PARALLEL POST SOURCE DECAY VERSUS PEPTIDE MASS FINGERPRINTING – ENHANCED PROTEIN IDENTIFICATION ON A MALDI TOF MS

ASMS 2005

Emmanuelle Claude[1]; Marten Snel[1]; Thomas Franz[2]; Anja Bathke[2]; Therese McKenna[1]; **James Langridge[1]**;
[1]Waters Corporation, Manchester, United Kingdom;
[2]European Molecular Biology Laboratory, Heidelberg, Germany;

OVERVIEW

- In this poster we use MALDI MS and PSD, to characterize 2D gel isolated proteins originating from *E.coli*
- The success rate for protein identification using Peptide Mass Fingerprinting is compared to that obtained using PSD MX
- The incorporation of PSD MX™ information is clearly shown to improve the success rate for protein identification by databank searching
- A significant increase in the confidence of protein assignment is observed, when PSD fragment ion information is incorporated into the databank search.

INTRODUCTION

MALDI TOF has been extensively used as a tool for protein identification. Typically proteins are separated from complex mixtures using gel based approaches. Isolated proteins are enzymatically digested and MALDI MS and/or MALDI Post Source Decay (PSD) experiments are performed. Despite the wide acceptance and success of this technique, PMF under certain circumstances will fail. For example, in cases where the protein sequence is poorly characterized; if a mixture of proteins with a wide dynamic range is present; or if the number of tryptic peptides produced by the proteolytic digestion is low, then PMF maybe be frustrated. The advantage of performing a PSD, or MS/MS analysis, of the sample is that additional specificity is provided. Even limited fragment ion information, can provide sufficient information to overcome the limitations of PMF and result in unambiguous protein identification.

Here we show a novel form of PSD analysis (Parallel PSD) allowing fragment ions from multiple precursor ions to be analyzed in parallel. In traditional PSD the selection of precursor ions with a timed electrostatic ion-gate is required. In the 'parallel' PSD approach, the ion gate is not required as fragment ions from all of the precursor ions are acquired simultaneously. A deconvolution algorithm is then used to match precursor ions with fragment ions. This novel approach simplifies the PSD experiment, as no decision has to be made on which precursor ions to select, reduces sample consumption, and increases the number of peptides analyzed. This approach improves sensitivity, avoids the need for precursor ion selection and is highly amenable to automation.

In this poster we compare and contrast results obtained by MALDI MS, to the data from PSD MX, for proteins originating from an *E. coli* cell lysate which have been separated by 2D-gel electrophoresis.

EXPERIMENTAL

Sample preparation

A 250 µg sample of a lyophilised *E. coli* protein sample (Bio-Rad, Hercules, CA) was separated by 2D-gel electrophoresis. The proteins were visualised by Coomassie G250 staining (Merck, Darmstadt, Germany).

Gel spots were excised from the 2D-gel using a ProteomeWorks Plus spot cutting robot (Bio-Rad, Hercules, CA). The gel pieces were deposited into one 96-well microtiter plate with between1-5 gel pieces per well.

The excised gel spots were processed using the MassPREP™ Station liquid handling robot (Waters, Milford, MA). The protocol used included gel spot de-staining, reduction, alkylation, tryptic digestion, and finally peptide extraction. The tryptic peptide solutions (1µL) were spotted with 1µL of alpha-cyano-4-hydroxycinnamic acid matrix, 3 mg/mL (1/1 v/v MeCN/0.1 % aqueous TFA).

Mass spectrometry

Acquisition

- All MS spectra were acquired on a MALDI micro MX™ mass spectrometer, (Waters, Manchester, UK). Data were acquired in the positive ion reflectron-mode over the m/z range 1000-3000.
- The instrument was operated in a fully automated manner, using the data to direct the settings.
- For each **MS spectrum,** 100 laser shots were summed.
- Six PSD segments, with each segment acquired at two reflectron voltages, were acquired per sample. For each segment 400 laser shots were summed, giving a total of 2,400 laser shots per **PSD MX** experiment.

Calibration

- A digest of alcohol dehydrogenase (Waters, Milford, MA) spotted at a concentration of 500 fmol on target was used to generate a multi-point external calibration.
- **MS** data was externally lock mass corrected using a synthetic polyproline peptide P14R (Sigma, St Louis, MO); $[MH]^+$ = 1533.8582 Da at a concentration of 250 fmol on target.
- **PSD MX** data were calibrated using the PSD fragments generated from 2pmoles of P14R on-target.

Data processing and searching

- ProteinLynx™ Global SERVER 2.2 (Waters, Manchester, UK) was used for data processing and database searching.
- **MS** (PMF) data were smoothed, background subtracted and de-isotoped using the MaxEnt Lite algorithm.
- **PSD MX** data were smoothed, background subtracted and centroided.
- Processed **MS** data generated were automatically submitted to a databank search against a Swiss-Prot database (v40). Database searching was performed using the Mascot search engine (Matrix Science Ltd, London, UK), with results collated and displayed using the ProteinLynx Global SERVER v2.2 interface.
- For **PSD MX** experiment, fragment ions were matched to their precursor ions using a deconvolution algorithm, implemented in MassLynx™ and ProteinLynx Global SERVER 2.2. A peak list (pkl) was generated for each spot and automatically submitted to a databank search, against a Swiss-Prot database (v40).

RESULTS

Databank search results

A summary of the proteins identified by the different techniques is illustrated graphically in Figure 1.

Two different types of databank search were performed on the MS or PSD data obtained from each sample well :-

1. MS data were submitted for a peptide mass fingerprint search
2. PSD data were submitted using the fragment ion information. The search results from the fragment ion search were formatted as a "protein report", described here as PSD MX.

The two different databank searches above, (1&2), have a 95 % confidence limit, which is governed by the search parameters used. The 95% confidence limit for PMF and PSD MX searches, in these experiments, is a score of 64.

- A summary of the resulting data indicates that using the traditional PMF approach, 53 proteins were confidently identified from 96 sample wells with search scores ≥ 64.
- With identification of protein mixtures from some 2D-gel spots, these search results represent successful identification from ***54 % of the samples***.
- In contrast, databank searching of the PSD MX data, in combination with the PMF information, resulted in confident identification of 81 proteins, equivalent to a protein identification success rate of ***82% from the samples*** analyzed.
- Of the 81 proteins identified, 30 proteins were only unambiguously identified when the PSD MX data was included. In addition two proteins were found exclusively by using the PMF approach.

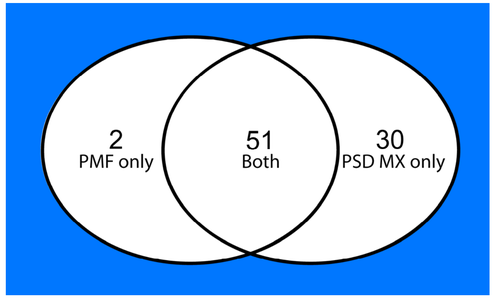

Figure 1: Results from the MALDI PMF and PSD MX analysis of the *E.coli* 2D gel samples The numbers indicate the proteins identified by the two techniques. The total number of samples analysed was 96.

Examples of when PMF may fail and the advantage of using PSD MX data.

1. Insufficient tryptic peptides

PMF analysis may fail when an insufficient number of tryptic peptides are produced by the enzymatic process. An example of this is shown in the MALDI mass spectrum below (Figure 2). Subsequent databank searching of the PMF information resulted in an ambiguous result.

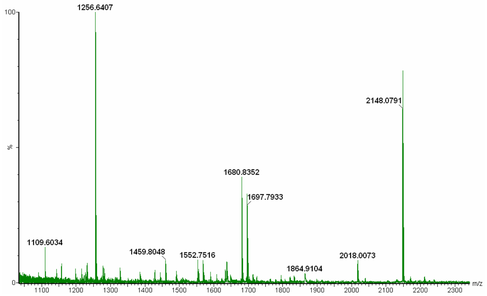

Figure 2: MS spectrum of sample A,7. PMF databank searching resulted in ambiguous result.

However, the PSD MX experiment provided specific fragment ion information and using this information in a databank search, it was possible to unambiguously identify Arginine-binding periplasmic protein 2 (ARTJ_ECOLI) (see figure 3).

Databank searching of the PSD spectra led to identification of the parent protein with a score of 75. When the PMF and PSD MX information were combined, the score increased up to 125; with only three peptides matched in total.

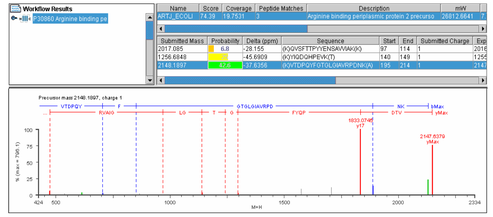

Figure 3: Results from the fragment ion PSD MX database search. Highlighted is the PSD MX spectrum for m/z 2148.1997.

2. The presence of protein mixtures in a sample

- Analysis of the 2D-gel spot contained in position D, 4 by MALDI PMF identified two proteins, SERC_ECOLI (MOWSE PMF score of 110) and TYRA_ECOLI (MOWSE PMF score of 74).
- Searching of the PSD MX data identified an additional protein as the top scoring hit, ALF_ECOLI.
- Figure 4 shows the MS spectrum of sample D,4, annotated with the matching peptides from the 3 proteins.

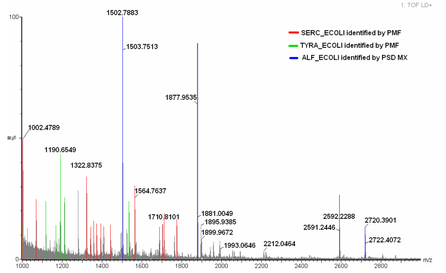

Figure 4: MALDI MS spectrum of sample D,4. Peaks highlighted in red have been matched to SERC_ECOLI by PMF. Peaks highlighted in green have been matched to TYRA_ECOLI by PMF. The peaks highlighted in blue have been matched to ALF_ECOLI with PSD MX.

Both SERC_ECOLI and TYRA_ECOLI were identified by several low-intensity peptides, 14 and 8 respectively. In comparison, ALF_ECOLI, was identified from just 4 peptides present in the mixture. This is below the minimum number of peptides, (typically 5), required for a confident identification from a PMF database search. However, the four peptides detected all gave fragments by PSD MX and ALF_ECOLI could be confidently assigned, due to the specificity of this information.

In addition, the power of PSD MX was used to correct the assignment of one peptide in this mixture. Initially, after the PMF search, the peptide at 1877.9535 was matched to TYRA_ECOLI. However, following the PSD MX databank search, the ion was re-assigned as a tryptic peptide from ALF_ECOLI (IFDFVKPGVITGDDVQK), with 9 fragment ions matched to the sequence (Figure 5).

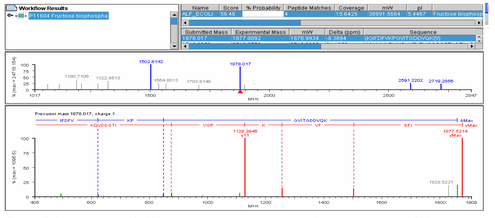

Figure 5: Results of the MALDI PSD MX database search from sample D,4. Highlighted is the PSD MX spectrum of m/z 1877.9535. Initially this peptides was assigned by PMF to a sequence from TYRA_ECOLI. However, the PSD information above assigned a different peptide sequence, originating from ALF_ECOLI.

3. Added confidence in protein assignment

The 2D gel spot located at position F5, gave rise to the identification of the E.coli protein Succinyl CoA synthetase alpha chain when analysed by PMF. If the PSD MX information is utilised then this gives rise to added confidence in the protein identification, with the score rising from 120 to 277. In this case, the 4 major peptides fragmented sufficiently to generate sequence information (figure 6).

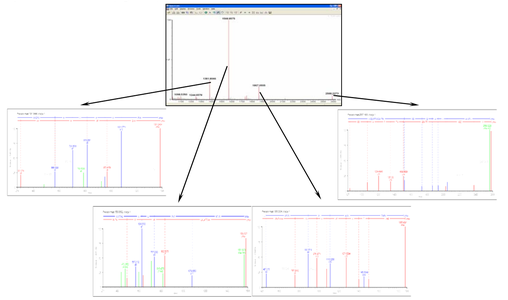

Figure 6: PLGS display of the fragment only PSD MX database search. It highlights the PSD MX spectrum of the 4 peptides which gave good PSD MX information on sample F,5.

CONCLUSIONS

- Peptide mass fingerprinting resulted in identification of 54 % of the 96 2D-gel samples analyzed.
- Analysis of the same samples by PSD MX increased the success rate from 54 % to 83 %.
- PSD MX provides additional, specific, information that can be vital in obtaining unambiguous protein identification. Examples that we have shown are:
  - If the digestion process produces only a limited number of peptides
  - When a 2D-gel spot is a mixture of three digested proteins.
- This approach can lead to highly confident protein identification for ambiguous PMF databank search results, providing good quality MS/MS information.

TO DOWNLOAD A COPY OF THIS POSTER VISIT WWW.WATERS.COM/POSTERS

Lit Code Number 720001244EN