
Article:
Couto, N., Davlyatova, L., Evans, C.A. orcid.org/0000-0003-4356-9216 et al. (1 more author) (2018) Application of the broadband collision-induced dissociation (bbCID) mass spectrometry approach for protein glycosylation and phosphorylation analysis. Rapid Communications in Mass Spectrometry, 32 (2). pp. 75-85. ISSN 0951-4198

https://doi.org/10.1002/rcm.8016

Application of the bbCID mass spectrometry approach for protein glycosylation and phosphorylation analysis

Journal:	*Rapid Communications in Mass Spectrometry*
Manuscript ID	RCM-17-0185.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a

Complete List of Authors:	Couto, Narciso; University of Sheffield, The ChELSI Institute, Department of Chemical and Biological Engineering Davlyatova, Liliya; University of Sheffield, The ChELSI Institute, Department of Chemical and Biological Engineering Evans, Caroline; University of Sheffield, The ChELSI Institute, Department of Chemical and Biological Engineering Wright, Phillip; University of Sheffield, The ChELSI Institute, Department of Chemical and Biological Engineering; Newcastle University, School of Chemical Engineering and Advanced Materials, Faculty of Science, Agriculture & Engineering

Keywords:	Data Independent Acquisition, Broad band collision-induced dissociation (bbCID), Glycosylation, Phosphorylation, Bio-engineering

Abstract:	Rationale Post-translational modified peptides analysis by mass spectrometry (MS) remains incomplete, in part due to incomplete sampling of all peptides which is inherent to traditional data-dependent acquisition (DDA). An alternative MS approach, data-independent acquisition (DIA), enables comprehensive recording of all detectable precursor and product ions, independent of precursor intensity. The use of broad band collision-induced dissociation (bbCID), a DIA method, was evaluated for the identification of protein glycosylation and phosphorylation. Methods bbCID was applied to identify glycopeptides and phosphopeptides generated from standard proteins using a high resolution Bruker maXis 3G mass spectrometer. In bbCID, precursor and product ion spectra were obtained by alternating low and high collision energy. Precursor ions were assigned manually based on the detection of diagnostic ions specific to either glycosylation or phosphorylation. The composition of the glycan modification was resolved in the positive ion mode, whilst the level of phosphorylation was investigated in the negative ion mode. Results The results in this study demonstrate for the first time that the use of a bbCID approach is suitable for the identification of glycopeptides and phosphopeptides based on the detection of specific diagnostic and

	associated precursor ions. The novel use of bbCID in negative ion mode, allowed the discrimination of singly and multiply phosphorylated peptides based on detection of phosphate diagnostic ions. The results also demonstrate the ability of this approach to allow the identification of glycan composition in N- and O-linked glycopeptides, in positive ion mode. Conclusions We contend that bbCID is a valuable addition to the existing toolkit for PTM discovery. Moreover, this technique could be employed to direct targeted proteomics methods, particularly where there is no a priori information on glycosylation or phosphorylation status. This technique is immediately relevant to the characterisation of individual proteins or biological samples of low complexity, as demonstrated for the analysis of the glycosylation status of a therapeutic protein.

Application of the bbCID mass spectrometry approach for protein glycosylation and phosphorylation analysis

Narciso Couto[1], Liliya Davlyatova[1], Caroline A Evans[1] and Phillip C Wright[1,2]

[1] The ChELSI Institute, Department of Chemical and Biological Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD UK

[2] Now at: School of Chemical Engineering and Advanced Materials, Faculty of Science, Agriculture & Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU UK

*Correspondent author: Narciso Couto, Department of Chemical and Biological Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD, United Kingdom. E-mail: n.couto@sheffield.ac.uk

Keywords: Data Independent Acquisition, Broad band collision-induced dissociation (bbCID), Glycosylation, Phosphorylation, Bio-engineering

ABSTRACT

Rationale

Post-translational modified peptides analysis by mass spectrometry (MS) remains incomplete, in part due to incomplete sampling of all peptides which is inherent to traditional data-dependent acquisition (DDA). An alternative MS approach, data-independent acquisition (DIA), enables comprehensive recording of all detectable precursor and product ions, independent of precursor intensity. The use of broad band collision-induced dissociation (bbCID), a DIA method, was evaluated for the identification of protein glycosylation and phosphorylation.

Methods

bbCID was applied to identify glycopeptides and phosphopeptides generated from standard proteins using a high resolution Bruker maXis 3G mass spectrometer. In bbCID, precursor and product ion spectra were obtained by alternating low and high collision energy. Precursor ions were assigned manually based on the detection of diagnostic ions specific to either glycosylation or phosphorylation. The composition of the glycan modification was resolved in the positive ion mode, whilst the level of phosphorylation was investigated in the negative ion mode.

Results

The results in this study demonstrate for the first time that the use of a bbCID approach is suitable for the identification of glycopeptides and phosphopeptides based on the detection of specific diagnostic and associated precursor ions. The novel use of bbCID in negative ion mode, allowed the discrimination of singly and multiply phosphorylated peptides based on detection of phosphate diagnostic ions.

The results also demonstrate the ability of this approach to allow the identification of glycan composition in N- and O-linked glycopeptides, in positive ion mode.

Conclusions

We contend that bbCID is a valuable addition to the existing toolkit for PTM discovery. Moreover, this technique could be employed to direct targeted proteomics methods, particularly where there is no a priori information on glycosylation or phosphorylation status. This technique is immediately relevant to the characterisation of individual proteins or biological samples of low complexity, as demonstrated for the analysis of the glycosylation status of a therapeutic protein.

Introduction

Post-translational modification (PTM) of proteins, such as phosphorylation and glycosylation, control many cellular processes.[1, 2] The physiological relevance of PTMs in proteins has resulted in the use of a wide range of techniques for their analysis, from simple spectrophotometric assays to mass spectrometry (MS)-based workflows. Of these, MS-based proteomics has emerged as the preferred technique for the investigation of PTMs.[3, 4]

Typical proteomics workflows, including hyphenation of liquid chromatography and collision-induced dissociation (CID) tandem mass spectrometry (MS/MS), have been proven to be powerful strategies for PTM identification and quantification. In these workflows, data dependent MS/MS acquisition (DDA) is frequently used in which two stages are employed. The first stage involves isolation of a specific peptide ion population and the second stage involves the fragmentation of a selected precursor ion. Upon fragmentation of PTM bearing peptides, unique fragment ions, the so-called diagnostic ions, can be generated and used to unequivocally identify a given tandem mass spectrum as related to a specific modification.[5-11]

Despite major advances in MS-based proteomics, PTM analysis is still difficult. For instance, the identification of a peptide modification site is challenging as some PTMs, such as phosphorylation and N-glycosylation, are labile and thus can be lost during the isolation and fragmentation process. To overcome this, different fragmentation techniques, such as electron transfer dissociation (ETD), have been used. For instance, ETD is known to preserve the phosphate groups on phosphopeptides during the isolation and fragmentation steps.[12] Additional difficulties may arise if multiple modifications exist on the same peptide. This can be

overcome through the use of different MS approaches. One strategy would be to employ targeted analysis such as multiple reaction monitoring using a triple quadrupole or parallel reaction monitoring using an Orbitrap, to monitor the transitions of precursor/fragment ions or precursor/PTM diagnostic ions.[7, 13, 14]

All solutions discussed so far for existing challenges employ the use of a traditional DDA approach. This approach severely limits the amount of information retrieved from the complex peptide mixtures, especially for low abundance peptides. This is because DDA is biased to selection of the most abundant ions for fragmentation and analysis. To overcome this, data independent acquisition (DIA) methods such as MS^E and SWATH have been introduced as a distinct, complementary strategy to DDA.[15, 16] However, MS^E and SWATH are techniques that are specific to platforms developed by Waters and ABSciex, respectively. Broadband collision induced dissociation (bbCID) is similar DIA technique that is available to the users of the Bruker maXis platform. In contrast to MS^E and SWATH, the use of bbCID for the analysis of PTMs of proteins has been relatively unexplored by the scientific community. We have previously described the utility of the bbCID technique for PTM analysis of acetylated peptides.[17] The key principle of this technique is to simultaneously fragment all precursor ions observed in the MS survey scan. Alternating between MS survey (low energy) and bbCID (higher energy) scans enables the generation of MS scans for all precursor and their fragment ions across the mass range. The presence of acetylated peptides was determined by detection of the specific diagnostic ion (*m/z* 126.091).[17]

In principle, this approach can be applied to other PTMs that produce diagnostic ions. Therefore, in this study, we sought to extend the methodology to evaluate phosphorylated and glycosylated peptides through identification of glycan and

phosphate diagnostic ions. Here, the task is more challenging than identification of a stable PTM such as acetylation, because glyco- and phospho- modifications are known to be labile during the isolation and fragmentation process in MS. In this study, we assess the merits and limitations of using bbCID for analysis of phosphopeptides and glycopeptides, using known phosphoprotein and glycoprotein standards.

Materials and Methods

Otherwise specified, all chemical and reagents were supplied by Sigma-Aldrich (Poole Dorset, UK) with the highest purity available. All solutions were prepared with high-performance liquid chromatography (HPLC) solvents (Fisher Scientific, Loughborough, UK) on 1.5 mL Eppendorf® protein Lobind microcentrifuge tubes (Eppendorf, Hamburg, Germany). All proteins were supplied as lyophilised powders and were used without further purification. Commercially available proteins with different levels of purity were used throughout this work such as lyophilised powders of human interferon alpha 2b (IFNα2b, purity ≥ 98%), fetal bovine serum fetuin A and bovine milk β-casein (purity ≥ 95%). A standard tryptic enolase (Waters, Manchester, UK) digest mixture with four synthetic enolase analogue phosphopeptides (1 nmol enolase peptides + 1 nmol each of synthetic phosphopeptides) was also used in this work. The enzyme used throughout this work was trypsin, proteomics grade supplied from Promega (Southampton, UK).

In solution Digestion

Glycoprotein and phosphoprotein protein standards (interferon alpha 2b, fetuin A and β-casein) were prepared in 50 mM ammonium bicarbonate solution (pH 8.0).

Samples (5 µg (IFNα2b) or 20 µg in total) were reduced using 4mM dithiothreitol (60 ℃ for 30 minutes) and alkylated using 8 mM iodoacetamide (30 min incubation in the dark at room temperature) prior to addition of trypsin at a protein:trypsin ratio of 50:1. Just prior liquid chromatography and mass spectrometry, samples were diluted to 100 fmol uL^{-1} by using HPLC water incorporating 0.1 % (v/v) trifluoroacetic acid (TFA).

Liquid chromatography and mass spectrometry

Ultra-high-performance liquid chromatography (uHPLC) was performed using an Ultimate 3000 series rapid separation liquid chromatography (RSLC) system (Thermo Fisher, Hemel Hempstead, UK) connected to an UHR maXis Q-ToF 3G (Bruker, Bremen, Germany) mass spectrometer equipped with an advanced CaptiveSpray source which operated in either positive mode (glycosylation analysis) or negative mode (phosphorylation analysis). The buffers used were loading buffer (98% (v/v) water, 2 % (v/v) acetonitrile with 0.1% (v/v) TFA), buffer A (98% (v/v) water, 2 % (v/v) acetonitrile with 0.1% (v/v) FA) and buffer B (98% (v/v) acetonitrile, 2% (v/v) water with 0.1% (v/v) FA). In total, 100 fmol of peptides were loaded on the column and peptide separation was achieved by a linear gradient and each sample was run as follows specified using an 80 minute program for phosphopeptide analysis and a 120 minute programme for glycopeptide analysis. The LC analytical column was a 75 µm x15 cm packed with C18, 5 µm, 100 Å particles (LC Packings, CA, USA) and was preceded by a 300 µm i.d x 5 cm trap column packed with C18, 5 µm, 100 Å particles (LC Packings, CA, USA). On the MS instrument, the electrospray source parameters were as follows: capillary voltage +1,700 V (positive ions) and -2,000 V (negative ions), endplate offset -500 V, nebulizer gas 0.4 bar, dry

gas 6.0 L min^{-1}, and dry temperature 180 ℃. Data acquisition rate was set to 1 Hz over the mass range of 50-1700 *m/z* in both MS and bbCID scan modes.

Before sample analyses, the MS was calibrated for both positive and negative ions by using a low concentration electrospray tuning mix (Agilent, West Lothian, UK), directly infused at 300 nL min^{-1} using a Hamilton syringe matching pump (KD Scientific, MA, USA). The UHR TOF MS was first optimised for mass transmission in order to detect glycan or phosphate diagnostic fragment ions using bbCID in positive and negative ion modes respectively.[18] Detection of low mass ions was improved by balancing guide voltages as well as storage times in the collision and ion cooler cells. In addition, pre-pulsed times were co-adjusted to enable improved transmission and sensitivity for low *m/z* ions. For positive ions, the transfer parameters were: funnel 1 RF 400.0 Vpp, ISCID energy, 0.0 eV, multiple RF 400 Vpp. In the quadrupole device, quadrupole ion energy, 5.0 eV and low mass for transmission was set at 100.00 *m/z*. The collision energy applied in the collision cell was set to 10.0 eV with a collision RF of 600 Vpp. In the cooling cell, the values of the ion cooler RF, 400.0 Vpp, transfer time 65 µs, pre pulse storage, 5.0 µs were set as listed. In negative ion mode, the applied voltages were the same but automatically adjusted upon switching the polarity from the positive mode.

Alternate collision energy was set as a scan mode to perform alternating low and high energy in bbCID to provide survey scan information on peptide precursors and fragments within the full mass range of 50-1700 *m/z*. bbCID was set to alternate between 10-100 eV for evaluation. The acquisition factors were set to 2 for high collision energy and 1 for low collision energy as the default. The optimal high bbCID energy was established to be between 35-45 eV for both glycoprotein and phosphoprotein samples, providing balance of low and high *m/z* values across the

MS spectra by use of collision sweeping. Compass DataAnalysis v4.1 software (Bruker, Bremen, Germany) was used to extract diagnostic ion information and to associate fragment ions to their respective precursor ions.

Results and Discussion

Representative glycoprotein and phosphoprotein samples were used to assess the utility of the bbCID method for PTM analysis in positive and negative ion modes, respectively. No pre-enrichment steps were included.

Glycopeptide analysis

Commercially available bovine fetuin A and human interferon alpha 2b (IFNα2b) were used as representative glycoproteins to evaluate asparagine N-linked and serine and threonine O-linked oligosaccharide modifications. Fetuin A was selected as an example of a highly glycosylated protein containing sialylated *N*-linked and *O*-linked glycans that exhibit heterogeneity on the glycans attached to the same amino acid.[8-11, 19-21] IFNα2b was selected as an example of a therapeutic protein, in which glycosylation is essential for biological activity and efficacy.

Analysis of tryptic digests of fetuin and IFNα2b was performed by bbCID without glycopeptide enrichment, in order to evaluate the utility of this method with respect to glycopeptide discovery. Glycosylated peptides were assigned based on the presence of specific glycan diagnostic ions.[8-11, 15] The diagnostic ions included the oxonium ions N-acetylglucosamine (HexNAc) at *m/z* 204.079, hexose+N-acetylglucosamine (Hex-HexNAc) at *m/z* 366.132, $HexNAc_2$ at *m/z* 407.159 rarely observed in MS/MS spectra, Hex_2-HexNAc at *m/z* 528.185, N-acetylneuraminic acid (SA) and SA-H_2O at *m/z* 292.095 and *m/z* 274.095, respectively. [8-11, 15] Upon

identification of the diagnostic ions in the high energy bbCID spectrum, the precursor *m/z* value and the charge state were used for glycopeptide identification. The glycan composition was manually interpreted and associated to a peptide. Representative mass spectra (low and high bbCID energy) for the N-glycosylated peptide, LCPDCPLLAPLN(152)DSR, of fetuin A are shown in Fig. 1. The high energy (45 eV) bbCID mass spectrum, (Fig. 1B), is dominated by the presence of saccharide diagnostic ions HexNAc (*m/z* 204.089), SA (*m/z* 292.104 and 274.094), Hex-HexNAc (*m/z* 366.140) and Hex-HexNAc-SA (*m/z* 657.235). In Table 1, all identified N-glycosylated peptides of fetuin A are shown, indicating that glycosylation variants could be distinguished.

The potential presence of several glycans attached to the same protein site makes glycopeptide analysis more complex and technically challenging. Nevertheless, using this approach of alternating low and high energy CID, it was possible to establish a relationship between glycan composition and the *m/z* value of the precursor peptide ions. The fragmentation of the oligosaccharide requires much less energy than the fragmentation of the peptide backbone, resulting in the glycan fragments dominating the high energy bbCID spectrum (Fig. 1B), to the detriment of peptide sequence information, but providing information on the glycan content. These findings are similar to that of Gilar et al.[20] using MS^E and to that of Geromanos et al.[22] where a 'fragment all' MS approach using a Synapt G2 platform was applied.

Figure 1 goes here

Table 1 goes here

The bbCID method was also used to characterise the glycan heterogeneity on recombinant IFNα2b. This protein does not possess the motif required for N-glycosylation and only the threonine (T-106) position of this protein is O-glycosylated. Previous sequence analysis by MS of this protein identified T-106 as the only O-glycosylation site, with limited peptide sequence information.[23, 24] Representative spectra (low and high bbCID energy) of the IFNα2b glycopeptide, FYTELYQQLNDLEACVIQGVGVT(106)ETPLMK, are shown in Fig. 2 and all identified glycoforms are shown in Table 2.

Figure 2 goes here

Table 2 goes here

In Fig. 2B, the presence of saccharide diagnostic ions were identified for HexNAc (*m/z* 204.089), SA (*m/z* 292.104) and its dehydrated form (SA-H_2O) (*m/z* 274.094), Hex-HexNAc (*m/z* 366.140), HexNAc-SA (*m/z* 495.182), Hex-HexNAc-SA (*m/z* 657.228) and Hex-HexNAc-SA_2 (*m/z* 948.329). Manual analysis of the corresponding low energy spectrum revealed the presence of distinct glycoforms (Fig. 2A). Hex-HexNAc-SA_2 modification was seen to be associated with the triply and quadruply protonated precursor ions of the peptide FYTELYQQLNDLEACVIQGVGVT(106)ETPLMK, at *m/z* 1436.326 and *m/z* 1077.500, respectively. In the same spectrum, a different glycan, Hex_2-$HexNAc_2$-SA2, was seen to be associated with the triply and quadruply protonated precursor ions of the same peptide sequence at *m/z* 1557.787 and *m/z* 1168.675 (Fig. 2A). The presence of two glycoforms in the same retention time suggests that one glycoform may derive from the other during gas-phase processes.

In addition, some ions present in Fig. 2B such as the doubly protonated ions, at *m/z* 1680.833 and *m/z* 1781.870, confirmed the assignment and were associated with the aglycosylated and the HexNAc forms. The presence of the peptide with attached HexNAc suggests that this monosaccharide is the one directly attached to the side chain of the threonine residue. It has been observed previously that cleavage between the first and second monosaccharides adjacent to the peptide is a dominant fragmentation route.[25] Further glycosylation isomers were also identified in small amounts, which is in agreement with results previously reported for this protein (Table 2). Importantly, our technique applied to INFα2b analysis revealed no additional glycosylation sites. This suggests a potential application of our technique to assess the glycosylation status of glycoproteins, especially for the analysis of therapeutic proteins. In addition, our technique is complementary to analysis of intact glycoproteins where extracting information regarding glycosite is not possible.

The bbCID method is thus of value to glycoengineering for evaluation of the glycosylation status.[26, 27] Our technique of using bbCID for glycoprotein analysis will also have value in assessing glycosylated protein production using recombinant systems, such as *Escherichia coli* engineered to produce specific glycosylated proteins.[28] Introduction of glycosylation machinery from other bacteria such as *Actinobacillus pleuropneumoniae* in *E. coli* can lead to engineering of novel (non-traditional) consensus glycosylation sequence binding sites (sequons). Examples include NXA, NXG, NXD, NXV AND QXT and QXA for N-linked glycosylation or SXT and SXV for novel O-linked glycosylation consensus sites.[29] Our technique is also compatible with the recently reported method of sequential fragment ion filtering and endoglycosidase assisted identification of intact glycopeptide glycoforms.[30] These examples demonstrate the value of our technique in discriminating glycoforms.

However, the applicability of bbCID for the analysis of glycopeptides with multiple glycosylation sites, or a complex system, where multiple glycopeptides can fall into a DIA window and be fragmented together, is still challenging. The major limitation to bbCID and similar DIA approaches such MS^E and SWATH is interference due to chimeric peptides. This interference persists even with the implementation of narrow m/z windows for peptide fragmentation. To overcome this, multiplexing narrowband and wideband DIA acquisitions in a single analytical workflow to generate highly selective, product ion spectra has been applied.[31] This approach may prove useful when applying bbCID approach to the evaluation of PTMs in complex systems.

Phosphopeptide analysis

Commercially available enolase spiked with synthetic phosphopeptide enolase analogues and bovine β-casein were used as representative phosphoproteins to evaluate phosphorylation levels, in negative ion mode. We used the negative ion mode to investigate the diagnostic ions $[PO_3]^-$ and $[H_2PO_3]^-$ at *m/z* 78.938 and *m/z* 96.968, respectively, because phosphate diagnostic ions are typically not observed in the positive ion mode.

The bbCID MS acquisition method was optimised using a standard peptide mixture of tryptic enolase peptides additionally containing equimolar amount of four synthetic enolase phosphopeptides, designed to incorporate a phosphoserine (HLADL<u>S</u>K), a phosphotyrosine (NVPL<u>Y</u>K), a phosphothreonine (VNQIG<u>T</u>LSESIK) and a double phosphoserine (VNQIGTL<u>S</u>E<u>S</u>IK). Detailed information regarding these synthetic phosphopeptides is presented in Table 3. The selectivity of the bbCID method to detect and differentiate phosphorylated from non-phosphorylated peptides was evaluated using this mixture. Representative bbCID mass spectra for a mono-

phosphorylated (NVPLYK) and di-phosphorylated (VNQIGTLSESIK) enolase peptides, at low and high bbCID energy, are shown in Fig. 3 and 4 respectively.

Table 3 goes here

Figures 3 and 4 go here

In Fig. 3A, under low energy regime, the singly deprotonated ion NVPLYK at *m/z* 811.352 is the dominant specie present in the mass spectrum. The phosphate diagnostic anions, $[PO_3]^-$ and $[H_2PO_3]^-$ at *m/z* 78.938 and *m/z* 96.968 respectively, were better observed at optimal bbCID collision energy of 35 eV (Figure 3B). For the di-phosphorylated peptide, VNQIGTLSESIK, both singly and doubly deprotonated ions are observed at *m/z* 1146.585 and *m/z* 722.789, where the doubly charged species is largely dominant (Fig. 4A). Phosphate diagnostic ions were identified using a bbCID collision energy of 35eV (Fig. 4B). However, additional information regarding the number of phosphate residues was obtained only upon increasing the bbCID collision energy to 45eV (Fig. 4C). Here, consecutive neutral loss of 2 phosphoric acid (H_3PO_3) molecules and phosphate related diagnostic ions, $[HP_2O_6]^-$ and $[H_3P_2O_7]^-$ at *m/z* 157.927 and *m/z* 176.937 respectively, were indicative of a doubly phosphorylated ion.

Further analysis of the equimolar mixture of phosphorylated and non-phosphorylated peptides demonstrates that these phosphopeptides are more hydrophobic than non-phosphorylated peptides, due to their higher retention times (Table 3). The presence of a phospho- group is generally expected to decrease the hydrophobicity of a peptide and lower their retention time during reverse-phase chromatography. However, an increased phosphopetide hydrophobicity has also been previously reported and attributed to polar interactions between the phospho- group and side

chains of neighbouring aminoacids.[32] This is consistent with the results observed in the present study.

Extracted ion chromatograms revealed similar ion populations of phosphopeptides and the non-phosphorylated analogues (Table 3). This confirms that in negative ion mode there are no significant changes in ionization efficiency for phosphopeptides. This implies that decreased detection of phosphopeptides in complex mixtures is not due to ion suppression as commonly believed, but due to the low abundance of phosphopeptides and/or their existence at sub-stoichiometric levels as discussed elsewhere.[33]

We extended this approach to evaluate β-casein, a multi-phosphorylated protein containing phosphate groups attached to a serine residue at positions 30, 32, 33, 34 and 50. The low and high bbCID energy mass spectra for the phosphopeptide FQS(50)EEQQQTEDELQDK are shown in Fig. 5A and 5B. In Fig. 5A, the doubly deprotonated ion dominates the low energy mass spectrum at *m/z* 1029.384. The low and high bbCID energy spectra for the phosphopeptide ELEELNVPGEIVES(30)LSSS(32-34)EESITR are shown in Fig. 6A and 6B. The triply deprotonated ion is the dominant charge state in the low energy mass spectrum at *m/z* 1039.726. A comparison of Fig. 5A and 6A suggests that there may be a relationship between peptide size and the dominant charge state, in negative ion mode.

Figure 5 here

Figure 6 here

In Figure 5B and 6B, H_3PO_4 neutral loss and the presence of the diagnostic ions is shown in the high energy bbCID spectra of β-casein phosphopeptides, FQS(50)EEQQQTEDELQDK and ELEELNVPGEIVES(30)LSSS(32-34)EESITR. The multiple phosphorylated peptide, ELEELNVPGEIVES(30)LSSS(32-34)EESITR, is associated with a more complex pattern of phosphate diagnostic ions incorporating phosphate dimerization products. The number of phosphate groups associated with the peptide is also extracted from the number of H_3PO_4 neutral losses.

In general, the ESI negative ion mode is less sensitive compared to the positive ion mode. Moreover, sequence information from negative ion MS/MS spectra of peptides is more difficult to extract, because an evenly spread backbone fragmentation is not observed. Fragmentation of peptide anions is often governed by neutral loss reactions involving side chains rather than peptide backbone fragmentation. Therefore, the body of experimental information on MS/MS spectra of peptide anions is still much smaller compared to positive ions.[5, 34-37] However, our results suggest that the higher energy bbCID approach proves to be useful for the assignment of phosphopeptides through the detection of phosphate diagnostic ions and H_3PO_4 neutral losses. A key feature of the data is that the presence of phosphate diagnostic ions and the level of neutral loss of H_3PO_4 from precursor ions enables assignment of the number of phosphate groups present in the phosphopeptide. Our results demonstrate that negative ion mode bbCID is practical and can be used as a phosphopeptide discovery and characterisation tool, which allows further characterisation of selected phosphopeptides using enrichment techniques such as sequential elution from immobilised metal ion affinity chromatography (SIMAC).[38] SIMAC employs enrichment and separation of mono- and multi-phosphorylated peptides. The phosphopeptide fractions are then analysed

separately by MS leading to improved phosphoproteomic coverage. We propose that the value of the bbCID approach is in identifying potential phosphopeptide precursors for further evaluation, particularly where the phosphorylation status (mono or multiple) is unknown. This has the potential to boost identification rates, by providing target precursor and product transitions for reaction monitoring based approaches.

It is interesting to note that phosphorylation and glycosylation can result in missed cleavages. In our results, missed cleavages can be especially seen in the case of glycosylated peptides (Table 1 and 2). In the case of glycosylation, missed cleavages following proteolysis has been reported with trypsin and the use of additional enzymes such as pepsin and thermolysin increases their discovery rate.[39] Missed cleavages have been reported in the literature also for phosphorylated peptides, not only with trypsin but also other proteases such as Lys-C, Asp-N, Glu-C, and chymotrypsin.[40]

Conclusion

In this study, bbCID has been used to successfully investigate glycopeptides and phosphopeptides in samples of low complexity using positive and negative ion mode, respectively. In this method, the concomitant detection of the *m/z* value of the precursor and the diagnostic ions unique to glycan and phosphate modifications, coupled with the use of high resolution MS allows the unequivocal identification of glycopeptides and phosphopeptides. This method is applicable for the investigation of relatively simple protein mixtures without a PTM enrichment step. The extension of bbCID for the evaluation of glycosylation or phosphorylation in PTM enriched

biological peptides is promising but limited by the current non-availability of algorithms for automated MS analysis.

Acknowledgements

Phillip Wright and Caroline Evans acknowledge the financial support from Engineering and Physical Sciences Research Council, ChELSI initiative (EP/E036252/1), Cancer Research UK (DS). Phillip Wright and Narciso Couto acknowledge the financial support from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement number 308518 (project CyanoFactory). Phillip Wright acknowledges support from the Biotechnology and Biological Sciences Research Council under the Bioprocess Research Industry Club (BRIC) (BB/K011200/1)

References

[1] Y.L. Deribe, T. Pawson, I. Dikic. Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol*. **2010,** *17*, 666-672.

[2] J.L. Parker, R.C. Lowry, N.A. Couto, P.C. Wright, G.P. Stafford, J.G. Shaw. Maf-dependent bacterial flagellin glycosylation occurs before chaperone binding and flagellar T3SS export. *Mol. Microbiol*. **2014,** *92*, 258-272.

[3] N. Chicooree, R.D. Unwin, J.R. Griffiths. The application of targeted mass spectrometry-based strategies to the detection and localization of post-translational modifications. *Mass Spectrom. Rev*. **2015,** *34*, 595-626.

[4] S. Doll, A.L. Burlingame. Mass spectrometry-based detection and assignment of protein posttranslational modifications. *ACS Chem. Biol*. **2014,** *10*, 63-71.

[5] M. Edelson-Averbukh, R. Pipkorn, W.D. Lehmann. Phosphate group-driven fragmentation of multiply charged phosphopeptide anions. Improved recognition of

peptides phosphorylated at serine, threonine, or tyrosine by negative ion electrospray tandem mass spectrometry. *Anal. Chem.* **2006,** *78*, 1249-1256.

[6] M.B. Trelle, O.N. Jensen. Utility of immonium ions for assignment of ε-N-acetyllysine-containing peptides by tandem mass spectrometry. *Anal. Chem*. **2008,** *80*, 3422-3430.

[7] J.R. Griffiths, N. Chicooree, Y. Connolly, M. Neffling, C.S. Lane, T. Knapman, D.L. Smith. Mass spectral enhanced detection of UBLs using SWATH acquisition: MEDUSA-simultaneous quantification of SUMO and ubiquitin-derived isopeptides. *J. Am. Soc. Mass Spectrom.* **2014,** *25*, 767-777.

[8] M.J. Huddleston, M.F. Bean, S.A. Carr. Collisional fragmentation of glycopeptides by electrospray ionization LC/MS and LC/MS/MS: methods for selective detection of glycopeptides in protein digests. *Anal. Chem.* **1993**, 65, 877-884.

[9] A.P. Hunter, D.E. Games. Evaluation of glycosylation site heterogeneity and selective identification of glycopeptides in proteolytic digests of bovine alpha 1-acid glycoprotein by mass spectrometry. *Rapid Commun. Mass Spectrom*. **1995**, 9, 42-56.

[10] J.J. Conboy, J. Henion. The determination of glycopeptides by liquid chromatography/mass spectrometry with collision-induced dissociation. *J. Am. Soc. Mass Spectrom*. 1992, 3, 804-814.

[11] A. Halim, U. Westerlind, C. Pett, M. Schorlemer, U. Rüetschi, G. Brinkmalm, C. Sihlbom, J. Lengqvist, G.r. Larson, J. Nilsson. Assignment of saccharide identities through analysis of oxonium ion fragmentation profiles in LC-MS/MS of glycopeptides. *J. Proteome Res*. **2014,** *13*, 6024-6032.

[12] A. Chi, C. Huttenhower, L.Y. Geer, J.J. Coon, J.E. Syka, D.L. Bai, J. Shabanowitz, D.J. Burke, O.G. Troyanskaya, D.F. Hunt. Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A*. **2007,** *104*, 2193-2198.

[13] R.D. Unwin, J.R. Griffiths, A.D. Whetton. A sensitive mass spectrometric method for hypothesis-driven detection of peptide post-translational modifications: multiple reaction monitoring-initiated detection and sequencing (MIDAS). *Nat. Protoc*. **2009,** *4*, 870-877.

[14] H. Tang, H. Fang, E. Yin, A.R. Brasier, L.C. Sowers, K. Zhang. Multiplexed parallel reaction monitoring targeting histone modifications on the QExactive mass spectrometer. *Anal. Chem*. **2014,** *86*, 5526-5534.

[15] H.L. Röst, G. Rosenberger, P. Navarro, L. Gillet, S.M. Miladinović, O.T. Schubert, W. Wolski, B.C. Collins, J. Malmström, L. Malmström. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat. Biotechnol*. **2014,** *32*, 219-223.

[16] R.S. Plumb, K.A. Johnson, P. Rainville, B.W. Smith, I.D. Wilson, J.M. Castro-Perez, J.K. Nicholson. UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Commun. Mass Spectrom*. **2006,** *20*, 1989-1994.

[17] C.A. Evans, S.Y. Ow, D.L. Smith, B.M. Corfe, P.C. Wright. Application of the CIRAD mass spectrometry approach for lysine acetylation site discovery. *Protein Acetylation: Methods Protoc*. **2013**, 13-23.

[18] S.Y. Ow, J. Noirel, M. Salim, C. Evans, R. Watson, P. Wright. Balancing robust quantification and identification for iTRAQ: Application of UHR-ToF MS. *Proteomics*. **2010,** *10*, 2205-2213.

[19] D. F. Zielinska, F. Gnad, J. R. Wisniewski, M. Mann. Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell*. **2010**, 141, 897-907.

[20] M. Gilar, Y. Q. Yu, J. Ahn, H. Xie, H. Han, W. Ying, X. Qian. Characterization of glycoprotein digests with hydrophilic interaction chromatography and mass spectrometry. *Anal. Biochem.* **2011,** *417*, 80-88.

[21] W. Ding, H. Nothaft, C.M. Szymanski, J. Kelly. Identification and quantification of glycoproteins using ion-pairing normal-phase liquid chromatography and mass spectrometry. *Mol. Cell. Proteomics*. **2009,** *8*, 2170-2185.

[22] S.J. Geromanos, J.P. Vissers, J.C. Silva, C.A. Dorschel, G.Z. Li, M.V. Gorenstein, R.H. Bateman, J.I. Langridge. The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics*. **2009,** *9*, 1683-1695.

[23] M. Loignon, S. Perret, J. Kelly, D. Boulais, B. Cass, L. Bisson, F. Afkhamizarreh, Y. Durocher. Stable high volumetric production of glycosylated human recombinant IFNalpha2b in HEK293 cells. *BMC Biotechnol*. **2008,** *8*, 65.

[24] G.R. Adolf, I. Kalsner, H. Ahorn, I. Maurer-Fogy, K. Cantell. Natural human interferon-α2 is O-glycosylated. *Biochem. J*. **1991,** *276*, 511-518.

[25] H. Jiang, H. Desaire, V.Y. Butnev, G.R. Bousfield. Glycoprotein profiling by electrospray mass spectrometry. *J. Am. Soc. Mass Spectrom*. **2004,** *15*, 750-758.

[26] A. Beck, F. Debaene, H. Diemer, E. Wagner-Rousset, O. Colas, A.V. Dorsselaer, S. Cianférani. Cutting-edge mass spectrometry characterization of originator, biosimilar and biobetter antibodies. *J. Mass Spectrom*. **2015,** *50*, 285-297.

[27] V. Dotz, R. Haselberg, A. Shubhakar, R.P. Kozak, D. Falck, Y. Rombouts, D. Reusch, G.W. Somsen, D.L. Fernandes, M. Wuhrer. Mass spectrometry for glycosylation analysis of biopharmaceuticals. *TrAC Trends in Anal. Chem*. **2015,** *73*, 1-9.

[28] J. Pandhal, L. Woodruff, S. Jaffe, P. Desai, S.Y. Ow, J. Noirel, R.T. Gill, P.C. Wright. Inverse metabolic engineering to improve Escherichia coli as an N-glycosylation host. *Biotechnol. Bioeng*. **2013,** *110*, 2482-2493.

[29] A. Naegeli, C. Neupert, Y.Y. Fan, C.W. Lin, K. Poljak, A.M. Papini, F. Schwarz, M. Aebi. Molecular analysis of an alternative N-glycosylation machinery by functional transfer from Actinobacillus pleuropneumoniae to Escherichia coli. *J. Biol. Chem*. **2014,** *289*, 2170-2179.

[30] Z. Yu, X. Zhao, F. Tian, Y. Zhao, Y. Zhang, Y. Huang, X. Qian, W. Ying. Sequential fragment ion filtering and endoglycosidase-assisted identification of intact glycopeptides. *Anal. Bioanal. Chem*. **2017**, 1-11.

[31] B.J. Williams, S.J. Ciavarini, C. Devlin, C, S.M. Cohn, R. Xie, J.P. Vissers, L.B. Martin, A. Caswell, J.I. Langridge, S.J. Geromanos. Multi-mode acquisition (MMA): An MS/MS acquisition strategy for maximizing selectivity, specificity and sensitivity of DIA product ion spectra. *Proteomics*, **2016**, *16*, 2284-2301.

[32] J. Kim, K. Petritis, Y. Shen, D.G. Camp II, R.J. Moore, R.D. Smith. Phosphopeptide elution times in reversed-phase liquid chromatography. *J. Chromatogr. A.* **2007**, 1172, 9-18.

[33] H. Steen, J.A. Jebanathirajah, J. Rush, N. Morrice, M.W. Kirschner. Phosphorylation analysis by mass spectrometry myths, facts, and the consequences for qualitative and quantitative measurements. *Mol. Cell. Proteomics*. **2006,** *5*, 172-181.

[34] T. Tran, T. Wang, S. Hack, P. Hoffmann, J.H. Bowie. Can collision-induced negative-ion fragmentations of [M-H]-anions be used to identify phosphorylation sites in peptides? *Rapid Commun. Mass Spectrom.* **2011,** *25*, 3537-3548.

[35] A. Tholey, J. Reed, W.D. Lehmann. Electrospray tandem mass spectrometric studies of phosphopeptides and phosphopeptide analogues. *J. Mass Spectrom*. **1999,** *34*, 117-123.

[36] M. Edelson-Averbukh, R. Pipkorn, W.D. Lehmann. Analysis of protein phosphorylation in the regions of consecutive serine/threonine residues by negative ion electrospray collision-induced dissociation. Approach to pinpointing of phosphorylation sites. *Anal. Chem.* **2007,** *79*, 3476-3486.

[37] H. Steen, B. Küster, M. Mann. Quadrupole time-of-flight versus triple-quadrupole mass spectrometry for the determination of phosphopeptides by precursor ion scanning. *J. Mass Spectrom*. **2001,** *36*, 782-790.

[38] T.E. Thingholm, M.R. Larsen. Sequential elution from IMAC (SIMAC): An efficient method for enrichment and separation of mono-and multi-phosphorylated peptides. *Phospho-Proteomics: Methods Protoc.* **2016**, 147-160.

[39] R. Chen, X. Jiang, D. Sun, G. Han, F. Wang, M. Ye, L. Wang, H. Zou. Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *Proteome Res*. **2009,** *8*, 651-661.

[40] P. Giansanti, T.T. Aye, H. van den Toorn, M. Peng, B. van Breukelen, A.J. Heck. An augmented multiple-protease-based human phosphopeptide atlas. *Cell Rep*. **2015,** *11*, 1834-1843.

Captions

Table 1 – Glycoform profiles of tryptic N-glycopeptides (N-99, N-156 and N-176) known to exist in bovine fetuin A where the consensus sequence (N-X-S/T, X not proline) exists.

Figure 1 – Representative example of the low and high energy (MS, bbCID MS) mass spectra of the tryptic fetuin A glycopeptide with sequence LCPDCPLLAPLNDSR. Underlined amino acids are modified; cysteine is artificially carbamidomethylated and asparagine (position 156) is naturally glycosylated with a reducing sugar ($HexNAc_4$-Hex_5-SA_2)+H_2O. A) several charged states (4, 5 and 6) are observed in the low energy MS with base peak at *m/z* 793.795 corresponding to the quintuple protonated ion. B) bbCID spectrum where glycan diagnostic ions are highlighted. The symbols, square, circle and diamond were used to represent mannose, N-acetylglucosamine and sialic acid, respectively.

Table 2 – Glycoform profile of tryptic IFNα2b O-glycopeptide at T-106. Different charge states and the existence of missed cleavages were observed.

Figure 2 – Low and high energy (MS, bbCID MS) mass spectra of the tryptic IFNα2b glycopeptide FYTELYQQLNDLEACVIQGVGVTETPLMK. Underlined amino acids are modified; cysteine is artificially carbamidomethylated and threonine (position 106) is naturally glycosylated. A) Low energy mass spectrum (5 eV) showing the triply and quadruple protonated ions at *m/z* 1436.326 and *m/z* 1077.500 corresponding to the glycan Hex-HexNAc-SA_2 attached to T(106). The triply and quadruple protonated ions at *m/z* 1557.787 and *m/z* 1168.675 corresponding to the same peptide with a more complex glycan Hex_2-$HexNAc_2$-SA_2 attached to T(106) is also shown. B) High energy mass spectrum (45 eV) showing the diagnostic ions

typical of glycan Hex-HexNAc-SA_2 and glycan Hex_2-$HexNAc_2$-SA_2 attached to T(106). The aglycosylayed *m/z* 1680.833 corresponding to the doubly protonated ions as well as the ion containing only HexNAc at *m/z* 1781.87 are also indicated. The symbols, square, circle and diamond were used to represent mannose, N-acetylglucosamine and sialic acid, respectively.

Table 3 – The amino acid sequence of enolase tryptic peptides and their synthetic phosphopeptide cognates is shown. The phosphosite, the *m/z* values and retention time (extracted ion chromatogram) are annotated. For peptides indicated on Table 3, the extracted ion chromatogram is highlighted.

Figure 3 – Low and high energy (MS, bbCID MS) mass spectra of T2p, a phosphotyrosine peptide (NVPLYK, *m/z* 811.352) in negative ion mode. Underlined amino acid tyrosine is phosphorylated. In Figure 1A and 1B, the low and high energy scans mass spectra, at 5 eV ion-source collision energy and 35 eV bbCID collision energy, are shown respectively.

Figure 4 – Low and high energy (MS, bbCID MS) mass spectra of T3pp, a phosphoserine peptide (VNQIGTLSESIK, *m/z* 722.789) in negative mode. Underlined amino acids are phosphorylated. In Figure 4A, the MS scan, at 5 eV ion-source collision energy, shows the dominance of the doubly deprotonated peptide anion. In Figure 4B and 4C, bbCID MS at 35 eV and 45eV collision energy respectively, are shown.

Figure 5 – bbCID MS scan at 5 eV ion-source collision energy and 45eV collision energy, 5A and 5B, respectively for β-casein doubly deprotonated phosphoserine peptide (FQS(50)EEQQQTEDELQDK, *m/z* 1029.383) in negative ion mode. Underlined amino acid is phosphorylated.

Figure 6 - bbCID MS scan at 5 eV ion-source collision energy and 45eV collision energy, 6A and 6B, for RELEELNVPGEIVES(30)LS(32)S(33)S(34)EESITR at *m/z* 1039.392, a tetra deprotonated phosphorylated peptide. Underlined amino acids are phosphorylated.

Tables

Table 1

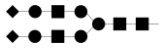
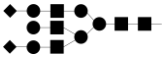
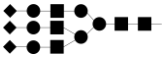
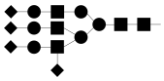
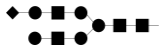
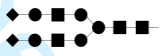
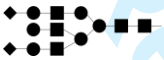
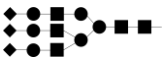
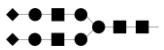
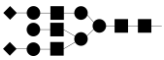
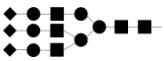
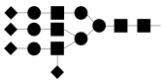
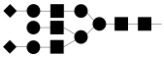
Peptide sequence	Missed Cleavages	Glycan composition	Glycan symbol nomenclature	Experimental m/z; charge state	Theoretical m/z; charge state	Error (ppm)	RT (min)
RPTGEVYDIEIDTLETTCHVLDPTPLAN(99)CSVR	1	Hex_5-$HexNAc_4$-SA_2		1176.101 $[M+5H]^{5+}$	1176.114 $[M+5H]^{5+}$	11	95.7
				980.252 $[M+6H]^{6+}$	980.263 $[M+6H]^{6+}$	11	
				840.363 $[M+7H]^{7+}$	840.369 $[M+7H]^{7+}$	7	
		Hex_6-$HexNAc_5$-SA_2		1249.126 $[M+5H]^{5+}$	1249.140 $[M+5H]^{5+}$	11	95.7
				1041.107 $[M+6H]^{6+}$	1041.118 $[M+6H]^{6+}$	11	
				892.529 $[M+7H]^{7+}$	892.531 $[M+7H]^{7+}$	2	
		Hex_6-$HexNAc_5$-SA_3		1307.349 $[M+5H]^{5+}$	1307.358 $[M+5H]^{5+}$	7	95.7
				1089.624 $[M+6H]^{6+}$	1089.634 $[M+6H]^{6+}$	9	
				934.109 $[M+7H]^{7+}$	934.116 $[M+7H]^{7+}$	7	
		Hex_6-$HexNAc_5$-SA_4		1365.571 $[M+5H]^{5+}$	1365.578 $[M+5H]^{5+}$	5	95.8
				1138.148 $[M+6H]^{6+}$	1138.150 $[M+6H]^{6+}$	2	
				975.715 $[M+7H]^{7+}$	975.701 $[M+7H]^{7+}$	-14	
LCPDCPLLAPLN(156)DSR	0	Hex_5-$HexNAc_4$-SA		1218.806 $[M+3H]^{3+}$	1218.844 $[M+3H]^{3+}$	31	85.0
				914.403 $[M+4H]^{4+}$	914.385 $[M+4H]^{4+}$	-20	
				731.745 $[M+5H]^{5+}$	731.709 $[M+5H]^{5+}$	-49	
		Hex_5-$HexNAc_4$-SA_2		1315.868 $[M+3H]^{3+}$	1315.876 $[M+3H]^{3+}$	6	85.5
				987.152 $[M+4H]^{4+}$	987.159 $[M+4H]^{4+}$	7	
				789.920 $[M+5H]^{5+}$	789.928 $[M+5H]^{5+}$	10	
		Hex_5-$HexNAc_4$-SA_2+H_2O		991.986 $[M+4H]^{4+}$	991.990 $[M+4H]^{4+}$	4	95.1
				793.790 $[M+5H]^{5+}$	793.799 $[M+5H]^{5+}$	11	
				661.659 $[M+6H]^{6+}$	661.649 $[M+6H]^{6+}$	-15	
		Hex_6-$HexNAc_5$-SA_2		1437.581 $[M+3H]^{3+}$	1437.587 $[M+3H]^{3+}$	4	95.0
				1078.432 $[M+4H]^{4+}$	1078.442 $[M+4H]^{4+}$	9	
				862.911 $[M+5H]^{5+}$	862.955 $[M+5H]^{5+}$	51	
		Hex_6-$HexNAc_5$-SA_3		1534.608 $[M+3H]^{3+}$	1534.618 $[M+3H]^{3+}$	7	95.5
				1151.212 $[M+4H]^{4+}$	1151.216 $[M+4H]^{4+}$	3	
				921.164 $[M+5H]^{5+}$	921.174 $[M+5H]^{5+}$	11	
KLCPDCPLLAPLN(156)DSR	1	Hex_5-$HexNAc_4$-SA_2		1358.571 $[M+3H]^{3+}$	1358.574 $[M+3H]^{3+}$	2	74.1
				1019.181 $[M+4H]^{4+}$	1019.182 $[M+4H]^{4+}$	1	
				815.528 $[M+5H]^{5+}$	815.547 $[M+5H]^{5+}$	23	
		Hex_6-$HexNAc_5$-SA_2		1480.273 $[M+3H]^{3+}$	1480.285 $[M+3H]^{3+}$	8	81.9
				1110.461 $[M+4H]^{4+}$	1110.465 $[M+4H]^{4+}$	4	
				888.525 $[M+5H]^{5+}$	888.574 $[M+5H]^{5+}$	55	
		Hex_6-$HexNAc_5$-SA_3		1577.299 $[M+3H]^{3+}$	1577.317 $[M+3H]^{3+}$	11	82.5
				1183.229 $[M+4H]^{4+}$	1183.239 $[M+4H]^{4+}$	8	
				946.788 $[M+5H]^{5+}$	946.793 $[M+5H]^{5+}$	5	
		Hex_6-$HexNAc_5$-SA_4		1674.350 $[M+3H]^{3+}$	1674.349 $[M+3H]^{3+}$	-1	94.5
				1256.006 $[M+4H]^{4+}$	1256.013 $[M+4H]^{4+}$	6	
				1005.006 $[M+5H]^{5+}$	1005.012 $[M+5H]^{5+}$	6	
VVHAVEVALATFNAESN(176)GSYLQLVEISR	0	Hex_6-$HexNAc_5$-SA_2		1118.094 $[M+5H]^{5+}$	1118.101 $[M+5H]^{5+}$	6	92.7
				931.901 $[M+6H]^{6+}$	931.919 $[M+6H]^{6+}$	19	

Table 2

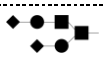
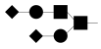
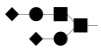
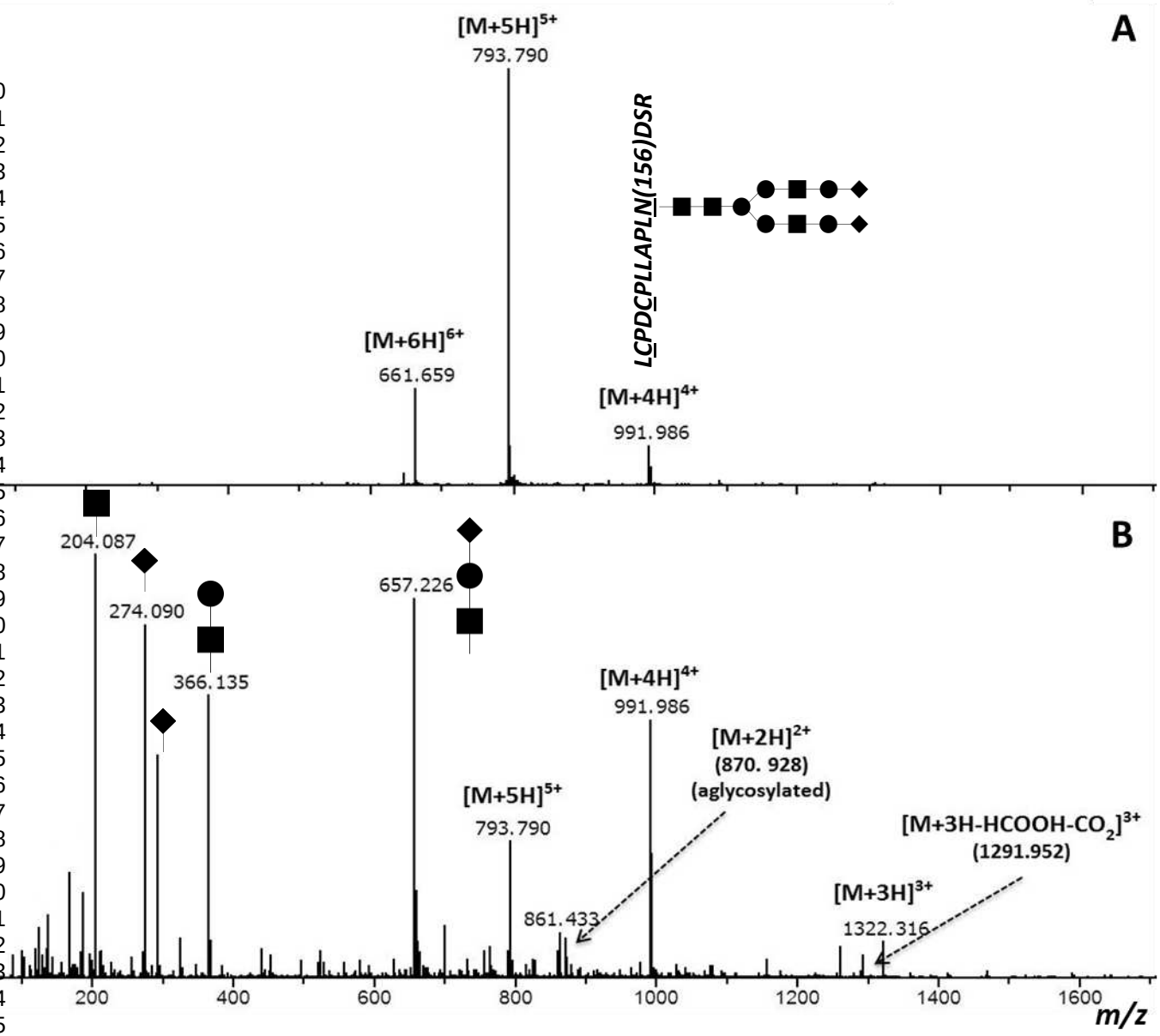
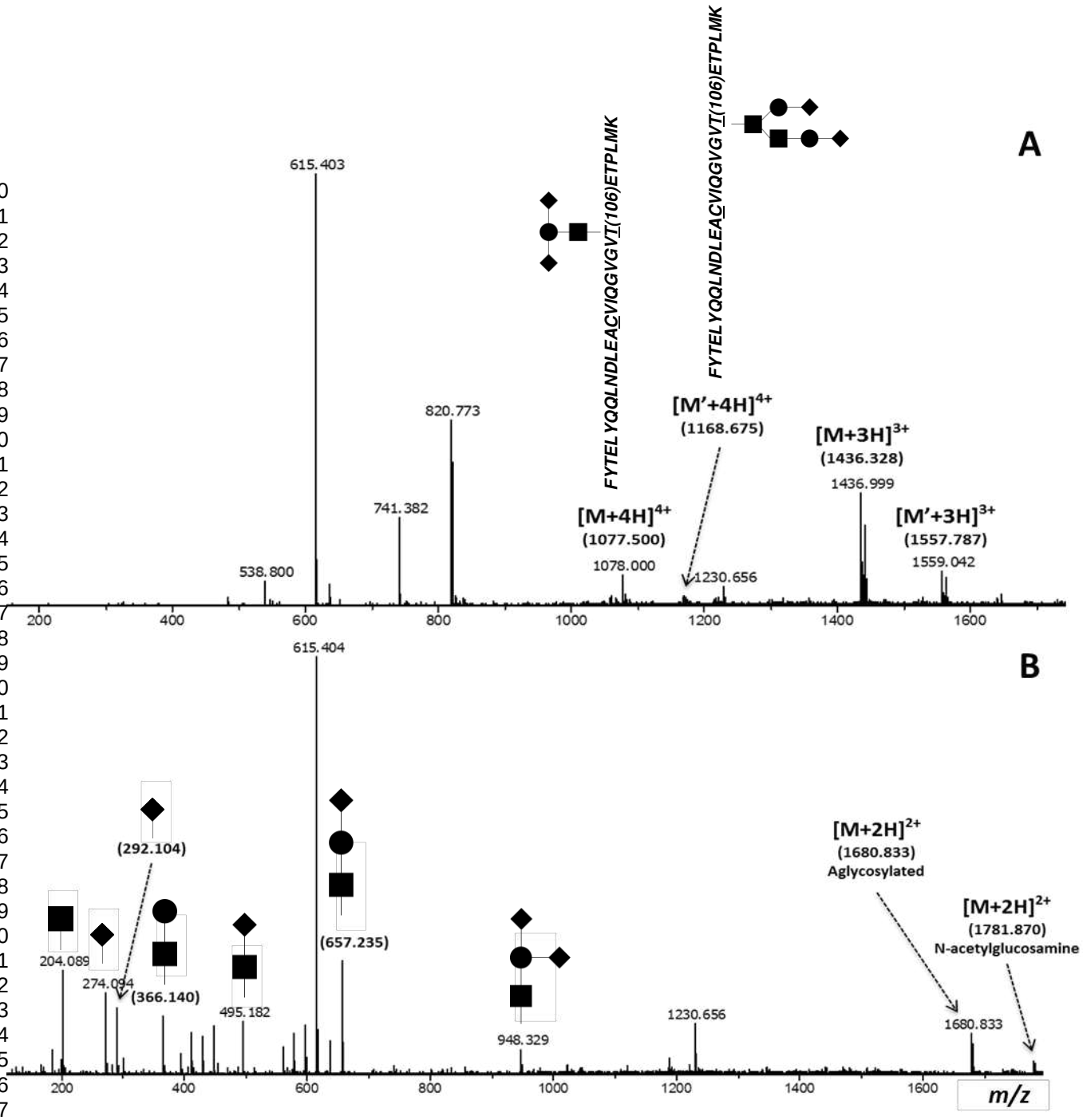
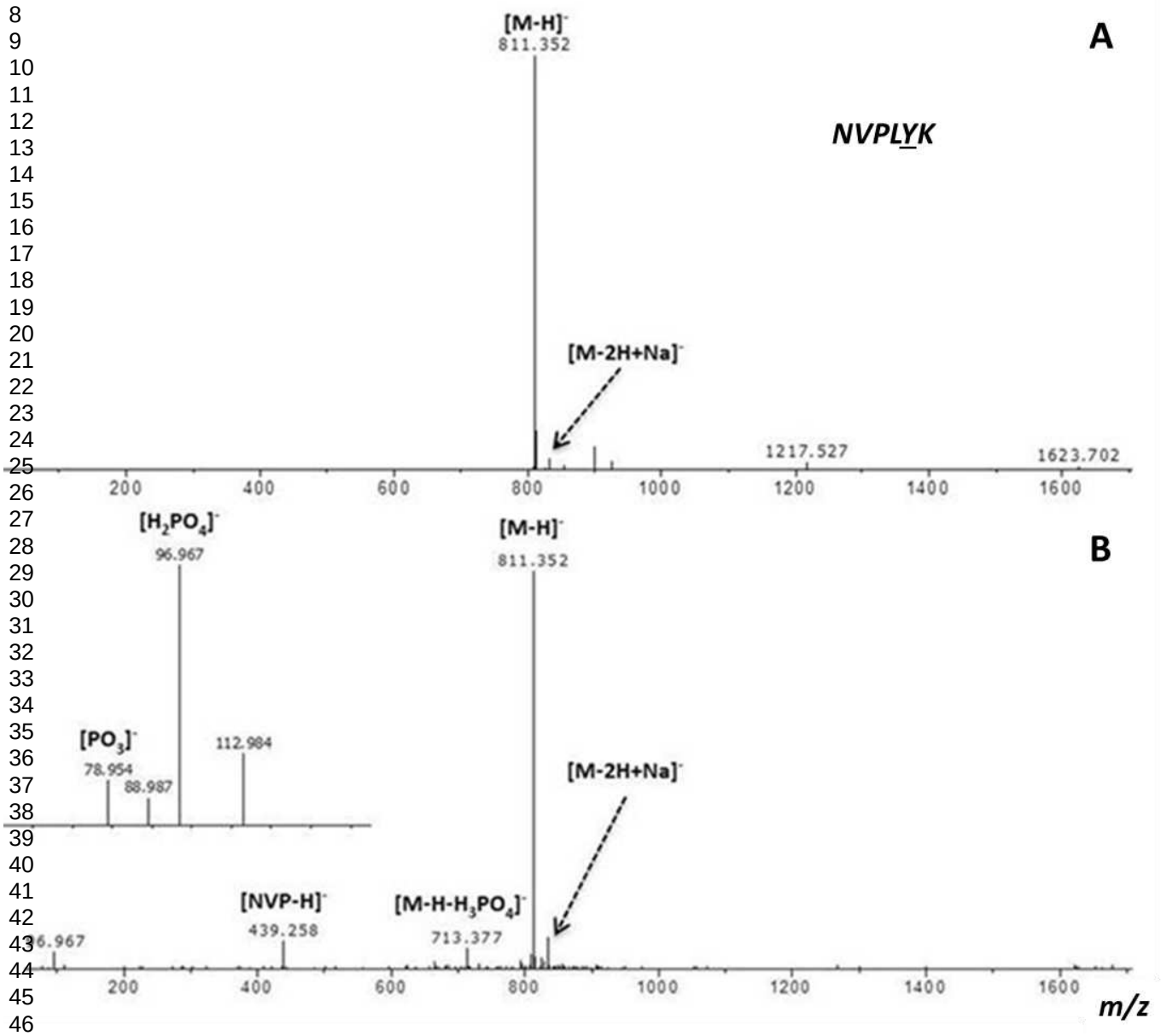
Peptide sequence	Missed Cleavages	Glycan composition	Glycan symbol nomenclature	Experimental m/z; charge state	Theoretical m/z; charge state	Error (ppm)	RT (min)
DSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTI(106)ETPLMK	1	Hex-HexNAc-SA_2		1435.414 $[M+4H]^{4+}$	1435.413 $[M+4H]^{4+}$	-1	87.6
				1148.545 $[M+5H]^{5+}$	1148.531 $[M+5H]^{5+}$	-12	
		Hex_2-$HexNAc_2$-SA_2		1526.695 $[M'+4H]^{4+}$	1526.696 $[M'+4H]^{4+}$	1	87.6
FYTELYQQLNDLEACVIQGVGVTI(106)ETPLMKEDSILAVR	1	Hex-HexNAc-SA_2		1298.262 $[M+4H]^{4+}$	1298.269 $[M+4H]^{4+}$	5	67.1
				1038.899 $[M+5H]^{5+}$	1038.896 $[M+5H]^{5+}$	-3	
DSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTI(106)ETPLMKEDSILAVR	2	Hex-HexNAc-SA_2		1656.281 $[M+4H]^{4+}$	1656.282 $[M+4H]^{4+}$	1	80.9
				1325.222 $[M+5H]^{5+}$	1325.227 $[M+5H]^{5+}$	4	
		Hex_2-$HexNAc_2$-SA_2		1747.574 $[M'+4H]^{4+}$	1747.565 $[M'+4H]^{4+}$	-5	80.9
				1398.257 $[M+5H]^{5+}$	1398.253 $[M+5H]^{5+}$	-3	
FYTELYQQLNDLEACVIQGVGVTI(106)ETPLMK	0	Hex-HexNAc-SA_2		1436.326 $[M+4H]^{3+}$	1436.330 $[M+4H]^{3+}$	3	70.9
				1077.500 $[M+4H]^{4+}$	1077.500 $[M+4H]^{4+}$	0	
		Hex_2-$HexNAc_2$-SA_2		1557.787 $[M+3H]^{3+}$	1558.789 $[M+3H]^{3+}$	1	70.7
				1168.675 $[M+4H]^{4+}$	1168.653 $[M+4H]^{4+}$	-19	

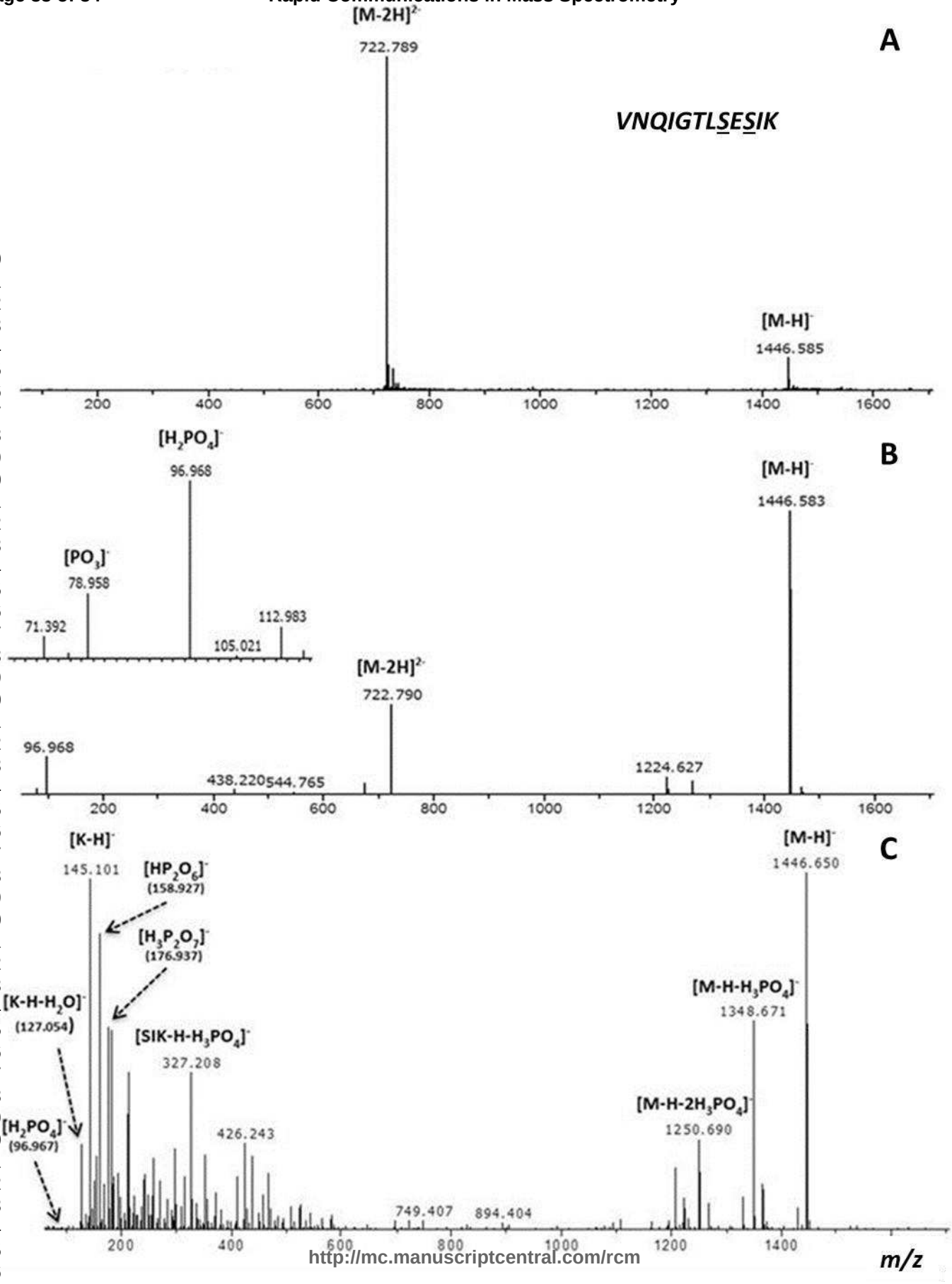
Table 3

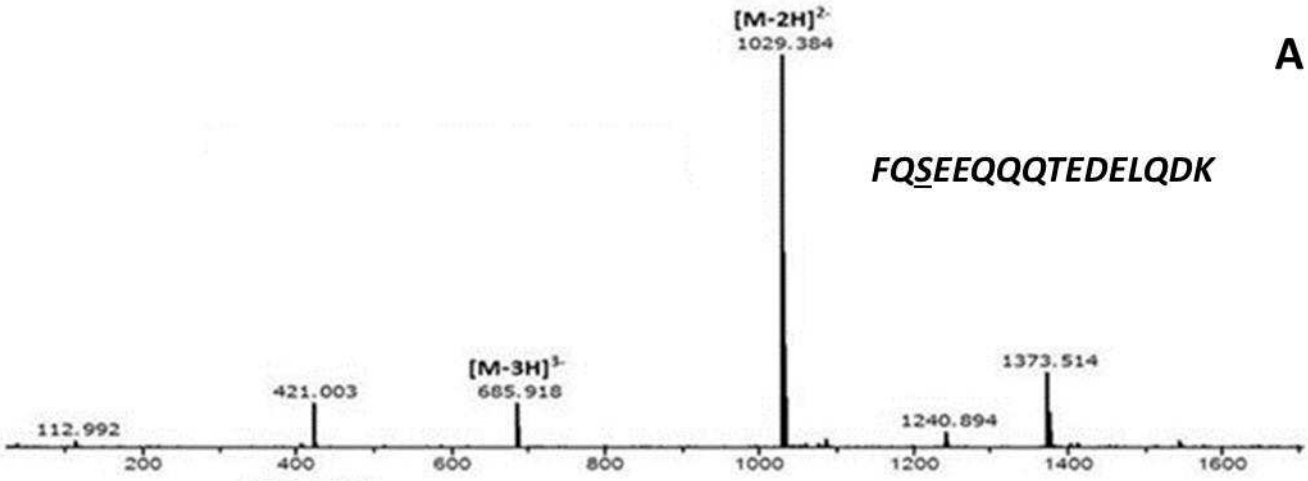
Peptide	*Sequence*	*Experimental m/z; charge state*	*Theoretical m/z; charge state*	*Error (ppm)*	*RT (min)*	*Intensity*
T1	*HLADLSK*	781.367 $[M-H]^-$	781.382	*19*	16.6	2.21E04
T1p	*HLADLSK*	861.327 $[M-H]^-$	861.342	*17*	17.9	5.20E04
T2	*NVPLYK*	731.358 $[M-H]^-$	731.366	*11*	19.5	3.20E04
T2p	*NVPLYK*	811.352 $[M-H]^-$	811.357	*6*	20.7	3.88E04
T3	*VNQIGTLSESIK*	1286.606 $[M-H]^-$	1286.609	*2*	26.1	1.29E04
		642.799 $[M-2H]^{2-}$	642.803	*6*		6.57E04
T3p	*VNQIGTLSESIK*	1366.567 $[M-H]^-$	1366.577	*7*	29.7	1.54E04
		682.779 $[M-2H]^{2-}$	682.791	*18*		1.03E05
T3pp	*VNQIGTLSESIK*	1446.585 $[M-H]^-$	1446.62	*24*	32.5	1.16E04
		722.789 $[M-2H]^{2-}$	722.795	*8*		1.75E05

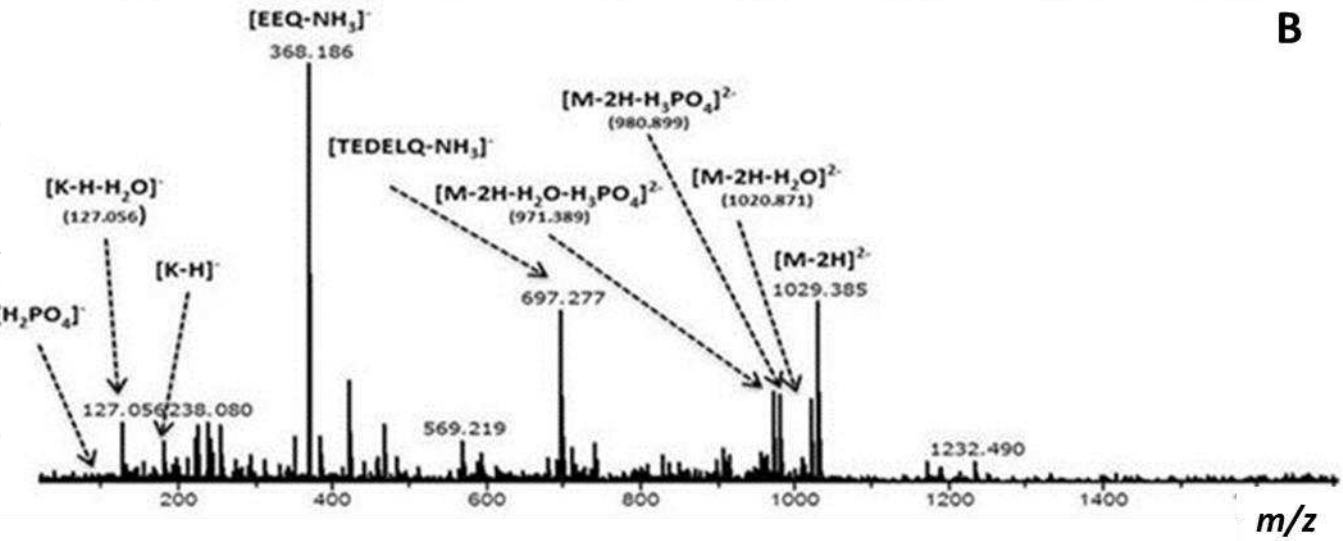

A
[M+5H]5+
793.790
LCPDCPLLAPLN(156)DSR
[M+6H]6+
661.659
[M+4H]4+
991.986
B
204.087
274.090
366.135
657.226
[M+4H]4+
991.986
[M+2H]2+
(870. 928)
(aglycosylated)
[M+5H]5+
793.790
[M+3H-HCOOH-CO2]3+
(1291.952)
[M+3H]3+
1322.316
861.433
200
400
600
800
1000
1200
1400
1600
m/z

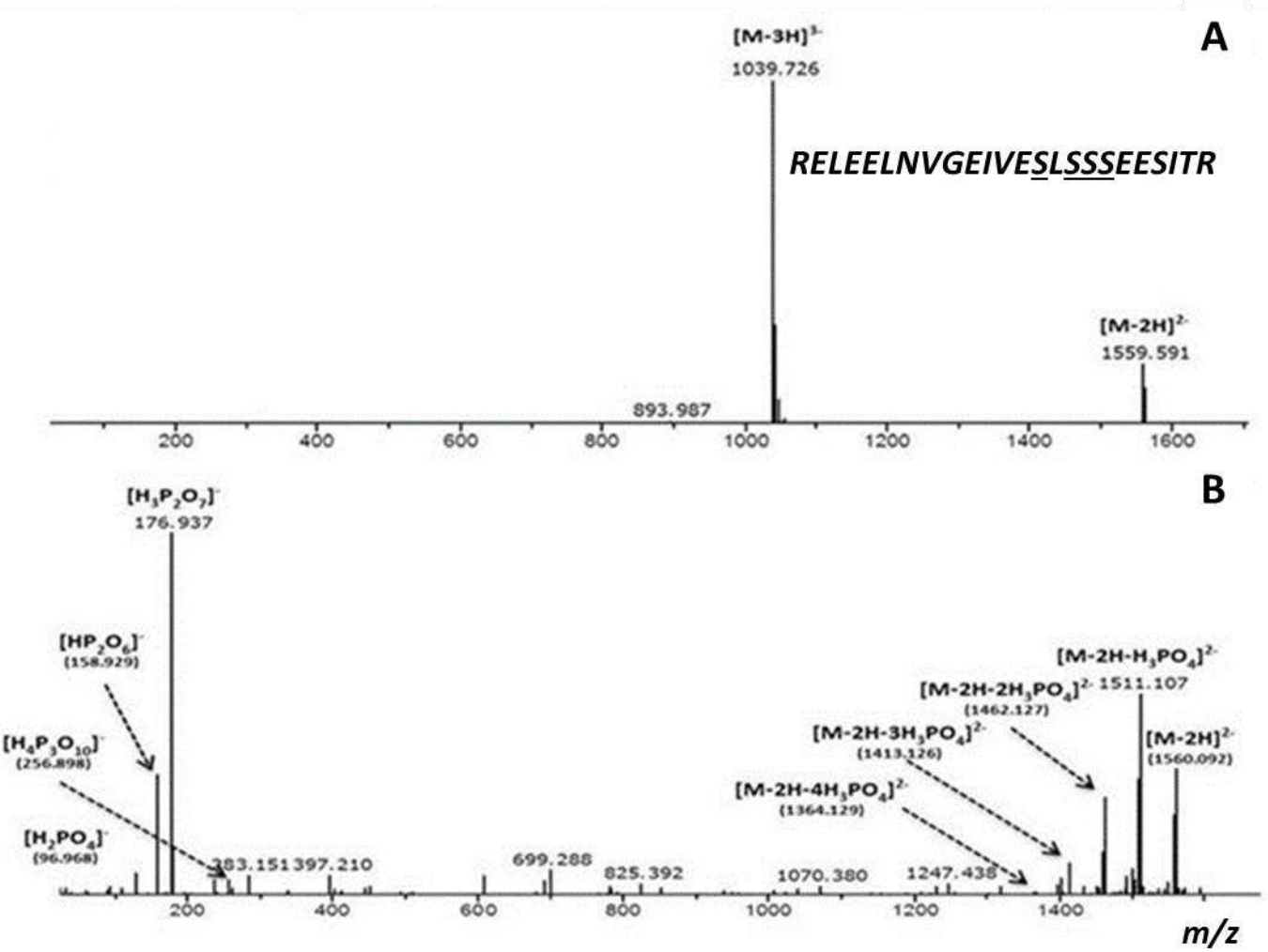

A
FYTELYQQLNDLEACVIQGVGVT(106)ETPLMK
FYTELYQQLNDLEACVIQGVGVT(106)ETPLMK
615.403
820.773
741.382
538.800
[M+4H]4+
(1077.500)
1078.000
[M'+4H]4+
(1168.675)
1230.656
[M+3H]3+
(1436.328)
1436.999
[M'+3H]3+
(1557.787)
1559.042
200
400
600
800
1000
1200
1400
1600
B
615.404
(292.104)
204.089
274.094
(366.140)
495.182
(657.235)
948.329
1230.656
[M+2H]2+
(1680.833)
Aglycosylated
[M+2H]2+
(1781.870)
N-acetylglucosamine
1680.833
200
400
600
800
1000
1200
1400
1600
m/z


[M-H]⁻
811.352
A
NVPLYK
[M-2H+Na]⁻
1217.527
1623.702
200
400
600
800
1000
1200
1400
1600
[H2PO4]⁻
96.967
[M-H]⁻
811.352
B
[PO3]⁻
78.954
88.987
112.984
[M-2H+Na]⁻
[NVP-H]⁻
439.258
[M-H-H3PO4]⁻
713.377
96.967
200
400
600
800
1000
1200
1400
1600
m/z


A
[M-2H]2-
722.789
VNQIGTLSESIK
[M-H]-
1446.585
200
400
600
800
1000
1200
1400
1600
B
[H2PO4]-
96.968
[M-H]-
1446.583
[PO3]-
78.958
71.392
112.983
105.021
[M-2H]2-
722.790
96.968
438.220
544.765
1224.627
C
[K-H]-
145.101
[M-H]-
1446.650
[HP2O6]-
(158.927)
[H3P2O7]-
(176.937)
[K-H-H2O]-
(127.054)
[M-H-H3PO4]-
1348.671
[SIK-H-H3PO4]-
327.208
[M-H-2H3PO4]-
1250.690
[H2PO4]-
(96.967)
426.243
749.407
894.404
m/z


[M-2H]2-
1029.384
A
FQSEEQQQTEDELQDK
[M-3H]3-
685.918
421.003
1373.514
112.992
1240.894
200
400
600
800
1000
1200
1400
1600


[EEQ-NH3]-
368.186
B
[M-2H-H3PO4]2-
(980.899)
[TEDELQ-NH3]-
[K-H-H2O]-
(127.056)
[M-2H-H2O-H3PO4]2-
(971.389)
[M-2H-H2O]2-
(1020.871)
[M-2H]2-
1029.385
[K-H]-
697.277
[H2PO4]-
127.056
238.080
569.219
1232.490
200
400
600
800
1000
1200
1400
m/z


A
[M-3H]3-
1039.726
RELEELNVGEIVESLSSSEESITR
[M-2H]2-
1559.591
893.987
200
400
600
800
1000
1200
1400
1600
B
[H3P2O7]-
176.937
[HP2O6]-
(158.929)
[H4P3O10]-
(256.898)
[H2PO4]-
(96.968)
383.151
397.210
699.288
825.392
1070.380
1247.438
[M-2H-H3PO4]2-
[M-2H-2H3PO4]2- 1511.107
(1462.127)
[M-2H-3H3PO4]2-
(1413.126)
[M-2H]2-
(1560.092)
[M-2H-4H3PO4]2-
(1364.129)
200
400
600
800
1000
1200
1400
m/z