In-Source Fragmentation and the Sources of Partially Tryptic Peptides in Shotgun Proteomics

Jong-Seo Kim[1, 2], Matthew E. Monroe[1], David G. Camp II[1], Richard D. Smith[1], and Wei-Jun Qian[1]
[1]Pacific Northwest National Laboratory, Richland, WA; [2]Institute for Basic Science, Seoul National University, Seoul, Korea

Pacific Northwest NATIONAL LABORATORY

Overview

- This study investigates the impact of in-source fragmentation on peptide identifications from a standard protein mixture and two complex biological samples.
- In-source fragments are distinguished from endogenous partially tryptic peptides based on elution time criteria.

Introduction

- In shotgun proteomics, numerous partially tryptic peptides are often observed, though ideally peptides are supposed to be fully tryptic.
- The partially tryptic peptides from chymotrypsin contamination and trypsin-autolysis issues were largely alleviated by the use of sequencing grade modified trypsin.[1]
- The other major sources for generating partially tryptic peptides are endogenous proteolysis and in-source fragmentation whereby fully tryptic peptides are fragmented within the electrospray ionization source.
- To clarify the sources of partially tryptic peptides and the specificity of trypsin, we investigated the impact of in-source fragmentation on peptide identifications from a standard protein mixture and two complex biological samples (mouse brain lysate and plasma).
- In-source fragments were distinguished from endogenous partially tryptic peptides based on elution time criteria.

Methods

- Three different peptide samples were analyzed using LC-MS/MS (Orbitrap mass spectrometer): 1) a mixture of six protein standards (bovine carbonic anhydrase II, bovine beta-lactoglobulin, *E. coli* beta-galactosidase, equine skeletal muscle myoglobin, chicken ovalbumin, and bovine cytochrome *c*); 2) mouse brain, and 3) mouse plasma.
- Parameters for ion optics were automatically tuned by Xcalibur Tune Plus to maximize ion gain. Tuning parameters were: ESI voltage 2.2 kV, heated capillary temperature 200 ℃, capillary voltage 30 V, and tube lens voltage 80 V.
- A false discovery rate of 0.05% for peptide identification was achieved by filtering peptides, using an MS-GF[2, 3] score threshold in combination with a 10 ppm mass accuracy cut-off.
- The first observed scan number for all peptides identified was plotted against predicted normalized elution times (NET) using an in-house NET prediction tool[4] (publicly available at http://omics.pnl.gov/software).
- These plots and individual total ion chromatographs of fully tryptic peptides and their potential daughter peptides were used to classify all partially tryptic peptide from in-source fragmentation based on elution time criterion.

Results

Partially tryptic peptides from in-source fragmentation

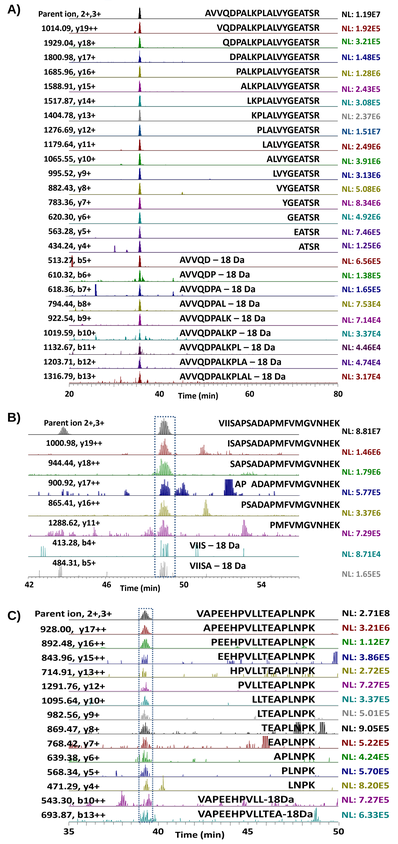
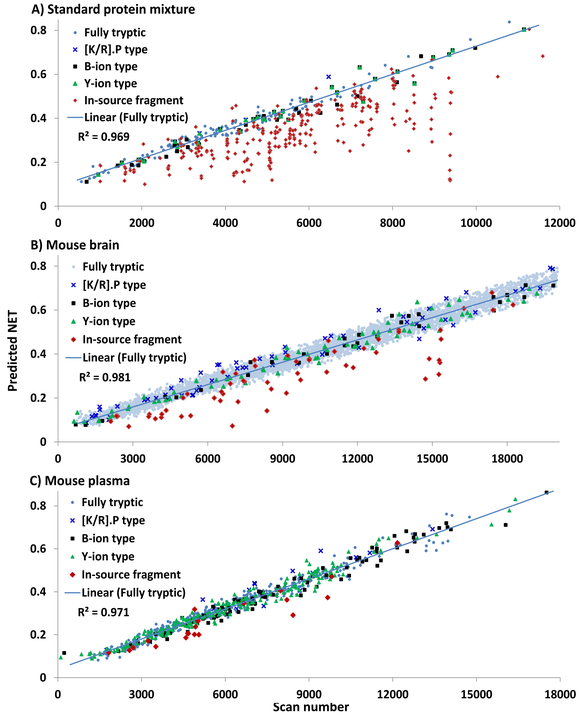


Fig. 1. Extracted ion chromatograms of AVVQDPALKPLALVYGEATSR from bovine carbonic anhydrase II in a standard protein mixture (A); VIISAPSADAPMFVMGVNHEK from glyceraldehyde -3-phosphate dehydrogenase (B) and VAPEEHPVLL -TEAPLNPK from actin, cytoplasmic 1 (C) in mouse brain lysate along with their detectable in-source fragments. NL: Signal intensity of the most intense peak in view.

Proportions of in-source fragmentation-generated peptides



Type of partially tryptic peptides	Description
[K/R].P type	[K/R].P---X[K/R].X' or [K/R].X---P.[K/R]
Y-ion type	[not K/R].X---X'[K/R].X''
B-ion type	[K/R].X---X'.[not K/R]

X indicates any amino acid residue

Fig. 2. Plots of first scan number of identified peptides vs. predicted normalized elution time (NET) in standard protein mixture (A), mouse brain (B), and mouse plasma (C). The observed elution times of in-source fragments appear later than predicted as in-source fragments have the same elution times as their parent peptides.

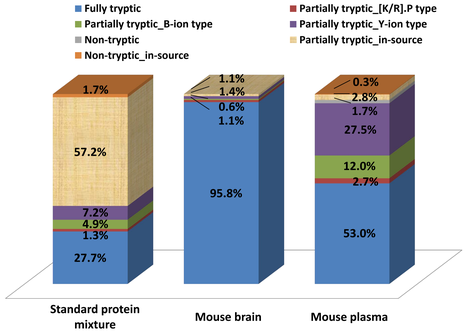


Fig. 3. Relative portions of each type of tryptic peptide at the unique peptide level in the three biological samples.

Example of sources of partially tryptic peptides in mouse plasma

Protein	Peptide	Origin
APOA1_MOUSE	S.PVAEEFR.D	in-source
	Q.LAPHSEQMR.E	in-source
	A.PLGAELQESAR.Q	in-source
	L.SPVAEEFR.D	protease
	D.FWDNLEK.E	protease
	D.TLGSTVSQLQER.L	protease
	W.DTLGSTVSQLQER.L	protease
	N.LNLLENWDTLGSTVSQLQER.L	protease
HEMO_MOUSE	N.PVTGEVPPR.Y	in-source
	R.WKNPITSVD.A	protease
	W.FWDFATR.T	protease
IGG2B_MOUSE	L.PPPAEQLSR.K	in-source
	S.VAHPASSTTVDK.K	protease
	C.SVAHPASSTTVDK.K	protease
	T.CSVAHPASSTTVDK.K	protease

Sequence motifs of partially tryptic peptides

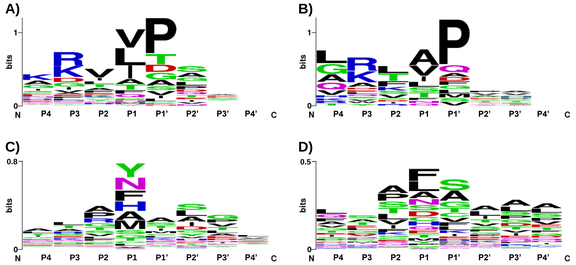


Fig. 4. Sequence motifs around cleavage sites in in-source fragmentation artifacts from mouse brain (A) and mouse plasma (B). Sequence motifs of partially tryptic peptides after excluding in-source artifacts from mouse brain (C) and mouse plasma (D).

Conclusions

- Our results indicate in-source fragmentation can be a major source of partially tryptic peptides in LC-MS/MS proteomics and that the impact of these artifacts is much less significant for data-dependent measurements of more complex mixtures due to the larger extent of under sampling of low-abundance species.
- It appears that the impact of trypsin-induced partially tryptic peptides is minimal (< 2%).
- The majority of the in-source fragments showed much later elution times compared to the predicted time, which is consistent with the larger size of their actual parent ions. By excluding such in-source fragmentation artifacts, one can then much more effectively identify in-vivo or in-vitro proteolytic products, as well as better characterize protease activities.

Acknowledgements

Portions of this work were supported by the NIH grants DP2OD006668 and 8P41 GM103493.

The experimental work described herein was performed in the Environmental Molecular Sciences Laboratory, a U.S. Department of Energy/BER national scientific user facility located at Pacific Northwest National Laboratory in Richland, Washington.

References

1. Picotti, P.; Aebersold, R.; Domon, B., *Mol. Cell. Proteomics* **2007,** 6, (9), 1589-1598.
2. Kim, S.; Gupta, N.; Pevzner, P. A., *J. Proteome Res.* **2008,** 7, (8), 3354-63.
3. Kim, S.; Mischerikow, N.; Bandeira, N.; Navarro, J. D.; Wich, L.; Mohammed, S.; Heck, A. J. R.; Pevzner, P. A., *Mol. Cell. Proteomics* **2010,** 9, (12), 2840-2852.
4. Petritis, K.; Kangas, L. J.; Yan, B.; Monroe, M. E.; Strittmatter, E. F.; Qian, W. J.; Adkins, J. N.; Moore, R. J.; Xu, Y.; Lipton, M. S.; Ii, D. G. C.; Smith, R. D., *Anal. Chem.* **2006,** 78, (14), 5026-5039

CONTACT: Jong-Seo Kim, Ph.D.
Institute for Basic Science
Seoul National University, South Korea
E-mail: jongseokim@snu.ac.kr

Wei-Jun Qian, Ph.D.
Biological Sciences Division
Pacific Northwest National Laboratory
E-mail: Weijun.Qian@pnnl.gov