

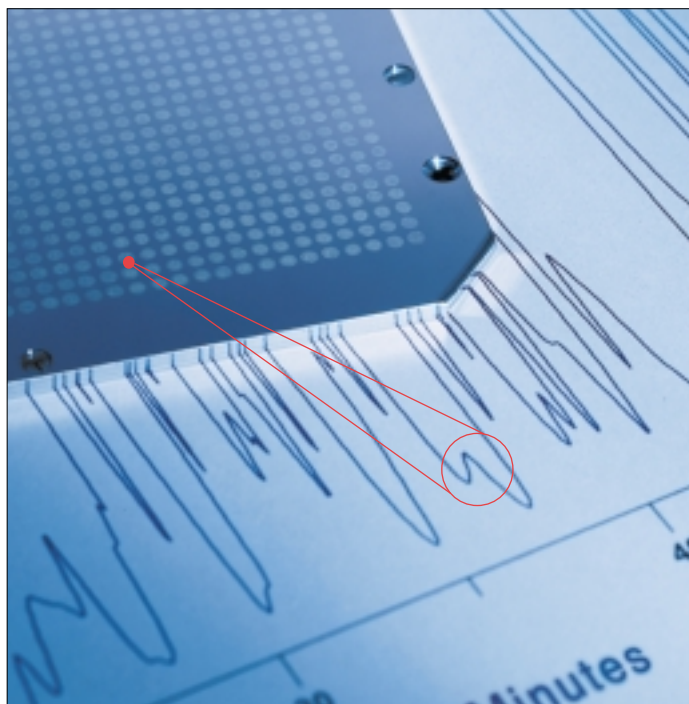
# The Power of LC MALDI: Identification of Proteins by LC MALDI MS/MS Using the Applied Biosystems 4700 Proteomics Analyzer with TOF/TOF™ Optics

## Purpose

The Applied Biosystems 4700 Proteomics Analyzer with GPS Explorer™ Software enables a fully automated and advanced LC MALDI MS/MS workflow for protein identification from complex mixtures. Novel MS peak detection identifies the presence of all peptides including those of closely related composition or isobaric peptides. Additionally, MS/MS spectra are collected at peak maxima increasing assay sensitivity and depth of sample coverage. This application note describes the analysis of a six-protein mixture using the LC MALDI workflow.

## Overview

Automated LC/MS/MS is a well-established technique for identification of components from complex mixtures. Although an electrospray ionization source is the interface of choice between the chromatograph and the mass spectrometer, there are significant advantages of interfacing chromatographic separations with a MALDI source. Using an LC MALDI approach, the rate of the collection of MS/MS data is decoupled from the chromatographic separation, allowing as much or as little time as necessary for acquiring spectra without fear of missing components as they elute from the HPLC column. In addition, more sophisticated selection of precursor ions ensures that all data is collected at the apex of the chromatographic response, thereby increasing



the number of components that can be identified from complex mixtures. Isobaric responses can also be automatically selected from ion chromatogram plots that are constructed *in silico* for every detected mass. This is often difficult to accomplish using electrospray ionization methods where such closely related components elute in a narrow chromatographic time window, and contemporary ion exclusion strategies are used to reduce data redundancy.

In this application note, features of the GPS Explorer Software and the 4700 Explorer™ Software developed to support the LC MALDI workflow are

demonstrated by the analysis of a tryptic digest of a six-protein mixture using the 4700 Proteomics Analyzer.

## Key Features

- Intelligent selection of precursor ions for MS/MS analysis ensures maximum depth of sample coverage.
- Use of *in silico* extracted ion chromatograms for each ion detected in the whole analysis enables identification of isobaric or closely related peptides that are components of complex mixtures in a single experiment.

- Collection of MS/MS spectra at chromatographic peak maxima improves the sensitivity of the automated LC MALDI MS/MS experiment.
- Sophisticated peak exclusion and adduct exclusion reduces collection of redundant data.
- Bi-directional communication between the 4700 Explorer Software and the GPS Explorer Software enables efficient peptide sequence elucidation and protein identification.
- Ability to archive plates containing chromatography and re-acquire data using modified data collection conditions increases confidence in peptide identifications without additional chromatography.

### Experimental Conditions

A trypsin digest of the six-protein mixture, comprised of beta-galactosidase, sero-transferrin, bovine serum albumin, alpha-lactalbumin, lysozyme and beta-lactoglobulin, was prepared for LC MALDI analysis using an LC Packings UltiMate™ HPLC system, Switchos™ column switching module, FAMOS™ autosampler and PROBOT™ microfraction collector. An aliquot (10  $\mu$ L) of the trypsin digest (representing 2  $\mu$ g of protein) was injected onto a peptide trap column from LC Packings using a 10 minute flush of the sample loop with solvent A (water:acetonitrile:TFA, 98:2:0.1 v/v/v). Following this injection procedure, the peptide trap column was switched into the solvent flow and peptides were separated using a 2-hour gradient from solvent A to solvent B (water:acetonitrile:isopropanol:TFA, 10:85:5:0.1 v/v/v/v, see Table 1). The mobile phase flow rate was 250 nL/minute.

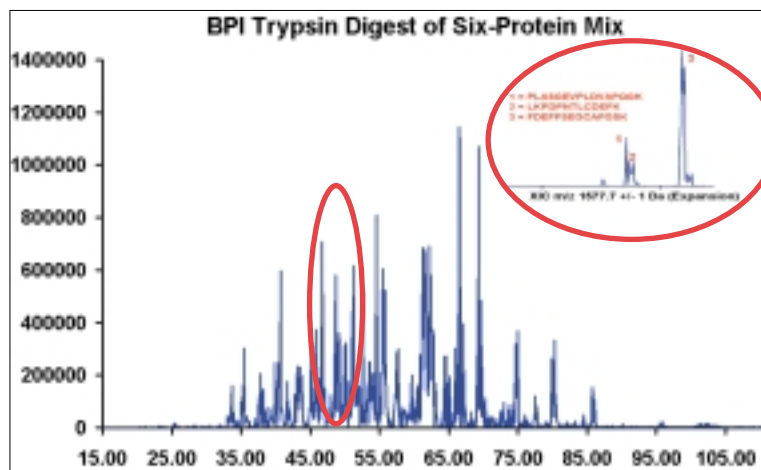


Figure 1. Base peak chromatogram for trypsin digested six-protein mix used in this experiment. The inset shows the extracted ion chromatogram of three peptides within a single Dalton of each other that were eluted within a four-minute chromatographic window.

TIME (min.)	% SOLVENT A	% SOLVENT B
0	95	5
120	65	35
122	5	95
125	5	95
127	95	5

Table 1. Timetable for Gradient Elution of Peptides

The column for the reversed-phase separation was a LC Packings PepMap C18, 75  $\mu$ m ID by 15 cm long. Pore size was 100 Å. PROBOT fractions were collected every 5 seconds using a 24-well by 24-well pattern across two MALDI plates. The eluent from the HPLC was mixed post column with matrix (7mg/mL alpha cyano-4-hydroxycinnamic acid in a solvent consisting of acetonitrile:water (75:25 v/v) before spotting onto the

MALDI plate. The matrix solution also contained 130  $\mu$ g/mL ammonium citrate to suppress matrix clusters, and the HPLC eluent was mixed with the matrix solution at a rate of one part HPLC eluent to two parts matrix solution using a low dead volume (25 nL) mixing tee. PROBOT fraction collection was started 10 minutes after commencement of the gradient.

The 4700 Proteomics Analyzer was used to analyze the samples. For MS data collection, 400 laser shots per spot were used. From these spectra, precursor ions with a signal to noise ratio of  $\geq 20$  were automatically selected for MS/MS using the Job Wide Peak Selection method that provides retention time, mass exclusion, exclusion of adduct responses

Protein	Peptides Identified	Protein Ion Score	Single Peptide Ion Score
beta-galactosidase	33	1500	118
sero-transferrin	22	1072	114
BSA	27	1187	106
beta-lactoglobulin	5	378	102
alpha-lactalbumin	3	177	102
lysozyme	3	89	63

Table 2. Proteins identified by an automated LC MALDI MS/MS experiment of a trypsin digest of a six-protein mixture. The MASCOT® database probability score reported that scores greater than 40 were significant ( $p < 0.05$ ).

**Table 3.**

Isobaric and closely related peptides identified in a single, automated LC MALDI MS/MS experiment of a trypsin digest of a six-protein mixture.

Precursor Ion (m/z)	Peptide Sequence	Retention Time (min.)	Protein
1064.6	TPHPALTEAK	31.17	beta-galactosidase
1065.6	VLVLDTDYK	48.00	beta-lactoglobulin
1283.6	EGYYGYTGAFR	50.08	sero-transferrin
1283.6	HPEYAVSVLLR	57.50	BSA
1428.7	DWENPGVTQLNR	46.75	beta-galactosidase
1428.7	FSENFNTQATNR	33.58	lysozyme
1478.7	MYLGYEYVTAIR	66.83	sero-transferrin
1478.7	ETYGDMADCCCK	29.50	BSA
1576.8	LKDPNTLCDEFK	47.50	BSA
1577.7	PLASGEVPLDVAPOGK	47.08	beta-galactosidase
1577.7	FDEFFSEGCAPGSK	51.67	sero-transferrin

(e.g. sodium or potassium etc.), and the ability to acquire less abundant peaks first. To reduce data redundancy, the selected precursor ions were excluded after analysis. MS/MS spectra were collected using 1,250 laser shots. The metastable suppressor was on with optimized precursor function and data was collected with the CID gas on. The laser attenuator was set to 4,200 for reflector MS data collection and 5,000 for collection of MS/MS spectra. A combined total of 2,419 MS/MS spectra were collected automatically across the two plates of the analysis (1,375 spectra from plate 1 and 1,034 spectra from plate 2). The data generated by this experiment were analyzed by GPS Explorer™ Software, which invoked a MASCOT® database search of the SwissProt database.

### Results and Discussion

Digestion of simple protein mixtures typically generates peptide samples of high complexity. This is exemplified in this study by the base peak plot extracted from LC/MS data collected for a tryptic digest of a six-protein mixture (Figure 1). Despite this complexity, all six proteins are confidently identified from a single

analysis in which MS/MS spectra were automatically collected using the Job Wide Peak Selection Method (Table 2).

Using the LC MALDI approach, sample coverage was aided by the ability to identify isobaric peptides (those that are closely related to within one or two Daltons). However, when traditional automated LC/ESI MS/MS algorithms are used, often the second eluting component of two closely related peptides is not selected for MS/MS analysis. This is due in part to the use of ion exclusion strategies to reduce data redundancy, and a reasonably wide exclusion mass window to reduce collection of MS/MS spectra of isotope peaks. In automated LC/ESI MS/MS experiments, there is always a trade-off between collection of redundant data (which also leads to a lower rate of data collection for other components in the mixture) and excluding ions of interest from MS/MS experiments (where the component elution time is unknown).

For proteomics target discovery experiments in which important components of the mixture are often

unknown, this is not optimal, and this highlights one of the many benefits of having the chromatography “frozen” on the MALDI plate. The time required for MS/MS analysis is decoupled from the speed of the chromatography. This permits a more sophisticated algorithm for the selection of precursors for MS/MS experiments. Precursor ions represented by all peaks detected in a single ion extracted chromatogram enables both the peak apex to be used for collection of MS/MS spectra and the selection of other peptides of the same or similar mass, enhancing the possibility of sequencing isobaric peptides. This is an exciting addition to the tool kit for the proteomics researcher since it helps tip the balance described above in favor of collection of more components of the mixture and improves the depth of coverage of the sample.

Examples of isobaric peptides and those closely related in mass that were identified from the automated LC MALDI MS/MS analysis of the trypsin digest of a six-protein mixture are shown in Table 3. Many of these responses were well resolved chromatographically, and MS/MS spectra

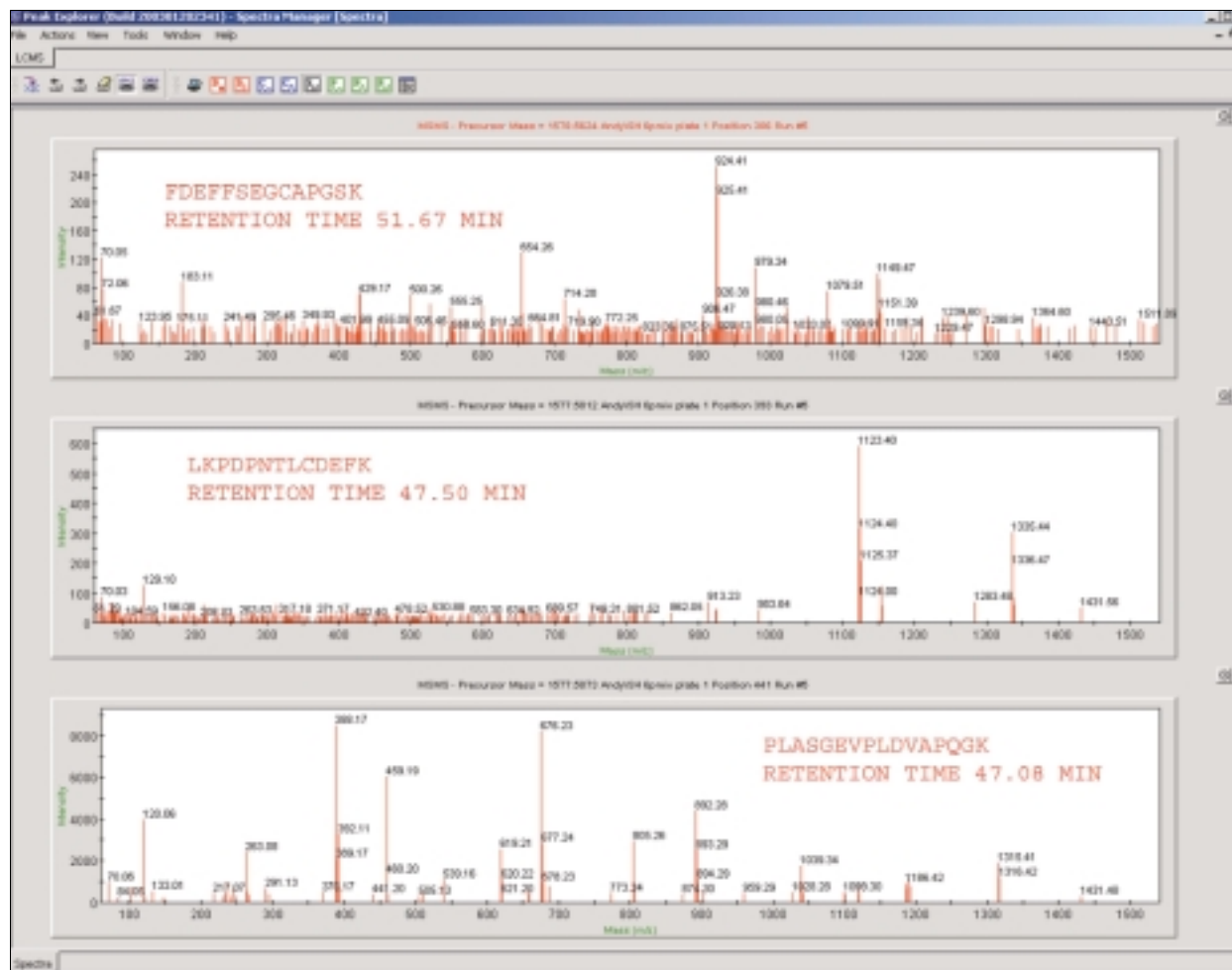


Figure 2. Similar mass and closely eluted trypsin fragments of BSA, beta-galactosidase, and sero-transferrin were automatically identified within a four-minute chromatographic time window. Each spectrum was acquired with 1,250 laser shots.

would be expected to be collected for each peptide in an automated analysis. However, the trypsin digest of this simple protein mixture was found to contain three peptides that were closely related in mass and eluted with similar retention times. These peptides were identified as trypsin fragments of BSA, beta-galactosidase and sero-transferrin. All three peptides were eluted in a 4-minute chromatographic window and were within a single Dalton of each other. In fact, two of the peptides eluted with retention times that were within 30 seconds of each other (inset, Figure 1). These peptides would have been difficult to analyze by contemporary automated LC/ESI MS/MS strategies as the second component would have

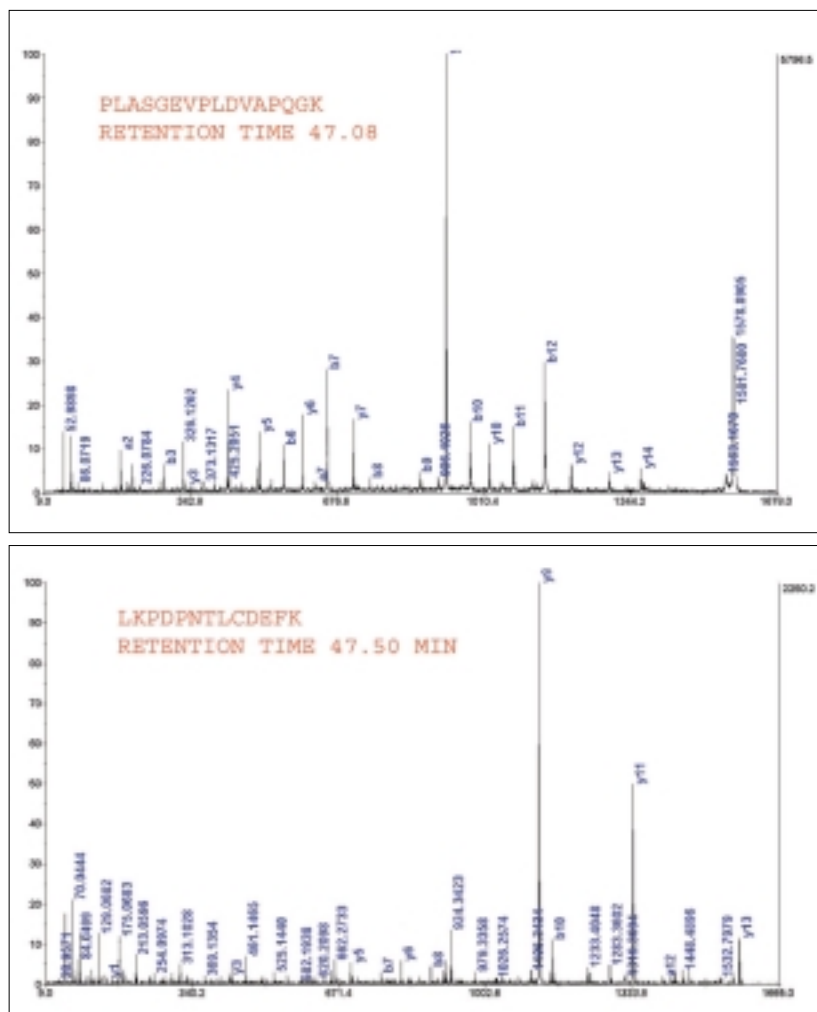
typically been excluded by both mass and retention time exclusion windows. The data collected automatically using the 4700 Proteomics Analyzer were of sufficient quality to confidently identify each peptide, and increase protein sequence coverage (Figure 2).

Another advantage of the LC MALDI approach is an ability to archive the sample plate and reacquire data after peptide identification. The ability to reacquire data is helpful in the confirmation of protein identification, or in selecting targeted components of interest for further analysis. In the current example, further MS/MS spectra were collected for the minor components with  $m/z$  values of 1577.7 and 1576.8 eluting

with retention times of 47.08 and 47.50 minutes respectively. In this experiment, 5,000 laser shots were collected for each spectrum. The MS/MS spectra collected were of enhanced spectral quality which increases the confidence in the identification of each peptide (Figure 3).

### Conclusions

The 4700 Explorer™ Software enables sophisticated collection of LC/MS/MS data from complex mixtures. Selection of the chromatographic peak apex ensures the most sensitive analysis of each component. Additionally, the use of such chromatographic data enables identification of closely eluting peptides that are isobaric or close in mass. This was exemplified by analysis of



#### **Worldwide Sales Offices**

Applied Biosystems vast distribution and service network, composed of highly trained support and applications personnel, reaches 150 countries on six continents. For international office locations, please call the division headquarters or refer to our Web site at [www.appliedbiosystems.com](http://www.appliedbiosystems.com).

Applera is committed to providing the world's leading technology and information for life scientists.

Applera Corporation consists of the Applied Biosystems and Celera Genomics businesses.

#### **Headquarters**

850 Lincoln Centre Drive  
Foster City, CA 94404 USA  
Phone: 650.638.5800  
Toll Free: 800.345.5224  
Fax: 650.638.5884

For Research Use Only.  
Not for use in diagnostic procedures.

AB (Design), Applera, Explorer, and TOF/TOF are trademarks and Applied Biosystems is a registered trademark of Applera Corporation or its subsidiaries in the US and/or certain other countries.

MASCOT is a registered trademark of Matrix Science Ltd.

FAMOS, PROBOT, Switchos, and UltiMate are trademarks of LC Packings.

©2003 Applied Biosystems. All rights reserved.

Printed in the USA. 5/2003, LD  
Publication 115AP18-01