Waters

ENHANCED IDENTIFICATION OF 2D-GEL ISOLATED PROTEINS FROM *E. COLI* USING THE NOVEL MALDI PSD MX

R. O'Malley[1], E. Claude[1], M. Snel[1], T. Franz[2], A. Bathke[2], T. McKenna[1] and J. Langridge[1]; [1] Waters Corporation, MS Technologies Centre, Manchester, UK [2]; European Molecular Biology Laboratory, Heidelberg, Germany

Overview

- In this poster we describe use of the new PSD MX technique for the identification of 2D-gel isolated proteins.
- The incorporation of PSD MX information is clearly shown to improve the success rate for protein identification.
- In this study PSD MX increased the number of sample spots from which protein identifications were obtained by 57%, compared to peptide mass fingerprinting alone.
- The overall success rate for protein identification across the 96 in-gel digested samples was 83%, using both the MS and PSD MX information.

Introduction

Peptide mass fingerprinting (PMF) is a well-established technique for the identification of proteins from MALDI TOF MS data.[1]

Despite the wide acceptance and success of this technique, PMF under certain circumstances will fail. In cases where the protein sequence is poorly characterised, if a mixture of proteins with a wide dynamic range is present in the sample or if the number of tryptic peptides produced by the proteolytic digestion is low then PMF maybe be frustrated.

In this study, we evaluate a new approach, parallel post source decay (PSD MX), that provides complementary structural information to MALDI MS information.

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 has been developed to match precursor ions with fragment ions.[2] 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 analysed. The PSD MX experiments are performed fully automatically.

In this poster we have compared the success rate for the identification of proteins using PSD MX data with that obtained when using PMF alone. The proteins analysed were obtained from an *E. coli* cell lysate that was separated by 2D-gel electrophoresis and subsequently subjected to *in-gel* tryptic digestion.

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 Coomasie staining.
- The gel pieces were excised and deposited in a 96-well microtiter plate with 1-5 gel pieces per well.
- The gel samples were processed using a MassPREP station liquid handling robot (Waters, Milford, MA).
- The extracted peptide solutions (1 µL) were spotted with 1 µL of alpha-cyano-4-hydroxycinnamic acid matrix, 3 mg/mL (1/1 v/v MeCN/0.1 % aqueous TFA).

Mass Spectrometry

PMF data acquisition and processing

- All MS spectra were acquired on a MALDI micro MX (Waters, Manchester, UK) mass spectrometer. Data were acquired in positive ion mode over the m/z range of 1000-3000 in reflectron mode.
- The instrument was operated in a fully automated manner, using the data to direct the settings. 100 laser shots were summed for each MS spectrum
- ProteinLynx Global SERVER 2.2 (Waters, Manchester, UK) was used for data processing and database searching.
- Processed 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.

PSD MX data acquisition and processing

- PSD MX spectra were acquired on a MALDI micro MX mass spectrometer (Waters, Manchester, UK). Data were acquired in positive ion mode over the m/z range100-3000 in PSD MX mode. The instrument was operated in a fully automated manner with laser settings and target plate positions determined automatically, using the data to direct the settings.
- ProteinLynx Global SERVER 2.2 (Waters, Manchester, UK) was used for data processing and database searching.
- Fragment ions were matched to their precursor ions using a deconvolution algorithm, implemented in MassLynx and ProteinLynx Global SERVER 2.2. A peaklist (pkl) was generated for each spot and automatically submitted to a databank search, against a Swiss-Prot database (v40).

Results

Databank search results

Databank search results obtained from the *E. coli* 2-D gel samples are summarised in Table 1. Two different types of databank search were performed on the MS or PSD data obtained from each sample well. MS data were submitted for a peptide mass fingerprint search whilst 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 types of score have a 95 % confidence limit, which is governed by the search parameters used, the limit for PMF and PSD MX scores is a score of 64. To facilitate comparing different types of databank search/score a relative score was devised. The relative score was obtained by using the following equation:-

$$\mathrm{Score}_{\mathrm{relative}} = \mathrm{Score}_{(\mathrm{indiv.\ Protein})} - \mathrm{Score}_{(95\%\ \mathrm{confidence})}$$

Therefore, the relative score can be calculated by subtracting the 95% confident protein identification level, (64 for both PMF and PSD MX) from the returned databank score for each individual protein. In the case of tenuous or ambiguous protein identification a negative relative score will be produced, whilst positive relative scores suggest statistically significant, or unambiguous protein identification.

Peptide Mass Fingerprinting

Using the traditional PMF approach 53 proteins were confidently identified from 96 sample wells (*cf.* Table 1). It was possible to match two proteins to the MS data from well F10 using PMF. These search results represent successful protein identification from 54 % of the samples. In our experience, this is a typical success rate for PMF protein identification.

Parallel PSD

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 analysed. Of the 81 proteins identified, 30 proteins were only unambiguously identified when PSD MX data were considered. In addition two proteins were found exclusively by using the PMF approach. A summary of the proteins identified by the different techniques used is illustrated graphically in Figure 1.

The benefit of using PSD MX over the PMF approach is clearly demonstrated in the case of sample D4. Analysis of this sample by PMF identified one protein, TYRA_ECOLI. However, using the additional information from the PSD MX experiment identification of a second protein, ALF_ECOLI could be made. Similarly for sample G12, DNAK_ECOLI was identified from the PSD MX data set and RS1_ECOLI was identified by PMF.

WELL	NAME	DESCRIPTION	MW	PMF		PSD MX	
				NO OF PEPTIDES	REL. SCORE	REL. SCORE	NO OF PEPTIDES PRODUCING SIGNIFICANT NUMBERS OF FRAGMENT IONS
A1	METE_ECO57	METHIONINE SYNTHASE	84538	10	34	34	-
C1	METE_ECO57	METHIONINE SYNTHASE	84538	7	13	20	1
D1	EFG_ECOLI	ELONGATION FACTOR G	77401	14	113	146	5
E1	PNP_ECOLI	POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE	77053	9	-25	30	2
F1	ILVD_ECOLI	DIHYDROXY ACID DEHYDRATASE	65359	7	13	78	1
H1	KPY1_ECOLI	PYRUVATE KINASE I	51039	3	-11	11	1
A2	DPPA_ECOLI	PERIPLASMIC DIPEPTIDE TRANSPORT PROTEIN PRECURSOR	60255	6	-24	18	1
B2	AHPF_ECOLI	ALKYL HYDROXIDE REDUCTASE SUBUNIT F	56484	9	6	50	1
C2	LEU1_ECOL57	2-ISOPROPYLMALATE SYNTHASE	57489	4	-7	27	3
D2	GLTD_ECOLI	GLUTAMATE SYNTHASE [NADPH] SMALL CHAIN	52478	-	-	21	-
F2	ASSY_ECOLI	ARGININOSUCCINATE SYNTHASE	49735	10	68	103	1
G2	TIG_ECOLI	TRIGGER FACTOR	48163	13	73	103	4
H2	GLYA_ECOLI	SERINE HYDROXYMETHYLTRANSFERASE	45288	9	-13	-4	1
A3	GLYA_ECOLI	SERINE HYDROXYMETHYLTRANSFERASE	45288	18	-8	45	2
B3	GLYA_ECOLI	SERINE HYDROXYMETHYLTRANSFERASE	45288	7	13	82	3
C3	DHE4_ECOLI	NADP-SPECIFIC GLUTAMATE DEHYDROGENASE	48778	11	45	97	2
E3	ENO_ECOLI	ENOLASE	45495	7	25	195	5
F3	EFTU_ECOLI	ELONGATION FACTOR TU	43155	5	30	59	3
G3	EFTU_ECOLI	ELONGATION FACTOR TU	43155	5	55	170	4
H3	G3P1_ECOLI	GLYCERALDEHYDE 3 PHOSPHATE DEHYDROGENASE A	35379	2	-11	171	5
A4	G3P1_ECOLI	GLYCERALDEHYDE 3 PHOSPHATE DEHYDROGENASE A	35379	3	-8	43	2
B4	AROG_ECOLI	PHOSPHO 2 DEHYDRO 3 DEOXYHEPTONATE ALDOLASE	37985	7	65	141	3
C4	CYSK_ECOLI	CYSTEINE SYNTHASE A	34337	5	46	202	3
D4	TYRA_ECOLI	T PROTEIN INCLUDES CHORISMATE MUTASE	42015	5	10	-8	-
	ALF_ECOLI	FRUCTOSE BISPHOSPHATE ALDOLASE CLASS II	38991	-	-	8	1
E4	ALF_ECOLI	FRUCTOSE BISPHOSPHATE ALDOLASE CLASS II	38991	4	11	94	2
F4	SERC_ECOLI	PHOSPHOSERINE AMINOTRANSFERASE	39626	8	43	43	-
G4	DHAS_ECOLI	ASPARTATE SEMIALDEHYDE DEHYDROGENASE	39992	5	11	77	2
H4	LIVJ_ECOLI	LEU ILE VAL BINDING PROTEIN PRECURSOR LIV BP	39052	5	4	15	-
A5	PGKL_ECOLI	PHOSPHOGLYCERATE KINASE	41133	7	26	90	3
B5	MDH_ECOLI	MALATE DEHYDROGENASE	32317	4	30	99	3
C5	OMPA_ECOLI	OUTER MEMBRANE PROTEIN A PRECURSOR	37177	4	25	225	4
D5	EFTS_ECOLI	ELONGATION FACTOR TS	30272	1	-27	37	1
E5	GLNA_ECOLI	GLUTAMINE SYNTHETASE	51739	6	-4	16	1
F5	SUCD_ECOLI	SUCCINYL COA SYNTHETASE ALPHA CHAIN	29627	7	60	164	4
G5	DAPD_ECOLI	THP SUCCINYLTRANSFERASE	29873	3	-18	61	4
H5	ZNUA_ECOLI	HIGH-AFFINITY ZINC UPTAKE SYSTEM PROTEIN ZNUA PRECURSOR	33870	-	-	31	1
A6	RS2_ECOLI	30S RIBOSOMAL PROTEIN S	26595	5	6	75	3
B6	SODM_ECOLI	SUPEROXIDE DISMUTASE MN	22951	4	27	149	2
C6	GPMA_ECOLI	2 3 BISPHOSPHOGLYCERATE DEPENDENT PHOSPHOGLYCERATE MUTASE	28407	9	94	95	1
D6	YODA_ECOLI	METAL-BINDING PROTEIN YODA PRECURSOR	24860	-	-	3	3
E6	GPMA_ECOLI	2 3 BISPHOSPHOGLYCERATE DEPENDENT PHOSPHOGLYCERATE MUTASE	28407	8	100	116	2
F6	YODA_ECOLI	METAL BINDING PROTEIN YODA PRECURSOR	24746	3	-11	54	3
G6	TRPA_ECO57	TRYPTOPHAN SYNTHASE ALPHA CHAIN	28878	10	0	38	2
H6	KAD_ECOLI	ADENYLATE KINASE	23571	6	4	-8	-
A7	ARTJ_ECOLI	ARGININE-BINDING PERIPLASMIC PROTEIN 2 PRECURSOR	26927	-	-	53	3
B7	HISJ_ECOLI	HISTIDINE-BINDING PERIPLASMIC PROTEIN PRECURSOR	28580	8	21	66	1
C7	FLIY_ECOLI	CYSTINE BINDING PERIPLASMIC PROTEIN PRECURSOR	29021	5	-5	97	2
D7	AHPC_ECOLI	ALKYL HYDROPEROXIDE REDUCTASE SUBUNIT C	20617	9	69	88	4
E7	IPYR_ECOLI	INORGANIC PYROPHOSPHATASE	19560	4	6	56	2
F7	TERE_SERMA	TELLURIUM RESISTANCE PROTEIN TERE	20435	-	-	65	3
G7	PTGA_ECOLI	PTS SYSTEM GLUCOSE SPECIFIC IIA COMPONENT	18108	2	-11	19	1
H7	TPX_ECOLI	THIOL PEROXIDASE	17864	-	-	21	2
A8	PPIB_ECOLI	PEPTIDYL-PROLYL CIS-TRANS ISOMERASE B	18256	-	-	-3	1
C8	CH10_ECOLI	10 KDA CHAPERONIN	10380	3	-5	37	2
E8	OSMC_ECOLI	OSMOTICALLY INDUCIBLE PROTEIN C	15062	-	-	29	1
F8	RL9_ECOLI	50S RIBOSOMAL PROTEIN L	15759	3	18	99	2
G8	GCH1_ECOLI	GTP CYCLOHYDROLASE I	24798	8	9	25	3
H8	GLNH_ECOLI	GLUTAMINE BINDING PERIPLASMIC PROTEIN PRECURSOR	27173	5	-9	44	2
A9	GLTI_ECOLI	GLUTAMATE/ASPARTATE PERIPLASMIC BINDING PROTEIN PRECURSOR	33513	12	14	14	-
B9	RS2_ECOLI	30S RIBOSOMAL PROTEIN S	26595	5	-14	-13	-
C9	SUCD_ECOLI	SUCCINYL-COA SYNTHETASE ALPHA CHAIN	29913	-	-	12	2
E9	CARA_ECOLI	CARBAMOYL-PHOSPHATE SYNTHASE SMALL CHAIN	41633	10	13	59	3
F9	G3P1_ECOLI	GLYCERALDEHYDE 3 PHOSPHATE DEHYDROGENASE A	35379	4	12	50	4
G9	DAPA_ECOLI	DIHYDRODIPICOLINATE SYNTHASE	31535	-	-	96	3
H9	OMPT_ECO57	PROTEASE VII PRECURSOR	35571	13	-8	46	-
A10	TRXB_ECOLI	THIOREDOXIN REDUCTASE	34470	8	38	68	2
B10	TALB_ECOLI	TRANSALDOLASE B	35066	5	13	79	1
C10	ATPA_ECOLI	ATP SYNTHASE ALPHA CHAIN	55187	6	13	70	2
D10	IMDH_ECOLI	INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE	52275	11	93	109	3
E10	EFTU_ECOLI	ELONGATION FACTOR TU	43155	9	-2	39	-
F10	PEPB_ECO57	PEPTIDASE B	46227	14	31	3	-
	EFTU_ECOLI	ELONGATION FACTOR TU	43155	10	11	15	1
G10	TRPB_ECO57	TRYPTOPHAN SYNTHASE BETA CHAIN	42838	11	57	57	-
H10	POTF_ECOLI	PUTRESCINE-BINDING PERIPLASMIC PROTEIN PRECURSOR	40928	-	-	6	2
A11	K6P1_ECOLI	6 PHOSPHOFRUCTOKINASE ISOZYME I	34819	6	-19	-	-
B11	YNCE_ECO57	HYPOTHETICAL PROTEIN YNCE PRECURSOR	38644	6	36	180	3
C11	METF_ECOLI	5 10 METHYLENETETRAHYDROFOLATE REDUCTASE	33081	4	7	28	1
D11	RHO_ECOLI	TRANSCRIPTION TERMINATION FACTOR RHO	47032	9	31	90	2
E11	ATOS_ECOLI	SENSOR PROTEIN ATOS	67747	5	-19	-17	-
F11	YFBU_ECOLI	PROTEIN YFBU	19638	7	22	34	-
B12	UPP_ECOLI	URACIL PHOSPHORIBOSYLTRANSFERASE	22518	3	-10	48	1
D12	UDP_ECOLI	URIDINE PHOSPHORYLASE	27182	5	2	19	-
E12	CH60_ECOLI	60 KDA CHAPERONIN	57161	14	104	224	5
F12	RS1_ECOLI	30S RIBOSOMAL PROTEIN S	61120	7	51	73	-
G12	RS1_ECOLI	30S RIBOSOMAL PROTEIN S	61120	8	38	152	4
H12	DNAK_ECOLI	CHAPERONE PROTEIN DNAK (HEAT SHOCK PROTEIN 70)	68998	-	-	42	-
	ATPB_ECOLI	ATP SYNTHASE BETA CHAIN	50220	10	59	113	2
*NO RESULTS WERE OBTAINED FROM B1, G1, E2, D3, B8, D8, D9, G11, A12 AND C12.							

Table 1: Table of the proteins identified by PMF and PSD MX using the MALDI micro MX. Green illustrates confident identification by PMF. Red illustrates confident identification from PSD MX.

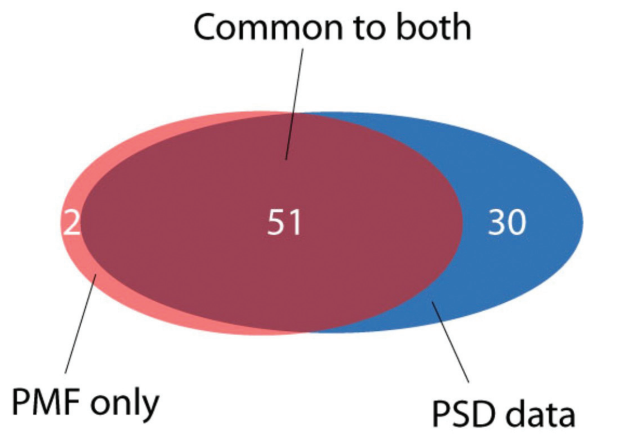

Figure 1: Diagram illustrating the number of proteins identified from the E. coli 2D gel samples by PMF and PSD MX data. The total number of samples analysed was 96.

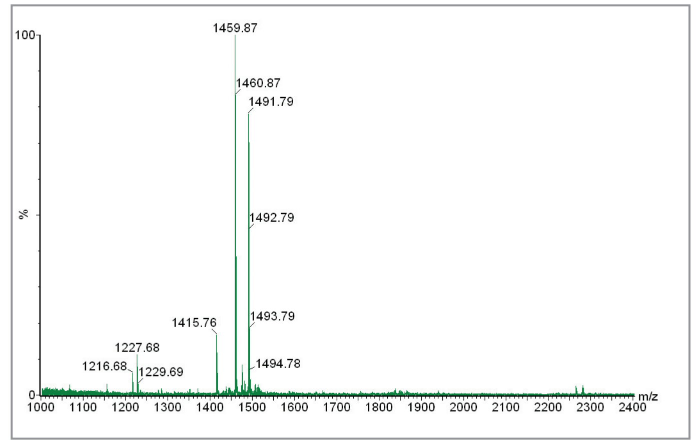

Figure 2: The MALDI MS data obtained from MALDI target position F6. Databank searching of this information, via PMF, did not result in identification of a protein sequence.

A typical example of PSD MX analysis resulting in successful protein identification

PMF analyses may fail due to an insufficient number of tryptic peptides produced by the enzymatic process or due to the poor ionisation of peptides in the mass spectrometer or a combination of these two factors. An example of this phenomenon is the mass spectrum shown in Figure 2, where only four intense peaks can be observed. Subsequent database searching of this information did not result in the unambiguous identification of a protein. However, a PSD MX experiment provided additional fragment ion information and, using this information in a databank search, it was possible to identify the protein as Metal-binding protein YodA from *E. coli*. This identification was based on PSD fragment ion data from three peptides. The search result is shown in Figure 3.

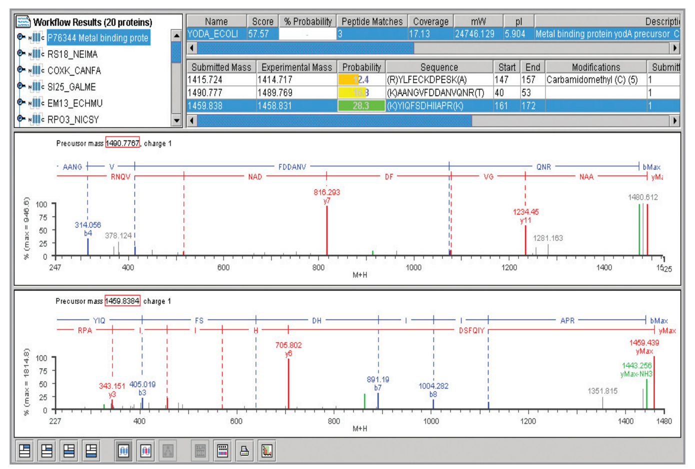

Figure 3: Result from the databank search of PSD MX data from MALDI target position F6. Shown inset are two matching PSD MX spectra annotated with the amino acid sequence from the databank.

Conclusions

- Peptide mass fingerprinting resulted in identification of 54% of the wells from 96 2D-gel separated E. coli samples
- PSD MX analysis of the same samples increased the hit rate from 54 % to 83 %, resulting in the identification of 83 proteins in total.
- Three of the 2D gel spots analysed were found to contain a mixture of two proteins, PSD MX was required to identify two of these mixtures.
- PMF and PSD MX were shown to be complementary techniques, which can be performed on the same mass spectrometer, MALDI micro MX.

[1] Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. C.; Gimley, C.; Watanabe, C. *Proceedings of the National Academy of Sciences*; 1993, 90, 5011-5015

[2] D. Kenny, J. Brown, M. Snel, *Technical note*: Multiplexed Post Source Decay (PSD MX) A novel technique explained