

SMASH: Software for Identification of Microwave-Assisted Hydrolyzed Proteins by MALDI-TOF MS

Weijie Yang¹; Cunjie Zhang²; Bin Ma¹; Gilles Lajoie²

¹Department of Computer Science; ²Department of Biochemistry, University of Western Ontario, London, ON, Canada

Overview

First automated software for analyzing MALDI-TOF MS of microwave assisted acid hydrolyzed proteins.

Introduction

Mass fingerprinting after proteolytic digestion is widely used for rapid protein identification with MALDI-TOF MS. However, mass fingerprinting often does not provide sufficient information to identify proteins with high confidence. Recently, a new protein identification method using MALDI-TOF MS has been proposed by Zhong et al. (Nature Biotech. 2004, 22:1291-1296). In this method, a protein is acid hydrolyzed by exposure to microwave irradiation. Two series of polypeptide ladders from N- and C- termini can be detected by MALDI-TOF MS. However, no software is available for analyzing the data and the interpretation has to be done manually. In this poster, a new software for the automated analyses of these data is described.

Methods

The microwave assisted acid hydrolysis was carried out as described by Zhong et al. The sample was spotted on the MALDI plate using the one layer method. MALDI-MS experiments were performed on either a MALDI-R (Micromass) or a Reflex IV MALDI (Bruker, linear mode) mass spectrometer. Data were exported as peak list text files.

The software processes the data in the following steps

1. *Peaks Selection*
 2. *Tags Formation*
- By examining the neighboring peaks, many length-4 sequence tags are generated.

3. *Candidate Finding*
The sequence tags are matched in a protein database to find 1000 protein candidates.

4. *Protein Identification*
Each candidate is further compared with the spectrum to find the best-scoring protein. The length-4 tags found in step 2 were assembled to longer tags (Fig 1 and 2). Candidates with longer tags are ranked higher.

Experimental Results

The algorithm was implemented in an automated computer program using java programming language. Nine sets of MALDI MS spectra were obtained from horse myoglobin, cytochrome C, lysozyme, bovine pancreatic ribonuclease A, and beta-casein bovine, by using MALDI-R and MALDI linear respectively. The following table is a list of the spectra obtained.

Protein	Method	Matched Peptides	
		N-Terminal	C-Terminal
MYO_HORSE	MALDI-R	WQQVL	KAVIDNR LELAKTAMAGQADAGTDGPH
MYO_HORSE	MALDI-L	KVEADI	ADAGFDG
CYTO_BOVIN	MALDI-R	KGGKHKTG GRKTGQ	
CYTO_BOVIN	MALDI-L	PNLHGLFGRK	
LYC_CHICK	MALDI-R	RHGLDN	DTGKCR
LYC_CHICK	MALDI-L	AAMKRHG	DTGZC
RNP_BOVIN	MALDI-R	MDSSTSA	PNGEC AQTTYK
RNP_BOVIN	MALDI-L		HKNAQT KSSGTERC
B_CASEIN_BOVIN	MALDI-R	ESLSSSEE	GLVPEQYL

SwissProt protein database was used to search against by the software. There are 163,496 proteins in the version we used. The software correctly identified all five proteins using either MALDI-R or MALDI linear. The sequence tags generated and used by the software to identify the proteins were given in the table. Additionally, the correct proteins were listed as the top scoring proteins in all but one case. As can be expected, homologous proteins from different organisms were also listed as matches. These can be disregarded if the source organism is known.

This initial test demonstrates that the automated software works well for protein identification by MALDI-TOF MS of microwave-assisted acid hydrolyzed proteins. It is worth noting that the bovine pancreatic ribonuclease A protein contains unknown amount of glycosylation and b-casein is a phosphoprotein but the software succeeded in its identification with the MALDI-R data.

Reference:

1. Zhong et al. Protein sequencing by mass analysis of polypeptide ladders after controlled protein hydrolysis. Nature Biotechnology (2004), 22, 1291-1296.
2. Zhong et al. Microwave-assisted acid hydrolysis of proteins combined with liquid chromatography MALDI MS/MS for protein identification. JASMS (2005), 16(4), 471-481.

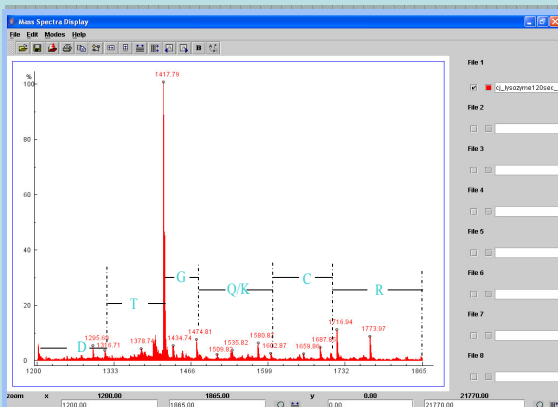


Fig 1

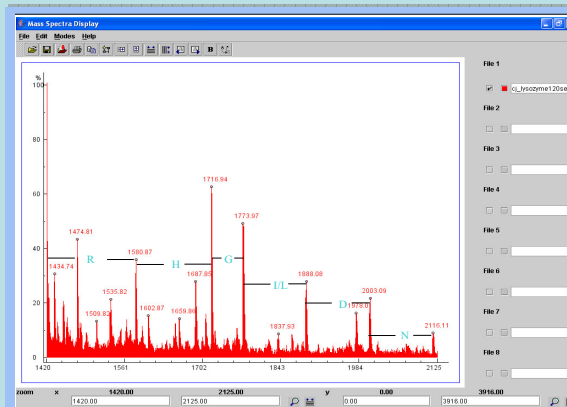


Fig 2