Biochemistry **2005,** *44,* 5453−5460

Identification of the Glycosylation Site of the Adenovirus Type 5 Fiber Protein

Gilles Cauet,*,‡ Jean-Marc Strub,§ Emmanuelle Leize,§ Elsa Wagner,§ Alain Van Dorsselaer,§ and Monika Lusky‡

Transgene SA, 11 rue de Molsheim, 67082 Strasbourg, France, and Laboratoire de Spectrométrie de Masse Bio-Organique ECPM, 25 rue Becquerel, 67087 Strasbourg, France

Received October 28, 2004; Revised Manuscript Received February 2, 2005

ABSTRACT: The fiber protein purified from the pool of nonincorporated viral protein after infection of cells with adenovirus type 5 exists as two forms separable by reverse-phase HPLC. As determined by mass spectrometry, this heterogeneity results from a difference in one O-linked *N*-acetylglucosamine (GlcNac). A western blot analysis using a monoclonal antibody directed against the GlcNac motif showed that only one of the two forms reacted with the antibody, suggesting that one form carries a single GlcNac and the other form has none. The ratio of glycosylated to nonglycosylated forms of fiber, which is about 1, is conserved in assembled viruses. After digestion of glycosylated fiber with endoproteinase GluC, isolation of the glycosylated peptide by reverse-phase HPLC, and chemical derivatization using dimethylamine, the site of glycosylation was located in the fiber shaft at serine 109 by mass spectrometry. Elimination of glycosylation by site-directed mutagenesis of fiber should help to understand the function of this postranslational modification.

Among the capsid proteins of adenovirus, the fiber protein plays a key role in virus adsorption to various cell receptors (*1*, *2*), allowing infection of a wide variety of cell types. The adenovirus fiber is located at each of the 12 vertices of the viral icosahedron. Fiber consists of a homotrimer divided into three domains, a C-terminal globular head, the knob, a long thin shaft, and an N-terminal tail that anchors the fiber to the capsid via direct interactions with the penton base protein (*3*−*5*). It has been shown previously that Ad5 and Ad2 fibers contained an O-linked *N*-acetylglucosamine (O-linked GlcNac)[1] (*6*−*8*). Fiber glycosylation seems to be a specific feature of Ad2 and Ad5 since fibers from other adenovirus serotypes do not apparently contain this modification (*7*). Although it has been suggested that fiber O-linked glycan might be important for fiber assembly or stabilization, the functional role of this modification remains obscure.

Proteins containing O-linked GlcNac are predominantly localized in the cytoplasm and nucleus of eukaryotic cells. Although little is known about the precise role of O-linked GlcNac modification of proteins, it seems important in a number of fundamental cellular events such as transcription, translation, nuclear transport, and cytoskeletal assembly (reviewed in ref *9*).

To date, no consensus sequence has been identified for the O-glycosylation of Ser or Thr residues. It usually occurs in Ser/Thr-rich domains containing proline residues. Thus, identification of the glycosylated sites is a challenging problem that requires a biochemical characterization of the glycosylated protein including protease treatment and isolation of the glycosylated peptides. Mass spectrometry is particularly suited for the rapid and sensitive sequencing of peptides, and direct identification of glycosylation sites has been achieved in a number of favorable cases (*10*, *11*). However, due to the lability of the glycosidic bond, a chemical modification of the glycosylated peptide is often necessary in order to generate more stable derivatives prior to mass spectrometry analysis (*12*−*15*).

In this study, it is shown that the Ad5 fiber is partially modified by an O-linked *N*-acetylglucosamine. Peptide mapping associated with mass spectrometry allowed the identification of a single glycosylation site.

MATERIALS AND METHODS

Chemicals. Modified porcine trypsin sequencing grade was from Promega. Endoproteinase GluC sequencing grade, α-cyano-4-hydroxycinnamic acid, dimethylamine, and ethanethiol were from Sigma.

Purification of Fiber. Fiber was purified from the lysate of infected cells essentially as previously described (*16*). Briefly, E1 complementation cells were infected with AdTG6401 (*17*) at MOI 5. Cells were harvested 48 h after infection and washed in 50 mM phosphate buffer, pH 7.0. Washed cells were resuspended in 50 mM phosphate buffer, pH 7.0, containing a mixture of proteases inhibitors (Complete, Boehringer), disrupted by successive freezing and thawing, and subsequently centrifuged at 3000*g* for 10 min to remove cell debris. The supernatant was loaded on a cesium chloride gradient and centrifuged for 2 h at 30000

---

* Address correspondence to this author at Biopep S.A., 471 rue Nungesser, Le Mas des Cavaliers II, 34130 Mauguio, France. Tel: +33(0)467132105. Fax: +33(0)467132101. E-mail: gilles.cauet@biopep.com.

‡ Transgene SA.

§ Laboratoire de Spectrométrie de Masse Bio-Organique ECPM.

[1] Abbreviations: RP-HPLC, reverse-phase high-performance liquid chromatography; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid; MALDI-TOF, matrix-assisted laser desorption/ionization time of flight; MS, mass spectrometry; Q-TOF, quadrupole time of flight; GlcNac, *N*-acetylglucosamine; GFP, Glu-fibrinopeptide.

10.1021/bi047702b CCC: $30.25

rpm in a SW40 rotor (Beckman). The material above the virion band was collected and dialyzed against 50 mM potassium phosphate buffer, pH 6.8 (buffer A) overnight at 4 °C. After dialysis, the supernatant was loaded at 1 mL/min on a Hitrap DEAE-FF column (Amersham Biosciences) equilibrated in buffer A. The flow-through was collected and directly applied onto a column of hydroxylapatite (macroprep ceramic type I, Bio-Rad) equilibrated in buffer A. Fiber was eluted with a linear gradient from 50 to 250 mM potassium phosphate buffer, pH 6.8, in 30 min.

SDS−PAGE. SDS−PAGE analysis was performed in 12% SDS−polyacrylamide gels in a Tris−glycine buffer (Novex). Proteins were stained with colloidal Coomassie blue G250 (*18*).

Western Blot. Proteins were denatured and resolved on a Nupage 10% polyacrylamide Bis-Tris gel using MOPS as the running buffer (Invitrogen). The fiber protein was identified with a polyclonal antiserum directed against the knob domain of the Ad5 fiber (serum E642, 1:3000; obtained from R. Gerard); the glycosylation state was identified with a monoclonal anti-glycosylation antibody (hybridoma supernatant RL2, 1:50; obtained from P. Boulanger). The secondary antibodies used were HRP-conjugated anti-rabbit IgG and HRP-conjugated anti-mouse IgG. The enhanced chemiluminescence detection system (ECL; Amersham) was used to detect the protein.

Proteolysis. HPLC-purified fiber was digested with either trypsin or endoproteinase GluC. HPLC fractions containing fiber were dried in a SpeedVac and resuspended in digestion buffer. Trypsin digestion were performed overnight at 28 °C in 25 mM ammonium bicarbonate, pH 8.0. The ratio of trypsin to fiber was 1/5 to 1/10. GluC digestions were performed at 32 °C in either 100 mM ammonium bicarbonate or 100 mM phosphate buffer, pH 8.0. The enzyme to fiber ratio was 1/20 to 1/40.

Reverse-Phase HPLC. Fiber was routinely analyzed by RP-HPLC (HP1090; Agilent) on a C4 Jupiter column (150 × 2.0 mm; Phenomenex) at 0.2 mL/min and 60 °C. Solution A was 0.1% trifluoroacetic acid (TFA) in water; solution B was acetonitrile/2-propanol (50:50 v/v) containing 0.1% TFA. Fiber was eluted with a linear gradient from 20% to 60% solution B in 40 min. UV absorption was monitored at 215 nm.

Tryptic or GluC digests were analyzed on a C18 Uptishere column (300 Å, 3 μm, 150 × 2.1 mm; Interchim) at 0.2 mL/min and 50 °C. Peptides were eluted with a linear gradient of aqueous acetonitrile containing 0.1% TFA from 5% to 65% in 60 min. UV absorption was monitored at 215 nm.

β-Elimination/Alkylation. Alkaline deglycosylation by dimethylamine and subsequent alkylation by ethanethiol were performed as described previously (*15*). After solvent evaporation using a Speed Vac, HPLC-purified peptides were incubated for 5−8 h in 100 μL of a mixture of 2-propanol/dimethylamine/ethanethiol (1:1:1 v/v/v) at 45 °C. The reaction was stopped upon elimination of the reagents by evaporating centrifugation in a Speed Vac.

Mass Spectrometry Analyses. (A) MALDI-MS. A saturated solution of the matrix α-cyano-4-hydroxycinnamic acid (0.5 μL) in acetone was deposited onto the sample probe. After drying, 0.5 μL of 2% TFA in water was added to the matrix surface and mixed with 0.5 μL of the tryptic digest and 0.5 μL of a saturated solution of the matrix in acetonitrile/water (50:50). Samples were dried in an air vacuum apparatus. Spectra were obtained on a MALDI time-of-flight instrument (Biflex, Bruker) operated in either reflectron or linear modes. Spectra were mostly externally calibrated using a mixture of synthetic peptides of known masses, Arg-Leu-enkephalin, angiotensin, substance P, bombesin, and ACTH, with monoisotopic masses respectively at m/z 712.37, 1046.54, 1347.73, 1619.82, and 2465.19. Measured masses were monoisotopic or average depending on the type of experiment (see text).

Before MALDI-MS, peptides were occasionally concentrated on ZipTip (Millipore) following the instructions of the manufacturer.

(B) MALDI-MS/MS. Samples were deposited on the target using the dried droplet method. The sample (0.5 μL) was mixed on target with 0.5 μL of a saturated solution of the matrix α-cyano-4-hydroxycinnamic acid in acetonitrile/water (50:50). MALDI-MS/MS spectra were recorded on an Ultraflex mass spectrometer (Bruker). The instrument was calibrated using the fragment ions of angiotensin with monoisotopic mass at m/z 1046.54.

(C) Liquid Chromatography−Mass Spectrometry. Analyses was performed on a triple-quadrupole mass spectrometer equipped with an electrospray source (QuattroII, Micromass), operating in the positive mode. During on-line LC-MS, spectra were scanned over m/z 400−1600 for 6 s, and the resolution was adjusted so that the peak at m/z 998 from heart myoglobin was 1.5 wide at the base. The mass spectrometer was calibrated using the multiply charged ions produced by a separate introduction of horse myoglobin (16951.48 Da). Fiber was eluted in the conditions described above using a C4 Jupiter column (150 × 2.0 mm) coupled to the mass spectrometer.

Nanoelectrospray Mass Spectrometry. (A) Q-TOF Mass Spectrometer. MS and MS/MS mass measurements were performed with a Q-TOF 2 hybrid quadrupole/time-of-flight mass spectrometer (Micromass) equipped with a Z-spray ion source. The instrument consists of an electrospray ionization source, a quadrupole mass filter operating as a variable band-pass device, a hexapole collision cell, and an orthogonal acceleration time-of-flight (TOF) mass analyzer. The MS and MS/MS analyses were performed with Protana medium NanoES spray capillaries. The instrument was calibrated using fragment ions of GFP.

(B) Ion-Trap Mass Spectrometer. Tandem mass spectra were acquired on an ESQUIRE 3000+ Ion-Trap mass spectrometer (Bruker) equipped with a nanospray ion source. MS/MS analysis was performed with Protana medium NanoES spray capillaries. Calibration of the Ion-Trap analyzer was performed by using the multiply charged ions produced by the following peptide mixture, Arg-Leu-enkephalin, angiotensin, substance P, bombesin, and ACTH, with monoisotopic masses respectively at $m/z = 712.37$ (1+) and 356.68 (2+), $m/z = 1046.54$ (1+) and 523.77 (2+), $m/z = 1347.73$ (1+) and 674.36 (2+), $m/z = 1619.82$ (1+) and 810.41 (2+), and $m/z = 1233.09$ (2+).

RESULTS

Ad5 Fiber Is Heterogeneous. Ad5 fiber was partially purified from the cellular pool of nonincorporated viral capsid

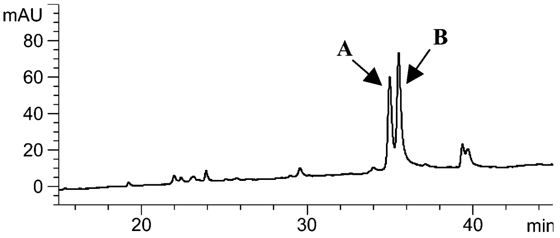


FIGURE 1: RP-HPLC analysis of partially purified fiber. Peaks A and B contained fiber as determined by MALDI-TOF mass spectrometry.

proteins by a two-step procedure using DEAE and hydroxylapatite chromatographic separations essentially as described by Boulanger and Puvion (*16*). A prominent protein band at about 62 kDa was identified as being fiber after in-gel trypsin digestion and analysis of the tryptic peptides by MALDI-TOF mass spectrometry. Trimerization of fiber was assessed by an SDS−PAGE separation of the partially purified protein without boiling it in the sample buffer (*19*). Under these conditions, fiber migrates at a M_r consistent with that of a trimer (not shown).

Purified fiber was further analyzed by RP-HPLC using a linear gradient of aqueous acetonitrile/2-propanol (50:50 v/v) at 60 °C (Figure 1). Peaks were collected, dried, digested by trypsin, and analyzed by MALDI-MS. Fiber was identified in both peak A and peak B.

To investigate whether the difference between the two fiber peaks on HPLC resulted from a difference in mass, the column was coupled to an electrospray mass spectrometer. The measured average masses of the A and B forms of fiber were 61842.51 ± 4.27 and 61637.82 ± 2.25 Da, respectively. The mass difference of 204.69 Da between these peaks was consistent with the presence of an *N*-acetylglucosamine (GlcNac) motif (expected difference 203.19 Da) already shown to be present on fiber (*6*−*8*). Thus a difference in glycosylation might explain the heterogeneity of fiber observed on RP-HPLC.

Identification of the O-Glycosylated Peptide. To confirm this result, fiber was purified by RP-HPLC using a steep gradient of acetonitrile/2-propanol at 60 °C. Fiber A and B forms were collected in the same tube, dried in a Speed Vac, and digested overnight with trypsin. The trypsin-generated peptides were analyzed by MALDI-MS searching for a mass increment of 203 Da between two signals indicative of the presence of a GlcNac motif. The data obtained, together with the data generated from previous in-gel digested fiber, resulted in about 40% protein sequence coverage up to *m*/*z* 3000 with no indication of the presence of a glycosylated peptide (not shown). Since trypsin should generate several large peptides from fiber, the analysis was performed also in the linear mode up to *m*/*z* 7000. To obtain a more accurate mass determination, spectra were internally recalibrated using fiber peptides at *m*/*z* 3646.15 and 6162.72. The mass signals of the expected six largest peptides which represented about 42% of the protein sequence were observed. Interestingly, a mass difference of 203.27 Da between *m*/*z* 4779.27 and *m*/*z* 4982.54 was observed (Figure 2). This indicated that the peptide of calculated *m*/*z* 4779.51 (average) whose sequence is comprised between Ser[95] and Lys[141] should carry a GlcNac motif. This observation is in agreement with previous work which localized the glycosylation site between residues 18 and 194 (*7*). This result was confirmed when the A and B forms were collected separately after RP-HPLC, digested with trypsin, and analyzed by MALDI-MS in the linear mode. The signal at *m*/*z* 4982 was only observed in the tryptic digest of the A form, whereas the signal at *m*/*z* 4779 was only observed in the tryptic digest of the B form (not shown.). This result was consistent with the fact that the A form has a molecular mass increment of about 203 Da with respect to the B form.

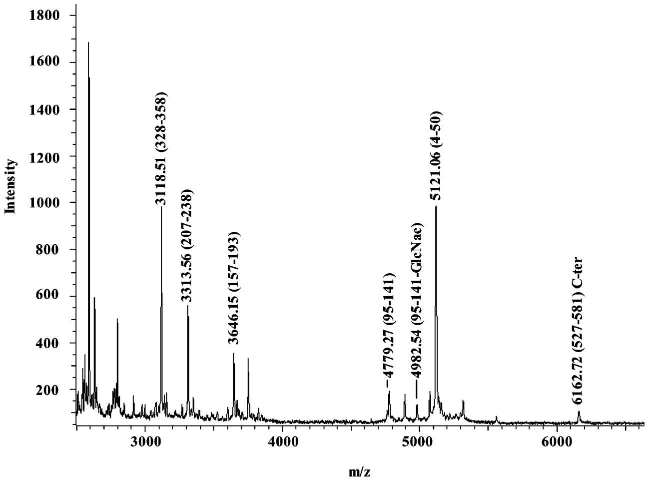


FIGURE 2: MALDI-TOF mass spectrum of Ad5 fiber purified by RP-HPLC obtained in the linear mode. Indicated *m*/*z* are averaged values. Numbers in parentheses are the amino acid position of the corresponding tryptic peptides.

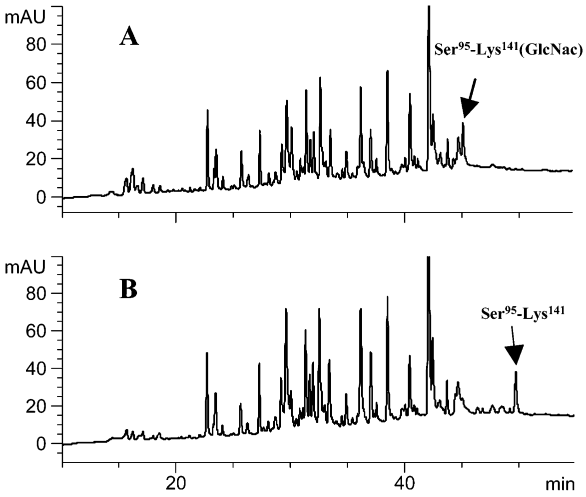


FIGURE 3: RP-HPLC analysis of the A form (A) and B form (B) of fiber digested with trypsin. Arrows indicate the peptide Ser[95]−Lys[141] of mass 4982 in (A) and 4779 in (B) measured by MALDI-TOF mass spectrometry.

Glycosylated and nonglycosylated forms of the peptide Ser[95]−Lys[141] were purified by RP-HPLC after digestion of fiber with trypsin. For this, the tryptic digests of the A and B forms of fiber were separately fractionated by RP-HPLC on a C18 column. The HPLC traces were identical except for the two peaks indicated by the arrows in Figure 3. These peaks were collected and identified by MALDI-MS. The peak collected from the tryptic digest of the A form of fiber (Figure 3A) corresponded to the glycosylated peptide Ser[95]−Lys[141] with a measured average mass at *m*/*z* 4982. The peak collected from the tryptic digest of the B form of fiber (Figure 3B) exhibited a signal 4779 corresponding to the nonglycosylated peptide Ser[95]−Lys[141]. This peptide contained 11

Table 1: Expected Monoisotopic Masses upon Cleavage of Ser^{95}−Lys^{141} Peptide by Endoproteinase GluC with and without the Presence of GlcNac^{a}

Endoproteinase GluC cleavage peptides	**Expected monoisotopic masses ([M+H]$^{+}$)**	
	NG	**+ GlcNac**
^{95}**SNINLE**^{100}	**689.34**	**892.42**
^{101}**ISAPLTVTSE**^{110}	**1017.54**	**1220.62**
^{111}**ALTVAAAAPLMVAGNTLTMQSQAPLTVHD**^{139}	**2892.49**	**3095.57**
^{111}**ALTVAAAAPLMVAGNTLTMQSQAPLTVHDSK**$^{141\ (*)}$	**3107.62**	**3310.70**

[a] NG = nonglycosylated. The asterisk indicates a missed cleavage at Asp139 was observed.

serine and threonine residues which could be potentially O-glycosylated.

To further localize the glycosylation site on the peptide Ser^{95}−Lys^{141}, we took advantage of the presence of glutamic acid residues to further cleave both the glycosylated and nonglycosylated forms with endoproteinase GluC. The glycosylated and nonglycosylated peptides purified as described above were incubated overnight in 100 mM sodium phosphate buffer, pH 7.8, at 32 °C. In this condition, endoproteinase GluC should cleave at the C-terminus of Glu and Asp. The digests were desalted, concentrated on a ZipTip (Millipore), and analyzed by MALDI-MS. The expected monoisotopic masses are listed in Table 1.

Digestion of nonglycosylated peptide Ser^{95}−Lys^{141} with GluC generated signals at *m*/*z* 689.30, 1017.26 (plus the corresponding Na^{+} and K^{+} adducts at *m*/*z* 1039.36 and 1055.28, respectively), and 3107.31 that corresponded to a missed cleavage at Asp^{139} (expected 3107.62) (Figure 4A). Signals at *m*/*z* 689.45 and 3107.86 were observed for the GluC digest of the corresponding O-glycosylated form (Figure 4B). No signal was observed either at *m*/*z* 1017.54 or at *m*/*z* 1220.62. However, signals at *m*/*z* 1242.55 and 1258.55 which might correspond to the Na^{+} and K^{+} adducts of *m*/*z* 1220.62 (expected *m*/*z* 1242.61 and 1258.58, respectively) were observed. These results strongly suggested that the peptide Ile^{101}−Glu^{110} contained the GlcNac group.

GluC treatment of the whole fiber was used to isolate peptide Ile^{101}−Glu^{110} in its glycosylated and nonglycosylated forms. For this, the A and B forms of fiber were purified separately by RP-HPLC and digested with GluC in 25 mM ammonium bicarbonate, pH 8.0, overnight at 32 °C. The peptides generated were separated by RP-HPLC on an Uptishere C18 column (Figure 5). The UV traces were identical except for those peaks indicated by the arrows at retention times 24 and 25 min. These peaks were collected, concentrated on ZipTip, and analyzed by MALDI-TOF MS for identification. The peak at RT = 25 min corresponded to the nonglycosylated peptide ^{101}ISAPLTVTSE^{110} with signals at *m*/*z* 1017.31 ([M + H]$^{+}$), 1039.36 ([M + Na]$^{+}$), and 1055.32 ([M + K]$^{+}$) (Figure 5B). The peak at RT = 24 min corresponded to the same peptide but glycosylated with signals at *m*/*z* 1242.60 ([M + Na]$^{+}$) and 1258.54 ([M + K]$^{+}$) (Figure 5A). A small signal at *m*/*z* 1017.48 was observed, resulting probably from a partial deglycosylation during analysis. The [M + H]$^{+}$ ion of the glycosylated peptide was not observed.

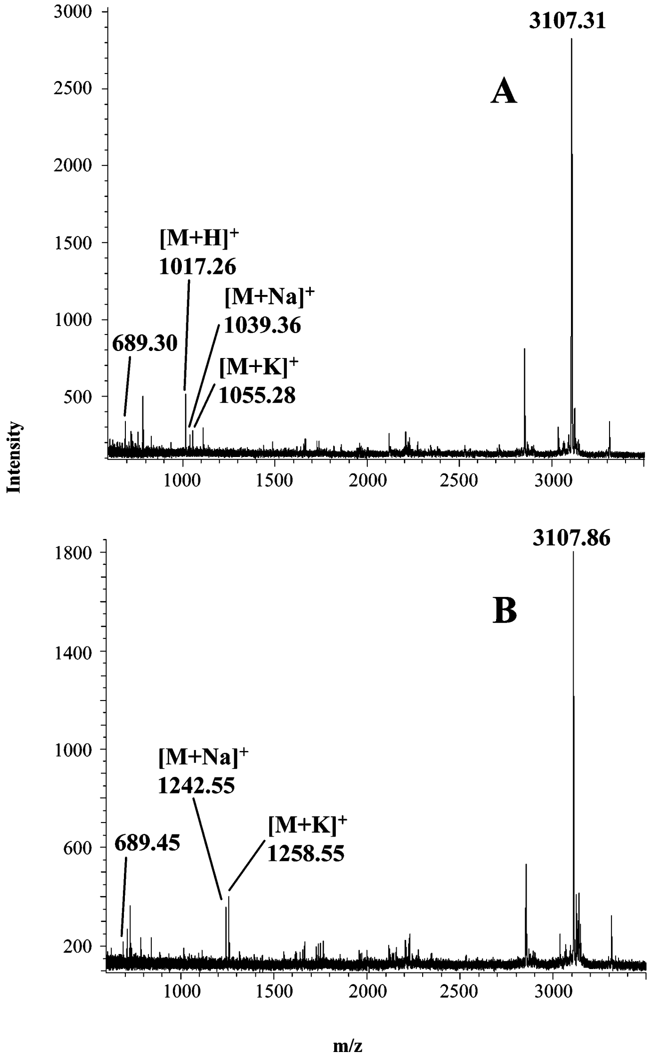


FIGURE 4: MALDI-TOF mass spectrum of nonglycosylated (A) and glycosylated (B) Ser^{95}−Lys^{141} peptide after digestion with endoproteinase GluC.

For further characterization, the A and B forms of fiber were no longer separated but collected together from RP-HPLC and digested by GluC, and the two peptides were isolated with a C18 column. Since the recovery of these peptides was very low with respect to the amounts of fiber treated, the conditions of digestion of fiber by GluC were modified. First, it was observed that digestion in phosphate buffer was at least two times more efficient than in

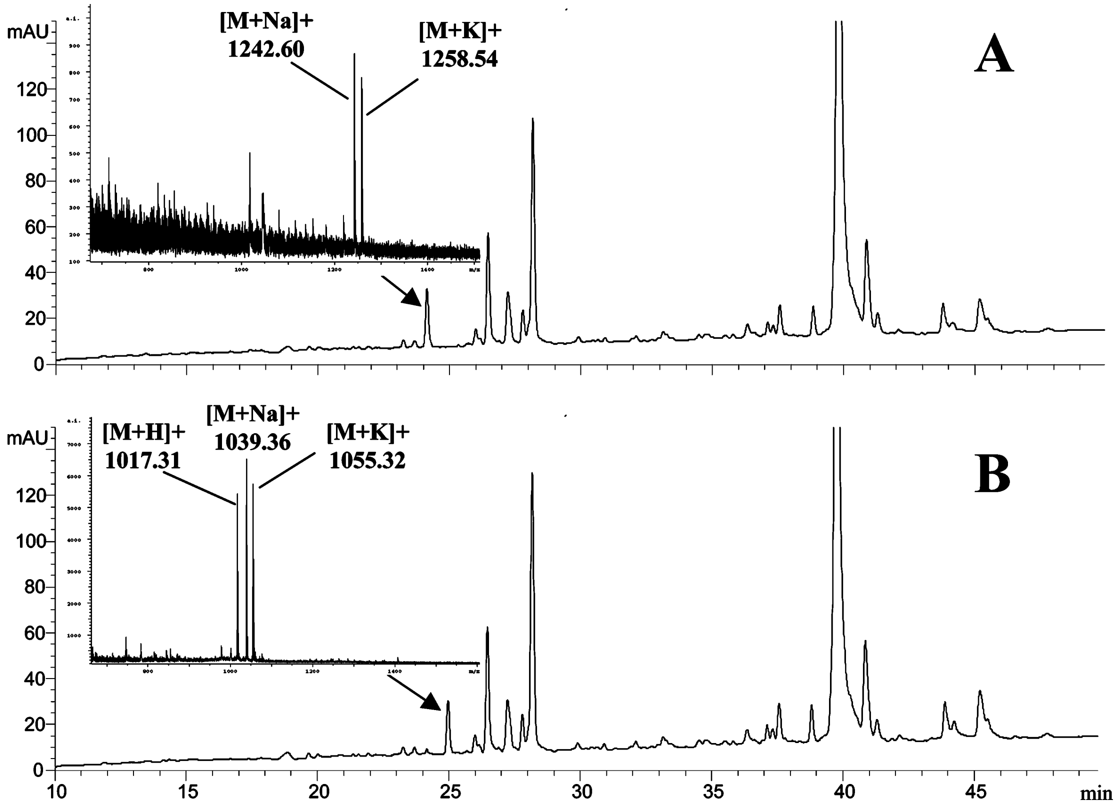


FIGURE 5: RP-HPLC analysis of the A form (A) and B form (B) of fiber digested with endoproteinase GluC. Inserts are MALDI-TOF mass spectra of peaks indicated by the arrows.

ammonium bicarbonate buffer. Second, the HPLC-purified fiber dried by centrifugal evaporation was poorly soluble when resuspended in GluC digestion buffer. Thus, dried HPLC-purified fiber was first solubilized in a small volume of a mixture of 25:75 acetonitrile/water containing 0.1% TFA. After a brief sonication, the solution was diluted with 100 mM phosphate buffer, pH 8.2, to a final concentration of 5% acetonitrile. The pH of the diluted solution was 8.0. Although some insoluble material appeared after dilution, the peptide recovery was improved by a factor of about 5 in these conditions.

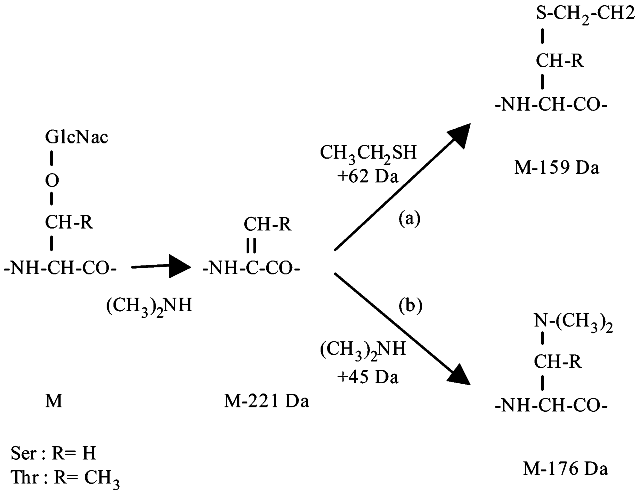
A direct identification of the glycosylation site by tandem mass spectrometry using a Q-TOF instrument was attempted. Glycosylated and nonglycosylated peptide were both submitted to fragmentation. MS/MS analysis of $[M + H]^+$ at *m*/*z* 1017 for nonglycosylated peptide generated a majority of b ions together with internal fragments that allowed a complete coverage of the sequence ISAPLTVTSE (not shown). Unfortunately, MS/MS analysis of the glycosylated peptide resulted in the release of the glycan moiety from the peptide generating an intense signal at *m*/*z* 204 (not shown). Thus, a chemical approach aiming to obtain more stable derivatives was attempted.

β-Elimination/Alkylation. Various chemical treatments based on a β-elimination strategy have been proposed (*12*−*14*). In a method developed by Czeszak and co-workers deglycosylation is performed by alkaline treatment with dimethylamine followed by subsequent tagging of the transient dehydro species with ethanethiol prior to tandem mass spectrometry (*15*). In this method, the identification of the glycosylation site is based on a mass increment of +44 Da of the originally glycosylated amino acid compared to the nonglycosylated amino acid.

Consequently, glycosylated and nonglycosylated Ile^{101}−Glu^{110} peptides isolated as described above were incubated with a mixture of 2-propanol/dimethylamine/ethanethiol for 5 h at 45 °C. After elimination of the reagents by centrifugal evaporation, samples were resuspended in aqueous acetonitrile containing 0.1% TFA and analyzed by MALDI-MS. The mass spectrum of the nonglycosylated peptide remained unchanged upon treatment, indicating that the reaction did not affect free Ser or Thr residues (not shown). From the mass spectrum of the glycosylated peptide, signals at *m*/*z* 1044.66, 1066.67, 1083.62, 1242.74, and 1258.74 were observed (not shown). The presence of signals at *m*/*z* 1242.74 and 1258.74, representing the $[M + Na]^+$and $[M + K]^+$ ions, indicated that the reaction was not complete. *m*/*z* 1083.62 might represent a $[M + Na]^+$ ion originating from a sodium adduct of the glycosylated peptide having reacted (i.e., 1242.61−221.09 for deglycosylation + 62.02 for thioethylation = 1083.54) according to Scheme 1a. Signals at *m*/*z* 1044 and 1066 could not be assigned according to the reaction described in Scheme 1a. However, as already described with other alkylamines (*12*, *14*), an addition of dimethylamine to the unsaturated amino acid intermediate instead of ethanethiol, according to Scheme 1b, might account for a signal at *m*/*z* 1044.

In this case, a mass difference of −176.03 Da (i.e., −221.09 for deglycosylation and +45.05 for addition of dimethylamine) should be observed. Accordingly, *m*/*z* 1066.67 might originate from the $[M + Na]^+$ ion at *m*/*z* 1242, *m*/*z* 1044.66 representing the corresponding $[M + H]^+$ ion. To test this hypothesis, ethanethiol was omitted from the reaction mixture, and the glycosylated peptide was incubated for 8 h at 45 °C with only dimethylamine and 2-propanol. In these conditions, only signals at *m*/*z* 1044 ($[M + H]^+$), 1066 ([M

Scheme 1: β-Elimination of GlcNac from Glycosylated Ser or Thr by Dimethylamine and Addition of Ethanethiol to the Dehydro Intermediate (*15*) (a) and Results Suggesting a Concomitant and Preferred Addition of Dimethylamine to the Dehydro Intermediate (This Work) (b)



+ Na]$^+$), and 1082 ([M + K]$^+$) were observed (Figure 6). *m*/*z* 1044, which was the most abundant ion produced, was further subjected to MS/MS analyses.

Identification of the Glycosylation Site. The molecular ions described above were subjected to MALDI-MS/MS analysis. First attempts on isolated [M + Na]$^+$ ions, i.e., *m*/*z* 1039, 1242, and 1083, resulted in noninformative fragmentation spectra (not shown). On the contrary, *m*/*z* 1044, which was expected to be a [M + H]$^+$ ion, showed predominantly fragmentation ions of the b and y series together with numerous internal fragments (Figure 7). According to the hypothesis mentioned above, *m*/*z* 1044 should result from an alkaline deglycosylation of the glycosylated peptide and subsequent addition of dimethylamine on the unsaturated intermediate formed. In such case, one would expect a mass increment of +27 Da of the formerly glycosylated amino acid with respect to the nonglycosylated amino acid. The ^{101}ISAPLTVTSE^{110} peptide contains four potential O-glycosylation sites. Fragment ions from b_2 through b_8 were found with values corresponding to unmodified amino acids that indicated that neither Ser^{102}, Thr^{106}, nor Thr^{108} was modified. By deduction, Ser^{109} should be the candidate. Unfortunately, fragment ion b_9 which corresponds to cleavage of peptide bond Ser^{109}−Glu^{110} was absent. However, all y fragment ions containing Ser^{109} (i.e., y_2 through y_9) were present and showed a mass increment of +27 Da identifying Ser^{109} as the O-glycosylation site.

Further confirmation of the location of the glycosylation site at Ser^{109} was achieved with nanoES-MS/MS experiments on the glycosylated peptide treated with dimethylamine using an Ion-Trap instrument. The [M + H]$^+$ ion at *m*/*z* 1044.5 subjected to an MS/MS analysis produced both b and y ions. Fragment ions from b_4 to b_8 were present and had expected mass values corresponding to unmodified amino acids. A specific mass increment of +27 Da was observed for b_9, confirming that Ser^{109} was formerly O-glycosylated. Fragment ions from y_3 to y_9 were observed; all contained Ser^{109} and had a mass increment of +27 Da (not shown).

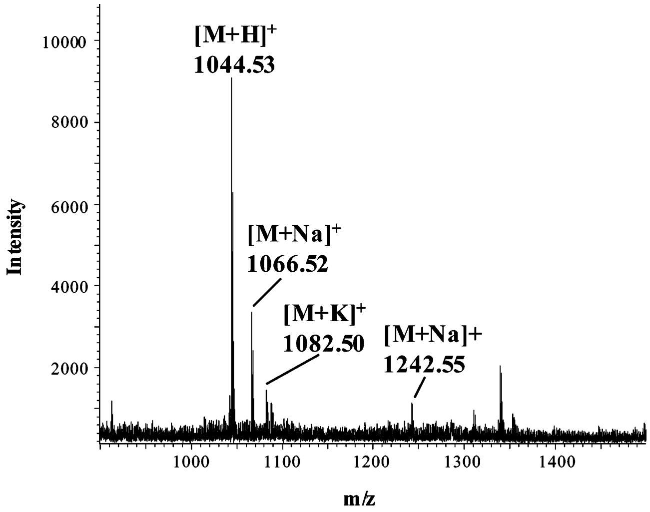


FIGURE 6: MALDI-TOF mass spectrum of O-glycosylated peptide ^{101}ISAPLTVTSE^{110} treated with dimethylamine for 8 h at 45 °C.

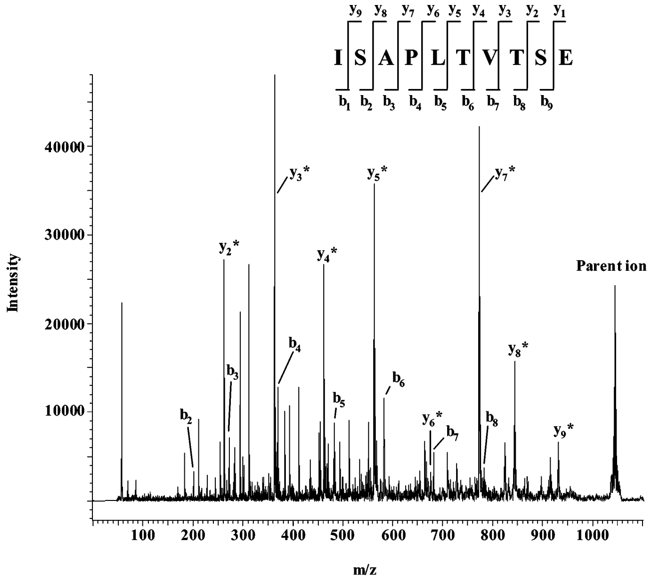


FIGURE 7: MALDI-TOF/TOF spectrum of O-glycosylated peptide ^{101}ISAPLTVTSE^{110} treated with dimethylamine. The asterisks indicate a mass increment of +27 Da. Nonlabeled peaks represent internal fragments.

The possibility that minor forms of the peptide in which glycosylation could have occurred on hydroxyamino acids other than Ser^{109} was investigated. Fragmentation spectra were examined for theoretical signals corresponding to fragments containing possibly deglycosylated/dimethylaminylated Ser^{102}, Thr^{106}, or Thr^{108}. Although a few signals, very low in intensity, that could correspond to discrete fragments in each spectrum, were observed, the remaining signals did not support the hypothesis of glycosylation at any other site than Ser^{109}.

Status of the Fiber in Adenovirus Capsid. Since the fiber that was studied here was purified from the pool of capsid proteins that was nonincorporated into virus particles, the modification status of fiber in the viral capsid had to be determined. To address this question, purified virions (AdTG6401) were injected on a C4 Jupiter column (150 × 2.1 mm) under the conditions described in the Material and Methods section for fiber analysis. Since the relative amount of fiber with respect to the other proteins of the viral capsid is low and since fiber coelutes partly with hexon, it was not possible to analyze the HPLC profile of fiber in whole virus

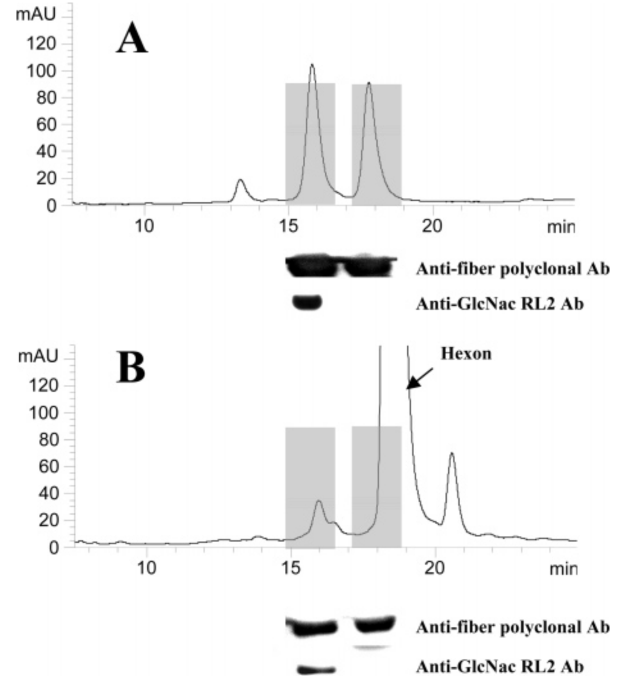


FIGURE 8: RP-HPLC analysis of purified fiber from the cellular pool of nonincorporated viral proteins (A) and whole virus AdTG6401 (B). Shaded areas indicate fractions collected and assayed by western blot for anti-fiber and anti-GlcNac antibodies.

from the UV trace. To overcome this difficulty, fractions corresponding to the retention time of the A and B forms of fiber on RP-HPLC were collected and analyzed by western blot using an anti-fiber antibody. As a control, the same experiment was performed with purified fiber. A and B forms of fiber were equally recognized by the antibody (Figure 8A). A similar profile was obtained from the whole virus (Figure 8B), indicating that the A and B forms of fiber were both present in viral particles. The same samples were assayed with a monoclonal antibody that recognizes O-linked GlcNac (*20*). Only the A form of purified fiber was recognized by the antibody, which indicated that the B form does not contain an O-linked GlcNac. The same observation was made for the whole virus. Thus the A form of purified fiber and fiber incorporated into virus contained most likely one GlcNac residue, whereas the B form had none.

DISCUSSION

In this study it has been shown that the adenovirus fiber protein purified from the pool of nonincorporated capsid proteins in infected cells exists as a mixture of O-glycosylated and non-O-glycosylated forms in roughly equal amounts. It has been shown that both forms of fiber were equally incorporated in the viral capsid since the same ratio of glycosylated and nonglycosylated forms is found in the purified viral particles.

The glycosylated fiber is likely a monoglycosylated form since only the A form, which differs from the B form by one GlcNac, is recognized by an antibody directed toward the GlcNac motif. Furthermore, whereas 82% of the fiber sequence was covered by MALDI-MS experiments, only one glycosylated peptide was found. Thus the relative number of GlcNac per polypeptide chain should be lower than 1. Previous studies reported an estimation of the number of GlcNac residues per polypeptide chain which varied from 3 to 4 and 1.7 to 2.5 for adenovirus type 2 and type 5 fibers, respectively (*8*). It would be interesting to check if these differences result from the cell line used to produce the viruses. Actually, the ratio of glycosylated to nonglycosylated fiber might be dependent on the activity of the galactosyl transferases which might vary from cell to cell.

MALDI-MS experiments on the A and B forms of fiber digested by trypsin resulted in the identification of one peptide sequence of 47 amino acids ($Ser^{95}-Lys^{141}$) whose measured mass was in agreement with the presence of a GlcNac residue. This peptide had 11 potential sites of O-glycosylation. However, its relatively high molecular mass (m/z 4982, average) was not suitable for an extensive MS/MS analysis. Digestion of the glycosylated and nonglycosylated forms of fiber with endoproteinase GluC followed by RP-HPLC separation allowed the identification of a peptide of 10 amino acids (^{101}ISAPLTVTSE^{110}) which contained the GlcNac residue. MALDI-MS analyses revealed that signals corresponding to almost exclusively cationized $[M + Na]^+$ and $[M + K]^+$ ions were observed.

Sites of attachment of the glycan moiety have usually been estimated by the presence of blank cycles following Edman degradation of O-glycosylated peptides (*21*−*23*). Recognition of GalNac-O-Ser/Thr-PTH derivatives during Edman degradation has been published (*24*). This requires, however, sequencing amounts of glycosylated peptides usually in the nanomolar range. Alternatively, tandem mass spectrometry has been proposed to elucidate O-glycosylation sites in proteins (*25*−*28*). The major drawback of this method is that the glycan moiety is labile and a substantial loss of O-linked sugar is observed during collision-induced or post-source decay fragmentation. The fragment ions retaining the glycan are usually of low abundance or absent, which makes the identification of the glycosylation sites difficult. To overcome this difficulty, several publications proposed a concomitant deglycosylation and tagging of the glycosylation site, resulting in derivatives more stable to analysis by MS/MS techniques (*12*−*15*).

Preliminary MS/MS analysis of the O-glycosylated form of this peptide resulted in a loss of the glycan moiety restoring the molecular ion corresponding to the nonglycosylated peptide and, thus, preventing any identification of the glycosylation site. Therefore, a strategy employing a chemical labeling of the glycosylation site after β-elimination was attempted. Deglycosylation by dimethylamine and concomitant labeling of the unsaturated intermediate by ethanethiol (*15*) resulted in the formation of the expected derivative which was actually a $[M + Na]^+$ ion. Together with this derivative, unexpected products resulting from the direct addition of dimethylamine to the unsaturated intermediate instead of ethanethiol were observed. The introduction of an alkylamine group in the peptide led to the production of an intense signal which corresponded to an $[M + H]^+$ ion. No informative fragmentation could be obtained from the thioethylated $[M + Na]^+$ ion subjected to MS/MS analysis. On the contrary, the fragmentation spectrum of the dimethylaminylated $[M + H]^+$ ion obtained with different techniques unambiguously locates the formerly O-linked GlcNac on Ser^{109}.

Mutation of serine 109 to alanine results in complete loss of fiber glycosylation in viral particles, confirming our identification of the glycosylation site (M. Lusky, unpublished observation). It is not known whether this partial fiber glycosylation plays a role in viral assembly or any other aspects of the viral life cycle. The functional analysis of the mutant is currently under evaluation and will help to decipher the role of fiber glycosylation.

Recombinant human adenoviruses are good candidates as gene therapy vectors. They are well tolerated and able to transfer therapeutic genes to a wide variety of cells (*29*). However, preexisting neutralizing antibodies to viral capsid proteins may reduce the efficacy of gene transfer. Interestingly, it was shown that a temperature-sensitive fiber mutant of type 5 adenovirus, which fails to glycosylate its fiber at the nonpermissive temperature, produces trimeric fibers with no immunoreactivity toward anti-fiber sera (*30*). Since fiber appears to have a key role in the production of neutralizing antibodies against the virus (*31*), a virus less sensitive to preexisting immunity might be of value for gene therapy.

ACKNOWLEDGMENT

We thank Catherine Ledoux and Laurence Laruelle for technical assistance.

REFERENCES

1. Bergelson, J. M., Cunningham, J. A., Droguett, G., Kurt-Jones, E. A., Krithivas, A, Hong, J. S., Horwitz, M. S., Crowell, R. L., and Finberg, R. W. (1997) Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5, *Science 275*, 1320−1323.
2. Hong, S. S., Karayan, L., Tournier, J., Curiel, D. T., and Boulanger, P. A. (1997) Adenovirus type 5 fiber knob binds to MHC class I alpha2 domain at the surface of human epithelial and B lymphoblastoid cells, *EMBO J. 16*, 2294−2306.
3. Green, N. M., Wrigley, N. G., Russel, W. C., Martin, S. R., and McLachlan A. (1983) Evidence for a repeating cross-beta sheet structure in the adenovirus fibre, *EMBO J. 8*, 1357−1365.
4. Van Oostrum, J., and Burnett, R. M. (1985) Molecular composition of the adenovirus type 2 virion, *J. Virol. 56*, 439−448.
5. Stouten, P., Sander, C., Ruigrok, R., and Cusack, S. (1992) New triple-helical model for the shaft of the adenovirus fibre, *J. Mol. Biol. 226*, 1073−1084.
6. Ishibashi, M., and Maizel, J. V. (1974) The polypeptides of adenovirus. VI. Early and late glycopolypeptides, *Virology 58*, 345−361.
7. Caillet-Boudin, M.-L., Strecker, G., and Michalski, J.-C. (1989) O-linked GlcNac in serotype-2 adenovirus fiber, *Eur. J. Biochem. 184*, 205−211.
8. Mullis, K. G., Haltiwanger, R. S., Hart, G. W., Marchase, R. B., and Engler, J. A. (1990) Relative accessibility of *N*-acetylglucosamine in trimers of the adenovirus types 2 and 5 fiber proteins, *J. Virol. 64*, 5317−5323.
9. Comer, F. I., and Hart, G. W. (2000) O-Glycosylation of nuclear and cytosolic proteins. Dynamic interplay between O-GlcNAc and O-phosphate, *J. Biol. Chem. 275*, 29179−29182.
10. Medzihradszky, K. F., Gillece-Castro, B. L., Settineri, C. A., Townsend, R. R., Masiarz, F. R., and Burlingame, A. L. (1990) Structure determination of O-linked glycopeptides by tandem mass spectrometry, *Biomed. Environ. Mass Spectrom. 19*, 777−781.
11. Settineri, C. A., Medzihradszky, K. F., Masiarz, F. R., Burlingame, A. L., Chu, C., and George-Nascimento, C. (1990) Characterization of O-glycosylation sites in recombinant B-chain of platelet-derived growth factor expressed in yeast using liquid secondary ion mass spectrometry, tandem mass spectrometry and Edman sequence analysis, *Biomed. Environ. Mass Spectrom. 19*, 665−676.
12. Mirgorodskaya, E., Hassan, H., Clausen, H., and Roepstorff, P. (2001) Mass spectrometric determination of O-glycosylation sites using β-elimination and partial acid hydrolysis, *Anal. Chem. 73*, 1263−1269.
13. Rademaker, G. J., Pergantis, S. A., Blok-Tip, L., Langridge, J. I., Kleen, A., and Thomas-Oates, J. E. (1998) Mass spectrometry determination of the sites of O-glycan attachment with low picomolar sensitivity, *Anal. Biochem. 257*, 149−160.
14. Hanisch, F.-G., Jovanovic, M., and Peter-Katalinic, J. (2001) Glycoprotein identification and localization of O-glycosylation sites by mass spectrometric analysis of deglycosylated/alkylaminylated peptide fragments, *Anal. Biochem. 290*, 47−59.
15. Czeszak, X., Ricart, G., Tetaert, D., Michalski, J. C., and Lemoine, J. (2002) Identification of substituted sites on MUC5AC mucin motif peptides after enzymatic O-glycosylation combining beta-elimination and fixed-charge derivatization, *Rapid Commun. Mass Spectrom. 16*, 27−34.
16. Boulanger, P., and Puvion, F. (1973) Large-scale preparation of soluble adenovirus hexon, penton and fiber antigens in highly purified form, *Eur. J. Biochem. 39*, 37−42.
17. Lusky, M., Christ, M., Rittner, K., Dieterle, A., Dreyer, D., Mourot, B., Schultz, H., Stoeckel, F., Pavirani, A., and Mehtali, M. (1998) In vitro and in vivo biology of recombinant adenovirus vectors with E1, E1/E2A, or E1/E4 deleted, *J. Virol. 72*, 2022−2032.
18. Neuhoff, V., Arold, N., Taube, D., and Ehrhardt, W. (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250, *Electrophoresis 9*, 255−262.
19. Novelli, A., and Boulanger, P. (1991) Assembly of adenovirus type 2 fiber synthesized in cell-free translation system, *J. Biol. Chem. 266*, 9299−9303.
20. Snow, C. M., Senior, A., and Gerace, L. (1987) Monoclonal antibodies identify a group of nuclear pore complex glycoproteins, *J. Cell Biol. 104*, 143−156.
21. Tomita, M., and Marchesi, V. T. (1975) Amino-acid sequence and oligosaccharide attachment sites of human erythrocyte glycophorin, *Proc. Natl. Acad. Sci. U.S.A. 72*, 2964−2968.
22. Rohrer, J. S., Cooper, G. A., and Townsend, R. R. (1993) Identification, quantification, and characterization of glycopeptides in reversed-phase HPLC separations of glycoprotein proteolytic digests, *Anal. Biochem. 212*, 7−16.
23. Stadie, T. R. E., Chai, W., Lawson, A. M., Byfield, P. G. H., and Hanisch, F. G. (1995) Studies on the order and site specificity of GalNAc transfer to MUC1 tandem repeats by UDP-GalNAc: polypeptide *N*-acetylgalactosaminyltransferase from milk or mammary carcinoma cells, *Eur. J. Biochem. 229*, 140−147.
24. Gerken, T. A., Owens, C. L., and Pasumarthy, M. (1997) Determination of the site-specific O-glycosylation pattern of the porcine submaxillary mucin tandem repeat glycopeptide. Model proposed for the polypeptide:galnac transferase peptide binding site, *J. Biol. Chem. 272*, 9709−9719.
25. Goletz, S., Thiede, B., Hanisch, F. G., Schultz, M., Peter-Katalinic, J., Müller S., Seitz O., and Kartsen, U. (1997) A sequencing strategy for the localization of O-glycosylation sites of MUC1 tandem repeats by PSD-MALDI mass spectrometry, *Glycobiology 7*, 881−896.
26. Müller, S., Goletz, S., Packer, N., Gooley, A., Lawson, A. M., and Hanisch, F. G. (1997) Localization of O-glycosylation sites on glycopeptide fragments from lactation-associated MUC1. All putative sites within the tandem repeat are glycosylation targets in vivo, *J. Biol. Chem. 272*, 24780−24793.
27. Alving, K., Paulsen, H., and Peter-Katalinic, J. (1999) Characterization of O-glycosylation sites in MUC2 glycopeptides by nanoelectrospray QTOF mass spectrometry, *J. Mass Spectrom. 34*, 395−407.
28. Hanisch, F.-G., Green, B. N., Bateman, R., and Peter-Katalinic, J. (1998) Localization of O-glycosylation sites of MUC1 tandem repeats by QTOF ESI mass spectrometry, *J. Mass Spectrom. 33*, 358−362.
29. Bramson, J., Graham, F., and Gauldie, J. (1995) The use of adenoviral vector for gene therapy and gene transfer *in vivo*, *Curr. Opin. Biotechnol. 6*, 590−595.
30. Cheng chee-Sheung, C., and Ginsberg, H. S. (1982) Characterization of a temperature-sensitive fiber mutant of type 5 adenovirus and effect of the mutation on virion assembly, *J. Virol. 42*, 932−950.
31. Gahéry-ségard, H., Farace, F., Godfrin, D., Gaston, J., Lengagne, R., Tursz, T., Boulanger, P., and Guillet, J.-G. (1998) Immune response to recombinant capsid proteins of adenovirus in humans: antifiber and anti-penton base antibodies have a synergistic effect on neutralizing activity, *J. Virol. 72*, 2388−2397.

BI047702B