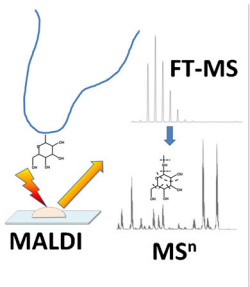
J. Am. Soc. Mass Spectrom. (2016) 27:735–747
DOI: 10.1007/s13361-015-1321-9

RESEARCH ARTICLE

Multi-Stage Mass Spectrometry Analysis of Sugar-Conjugated β-Turn Structures to be Used as Probes in Autoimmune Diseases

Chiara Giangrande,[1,2] Nicolas Auberger,[3,4,5] Cédric Rentier,[6,7,8] Anna Maria Papini,[6,7,8] Jean-Maurice Mallet,[3,4,5] Solange Lavielle,[3,4,5] Joëlle Vinh[1,2]

[1]Laboratory of Biological Mass Spectrometry and Proteomics, ESPCI ParisTech, PSL Research University, Paris, France
[2]CNRS USR 3149 SMBP, Paris, France
[3]Département de Chimie, École Normale Supérieure-PSL Research University, 24 rue Lhomond, 75005, Paris, France
[4]Sorbonne Universités, UPMC Univ Paris 06, LBM, 4 place Jussieu, F-75005, Paris, France
[5]CNRS, UMR 7203 LBM, F-75005, Paris, France
[6]Laboratory of Peptide and Protein Chemistry and Biology – PeptLab, Sesto Fiorentino, Italy
[7]Department of Chemistry "Ugo Schiff", University of Florence, Via della Lastruccia 13, 50019, Sesto Fiorentino, Italy
[8]PeptLab@UCP Platform and Laboratory of Chemical Biology EA4505, University of Cergy-Pontoise, 5 mail Gay-Lussac, 95031, Cergy-Pontoise CEDEX, France



Abstract. Synthetic sugar-modified peptides were identified as antigenic probes in the context of autoimmune diseases. The aim of this work is to provide a mechanistic study on the fragmentation of different glycosylated analogs of a synthetic antigenic probe able to detect antibodies in a subpopulation of multiple sclerosis patients. In particular the *N*-glucosylated type I′ β-turn peptide structure called CSF114(Glc) was used as a model to find signature fragmentations exploring the potential of multi-stage mass spectrometry by MALDI-LTQ Orbitrap. Here we compare the fragmentation of the glucosylated form of the synthetic peptide CSF114(Glc), bearing a glucose moiety on an asparagine residue, with less or non- immunoreactive forms, bearing different sugar-modifications, such as CSF114(GlcNAc), modified with a residue of *N*-acetylglucosamine, and CSF114[Lys^7(1-deoxyfructopyranosyl)], this last one modified with a 1-deoxyfructopyranosyl moiety on a lysine at position 7. The analysis was set up using a synthetic compound specifically deuterated on the C-1 to compare its fragmentation with the fragmentation of the undeuterated form, and thus ascertain with confidence the presence on an Asn(Glc) within a peptide sequence. At the end of the study, our analysis led to the identification of signature neutral losses inside the sugar moieties to characterize the different types of glycosylation/glycation. The interest of this study lies in the possibility of applyimg this approach to the discovery of biomarkers and in the diagnosis of autoimmune diseases.

Keywords: Glycosylation, Autoimmune diseases, Mass spectrometry, β-Turn structures, CID fragmentation

Received: 17 September 2015/Revised: 25 November 2015/Accepted: 26 November 2015/Published Online: 4 January 2016

Introduction

Carbohydrates are key constituents of all living organisms. Nearly 1% of each genome is devoted to highly conserved sugar processing enzymes, namely glycosyltransferases and glycosidases, which in a concerted work are able to create thousands of structures not only in eukaryotic cells but also in the microbial world [1]. Conjugation of sugars to proteins considerably increases variation among gene products with a total of 13 monosaccharides and eight amino acids forming at least 41 types of different glycosidic linkages [2]. Glycoproteins play a wide range of roles in biological processes. Glycans covering cell surfaces can be considered as a molecular code for cells to communicate with each other, a

Electronic supplementary material The online version of this article (doi:10.1007/s13361-015-1321-9) contains supplementary material, which is available to authorized users.

Correspondence to: Chiara Giangrande; *e-mail:* chiara.giangrande@espci.fr

"glycocode", which controls several biological functions [3]. There is strong evidence that aberrant post-translational modifications may play a central role in the onset and progression of autoimmune diseases [4] where glycoproteins/glycolipids can trigger an immunogenic response as a consequence of a modification of the glycosylation pattern. However, the non-template driven synthesis of glycans and their heterogeneity complicate the analyses of glycoconjugates.

Autoimmune diseases are a large group of disorders characterized by humoral and/or cell-mediated immune response to one or more self-antigens. The serum of autoimmune disease patients contains a wide range of autoantibodies, fluctuating with disease activity. Enzymatic post-translational modifications, like glycosylation and phosphorylation, are not the unique chemical modifications responsible for the onset of autoimmune diseases. Even non-enzymatic chemical modifications can lead to the formation of autoantigens. Protein glycation is an example of this type of modification and, though it is generally a physiological process leading to biological aging, it is responsible for the accumulation of immunogenic products that are recognized as non-self in certain conditions, particularly in the case of type I diabetes [5].

Multiple sclerosis is one of the most known inflammatory, demyelinating diseases of the central nervous system (CNS) possibly triggered by an autoimmune reaction against different CNS myelin components, like myelin basic protein (MBP) [6], proteolipid protein, and myelin oligodendrocyte glycoprotein [7].

In order to probe the nature of the modifications involved in the triggering onset of multiple sclerosis, a library of glycopeptide mimetics was screened following a "chemical reverse approach" [8] leading to the selection of post-translationally modified peptides, as optimized antigenic probes [9]. In particular, a type I′ β-turn peptide structure characterized by 21 amino acids named CSF114(Glc) was able to identify high titers of autoantibodies in patients' sera [10]. This *N*-glucosylated peptide was characterized by the presence of a β-D-glucopyranosyl moiety linking an asparagine residue on the tip of the β-turn structure by an amide bond [11]. There is no evidence for the existence of such a modification in eukaryotic cells, but recently a bacterial glycosyltransferase *Haemophilus influenzae* was described to be able to transfer hexose units (mainly glucose but possibly also galactose as mono- or dihexose molecules) to *N*-glycosylation sites [12–14]. Recently, the *Actinobacillus pleuropneumoniae N*-glycosyltransferase ApNGT was expressed in *E. coli* together with its substrate proteins and the *N*-glucosylation was detected by using specific anti-*N*-glucosylation human antibodies present in multiple sclerosis patients' sera. The approach allowed the detection of the target proteins, and their *N*-hexosylated residues were identified by LC-MS/MS, using ETD as fragmentation mode. Some endogenous *E. coli* proteins were also demonstrated to be modified by the enzyme [15]. Glucosylation by bacterial enzymes was already evidenced in other pathological states. For example *Clostridium difficile,* one of the major causes of intestinal infections in hospitals, was described to produce virulence factors that are indeed *O*-glucosylating toxins, acting on serine or threonine sites [16]. This observation led some of the authors to hypothesize that an aberrant *N*-glucosylation of myelin proteins can be one of the etiological causes for neoantigens formation triggering an early autoimmune response in multiple sclerosis.

Further studies based on the screening of differently glycosylated peptides and unmodified ones were able to demonstrate that the glucosylated asparagine is specifically recognized by the sera of multiple sclerosis patients' subpopulation, whereas peptides carrying different sugars or different modification sites had a lower or no affinity for the autoantibodies present in patients' sera [17]. A bioinformatic approach was explored to select the human myelin proteins mimicked by the synthetic antigenic probe CSF114(Glc), and thus sharing sequences characterized by a high degree of structural homology and the presence of a sequon (i.e., an *N*-glycosylation site). The study led to the identification of five proteins containing at least one peptide aligning CSF114(Glc) [18]. Therefore, the characterization of these glycopeptide mimetics is important from the analytical point of view addressing the issue of unveiling anti-sugar peptide structure antibodies as biomarkers that can be useful as diagnostic and/or prognostic tools. A recent study based on solid-phase ELISA using CSF114(Glc) as a probe has shown the presence of specific IgM autoantibodies in patients affected by Rett syndrome [19]. Even though shorter synthetic linear and cyclic sequences (eight residues) showed a nanomolar affinity, the 21-mer CSF114(Glc) still remains the best antigenic probe to detect IgMs as biomarkers of antibody-mediated forms of multiple sclerosis and Rett syndrome [20].

The aim of this work is to provide a mechanistic study on the fragmentation by multi-stage mass spectrometry of different sugar-conjugated forms at position 7 of the type I′ beta-turn structural probe CSF114 in order to find signature cleavages able to unequivocally identify either the type or the site of the modification. During CID fragmentations of glycopeptides, the sugar moiety tends to fragment more easily than the peptide backbone, giving poor information on the peptide sequence and on the modification site. The spectrum is strongly dominated by neutral losses occurring on the glycan moiety [21, 22]. The clarification of the fragmentation pathway is a common issue in the study of biomarkers containing sugar-peptide structures. Attempts to achieve more structural information were carried out using a multi-stage mass spectrometry approach on a MALDI-LTQ Orbitrap mass spectrometer (ThermoScientific, Bremen, Germany). The design of this instrument is characterized by the coupling of a MALDI source with an accurate mass, high-resolution hybrid mass spectrometer [23]. This instrument is suitable for obtaining structural information because of its ability to trap ions in the LTQ analyzer and, subsequently, to perform MS^n experiments.

Experimental

Reagents

Rink amide (*p*-methylbenzhydrylamine)-resin (MBHA-resin, 100–200 mesh, 0.52 mmol/g), *O*-(7-azobenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU), Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, and Fmoc-Ser-tert-butyl (*t*-Bu)-OH were purchased from Iris Biotech GmbH, Marktredwitz, Germany. *O*-(benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and Fmoc-His(Trt)-OH were purchased from Novabiochem (La Jolla, CA, USA). (1-2H)-D-glucose (98% D) was purchased from Eurisotop (Saint-Aubin, France). Fmoc-*L*-Lys(Boc)(2,3:4,5-di-*O*-isopropylidene-1-deoxyfructopyranosyl)-OH was purchased from Polypeptide Laboratories (Strasbourg, France). 2,5-Dihydroxybenzoic acid (DHB) matrix was purchased from LaserBio Labs (Sophia-Antipolis, France).

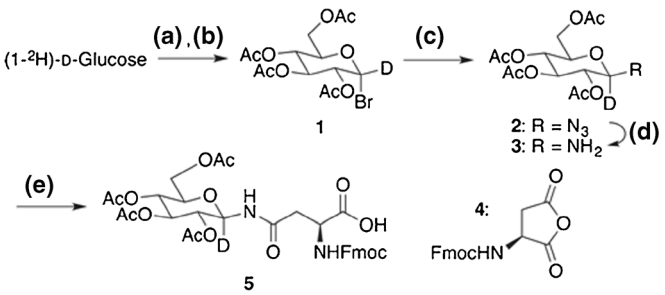
Synthesis of Fmoc-Asn[Glc(OAc)$_4$]-OH

The building block was prepared according to a previously reported procedure [24]. The corresponding deuterated building block 5 was synthesized following the same protocol and starting from (1- ^{2}H)-D-glucose (Scheme 1).

(1- ^{2}H)-2,3,4,6-Tetra-O-Acetyl-α-D-Glucopyranosyl Bromide (1) [25]

To a stirred suspension of (1-^{2}H)-D-glucose (992 mg, 5.5 mmol) in pyridine (6 mL) was added dropwise acetic anhydride (3 mL). The reaction mixture was stirred at room temperature (rt) for 16 h and concentrated. The residue was taken up in EtOAc, successively washed with 1 M aqueous HCl solution and saturated aqueous $NaHCO_3$ solution, dried ($MgSO_4$), and concentrated.

The residue was then dissolved in CH_2Cl_2 (3 mL) and 33% hydrogen bromide solution in acetic acid (6 mL) was added dropwise at 0 °C. The reaction mixture was slowly warmed to rt and stirred for 3 h, then washed with ice-water. Organic phase was washed with cold saturated aqueous $NaHCO_3$ solution, dried ($MgSO_4$), and concentrated to give glucosyl bromide 1 used in the next step: ^{1}H NMR (300 MHz, $CDCl_3$, δ) 5.55 (dd, 1H, $J_{H3\text{-}H2}$ = 9.9 Hz, $J_{H3\text{-}H4}$ = 9.3 Hz, H_3), 5.16 (dd, 1H, $J_{H4\text{-}H5}$ = 9.9 Hz, $J_{H4\text{-}H3}$ = 9.3 Hz, H_4), 4.83 (d, 1H, $J_{H2\text{-}H3}$ = 9.9 Hz, H_2), 4.36-4.25 (m, 2H, H_5, H_{6a}), 4.12 (d, $J_{H6b\text{-}H6a}$ = 10.8 Hz, H_{6b}), 2.09, 2.05, 2.03 (3s, 12H, Ac); ^{13}C NMR (75 MHz, $CDCl_3$, δ) 170.6, 170.0, 169.9, 169.6 (CO), 86.4 (t, $J_{C1\text{-}D}$ = 28 Hz, C_1), 72.3 (C_5), 70.7 (C_2), 70.3 (C_3), 67.3 (C_4), 61.1 (C_6), 20.8, 20.8, 20.7, 20.7 (Ac).



Scheme 1. Reagents and conditions for synthesis of building block 5: **(a)** Ac_2O, pyridine, 16 h; **(b)** 33% HBr in AcOH, CH_2Cl_2, 0 °C to rt, 3 h; **(c)** NaN_3, CH_3CN, reflux, 16 h, 61% in 3 steps; **(d)** H_2, Pd/C 10%, Et_3N, THF, 1h, 91%; **(e)** 4, DMSO, 1 h, 86%

(1-^{2}H)-2,3,4,6-Tetra-O-Acetyl-β-D-Glucopyranosyl Azide (2)

A solution of 1 and sodium azide (1.2 g, 19 mmol, 3.5 eq.) in acetonitrile (15 mL) was refluxed for 16 h, then filtered and concentrated. Crystallization in absolute ethanol afforded glucosyl azide 2 as a white solid (1.27 g, 61% in three steps): ^{1}H NMR (300 MHz, $CDCl_3$, δ) 5.22 (dd, 1H, $J_{H3\text{-}H4}$ = 9.6 Hz, $J_{H3\text{-}H2}$ = 9.3 Hz, H_3), 5.10 (dd, 1H, $J_{H4\text{-}H5}$ = 9.9 Hz, $J_{H4\text{-}H3}$ = 9.6 Hz, H_4), 4.96 (d, 1H, $J_{H2\text{-}H3}$ = 9.3 Hz, H_2), 4.28 (dd, 1H, $J_{H6a\text{-}H6b}$ = 12.6 Hz, $J_{H6a\text{-}H5}$ = 4.8 Hz, H_{6a}), 4.17 (dd, 1H, $J_{H6b\text{-}H6a}$ = 12.6 Hz, $J_{H6b\text{-}H5}$ = 2.1 Hz, H_{6b}), 3.79 (ddd, 1H, $J_{H5\text{-}H4}$ = 9.9 Hz, $J_{H5\text{-}H6a}$ = 4.8 Hz, $J_{H5\text{-}H6b}$ = 2.1 Hz, H_5), 2.11, 2.08, 2.03, 2.01 (4s, 12H, Ac); ^{13}C NMR (75 MHz, $CDCl_3$, δ) 170.8, 170.3, 169.5, 169.4 (CO), 87.7 (t, $J_{C1\text{-}D}$ = 23 Hz, C_1), 74.1 (C_5), 72.7 (C_3), 70.7 (C_2), 68.0 (C_4), 61.8 (C_6), 20.9, 20.7, 20.7 (Ac).

(1-^{2}H)-2,3,4,6-Tetra-O-Acetyl-β-D-Glucopyranosyl Amine (3)

To a stirred solution of 2 (1.25 g, 3.4 mmol) in THF (10 mL) were added triethylamine (480 µL, 3.4 mmol, 1 eq.) and Pd/C 10% (35 mg, 34 µmol, 1% mol). The reaction vessel was purged from air and filled with H_2 (3 bar). The reaction was stirred at rt for 1 h, then diluted with ethyl acetate, filtered through Celite, and concentrated, to give crude glucosyl amine 3 (1.06 g, 91%): ^{1}H NMR (300 MHz, $CDCl_3$, δ) 5.22 (dd, 1H, $J_{H3\text{-}H2}$ = 9.7 Hz, $J_{H3\text{-}H4}$ = 9.5 Hz, H_3), 5.02 (dd, 1H, $J_{H4\text{-}H5}$ = 10.0 Hz, $J_{H4\text{-}H3}$ = 9.5 Hz, H_4), 4.81 (d, 1H, $J_{H2\text{-}H3}$ = 9.7 Hz, H_2), 4.21 (dd, 1H, $J_{H6a\text{-}H6b}$ = 12.3 Hz, $J_{H6a\text{-}H5}$ = 4.8 Hz, H_{6a}), 4.08 (dd, 1H, $J_{H6b\text{-}H6a}$ = 12.3 Hz, $J_{H6b\text{-}H5}$ = 2.2 Hz, H_{6b}), 3.67 (ddd, 1H, $J_{H5\text{-}H4}$ = 10.0 Hz, $J_{H5\text{-}H6a}$ = 4.8 Hz, $J_{H5\text{-}H6b}$ = 2.2 Hz, H_5), 2.07, 2.05, 2.00, 1.99 (4s, 12H, Ac); ^{13}C NMR (75 MHz, $CDCl_3$, δ) 170.8, 170.3, 169.7 (CO), 84.7 (t, $J_{C1\text{-}D}$ = 23 Hz, C_1), 73.2 (C_3), 72.8 (C_5), 72.1 (C_2), 68.9 (C_4), 62.4 (C_6), 20.9, 20.9, 20.7, 20.7 (Ac).

N-Fmoc- L-Aspartic Anhydride (4)

Anhydride 4 was prepared as previously described [24] in two steps starting from L-aspartic acid.

N-α-Fmoc-N-γ-(1-^{2}H)-2,3,4,6-Tetra-O-Acetyl-β-D-Glucopyranosyl]- L-Asparagine (5)

To a stirred solution of 3 (1.06 g, 3.05 mmol, 1 eq.) in DMSO (2 mL) was added anhydride 4 (1.02 g, 3.05 mmol, 1 eq.). The reaction mixture was stirred at rt for 1 h, diluted with water, and extracted with EtOAc. Organic phase was washed twice with saturated aqueous NaCl solution, dried ($MgSO_4$), and

concentrated, to give 5 as a white solid (1.80 g, 86%), which was used without further purification step: ^{1}H NMR (300 MHz, $CDCl_3$, δ) 7.73 (d, 2H, J = 7.3 Hz, H_{ar}), 7.57 (d, 2H, J = 7.3 Hz, H_{ar}), 7.36 (dd, 2H, J = 7.3 Hz, J = 7.2 Hz, H_{ar}), 7.31-7.23 (m, 2H, H_{ar}), 7.04 (broad s, 1H, $H^{\delta 2}$), 6.26 (d, 1H, $J_H{}^N{}_{-H}{}^{\alpha}$ = 7.9 Hz, H^N), 5.33 (dd, 1H, J_{H3-H4} = 9.6 Hz, J_{H3-H2} = 9.5 Hz, H_3), 5.07 (dd, 1H, J_{H4-H5} = 9.8 Hz, J_{H4-H3} = 9.6 Hz, H_4), 4.95 (d, 1H, J_{H2-H3} = 9.5 Hz, H_2), 4.65-4.57 (m, 1H, H^{α}), 4.41-4.30 (m, 2H, CH_2 Fmoc), 4.27 (dd, 1H, $J_{H6a-H6b}$ = 12.0 Hz, J_{H6a-H5} = 3.9 Hz, H_{6a}), 4.24-4.15 (m, 1H, CH Fmoc), 4.04 (d, 1H, $J_{H6b-H6a}$ = 12.0 Hz, H_{6b}), 3.88-3.80 (m, 1H, H_5), 2.91 (dd, 1H, $J_H{}^{\beta}{}_{-H}{}^{\beta'}$ = 16.5 Hz, $J_H{}^{\beta}{}_{-H}{}^{\alpha}$ = 3.9 Hz, H^{β}), 2.77 (dd, 1H, $J_H{}^{\beta'}{}_{-H}{}^{\beta}$ = 16.5 Hz, $J_H{}^{\beta'}{}_{-H}{}^{\alpha}$ = 4.2 Hz, $H^{\beta'}$), 2.01, 2.00 (2s, 12H, Ac); ^{13}C NMR (75 MHz, $CDCl_3$, δ) 173.7 (C'), 171.4, 170.8, 170.1, 169.7 (CH_3CO), 156.6 (CO Fmoc), 143.8, 141.4 (Cq Fmoc), 127.9, 127.2, 125.3, 120.1 (CH_{ar}), 73.8 (C_5), 72.9 (C_5), 70.8 (C_2), 68.2 (C_4), 67.6 (CH_2 Fmoc), 61.8 (C_6), 50.6 (C^{α}), 47.1 (CH Fmoc), 37.7 (C^{β}), 20.8, 20.7 (Ac).

Synthesis of the Sugar-Modified Peptide Analogs of CSF114 (I-V)

Both the pentapeptide I corresponding to the central beta-turn motif included in CSF114(Glc) and its deuterated analog II (Table 1) were synthesized manually by solid-phase Fmoc/*t*-Bu strategy on a Rink amide resin in a fritted syringe (0.05 mmol scale).

Ac-RN(Glc)GHS-NH$_2$ (I)

Fmoc protection was removed with a solution of piperidine in DMF (20%, 5 min then 15 min). Stepwise coupling reactions were performed in DMF with Fmoc-protected natural amino acids, HBTU and NMM (5/5/7 eq.) for 20 min at rt. Coupling reaction with Fmoc-protected synthesized amino acid was performed with HATU and NMM (3/3/5 eq.) for 1 h at rt. After removal of the last Fmoc protecting group, the peptide was *N*-acetylated with a solution of acetic anhydride/NMM (1:1) in CH_2Cl_2 (20%, 2 × 45 min). The resin was washed with CH_2Cl_2, CH_3OH and dried under vacuum. It was then treated with TFA/H_2O/TIS (95:2.5:2.5; 25 mL/g resin, 3 h, rt), and the cleaved peptide was precipitated in diethyl ether and lyophilized.

Crude peptide was dissolved in MeOH/H_2O (5:1) and K_2CO_3 was added (pH ≈ 8–9) to remove the *O*-acetyl protecting groups. After completion of the reaction (2–3 h), the mixture was neutralized, concentrated, and lyophilized. Purification by RP-HPLC on a C18 semi-preparative column using a 30 min 0%–5% linear acetonitrile (0.1% TFA) gradient in an aqueous solution (0.1% TFA) afforded I as a white powder (14.3 mg, 28%) after lyophilization. Retention time of the product was 8.3 min. The peptide was characterized by MALDI-TOF MS (*m/z*): $[M + H]^+$ calcd, 773.35; found, 773.40.

Table 1. Sequences of the Synthetic Sugar-Modified Peptide Analogs of the Type I′ β-turn structure CSF114

N°	Name	Sequence
I	CSF114(Glc)(6–10)	Ac-RN(Glc)GHS-NH_2
II	CSF114(1-^{2}H)(6–10)	Ac-RN[(1-^{2}H)-Glc]GHS-NH_2
III	CSF114(Glc) [9]	TPRVERN(Glc)GHSVFLA PYGWMVK
IV	CSF114(GlcNAc) [9]	TPRVERN(GlcNAc)GHSVF LAPYGWMVK
V	[Lys^7(1-deoxyfructopyranosyl)] CSF114	TPRVERK(1-deoxyfructopyranosyl) GHSVFLAPYGWMVK

Ac-RN[(1-^{2}H)Glc]GHS-NH$_2$ (II)

The same procedure gave II as a white powder (10.8 mg, 21%). Retention time of the product was 8.6 min. The peptide was characterized by MALDI-TOF MS (*m/z*): $[M + H]^+$ calcd, 774.36; found, 774.49.

TPRVERN(Glc)GHSVFLAPYGWMVK (III) and TPRVERN(GlcNAc)GHSVFLAPYGWMVK (IV)

Peptides CSF114(Glc) (III) and CSF114(GlcNAc) (IV) were synthesized as previously described by Lolli et al. [9], including, respectively, at Asn-7 a β-D-glucopyranosyl and an *N*-acetyl-β-D-glucosylamine moiety.

TPRVERK(1-Deoxyfructopyranosyl) GHSVFLAPYGWMVK (V)

The glycated peptide [Lys^7(1-deoxyfructopyranosyl)]CSF114 (V) was synthesized introducing the building-block Fmoc-*L*-Lys(Boc)(2,3:4,5-di-*O*-isopropylidene-1-deoxyfructopyranosyl)-OH reported for the first time by Carganico et al. [26]

Briefly, the peptide [Lys^7(1-deoxyfructopyranosyl)]CSF114 was synthesized by SPPS, using a microwave-assisted peptide synthesizer (CEM LibertyBlue; CEM Corporation, Matthews, NC, USA), following the fluorenylmethoxycarbonyl (Fmoc)/(*t*-Bu) strategy. The peptide was prepared starting from a Fmoc-Lys(Boc)-Wang resin (0.31 mmol/g). All the reactions were performed in a Teflon vessel, where mixing and filtration are provided by a N_2 flow. Predissolved reagents are 0.2 M orthogonally protected Fmoc-amino acids, 0.5 M DIC, and 1 M Oxyma solutions in DMF. Fmoc-deprotection cycles were, respectively, of 15 s (75 °C, 155 W) and 30 s (90 °C, 30 W). Couplings of Arg residues were performed in 1500 s (25 °C, 0 W) + 300 s (75 °C, 30 W). Coupling of His residue was performed in 120 s (25 °C, 0 W) + 240 s (50 °C, 35 W). Couplings of other residues were performed in 15 s (90 °C, 170 W) + 110 s (90 °C, 30 W).

The protected building-block Fmoc-*L*-Lys(Boc)(2,3:4,5-di-*O*-isopropylidene-1-deoxyfructopyranosyl)-OH was introduced by manual batch SPPS with PLS 4×4 (Advanced ChemTech, Louisville, KY, USA) with an excess of 2.5 eq of HATU as activator, and 3.5 eq of NMM, for 1 h at rt.

The protected building-block Fmoc-*L*-Lys(Boc)(2,3:4,5-di-*O*-isopropylidene-1-deoxyfructopyranosyl)-OH was

introduced by manual batch SPPS with PLS 4×4 (Advanced ChemTech) with an excess of 2.5 eq, 2.5 eq of HATU as activator, and 3.5 eq of NMM, for 1 h at rt.

Cleavage from the resin and side-chain deprotection was performed at rt for 6 h using a mixture of TFA/TIS/H_2O (95:2.5:2.5 v/v/v) (1 mL/100 mg resin). The resin was filtered off and the solution was concentrated by flushing N_2. The peptide was precipitated from cold Et_2O, centrifuged, and lyophilized. The peptide was purified by semi-preparative RP-HPLC performed on a Waters (Waters Corporation, Milford, MA, USA) apparatus using a Phenomenex (Phenomenex Inc., Torrance, CA, USA) Jupiter column C_{18} (10 μm, 250 × 10 mm), with a gradient 25%–40% B in 30 min at 4 mL/min with solvent system A (0.1% TFA in H_2O), and B (0.1% TFA in CH_3CN). Peptide was obtained with a purity of at least 95%, as a white powder (yield of 4%).

Peptide characterization was performed by analytical HPLC-MS using an Alliance Chromatography (Waters) with a Phenomenex Kinetex C_{18} column (2.6 μm , 3.0 × 100 mm), with a gradient 20%–70% B in 5 min at 0.6 mL/min, coupled to a single quadrupole ESI-MS (Micromass ZQ). Rt: 3.99 min. *m/z* calcd. 2620.4 $[M + H]^+$, found 874.8 $[M + 3H]^{3+}$.

Mass Spectrometry

The mass spectrometry experiments were performed on a ThermoScientific MALDI LTQ Orbitrap XL mass spectrometer. DHB was used as matrix at a concentration of 30 mg/mL in EtOH/H_2O/TFA, (7:3:0.1, v/v). Samples were dissolved in aq. 0.1% formic acid (v/v) to a final concentration of 0.5 μM. All full scan analyses were performed in positive ion mode, using the Orbitrap as detector and a resolving power of 60,000. Automatic gain control (AGC) allowed the control of the filling status of the linear ion trap. Five microscans were recorded for each cycle. Crystal positioning system (CPS) was used to optimize the choice of shot position, and one microscan was registered for each step. Laser intensity was set between 15 and 25 kV in order to avoid in-source fragmentation. Fragmentation experiments were conducted in the linear ion trap with ion trap detection manually choosing the best normalized collision energy and using Q = 0.250 and T = 30 ms.

Results and Discussion

The mechanistic study has been focused on a series of sugar-modified peptides (Table 1).

First we analyzed two *N*-glucosylated pentapeptides: the *N*-glucosylated peptide I corresponding to the central beta-turn motif included in CSF114(Glc) and its deuterated form on C-1 of the glucose moiety, the analog II. This short sequence with only 5 amino acids was chosen to aid spectral interpretation focusing on the cleavages at the level of the sugar moiety and of the modification site. *C*-terminus amidation, mimicking an amide bond within a larger sequence, was useful to reduce water losses in the peptide backbone, thus facilitating the attribution of neutral losses occurring in the sugar moiety.

Fragmentation experiments on the two parent ions of I and II at *m/z* 773.39 and 774.39, respectively, were conducted using different values of normalized collision energy (CE), using the linear ion trap as collision cell and detector (data not shown). CID fragments were detected in the LTQ trap because of its higher sensitivity. The best collision energy was chosen to obtain sequence information as well as the modification site.

The MS/MS spectra obtained at lower levels of CE (between 5% and 20%) displayed the presence of peaks probably corresponding to the cleavage at the level of the most labile bonds giving rise to fragments at *m/z* 653.42 and 594.42 (data not shown). Neutral losses frequently occur in the sugar moiety of glycoconjugates where analytes undergo water and formaldehyde losses forming detectable intermediary oxonium ions. The mechanism has been extensively investigated in the case of protein glycation [27], where the elimination of water molecules was described with a proton-driven mechanism leading to the formation of oxonium ions.

The same neutral losses are the most intense peaks even in the MS^2 spectrum obtained at CE 27% (Supplementary Figure 1). Table 2 summarizes the products arising from MS^2 and MS^3 experiments carried out on peptides I and II. MS^2 spectra are populated by successive water losses, with the loss of three water units already described in glycopeptide esters [28]. We can interpret the peak at *m/z* 653.42 as the loss of a unit of $C_4H_8O_4$ (MH^+ – 120 Da) on the base of the chemical formula for glucose, whereas the peak at 594.42 can be explained with the loss of $C_6H_{10}O_5 + NH_3$ (MH^+ – 179 Da). The neutral loss of 120 Da in the deuterated *N*-glucosylated pentapeptide (II) still exhibits a mass shift of 1 Da if compared with the corresponding undeuterated I, whereas the fragment corresponding to the neutral loss of 179 Da has no mass shifts, clearly showing that the C-1 of the sugar moiety is not implicated in the loss of 120 Da but it is lost in the second case when the entire sugar moiety is lost together with an ammonia molecule. The inspection of the spectrum allows also the visualization of additional neutral losses. The signal at *m/z* 743.33 can be easily interpreted as the elimination of a formaldehyde unit (30 Da), whereas peaks at *m/z* 713.25 and 683.33 can be due to the elimination of two CH_2O and three CH_2O units, with a nominal mass of 60 and 90 Da, respectively. Another possible interpretation can be found in the loss of aldose units of $C_x(H_2O)_x$. The fragment at *m/z* 683.33 might have undergone the elimination of a $C_2H_4O_2$ (glycolaldehyde) unit followed by the elimination of a formaldehyde unit, or of the elimination of a $C_3H_6O_3$ (glyceraldehyde) unit. All these species still contain the deuterium on C-1. The elimination of the whole glucose moiety, corresponding to a loss of 162 Da at *m/z* 611.33, shows no mass difference among the deuterated peptide II and the undeuterated I. It is worth noting that this loss is of lower intensity if compared to the losses of 120 and 179 Da. This observation has induced us to explore the

Table 2. List of the Monoisotopic Mass of the Most Significant Ions Observed in Product Ion Spectra (MS^2 and MS^3) of Peptide I and Peptide II Recorded on MALDI-LTQ Orbitrap, Using CID Fragmentation in the LTQ Orbitrap. Observed and Calculated Values are Listed in the Table

MS^2	Peptide I		Peptide II	
	Observed (Da)	Calculated (Da)	Observed (Da)	Calculated (Da)
MH^+	773.42	773.35	774.42	774.36
$-H_2O$	755.42	755.34	756.42	756.35
$-CH_2O$	743.33	743.34	744.42	744.35
$-H_2O-NH_3$	738.33	738.31	739.33	739.32
$-2H_2O$	737.42	737.33	738.42	738.34
$-CH_2O-H_2O$	725.42	725.33	726.33	726.34
$-2H_2O-NH_3$	720.33	720.30	721.33	721.31
$-C_2H_4O_2$	713.25	713.33	714.33	714.34
$3H_2O-NH_3$	702.33	702.29	703.25	703.30
$-C_3H_6O_3$	683.33	683.32	684.42	684.33
b_4	669.33	669.29	670.33	670.30
$-C_4H_8O_4$	653.42	653.31	654.42	654.32
$-C_4H_8O_4-H_2O$	636.33	636.30	637.33	637.31
$-C_4H_8O_4-H_2O-NH_3$	618.33	618.27	619.33	619.28
$-C_6H_{10}O_5$	611.33	611.30	611.33	611.30
$-C_6H_{10}O_5-NH_3$	594.42	594.27	594.33	594.27
$-C_6H_{10}O_5-2NH_3$	577.25	577.24	577.33	577.24
y_4	575.25	575.24	576.33	576.25
b_3	532.25	532.23	533.23	533.24
$b_4- C_6H_{10}O_5-NH_3$	490.33	490.21	490.17	490.21
b_2	475.25	475.21	476.25	476.22
$y_4- C_4H_8O_4$	455.33	455.21	456.25	456.22
$y_4- C_6H_{10}O_5$	413.17	413.19	413.25	413.19
$y_4- C_6H_{10}O_5-NH_3$	396.25	396.17	396.17	396.17
$b_3- C_6H_{10}O_5-NH_3$	353.17	353.16	353.17	353.16
$b_2- C_6H_{10}O_5$	313.25	313.06	313.08	313.06
y_3	299.17	299.15	299.08	299.15
y_2	242.17	242.12	242.08	242.12
MS^3 ($-C_6H_{10}O_5$)				
$MH^+-C_6H_{10}O_5$	611.42	611.30	611.33	611.30
$b_4-C_6H_{10}O_5$	507.42	507.24	507.50	507.24
$y_4-C_6H_{10}O_5$	413.08	413.19	413.25	413.19
$b_3-C_6H_{10}O_5$	369.92	370.18	370.25	370.18
$b_2-C_6H_{10}O_5$	313.25	313.16	313.00	313.16
y_2	241.17	242.12	241.83	242.12
b_1	198.83	199.12	199.12	199.12
$MS^3(-C_6H_{10}O_5-NH_3)$				
$MH^+-C_6H_{10}O_5-NH_3$	594.33	594.27	594.58	594.27
$MH^+-C_6H_{10}O_5-NH_3-CH_2O$	564.43	564.26	564.33	564.26
$b_4-C_6H_{10}O_5-NH_3$	490.17	490.21	490.25	490.21
$b_4-C_6H_{10}O_5-2NH_3$	473.17	473.18	473.17	473.18
$y_4-C_6H_{10}O_5-NH_3$	396.25	396.16	396.17	396.16
$b_3-C_6H_{10}O_5-NH_3$	353.00	353.15	353.00	353.15
$b_2-C_6H_{10}O_5-NH_3$	299.08	299.14	299.08	299.14
y_3	296.08	296.13	296.17	296.13
y_2	242.17	242.12	242.00	242.12
b_1	198.92	199.12	199.17	199.12
$MS^3(-C_4H_8O_4)$				
$MH^+-C_4H_8O_4$	653.42	653.31	654.33	654.32
$MH^+-C_4H_8O_4-CH_2O$	623.42	623.30	624.42	624.31
$MH^+-C_6H_{10}O_5$	611.33	611.30	611.33	611.30
$MH^+-C_6H_{10}O_5-NH_3$	594.33	594.27	594.33	594.27
$b_4-C_4H_8O_4$	549.33	549.25	550.17	550.26
$y_4-C_4H_8O_4$	455.17	455.20	456.25	456.21
$b_3-C_4H_8O_4$	412.25	412.19	413.25	413.20
$b_2-C_4H_8O_4$	355.17	355.17	356.17	356.18
y_3	299.17	299.14	299.17	299.14
y_2	242.08	242.12	242.08	242.12
b_1	199.00	199.12	199.08	199.12

potential of multi-stage tandem mass spectrometry to validate our hypotheses. MS^3 experiments were carried out on the fragments corresponding to the neutral losses of 120, 162, and 179 Da.

The fragment at *m/z* 611.33 is not very intense in the MS/MS spectrum compared with the other neutral losses, but its MS^3 fragmentation allows the reconstruction of the peptide sequence. The spectra of the undeuterated and deuterated

species do not display any differences, underlying that the deuterium is lost during the previous fragmentation event. Both the b- and y-series are represented in the figure, giving complete sequence coverage of this small peptide.

MS^3 experiments on the species losing the entire glucose moiety and an ammonia molecule (Supplementary Figure 3) are obtained by a more intense precursor compared with the neutral loss of 162 Da. Most probably the additional neutral loss of 17 Da occurs on the amide of the asparagine residues. In the MS^3 spectrum, further eliminations of ammonia units are possible, for example the peak at *m/z* 473.17 can be explained by the loss of two ammonia units and a glucose unit from the b_4 ion. The fragment at *m/z* 594.33 even undergoes a neutral loss of 30 Da forming *m/z* 564.42. We can easily interpret this peak as a formaldehyde loss due to the fragmentation at the level of the carbonyl of the asparagine residue. In the MS^2 spectrum, the signal corresponding to the loss of 179 Da is higher than the signal corresponding to the loss of only the glucose moiety and, for this reason, we consider that this loss is indeed a signature of this modification.

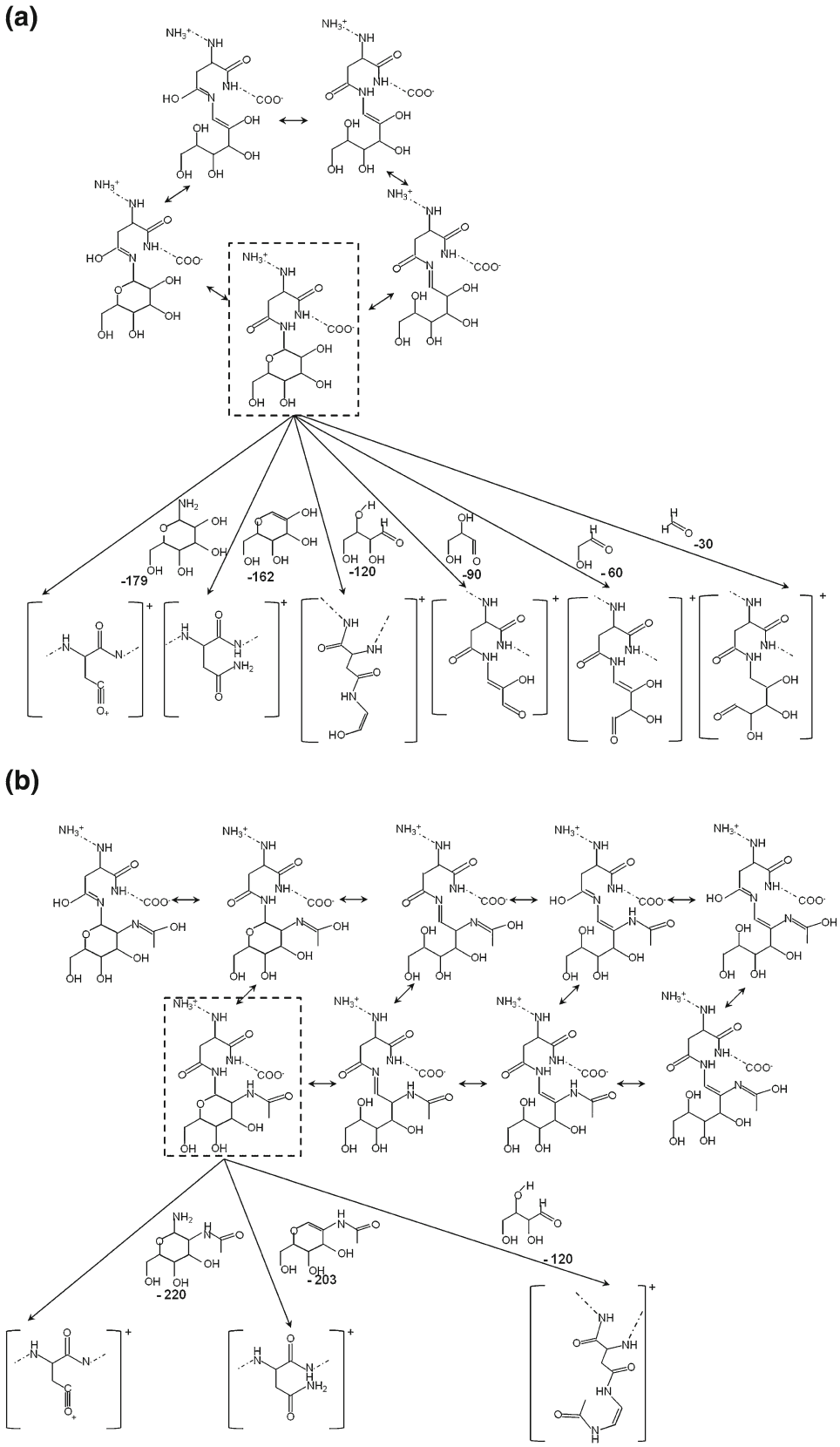
The most intense diagnostic fragmentation is the neutral loss of 120 Da (Supplementary Figure 4). This species still contains the glucose C-1, as is clearly demonstrated by the increase of 1 Da on many peaks in the spectrum corresponding to the deuterated species peptide II. This cleavage can be explained by the loss of four consecutive formaldehyde molecules, two $C_2H_4O_2$ (glycolaldehyde) units, or a $C_4H_8O_4$ (erythrose) from the open ring structure as shown in Scheme 2A, summarizing the most significant sugar ring cleavages described above. Scheme 2A also displays the possibility to consider several resonance structures for the parent ion, thus explaining the relative stability of the amide bond between asparagine and glucose, as well as the coexistence of the open-ring conformation together with the pyranosidic ring.

Exploring the MS^3 spectrum, it was surprising not to detect the loss of the residual modification leading to the theoretical mass of 611.3 Da (unglucosylated form), whereas a loss of 59 Da can be interpreted as the loss of a unit containing the residual modification as well as an ammonia molecule (NH_2CH_2COH). A fragment attributable to the loss of a unit containing five carbon atoms is observed at *m/z* 623.42 and 624.42 in peptides I and II, respectively. The difference of 1 Da between the two species indicates that this fragment still contains the deuterium bound to the C-1. This peculiar fragmentation leading to the complete degradation of the glucose moiety from its open ring structure was already evidenced by Jerić et al. in the case of glycopeptide esters [28].

Sugar-Peptide Analogs of CSF114

Starting from this characterization, we have exploited the potential of multi-stage mass spectrometry by using the MALDI-LTQ Orbitrap to study the fragmentation behavior of the type I′ β-turn structure CSF114(Glc) (III). This peptide was responsible for specific detection of serum autoantibodies in a multiple sclerosis patient subpopulation [10]. In the attempt to identify signature fragmentations for the *N*-glucosylation motif, multi-stage mass spectrometry experiments were carried out. The pattern of fragmentation obtained for this peptide was first compared with the fragmentation of standard peptides, and then with the series of CSF114 peptide structures bearing different sugar modifications, namely the peptide CSF114(GlcNAc) (IV), modified with a residue of *N*-acetylglucosamine on an Asn residue at position 7, and the peptide [Lys^7(1-deoxyfructopyranosyl)]CSF114 (V) obtained by introducing during the solid-phase synthesis the protected building block Fmoc-*L*-Lys(Boc)(2,3:4,5-di-*O*-isopropylidene-1-deoxyfructopyranosyl)-OH (26) at position 7 of the peptide structure. This last modification, known as protein glycation, is due in vivo to the non-enzymatic reaction of glucose with free amino groups, giving a Schiff base that is easily rearranged forming a ketoamine named Amadori product [29].

Figure 1 shows the multi-stage mass spectrometry characterization of the peptide CSF114(Glc) (III). Figure 1a displays the Orbitrap detection of the monoisotopic mass of this peptide at *m/z* 2606.33 Da. The model of fragmentation proposed for the standard pentapeptides I and II is clearly confirmed in Figure 1b, representing its MS^2 spectrum. The diagnostic neutral losses of 162, 120, and 179 Da are all present at *m/z* 2444.18, at *m/z* 2486.18, and at *m/z* 2427.09, respectively, as well as the loss of single formaldehyde units and aldose moieties. The spectrum is registered in the mass range *m/z* 715–2800 because of the low-mass cutoff of the LTQ trap. In addition to the signals already present in the previous analyses, here we find a further signal in the high mass region of the spectrum at *m/z* 2562.18 corresponding to a neutral loss of 44 Da. One of the positions where the loss can take place is the *C*-terminal end of the peptide where CO_2 units can rarely be lost. This loss was not detected in the peptides I and II with a *C*-terminal amide. But, there is a second possibility where the loss can be located to the *N*-terminal threonine residue. It has already been reported [30] that threonine residues can undergo the loss of CH_3CHO units in peptides containing the sequence TT- or TS- at the *N*-terminal side. The condition for this is a previous water loss in the side chain of the amino acid in second position generating a cyclic intermediate, which may eliminate an acetaldehyde fragment. In this case the cyclic intermediate was mimicked by proline, whose presence may promote the loss of 44 Da without any previous water loss in the peptide backbone. To confirm this hypothesis, the MS^3 spectra of the daughter ion at *m/z* 2562.18 (Supplementary Figure 5) was carried out. The spectrum clearly confirms the second hypothesis: in fact all the b-fragments are detected with the loss of 44 Da. For example, the b_7 fragment has mass to charge ratio of 971.39 instead of 1015.55 in the MS^2 spectrum. Conversely, the y-series remain unmodified (as shown for fragment y_{16} at *m/z* 2023.82) until the fragment bears the neutral loss, which is the *N*-terminal amino acid. Another confirmation of this hypothesis is given in the MS^3 spectrum of the fragment at *m/z* 2444.30, corresponding to the loss of the entire sugar moiety (Figure 1c). In this spectrum, we should be



Scheme 2. Hypothetical fragmentation mechanism of the asparagine-modified peptides: resonance structures for the precursor ions are displayed. **(a)** Fragmentation of the *N*-glucosylated peptide: the precursor *N*-glucosylated-asparagine is evidenced by the dotted rectangle. **(b)** Fragmentation of the *N*-acetylglucosamine-modified peptide: the precursor is evidenced by the dotted rectangle

able to detect a fragment corresponding to the loss of 44 Da on the threonine residue, and indeed this fragment is present at *m/z* 2400.09, clearly showing that the loss was located on the *N*-terminal end of the peptide backbone. Thus, the MS^3 spectrum of the neutral loss of 162 allowed not only the elucidation of the peptide sequence but even the loss of 44 Da was explained. The MS^2 spectrum exhibits the peak at *m/z* 2427.09 corresponding to the loss of the entire sugar moiety and of the ammonia molecule (−179 Da) already described for the standard peptides. This peak is more abundant than the one corresponding to the loss of only the sugar moiety even in this case.

As for the cross-ring cleavages in the glucose moiety, the loss of a formaldehyde unit can be detected in the MS^2 spectrum at *m/z* 2576.18, as well as the loss of a unit of 60 Da, corresponding to the loss of two formaldehyde units or of a glycolaldehyde unit (*m/z* 2546.18). In addition to this, at *m/z* 2516.18 the loss of 90 Da can be easily explained by the loss of three formaldehyde units, or of the glycolaldehyde fragment and a formaldehyde molecule, or of a glyceraldehyde moiety. The peak at *m/z* 2486.09 corresponding to the loss of 120 Da is more intense compared with the above-mentioned neutral losses. As it was previously explained, our interpretation for this loss leading to the cross-ring $^{0,2}X$-type fragment is the loss of two glycolaldehyde units or, most probably, of an erythrose molecule. In analogy with the previous analyses on the *N*-glucosylated pentapeptides I and II, the fragmentation was investigated for the species at *m/z* 2516.18 and 2486.18 (Supplementary Figures 6 and 7). This last one is higher than the peak corresponding to the loss of the sugar moiety, as already explained in the analyses of I and II.

Finally, the CSF114(Glc) (III) presents the typical fragmentation pattern for the glucose moiety that can be considered the signature of this modification in the research for biomarkers in autoimmune neurodegenerative diseases.

The following step was used to assess if the type of sugar-peptide linkage and the kind of monosaccharide moiety can influence the fragmentation behavior and then the characterization of the sugar-modification. In this respect, we report the fragmentation spectrum of the peptide CSF114(GlcNAc) (IV) bearing a unit of *N*-acetylglucosamine on asparagine 7. The aim was to establish if the presence of an *N*-acetyl-group on the C-2 was responsible for a different signature fragmentation. This modification induces a mass increment of 203.08 Da on the asparagine residue in position 7. The monoisotopic mass of this peptide was accurately detected by Orbitrap analysis and displayed at *m/z* 2647.35 in Figure 2a. Exploring the MS^2 spectrum of this precursor in Figure 2b, the loss of the CH_3CHO unit on the *N*-terminal threonine can be easily recognized at *m/z* 2603.55, as previously noticed. This peak is related to the peak at *m/z* 2585.45 with the loss of an extra water

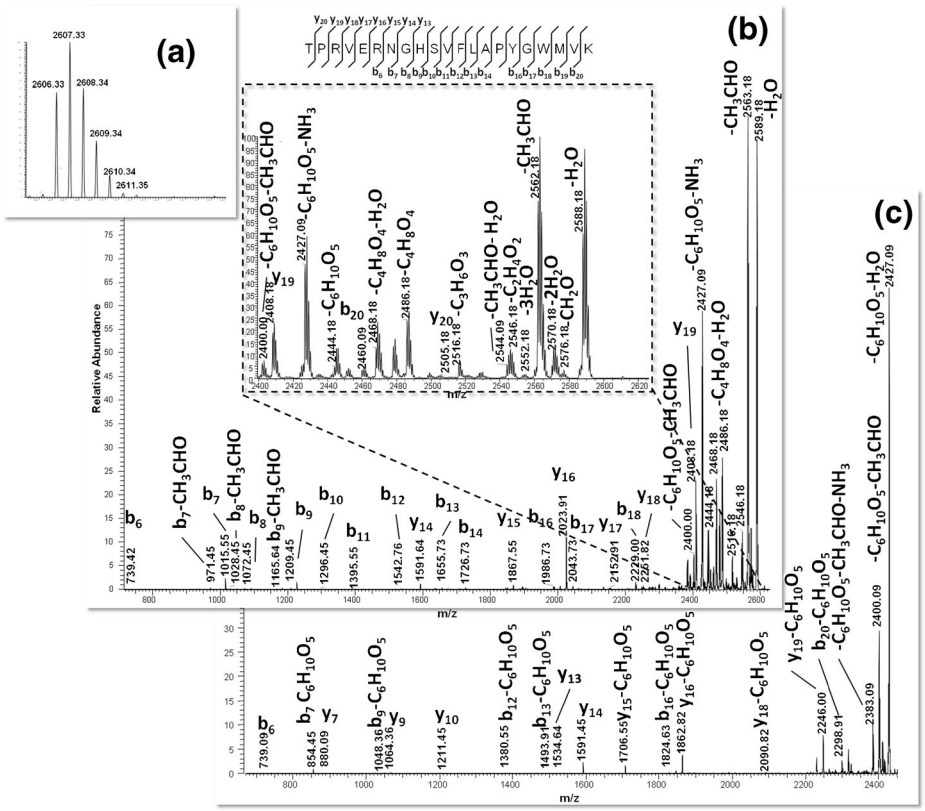


Figure 1. Multi-stage mass spectrometry analysis of peptide III [CSF114(Glc)]. **(a)** MALDI-Orbitrap detection of precursor ion at *m/z* 2606.33. **(b)** MS/MS fragmentation of precursor ion obtained at CE 40% in the LTQ trap. The high mass region of the spectrum was magnified in order to highlight neutral losses in the sugar moiety. **(c)** MS^3 spectrum of the peak at *m/z* 2444.16 corresponding to $[MH^+ - C_6H_{10}O_5]$

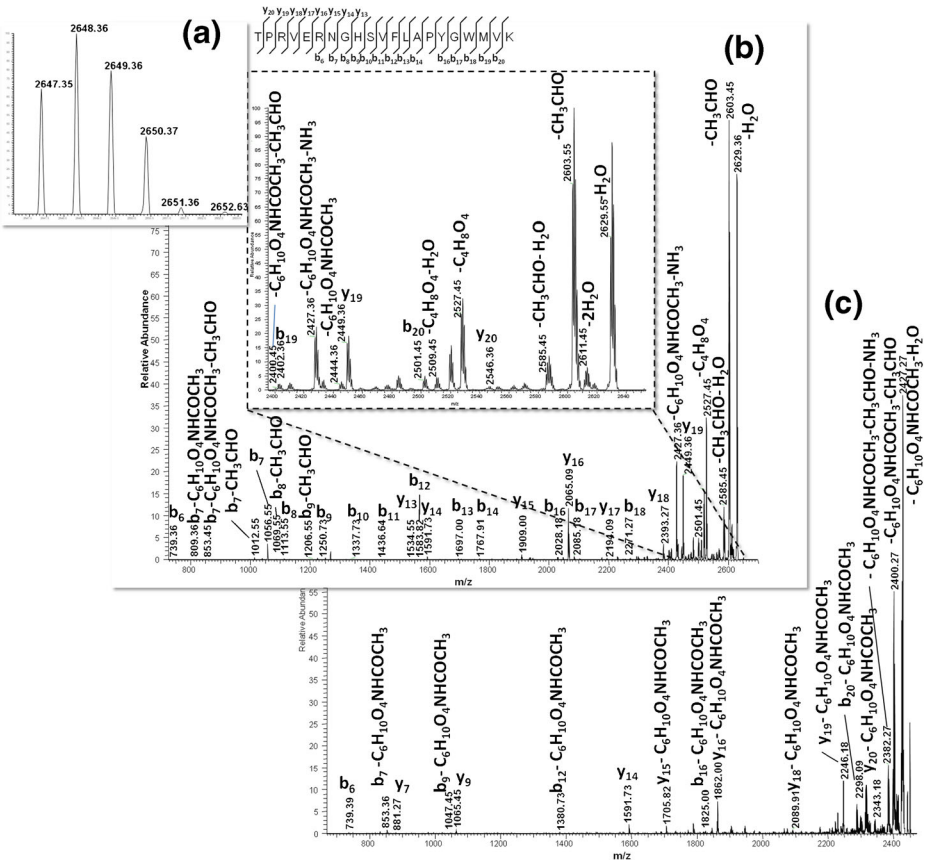


Figure 2. Multi-stage mass spectrometry analysis of peptide IV [CSF114(GlcNAc)]. **(a)** MALDI-Orbitrap detection of precursor ion at *m/z* 2647.35. **(b)** MS/MS fragmentation of precursor ion obtained at CE 40%in the LTQ trap. The high mass region of the spectrum was magnified in order to highlight neutral losses in the sugar moiety. **(c)** MS^3 spectrum of the peak at *m/z* 2444.36 corresponding to $[MH^+ - C_6H_{10}O_4NHCOCH_3]$

molecule. The neutral loss of two water molecules is detected at *m/z* 2611.45, and it is the maximum number of water molecules lost during the experiment. The analysis of the spectrum shows that the elimination of the *N*-acetyl-hexose generates the peak at *m/z* 2444.36, as previously reported, for the peptide CSF114(Glc). However, the abundance of this fragment is very low compared with the losses of 120 and 220 Da. The latter, corresponding to the elimination of the sugar moiety accompanied by the loss of an ammonia molecule from the amide of the asparagine residue, is represented by a very abundant peak at *m/z* 2427.36.

Figure 2c represents the MS^3 spectrum of the species at *m/z* 2444.36 that displays a profile very similar to Figure 1c, with the complete fragmentation of the peptide backbone.

At *m/z* 2527.45 the neutral loss of 120 Da can be detected. Considering the intensity of this peak, we can infer that the cleavage is highly favored during the fragmentation process, probably because of the presence of the *N*-acetyl group on the C-2. However, our interpretation for this fragmentation differs from the same neutral loss mechanism proposed for the peptide CSF114(Glc) (III). The first difference lies in the fact that the losses of formaldehyde and aldose units were not detected in the MS^2 spectrum of the peptide CSF114(GlcNAc) (IV). The delocalization of the lone pairs of electrons of nitrogen and oxygen of the *N*-acetyl-group allows the representation of the structure as a resonance hybrid with more resonance structures compared with the CSF114(Glc) (III). The *N*-acetyl-group extends electron delocalization promoting the cleavage of the sugar ring to generate an $X^{0,2}$-type cross-ring cleavage during CID fragmentation, as summarized in Scheme 2B, where resonance structures are also displayed for the parent ion. Conversely, the peptide CSF114(Glc) tends to fragment by consecutively eliminating formaldehyde and aldose molecules.

Asparagine residues modified in animal proteins by an *N*-acetylglucosamine in the consensus sequence for *N*-glycosylation were recently found in plants. *A. thaliana* proteins bearing this modification were explained as the product of the reaction of the enzyme endo-β-*N*-acetylglucosaminidase on larger high mannose *N*-glycans [31]. In animal cells, endoglycosidases are encoded in the genome even if it is not clear if their activity is directed towards glycoproteins, glycopeptides, or free oligosaccharide [32, 33]. In fact, little evidence exists confirming the actual in vivo activity of these endoglycosydases. Chalkley et al. [34] proposed that the identification of GlcNAc peptides in post-synaptic density preparations from murine brains might be due to an artifact during sample preparation.

In contrast to the enzymatic attachment of sugar units to proteins catalyzed by oligosaccharyl transferases (OST), there

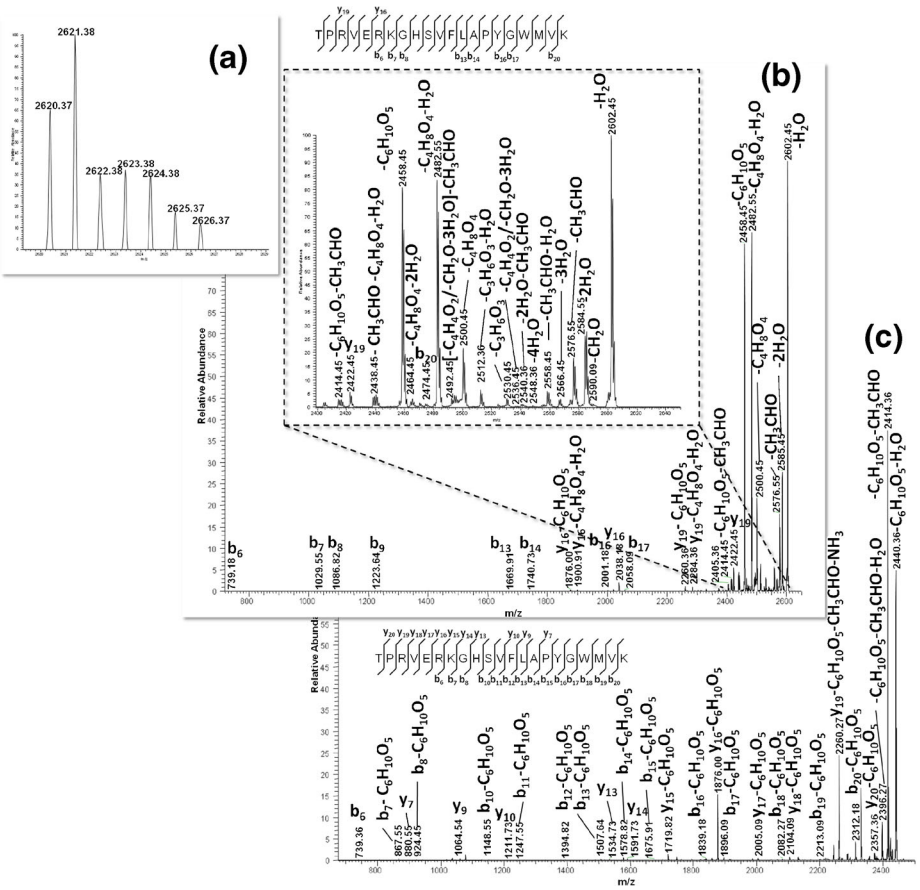


Figure 3. Multi-stage mass spectrometry analysis of peptide V [Lys^7(1-deoxyfructopyranosyl)]CSF114. **(a)** MALDI-Orbitrap detection of precursor ion at *m/z* 2620.37. **(b)** MS/MS fragmentation of precursor ion obtained at CE 40%in the LTQ trap. The high mass region of the spectrum was magnified in order to highlight neutral losses in the sugar moiety. **(c)** MS^3 spectrum of the peak at *S* 2458.45 corresponding to [$MH^+ - C_6H_{10}O_5$]

is another pathway for addition of sugars to proteins: the glycation reaction. Glycation of proteins occurs through spontaneous and reversible condensation of a reducing sugar and the free amino group of a protein at the *N*-terminus and/or on lysine side chains, to form a Schiff base. This then undergoes a rearrangement into the more stable ketoamine known also as

Table 3. List of the Monoisotopic Mass of the Most Significant Neutral Losses Observed in Product Ion Spectra (MS^2) of Peptide III, IV, and V recorded on MALDI-LTQ Orbitrap, Using CID Fragmentation in the LTQ Orbitrap. Observed and Calculated Values are Listed in the Table. The Neutral Losses that are Not Detected in a Particular Case are Indicated with ND (Not Detected)

	Peptide III		Peptide IV		Peptide V	
	Observed (Da)	Calculated (Da)	Observed (Da)	Calculated (Da)	Observed (Da)	Calculated (Da)
MS (FT detection)						
MH^+	2606.33	2606.31	2647.37	2647.35	2620.36	2620.37
MS2 (LTQ detection)						
$-H_2O$	2588.18	2588.3	2629.55	2629.34	2602.45	2602.36
$-CH_2O$	2576.18	2576.3	ND	2617.34	2590.09	2590.36
$-2H_2O$	2570.18	2570.29	2611.45	2611.33	2584.55	2584.35
$-CH_3CHO$	2562.18	2562.28	2603.55	2603.33	2576.55	2576.35
$-3H_2O$	2552.18	2552.28	ND	2593.32	2566.45	2566.34
$-C_2H_4O_2$	2546.18	2546.29	ND	2587.33	2560.45	2560.35
$-CH_3CHO- H_2O$	2544.09	2544.27	2585.45	2585.32	2558.45	2558.34
$-4H_2O$	ND	2534.27	ND	2575.31	2548.36	2548.33
$-C_3H_6O_3$	2516.18	2516.28	ND	2557.32	2530.45	2530.34
$-C_4H_8O_4$	2486.18	2486.27	2527.45	2527.31	2500.45	2500.33
$-C_4H_8O_4-H_2O$	2468.18	2468.26	2509.55	2509.3	2482.55	2482.32
-Sugar moiety	2444.18	2444.27	2444.36	2444.27	2458.45	2458.32
-Sugar moiety $-NH_3$	2427.09	2427.23	2427.36	2427.23	ND	2441.29
-Sugar moiety $-CH_3CHO$	2400.00	2400.25	2400.45	2400.25	2414.18	2414.30

the Amadori product [29]. In the case of glucose, the initially formed Schiff base rearranges into the more stable 1-deoxyfructopyranosyl moiety. Subsequent dehydration, condensation, fragmentation, oxidation, and cyclization reactions lead to the irreversible formation of advanced glycation end products (AGEs). Glycation of CSF114 was achieved introducing at position 7 the completely protected building block Fmoc-*L*-Lys(Boc)(2,3:4,5-di-*O*-isopropylidene-1-deoxyfructopyranosyl)-OH, previously described [26]. Glycation induces a mass shift of 162 Da on the modified residue. The monoisotopic mass of the glycated CSF114 is displayed in Figure 3a at *m/z* 2620.36. In the fragmentation spectrum in Figure 3b, we can notice a peak at *m/z* 2458.45 corresponding to the precursor after the loss of the entire carbohydrate. This peak is more intense than the peak at *m/z* 2500.45 corresponding to the loss of 120 Da; both of them were already reported [35], as well as the presence of water losses. A typical signature fragmentation for the glycated structure is represented by the peak at *m/z* 2536.45, describing the loss of a partial dehydrated glucose ($C_4H_4O_2$) as previously reported [36]. The same neutral loss was interpreted as a loss of three water molecules and a formaldehyde unit in previous works, with the formation of furylium ions [27, 37]. It is worth noting that here the loss of the entire sugar moiety at *m/z* 2458.45 is represented by a more intense peak compared with the neutral loss of 120 Da at *m/z* 2500.45, and this behavior is completely different if compared with the *N*-modified glucopeptides. Furthermore, the loss of the sugar moiety together with the amine group, which was predominant in the above-mentioned forms of CSF114, is completely absent here, probably because it could generate an unstable primary carbocation on the residual group of the lysine. The spectrum is dominated by the cleavages inside the sugar moiety and b-ions are very weak or even absent, as previously reported; [38] consequently, it is very difficult to obtain sequence information from this spectrum. For this reason, MS^3 fragmentation was used to have a better sequence coverage for the peptide. In Figure 3b, the spectrum corresponding to the neutral loss of 162 Da, the entire sugar moiety, displays the improvements achieved for sequence information. However, very limited information is obtained on the modification site, probably because of the difficult ionization conditions of the fragments bearing the sugar moiety. To summarize, in Table 3 all the most significant neutral losses observed in product ion spectra of peptides III, IV, and V are listed in order to better compare the fragmentation pattern of these glycopeptides.

Conclusions

Sugar-conjugated peptides present a very characteristic pattern of neutral losses during CID fragmentation experiments. Herein, we propose the analysis of synthetic modified peptides with different sugar moieties by the use of multi-stage mass spectrometry on a MALDI-LTQ Orbitrap. The study allowed the identification of signature fragmentations to trace the nature of the monosaccharide moiety and the linking amino acid introduced at position 7 on the tip of the type I′ β-turn model peptide CSF114 in three different variants (III–V). To the best of our knowledge, CID fragmentation of a peptide with a β-D-glucopyranosyl moiety linked to an asparagine residue is deeply investigated for the first time. The inspection of spectra revealed a different behavior when comparing the fragmentation of this glucose-modified peptide to the *N*-acetylglucosamine-modified peptide the same position (position 7). The main difference was the more complex fragmentation pattern of the glucose-conjugated peptide comprising the neutral losses of formaldehyde and aldose units, even though some similarities are also evident, as the preferential cross-ring cleavage leading to the neutral loss of 120 Da.

The more common non-enzymatic glucose modification represented by protein glycation was also investigated by fragmenting a peptide carrying the isobaric modification of lysine residue at the same position 7 by the 1-deoxyfructopyranosyl moiety. Even in this case, the comparison revealed peculiar signature neutral losses, concerning the type of fragments and the relative abundance among them. It is worth noting that in this case the peak corresponding to the complete loss of the sugar moiety (162 Da) was preferred to the cross-ring cleavage (loss of 120 Da).

All these consideration might constitute some preliminary insights into the development of simple analytical methods to analyze aberrantly glycosylated antigens in autoimmune diseases. MALDI-MS has emerged as a rapid and reliable tool, already used in other biomarkers and diagnostics analytical methods. Here, the coupling of a MALDI source to the highly accurate FT analyzer allowed not only the univocal identification of autoantigens but also their fragmentation exploiting the efficiency of the linear ion trap fragmentation and the possibility of performing multi-stage fragmentation events.

Acknowledgments

The authors acknowledge support for this work by the "Fondation Pierre-Gilles de Gennes" for a grant "Immune-mediated diseases: Glycopeptidomic-based diagnostics to follow-up disease activity." Moreover, ANR Chaire d'Excellence PeptKit 2009–2014 (grant no. ANR-09-CEXC-013-01), French-Italian University (Vinci Project grant no. C2-133), and Ente Cassa di Risparmio di Firenze are gratefully acknowledged for their financial support to PeptLab.

References

1. Kiessling, L.L., Splain, R.A.: Chemical approaches to glycobiology. Annu. Rev. Biochem. **79**, 619–653 (2010)

2. Gabius, H.J., André, S., Jiménez-Barbero, J., Romero, A., Solís, D.: From lectin structure to functional glycomics: principles of the sugar code. Trends Biochem. Sci. **36**(6), 298–313 (2011)
3. Kulkarni, A.A., Weiss, A.A., Iyer, S.S.: Glycan-based high-affinity ligands for toxins and pathogen receptors. Med. Res. Rev. **30**(2), 327–393 (2010)
4. Doyle, H.A., Mamula, M.J.: Post-translational protein modifications in antigen recognition and autoimmunity. Trends Immunol. **22**(8), 443–449 (2001)
5. Shweta, B., Sheon, M., Reema, B., Ashok, P.G., Mahesh, J.K.: Immune response to chemically modified proteome. Proteom Clin. Appl. **8**, 19–34 (2014)
6. Kim, J.K., Mastronardi, F.G., Wood, D.D., Lubman, D.M., Zand R., Moscarello M.A.: Multiple sclerosis: an important role for post-translational modifications of myelin basic protein in patho-genesis. Mol. Cell. Proteomics **2**, 453–462 (2003)
7. McLaughlin, K.A., Chitnis, T., Newcombe, J., Franz, B., et al.: Age-dependent B cell autoimmunity to a myelin surface antigen in pediatric multiple sclerosis. J Immunol **183**, 4067–4076 (2009)
8. Papini, A.M.: The use of post-translationally modified peptides for detection of biomarkers of immune-mediated diseases. J. Pept. Sci. **15**, 621–628 (2009)
9. Lolli, F., Mulinacci, B., Carotenuto, A., Bonetti, B., Sabatino, G., Mazzanti, B., D'Ursi, A.M., Novellino, E., Pazzagli, M., Lovato, L., Alcaro, M.C., Peroni, E., Pozo-Carrero, M.C., Nuti, F., Battistini, L., Borsellino, G., Chelli, M., Rovero, P., Papini, A.M.: An *N*-glucosylated peptide detecting disease-specific autoantibodies, biomarkers of multiple sclerosis. Proc. Natl. Acad. Sci. U. S. A. **102**(29)), 10273–10278 (2005)
10. Lolli, F., Mazzanti, B., Pazzagli, M., Peroni, E., Alcaro, M.C., Sabatino, G., Lanzillo, R., Brescia Morra, V., Santoro, L., Gasperini, C., Galgani, S., D'Elios, M.M., Zipoli, V., Sotgiu, S., Pugliatti, M., Rovero, P., Chelli, M., Papini, A.M.: The glycopeptide CSF114(Glc) detects serum antibodies in multiple sclerosis. J. Neuroimmunol. **167**(1/2), 131–137 (2005)
11. Carotenuto, A., Alcaro, M.C., Saviello, M.R., Peroni, E., Nuti, F., Papini, A.M., Novellino, E., Rovero, P.: Designed glycopeptides with different β-turn types as synthetic probes for the detection of autoantibodies, biomarkers of multiple sclerosis. J. Med. Chem. **51**, 5304–5309 (2008)
12. Gross, J., Grass, S., Davis, A.E., Gilmore-Erdmann, P., Townsend, R.R., St. Geme III, J.W.: The Hemophilus influenzae HMW1 adhesin is a glycoprotein with an unusual N-linked carbohydrate modification. J. Biol. Chem. **283**(38), 26010–26015 (2008)
13. Grass, S., Buscher, A.Z., Swords, W.E., Apicella, M.A., Barenkamp, S.J., Ozchlewski, N., St Geme III, J.W.: The Hemophilus influenzae HMW1 adhesin is glycosylated in a process that requires HMW1C and phosphoglucomutase, an enzyme involved in lipo-oligosaccharide biosynthesis. Mol. Microbiol. **48**(3), 737–751 (2003)
14. Choi, K.J., Grass, S., Paek, S., St. Geme III, J.W., Yeo, H.J.: The Actinobacillus pleuropneumoniae HMW1C-like glycosyltransferase mediates N-linked glycosylation of the Hemophilus influenzae HMW1 adhesin. PLoS One **5**(12), e15888 (2010)
15. Naegeli, A., Neupert, C., Fan, Y.Y., Lin, C.W., Poljak, K., Papini, A.M., Schwarz, F., Aebi, M.L.: Molecular analysis of an alternative *N*-glycosylation machinery by functional transfer from *Actinobacillus pleuropneumoniae* to *Escherichia coli*. J. Biol. Chem. **289**(4), 2170–2179 (2014)
16. Zeiser, J., Gerhard, R., Just, I., Pich, A.: Substrate specificity of clostridial glucosylating toxins and their function on colonocytes analyzed by proteomics techniques. J. Proteome Res. **12**(4), 1604–1618 (2013)
17. Nuti, F., Peroni, E., Real-Fernández, F., Bonache, M.A., Le Chevalier-Isaad, A., Chelli, M., Lubin-Germain, N., Uziel, J., Rovero, P., Lolli, F., Papini, A.M.: Post-translationally modified peptides efficiently mimicking neoantigens: a challenge for theragnostics of autoimmune diseases. Biopolymers **94**(6), 791–799 (2010)
18. Pandey, S., Alcaro, M.C., Scrima, M., Peroni, E., Paolini, I., Di Marino, S., Barbetti, F., Carotenuto, A., Novellino, E., Papini, A.M., D'Ursi, A.M., Rovero, P.: Designed glucopeptides mimetics of myelin protein epitopes as synthetic probes for the detection of autoantibodies, biomarkers of multiple sclerosis. J. Med. Chem. **55**(23), 10437–10447 (2012)
19. Papini, A.M., Nuti, F., Real-Fernandez, F., Rossi, G., Tiberi, C., Sabatino, G., Pandey, S., Leoncini, S., Signorini, C., Pecorelli, A., Guerranti, R., Lavielle, S., Ciccoli, L., Rovero, P., De Felice, C., Hayek, J.: Immune dysfunction in Rett syndrome patients revealed by high levels of serum anti-N(Glc) IgM antibody fraction. J. Immunol. Res. **2014**, 260973 (2014)
20. Real Fernández, F., Di Pisa, M., Rossi, G., Auberger, N., Lequin, O., Larregola, M., Benchohra, A., Mansuy, C., Chassaing, G., Lolli, F., Hayek, J., Lavielle, S., Rovero, P., Mallet, J.M., Papini, A.M.: Antibody recognition in multiple sclerosis and Rett syndrome using a collection of linear and cyclic *N*-glucosylated antigenic probes. Biopolymers. **104**(5), 560–576 (2015)
21. Ko, B.J., Brodbelt, J.S.: Comparison of glycopeptide fragmentation by collision induced dissociation and ultraviolet photodissociation. Int. J. Mass Spectrom. **377**(1), 385–392 (2015)
22. Ruan, E.D., Wang, H., Ruan, Y., Juárez, M.: Study of fragmentation behavior of amadori rearrangement products in lysine-containing peptide model by tandem mass spectrometry. Eur. J. Mass Spectrom. (Chichester, Engl) **19**(4), 295–303 (2013)
23. Strupat, K., Kovtoun, V., Bui, H., Viner, R., Stafford, G., Horning, S.: MALDI produced ions inspected with a linear ion trap-Orbitrap hybrid mass analyzer. J. Am. Soc. Mass Spectrom. **20**, 1451–1463 (2009)
24. Ibatullin, F.M., Selivanov, S.I.: Reaction of N-Fmoc aspartic anhydride with glycosylamines: a simple entry to *N*-glycosyl asparagines. Tetrahedron Lett **50**, 6351–6354 (2009)
25. Hosie, L., Sinnott, M.L.: Effects of deuterium substitution alpha and beta to the reaction center, 18O substitution in the leaving group, and aglycone acidity on hydrolyses of aryl glucosides and glucosyl pyridinium ions by yeast alpha-glucosidase. A probable failure of the antiperiplanar-lone-pair hypothesis in glycosidase catalysis. Biochem. J. **226**, 437–446 (1985)
26. Carganico, S., Rovero, P., Halperin, J.A., Papini, A.M., Chorev, M.: Building blocks for the synthesis of post-translationally modified glycated peptides and proteins. J. Pept. Sci. **15**(2), 67–71 (2009)
27. Stefanowicz, P., Kapczynska, K., Jaremko, M., Jaremko, Ł., Szewczuk, Z.: A mechanistic study on the fragmentation of peptide-derived Amadori products. J. Mass Spectrom. **44**(10), 1500–1508 (2009)
28. Jerić, I., Versluis, C., Horvat, S., Heck, A.J.: Tracing glycoprotein structures: electrospray ionization tandem mass spectrometric analysis of sugar-peptide adducts. J. Mass Spectrom. **37**(8), 803–811 (2002)
29. Baynes, J.W., Watkins, N.G., Fisher, C.I., Hull, C.J., Patrick, J.S., Ahmed, M.U., Dunn, J.A., Thorpe, S.R.: The Amadori product on protein: structure and reactions. Prog. Clin. Biol. Res. **304**, 43–67 (1989)
30. Neta, P., Pu, Q.L., Yang, X., Stein, S.E.: Consecutive neutral losses of H_2O and C_2H_4O from N-terminal Thr-Thr and Thr-Ser in collision-induced dissociation of protonated peptides. Position-dependent water loss from single Thr or Ser. Int J Mass Spectrom **267**, 295–230 (2007)
31. Kim, Y.C., Jahren, N., Stone, M.D., Udeshi, N.D., Markowski, T.W., Witthuhn, B.A., Shabanowitz, J., Hunt, D.F., Olszewski, N.E.: Identification and origin of N-linked β-D-*N*-acetylglucosamine monosaccharide modifications on Arabidopsis proteins. Plant Physiol. **161**(1), 455–464 (2013)
32. Suzuki, T., Yano, K., Sugimoto, S., Kitajima, K., Lennarz, W.J., Inoue, S., Inoue, Y., Emori, Y.: Endo-β-*N*-acetylglucosaminidase, an enzyme involved in processing of free oligosaccharides in the cytosol. Proc. Natl. Acad. Sci. U. S. A. **99**(15), 9691–9696 (2002)
33. Chantret, I., Fasseu, M., Zaoui, K., Le Bizec, C., Yayé, H.S., Dupré, T., Moore, S.E.: Identification of roles for peptide: *N*-glycanase and endo-β-*N*-acetylglucosaminidase (Engase1p) during protein *N*-glycosylation in human HepG2 cells. PLoS One **5**(7), e11734 (2010)
34. Chalkley, R.J., Thalhammer, A., Schoepfer, R., Burlingame, A.L.: Identification of protein O-GlcNAcylation sites using electron transfer dissociation mass spectrometry on native peptides. Proc. Natl. Acad. Sci. U. S. A. **106**(22), 8894–8899 (2009)
35. Montgomery, H., Tanaka, K., Belgacem, O.: Glycation pattern of peptides condensed with maltose, lactose, and glucose determined by ultraviolet matrix-assisted laser desorption/ionization tandem mass spectrometry. Rapid Commun. Mass Spectrom. **24**(6), 841–848 (2010)
36. Gadgil, H.S., Bondarenko, P.V., Treuheit, M.J., Ren, D.: Screening and sequencing of glycated proteins by neutral loss scan LC/MS/MS method. Anal. Chem. **79**(15), 5991–5999 (2007)
37. Zhang, Q., Petyuk, V.A., Schepmoes, A.A., Orton, D.J., Monroe, M.E., Yang, F., Smith, R.D., Metz, T.O.: Analysis of non-enzymatically glycated peptides: neutral-loss-triggered MS(3) versus multi-stage activation tandem mass spectrometry. Rapid Commun. Mass Spectrom. **22**(19), 3027–3034 (2008)
38. Frolov, A., Hoffmann, P., Hoffmann, R.: Fragmentation behavior of glycated peptides derived from D-glucose, D-fructose, and D-ribose in tandem mass spectrometry. J. Mass Spectrom. **41**(11), 1459–1469 (2006)