Peptide sequencing using chemically assisted fragmentation (CAF) and Ettan MALDI-ToF Pro mass spectrometry

Key words: MALDI-ToF, PSD, peptide sequencing, chemically assisted fragmentation

Introduction

Peptide mass fingerprinting using matrix-assisted laser desorption/ionization (MALDI) time-of-flight (ToF) mass spectrometry has become a major tool for identifying proteins in proteomics research. The method provides high speed, sensitivity, and mass accuracy, but for a substantial fraction of the proteins analyzed, identification is not definitive. Under these circumstances, amino acid sequence information from one or more peptides is required for unambiguous identification. Such information can be readily obtained using the Ettan™ CAF-MALDI Sequencing Kit in conjunction with Ettan MALDI-ToF Pro mass spectrometer (Fig 1).

Ettan MALDI-ToF Pro has a quadratic field reflectron that allows fast post-source decay (PSD) analysis, focusing all fragments in a single run, independent of size. The accompanying software includes tools for automated protein identification from data generated using the Ettan CAF-MALDI Sequencing Kit.

The chemistry of the Ettan CAF-MALDI Sequencing Kit, based on a new class of water-stable sulfonation reagents, eliminates the most common problems associated with PSD analysis. Ettan CAF-MALDI chemistry greatly improves the fragmentation efficiency of the peptides and also simplifies the interpretation of their fragmentation spectra.

The special features of the Ettan MALDI-ToF Pro mass spectrometer and the simple and robust derivatization protocols using the Ettan CAF-MALDI Sequencing Kit combine to provide the proteomics researcher with:

- Identification of increased numbers of proteins
- Rapid, sensitive, and precise peptide sequencing
- Characterization of phosphorylation sites

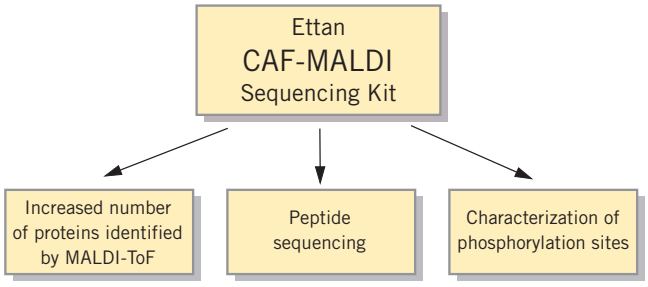


Fig 1. Ettan CAF MALDI Sequencing Kit in conjunction with Ettan MALDI-ToF Pro offers a multifaceted approach to protein identification.

Strategy

CAF-MALDI chemistry is based on the introduction of a negatively charged group to the N-terminus of peptides generated by tryptic digestion. The theory is that, after derivatization, the formation of a positively charged ion (net charge) requires that two protons be introduced into the peptide. One of these protons will primarily reside in the basic C-terminal side chain, while the other has a higher degree of freedom to resonate in the peptide backbone, assisting fragmentation. In the appropriate analysis mode, only the y-ions, which retain a net positive charge, will become separated in the reflectron. The N-terminal fragment ions will be neutral and non-detectable.

The derivatization reaction is divided into two steps (Fig 2). The first step converts the ε-amino group of each lysine side chain to homo-arginine (generating a mass addition of 42 atomic mass units [amu]). This step is necessary to protect the lysine side chains from sulfonation in the following step. The second step introduces a sulfonic acid group at the N-terminus (generating a mass addition of 136 amu).

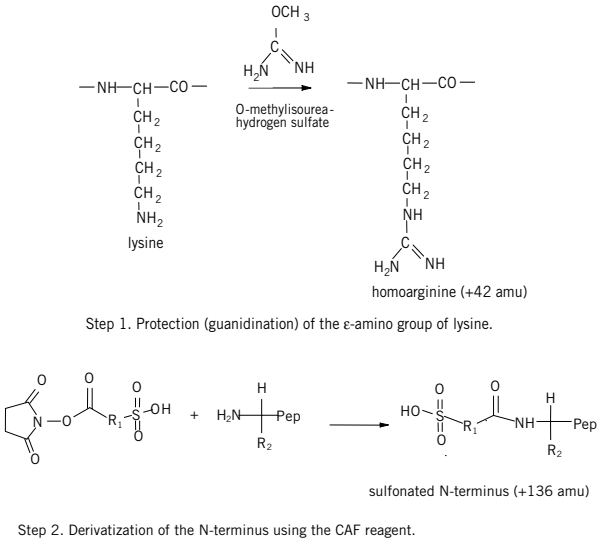


Fig 2. The two steps of the derivatization reaction.

Products used	**Code no.**
Ettan CAF-MALDI Sequencing Kit	17-6002-97

Kit contents:
Lysine modifier: O-methylisoureahydrogen sulfate
CAF reagent: NHS-ester containing a sulfonic acid group
Stop solution: 50% hydroxylamine solution
Lysine modifier buffer: 0.25 M sodium carbonate, pH 11.7
CAF buffer: 0.25 M sodium bicarbonate, pH 9.4
Control peptides: Lys-peptide: HGTVVLTALGGILK (M_r 1378)
Arg-peptide: ADSGEGDFLAEGGGVR (M_r 1536)

ACH-Cinnamic Acid, Ettan Chemicals	17-6002-80
Trifluoroacetic Acid, Ettan Chemicals	17-6002-76

Methods and Instrumentation

Derivatization reactions of tryptic peptides were performed using the instructions and reagents in the Ettan CAF-MALDI Sequencing Kit. Additional reagents included C_{18} ZipTip (Millipore), ACH-Cinnamic Acid, Trifluoroacetic Acid, acetonitrile (Sigma), and ultrapure water (18 MΩ/cm). Samples and matrix were loaded on the Ettan MALDI-ToF Pro target as described in the kit instructions.

All spectra were obtained with the ion gate set for the precursor ion (i.e. parent ion) in PSD analysis mode.

The fragment masses of CAF-labelled peptides were determined using the Ettan MALDI-ToF Pro software package. Proteins were identified using an automated protein identification program included in the system.

Several types of analysis were performed to show examples of the performance of the Ettan CAF-MALDI Sequencing Kit in peptide sequencing and identification of phosphorylation sites, as well as to provide data on sensitivity with different samples.

Results and Discussion

Peptide sequencing

The simplified PSD spectra, showing y-ions only, allow fast and accurate sequence determination, where the distance between two consecutive peaks corresponds to the mass of an amino acid. Figure 3 shows an example of the complete peptide sequencing of a model peptide comprising 16 amino acids.

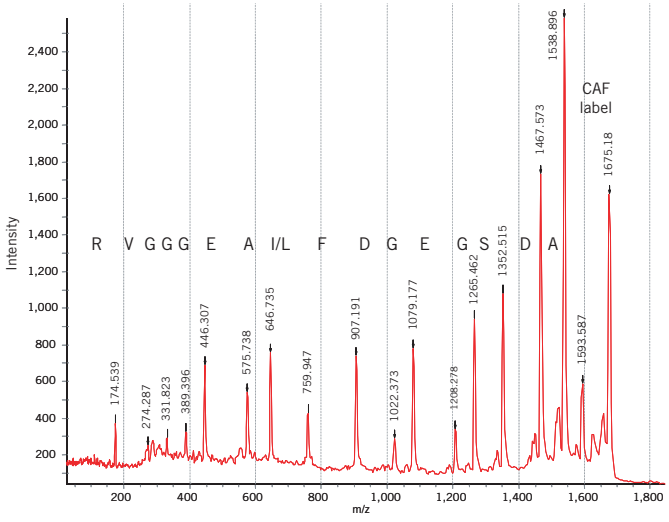


Fig 3. Peptide sequencing of a model peptide comprising 16 amino acids.

Identification of phosphorylation sites

The Ettan CAF-MALDI Sequencing Kit can also be used for rapid identification of phosphorylation sites. An example of peptide sequencing of a phosphorylated peptide is shown in Figure 4. The peptide contains two phosphorylated tyrosine residues. The phosphorylated tyrosine residues remain intact during PSD and can easily be identified by a mass difference of 243 amu (163 + 80, HPO_3) between them.

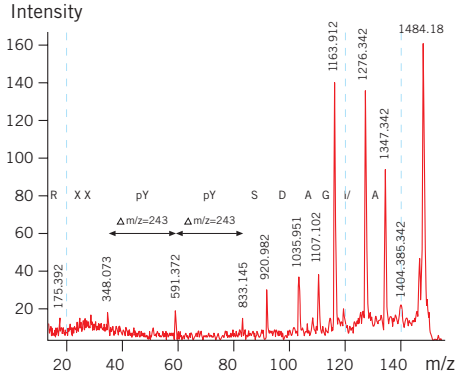


Fig 4. Peptide sequencing of a peptide containing two phosphorylated tyrosine residues. The sequence is ALGADSpYpYTAR (two fragment peaks are missing from the spectrum, each indicated by an X).

Sensitivity

The sensitivity limit, defined as the lowest amount of sample needed for protein identification from CAF-MALDI data, was determined for both lysine- and arginine-terminated peptides from different sources. Model peptides, in-solution, and in-gel digests were all investigated. Figure 5 shows the results when 40 fmol of a lysine-terminated peptide was analyzed. Figure 6 shows the results with 40 fmol of an arginine-terminated peptide. In both cases, the sequence could be determined and the protein identified.

The sensitivity limit for an in-solution digest of horse myoglobin was 40 fmol on the MALDI target for both lysine- and arginine-terminated peptides (see Figs 7A and B).

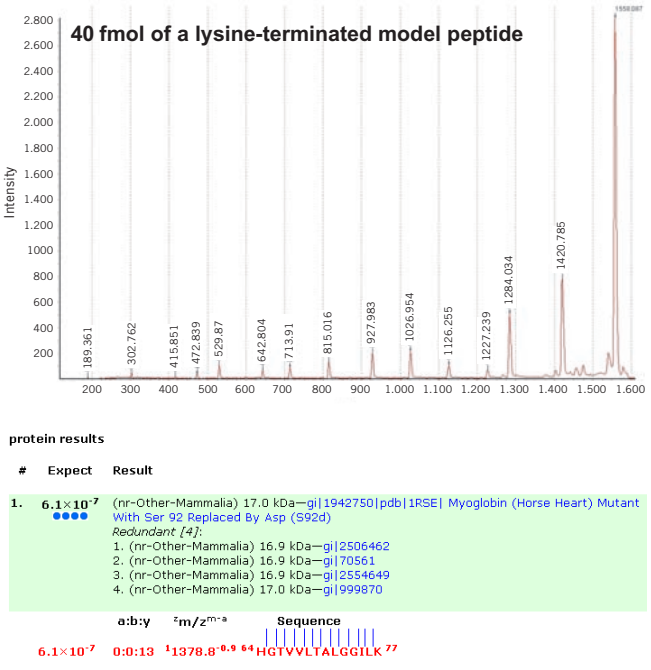


Fig 5. Sensitivity of a Lys-terminated model peptide (containing the same sequence as a Lys-terminated tryptic peptide of horse myoglobin). Four hundred fmol were derivatized, and one-tenth of the sample (40 fmol) was analyzed.

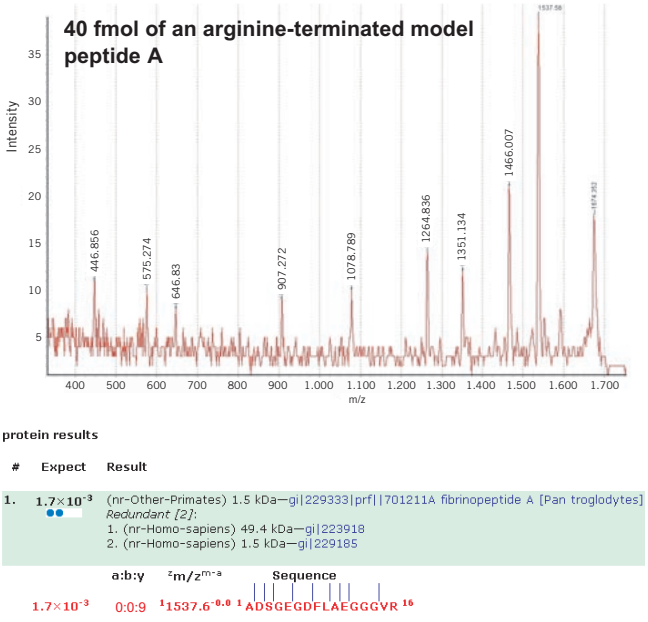
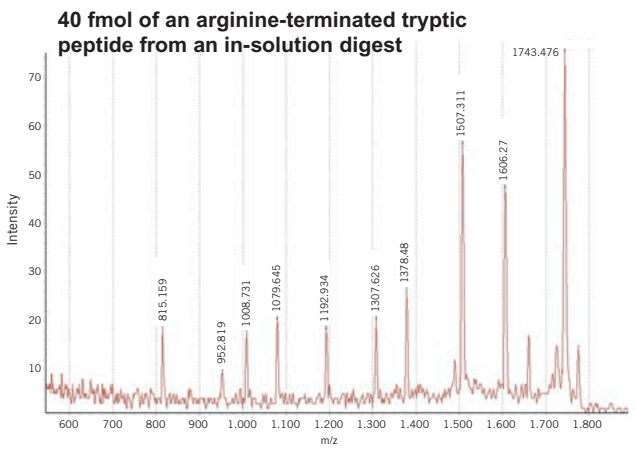


Fig 6. Sensitivity of an Arg-terminated model peptide (containing the same sequence as an Arg-terminated tryptic peptide of fibrinopeptide A). Four hundred fmol were derivatized, and one-tenth of the sample (40 fmol) was analyzed.

Finally, phosphorylase B was used to investigate the sensitivity limit for an in-gel digest. Seven-hundred and fifty fmol of protein was spiked into a gel plug. The protein was digested with trypsin and the peptides were thereafter extracted from the gel plug, followed by lysine protection and CAF-derivatization. Figure 8 shows the PSD spectrum of one of the peptides of the digest, when one tenth of the sample was loaded onto the MALDI target.

Fig 7a.

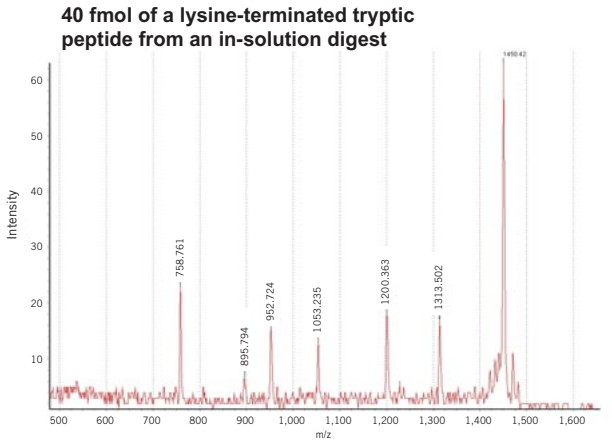


protein results

#	Expect	Result
1.	2.0×10^{-2}	(nr-Other-Mammalia) 17.1 kDa—gi|127680|sp|P02173|MYG_ORCOR MYOGLOBIN *Redundant [1]*: 1. (nr-Other-Mammalia) 17.1 kDa—gi|127655

a:b:y	$^{z}m/z^{m-a}$	Sequence
2.0×10^{-2}	0:0:8	^{1}1606.3$^{-1.5}$ ^{17}VEADLAGHGQDILIR^{31}

Fig 7b.

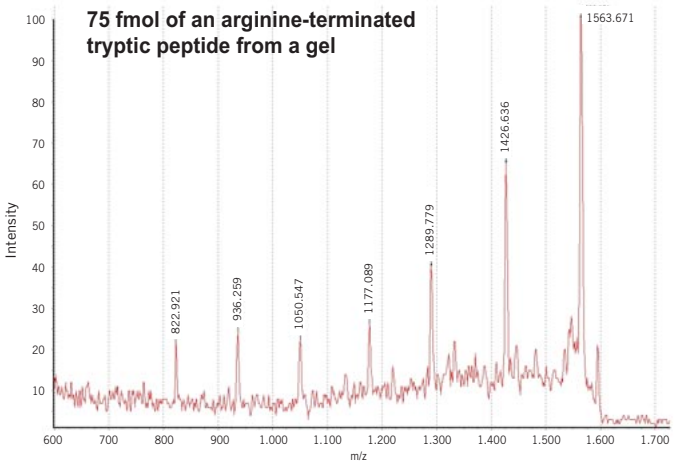


protein results

#	Expect	Result
1.	0.08	(nr-Other-Mammalia) 17.1 kDa—gi|127638|sp|P02192|MYG_BOVIN MYOGLOBIN *Redundant [13]*: 1. (nr-Other-Primates) 16.9 kDa—gi|127667 2. (nr-Other-Mammalia) 17.0 kDa—gi|494711 3. (nr-Other-Mammalia) 16.9 kDa—gi|2506462 4. (nr-Other-Mammalia) 16.9 kDa—gi|127643 5. (nr-Other-Mammalia) 16.9 kDa—gi|127694 6. (nr-Other-Mammalia) 16.9 kDa—gi|70561 7. (nr-Other-Mammalia) 17.0 kDa—gi|999870 8. (nr-Other-Mammalia) 17.0 kDa—gi|127668 9. (nr-Other-Mammalia) 17.0 kDa—gi|70559 10. (nr-Other-Mammalia) 16.9 kDa—gi|2914321 11. (nr-Other-Mammalia) 17.0 kDa—gi|2811090 12. (nr-Other-Mammalia) 16.9 kDa—gi|2554649 13. (nr-Other-Mammalia) 17.0 kDa—gi|1942750

a:b:y	$^{z}m/z^{m-a}$	Sequence
0.08	0:0:5	^{1}1271.5$^{-0.9}$ ^{33}LFTGHPETLEK^{43}

Fig 7a-b. Analysis of 40 fmol of a trypic peptide of horse myoglobin from an in-solution digest. In Panel A, the peptide is Arg-terminated, and in Panel B it is Lys-terminated. In both cases, 400 fmol were derivatized, and one-tenth (40 fmol) was analyzed.



protein results

#	Expect	Result
1.	0.50	(nr-Other-Mammalia) 96.0 kDa—gi|442605|pdb|1ABB|A Chain A, Glycogen Phosphorylase (E.C.2.4.1.1) Complex With Pyridoxal-5'-Diphosphate *Redundant [18]:* 1. (nr-Other-Mammalia) 97.7 kDa—gi|8569507 2. (nr-Other-Mammalia) 97.7 kDa—gi|14916625 3. (nr-Other-Mammalia) 97.6 kDa—gi|223003 4. (nr-Other-Mammalia) 97.8 kDa—gi|1664 5. (nr-Homo-sapiens) 97.5 kDa—gi|5032009 6. (nr-Other-Mammalia) 97.7 kDa—gi|8569398 7. (nr-Mus-musculus) 97.7 kDa—gi|6755256 8. (nr-Rattus) 97.8 kDa—gi|1730556 9. (nr-Other-Mammalia) 96.2 kDa—gi|6730143 10. (nr-Other-Mammalia) 97.7 kDa—gi|1633147 11. (nr-Homo-sapiens) 97.6 kDa—gi|87576 12. (nr-Other-Mammalia) 97.7 kDa—gi|14916628 13. (nr-Other-Mammalia) 97.6 kDa—gi|3318708 14. (nr-Homo-sapiens) 97.4 kDa—gi|225897 15. (nr-Other-Mammalia) 96.4 kDa—gi|231300 16. (nr-Rattus) 97.6 kDa—gi|479548 17. (nr-Other-Mammalia) 97.7 kDa—gi|8569323 18. (nr-Other-Mammalia) 97.8 kDa—gi|6093713

	a:b:y	$^{z}m/z^{m-a}$	Sequence
0.50	0:0:5	$^{1}1426.0^{-1.6}$ 390	HLQIIYEINQR 400

Fig 8. Sequencing of an Arg-terminated tryptic peptide of phosphorylase B isolated from a 2-D gel. Seven-hundred and fifty fmol were derivatized, and one-tenth of the protein from the gel was analyzed.

Conclusions

Ettan CAF-MALDI Sequencing Kit in combination with single-run PSD analysis using Ettan MALDI-ToF Pro and its automated software for protein identification from fragmentation data enables sensitive and rapid peptide sequencing of tryptic peptides. Key advantages of the kit include:

- Specific and complete N-terminal derivatization under aqueous conditions.
- Optimized derivatization protocol for fast and robust sample preparation.
- Increased fragmentation in PSD.
- PSD spectra allow easier interpretation of data.
- Precise identification of phosphorylation sites.
- High sensitivity. Low fmol levels of model peptides, in solution or in-gel digests can be used for protein identification by PSD using the Ettan CAF-MALDI Sequencing Kit.

References

1. Keough, T. et al., A method for high-sensitivity peptide sequencing using post-source decay matrix-assisted laser desorption ionization mass spectrometry. *Proc. Natl. Acad. Sci. USA.* **96**, 7131–7136 (1999).
2. Liminga, M. et al., Water-stable chemistry for improved amino acid sequencing by derivatization post-source decay (dPSD) using Ettan MALDI-ToF with a quadratic field reflectron. *Proc. 49th Conf. Mass Spectrometry and Allied Topics,* Chicago, IL (2001).
3. Keough, T. et al., Atmospheric pressure matrix-assisted laser desorption/ionization ion trap mass spectrometry of sulfonic acid derivatized tryptic peptides. *Rapid Commun. Mass Spectrom.* **15**, 2227–2239 (2001).

to order:

Asia Pacific Tel: +852 2811 8693 Fax: +852 2811 5251 **Australasia** Tel: +61 2 9899 0999 Fax: +61 2 9899 7511 **Austria** Tel: 01 576 0616 20 Fax: 01 576 0616 27 **Belgium** Tel: 0800 73 888 Fax: 03 272 1637 **Canada** Tel: 1 800 463 5800 Fax: 1 800 567 1008 **Central, East, South East Europe** Tel: +43 1 982 3826 Fax: +43 1 985 8327 **Denmark** Tel: 45 16 2400 Fax: 45 16 2424 **Finland & Baltics** Tel: +358 (0)9 512 3940 Fax: +358 (0)9 512 1710 **France** Tel: 0169 35 67 00 Fax: 0169 41 9677 **Germany** Tel: 0761 4903 401 Fax: 0761 4903 405 **Italy** Tel: 02 27322 1 Fax: 02 27302 212 **Japan** Tel: 81 3 5331 9336 Fax: 81 3 5331 9370 **Latin America** Tel: +55 11 3667 5700 Fax: +55 11 3667 87 99 **Middle East and Africa** Tel: +30 (1) 96 00 687 Fax: +30 (1) 96 00 693 **Netherlands** Tel: 0165 580 410 Fax: 0165 580 401 **Norway** Tel: 2318 5800 Fax: 2318 6800 **Portugal** Tel: 21 417 7035 Fax: 21 417 3184 **Russia & other C.I.S. & N.I.S.** Tel: +7 (095) 232 0250,956 1137 Fax: +7 (095) 230 6377 **South East Asia** Tel: 60 3 8024 2080 Fax: 60 3 8024 2090 **Spain** Tel: 93 594 49 50 Fax: 93 594 49 55 **Sweden** Tel: +46 18 612 19 00 Fax: +46 18 612 19 10 **Switzerland** Tel: 01 802 81 50 Fax: 01 802 81 51 **UK** Tel: 0800 616 928 Fax: 0800 616 927 **USA** Tel: +1 800 526 3593 Fax: +1 877 295 8102

Produced by Wikstroms, Sweden 1020379, 06.2002
Printed matter. Licence 341 051

Amersham Biosciences