AXIMA

Post Source Decay analysis of large peptides (>2,400 Da) and its value in protein identification

- Seamless PSD generated in as little as one minute
- Excellent sequence coverage with fragments down to <10% of the parent mass in a single data acquisition window
- Seamless PSD data were used to successfully identify proteins using broad search parameters within the Mascot™ MS/MS search engine

SHIMADZU BIOTECH

bringing analysis to life

PSD Analysis of Large Peptides

Introduction

The purpose of this work was:

1 to assess the efficacy of reflectron time-of-flight to produce high quality post source decay (PSD) spectra for large peptides (>2,400 Da),

2 to evaluate the usefulness of the PSD data generated in peptide and protein characterization, for example, sequence identification from a database search.

A MALDI TOF mass spectrometer with a curved field reflectron[1] was used to enable acquisition of a fully focused PSD spectrum in a single analysis - seamless PSD (sPSD).

Method

A curved field reflectron MALDI TOF mass spectrometer (AXIMA-CFR, Kratos Analytical, Manchester) was used, which comprised a UV nitrogen laser (337 nm), near ultra high vacuum system with a novel laser beam focusing mechanism for increased ion extraction efficiency and increased sensitivity, an integrated 1 GHz transient recorder, optimized ion optics for high resolution and accuracy, a high resolution ion gate for parent ion selection for sPSD analyses, and a microplate-type target typically holding 384 samples.

Database searches were performed using the Mascot™ MS/MS search engine (Matrix Science Limited, London, UK) utilizing the following parameters:

Database	NCBInr (680,000 sequences)
Taxonomy	all species
Enzyme	none
Mass values	average
Protein mass	unrestricted
Peptide mass tolerance	+/- 1 Da
Fragment mass tolerance	+/- 1 Da
Missed cleavages	0

Some of the synthetic peptides analyzed were C-terminal amidated (the C-terminal carboxylic acid had been modified to an amide group resulting in the loss of 1 Da). With the exception of C-terminal glycine, Mascot™ database searches cannot be modified to account for amidated peptides. Therefore, to allow for the presence of amidation, the peptide mass tolerance and fragment mass tolerance windows were increased to +/- 2 Da.

Synthetic peptides (Sigma-Aldrich, St. Louis, MO), all of which were based on sequences from naturally occurring proteins (see Table), were used to assess the performance of the curved field reflectron MALDI mass spectrometer. The instrument was used to generate seamless post source decay (sPSD) fragmentation for peptide characterization.

				database search results	1st place score	2nd place score
peptide	**sequence**	**mol wt ***	**fragment types**	**1st place protein ID****		
ACTH 18-39, human	RPVKVYPNGAEDESAEAFPLEF	2466.72	b	corticotropin	145	29
ACTH 8-37, human	FRWGKPVGKKRRPVKVYPNGAEDESAEAFPLE	3660.18	b	corticotropin	90	35
amylin fragment 8-37, human	ATQRLANFLVHSSNNFGAILSSTNVGSNTY-amide	3184.51	a, a-17, b, b-17	amyloid protein	83	32
amyloid B protein fragment 1-28, human	DAEFRHDSGYEVHHQKLVFFAEDVGSNK	3263.52	a, a-17, b, b-17, y	amyloid beta protein	95	18
amyloid B peptide precursor, human	DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV	4330.88	a, b	amyloid beta protein	122	22
atrial natriuretic peptide, chicken	MMRDSGCFGRRIDRIGSLSGMGCNGSRKN	3163.69	ND	NA	NA	NA
atrial natriuretic peptide, human	SLRRSSCFGGRMDRIGAQSGLGCNSFRY	3083.51	ND	NA	NA	NA
CGRP, fragment 8-37, human	VTHRLAGLLSRSGGVVKNNFVPTNVGSKAF- amide	3126.65	a, a-17, b, b-17	CGRP	57	35
cecropin P1, porcine	SWLSKTAKKLENSAKKRISEGIALAIQGGPR	3339.92	y	cecropin P1	120	13
β-endorphin, bovine	YGGFMTSEKSQTPLVTLFKNAIIKNAHKKGQ	3439.02	y, (b)	lipotropin beta, bovine	185	19
β-endorphin, human	YGGFMTSEKSQTPLVTLFKNAIIKNAYKKGQ	3466.05	y, (b)	lipotropin beta, human	180	21
galanin, human	GWTLNSAGTLLGPHAVGNHRSFSDKNGLTS	3158.47	y	pro-galanin	91	21
big gastrin I, human	LGPQGPPHLVADPSKKQGPWLEEEEEAYGWMDF-amide	3850.27	b, (y)	NI	NI	NI
glucagon	HSQGTFTSDYSKYLDSRRAQDFVQWLMNT	3483.81	(b), (y)	glucagon precursor	43	14
GHRF, fragment 1-29 amide	YADAIFTNSYRKVLGQLSARKLLQDIMSR-amide	3358.95	y, y-17	somatoliberin	65	25
histatin 5, human	DSHAKRHHGYKRKFHEKHHSHRGY	3037.35	y, (b)	histatin 3	83	16
insulin oxidized B chain, bovine	FVNQHLBGSHLVEALYLVBGERGFFYTPKA	3496.96	y, (b)	NA	NA	NA
neuropeptide Y fragment 3-36, porcine	SKPDNPGEDAPAEDLARYYSALRHYINLITRQRY-amide	3994.43	(y), (y-17)	neuropeptide Y	73	15
protein kinase C fragment 530-558	LLYEMLAGQAPFEGEDEDELFQSIMEHNV-amide	3355.73	y, y-17	protein kinase C	112	49
secretin, human	HSDGTFTSELSRLREGARLQRLLQGLV-amide	3040.47	(y), (y-17)	secretin	69	22
somatostatin 28	SANSNPAMAPRERKAGCKNFFWKTFTSC	3151.63	y, (b)	NA	NA	NA
vasoactive intestinal peptide, human	HSDAVFTDNYTRLRKQMAVKKYLNSILN-amide	3326.86	y, (b)	vasoactive intestinal precursor	93	21
urodilatin, human	TAPRSLRRSSCFGGRMDRIGAQSGLGCNSFRY	3509.00	ND	NA	NA	NA

* average molecular weight, MH^+

** a score greater than 68 was significant
NI = not identified in the database

Table. Properties of the synthetic peptides analyzed by MALDI MS and search results achieved with Mascot™ MS/MS search engine

Each peptide was prepared as a stock solution of 100 or 1000 pmol/µl and 10-fold serially diluted to 10 fmol/µl. The dried droplet method for MALDI target preparation was employed using 0.5 µl sample with α-cyano-4-hydroxy cinnamic acid (CHCA) as matrix (0.3 µl). Each peptide was analyzed in terms of:

1 limit of detection of the precursor (signal-to-noise of 5:1 and isotopically resolved signals)

2 limit of detection of sPSD (more than 10 fragments with a signal-to-noise of 3:1)

3 type of fragmentation generated (comparison of experimental data versus theoretical lists)

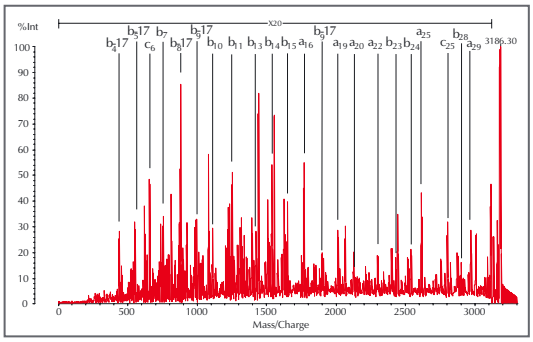
4 database search result.

Results

Precursor and sPSD spectra were generated for all 23 peptides at 10 pmoles/µl (10000 fmol/µl). Precursor spectra demonstrated that all the peptides contained a single component (data not shown). The sPSD data showed considerable variation in the type of fragmentation and in sequence coverage (see Figures 1 to 9). Those peptides with an arginine residue near the N-terminus (ACTH and amylin related peptides - Figures 1, 3 and 5) or the C-terminus (cecropin P1 and bovine beta endorphin - Figures 2 and 4) generated the cleanest spectra with extensive patterns of b or y ions respectively. This was consistent with observations made for MS/MS fragmentation methods. Those peptides with arginine residues central to the sequence (such as glucagon - Figure 6) produced a few strong signals, but significant gaps in the fragmentation pattern. This was compounded in examples where an internal disulphide bridge was present creating a cyclized peptide (somatostatin 28 - Figure 7). In such cases it was not possible to correlate the experimental and theoretical fragment masses. Other fragment types, a, a-17, b-17 and y-17 ions, were present in some peptides though this seemed to be sequence specific.

Some of the peptides were analyzed to assess both precursor ion and sPSD fragmentation sensitivity which demonstrated a 1000-fold variation. The range was most likely to be due to variation in amino acid composition. This was particularly apparent in the 10-fold and 100-fold increase in parent and sPSD sensitivity, respectively, of bovine beta endorphin compared to the human form, which differed by a single amino acid - H 27 (bovine) <-> Y 27 (human) (Data not shown). The best sensitivity was achieved for insulin oxidized B chain, (Figure 9), amyloid B fragment 1-28, CGRP fragment 8-37, cecropin P1 and bovine beta endorphin.

sPSD data from all of the peptides, except those with internal disulphide bonds, was used to search a protein sequence database for the identity of the naturally occurring protein from which the peptide was originally derived. With the exception of big gastrin I (Figure 8), all the proteins were identified unequivocally (summarized in the Table). Even peptides that produced poor sequence coverage during fragmentation, such as glucagon (Figure 6), enabled the identification of the original protein. The Probability Based Mowse Scores of Mascot™ are shown for amylin, cecropin P1, ACTH 18-39, bovine beta-endorphin, amyloid B peptide precursor and glucagon beside their respective sPSD spectra (Figures 1 to 6).



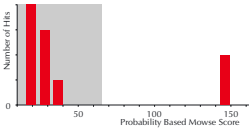
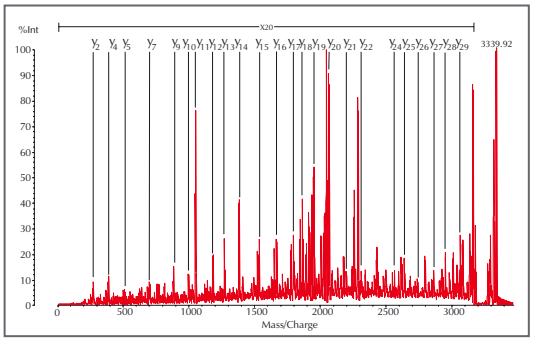


Figure 1. sPSD and database search result of ACTH 18-39



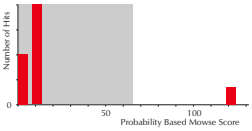
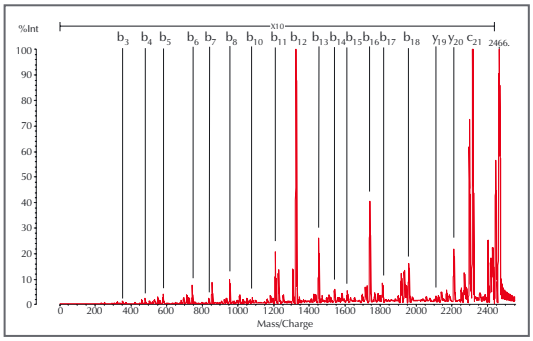


Figure 2. sPSD and database search result of cecropin P1



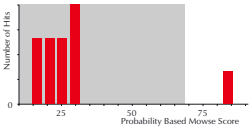
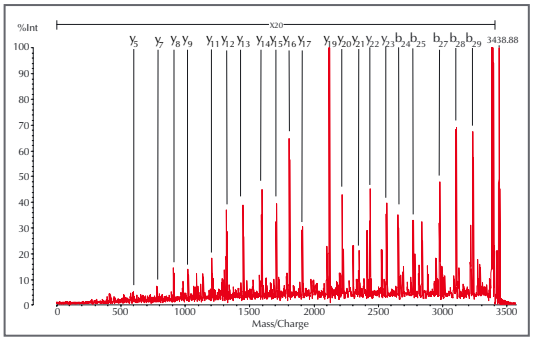


Figure 3. sPSD and database search result of amylin



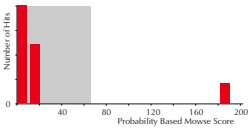


Figure 4. sPSD and database search result of bovine beta endorphin

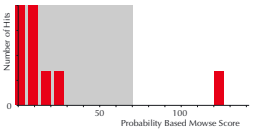


Figure 5. sPSD and database search result of amyloid B peptide precursor

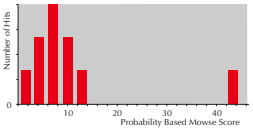
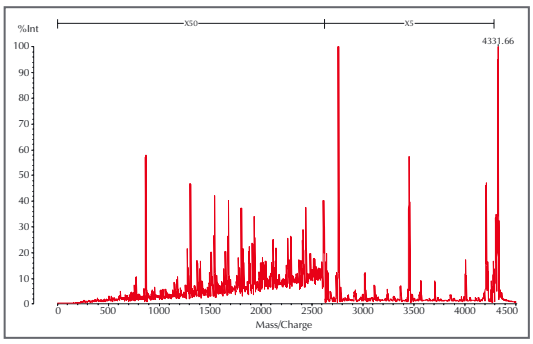
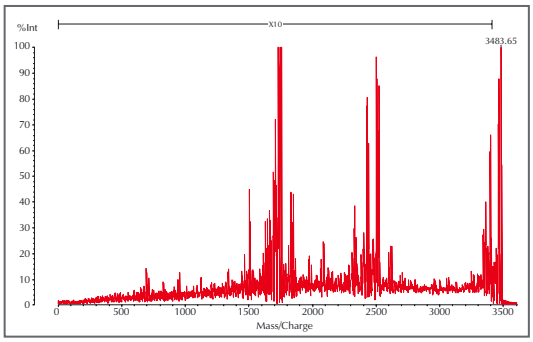


Figure 6. sPSD and database search result of glucagon





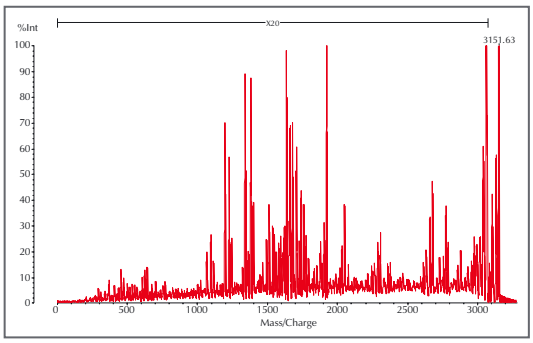


Figure 7. sPSD of somatostatin 28

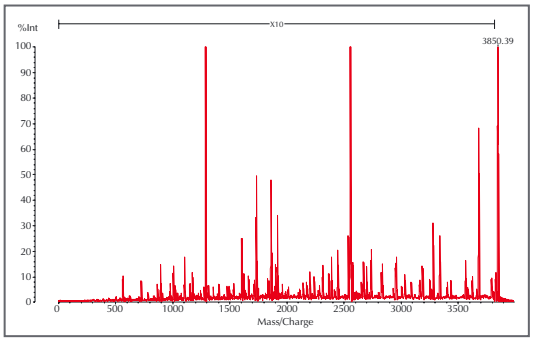


Figure 8. sPSD of big gastrin I

Conclusions

- sPSD spectra were generated rapidly with a curved field reflectron mass spectrometer – typically 60 to 90 seconds per spectrum
- fragments were observed to less than 10% of the parent mass in a single window of data acquisition
- substantial fragmentation was obtained with large peptides and in most cases excellent sequence coverage obtained
- sPSD mass data was used to successfully identify the original proteins using the Mascot™ MS/MS search engine, with broad search parameters – mass tolerance (+/- 1 or 2 Da), no species filter and no digest conditions

These factors make sPSD an ideal candidate for protein identification using a shotgun approach based on fragmentation and database searching using only high mass peptides from protein digests.

REFERENCES

1. Cornish, T., Cotter, R. (1993) Rapid Comm. Mass Spectrom., 7, 1037

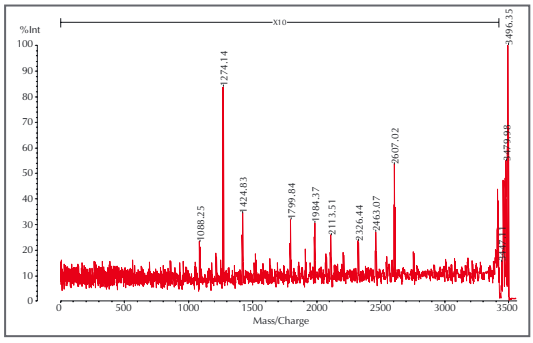


Figure 9. sPSD of 10 fmol/µl insulin oxidized B chain

SHIMADZU BIOTECH
bringing analysis to life