Published in final edited form as:

J Mass Spectrom. 2011 October ; 46(10): 993–1003. doi:10.1002/jms.1980.

Determination of the glycation sites of *Bacillus anthracis* neoglycoconjugate vaccine by MALDI-TOF/TOF-CID-MS/MS and LC-ESI-QqTOF-tandem mass spectrometry

Farid Jahouh^a, Shu-jie Hou^b, Pavol Kováč^b, and Joseph H. Banoub^{a,c,*}

^aDepartment of Chemistry, Memorial University of Newfoundland, Saint John's, NL, Canada

^bNIHDK, LBC, National Institutes of Health, Bethesda, MD 20892-0815, USA

^cScience Branch, Special Projects, Department of Fisheries and Oceans Canada, Saint John's, NL A1C 5X1, Canada

Abstract

We present herein an efficient mass spectrometric method for the localization of the glycation sites of a model neoglycoconjugate vaccine formed by a construct of the tetrasaccharide side chain of the *Bacillus anthracis* exosporium and the protein carrier bovine serum albumin. The glycoconjugate was digested with both trypsin and GluC V8 endoproteinases, and the digests were then analyzed by MALDI-TOF/TOF-CID-MS/MS and nano-LC-ESI-QqTOF-CID-MS/MS. The sequences of the unknown peptides analyzed by MALDI-TOF/TOF-CID-MS/MS, following digestion with the GluC V8 endoproteinase, allowed us to recognize three glycopeptides whose glycation occupancies were, respectively, on Lys 235, Lys 420, and Lys 498. Similarly, the same analysis was performed on the tryptic digests, which permitted us to recognize two glycation sites on Lys 100 and Lys 374. In addition, we have also used LC-ESI-QqTOF-CID-MS/MS analysis for the identification of the tryptic digests. However, this analysis identified a higher number of glycopeptides than would be expected from a glycoconjugate composed of a carbohydrate–protein ratio of 5.4:1, which would have resulted in glycation occupancies of 18 specific sites. This discrepancy was due to the large number of glycoforms formed during the synthetic carbohydrate–spacer–carrier protein conjugation. Likewise, the LC-ESI-QqTOF-MS/MS analysis of the GluC V8 digest also identified 17 different glycation sites on the synthetic glycoconjugate.

Keywords

neoglycoconjugate; proteinases; glycated peptides; MALDI-TOF/TOF-CID-MS/MS; LC-ESI-QqTOF-CID-MS/MS

INTRODUCTION

Bacterial lipopolysaccharides have been extensively studied during the last decades, and their use in the synthesis of neoglycoconjugates as potential vaccines has been successful.^[1–7] Different parameters have been found to influence the efficacy of a

Copyright © 2011 John Wiley & Sons, Ltd.

*Correspondence to: Joseph H. Banoub, Department of Fisheries and Oceans Canada, Science Branch, Special Projects, St. John's, NL A1C 5X1, Canada. banoubjo@dfo-mpo.gc.ca.

Supporting Information

Supporting information can be found in the online version of this article.

glycoconjugate vaccine, such as the saccharide size, the average number of saccharide chains per conjugate molecule, the nature of the carrier, and the distance between the saccharide and the protein in the formed glycoconjugate.^[8–11] Until very recently, the covalent attachment sites of the antigenic saccharide to the protein in all neoglycoconjugate have not been determined. We were the first to attempt such an investigation and have unraveled some of the glycation sites in the neoglycoconjugates made from bovine serum albumin (BSA) and the O-specific polysaccharide (O-PS) of *Vibrio cholerae* O1, serotype Ogawa.^[12]

Bacillus anthracis belongs to the Gram-positive bacterium family and is the cause of anthrax, a disease which can affect animals and humans.^[13] The bacterium is known to form endospores,^[14] which when mature are resistant to extreme conditions (temperature, radiations, desiccation, physical damage, harsh chemicals).^[15] The currently increased interest in anthrax is because some form of *B. anthracis*, which is the etiologic agent of anthrax,^[16] including anthrax spores, can be used as a biological weapon. Several studies have been carried out on the *B. anthracis* to understand its pathogenesis, to discover new markers for its detection, and also to identify constituents of its spore that could be targeted for its inactivation.^[17]

Different synthetic vaccines have been developed against anthrax; these were based on the immunogenic conjugates of the capsular polypeptide (polyglutamic acid) of *B. anthracis*, and anthrax capsular vaccine has also been developed.^[18]

A better understanding of the anthrax spores composition and structure was necessary before a vaccine for anthrax could be developed on the basis of a component of anthrax spores. Recently, the structure of the tetrasaccharide side chain of the collagen like region of the major glycoprotein of the *B. anthracis* exosporium was determined by Daubenspeck *et al.*^[19] The upstream terminal of the tetrasaccharide was found to be the new sugar anthrose [4,6-dideoxy-4-(3-hydroxy-3-methylbutyramido)-2-*O*-methyl- β -glucopyranose]. By comparing the NMR data of the synthetic tetrasaccharide^[20] with those recorded^[19] for the tetrasaccharide isolated from *B. anthracis* exosporium, it was confirmed that the structure of the tetrasaccharide was correct as proposed.^[19] Kovac *et al.* synthesized the tetrasaccharide side chain and linked it to the BSA using squaric acid chemistry.^[21] They obtained synthetic glycoconjugates having an average carbohydrate–protein ratio of 3.5:1 and 5.7:1. Figure 1 displays a schematic representation of the general structure of the carbohydrate–protein neoglycoconjugate construct from the tetrasaccharide side chain of the *B. anthracis* exosporium.

The analysis and the localization of the glycation sites in carbohydrate–protein neoglycoconjugates are very important features for the quality control of the conjugate vaccines. Different strategies have been developed for the analysis of glycoproteins, mainly based on the mass spectrometry analysis of enzymatic digests.^[22,23] In a previous study on the determination of the glycation sites in synthetic neoglycoconjugate models of *V. cholerae* O1, we have shown that out of the 59 possible glycation sites on BSA, we were only able to characterize three, when using a standard method that consists on the trypsin digestion of the glycoconjugate. We hypothesized that this protease may not be the best to release all the glycosylated peptides from this neoglycoconjugate.^[12] In addition, the unraveling of some of the glycation sites in this neoglycoconjugate was achieved by performing tandem mass spectrometry analysis using a MALDI-TOF/TOF-MS/MS instrument, which benefit of high-energy collision dissociation.^[12]

In this manuscript, we present an efficient method for the localization of the glycation sites of the neoglycoconjugate formed by the antigenic tetrasaccharide side chain of the *B.*

anthracis exosporium covalently attached to the BSA protein carrier. The digestion of the glycoconjugate was carried out with two different proteases: trypsin and GluC V8. The analysis and sequencing of the resulting peptides and glycopeptides was performed separately by using two different state-of-the-art mass spectrometric and tandem mass spectrometric methods.

MATERIALS AND METHODS

Preparation of the hapten–BSA neoglycoconjugate

The synthesis of the hapten–BSA glycoconjugate has been described previously.^[21] Thus, the synthetic 2-*O*-methyl- β -D-glucopyranosyl-(1 $\rightarrow$ 3)- α -L-rhamnopyranosyl-(1 $\rightarrow$ 3)- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- α -L-rhamnopyranoside corresponding to the *B. anthracis* exosporium tetrasaccharide was first treated with ethylenediamine to give the 2-aminoethyl amide, which when treated with diethyl squarate afforded the squaric acid monoethylester. The conjugation of this hapten to BSA (Sigma Aldrich, St Louis, MI, USA) in 0.5 M borate buffer (pH 9.0) was followed by SELDI-TOF-MS,^[27] which made it possible to obtain the hapten–BSA glycoconjugates, with a predetermined carbohydrate–protein ratio of ~5:1. The isolation of the neoglycoconjugate was carried out by ultrafiltration with the aid of a centrifugal device with a cutoff 30,000 Da.

Digestion

The digestions of the hapten–BSA glycoconjugates were carried out with trypsin and GluC V8 protease (Sigma Aldrich). Thus, 100 μ g of the glycoconjugate (1.4 nmol) was dissolved in a mixture of 0.1% RapiGest SF Surfactant (1 μ g, Waters, USA) in 50 mM of NH_4HCO_3 (100 μ l) at a pH of 8.0 and reduced by treatment with 2 μ l of 10 mM dithiothreitol (Sigma Aldrich) for 30 min at room temperature, followed by alkylation with 2 μ l of a 50-mM iodoacetamide (Sigma Aldrich) for 1 h at room temperature. A portion (50 μ g, 0.7 nmol) of the BSA glycoconjugate was digested with trypsin using a 20 ng/ml of trypsin dissolved in NH_4HCO_3 (50 mM, 1 ml) at a trypsin–glycoprotein ratio of 1:25 (w/w) and incubated at 37 °C overnight with shaking. The other 50 μ g (0.7 nmol) of the BSA glycoconjugate was digested using the GluC V8 endoprotease dissolved in NH_4HCO_3 (50 mM, 1 ml) at a protease–glycoprotein ratio of 1:25 (w/w) and incubated at 37 °C overnight with shaking. The sample was then dried under vacuum, and the residue was dissolved in 20 ml of 1% acetic acid (Sigma-Aldrich, Oakville, ON, Canada). An aliquot of each sample (10 μ l) was then cleaned up using ZipTip C18 (Millipore, Bedford, MA, USA) before mass spectral analysis.

MALDI-TOF/TOF-MS and MALDI-TOF/TOF-CID-MS/MS analyses

Mass spectral analyses were carried out with a 4700 Proteomics analyzer equipped with TOF-TOF optics (Applied Biosystems Foster City, CA) and a 200-Hz frequency-tripled Nd:YAG laser. α -Cyano-4-hydroxycinnamic acid (α -CHCA) was used as matrix for the analysis of BSA, hapten–BSA conjugate, and the peptides resulting from their digestions, with an average of 5000 to 8000 laser shots per spectra. Basically, 1 μ l of a 20-mg/ml solution of α -CHCA (dissolved in acetone, 0.1% trifluoroacetic acid) was spotted on the MALDI plate and dried at room temperature. Then, an aliquot of 1 μ l of sample was spotted on the top of the dried matrix and let to dry before the MALDI-MS experiments.

The tandem mass spectrometric analyses were achieved with air as collision gas and at collision energy of 1 kV, which correspond to the difference between the accelerating potential (8 kV) and the floating cell (7 kV).^[28] The internal calibration of the digested glycoproteins and BSA was performed in the range m/z 600–3200 using known peptides of the BSA.

LC-ESI-QqTOF-CID-MS/MS analysis

The peptides were separated on a DIONEX UltiMate3000 Nano LC System (Germering, Germany). Two hundred fifty femtomoles of digested glycoprotein was dissolved in 0.1% TFA and loaded onto a precolumn (300 μm i.d. $\times$ 5 mm, C18 PepMap100, 5 μm (LC Packing, Sunnyvale, CA)) to desalt and concentrate the sample. After their elution from the precolumn, the mixture of peptides and glycopeptides was separated on a nanoflow analytical column (75 μm i.d. $\times$ 15 cm, C18 PepMap 100, 3 μm , 100 Å; LC Packing, Sunnyvale, CA) at a flow rate of 180 nL/min. The mobile phase eluents used were composed of 0.1% FA/0.01% TFA/2% ACN (A) and 0.08% FA/0.008% TFA/98% ACN (B). The elution started with 0% B for 10 min, followed by a gradient of 0%–60% B in 55 min and 60%–90% B in 3 min and was kept at 90% B for 3 min. The tandem mass spectrometry analysis of the eluted peptides and glycopeptides was accomplished using an Applied Biosystems API-QSTAR XL quadrupole orthogonal time-of-flight (QqTOF)-MS/MS hybrid tandem mass spectrometer (Applied Biosystems International-MDS Sciex, Foster City, CA, USA) equipped with a nano-electrospray source (Protana XYZ manipulator), which produces the electrospray through a PicoTip needle (10 μm i.d.; New Objectives, Woburn, MA, USA) carrying a voltage of 2400 V.

RESULTS

As mentioned earlier, the aim of this study was to localize the glycation sites for the neoglycoconjugate composed of the antigenic tetrasaccharide side chain of *B. anthracis* exosporium covalently attached to the BSA protein carrier. In this rationale, we have aimed to search the best method for the characterization of the glycopeptides. The digestion of the glycoconjugate was carried out with two different proteases: trypsin and GluC V8. Although trypsin cleaves proteins at the carboxyl side of lysine and/or arginine residues,^[24] GluC V8 endoproteinase is known to cleave protein at the C-terminal of the peptide bond on glutamic acid residues at pH 4.0 and on both glutamic acid and aspartic acid residues at pH 7.8.^[25,26]

Furthermore, for the analysis and sequencing of the resulting peptides and glycopeptides, we have used two different mass spectrometric approaches. The first one involved the analysis of the digests by MALDI-TOF/TOF-MS (MALDI-MS), followed by peptides sequencing using high collision energy MALDI-TOF/TOF-CID-MS/MS (MALDI-MS/MS) mass spectrometry. The second approach involved analysis of the mixed peptide digests by LC-ESI-QqTOF-MS (LC-MS) followed by peptides sequencing using low-energy collision-induced dissociation CID-QqTOF-MS/MS analysis (LC-MS/MS). A scheme, indicating the general strategy used in this manuscript, is illustrated in Fig. 2.

MALDI-TOF/TOF-MS analysis of BSA and the hapten–BSA neoglycoconjugates

A single-stage MALDI-MS analysis of the BSA and the carbohydrate hapten–BSA glycoconjugate allowed us to determine the average number of carbohydrate–spacer covalently attached to the BSA by comparison of the observed molecular weight of the glycoconjugate to the molecular weight of the native BSA. Thus, the MALDI-MS analysis of the hapten–BSA glycoconjugate allowed us to identify the protonated molecule $[M + H]^+$ at m/z 71,448.36 and $[M + 2H]^{2+}$ at m/z 35,730.08. Using these data, we calculated the molecular weight of the hapten–BSA glycoconjugate, which was 71,448.36 Da. Finally, the comparison of the observed molecular weight of the BSA and the hapten–BSA enabled us to calculate the average hapten–BSA ratio on the glycoconjugate. As previously reported, during the conjugation, the squaric acid is covalently linked to the 2-amino groups of the lysine residues in BSA with loss of a molecule of ethanol.^[9,12,21] Knowing the molecular weight of the tetrasaccharide-squaric acid unit (950 Da), we were able to calculate the carbohydrate hapten–BSA ratio, which was found to be 5.4:1. The glycated peptides on the

BSA were identified by the peptide mass fingerprinting (PMF) experiment on the hapten–BSA conjugate digests and the CID-MS/MS analysis of the glycated peptides permitted to localize the glycation site.

PMF analysis using MALDI-TOF/TOF-MS of the digested hapten–BSA neoglycoconjugates

MALDI-MS analysis was carried out on the digested (trypsin and GluC V8) hapten–BSA glycoconjugate, and the obtained MS spectra were submitted to the MASCOT library to identify by PMF the peptides matching to the BSA (Figs 3a and 3b and Tables S1 and S2). The following parameters were applied for the MASCOT library searching: carbamidomethyl (C) as fixed modifications, oxidation (M) as variable modifications, a peptide mass tolerance of ± 0.2 Da, and a maximum of missed cleavages of 1.

The MASCOT report of the tryptic and GluC V8 digests spectra gave a match to two isoforms of the serum albumin protein from the *Bos taurus* species, namely, the serum albumin precursor (gi|1351907) and the serum albumin (gi|74267962). Thus, we postulated that the obtained peptides that were not a match with the peptides of the BSA had to be conjugated with the hapten carbohydrate. Consequently, these were analyzed by MALDI-MS/MS analysis.

The MALDI-MS analysis of the carbohydrate hapten–BSA glycoconjugate digested with trypsin (Fig. 3a, Table S1) afforded only three identifiable glycopeptides having different glycation sites. These diagnostic three glycopeptides corresponded to the BSA peptides whose molecular masses have increased by 950 Da were identified. These were verified identified to have the following structures: ALK*AWSVAR at m/z 1951.0130, VTK*CTESLVNR at m/z 2416.1372, and QNCDQFEK*LGEYGFQNALIVR at m/z 3478.6429 (glycation represented with an asterisk on the lysine residue).

The MALDI-MS analysis of the glycopeptides resulting from the digestion of the hapten–BSA glycoconjugate with the GluC V8 protease (Fig. 3b, Table S2) afforded only three glycopeptides: LCK*VASLRE at m/z 2025.0186, YAVSVLLRLAK*E at m/z 2311.2476, and YAVSVLLRLAK*EYE at m/z 2603.3484.

High-energy MALDI-TOF/TOF-CID-MS/MS analyses of the glycated peptides obtained after trypsin digestion of the carbohydrate hapten–BSA glycoconjugate

The sequence and the determination of the glycation sites of the trypsin obtained glycopeptides were verified by tandem mass spectrometry. Thus, we carefully analyzed by MALDI-MS/MS analyses the precursor protonated molecules $[M + H]^+$ extracted from the following glycopeptides: ALK*AWSVAR (Lys 235) at m/z 1951.0130 presented in Fig. 4a (Table 1); VTK*CTESLVNR at m/z 2416.1372 (Fig. S1), and QNCDQFEK*LGEYGFQNALIVR at m/z 3478.6429 (Fig. S2).

During the high-energy CID-MS/MS analysis of the glycopeptide ALK*AWSVAR (Lys 235) at m/z 1951.0130 (Fig. 4a, Table 1), we observed a series of consecutive losses because of the various carbohydrate fragments. These losses seem to indicate that during CID-MS/MS analyses, the glycosyl hapten gas-phase fragmented, and this fact permits the unambiguous identification and localization of the glycation site.

For the nomenclature of the glycopeptide product ions obtained, we have used the nomenclature established by Domon and Costello^[29] for the carbohydrate portion as A, B, C, X, Y, and Z. Thus, we observed that the product ions generated by the fragmentation of the carbohydrate hapten, occurred by the successive losses of either each of the monosaccharide B₁ (–259 Da), disaccharide B₂ (–405 Da), trisaccharide B₃ (–551 Da), and/or the tetrasaccharide B₄ (–697 Da), leading, respectively, to the formation of the following

product ions: Y_3^+ at m/z 1691.8829, Y_2^+ at m/z 1545.8376, Y_1^+ at m/z 1399.7768, and Y_0^+ at m/z 1253.7010. We also noted that the difference between Y_3^+ and Y_2^+ , Y_2^+ and Y_1^+ , Y_1^+ and Y_0^+ was 146 Da (Fig. 4a, Table 1), and this difference corresponded to one α -L-rhamnopyranosyl unit. We therefore propose in this study, that the loss of a glycosyl unit represents a diagnostic fragmentation signature of the synthetic tetrasaccharide. This attribute is extremely important for the confirmation of the identity of the carbohydrate hapten covalently linked to the peptide. In addition, different carbohydrate product ions resulting from other fragmentation routes were also identified, and these were assigned as follows: $C_{17}H_{26}N_3O_4^+$ at m/z 336.1242, B_1^+ at m/z 260.0739, $[B_1-H_2O]^+$ at m/z 242.0709, $^{2,5}A_1^+$ at m/z 230.0719, $C_{11}H_{16}N_3O_2^+$ at m/z 222.0670, and $[C_1-C_5H_{11}NO_2]^+$ at m/z 159.0486 (Fig. 4b).

To minimize the ambiguity in the nomenclature used in this study, we have also used the preexisting peptides nomenclature involving the detection of the x-, y-, and z-ions and the a-, b-, and c-ions for the individual peptide sequence.^[30,31] In consequence, it is important to point out that the b- and y-product ions obtained, included the glycosylated lysine, which have lost part of the carbohydrate moiety. However, it is important to comprehend that these peptides still preserved the portion of the attached spacer-squaric acid chain. The following are examples of the b- and y-product ions that we have observed: $[b_3-B_4]^+$ at m/z 565.2661, $[b_4-B_4]^+$ at m/z 636.2715, and $[y_7-B_2]^+$ at m/z 1361.7265.

The high-energy CID-MS/MS analysis of the $[M + H]^+$ precursor ion at m/z 2416.1372 of the glycopeptide VTK*CTESLVNR (Fig. S1) permitted us to pinpoint the tetrasaccharide hapten location site on the Lys 498. This CID-MS/MS analysis also showed the consecutive losses of the glycosyl portions of the tetrasaccharide B_{1-4} to afford the product ions Y_{0-3} and other oligosaccharide product ions (Fig. S1). Finally, we were able to identify the following glycopeptide as QNCDQFEK*LGEYGFQNALIVR at m/z 3478.6429 (Fig. S2), which contained the carbohydrate hapten covalently attached to the Lys 420 residue.

It is noteworthy to mention that during high-energy CID-MS/MS analysis, we were subjected to a small amount of flawed results. For instance, what we presumed to be a glycosylated peptide at m/z 2338.1690 turned unfortunately to be the intact entire nonglycosylated peptide [MPCTEDYLSLILNR] product ion. This discrepancy was most probably because during MALDI-MS/MS analysis the instrument ion gate selection of the precursor ion allowed also the transmission of nonglycosylated peptide, which was isobaric to the selected precursor ion at m/z 2338.1690. In addition, it has been recently reported by Sato *et al.* that the ions produced in the MALDI source cause post source decay^[32,33] and that they may overlap the product ions resulting from the high-energy CID analysis of selected precursor ions, which complicate the product ions spectra.^[34]

High-energy MALDI-TOF/TOF-CID-MS/MS analyses of the glycosylated peptides obtained after the GluC V8 digestion of the carbohydrate hapten-BSA conjugate

The digested GluC V8 endoproteinase glycosylated peptides were detected by single-stage MALDI-MS analysis and analyzed by MALDI-MS/MS. The high-energy CID-MS/MS analysis of the precursor ion $[M + H]^+$ at m/z 2025.0186 extracted from the glycopeptide LCK*VASLRE (Fig. S3, Table S3) allowed us to confirm the sequence of the uncleaved complete glycopeptide ion and to localize the carbohydrate occupancy on the covalently attached Lys 100 residue. Once more, the precursor ion protonated glycopeptide at m/z 2025.0186 afforded various peptide product ions containing the full intact peptide portion, which have lost sections of the intact tetrasaccharide hapten. These product ions were identified as Y_3^+ at m/z 1765.8924, Y_2^+ at m/z 1619.8220, Y_1^+ at m/z 1473.7520, Y_0^+ at m/z 1327.7091, and $[Y_0-NH_3]^+$ at m/z 1310.6732. The sequence coverage of this glycopeptide is shown in Fig. S3, and as expected, some peptide product ions were observed with the loss of

the entire tetrasaccharide B₄ moiety: [b₆-B₄]⁺ at *m/z* 911.4274, [b₅-B₄]⁺ at *m/z* 824.3805, [b₄-B₄]⁺ at *m/z* 753.3832, [b₃-B₄]⁺ at *m/z* 654.3269, and [K*V-B₄]⁺ at *m/z* 452.3055. In addition, the following diagnostic carbohydrate-spacer product ions created by fragmentation of the tetrasaccharide spacer portion were also detected: C₁₇H₂₆N₃O₄⁺ at *m/z* 336.2023, B₁⁺ at *m/z* 260.1505, [B₁-H₂O]⁺ at *m/z* 242.1599, C₁₁H₁₆N₃O₂⁺ at *m/z* 222.1627, [C₁-C₃H₇O]⁺ at *m/z* 202.1093, and C₇H₁₃O₂⁺ at *m/z* 129.1094.

High-energy CID-MS/MS analysis of the precursor ions [M + H]⁺ at *m/z* 2311.2476 extracted from the glycopeptide YAVSVLLRLAK*E (Fig. S4) and at *m/z* 2603.3484 for the glycopeptide YAVSVLLRLAK*EYE (Fig. S5) revealed that the second glycation site was located on the Lys 374 residue. To sum up, the MALDI-MS/MS analysis of the digested carbohydrate-BSA conjugate with trypsin and GluC V8 allowed us to identify five glycation sites: Lys 100, Lys 235, Lys 374, Lys 420, and Lys 498. It is crucial to point out that the detection of carbohydrate product ions are diagnostic of the presence of the glycated peptide and also allow the validation of the location of the covalently attached carbohydrate hapten sites. However, we report herein a conjecture on the uses of the single-stage MALDI-MS analyses, which in our opinion can be attributed to the low sensitivity of detection of the mixture of glycopeptides, and accordingly some glycated peptides were not detected at all. Moreover, the low precursor ion selectivity of the linear TOF in addition to the overlapping of the PSD and CID fragment ions can complicate the interpretation of the product ions spectrum.^[34,35]

LC-ESI-QqTOF-CID-MS/MS analysis of the digested carbohydrate-BSA conjugate

In this rationale, as mentioned before, we have used a second approach, involving the analysis of the mixed peptide digests by LC-ESI-QqTOF-MS (LC-MS) followed by peptides sequencing using low-energy collision-induced dissociation CID-QqTOF-MS/MS analysis (LC-MS/MS). Therefore, as expected, we have found that the Nano-LC-MS/MS analysis of the digested hapten-BSA glycoconjugate, presented several advantages to the MALDI-MS/MS analysis of the glycated peptides. This was mainly due to the chromatographic separation, which restrains the ionization suppression effect present during the MALDI-MS analysis.^[36,37] Moreover, during the tandem mass spectrometry, the CID-MS/MS analyses are low-energy compared with the high-energy CID-MS/MS experiments realized with the TOF/TOF-MS/MS instrument. For that reason, it is well known that the low-energy CID-MS/MS analysis of peptides generally produces unique peptide diagnostic b- and y-ions, which provide a bidirectional analysis of their sequence.^[38]

Nano LC-MS/MS was carried out on both tryptic digests and GluC V8 digests. The results were submitted to the Mascot library to identify the peptides belonging to the BSA. For the analysis of the tryptic digests, the Mascot library gave a match with two serum albumin protein isoforms from the *B. taurus* species: serum albumin precursor (gi|1351907) and serum albumin (gi|74267962).

CID-MS/MS of the tryptic digested peptides

The detected peptides that matched to these two isoforms are reported in Table S4. The detected peptides covered 58% of the sequence of the serum albumin protein from *B. taurus* (gi|74267962), and in addition four unique peptides for this protein were identified [Table S4]: TVMENFVAFVGK at *m/z* 671.3491 (+2), ETYGDMADECCAK at *m/z* 710.7665 (+2), LGEYGFQNELIVR at *m/z* 769.4033 (+2), and MPCAEDYLSLILNR at *m/z* 847.9184 (+2). Similarly, the digested peptides covered 57% of the sequence of the serum albumin precursor from *B. taurus* (gi|1351907), and one unique peptide was detected for this protein: ETYGDMADECCKEQEPER at *m/z* 706.6155 (+3). All the peptides that could not be assigned to any protein hits were presumed to be potentially glycated peptides.

The low-energy CID-MS/MS analysis of the extracted precursor ions extracted from the tetrasaccharide-spacer-BSA neoglycoconjugate and the corresponding glycopeptide product ions identified are shown in Table 2. Thus, the LC-MS/MS analyses permitted us to identify the carbohydrate occupancy located on the following 18 lysine glycation sites indicated as follows: Lys 140, Lys 155, Lys 156, Lys 204, Lys 211, Lys 228, Lys 235, Lys 304, Lys 374, Lys 401, Lys 420, Lys 437, Lys 455, Lys 463, Lys 495, Lys 498, Lys 547, and Lys 559. In this justification and for simplification purposes, we only decided to describe four examples of the low-energy CID-MS/MS analyses of the glycopeptides (spectra and tables for the other glycated peptides are available upon request). Figures 5a (Table S5) and 5b exhibit two examples of the gas-phase fragmentation resulting from the CID-MS/MS analysis of the glycopeptides as well as the product ion calculated masses and the mass deviation of the experimental masses compared with the calculated ones, and Figs 5a and 5b display the LC-MS/MS spectra of the glycopeptides. In these low-energy CID-analyses, we observed the consecutive losses of carbohydrate moieties from the intact glycopeptides as well as the formation of peptide product ions characterized by either the loss of the entire carbohydrate moiety and/or the partially fragmented carbohydrate product ions.

For example, the CID-MS/MS analysis of the glycopeptides precursor ion $[M + 2H]^{2+}$ at m/z 1058.4652 extracted from CCTK*PESER (Lys 463) (Fig. 5a, Table S5), showed the following peptide ions, which were assigned to be formed by the loss of the carbohydrate moieties: Y_3^+ at m/z 1856.7628, Y_2^+ at m/z 1710.7332, Y_1^+ at m/z 1564.6742, and Y_0^+ at m/z 1418.6321. As observed for the previous MALDI-MS/MS analysis of the glycated peptides, we were able to identify peptide fragment ions presenting a loss of a carbohydrate moiety: $[y_7-B_4]^+$ at m/z 1098.5412, $[y_6-B_4]^+$ at m/z 997.4909, and $[b_4-B_4]^+$ at m/z 802.3246. The following carbohydrate-spacer product ions were also detected: B_2^+ at m/z 406.2071, $C_{17}H_{26}N_3O_4^+$ at m/z 336.1908, C_1^+ at m/z 276.1435, B_1^+ at m/z 260.1460, $[B_1-H_2]^+$ at m/z 258.1313, $[B_1-H_2O]^+$ at m/z 242.1354, $C_{11}H_{16}N_3O_2^+$ at m/z 222.1184, $[C_1-C_3H_7O]^+$ at m/z 202.1034, $^{2,4}A_1^+$ at m/z 184.0928, and $[B_1-C_5H_8O_2]^+$ at m/z 160.0945. The further continuous coverage of the peptide sequence was also achieved by detection of the y - and b -ions.

The CID-MS/MS analysis of the precursor ion $[M + 2H]^{2+}$ at m/z 1073.0144 isolated from the glycopeptide CASIQK*FGER (Lys 228) followed the same gas-phase fragmentation pathway as that of $[M + 2H]^{2+}$ precursor ion at m/z 1058.4652 described earlier (Fig. 5b).

CID-MS/MS of the GluC V8 digested peptides

The Mascot library search of the product ions obtained by the LC-MS/MS analysis of the GluC V8 digests matches the same serum albumin proteins as observed for the analysis of the tryptic digests, namely, serum albumin precursor (*B. taurus*) (gi|1351907) and serum albumin (*B. taurus*) (gi|74267962). The sequences of the peptides that matched these two isoforms are displayed in Table S6. The protein sequence coverage was 42% for the serum albumin (gi|74267962) and 45% for the precursor serum albumin (gi|1351907), which is consistently lower than the coverage obtained with the tryptic digests.

The following 17 glycation sites were also confirmed by low-energy CID-MS/MS analyses: Lys 65, Lys 75, Lys 88, Lys 100, Lys 117, Lys 151, Lys 183, Lys 197, Lys 256, Lys 266, Lys 304, Lys 309, Lys 336, Lys 374, Lys 420, Lys 455, and Lys 495 (Table 3). To illustrate the fragmentation of the glycated peptides identified after the digestion of the hapten-BSA glycoconjugate with the endoproteinase GluC V8, two CID spectra (Figs 5c and 5d) are displayed.

The LC-MS/MS analysis of the glycopeptides precursor ion $[M + 2H]^{2+}$ at m/z 987.9658 isolated from the glycopeptide K*VTKCCTE (Lys 495) (Fig. 5c, Table S7) enabled us to observe peptide product ions obtained by loss of the fragmented tetrasaccharide portion of the glycoconjugate: Y_2^+ at m/z 1569.6618, Y_1^+ at m/z 1423.6440, Y_0^+ at m/z 1277.5919, and $[Y_0-H_2O]^+$ at m/z 1259.5710. The y- and b- product ions permitted us to cover the peptide sequences and to detect fragmented peptide formed by loss of the tetrasaccharide moiety helped for the localization of the carbohydrate hapten on the Lys 495 residue: $[b_7-B_4]^+$ at m/z 1130.5222, $[b_5-B_4]^+$ at m/z 869.4469, $[b_4-B_4]^+$ at m/z 709.4351, $[b_3-B_4]^+$ at m/z 581.3174, and $[b_2-B_4]^+$ at m/z 480.2884. In addition, we noted that the same fragmented carbohydrate product ions, which were detected for the trypsin digestion of the hapten-BSA glycoconjugate, were also present: B_2^+ at m/z 406.2084, $C_{17}H_{26}N_3O_4^+$ at m/z 336.1892, C_1^+ at m/z 276.1415, B_1^+ at m/z 260.1460, $[B_1-H_2]^+$ at m/z 258.1307, $[B_1-H_2O]^+$ at m/z 242.1380, $[C_1-C_3H_7O]^+$ at m/z 202.1047, $^{2,4}A_1^+$ at m/z 184.0948, and $[B_1-C_5H_8O_2]^+$ at m/z 160.0912.

The other example that we selected to discuss was the glycopeptide K*QEPERNE (Lys 117) (Fig. 5d). The low-energy CID-MS/MS analysis of this glycopeptide precursor ion $[M + H]^{2+}$ at m/z 989.9697 extracted from K*QEPERNE (Lys 117) exhibited also the expected y- and b-product ions, which cover the peptide sequence, the carbohydrate product ions, and also the diagnostic peptide product ions, which have lost the tetrasaccharide B_4 portion.

During the LC-MS/MS analysis of the digested hapten-BSA glycoconjugate with both trypsin and GluC V8, we were able to identify more glycation sites than we did during the MALDI-TOF/TOF-MS/MS analysis of the digested glycoprotein. LC-MS/MS analysis allowed us to confidently identify 30 glycation sites on the lysine residues with total sequence coverage of 92%, whereas only 5 glycation sites were identified during the MALDI-MS/MS analysis of the digested glycoconjugate. The reason for this discrepancy seems to lie entirely on the formation of various glycoforms because of the protocol for the carbohydrate conjugation involving, intentionally, the amount of the glycation agent insufficient to effect complete the derivatization of the protein carrier. Moreover, during the study, we also observed identically located lysine residues in their derivatized and underivatized form. For example, the Lys 183 was detected through the underivatized doubly charged peptide ion LLYYANKYNGVFQE at m/z 861.4378 and also through its derivatized form on the doubly charged peptide ion LLYYANK*YNGVFQE at m/z 1336.1446 obtained from the GluC V8 digestion of the neoglycoconjugate. These observations support the fact that the sample is composed of a mixture of glycoforms.

The total glycation sites identified during this study are summarized in Fig. 6a. In addition, Fig. 6b displays the 3-D representation of the BSA obtained by homology modeling using the Swiss-Pdb Viewer software, in which the glycated lysine residues are highlighted in red.^[39,40] We can observe, that most the glycated lysines are located on the outer-surface region of the protein. Thus, we can presume that the carbohydrate hapten functionalization and occupancy on the outer surface of the BSA will depend entirely on its tertiary structure.

CONCLUSION

We have analyzed the *B. anthracis* neoglycoconjugate vaccine using the MALDI- and LC-ESI-tandem mass spectrometry. This was achieved to localize the carbohydrate haptens occupancies and their conjugation sites in the synthetic tetrasaccharide *B. anthracis* exosporium-squaric acid linker-BSA protein carrier neoglycoconjugate. The single-stage MALDI-MS analysis of the glycoconjugate was carried out to verify the average hapten-BSA ratio, which was found to be 5.4.

The high-energy MALDI-MS/MS analyses were carried out on tryptic and GluC V8 digests of the conjugate, and these allowed us to sequence and reveal only five glycation sites of the identified glycopeptides. It is noteworthy to mention that the MALDI-MS/MS experiments of the digested GluC V8 endoproteinase enabled the identification of three new glycation sites, which were not revealed during the analysis of the tryptic digests. The nano-LC-MS/MS analysis of the tryptic and GluC V8 digests permitted us to identify many more glycation sites (30 were identified). Fewer carbohydrate–spacer fragment ions were observed during the LC-MS/MS analysis when compared with the MALDI-MS/MS analyses, which are performed with higher collision energies that enhance carbohydrate cleavages.

The use of GluC V8 endoproteinase for the neoglycoconjugate digestion enabled us to identify more carbohydrate haptens occupancies and their respective glycation sites than when only trypsin was used. Thus, a total of 18 glycation sites were identified in the CID-MS/MS analysis of the tryptic digest, and 12 novel glycation sites were observed through the LC-MS/MS analysis of the GluC V8 digest of the carbohydrate–BSA conjugate.

Finally, the central conclusion drawn from this work, as expected, the nano-LC-MS/MS analysis of tryptic and GluC V8 digests is a reliable tool for the determination of carbohydrate occupancy and glycation sites of carbohydrate–protein conjugate vaccines.

Furthermore, as presented in this manuscript, we were able to use successfully ESI-CID-MS/MS to circumvent the complicated interpretation of the MALDI-CID fragmentation of the product ions containing the carbohydrate moieties. We are proposing to use in our future work, some more advanced activation methods for ESI-CID-MS/MS analysis, such as electron transfer dissociation and electron capture dissociation, measured by Fourier transform ion cyclotron resonance tandem mass spectrometry.^[42–44] These non-ergodic activation methods are known specifically to lead dissociation of the peptide backbone solely and to leave intact any further chemical modifications in the selected precursor ion.^[41,42]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES

1. Weintraub A. Immunology of bacterial polysaccharide antigens. *Carbohydr. Res.* 2003; 338:2539. [PubMed: 14670715]
2. Rietschel, ET.; Wollenweber, HW.; Brade, H.; Zahring, U.; Linder, B.; Seydel, U.; Bradaezek, H.; Barnickel, G.; Labischinski, H.; Giesbrecht, P. Handbook of Endotoxins, Chemistry of Endotoxins. Rietschel, ET., editor. Amsterdam, The Netherlands: Elsevier; 1984. p. 187
3. Holst, O.; Brade, H. Molecular Biochemistry and Cellular Biology. In: Morrison, DC.; Ryan, JL., editors. Bacterial Endotoxic Lipopolysaccharides. CRC Press: Boca Raton, Florida, USA; 1992. p. 135
4. Raetz CRH. Biochemistry of Endotoxins. *Ann. Rev. Biochem.* 1990; 59:129.
5. Kubler-Kielb J, Vinogradov E, Ben-Menachem G, Pozsgay V, Robbins JB, Schneerson R. Saccharide/protein conjugate vaccines for *Bordetella* species: Preparation of saccharide, development of new conjugation procedures, and physico-chemical and immunological characterization of the conjugates. *Vaccin.* 2008; 26:3587.
6. Pozsgay V. Recent Developments in Synthetic Oligosaccharide-Based Bacterial Vaccines. *Curr. Top. Med. Chem.* 2008; 8:126. [PubMed: 18289082]

7. Grandjean C, Boutonnier A, Dassy B, Fournier JM, Mulard LA. Investigation towards bivalent chemically defined glycoconjugate immunogens prepared from acid-detoxified lipopolysaccharide of *Vibrio cholerae* O1, serotype Inaba. *Glycoconj. J.* 2009; 26:41. [PubMed: 18648931]
8. Daum RS, Hogerman D, Rennels MB, Bewley K, Malinoski F, Rothstein E, Reisinger K, Block S, Keyserling H, Steinhoff M. Infant immunization with pneumococcal CRM₁₉₇ vaccines: effect of saccharide size on immunogenicity and interactions with simultaneously administered vaccines. *J. Infect. Dis.* 1997; 176:445. [PubMed: 9237711]
9. Lefebvre DJ, Kamerling JP, Vliegenthart JFG. Synthesis of *Streptococcus pneumoniae* Type 3 Neoglycoproteins Varying in oligosaccharide chain length, loading and carrier protein. *Chem. Eur. J.* 2001; 7:4411. [PubMed: 11695675]
10. Paoletti LC, Kasper DL, Michon F, DiFabio J, Jennings HJ, Tosteson TD, Wessels MR. Effects of chain length on the immunogenicity in rabbits of group B *Streptococcus* type III oligosaccharidetetanus toxoid conjugates. *J. Clin. Invest.* 1992; 89:203. [PubMed: 1729272]
11. Chernyak A, Kondo S, Wade TK, Meeks MD, Alzari PM, Fournier JM, Taylor RK, Ková P, Wade WF. Induction of protective immunity by synthetic *Vibrio cholerae* hexasaccharide derived from *V. cholerae* O1 Ogawa lipopolysaccharide bound to a protein carrier. *J. Infect. Dis.* 2002; 185:950. [PubMed: 11920320]
12. Jahouh F, Saksena R, Aiello D, Napoli A, Sindona G, Ková P, Banoub JH. Glycation sites in neoglycoconjugates from the terminal monosaccharide antigen of the O-PS of *Vibrio cholerae* O1, serotype Ogawa, and BSA revealed by Matrix-Assisted-Laser-Desorption-Ionization tandem mass spectrometry. *J. Mass Spectrom.* 2010; 45:1148. [PubMed: 20860010]
13. Mock M, Fouet A. Anthrax. *Annu. Rev. Microbiol.* 2001; 55:647. [PubMed: 11544370]
14. Priest, FG. *Bacillus subtilis* and other Gram-positive bacteria: Biochemistry, Physiology, and Molecular Biology. Sonenshein, AL.; Hoch, JA.; Losick, R., editors. Washington, D. C., USA: American Society for Microbiology; 1993. p. 3
15. Nicholson WL, Munakata N, Horneck G, Melosh HJ, Setlow P. Resistance of *Bacillus* endospores to extreme terrestrial and extraterrestrial environments. *Microbiol. Mol. Biol. Rev.* 2000; 64:548. [PubMed: 10974126]
16. Boutiba-Ben Boubaker I, Ben Redjem S. *Bacillus anthracis*: causative agent of anthrax. *Tunis. Med.* 2001; 79:642. [PubMed: 11892436]
17. Williams DD, Benedek O, Turnbough CL Jr. Species-specific peptide ligands for the detection of *Bacillus anthracis* spores. *Appl. Environ. Microbiol.* 2003; 69:6288. [PubMed: 14532093]
18. Chabot DJ, Scorpio A, Tobery SA, Little SF, Norris SL, Friedlander AM. Anthrax capsule vaccine protects against experimental infection. *Vaccine.* 2004; 23:43. [PubMed: 15519706]
19. Daubenspeck JM, Zeng H, Chen P, Dong S, Steichen CT, Krishna NR, Pritchard DG Jr, Turnbough CL. Novel oligosaccharide side chains of the collagen-like region of BclA, the major glycoprotein of the *Bacillus anthracis* exosporium. *J. Biol. Chem.* 2004; 279:30945. [PubMed: 15152001]
20. Adamo R, Saksena R, Ková P. Studies towards a conjugate vaccine for Anthrax: synthesis of the tetrasaccharide side chain of the *Bacillus anthracis* exosporium. *Helv. Chim. Acta.* 2006; 89:1075.
21. Saksena R, Adamo R, Ková P. Immunogens related to the synthetic tetrasaccharide side chain of the *Bacillus anthracis* exosporium. *Bioorg. Med. Chem.* 2007; 15:4283. [PubMed: 17412599]
22. Dell A, Morris HR. Glycoprotein structure determination by mass spectrometry. *Science.* 2001; 291:2351. [PubMed: 11269315]
23. Hui J, Desaire H, Butnev VY, Bousfield GR. Glycoprotein profiling by electrospray mass spectrometry. *J. Am. Soc. Mass Spectrom.* 2004; 15:750. [PubMed: 15121204]
24. Rawlings ND, Barrett AJ. Families of serine peptidases. *Meth. Enzymol.* 1994; 244:19.
25. Drapeau GR, Boily Y, Houmard J. Purification and properties of an extracellular protease of *Staphylococcus aureus*. *J. Biol. Chem.* 1972; 247:6720. [PubMed: 4627743]
26. Birktoft JJ, Breddam K. Proteolytic enzymes: Glutamyl endopeptidases. *Methods Enzymol.* 1994; 244:114. In. [PubMed: 7845201]
27. Chernyak A, Karavanov A, Ogawa Y, Ková P. Conjugating oligosaccharides to proteins by squaric acid diester chemistry; rapid monitoring of the progress of conjugation, and recovery of the unused ligand. *Carbohydr. Res.* 2001; 330:479. [PubMed: 11269399]

28. Yergey AL, Coorssen JR Jr, Backlund PS, Blank PS, Humphrey GA, Zimmerberg J. De novo sequencing of peptides using MALDI/TOF-TOF. *J. Am. Soc. Mass Spectrom.* 2002; 13:784. [PubMed: 12148803]
29. Domon B, Costello C. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J.* 1988; 5:397.
30. Roepstorff P, Fohlman J. Letter to the editors. *Biol. Mass Spectrom.* 1984; 11:601.
31. Johnson RS, Martin SA, Biemann K, Stults JT, Watson JT. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal. Chem.* 1987; 59:2621. [PubMed: 3688448]
32. Cornish TJ, Cotter RJ. A curved field reflectron time-of-flight mass spectrometer for the simultaneous focusing of metastable product ions. *Rapid Commun. Mass Spectrom.* 1994; 8:781. [PubMed: 7949338]
33. Kaufmann R, Kirsch D, Spengler B. Sequencing of peptides in a time-of-flight mass spectrometer: evaluation of post source decay following Matrix-Assisted Laser Desorption Ionization (MALDI). *Int. J. Mass Spectrom. Ion Process.* 1994; 131:355.
34. Satoh T, Sato T, Kubo A, Tamura J. Tandem Time-of-Flight mass spectrometer with high precursor ion selectivity employing spiral ion trajectory and improved offset parabolic reflectron. *J. Am. Soc. Mass Spectrom.* 2011; 22:797. [PubMed: 21472514]
35. Papayannopoulos IA. The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrom. Rev.* 1995; 14:49.
36. Burkitt WI, Giannakopoulos AE, Sideridou F, Bashir S, Derrick PJ. Discrimination effects in MALDI-MS of mixtures of peptides-analysis of the proteome. *Austr. J. Chem.* 2003; 56:369.
37. Kratzer R, Eckerskorn C, Karas M, Lottspeich F. Suppression effects in enzymatic peptide ladder sequencing using ultraviolet-matrix assisted laser desorption/ionization-mass spectrometry. *Electrophor.* 1998; 19:1910.
38. Schlosser A, Lehmann WD. Five-membered ring formation in unimolecular reactions of peptides: a key structural element controlling low-energy collision-induced dissociation of peptides. *J. Mass Spectrom.* 2000; 35:1382. [PubMed: 11180628]
39. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinform.* 2006; 22:195.
40. Peitsch MC. Protein modeling by E-mail. *Biol. Technol.* 1995; 13:658.
41. Zubarev RA, Kelleher NL, McLafferty FW. ECD of multiply charged protein cations. A non-ergodic process. *J. Am. Chem. Soc.* 1998; 120:3265.
42. Stensballe A, Jensen ON, Olsen JV, Haselmann KF, Zubarev RA. Electron capture dissociation of singly and multiply phosphorylated peptides. *Rapid Commun. Mass Spectrom.* 2000; 14:1793. [PubMed: 11006587]
43. Mirgorodskaya E, Roepstorff P, Zubarev RA. Localization of *O*-glycosylation sites in peptides by electron capture dissociation in a Fourier transform mass spectrometer. *Anal. Chem.* 1999; 71:4431. [PubMed: 10546526]
44. Hakansson K, Cooper HJ, Emmett MR, Costello CE, Marshall AG, Nilsson CL. Electron capture dissociation and infrared multiphoton dissociation MS/MS of an *N*-glycosylated tryptic peptide to yield complementary sequence information. *Anal. Chem.* 2001; 73:4530. [PubMed: 11575803]

Figure 1. Schematic representation of the general structure of carbohydrate-protein constructs from oligosaccharide fragments of the antigenic tetrasaccharide side chain of the *Bacillus anthracis* exosporium.

Figure 2. General strategy applied for the mass spectrometry determination of the glycation sites on the hapten-BSA glycoconjugate.

Figure 3. (a) MALDI-MS analysis of the tryptic carbohydrate hapten-BSA glycoconjugate digests. (b) MALDI-MS analysis of the digested carbohydrate hapten-BSA glycoconjugate with the GluC V8 endoproteinase.

Figure 4.

(a) MALDI-TOF/TOF-MS/MS spectra of the glycosylated peptide ALK*AWSVAR (Lys 235) at m/z 1951.0130. (b) Different product ions involving the fragmentation of the carbohydrate hapten observed during the MALDI-TOF/TOFMS/MS analysis of the glycosylated peptide ALK*AWSVAR (Lys 235) at m/z 1951.0130.

Figure 5.

LC-QqTOF-MS/MS spectra of the tryptic glycosylated peptides (a) CCTK*PESER (Lys 463) at m/z 1058.4652 (+2), (b) CASIQK*FGER (Lys 228) at m/z 1073.0144 (+2) and GluC V8 digests, (c) K*VTKCCTE (Lys 495) at m/z 987.9658 (+2) and (d) K*QEPERNE (Lys 117) at m/z 989.9697 (+2).

(a)

```

1 MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA
51 FSQYLQQCPF DEHVK*LVNEL TEFAK*TCVAD ESHAGCEK*SL HTLFGDELK*
101 VASLRETYGD MADCCAK*QEP ERNECFLSHK DDSPDLPKLK* PDPNTLCDEF
151 K*ADEK*K*FWGK YLYEIAARRHP YFYAPELLYY ANK*YNGVFQE CCQAEDK*GAC
201 LLPK*IETMRE K*VLTSSARQR LRCASIQK*FG ERALK*AWSVA RLSQKFPKAE
251 FVEVTK*LVTD LTKVHK*ECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
301 CCDK*PLLEK*S HCIAEVEKDA IPENLPPLTA DFAEDK*DVCK NYQEAKDAFL
351 GSFLYEYSRR HPEYAVSVLL RLAK*EYEATL EECCAKDDPH ACYSTVFDKL
401 K*HLVDEPQNL IKQNCDQFEK* LGEYGFQNEL IVRYTRK*VPQ VSTPTLVEVS
451 RSLGK*VGTRC CTK*PESERMP CAEDYLSLIL NRLCVLHEKT PVSEK*VTK*CC
501 TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIK*KQT
551 ALVELLKHK*P KATEEQLKTV MENFVAFVGK CCAADDKEAC FAVEGPKLVV
601 STQTALA

```

(b)

Figure 6.

(a) BSA sequence where the glycation sites are indicated by an asterisk (red = identified on tryptic digests, blue = identified on GluC V8 digests and red and underlined = identified on both tryptic and GluC V8 digests) and (b) 3D-structure of the BSA. The glycosylated lysine residues are highlighted in red (Swiss-Pdb Viewer software).

Table 1MALDI-TOF/TOF-MS/MS analysis of the glycosylated peptide ALK*AWSVAR (Lys 235) at m/z 1951.0130

Molecular ion	Calculated m/z	Experimental m/z	Deviation (Da)
[M + H] ⁺	1951.0158	1951.0180	0.002
[M + H - CH ₂ O] ⁺	1921.0052	1921.0932	0.088
[M + H - CH ₁₀ O ₃] ⁺	1880.9528	1881.1765	0.224
Y ₃ ⁺	1691.8738	1691.8829	0.009
Y ₂ ⁺	1545.8159	1545.8376	0.022
Y ₁ ⁺	1399.758	1399.7768	0.019
[y ₇ - B ₂] ⁺	1361.6947	1361.7265	0.032
Y ₀ ⁺	1253.7001	1253.7010	0.001
[Y ₀ - NH ₃] ⁺	1236.6735	1236.6813	0.008
[b ₃ - C ₅ H ₁₁ NO ₂] ⁺	1145.5711	1145.5757	0.005
[b ₄ - B ₄] ⁺	636.3715	636.2715	-0.100
y ₅ ⁺	618.3558	618.2805	-0.075
[b ₃ - B ₄] ⁺	565.3442	565.2661	-0.078
y ₃ ⁺	345.2245	345.1345	-0.090
C ₁₇ H ₂₆ N ₃ O ₄ ⁺	336.1918	336.1242	-0.068
B ₁ ⁺	260.1492	260.0739	-0.075
y ₂ ⁺	246.1561	246.0775	-0.079
[B ₁ - H ₂ O] ⁺	242.1387	242.0709	-0.068
2,5A ₁ ⁺	230.1387	230.0719	-0.067
z ₂ ⁺	229.1306	229.0569	-0.074
C ₁₁ H ₁₆ N ₃ O ₂ ⁺	222.1237	222.0670	-0.057
[C ₁ - C ₃ H ₇ O] ⁺	202.1074	202.0350	-0.072
y ₁ ⁺	175.119	175.0651	-0.054
[C ₁ - C ₅ H ₁₁ NO ₂] ⁺	159.0652	159.0486	-0.017
[y ₁ - CH ₂ O ₂] ⁺	129.1134	129.0679	-0.045
[z ₁ - CH ₂ O ₂] ⁺	112.0869	112.0527	-0.034
[z ₁ - CH ₄ O ₂] ⁺	110.0713	110.0414	-0.030

Tryptic glycopeptides identified in the bovine serum albumin protein by LCESI-QqToF-MS/MS analysis of the haptin-BSA glycoconjugate

Table 2

Precursor ion <i>m/z</i> (charge)	Mr (expt)	Mr (calc)	Deviation Da	Missed cleavage	Peptide
729.8826 (+2)	1457.7506	1457.7389	0.012	1	KHK*P (Lys 559)
733.3170 (+2)	1464.6194	1464.7698	-0.150	1	QIK*K (Lys 547)
770.3757 (+2)	1538.7369	1538.7338	0.003	1	ADEK*K (Lys 155)
807.9138 (+2)	1613.8129	1613.7964	0.016	0	K*FWGK (Lys 156)
828.0811 (+3)	2481.2213	2481.2005	0.021	2	LKECCDK*PLLEK (Lys 304)
832.7770 (+3)	2495.3092	2495.3146	-0.005	1	LK*HLVDEPQNLIK (Lys 401)
863.7965 (+3)	2588.3677	2588.3572	0.010	0	K*VPQVSTPTLVEVSR (Lys 437)
883.9655 (+2)	1765.9164	1765.9085	0.008	1	SLGK*VGTR (Lys 455)
922.0941 (+3)	2763.2605	2763.2493	0.011	1	LAK*EYEATLEECCAK (Lys 374)
969.5086 (+2)	1937.0026	1936.9868	0.016	1	TPVSEK*VTK (Lys 495)
970.4900 (+2)	1938.9738	1938.9772	-0.003	1	EK*VLTSSAR (Lys 211)
976.0173 (+2)	1950.0201	1950.0085	0.012	1	ALK*AWSVAR (Lys 235)
990.4710 (+3)	2968.3911	2968.3886	0.002	2	LK*PDPNTLCDEFKADEK (Lys 140)
1073.0144 (+2)	2144.0143	2144.0082	0.006	1	CASIQK*FGER (Lys 228)
1058.4652 (+2)	2114.9159	2114.9123	0.004	1	CCTK*PESER (Lys 463)
1208.5857 (+2)	2415.1569	2415.1284	0.028	1	VTK*CCTESLVNR (Lys 498)
1160.2172 (+3)	3477.6297	3477.6385	-0.009	1	QNCDQFEK*LGEYGFQNALIVR (Lys 420)
1169.5978 (+2)	2337.1810	2337.1583	0.023	1	GACLLPK*ITETMR (Lys 204)

Table 3

Glycopeptides identified in the bovine serum albumin protein by LC-ESI/MS analysis of the hapten-BSA glycoconjugate digested with the endoproteinase GluC V8

Precursor ion <i>m/z</i> (charge)	Mr(expt)	Mr(calc)	Deviation Da	Missed cleavage	Peptide
698.3598 (+2)	1394.7051	1394.6803	0.025	0	K*LGE (Lys 420)
790.3774 (+2)	1578.7403	1578.7288	0.011	1	K*QEPE (Lys 117)
844.3721 (+2)	1686.7297	1686.7499	-0.020	1	EFK*ADE (Lys 151)
894.4637 (+2)	1786.9128	1786.8976	0.015	0	HVK*L VNE (Lys 65)
862.9539 (+2)	1723.8933	1723.8754	0.018	1	VTK*L VTD (Lys 256)
897.4200 (+2)	1792.8255	1792.8176	0.008	0	K*SHCIAE (Lys 309)
902.4801 (+2)	1802.9456	1802.9289	0.017	0	LTKVHK*E (Lys 266)
962.1552 (+2)	2883.4439	2883.4093	0.035	0	VSRSLGK*VGTRCCTKPE (Lys 455)
987.9658 (+2)	1973.9171	1973.8949	0.022	0	K*VTKCCCTE (Lys 495)
989.9697 (+2)	1977.9248	1977.9154	0.009	2	K*QEPERNE (Lys 117)
992.4568 (+2)	1982.8989	1982.884	0.015	1	CCDK*PLLE (Lys 304)
995.4702 (+2)	1988.9258	1988.8912	0.035	1	FAK*TCVADE (Lys 75)
1013.0142 (+2)	2024.0138	2024.0123	0.001	0	LCK*VASLRE (Lys 100)
1048.5225 (+2)	2095.0305	2094.9984	0.032	1	K*SLHTLFGDE (Lys 88)
1097.0595 (+2)	2192.1045	2192.0909	0.014	0	DK*GACLLPKIE (Lys 197)
1124.5018 (+2)	2246.9890	2246.9876	0.001	1	DK*DVCKNYQE(Lys 336)
1336.1446 (+2)	2670.2746	2670.2728	0.002	0	LLYYANK*YNGVFQE (Lys 183)
1156.132 (+2)	2310.2495	2310.2345	0.015	0	YAVSVLLRLAK*E (Lys 374)