FOCUS: MS/MS PEPTIDE IDENTIFICATION: RESEARCH ARTICLE

Unusual Fragmentation of Pro-Ser/Thr-Containing Peptides Detected in Collision-Induced Dissociation Spectra

Katalin F. Medzihradzsky, Jonathan C. Trinidad

Mass Spectrometry Facility, School of Pharmacy, University of California San Francisco, 600 16th Street Genentech Hall, Suite N472A, Box 2240, San Francisco, CA 94158-2517, USA

Abstract

During collision-induced dissociation (CID)-, phosphoserine- and phosphothreonine-containing peptides frequently undergo neutral loss of phosphoric acid. Subsequent amide bond cleavage N-terminal to the site of phosphorylation results in a y ion with a mass 18 Da lower than the corresponding unmodified y fragment. We report here that when the phosphoserine or phosphothreonine is directly preceded by a proline, an unusual fragment with a mass 10 Da higher than the corresponding unmodified y ion is frequently observed. Accurate mass measurements are consistent with elimination of the phosphoric acid followed by fragmentation between the α carbon and the carbonyl group of the proline residue. We propose a cyclic oxazoline structure for this fragment. Our observation may be explained by the charge-directed S_N2 neighboring group participation reaction proposed for the phosphoric acid elimination by Palumbo et al. [Palumbo, A. M., Tepe, J. J., Reid, G. E. Mechanistic Insights into the Multistage Gas-Phase Fragmentation Behavior of Phosphoserine- and Phosphothreonine-Containing Peptides. *J. Protein Res.* 7(2), 771–779 (2008)]. Considering such specific fragment ions for confirmation purposes after regular database searches may boost the confidence of peptide identifications as well as phosphorylation site assignments.

Key words: CID, HCD, Fragmentation, Phosphorylation, Database search

Introduction

High throughput proteomics experiments have been used for qualitative and quantitative characterization of protein complexes, cell organelles, as well as cell lysates. With the advent of ever faster and more sensitive new instrumentation, over 10,000 MS/MS spectra can be acquired in a 90 min LC-MS/MS run. Thus, the identification of these spectra has to be performed in an automated fashion and the reliability of the resulting peptide assignments needs to be addressed. Collision-induced dissociation (CID) of peptides may yield a wide variety of fragment ions [nomenclature: [1]; review: [2]]. Despite the abundance of fragment ion types, search engines usually consider only a

few “favorites” when matching spectra against a database. The products of the peptide bond cleavage, i.e. the N-terminal b and C-terminal y fragments are considered rightfully the most important for CID spectra. In addition, some neutral losses from these fragments are permitted, such as CO losses from b ions (a fragment formation), as well as ammonia- and water losses from both b and y fragments. Such losses are usually linked to the amino acid composition of the fragment ion: the presence of Asn, Gln, Lys, and Arg promotes NH_3 loss, while Asp, Glu, Ser, and Thr-containing fragments may lose H_2O .

Certain search engines, such as Mascot, Spectrum Mill and Protein Prospector will consider instrument-dependent fragmentation. For example, they may consider internal ion formation in CID-experiments performed in quadrupole collision cells (such as triple quadrupole and QTOF instruments, as well as HCD in Orbitraps). However, permitting

Correspondence to: Katalin F. Medzihradzsky; e-mail: folk1@cgl.ucsf.edu

too many fragment-types during the database searches may slightly increase the search time and increase the rate of false positive identifications [3]. At the same time, judiciously considering “minor” fragment-types, as well as unusual, sequence-specific fragment ions may increase the confidence in peptide assignments and eliminate false positives, if applied as a confirmation measure of the initial database search results.

While they share many common features, ion trap CID spectra are distinct from quadrupole collision cell CID data. However, the fact that they are both referred to as “CID” spectra leads to some confusion. Resonance excitation in ion traps generally leads to excitation and fragmentation of a specific m/z range. Therefore, fragment ions (which typically have m/z values distinct from the precursor) do not undergo secondary fragmentation. In terms of gas pressure and peptide collision energy, the design of quadrupole collision cells promotes multiple fragmentation events. Thus, one would expect more ions derived from secondary fragmentation using a quadrupole collision cell relative to an ion trap. Internal ion formation is a well known, though frequently ignored phenomenon. Other unusual fragments also have been reported. For example, it has been described that sequences featuring a Gln residue in their second position usually produce an abundant c_1 fragment [4]. Rearrangement reaction of b-type fragments also has been documented [5].

Software such as Peptide Prophet or Scaffold statistically reevaluate peptide assignments after the database search is finished, however neither of them utilize any new fragments in the confirmation process [6]. Protein Prospector may be

manually prompted to assign fragment ion types that were not considered during the database search, for example, ions formed via multiple neutral losses. However, we are not aware of any software that will consider specific minor fragments during the confirmation/evaluation process.

In this paper, we report an unusual fragment formation observed prominently from phosphoserine or phosphothreonine-containing peptides that contain a proline immediately preceding the site of modification. This fragmentation is observed, albeit less intensely, when the ProSer or ProThr sequence is not phosphorylated. Such fragmentation was detected in CID acquired in QTOF instruments as well as in HCD on an LTQ-Orbitrap, but was not observed in ion trap CID.

Experimental

Samples were tryptic digests from different projects, with or without iTRAQ-derivatization, unmodified and phosphorylated peptides separated by an enrichment step.

The synthetic peptide, AGSPS(phospho)ADPFR was a gift from Nick Agard (UCSF, School of Pharmacy), and was analyzed from the raw synthetic mixture without further purification.

All MS/MS data were acquired in LC-MS experiments in a data dependent fashion, when multiply charged precursor ions were computer-selected from the MS survey scans, and the activation conditions were adjusted to the m/z and z of the precursor ions. Mass spectrometers used for data acquisition were a QSTAR Pulsar, QSTAR Elite (MDS

Figure 1. CID spectrum of an iTRAQ-labeled phosphopeptide acquired in LC/MS mode on a QSTAR mass spectrometer from precursor ion m/z 649.3(3+). The peptide was identified as human dystrobrevin peptide TQFEDLVPS(phospho)PTSEK. The fragment designated “ x_6 ” is 97 Da higher than y_5 . In this spectrum, it was detected both singly and doubly charged. ^ indicates elimination of H_3PO_4 .

Table 1. List of phosphopeptides featuring the unusual fragment ion [7–9]

Gene	Protein	Peptide sequence	BP % ^a	x/y-98 ^b
PS(phospho)-containing sequences				
A3KGU4_MOUSE	Spectrin α 2	TSSKE(ss)PVPsPTLDR ^c	50	4.4
A3KGU4_MOUSE	Spectrin α 2	E(ss)PVPsPTLDR	28	3.5
A3KGU4_MOUSE	Spectrin α 2	RPPsPDPNTK	5	6.7
SPTB2_MOUSE	Spectrin β chain, brain 1	(tssKEss)PVPsPTLDR	80	4
ANK2_MOUSE	Ankyrin-2	TEWQPSTDIPLQK	20	4.7
B2RQQ5_MOUSE	Microtubule-associated protein 1B	SPSLSPSPsPIEK	30	0.3
MAP1A_MOUSE	Microtubule-associated protein 1A	AELEEMEEVHPsDEEEETK	9	1.8
NFH_MOUSE	Neurofilament heavy polypeptide	IGFGPsPFSLTEGLPK	14	1.8
B1ART6_MOUSE	Microtubule-actin crosslinking factor 1	LLDAEDVDVPsPDEK	BP	8.9
SYGPI_RAT	Ras GTPase-activating protein SynGAP	LPsPTK	BP	6.4
LRRC7_MOUSE	Leucine-rich repeat-containing protein 7	TPsPFEDR	43	2
KPCG_MOUSE	Protein kinase C gamma	MGP(sss)PIPSPPsPTDSKR	(2+)BP	3.3
			(+) 69	3
LMNA_MOUSE	Prelamin-A/C	LSPsPTSQR	BP	4.6
STE20_CANAL	Serine/threonine-protein kinase CST20	IPsNPTSTR	52	1.8
STE20_CANAL	Serine/threonine-protein kinase CST20	TPsSQSLPR	18	0.4
PMA1_CANAL	Plasma membrane ATPase 1	GYLVAMTGDGVNDAPsLK	29	7.3
ACT_CANAL	Actin	EITALAPsSMK	18	2.5
TSA1_CANAL	Peroxiredoxin TSA1	YGEVCPANWHPGDETIKPsPEASK	7	y-98 not detected
TSA1_CANAL	Peroxiredoxin TSA1	APVVQQPAPsFKK	16	5
RS13_CANMA	40S ribosomal protein S13	NAPsWFK	9	3
RADI8_CANAL	Postreplication repair E3 ubiquitin-protein ligase RADI8	ANVASPsPVAQSTVHK	12	y-98 not detected
PT(phospho)-containing sequences				
B2RQQ5_MOUSE	Microtubule-associated protein 1B	DLTGQVPtPPVK	52	2.8
RBM14_MOUSE	RNA-binding protein 14	QPtPPFFGR	5	0.3
TBA1B_MOUSE	Tubulin α -1B chain	sIQFVDWCPtGFK ^d	13	2.8
LRFN6_MOUSE	Leucine-rich repeat and fibronectin type-III domain-containing protein 6	VFSLDVDPHPtPTGLAK ^d	16	1.5
SSN2_CANAL	Mediator of RNA polymerase II transcription subunit 13	SNDYAPtPMIQDK	5	y-98 not detected
CHS5_CANAL	Chitin biosynthesis protein CHS5	IESVPtDEIDTK	4	3
STE7_CANAL	Serine/threonine-protein kinase STE7 homolog	ILHIPtQK	23	1

^aThe relative intensity of the ‘x’ fragment within the spectrum; BP = base peak.
^bThe intensity ratio of the ‘x’ fragment and the normal y-H₃PO₄ fragment.
^cAmino acids in lower case indicate phosphorylation sites; multiple residues in parentheses indicate ambiguity of the site assignment.
^diTRAQ-labeled peptides.

Figure 2. From the CID spectrum of an iTRAQ-labeled phosphopeptide, KDYLIENEELPS(phospho)P displaying the PS-specific fragment ion, ‘*x₂’. Data were acquired on a QSTAR Elite mass spectrometer. The spectrum was recalibrated using *m/z* 185 and 282 as internal standards (see Table 2). There is only one potential elemental composition for ‘*x₂’ within 20 ppm that also may fit a peptide fragment: C₉ H₁₃ N₂ O₄. Once the elemental composition of the C-terminal Pro is subtracted, C₄ H₃ N₁ O₂ is left

Table 2. Mass accuracy for the ions in Figure 2. Fragments m/z 185 and 282 served as internal standards

m/z exp.	m/z calc.	Dev [ppm]	Assignment
185.0921	185.0921	Ref 1	$y_2\text{-H}_3\text{PO}_4$
213.0909	213.0870	18	x_2
226.0886	226.0822	28	EN/NE-H ₂ O
244.0947	244.0928	8	EN/NE
259.0963	259.0925	15	EE
265.0605	265.0584	8	PS(phospho)
282.1448	282.1448	Ref 2	$y_3\text{-H}_3\text{PO}_4$
348.2084	348.2020	18	$b_3(2+)$
355.1240	355.1248	-2	ENE/NEE-H ₂ O
380.1256	380.1217	10	y_3
404.7454	404.7440	3	$b_4(2+)$

SCIEX, Concord, Canada) well as an LTQ-Orbitrap Velos (Thermo Fisher, San Jose, CA, USA) for the synthetic peptide.

Database searches were performed by Protein Prospector adjusting the search parameters according to the experiments, and the search results were manually inspected.

Instrument-specific fragment ion lists were obtained using MS-Product of Protein Prospector, <http://prospector.ucsf.edu/prospector/mshome.htm>.

Results and Discussion

Manual inspection of the quadrupole collision cell CID data of phosphopeptides from various enrichment experiments revealed a novel, frequently abundant ion for peptides containing a phosphorylated ProSer or ProThr sequence. The ProSer sequence is very common in phosphorylation motifs. Thus, it was first noticed that in ProSer(phospho)-containing peptides fragmentation between the proline and phosphorylated serine residues resulted in a “y” ion where the serine had an apparent mass of 97 Da (rather than 87) (Figure 1). In a small dataset of 176 biologically “interesting” phosphopeptides 16 unique, unrelated sequences featured the Pro followed by a phosphorylated Ser residue, and 8 of the CID spectra displayed this fragment (data not shown). However, the expected y ion that lost the phosphoric acid was also detected in each instance. We further investigated the phenomenon by examining other phosphopeptide datasets generated in our laboratory. Our observation was further supported by numerous quadrupole collision cell CID spectra featuring this unusual fragment ion (Table 1, [7–9]). In addition, the same phenomenon was detected in some ProThr(phospho)-containing pep-

tides: the apparent mass of the Thr was 111 Da, i.e., an unusual fragment was detected at an m/z value 10 Da higher than a y fragment without any modification (Table 1, [7–9]). Observation of this unusual fragment ion seemed to require an intense y ion series. In general, y ion intensity decreases with distance from the carboxy-terminus, and observation of this fragment ion was mostly limited to peptides where the ProSer/ProThr pair was within 12 amino acids of the C-terminus. Accurate mass measurements revealed that the 97 Da mass increment corresponds to an elemental composition of $C_4H_3NO_2$ (the elemental composition of proline is C_5H_7NO) (Figure 2, Table 2). After neutral loss of the phosphoric acid, cleavage between the proline’s α carbon and its carbonyl group (similar to x fragment formation) would result in the observation of a new fragment ion 10 Da higher in mass than a y fragment corresponding to an unmodified serine. The potential pathways for the elimination of phosphoric acid from Ser/Thr-phosphorylated peptides have been studied by Palumbo *et al.* using a linear ion trap [10]. Three different pathways were suggested. Pathway A is the canonical (charge-remote) β -elimination pathway, while pathways B and C represent alternative, charge-directed reaction mechanisms: E2 elimination reaction and S_N2 neighboring group participation reaction. They found that the charge-directed reactions represent the dominant process, and they provided experimental evidence for the phosphoric acid elimination via the S_N2 neighboring group participation reaction, i.e., pathway C (Scheme 1). Pathway C results in a structure that makes normal y ion formation unlikely because three bonds would need to be broken to observe that y ion. Since in most quadrupole CID spectra, the y ion resulting from fragmentation of the peptide bond at the N-terminal side of the phosphoSer/Thr residue is observed, we speculate that the other pathways must be preferred during quadrupole CID of multiply charged peptides. However, phosphoric acid elimination via pathway C would explain the formation of the novel fragment described here (Scheme 1). While this fragmentation was originally identified from QSTAR data, examination of LTQ-Orbitrap data of such peptides revealed that the phenomenon occurs in HCD analysis, but was not observed in ion trap CID.

We confirmed this observation using a synthetic peptide containing the Pro-Ser motif (AGSPS(phospho)ADPFR). This peptide was analyzed in LC-MS/MS using an LTQ-Orbitrap Velos instrument, subjecting the peptide to both ion trap CID and HCD analysis. The HCD spectrum of the phosphopeptide indeed showed the presence of the above described unusual

Scheme 1. Fragmentation mechanism proposed. The charge-directed S_N2 neighboring group participation reaction for phosphoric acid elimination was presented as pathway C [10]

Figure 3. (a) HCD spectrum of synthetic peptide, AGSPS(phospho)ADPFR-amide acquired from m/z 542.2375(2+) in LC/MS/MS mode using an LTQ-Orbitrap Velos. ^ indicates phosphoric acid elimination, '*x₆' stands for the Pro-Ser-sequence-specific fragment. Its calculated mass is 701.3366 that is 4 ppm lower than the experimental value. Fragments y₅ and y₇-H₃PO₄ were also measured with the same mass accuracy. (b) Ion trap CID of the same precursor ion acquired in the next scan. No Pro-Ser-sequence-specific fragment ion was detected in this experiment. The insert shows the region of interest magnified tenfold. (c) HCD spectrum acquired from the doubly charged ion (m/z 502.25) of the unmodified peptide, AGSPSADPFR-amide. Both the expected y₆ fragment and the Pro-Ser-sequence-dependent '*x₆' ion was observed in the experiment

Table 3. Some unmodified peptides yielding the unusual fragment ion^a

Sequence	'x'		Δ ppm
	<i>m/z</i> exp. ^b	<i>m/z</i> calc.	
VLALPEPSAAPTLLR	822.4538	822.4464	9
DTHDQLSEPEVR	500.2472	500.2463	2
<QLMHNGHPSEK	373.1753	373.1717	9
TPDSLEPSPLK	454.2625	454.2659	7
<QTPSGTASR	588.2673	588.2736	11

^aData were acquired on a QSTAR Pulsar mass spectrometer.^bAfter internal calibration.

fragment ion (Figure 3a), the corresponding ion trap CID did not display such fragmentation (Figure 3b). To our surprise, the unmodified version of the peptide also formed this fragment ion, albeit to a lesser extent (Figure 3c). Further investigation revealed that indeed non-modified Pro-Ser containing peptides also may undergo such fragmentation (Table 3).

Conclusions

In quadrupole collision-cell CID experiments, a novel fragment ion may be formed from peptides containing Pro-Ser/Thr sequences. The novel fragment is most likely formed in two consecutive steps: first the phosphoric acid/ or water is eliminated via a charge-directed neighboring group assisted S_N2 reaction, and a cyclic oxazoline structure is formed [10]. Then a cleavage occurs between this ring and the α carbon of the proline residue. When the hydroxy amino acid is not phosphorylated, this fragment ion is less intense than the expected y ion. However, in the case of a phosphorylated serine or threonine, this fragment may be very abundant. Thus, this fragment ion could be utilized after regular database searching for peptide assignment confirmation as well as aiding phosphorylation site assignment. In general, we believe that considering sequence specific "minor" fragmenta-

tion may improve the reliability of peptide identifications in large scale proteomics experiments.

Acknowledgments

The authors acknowledge support for this work by NIH grants NCRR P41RR001614 and 1S10RR012961, and the Howard Hughes Medical Institute (to the Bio-organic Biomedical Mass Spectrometry Resource at UCSF, director A. L. Burlingame).

References

1. Biemann, K.: Nomenclature for peptide fragment ions (positive-ions). *Methods Enzymol.* **193**, 886–887 (1990)
2. Medzihradsky, K.F.: Peptide Sequence Analysis. *Methods Enzymol.* **402**, 209–44 (2005)
3. Baker, P.R., Medzihradsky, K.F., Chalkley, R.J.: Improving software performance for peptide electron transfer dissociation data analysis by implementation of charge state- and sequence-dependent scoring. *Mol. Cell. Proteom.* **9**(9), 1795–1803 (2010)
4. Lee, Y.J., Lee, Y.M.: Formation of c₁ fragment ions in collision-induced dissociation of glutamine-containing peptide ions: A tip for de novo sequencing. *Rapid Commun. Mass Spectrom.* **18**(18), 2069–2076 (2004)
5. Erlekam, U., Bythell, B.J., Scuderi, D., Van Stipdonk, M., Paizs, B., Maitre, P.: Infrared spectroscopy of fragments of protonated peptides: Direct evidence for macrocyclic structures of b₅ ions. *J. Am. Chem. Soc.* **131**(32), 11503–11508 (2009)
6. Keller, A., Nesvizhskii, A. I., Kolker, E., Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**(20), 5383–5392 (2002); <http://www.proteomesoftware.com/>
7. Trinidad, J.C., Thalhammer, A., Specht, C.G., Lynn, A.J., Baker, P.R., Schoepfer, R., Burlingame, A.L.: Quantitative Analysis of Synaptic Phosphorylation and Protein Expression. *Mol. Cell. Proteom.* **7**(4), 684–696 (2008)
8. Trinidad, J.C., Thalhammer, A., Specht, C.G., Schoepfer, R., Burlingame, A.L.: Comprehensive Identification of Phosphorylation Sites in Postsynaptic Density Preparations. *Mol. Cell. Proteom.* **5**(5), 914–922 (2006)
9. Beltrao, P., Trinidad, J.C., Fiedler, D., Togoev, A., Lim, W.A., Shokat, K.M., Burlingame, A.L., Krogan, N.J.: Evolution of Phosphoregulation: Comparison of Phosphorylation Patterns across Yeast Species. *PLoS Biol.* **7**(6), e1000134 (2009)
10. Palumbo, A.M., Tepe, J.J., Reid, G.E.: Mechanistic Insights into the Multistage Gas-Phase Fragmentation Behavior of Phosphoserine- and Phosphothreonine-Containing Peptides. *J. Prot. Res.* **7**(2), 771–779 (2008)