows distinction between internal and C-terminal peptides based on the isotope pattern [28]. For the light chain, digestion with endoprotease Lys-C in buffer containing 50% $H_2^{18}O$ and analysis by MALDI-MS allow identification of the predicted C-terminal peptide S^{213}–C^{219} (m/z 869.36) (Fig. 9a). The peptide map generated by endoprotease Lys-C digestion resulted in 94.0% sequence coverage.

Digestion of the heavy chain with endoprotease Glu-C results in the detection of two forms of the C-terminal peptide: A^{437}–K^{453} (m/z 1880.8) and A^{437}–G^{452} (m/z 1752.79) indicating that C-terminal truncation had taken place (Fig. 9b). Several signals of low intensity could be observed, which could not be assigned to any sequence taking the expected specificity of endoprotease Glu-C in consideration. They probably represent peptides produced by unspecific cleavage. Additionally, two signals could be assigned to peptides derived from non-previously verified segments of the sequence (L^{240}–E^{264}, Q^{301}–E^{324}). Their observation increase the total sequence coverage of the heavy chain to 85.4%

MALDI-MS analysis of the carbohydrate pool produced three signals at m/z 1485.2, 1647.3, 1809.67 ([M+Na^+] species) with mass differences of 162 (Fig. 10), which can be explained in terms of different content of hexoses. Based on the typical complex type glycan structures observed in mammals possible structures could be proposed (Table 2). The glycopeptide was identified to be present in fraction 1 in the HPLC-separated tryptic digest based also on observed mass differences of 162. Upon sequential digestion with specific exoglycosidases it was possible to confirm the presence of the previously proposed structures (Fig. 11) and to assign the glycosylated peptide to E^{299}–R^{307}.

The digestion of deglycosylated heavy chain with trypsin followed by analysis by MALDI-MS identified a signal at m/z 1190.2 corresponding to the peptide E^{299}–R^{307} (m/z 1189.5) with the Asn 303 residue converted to Asp as expected by treatment with PGNase F (data not shown).

4. Discussion

Several reports dealing with the characterization of antibodies have been published over the last few years. The most complete characterization was obtained [20] by combined use of LC–ESI, MALDI-MS and Edman degradation for characterization of the amino acid sequence, and high performance ion exchange chromatography and digestion with specific exoglycosidases for characterization of the carbohydrate residues. In this report, we have used a

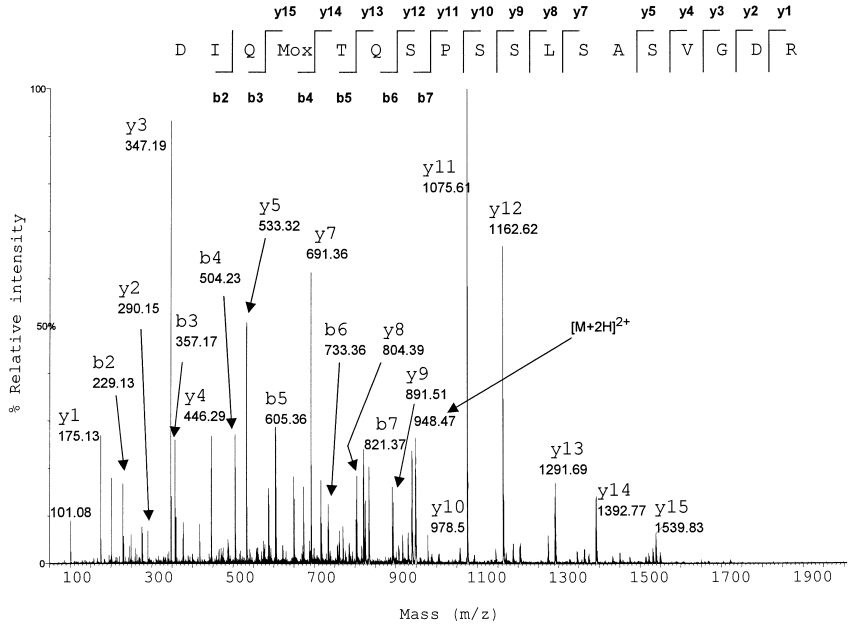

Fig. 7. ESI-MS–MS spectrum of $[M+2H]^{2+}$ for the light chain N-terminal peptide D^{1}–R^{18}. An oxidized methionine residue is identified based on b_4 and Y_{15} ions.

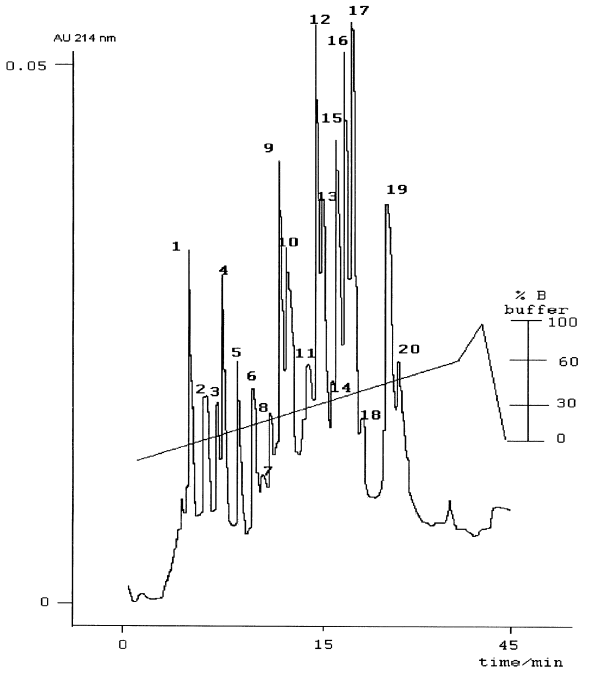

Fig. 8. Reversed-phase chromatogram of the tryptic digest of the reduced and carboxyamidomethylated heavy chain. Flow-rate=0.9 ml min^{-1}, wavelength=214 nm, oven temperature 45°C. Gradient: 0% B/10 min, 0–40% B/40 min, 40–60% B/20 min, 60–80% B/1 min. The identified peptides are shown in Table 2.

simple and faster strategy based on a mass determination of peptides derived by enzymatic digestion combined when needed with sequencing by MS–MS that has allowed us to obtain the same level of structural information.

Mass determination of the peptides derived by enzymatic digestion of the protein in itself does not confirm the sequence since the determined mass reflects the amino acid composition of a peptide independently of the sequence. For antibodies prepared by recombinant technology the assumed sequence is known and in such cases the sequence can be considered confirmed by the information derived by mass determination of the peptides produced by proteolysis with specific enzymes. Sequencing is only indicated for peptides with masses deviating from the expected value.

The main limitation for direct analysis of peptide mixtures by MALDI-MS is the ion specific suppression effect. In this study, the most striking example of suppression effect is the observation of only four peptides signals in direct analysis by MALDI-MS of the mixture derived by tryptic digestion of the light chain, while up to nine peptides could be isolated and identified by MS after reversed-phase chromatography.

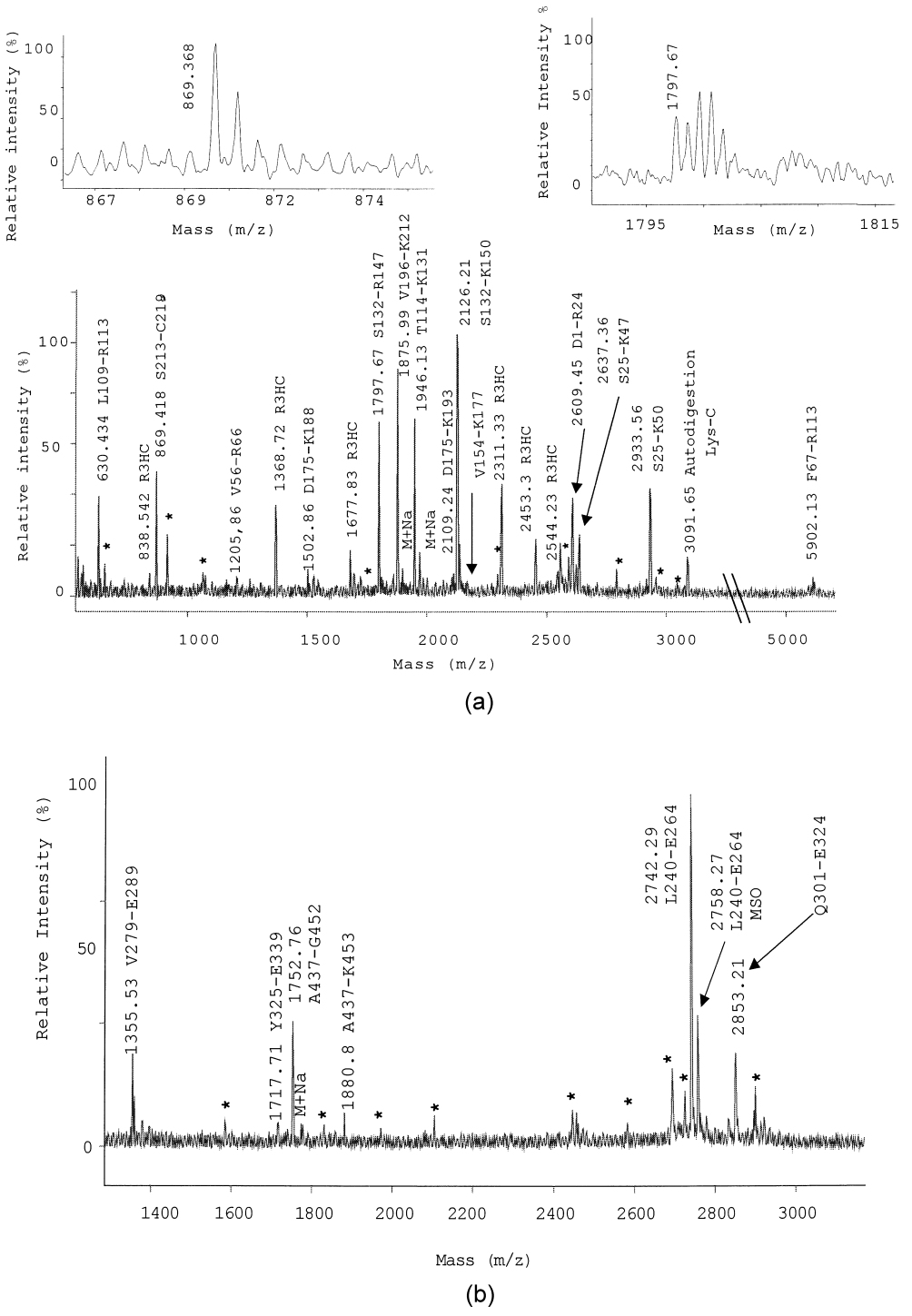

Fig. 9. (a) MALDI spectra of the lysyl endopeptidase digest of the reduced and carboxyamidomethylated light chain with buffer containing 50% (v/v) $H_2^{18}O$: insets:(top left) the C-terminal peptide, (top right) an internal peptide, (bottom) complete spectrum (R3HC are contaminant peptides from the heavy chain), and (b) the Glu-C endopeptidase digest of the reduced and carboxyamidomethylated heavy chain with buffer containing buffer containing 50% (v/v) $H_2^{18}O$. Non identified signals are labeled with an asterisk.

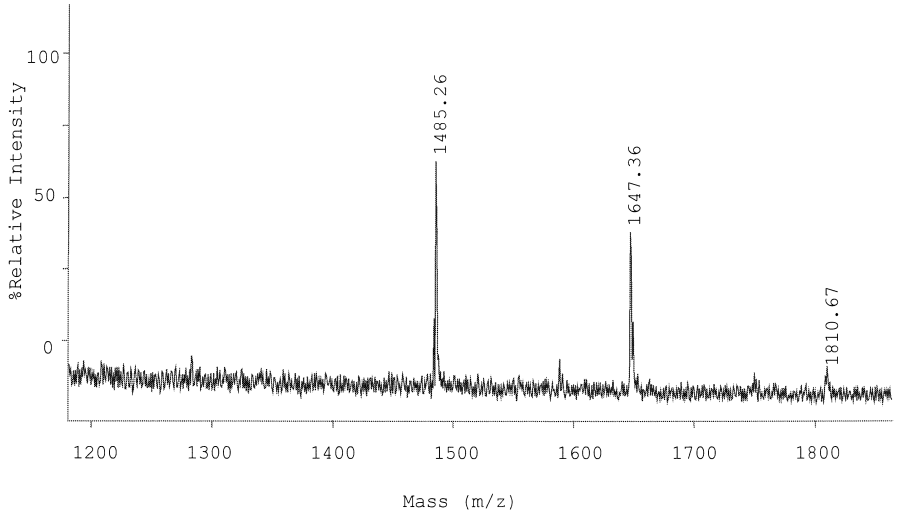

Fig. 10. MALDI spectrum of the glycans liberated by glycosidase F treatment ($[M+Na]^+$ species), see Table 3 for proposed structures.

Use of reversed-phase chromatography permitted the isolation of the modified light chain N-terminal peptide and allowed assignment of the modification by ESI-MS–MS. The obtained fragments ion spectrum confirmed the presence of an oxidized methionine residue on the basis of the molecular mass difference of 16 between the theoretical values of the b_4 (m/z 488.3) and Y_{15} (m/z 1523.8) ions and the observed masses (m/z 504.3 and m/z 1539.83). Out of the two peptides with an M_r of 749.29 we assigned the signal with m/z 749.39 detected in the RP-HPLC fractions 3 and 4 to the peptide L^{60}–R^{66} on the basis of the predicted elution time of the peptide. For the light chain RP-HPLC fractions 1, 2 and 20 no m/z signal was observed upon analysis by MALDI-MS. Most likely fractions 1 and 2 are UV-absorbing impurities from the chromatographic solvents. Fraction 20 can be a large hydrophobic peptide produced by incomplete cleavage during the tryptic digestion.

In the heavy chain an N-terminal pyroglutamic acid residue and a partial truncation of the C-terminal were observed. These observations agree with previous reports [17,19,20]. In addition, an oxidized tryptophan residue in position 102 of the heavy chain was identified. Although tryptophan oxidation have been recently reported for proteins [29–32] it have not been reported previously for antibodies. In this work we found that the most intensive peak corresponds to the single oxidized species (+16) and that the non-oxidized peptide was only present at a very low level. This contrasts previous observations where the oxidized species was observed at low level [30] and different oxidation states in all the tryptophan residues of the molecule were observed [31].

Oxidation of tryptophan is a well-known artifact generated during sample preparation for MALDI-MS [26] and the amount of oxidation can be reduced by careful sample preparation methods as used in this work. However, the fact that only Trp 102 out of 9 in the heavy chain was observed to be modified seems to indicate that the sample preparation method is not the primary cause. We therefore believe that this specific Trp residue is oxidized during the fermentation or the purification process.

Inclusion of $H_2^{18}O$ in the digestion buffer for distinction between the C-terminal peptide and internal peptides proved to be useful since without this precaution the small signal of the non-processed C-terminal peptide most likely could have been ignored.

Treatment of the MoAb R3Hu with PGNase F and subsequent analysis of the glycan pool by MALDI-MS identified three oligosaccharide structures of the bianntenary complex type. Digestion of the tryptic glycopeptide with exoglycosidases followed by MS analysis confirmed the oligosaccharides structures suggested by mass spectrometric analysis of the glycan pool. The observed non-sialylated,

Table 2
Tryptic peptides identified by HPLC and mass spectrometry

HPLC fraction	Peptide	Theoretical (M_r)	Observed (M_r)	Mass difference
Light chain				
9	D1–R18	1878.88	1878.7	−0.18
9	D1–R18	1878.88	1894.7	15.82*
11	S25–K47	2637.33	2637.4	0.14
3/4	F60–R66	749.39	749.36	−0.03
18/19	F67–K108	4716.14	4717.46	−1.32
14/15	T114–K131	1946.09	1946.24	0.15
16/17	S132–K147	1797.67	1797.09	0.59
5/6	V155–K174	2135.96	2135.66	0.3
10	D175–K188	1502.75	1502.8	0.05
8/9	V196–K212	1875.92	1876.03	0.11
Heavy chain				
5	E1–K12	1314.70	1297.59	−17.11
14	Q39–K63	2558.22	2558.26	0.04
4	A68–R77	2175.04	2175.09	−0.17
10	S88–R98	1396.48	1396.59	0.11
7/8	Q99–R108	1164.54	1180.56	16.02*
7/8	G128–K139	1186.64	1186.75	0.11
19	T229–K254	2844.45	2844.49	0.04
3/4	D255–R261	835.50	835.20	0.03
2	D255–R261	835.50	851.42	15.92
13	T262–K280	2139.02	2138.97	−0.050
12	F281–K294	1677.80	1677.77	−0.032
1	E299–R307	1189.51	2634.80	1445.20*
1	E299–R307	1189.51	2797.02	1607.50*
1	E299–R307	1189.51	2958.50	1769.00*
19/20	V308–K323	1808.00	1807.75	−0.25
15	E324–K340	1975.06	1975.93	0.87
2	A333–K340	838.50	838.20	0.3
6	E351–R361	1286.67	1286.69	0.02
9	E351–K366	1872.97	1872.96	−0.01
7	N367–K376	1161.63	1161.43	0.20
17	G377–K398	2544.13	2544.08	−0.05
17	T399–K415	1873.90	1873.50	−0.4
14	W423–K445	2801.26	2801.32	0.06

*These mass differences can all be explained by the reported modifications.

fucosylated bianntenary structures with one, two and three galactose residues are in agreement with previous observations for murines antibodies [19] and humanized monoclonal antibodies expressed in NSO cells [21]. Although expression of human proteins in murine cells like NSO is a well documented and characterized process, it have the potential risk of producing aberrant glycosylation and thereby of producing immunogenic glycoforms. Such immunogenic epitopes in the form of Galα(1–3)βGa(1–4) motif have been recently reported for human antibodies expressed in NSO cells [21]. In our work no evidence was found of the presence of any such epitopes.

The non-confirmed regions of the molecule contain several tryptic cleavage sites, which produced small peptides, typically di- to tetrapeptides. These were not detected either because their mass was below the mass gate value selected, because they were suppressed by matrix related ions or because they eluted in the non-bound fraction in the RP-HPLC analysis.

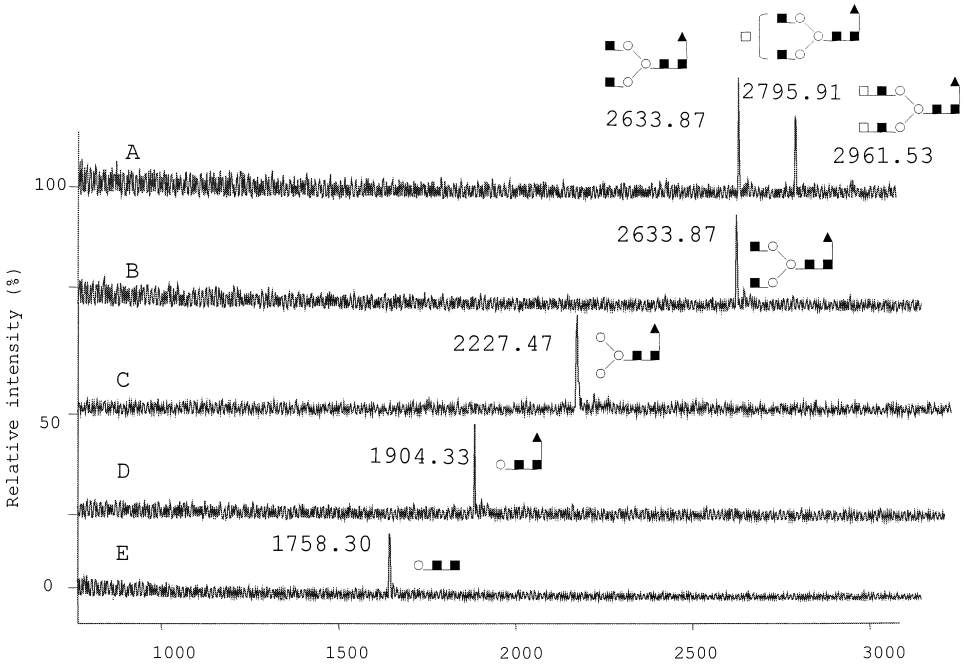

Fig. 11. MALDI-MS spectrum obtained upon sequential exoglycosidase digestions of the tryptic glycopeptides from the heavy chain. (A) Prior to treatment, (B) 1 mU β-galactosidase digest followed by, (C) 1 mU *N*-acetyl-β-D-glucosaminidase digestion, (D) 100 mU α-manosidase digestion, and (E) 10 mU α-fucosidase digestion. The suggested glycan structures are indicated. Galactose □, *N*-acetylglucosamine ■, manose ○, fucose ▲.

In the work described here we have identified N-terminal blocking, C-terminal processing and oxidation of amino acid residues as the main source of microheterogeneity in the MoAb R3Hu. Presence of additional high mass modifications in the light chain are not likely to be present since we obtain a good correlation between the theoretical and the observed M_r values and up to 94.0% coverage of the sequence. For the glycosylated heavy chain and for the intact MoAb the difference between the observed and the theoretical values of M_r are close to the mass of the main carbohydrate structure (Table 3) found in this work. This agreement is even better if the truncated C-terminal in the heavy chain is taken in account, reducing the theoretical values of the heavy chain and the intact monoclonal antibody by 128.05 and 256.1, respectively.

The mass difference of 813.0 between the theoretical value and the observed value for the intact deglycosylated monoclonal antibody in part reflect the lack of the two C-terminal Lys residues of the heavy chains. Taking these in account the mass

Table 3
Suggested carbohydrate structures observed in MoAb R3Hu[a]

M_r value[b]	Proposed structure
1809.4	Galβ1-4GlcNAcβ1-2Manα1,3 / Fucα1,6 / Manβ1-4GlcNAcβ1-4GlcNAc / Galβ1-4GlcNAcβ1-2Manα1,6
1647.39	Galβ1-4 { GlcNAcβ1-2Manα1,3 / GlcNAcβ1-2Manα1,6 } / Fucα1,6 / Manβ1-4GlcNAcβ1-4GlcNAc
1485.2	GlcNAcβ1-2Manα1,3 / Fucα1,6 / Manβ1-4GlcNAcβ1-4GlcNAc / GlcNAcβ1-2Manα1,6

[a] Mass values obtained by direct analysis by MALDI-MS of the carbohydrate pool obtained by enzymatic deglycosylation with PGnase F of the MoAb R3Hu heavy chain.

[b] The observed masses all represent $[M+Na]^+$.

difference is 558 (0.4%) between the calculated and observed values. Such mass accuracy is within the expected for large protein taking in consideration that the mass determination is based on external mass calibration of the system using BSA as calibrant.

Acknowledgements

L.M.F. was financially supported for 9 months by the Center of Molecular Immunology. D.E.K. was supported by a grant from the Brazilian Postgraduate Federal Agency (CAPES). This work is part of the activities of the Center for Experimental Bioinformatics sponsored by the Danish National Research Foundation.

References

[1] L. Snelling, C.T. Miyamoto, H. Bender, L.W. Brady, Z. Steplewski, R. Class, J. Emrich, M.A. Rackover, Hybridoma 14 (1995) 111.
[2] R. Pérez, M.R. Pascual, A. Macias, A. Lage, Breast Cancer Res. Treat. 4 (1989) 189.
[3] M.A. Rios, A. Fernández, B.R. Tormo, S. Quintero, I. Pérez, L. Skoog, R. Pérez, Anticancer Res. 12 (1992) 205.
[4] D.E. Neal, C. March, M.K. Benett, P.D. Abel, R.R. Hall, J.R.C. Sainsbury, A.L. Harris, Lancet 1 (1985) 346.
[5] I. Magnusson, A.V. Rosen, R. Nilson, A. Macias, R. Pérez, L. Skoog, Anticancer Res. 9 (1989) 299.
[6] D. Veale, T. Ascroft, C. March, G.J. Gibson, A.L. Harris, Br. J. Cancer 55 (1987) 513.
[7] C.R. Divgi, S. Welt, M. Kris, F.X. Real, S.D. Yeh, R. Gralla, B. Merchant, S. Schweighart, M. Unger, S.M. Larson et al., J. Natl. Cancer Inst. 83 (1991) 97.
[8] A. Fernández, R. Peréz, A. Macias, A. Velandia, I. Alvarez, M. Ramos, A. Velozo, Interferon Biotecnol. 6 (1989) 289.
[9] A. Fernández, E. Spitzer, R. Peréz, F.D. Boehmer, K. Eckert, W. Zschiesche, R. Grosse, J. Cell Biochem. 49 (1992) 157.
[10] N. Rodriguez, O. Torres, M. Ramos, J.A. Gomez, I. Marrero, M. Catalá, M. Sanchez, Y. Perdomo, A. Fernández, Rev. Esp. Med. Nucl. 14 (1995) 285.
[11] A. Morales-Morales, G. Nuñez-Gandolff, N. Peréz, B.C. Veliz, I. Caballero-Torres, J. Duconge, J. Fernández, F.Z. Crespo, A. Veloso, N. Iznaga-Escobar, Nucl. Med. Biol. 26 (1999) 717.
[12] M. Ramos, N. Rodriguez, J.P. Oliva, N. Escobar, A. Perera, A. Morales, N. Gonzales, M. Cordero, L. Torres, G. Pimentel, M. Borron, J. Gonzales, O. Torres, T. Rodriguez, R. Peréz, J. Nucl. Med. 40 (1999) 768.
[13] A. Morales, J. Duconge, I. Caballero, G. Nuñez, E. Fernández, N. Iznaga, Nucl. Med. Biol. 40 (1999) 275.
[14] C. Mateo, E. Moreno, K. Amour, J. Lombardero, W. Harris, R. Pérez, Immunotechnology 3 (1997) 71.
[15] L. Cano, K.M. Swiderek, J. Shively, in: J.W. Crabb (Ed.), Techniques in Protein Chemistry VI, Academic Press, New York, 1995, p. 21.
[16] S. Akashi, K. Noguchi, R. Yuji, U. Tagami, K. Hirayama, K. Kato, H. Kim, K. Tokioka, I. Shimada, Y. Arata, J. Am. Soc. Mass Spectrom. 7 (1996) 707.
[17] S. Akashi, K. Kato, T. Torizawa, N. Dohmae, H. Yamaguchi, M. Kamachi, A. Harada, T. Imanaka, I. Shimada, K. Takio, Biochem. Biophys. Res. Commun. 240 (1997) 566.
[18] K. Hirayama, R. Yuji, N. Yamada, K. Kato, Y. Arata, I. Shimada, Anal. Chem. 70 (1998) 2718.
[19] C.A. Lewis, A.W. Guzetta, W.S. Hancock, Anal Chem. 66 (1994) 585.
[20] G.D. Roberts, W.P. Johnson, S. Burman, K.R. Anumula, S.A. Carr, Anal. Chem. 67 (1995) 3613.
[21] D.M. Sheeley, B.M. Merril, L.C.E. Taylor, Anal. Biochem. 247 (1997) 102.
[22] U.K. Laemmli, Nature 227 (1970) 680.
[23] H. Towbin, T. Staehelin, J. Gordon, Proc. Natl. Acad. Sci. USA 76 (1979) 4350.
[24] R.D. Appel, A. Bairoch, D.F.A. Hochstrasser, Trends Biochem. Sci. 19 (1994) 258.
[25] J. Gobom, E. Nordhoff, E. Mirgorodskaya, R. Ekman, P. Roepstorff, J. Mass Spectrom. 34 (1999) 105.
[26] M. Kusmann, E. Nordhoff, H. Rahbek-Nielsen, S. Haebel, M. Rossel-Larsen, L. Jakobsen, J. Gobom, E. Mirgorodskaya, A. Kroll-Kristensen, L. Palm, P. Roepstorff, J. Mass Spectrom. 32 (1997) 593.
[27] W.C. Barker, J.S. Garavelli, P.B. McGarvey, C.R. Marzec, B.C. Orcutt, G.Y. Srinivasarao, L.L. Yeh, R.S. Ledley, H.-W. Mewes, F. Pfeiffer, A. Tsugita, C. Wu, Nucleic Acids Res. 27 (1999) 39.
[28] M. Schnølzer, P. Jedrzejewski, D.W. Lehmann, Electrophoresis 17 (1996) 945.
[29] E. Krause, H. Wenscuch, P.R. Jungblut, Anal. Chem. 71 (1999) 4160.
[30] B. Thiede, S. Lamer, J. Mattow, F. Siejak, C. Dimmler, T. Rudel, P.R. Jungblut, Rapid Commun. Mass Spectrom. 14 (2000) 496.
[31] E.L. Finley, J. Dillon, R.K. Crouch, K.L. Schey, Prot. Sci. 7 (1998) 2391.
[32] C.Y. Yang, Z.W. Gu, M. Yang, S.N. Lin, G. Siuzdak, C.V. Smith, Biochemistry 38 (1999) 15903.