

Journal of Chromatography B, 658 (1994) 31-38

JOURNAL OF CHROMATOGRAPHY B: BIOMEDICAL APPLICATIONS

Combination of high-performance anion-exchange chromatography and electrospray mass spectrometry for analysis of the *in vitro* O-glycosylated mucin motif peptide

Daniel Tetaert^{*,a}, Benoît Soudan^a, Jean-Marc Lo-Guidice^a, Colette Richet^a, Pierre Degand^a, Guy Boussard^b, Christophe Mariller^c, Geneviève Spik^c

^aUnité INSERM No. 377, Place de Verdun, 59045 Lille Cedex, France

^bLaboratoire de Chimie-Physique Macromoléculaire, CNRS UA 494, Ecole Nationale Supérieure des Industries Chimiques, B.P. 451, 1 rue Granville, 54001 Nancy, France

^cLaboratoire de Chimie Biologique de l'Université des Sciences et Technologies de Lille I, CNRS Unité Mixte de Recherche 111, 59655 Villeneuve d'Ascq Cedex, France

First received 21 December 1993; revised manuscript received 12 April 1994

Abstract

Reversed-phase high-performance liquid chromatography (HPLC) and high-performance anion-exchange chromatography (HPAEC) with pulsed amperometric detection were developed for the study of products obtained from the *in vitro* O-glycosylation of a mucin motif peptide, TTSAPTTS, the most representative sequence encoded by the human gene MUC_{sc} . After incubation of the peptide, which is rich in clustered hydroxyamino acids, by both human colonic and gastric microsomal homogenates, the glycosylated products were separated by HPLC and HPAEC and analysed by electrospray mass spectrometry (ES-MS). The combination of HPAEC and ES-MS was the approach used for evaluating the differences between the polypeptide N-acetylgalactosaminyltransferase activity in different digestive tissues.

1. Introduction

For many glycoproteins, the carbohydrate units confer important physico-chemical properties such as conformational stability, protease resistance, charge and water-binding capacity, and also biological recognition, where sequence diversity provides a signal for protein targeting, differentiation, bacterial and viral adhesion and cell-cell interactions [1-3].

The minimum amino acid sequence require-

ment for N-glycosylation of asparagine residues is well defined: Asn Xaa (Ser/Thr), where Xaa can be any amino acid except proline [4,5]. Conversely, no consensus sequence for Oglycosylation has been found so far. Nevertheless, much attention has been focused on whether the amino acid sequences surrounding the serine and threonine residues in peptide acceptor substrates influence the glycosylation and its extension on specific hydroxyamino acid residues [6–9]. It has been reported that the most prominent feature in the vicinity of Oglycosylated sites is a high frequency of proline

^{*} Corresponding author.

and alanine residues [10,11], near clusters of several hydroxyamino acids [9].

The mucins are the most representative Oglycoproteins and these macromolecules are characterized by tandem repeat peptide motifs in their structure such as Thr-Thr-Ser-Ala-Pro-Thr-Thr-Ser (TTSAPTTS in the one-letter code). This sequence, TTSAPTTS, corresponds to the most appropriate motif for a basic study of O-glycosylation: this tandem repeat peptide, encoded by the gene MUC_{5C} [12], is found naturally and predominantly expressed in the stomach; there are six clustered hydroxyamino acids, one of which is possibly O-glycosylated with a possible influence of a cis-trans prolyl isomerization, and this phenomenon could be easily detected by changing the temperature conditions of HPLC fractionation [13].

However, the analysis by chromatography of in vitro O-glycosylation reactions using this type of peptide acceptor has proved to be complex, because of the very hydrophilic nature of the resulting O-glycosylated products. Moreover, the effectiveness and reliability of the chromatographic separation techniques have not been well documented. Our experimental approach was therefore to study the in vitro O-glycosylation of the octapeptide substrate TTSAPTTS using highperformance anion-exchange chromatography (HPAEC) with pulsed amperometric detection (PAD). A protocol for separation of O-glycosylated peptide products was developed, and the analysis of reaction products was performed by electrospray mass spectrometry (ES-MS).

2. Experimental

2.1. Materials

AG1-X8 (100–200 mesh, Cl⁻ form) was purchased from Bio-Rad (Vitry s/Seine, France), MES [2-(N-morpholino)ethanesulfonate], MnCl₂, PMSF (phenylmethanesulfonyl fluoride), AMP (adenosine monophosphate) and Triton X-100 from Sigma (St. Louis, MO, USA) and UDP-N-[1-³H]acetylgalactosamine (8.7 Ci/mmol) from New England Nuclear Chemicals (DuPont de Nemours, Les Ulis, France). The synthetic peptide substrate NH_2 -TTSAPTTS-COOH was obtained from Neosystem (Strasbourg, France) and the amino acid sequence was further checked using an Applied Biosystems Model 477A gas-phase sequencer.

2.2. Enzyme preparation

Fundic human gastric (3 g) and colon (3-10 g)samples were obtained from patients after surgery. The tissues were cleaned with 0.125 Mphosphate-buffered saline (PBS) (pH 7.4); the overlying mucus was removed with a glass slide and the mucosa was scraped off with a scalpel blade. Microsome suspensions were prepared by Potter-Elvehjem homogenization of tissue in ice-cold 0.25 M sucrose-0.2 M NaCl and centrifugation at 16000 g for 10 min followed by ultracentrifugation of the supernatant at 140 000 g for 20 min to produce the microsomal pellet. The pellet was resuspended in a minimum amount of NaCl-sucrose solution to provide a final protein concentration of 3-4 g/l; the solution was stored at -80° C until used.

Human stomach and colon samples were healthy tissues from cancer patients who had undergone resection.

2.3. Analytical methods

The protein concentration was determined on a Hitachi 717 multiparametric automated chemistry analyser (Boehringer, Mannheim, Germany) using the Biotrol (Chennevières les Louvres, France) kit protein assay reagent.

2.4. Assay for UDP-N-acetylgalactosamine: polypeptide N-acetylgalactosaminyl transferase activity

Assay mixtures in a total volume of 84 μ l contained the following components at final concentrations: acceptor 6 mM TTSAPTTS (15 μ l); 125 mM MES buffer (pH 7.0) (21 μ l) containing 0.1% (v/v) Triton X-100, 12.5 mM MnCl₂, 1 mM PMSF; 1.25 mM AMP (15 μ l); 1 mM UDP-N-acetylgalactosamine-2.5-5 nM

UDP-N-acetyl[1-³H]galactosamine $(0.7 \cdot 10^6 \cdot 10^6$ dpm) (15 μ l), and 30 μ g of microsomal preparation (18 μ l). After the addition of the enzyme preparation, samples were incubated for 3 h at 37°C. The reaction was stopped by the addition of 8 volumes (672 μ l) of 20 mM sodium borate– 1 mM ethylenediaminetetraacetic acid (pH 9.1). Products were separated on an AG1-X8 anionexchange resin column (2 ml) and eluted with water (3 × 1.0 ml).

2.5. Reversed-phase high-performance liquid chromatography

HPLC methods were developed using a Varian (Les Ulis, France) Model 5000 system. For RP-HPLC, a Beckman Ultrasphere ODS column (5 μ m, 80 Å pore size; 250 × 4.6 mm I.D.) and a guard column (50×4.6 mm I.D., filled with Corasil C_{18} , 50 μ m pore size) were used (Altex Beckman, Berkeley, CA, USA). Eluates were detected with a Model 2238 Uvicord S II UV absorbance detector (LKB, Uppsala, Sweden) set at 206 nm and a sensitivity of 0.2 absorbance full-scale. Samples in 50 µl [NH₂-TTSAPTTS-COOH (5-20 μ g) and glycosylated related peptides (12000 dpm)] were introduced into the column and eluted at a flow-rate of 1.2 ml/min. Different elution conditions were employed: isocratic elution with 18% acetonitrile (Carlo Erba, Rueil Malmaison, France) in water containing the counter ion trifluoroacetic acid [0.1%](v/v) TFA] or a gradient of 0% in 5 min then from 0 to 20% acetonitrile (0.1% TFA) in 20 min; finally, the column was washed with 50% acetonitrile (0.1% TFA). The NH_2 column (Beckman Ultrasphere NH₂, 5 μ m, pore size 80 Å) was also tested by elution using a linear gradient over 35 min from 90 to 30% acetonitrile containing 0.1% TFA. A flow-rate of 1.0 ml/min was used, and the column was operated at room temperature.

2.6. High-performance anion-exchange chromatography

The glycopeptides were fractionated by HPAEC on a Dionex (Sunnyvale, CA, USA) LC

system consisting of a BioLC GP-M2 quaternary gradient module, a triple-pulse amperometric detector, equipped with a gold electrode (PAD II), and a CarboPac PA-100 guard column (50 \times 4 mm I.D.) followed by a CarboPac PA-100 analytical column (250×4 mm I.D.). Elution was performed at a flow-rate of 1.0 ml/min using buffer A [100 mM NaOH (J.T. Baker, Noisy le Sec, France) and buffer B [1 M sodium acetate (Merck, Darmstadt, Germany) containing 100 mM NaOH] under gradient conditions: buffer A, 10 min; 0 to 30% buffer B in 22 min; 30 to 60% buffer B in 3 min. Finally, the column was washed for 5 min with 60% buffer B. The Na⁺ ions produced from NaOH and sodium acetate were removed by connection of an anion selfregenerating suppressor (ASRS-1; Dionex) to the detector outflow. Analysis by flame spectrometry confirmed the absence of sodium ions (<0.1 mequiv./l compared with 0.5 mequiv/l for)buffer A).

2.7. Mass spectrometry

The molecular mass of the *in vitro* glycosylated products was determined by electrospray mass spectrometry (ES-MS) using a Perkin-Elmer mass spectrometer (Perkin-Elmer SCIEX, Toronto, Canada) with an electrospray ion source operating at atmospheric pressure followed by a quadrupole mass analyser with an upper mass limit of m/z = 2400. The samples (10 μ g in 100 μ l) were introduced into the source at a flow-rate of 3 μ l/min (carrier solvent, 20% acetonitrile in water containing 0.1% formic acid, pH 1.0). The sampling cone voltage was adjusted to 90 V to obtain the best sensitivity.

3. Results

The peptide substrate NH_2 -TTSAPTTS-COOH was subjected to the O-glycosylation reaction by the gastric and colonic microsomal preparations. The enzymatic activity was measured by the transfer of radiolabelled [³H]GalNAc from UDP-GalNAc to the peptide motif TTSAPTTS and the unemployed UDP- GalNAc was discarded by passing the reaction products through AG1-X8 resin. The apparent $K_{\rm m}$ and $V_{\rm max}$ values were calculated by determining the rate of incorporation that occurred at various substrate concentrations. The results showed that the affinity of GalNAc transferase(s) for the acceptor was higher in the colon ($K_{\rm m} =$ 0.56 mM; $V_{\rm max} = 38.8 \ \mu {\rm mol/h} \cdot {\rm g}$) than in the stomach ($K_{\rm m} = 9.75 \ {\rm m}M$; $V_{\rm max} = 39.9 \ \mu {\rm mol/h} \cdot {\rm g}$).

The products of the in vitro O-glycosylation reactions [using microsomes of human colon (HCR) and human stomach (HSR)] were separated on the C₁₈ HPLC column. Several injections of the reaction products [obtained from HCR (Fig. 1) and HSR] were made and the two profiles were similar. The radiolabelled fractions (fractions 1 and 2) were studied by ES-MS: (i) the column-excluded fraction 1 consisted of Nacetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc) and low-m/z components that corresponded to contaminants (determination by ES-MS) such as small peptides and/or glycopeptides obtained by the action of peptidase(s) in microsomal homogenates on the parent peptide or on the formed glycopeptides during the in vitro O-glycosylation incubation; (ii) fraction 2 was heterogeneous, but it contained the glycopeptidic fractions. Indeed, the ES mass spectra of HCR fraction 2 revealed the presence of one and two GalNAc residues in the glycopeptides detected at m/z = 968.0and 1170.6, respectively, in addition to the unmodified octapeptide TTSAPTTS $[M + H]^+$, m/ z = 765.4 (Fig. 2A). Furthermore, we observed during the UDP-GalNAc incorporation a phenomenon of capture of ions $(Na^+ and K^+)$ because $m/z = 787.4 (+23) [M + Na]^+$, m/z = $803.2 (+39) [M+K]^+$ were found and the molecular masses for the glycopeptides were recovered as a triplet in the singly charged forms (+1 GalNAc = 968.0, 989.4, 1005.6 and +2GalNAc = 1170.6, 1191.6, 1207.6) and also in the doubly charged forms (+1 GalNAc = 483.0,493.0, 503.0 and +2 GalNAc = 585.0, 593.2, 603.8).

Concerning the related radiolabelled HPLC fractions obtained from HSR, the ES mass



Fig. 1. Reversed-phase HPLC separation (C₁₈ column) of *in vitro* O-glycosylated products using human digestive mucosa homogenates (stomach) and NH₂-TTSAPTTS-COOH peptide as acceptor by gradient elution (0-20%) with acetonitrile in water containing 0.1% TFA. Radioactive product (dashed line) and absorbance at 206 nm (solid line) are indicated.

spectra also showed the peptide TTSAPTTS and glycopeptides as a triplet (m/z = 967.2, 989.4, 1005.2 and 1170.0, 1192.0, 1208.0) (Fig. 2B). The main difference between the spectra obtained from the two tissues was observed in the low-molecular-mass products, which presumably resulted from a degradative process related to different peptidase(s) still active in presence of the PMSF [e.g., the component relevant to m/z664.2 could be assigned to TSAPTTS, m/z 563.0 to SAPTTS and m/z 405.2 to PTTS (Fig. 2A)



Fig. 2. ES mass spectra of HPLC peak 2: (A) after reaction using human colonic microsome preparation (HCR); (B) after reaction using human gastric microsome preparation (HSR). In the insets, expanded mass scales are shown between m/z 800 and 1250 to show the capture of Na⁺ and K⁺ ions by the peptide [M + H]⁺ and also the glycopeptides, which are found as triplets.

and m/z 664.2 to TSAPTTS and m/z 577.1 to TSAPTT or TTSAPT (Fig. 2B)]. However, the RP-HPLC procedure developed seemed unsatisfactory in this precise study for a very accurate analysis of the *in vitro* O-glycosylation reaction, because peptide and glycopeptides were coeluted from the column. On the other hand, the use of another support such as normal-phase (NH_2-HPLC) column also allowed the slight separation of peptide TTSAPTTS and related glycopeptides, but the mono- and/or poly-glycosylated compounds were found in an unique fraction (data not shown).

Therefore, HPAEC was developed with pulsed



Fig. 3. Separation by HPAEC (Dionex system and Carbopac PA 100 column) of the *in vitro* O-glycosylation products obtained by incubation with digestive microsomal homogenates and the peptide TTSAPTTS as substrate: (A) by human colonic preparation (HCR); (B) by human gastric preparation (HSR). Peaks 1-6 = fractions obtained during isocratic elution with 100 mM NaOH and peaks a-d = fractions obtained during gradient elution (100 mM NaOH-0 to 1 M sodium acetate).

amperometric detection (PAD) and the Dionex system. This method was found to be a powerful tool for the analysis of less than 100 nmol of the *in vitro* O-glycosylated material. Our study was focused on the fractions where radioactivity was detected and these fractions were studied by ES-MS.

The elution profile in Fig 3 can be divided into

two parts. During the isocratic elution with 100 mM NaOH in 10 min, free sugars (GlcNAc and GalNAc; peaks 1 and 2, respectively) were eluted at the beginning of the chromatogram, whereas deacetylated O-glycopeptides and products of degradation of peptides and glycopeptides were slightly retained (peaks 3-6). Peak 6 was predominantly formed by the sucrose arising from the homogenate and distinctly detected by PAD.

The second part of the pattern obtained by elution of the sodium acetate gradient (peaks a-d; Fig. 3A and B) consisted primarily of undegraded glycopeptides. The one- (m/z 968)and two-[³H]GalNAc (m/z 1170) glycopeptides were found from peaks a-d when the unlabelled peptide was found in peak a (m/z 765.4). Moreover, the HPAEC pattern provided information concerning differences in O-glycosylation according to the choice of microsomal preparations (HSR or HCR): the in vitro O-glycosylated products of the peptide TTSAPTTS obtained with human colon reaction (HCR) gave at least four peaks (Fig. 3A) whereas a major single peak was obtained with human gastric microsomes (HSR) (Fig. 3B). For the colon, the later elution from the HPAEC column reflected the high percentage of polyglycosylated peptides as found by ES-MS analysis (data not shown).

4. Discussion

The microsomal homogenates isolated from human stomach and colon contain activity(ies) of UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase (GalNAc transferase) (EC 2.4.1.41) (the enzymes that initiated mucin-type O-glycosylation by addition of GalNAc to a hydroxyl group of serine or threonine residue). Several lines of evidence indicate that GalNAc transferase(s) reside in the *cis*-Golgi and/or the endoplasmic reticulum [14], and that the enzymes usually prepared from bovine colostrum also exist in a soluble form [15]. Special attention has also been paid to a GalNAc transferase extracted from porcine submaxillary glands and that is specific for glycosylation of threonine residues [16].

However, there is now a precedence for more than one mRNA species for at least four glycosyl transferases, and the genes coding for these enzymes appear to be under the control of more than one promoter [17-19]. In our experiments the HPAEC profiles obtained from the two homogenates account for the diversity and multiplicity of the GalNAc transferases in the human digestive tissues because the in vitro O-glycosylation products obtained from colonic microsomal reaction give at least four peaks of glycopeptides containing one and two GalNAc residues with regard to the major peak found from gastric products. Nevertheless, the possibility cannot be excluded that some peaks reflect a variable specificity of a single transferase toward the substrate according to a cis-trans isomerization and/or that, in the colonic microsomal homogenate, the prolyl isomerase activity could be more important with an increase in the accessibility of preferential glycosylation sites. Moreover, using a hydrophobic form of the peptide (N-t-BOC-TTSAPTTS-CO-NH-CH₁) and at different column temperatures, the separation of the cistrans-prolyl configurations of the octapeptide is more closely obtained (data not shown).

The acceptor specificity is not yet completely understood for any of these transferases. Recently, the study of a flanking-sequence requirement for O-glycosylation included all possible combinations of residue(s) surrounding the hydroxy amino acid. In an attempt to define a consensus signal, different motif peptides have been proposed for use in in vitro O-glycosylation, using sequences based on those (i) inferred from the 196 different O-glycoproteins extracted from the National Biomedical Research Foundation Protein Database and (ii) known glycopeptide motifs such as human von Willebrand factor (HVF) or myelin basic protein (MBP) [8,18–21]. The compilation of the results emphasizes the importance of proline and alanine residues near the site which induces a β -turn, and that the presence of clustered hydroxyamino acids and an eight amino acid segment is generally required.

The analysis of O-glycosylated products from a short and very hydrophilic peptide is complex and the characterization of such products, the kinetics of reaction and the purification of the resulting glycopeptides remain limited. The synthetic peptide substrate chosen for our study is based on the mucin tandem-repeat peptide motif MUC_{5C} (TTSAPTTS), containing six potential O-glycosylation sites in two clusters of hydroxy-amino acids spaced by a proline. This octapeptide is naturally found and strongly expressed in the human stomach mucins, whereas for those in colon, *in situ* hybridization showed a diffuse signal by the corresponding MUC_{5C} RNA probe [12,22].

Hence, according to the tissue origin of microsomal preparations, the glycosylation of TTSAPTTS could give rise to several possible products of monoglycosylated and two-glycosylated peptides revealed by studying the HPAEC elution profiles. Indeed, it appears that the GalNAc transferases in the stomach presumably have a close specificity, because preferential positions seem to accept GalNAc; however, in the colon, the linkage is randomly achieved.

The effects of the flanking sequence on the incorporation rate of GalNAc into Thr or Sercontaining substrates differ significantly according to different workers, although the role of the surrounding residues appear to be prominent [8,11,17,23]. The positions -1 and +3 relative to the site are identified as particularly significant, and a proline residue is often associated with positions +3 and -1. Nevertheless, these conclusions are based on results concerning motif peptides where hydroxyamino acids are poorly represented. Further investigations are in progress in order to state precisely (i) the activity(ies) of GalNAc transferases (linkage to Ser and/or Thr) and (ii) the positions of glycosylation with the possible role of a cis-trans isomerization. Therefore, the combination of HPAEC and ES-MS would become very efficient for purifying and analysing the glycopeptides that are usually co-eluted with the other reaction products using other chromatographic procedures such as HPLC on C_{18} and NH_2 columns.

Acknowledgements

We thank Drs. G. Briand and M. Kouach (Service Commun de Spectrométrie de Masse de Lille II, Faculté de Médecine, Lille, France) for running samples and for interpretation of the ES mass spectra and Professor P. Quandalle (Service de Chirurgie Adulte Ouest, Hôpital C. Huriez, Lille, France) for providing tissues. We are indebted to Professor J. Gagnon (Laboratoire de Biologie Structurale, Commissariat à l'Energie Atomique, CEN-G, Grenoble, France) for protein sequencing. We are grateful to Professor A. Verbert (Directeur de l'Unité Mixte CNRS 111) for generous support and advice.

References

- J. Montreuil, in A. Neuberger and L.L.M. van Deenen (Editors), *Comprehensive Biochemistry*, Vol. 19BII, Elsevier, Amsterdam, 1982, p. 1.
- [2] A. Kobata, Eur. J. Biochem., 209 (1992) 483.
- [3] H. Lis and N. Sharon, Eur. J. Biochem., 218 (1993) 1.
- [4] T. Osawas and T. Tsuji, Annu. Rev. Biochem., 56 (1987) 21.
- [5] K.K. Han and A. Martinage, Int. J. Biochem., 24 (1992) 19.
- [6] P. Degand, R. Havez and G. Biserte, C.R. Acad. Sci., 275 (1972) 491.
- [7] H.D. Hill, M. Schwyzer, H.M. Steiman and R.L. Hill, J. Biol. Chem., 252 (1977) 3799.
- [8] B.C. O'Connell, F.K.L.A. Tabak and N. Ramasubbu, Biochem. Biophys. Res. Commun., 180 (1991) 1024.

- [9] A.P. Elhammer, R.A. Poorman, E. Brown, L.L. Maggiora, J.G. Hoogerheide and F.J. Kézdy, J. Biol. Chem., 268 (1993) 10029.
- [10] I.B.H. Wilson, Y. Gavel and G. von Heijne, Biochem. J., 275 (1991) 529.
- [11] A.A. Gooley, B.J. Classon, R. Marschalek and K.C. Williams, Biochem. Biophys. Res. Commun., 178 (1991) 1194.
- [12] N. Porchet, J. Nguyen Van Cong, J. Dufosse, J.P. Audié, V. Guyonnet-Duperat, M.S. Gross, C. Denis, P. Degand, A. Bernheim and J.P. Aubert, *Biochem. Biophys. Res. Commun.*, 175 (1991) 414.
- [13] J.C. Gesquière, E. Diesis, M.T. Cung and A. Tartar, J. Chromatogr., 478 (1989) 121.
- [14] R.D. Cummings, S. Kornfeld, W.J. Schneider, K.K. Hobgood, M. Tolleshang, M.S. Brown and J.L. Goldstein, J. Biol. Chem., 258 (1983) 15261.
- [15] A. Elhammer and S. Kornfeld, J. Biol. Chem., 261 (1986) 5249.
- [16] Y. Wang, J.L. Abernethy, A.E. Eckhardt and R.L. Hill, J. Biol. Chem., 267 (1992) 12709.
- [17] D.H. Joziasse, Glycobiology, 2 (1992) 271.
- [18] F.L. Homa, T. Hollander, D.J. Lehman, D.R. Thomsen and A.P. Elhammer, J. Biol. Chem., 268 (1993) 12609.
- [19] F.K. Hagen, B. VanWuyckhuyse and L.A. Tabak, J. Biol. Chem., 268 (1993) 18960.
- [20] B.C. O'Connell, F.K. Hagen and L.A. Tabak, J. Biol. Chem., 267 (1992) 25010.
- [21] J.M. Cotrell, R.L. Hall, G. Sturton and P.W. Kent, Biochem. J., 283 (1992) 299.
- [22] J.P. Audié, A. Janin, N. Porchet, M.C. Copin, B. Gosselin and J.P. Aubert, J. Histochem. Cytochem., 41 (1993) 1479.
- [23] Y. Wang, N. Agrawal, A.E. Eckhardt, R.D. Stevens and R.L. Hill, J. Biol. Chem., 268 (1993) 22979.