Proteomics:
Biomarker Discovery and Validation
An Application Compendium

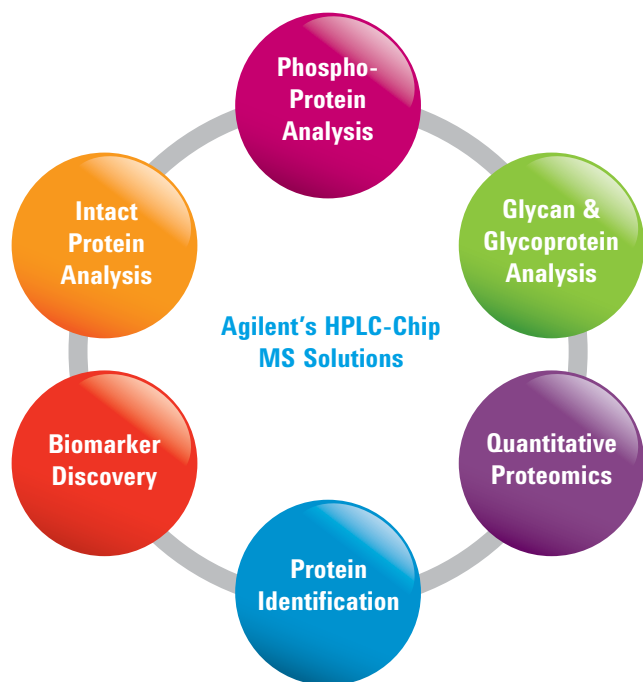**Agilent Technologies**

Table of Contents

Introduction

The field of proteomics has rapidly expanded to cover just about every aspect of protein research—from identification and characterization, to biomarker discovery and quantitation. Because protein analysis presents many distinct challenges, your proteomics research goals need to be met by complete, optimized, and accessible workflows for fast, accurate, and reproducible results.

Agilent's HPLC-Chip sits at the center of an integrated proteomics workflow that provides the highest analytical performance with unprecedented plug-and-play flexibility. Our robust, interchangeable workflows simplify setup and let you quickly switch between different methodologies to complete your research faster, with greater confidence.

These integrated proteomics workflows include Agilent's advanced LC/MS platforms. With the addition of innovative, proteomics-specific sample preparation options and high-productivity software packages, Agilent is uniquely able to deliver complete, end-to end solutions for all of your proteomics analysis needs.

Agilent 1200 Series HPLC-Chip II and 6000 Series LC/MS—clearly better together
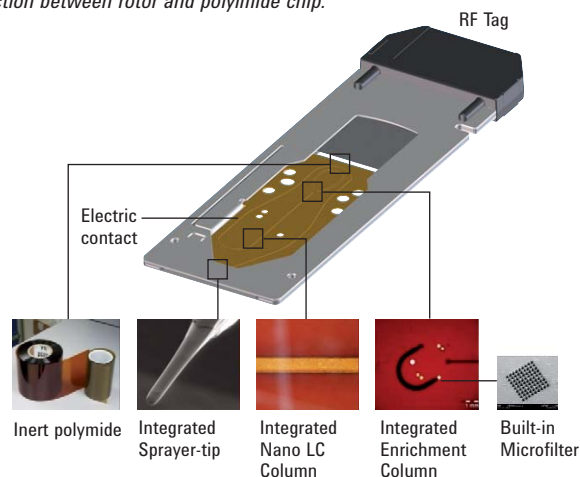
The Agilent 1200 Series HPLC-Chip/MS system is based on revolutionary microfluidic chip technology specifically designed for nanospray LC/MS. The system includes 1200 Series Capillary and Nanoflow Pumps, Micro-Well Plate Sampler with Thermostat, Chip Cube MS interface, and any Agilent 6000 Series mass spectrometer. System control is through either Agilent ChemStation or MassHunter software.

The second generation HPLC-Chip technology incorporates a carbon-ion-implanted filter, improving surface characteristics dramatically for optimal contact and sealing, as well as reducing friction between rotor and polyimide chip. These improvements double chip lifetime and lower the cost per analysis as well as improve chip-to-chip and run-to-run reproducibility. With Agilent's HPLC-Chip II, the Agilent HPLC-Chip/MS system takes you to new levels of nanospray MS reliability, robustness, sensitivity, and ease of use, allowing you to enter applications such as:

- Biomarker discovery and validation
- Intact monoclonal antibody characterization
- Small molecule analysis, such as DMPK
- Phosphopeptide analysis in post-translational modification (PTM) studies

HPLC-Chips are available for a wide range of application needs. Second-generation technology with new carbon-ion-implanted filter gives improved surface characteristics for optimal contact and sealing, as well as reduced friction between rotor and polyimide chip.

Agilent's HPLC-Chip MS Solutions

- Phospho-Protein Analysis
- Glycan & Glycoprotein Analysis
- Quantitative Proteomics
- Protein Identification
- Biomarker Discovery
- Intact Protein Analysis

RF Tag

Electric contact

Inert polymide Integrated Sprayer-tip Integrated Nano LC Column Integrated Enrichment Column Built-in Microfilter
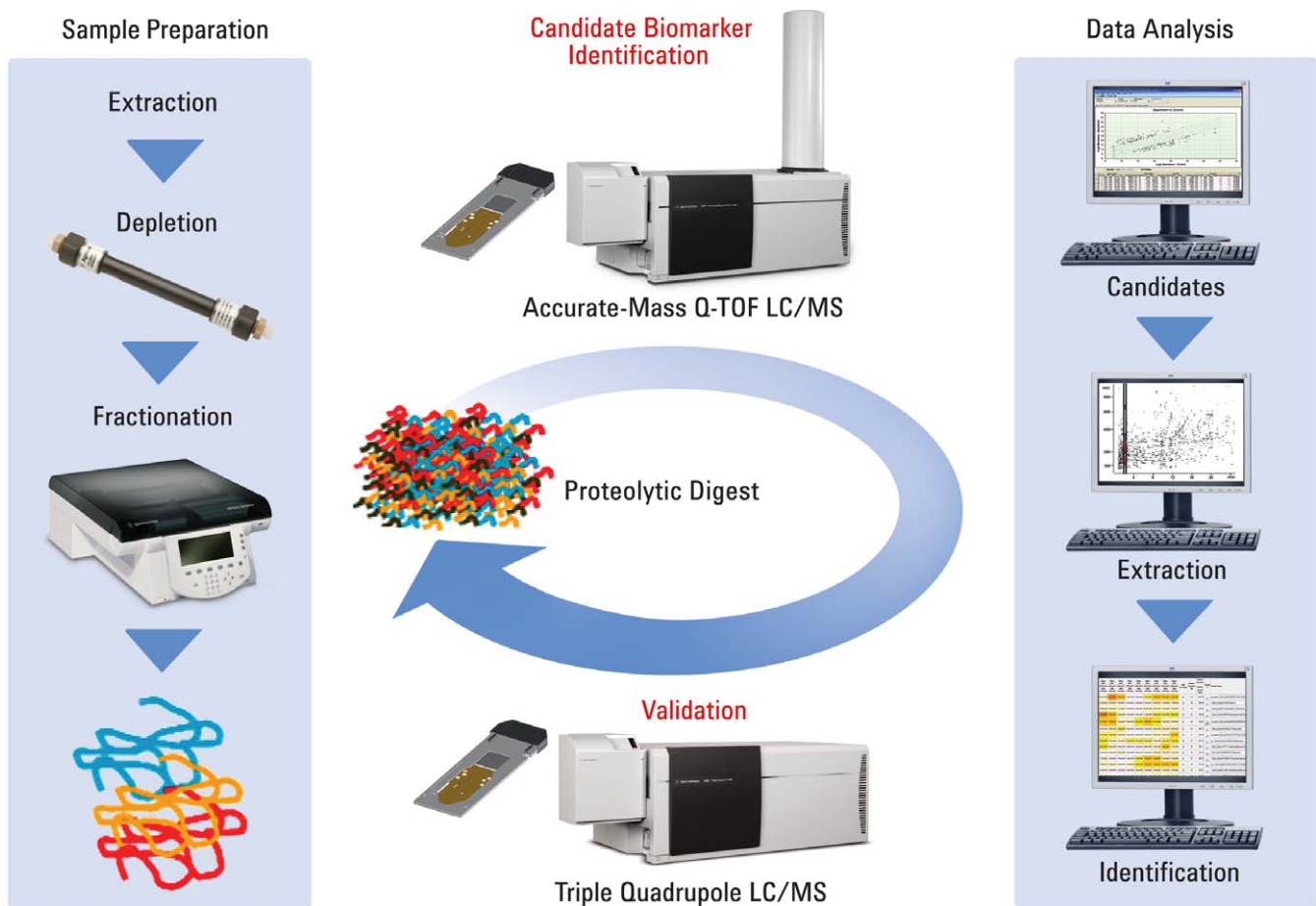
Protein biomarker discovery and validation workflow

Biomarker discovery and validation is the first step in the new paradigm of drug development involving biotechnology. LC/MS-based workflows play an important role in this field. Classical shotgun proteomics approaches relying on data-dependent acquisition are generally acknowledged to provide confident identification of only a subset of the actual proteins present in the sample.

Agilent's approach is to vastly reduce sample complexity using intelligent, two-step profiling before MS-MS. Differential protein expression analysis is employed in putative protein biomarker determination in clinical research or pharmaceutical development. Featuring HPLC-Chip/MS, Agilent offers a label-free identification and relative quantification workflow for rapid screening and identification (Q-TOF LC/MS) of potential biomarker and subsequent targeted validation (triple quadrupole LC/MS) of biomarker candidacy.

Reproducibility is key for biomarker discovery

- **MS abundance reproducibility** is necessary for ensuring that analytical variability remains well within the biological variability. The integrated spray design on the HPLC-Chip facilitates a stable nanospray and in combination with the 6500 Series Q-TOF design provides for excellent abundance reproducibility.

- **Retention time (RT) reproducibility** is also critical to ensure that molecular features align from replicate LC/MS runs. The HPLC-Chip delivers minimal retention time variation of the order of a few seconds over an hour-long gradient.

- **More accurate statistical analysis**, such as principal component analysis (PCA), can be achieved from data obtained from complex sample analysis—a direct benefit from the combined abundance and RT reproducibilities of HPLC-Chip/MS.
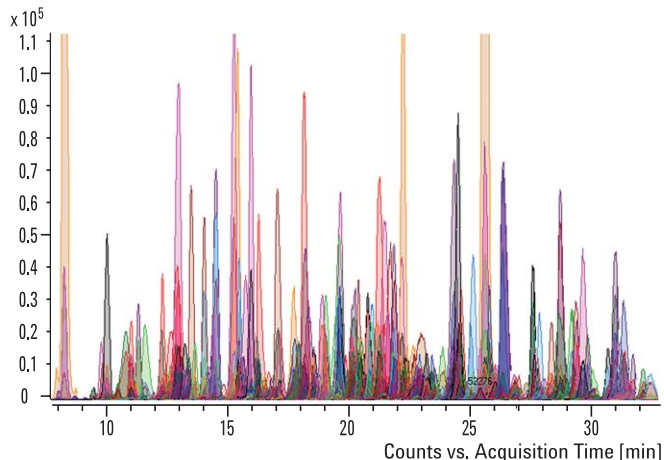

Protein biomarker discovery and validation workflow.

Intensity in run 10

Intensity: run 1 vs. run 10
Linear fit: $1.031988 \cdot x$
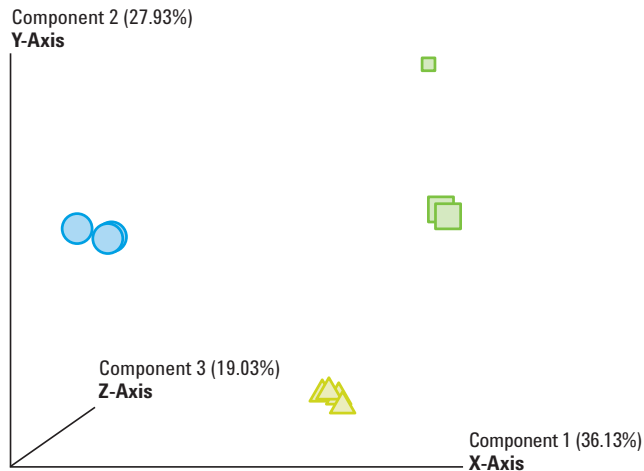Correlation: 0.998783
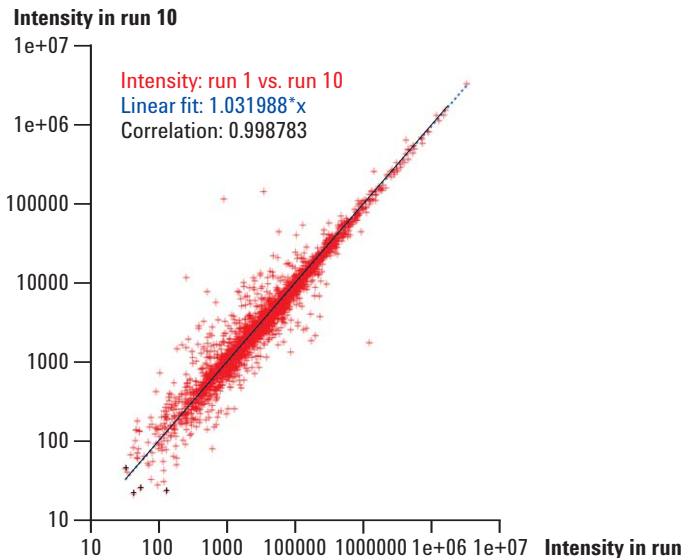
Scatter plot of ion intensities from two runs of 5000 glycopeptides. (Data courtesy of Prof. Rudi Aebersold, Institute of Molecular Systems Biology, Switzerland)

PCA analysis of complex digest samples reveal differentiated features among sample sets.

Confirmation and validation of putative protein biomarkers in complex biological samples requires an instrumental method that is fast, highly selective, and sensitive. Peptide quantification using multiple reaction monitoring (MRM) has been established as an important methodology for biomarker validation.

More than 2000 transitions can be acquired in a single dynamic MRM experiment using HPLC-Chip coupled to a triple quadrupole mass spectrometer.

Easy transition from discovery to validation

- The plug-and-play aspect of the Agilent HPLC-Chip allows the use of the same chip for discovery and validation, thus ensuring minimal retention time variation.
- The shared axial-acceleration collision cell between Agilent Q-TOF and Agilent Triple Quadrupole LC/MS systems permits the data from the discovery phase to be used to set up the dynamic MRM process on the triple quadrupole system for the validation phase.
- The high-quality retention time reproducibility allows the use of very narrow time windows within the dynamic MRM feature. This further enhances sensitivity by up to a factor of 10 over the classic MRM technique.

Profiling Approach for Biomarker Discovery using an Agilent HPLC-Chip Coupled with an Accurate-Mass Q-TOF LC/MS

Sudha Rajagopalan and Ravindra Gudihal, Agilent Technologies India pvt. Ltd., Bangalore, India; and Keith Waddell, Agilent Technologies, Inc., Santa Clara, California, U.S.A.

Abstract

A profile-directed biomarker discovery approach involves initial differential expression analysis followed by targeted identification of differentially expressed proteins. In the present study, a very small amount (100 femtomole and 10 femtomole) of horseradish peroxidase was spiked into human plasma and a profile-directed approach was used for the identification of the spiked protein. Excellent reproducibility in retention time and intensity measurement, and outstanding accuracy in mass measurement and sensitivity over a wide dynamic range, were achieved using various Agilent platforms such as a Multiple Affinity Removal (MARS) column, a 1200 Series HPLC-Chip/MS and a 6520 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) LC/MS. Differential analyses using MassHunter Mass Profiler and MassHunter Mass Profiler Professional (MPP) software enabled the comparison between spiked and non-spiked plasma samples. Subsequent targeted MS/MS analysis on a Q-TOF LC/MS followed by a database search using Spectrum Mill helped identify the spiked protein.

Introduction

Human plasma is extensively used for clinical diagnosis and early disease detection. Plasma is composed of classical blood proteins as well as tissue leakage proteins present in a wide dynamic range of concentrations. Highly sensitive, accurate and reproducible techniques are required to identify biomarkers in such complex biological samples. Mass spectrometry is used extensively in biomarker discovery and validation processes, especially in cancer biomarker studies.[1] Due to the volume and complexity of data generated in a biomarker study, sophisticated software tools are necessary for data analysis and visualization. This application note describes a profile-directed label free quantitation approach for the identification of a spiked plant protein in a human plasma sample. Excellent reproducibility in retention time and intensity measurements, along with accuracy in mass measurement and sensitivity over a wide dynamic range of concentrations offered by Agilent's 1200 Series HPLC-Chip/MS and 6520 Accurate-Mass Q-TOF LC/MS, have been demonstrated in the present work. Agilent also offers advanced software tools required for data analysis and statistical comparison.

Study design

Human plasma was depleted to remove 14 highly abundant proteins before digestion. A digest of a plant protein, horseradish peroxidase, was spiked at different concentration levels into the digested plasma and analyzed in four replicates using an Agilent 6520 Accurate-Mass Q-TOF LC/MS configured with a 1200 Series HPLC-Chip/MS. Replicate plasma digest samples without spiked peroxidase were also analyzed as controls. Molecular features extracted from peroxidase-spiked and non-spiked LC/MS runs were compared using Mass Profiler and MassHunter MPP. These features that were statistically differentiated in the spiked samples were identified using targeted MS/MS analysis on a Q-TOF LC/MS followed by database searching against the SwissProt database using Spectrum Mill. The experimental work flow and various Agilent platforms used in the study are shown in Figure 1.

Experimental

Sample preparation

Human plasma samples and peroxidase were obtained from Sigma. Fourteen of the highly abundant proteins in 125 µL of human plasma were depleted using Agilent's 10 x 100 mm MARS column as described.[2] Proteins that did not bind to the affinity column (eluted in the flow-through) were collected and buffer-exchanged with 100 mM ammonium bicarbonate using 5 kDa molecular weight cut-off spin concentrators.The protein concentration in the depleted plasma was measured using a DC protein assay kit (Bio-Rad) as described by the vendor. A 50 µg aliquot of the depleted plasma was reduced with dithiothreitol (DTT), alkylated using iodoacetamide and digested using 25:1 (v/v) trypsin (Agilent). A quantityof 250 picomole of peroxidase was similarly reduced, alkylated and digested. Equal amounts (2 µg) of digested plasma samples were spiked with varying amounts (0, 100 femtomole and 1 picomole) of peroxidase digest. A 200 ng aliquot of the plasma digest (spiked or non-spiked) was loaded onto the chip in each LC/MS analysis.
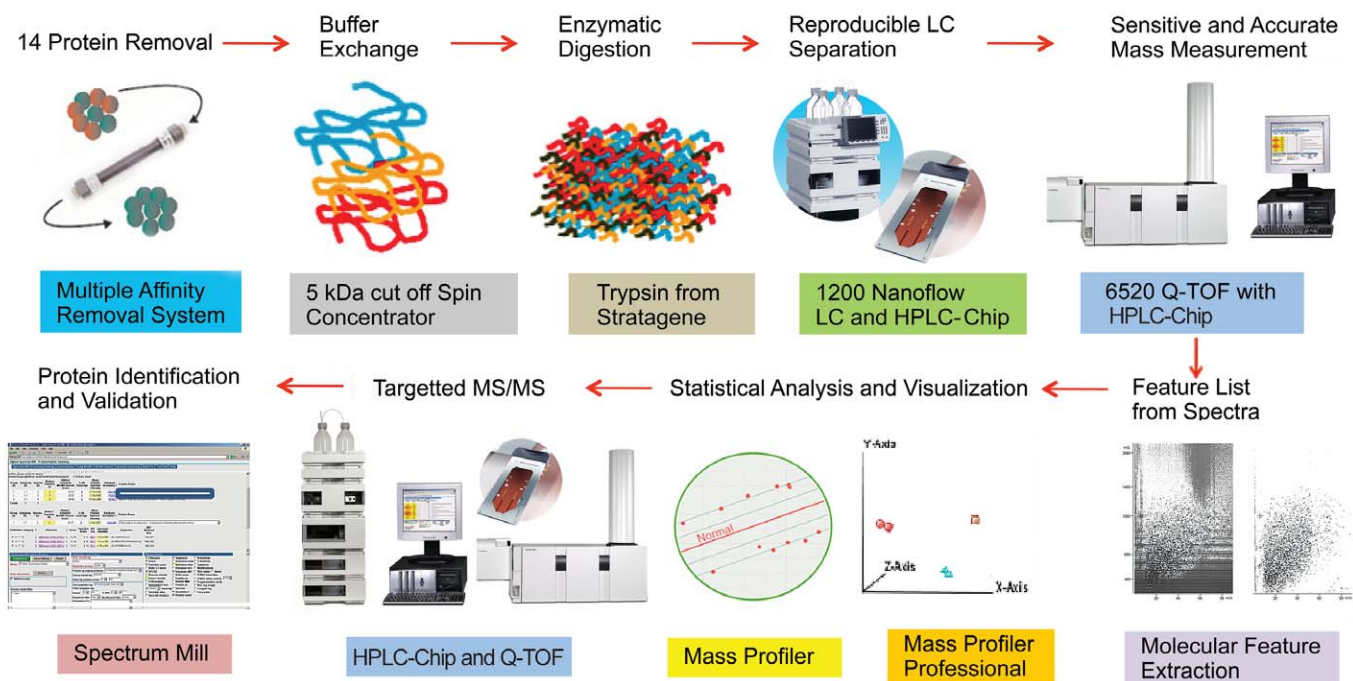
Figure 1. Workflow used for the biomarker model study.

LC/MS analysis

All LC/MS experiments were performed on an Agilent 1200 Series HPLC-Chip/MS interfaced to an Agilent 6520 Accurate-Mass Q-TOF LC/MS. The LC system consisted of a capillary pump for sample loading, a nanoflow pump and a thermostated microwell-plate autosampler. The HPLC-Chip configuration consisted of a 160 nL enrichment column and a 150 mm x 75 µm analytical column (Zorbax 300SB-C18). Mobile phases employed were: A). 0.1% formic acid in water and B). 90% acetonitrile with 0.1% formic acid. An 80 min long gradient method was used for the LC separation. Sample loading onto the enrichment column was done at 3% B. The gradient used for the analytical column began at 3% B, was raised to 45% B at 45 min, 70% B at 55 min, maintained at 70% B until 58 min, raised to 95% B at 68 min and then brought back to 3% B at 75 min. The column was equilibrated for 5 min before subsequent runs. Samples were loaded at 4 µL/min flow rate and eluted at 400 nL/min.

An Agilent 6520 Accurate-Mass Q-TOF LC/MS operating in high resolution (4 GHz) positive ion mode was used for all experiments. The MS source conditions were:

source temperature: 305°C

capillary voltage: 2,000 V

fragmentor voltage: 175 V

drying gas flow rate: 5 L/min

Data was acquired between m/z 300–3,000 at a scan rate of 1 spectra/sec for all samples in the profiling experiments. The number of LC/MS runs performed and samples used in the profiling experiment are described in Table 1. In the targeted MS/MS experiment, data was acquired from 300–3,000 m/z with an acquisition rate of 4 spectra/sec in MS mode and m/z region 50–3,000 with an acquisition rate of 5 spectra/sec in MS/MS mode. Complete system control was achieved using Agilent MassHunter data acquisition software (B.02.00).

Table 1. Details of LC/MS runs performed in the profiling experiment.

Sample No	Amount of peroxidase spiked on-column	Plasma digest loaded on chip in each LC/MS run	Number of replicates
1	0	200 ng	4
2	10 femtomole	200 ng	4
3	100 femtomole	200 ng	4

Data analysis

Data analysis was performed using the following software packages: Agilent MassHunter Qualitative Analysis (version B02.00), Mass Profiler (version B01.00), MPP (version B02.00), and Spectrum Mill(version B03.00). LC/MS data were extracted and evaluated using a specialized molecular feature extractor (MFE) algorithm in MassHunter Qualitative Analysis software. Mass Profiler and MPP were used for statistical evaluation of technical reproducibility and for comparison of spiked versus non-spiked (differential) samples. Spectrum Mill was used for protein identification from differential features.

Figure 2. Total ion chromatogram of four replicate LC/MS runs of plasma digest.

Figure 3. Mass versus retention time plot of the molecular features observed in LC/MS runs of plasma digest.

Results and Discussion

Total ion chromatograms of four replicate LC/MS analyses of sample 1 are shown in Figure 2 demonstrating the excellent reproducibility achieved in the LC/MS runs.

MFE takes raw data as the input and outputs a list of molecular features. A molecular feature represents a chemical entity such as a compound or a peptide. MFE reports a feature combining the abundance information of all its isotope clusters, different ion species such as multiple charge states, as well as dimers and adducts. Nearly 25,000 molecular features were observed in the LC/MS runs of plasma digest demonstrating the complexity of the sample (Figure 3). Retention times, masses and abundances of all features from each LC/MS run were compiled by MFE into feature lists.

Statistical analyses were performed on molecular features to check the reproducibility of the LC/MS analysis. A limited portion of the molecular feature list from four replicate LC/MS runs of plasma digest is shown in Table 2. Retention time deviation among the four replicate LC/MS runs is within ±4 sec (maximum being 0.06 min), which demonstrates the excellent retention time reproducibility achieved using Agilent's microfluidic-based HPLC-Chip. Precision of ±3 ppm is achieved in mass measurement. The relative standard deviation of abundance values is less than1% (maximum being 0.67%), which shows that the variation in measured intensities of the features between replicate runs is minimal. Excellent technical reproducibility achieved in the LC/MS analysis makes it suitable for comparison of samples with small biological differences.
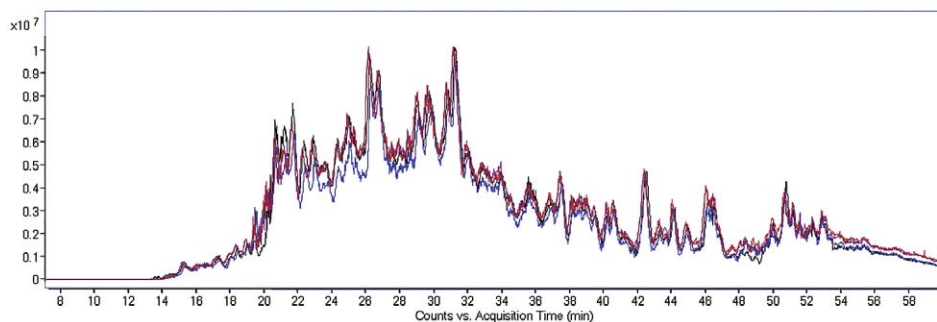
Table 2. Portion of the molecular feature list obtained from four replicate LC/MS runs of plasma digest.

ID	RT Value (min)	RT S.D. (min)	Mass Value (Da)	Mass S.D. (Da)	Abundance Value	Abundance R.S.D.	# of replicates in which feature is observed
1	42.566	0.048	3676.683	0.0091	36736530	0.15	4
2	46.208	0.051	2365.185	0.0053	32621650	0.44	4
3	46.671	0.043	2237.092	0.005	24123180	0.09	4
4	44.294	0.044	3520.587	0.0088	21518870	0.11	4
5	50.837	0.039	3065.391	0.0059	17312580	0.17	4
6	46.537	0.039	2612.318	0.0055	10599180	0.14	4
7	29.708	0.027	1565.76	0.0044	8590402	0.67	4
8	36.686	0.037	1418.698	0.004	8310790	0.08	4
9	47.022	0.047	4034.905	0.0088	8189842	0.15	4
10	37.091	0.060	3079.582	0.0082	7957212	0.1	4
11	39.577	0.057	1473.818	0.0043	7098632	0.05	4
12	36.914	0.045	2689.265	0.0073	5942096	0.08	4
13	42.693	0.037	3345.398	0.0086	5780965	0.13	4
14	50.078	0.034	2097.069	0.0038	5372915	0.18	4
15	35.485	0.019	1489.658	0.0043	4678271	0.07	4

MassHunter MPP for checking the technical reproducibility

In a biomarker discovery process, the technical variations (variations during sample preparation and analysis) should be small to ensure that differences between samples reflect true biological differences between them. In the present study, there should be clear differentiation between the replicates of plasma digest with equal amounts of peroxidase than between these with a different amount of peroxidase.

Agilent's MPP software offers advanced visualization and statistical tools for this differential analysis of multiple samples. Within MPP, various clustering algorithms and visualization tools are available to identify samples with similar mass abundance patterns. Principal component analysis (PCA) is a clustering tool often applied to reduce the dimensionality of complex data sets. When samples are plotted according to their correlation coefficient for the first three PCs on a 3D scatter plot, clustering of the sample can be seen as shown in Figure 4. The brown squares in Figure 4 represent the replicates of sample 1, while the blue and red dots represent replicates of samples 2 and 3, respectively. Clustering of replicates in a PCA scatter plot confirms that the replicates of each sample are more similar to one another than replicates of other samples.
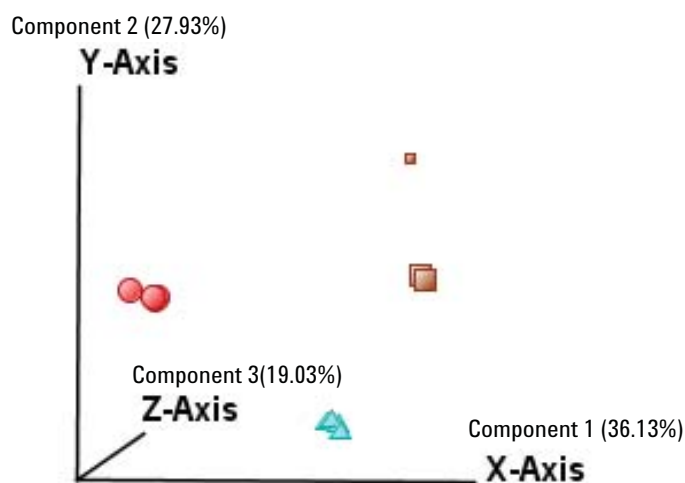

Figure 4. Principal Component Analysis: 3D Scatter Plot.

Differential analysis

The protein profiling approach in biomarker discovery comprises of two steps: rapid differential expression analysis of samples in MS-only mode followed by identification from targeted MS/MS data of differentiallyexpressed putative markers. Mass Profiler software enables retention time and m/z alignment for features across samples, intensity normalization, and t-test statistics for identification of significant differences between samples. Molecular features extracted from replicate LC/MS analyses of sample 1 were compared with features from replicates of sample 2 using Mass Profiler. Differential features were searched against the SwissProt database using peptide mass fingerprints of Spectrum Mill software. A screen shot of search results is shown in Figure 5. Peroxidase C1A precursor from horseradish is listed at the top among the identified proteins with a dynamic probability score of 0.0305.

A peptide mass fingerprint search against the SwissProt database using Spectrum Mill software on the differential features from the comparison of samples 1 and 3 also identified the protein peroxidase (data not shown). Six of the peroxidase peptides identified in the database search are shown in Table 3.

The concentration of spiked peroxidase is ten-fold higher in sample 3 as compared to sample 2 described in Table 1. Hence, abundance levels of peroxidase peptides are expected to be up regulated in sample 3. Features extracted from samples 2 and 3 were compared using Mass Profiler. Log_2 abundance of 10 femtomole spiked samples was plotted against log_2 abundance of 100 femtomole spiked samples (Figure 6). Filtering the results for an 8-fold difference in abundance identified many of the peroxidase peptides, some of which are marked in Figure 6.

Sample ID (comment): **10fmdif**
Database searched: **SwissProt**   Parameters used in Search
Molecular weight search (**10000 - 40000 Da**) selects **195305** entries.
Full pI range: **356194** entries.
Combined molecular weight and pI searches select **195305** entries.
PMF search selects **5** entries.

Rank	Dynamic Probability Score	Static Probability Score	# (%) Masses Matched	Mass Error Mean (Std Dev) (ppm)	Protein Coverage	Protein MW (Da)/pI	Species	Accession #	Protein Name
1	0.0305	4.21	4/26 (15%)	-29.5 (4.0)	9%	38825.3/5.67	ARMRU	P00433	Peroxidase C1A precursor - Armoracia rusticana (Horseradish) (Armoracia laphatifolia)
2	0.0376	5.15	4/26 (15%)	-29.5 (4.0)	9%	38847.3/6.21	ARATH	Q9LHB9	Peroxidase 32 precursor - Arabidopsis thaliana (Mouse-ear cress)
3	0.0528	0.659	4/26 (15%)	-37.8 (11.1)	41%	10846.7/9.52	ECOLI	Q47149	Uncharacterized protein yafQ - Escherichia coli (strain K12)
4	1.04	12.3	4/26 (15%)	-31.8 (10.4)	11%	38298.3/5.35	BACSU	P42319	Uncharacterized protein yxjH - Bacillus subtilis
5	1.29	4.21	4/26 (15%)	-23.5 (12.1)	17%	28259.2/5.65	HYPNA	Q0C1C1	30S ribosomal protein S2 - Hyphomonas neptunium (strain ATCC 15444)

Figure 5. Database search of differential features in the comparison of samples 1 and 2.

Table 3. Differential features obtained in the comparison of samples 1 and 3.

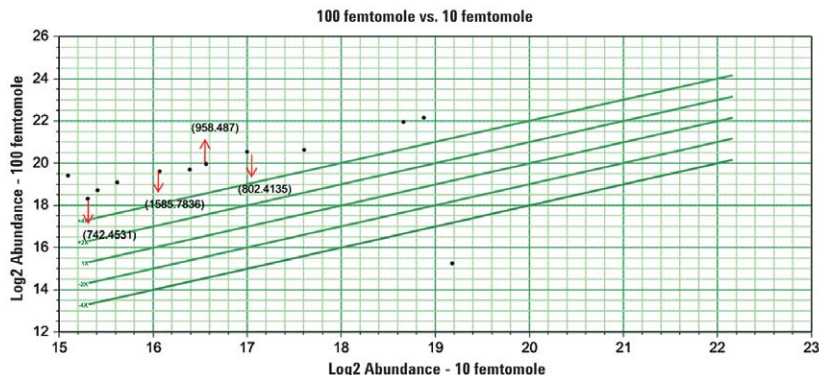| Feature No.	RT Value (min)	RT S.D. (min)	Mass Value (Da)	Mass S.D. (Da)	Abundance Value	Abundance R.S.D.	No. of runs observed
1	15.827	0.036	742.4529	0.0013	298859.6	1.64	4
2	18.849	0.025	802.4151	0.0016	816061	1.15	4
3	16.006	0.028	934.4484	0.0007	48349.07	1.51	4
4	18.275	0.056	958.4867	0.0006	539690.5	1.18	4
5	20.257	0.02	1474.707	0.003	211224.4	1.09	4
6	20.006	0.018	1585.787	0.0033	435146.3	1.19	4

Figure 6. Comparison of sample 2 and 3.

The mass spectrum of sample 3 between 18.09–18.24 min is shown in Figure 7. The doubly charged ion of a peroxidase peptide at m/z 480.2452 (mass value 958.49) in samples 2 and 3 are also shown in Figure 7. As shown in inset b, the intensity of the peak in a 10 femtomole peroxidase spiked sample is ten-fold less than the corresponding peak in a 100 femtomole peroxidase spiked sample (inset a).
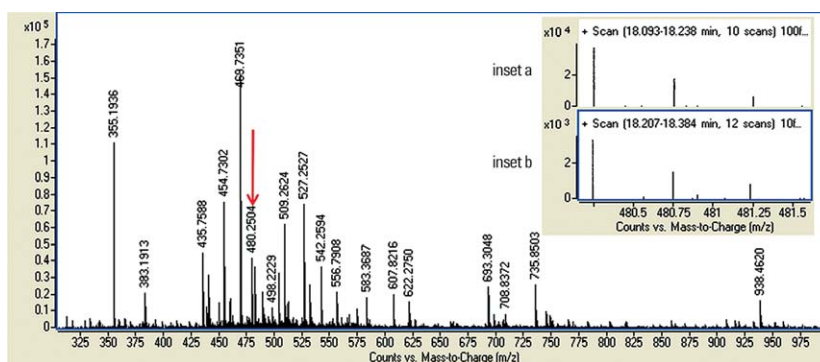

Figure 7. Mass spectrum of a 100 femtomole peroxidase spiked sample between 18.09 and 18.24 min. Inset shows the doubly charged ion at m/z 480.2 in the spectra of (a) 100 femtomole peroxidase spiked sample and (b) 10 femtomole spiked sample.

Targeted MS/MS analysis of sample 3 was performed using the differential features as an inclusion list for precursor selection. MS/MS data from the targeted analysis was searched against the Swiss Prot database using Spectrum Mill software. A screen shot of the search results is shown in Figure 8. The occurrence of peroxidase at the top of the list of identified proteins confirms the identification of perox-idase. Sequences of three of the identified peroxidase peptides are shown in this figure.

Figure 8. Database search of MS/MS data from targeted analysis.

The MS/MS spectrum of the perox-
idase peptide DTIVNELR (mass value
958.49) is shown in Figure 9. Even
though the peak at m/z 480.2504 is
not the most intense peak in the
spectrum of sample 3 (shown in
Figure 7), a strong MS/MS spectrum is
obtained in the targeted MS/MS
analysis. Nearly 20 ions are co-eluting
at the time of peak elution at m/z
480.25 as seen in Figure 7. As this peak
is the seventh in intensity, under classic
data dependent conditions (top six from a MS), this peak
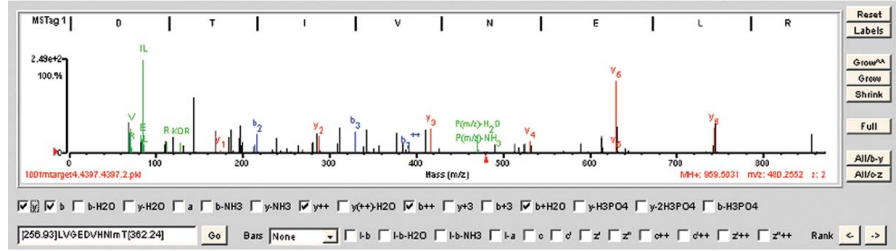may not get selected for MS/MS fragmentation.

Figure 9. MS/MS spectrum for the peptide DTIVNELR in targeted MS/MS analysis.

In the biomarker discovery process, it is necessary to be
able to identify small changes in the concentration of a few
proteins in the presence of a large number of proteins in
the sample. In this study, digest of a single plant protein
spiked at a low concentration of 10 femtomole in 200 ng of
human plasma digest could be identified, which demon-
strates the outstanding sensitivity achieved using Agilent's
6520 Accurate-Mass Q-TOF LC/MS configured with the
HPLC-Chip. Concentrations of the spiked peroxidase
peptides are comparable to the natural abundance levels of

tissue leakage proteins as shown in Figure 10. In the
figure, the normal range of abundance of classical plasma
proteins is marked in black, tissue leakage proteins are
blue and interleukins etc. are in purple boxes. 14 of the
high abundant classical plasma proteins (removed using
Agilent's MARS column in this study) are enclosed in a
green box. The concentration of a spiked peroxidase
peptide level is marked with a red dotted line showing that
it is in the tissue leakage region where biomarkers can be
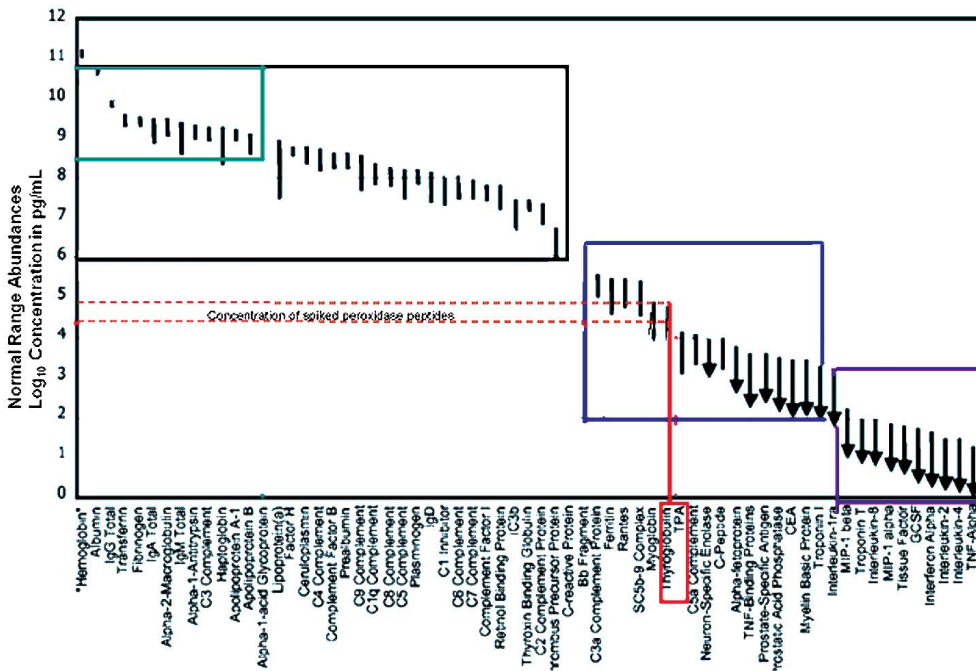expected to be present.

Figure 10. Plasma proteome map
showing the abundance range of some
known proteins used in clinical
diagnosis. (Figure adopted from
Anderson et al.)

Conclusions

The profiling approach described in this model study demonstrates outstanding sensitivity, exceptional accuracy and excellent reproducibility achieved using Agilent's 6520 Accurate-Mass Q-TOF LC/MS system coupled with Agilent's 1200 HPLC-Chip/MS system.

Removal of 14 highly abundant proteins using an Agilent MARS column enables in-depth analysis of low abundant proteins.

The MFE algorithm in the MassHunter Qualitative Analysis software helps detect low abundant peaks in the mass spectrum, increasing the sensitivity of detection.

A quality control check using MPP software is useful in evaluation of the technical variation and improves the reliability of differential analysis.

Differential analysis using Mass Profiler and MPP software enables the identification of a few differentially expressed features among a large number of similarly expressed features making it ideal for the biomarker discovery process.

Targeted MS/MS analysis on the differential features using Agilent's 6520 Accurate-Mass Q-TOF LC/MS demonstrates the ability to select the less abundant ions for fragmentation in the presence of highly abundant ions, which may not be possible in the conventional data dependent proteomic approach.

Low amounts (10 femtomole) of a protein spiked in a highly complex human plasma sample can be identified using the LC/MS based label free quantitation method.

References

1. Isaaq HJ, Conrads TP, Prieto DA, Tirumalai R, Veenstra TD, SELDI-TOF MS for diagnostic proteomics. *Anal Chem* 2003, 75, 148–155A.
2. Agilent publication 5989-7839EN: Human Serum and Plasma Protein Depletion – Novel High-Capacity Affinity Column for the Removal of the "Top 14" Abundant Proteins.
3. Anderson NL and Anderson NG, The human plasma proteome, *Molecular & Cellular Proteomics* 2002, 845–867.

The Agilent HPLC-Chip/6210 TOF LC/MS Enables Highly Accurate Profiling of Peptide Maps for Differential Expression Studies

Dr. Pierre Thibault, Institute for Research in Immunology and Cancer, University of Montreal, Canada

The Situation

Currently, no single LC/MS workflow has been adopted as the gold standard for biomarker discovery and validation. Shotgun proteomic techniques utilizing typical data-dependent MS-MS acquisition strategies provide meaningful identification on only a subset of the actual proteins present in a given sample.

The Solution

The Agilent HPLC-Chip/6210 TOF LC/MS is the perfect solution. Now Dr. Pierre Thibault and his team use a profiling strategy to confidently reveal putative biomarkers often missed with the shotgun approach, and subsequently target these putative biomarkers for identification.

Protein biomarker discovery has gained a lot of interest in drug development and early detection for the prevention and treatment of disease.

Liquid chromatography/mass spectrometry (LC/MS) based workflows, with their analytical power and potential throughput, play an important role in the discovery and validation of protein-based biomarkers. Current LC/MS workflows for biomarker discovery range from the classical shotgun proteomics approach to protein profiling strategies. However, no single LC/MS workflow has been adopted by the scientific community as the gold standard for biomarker research.

Shotgun proteomic approaches utilize data-dependent MS/MS acquisition, which identifies only a subset of the actual proteins present in the entire sample; these proteins usually represent the higher abundance proteins that are not necessarily of biological significance in disease. As a result, extensive fractionation of complex samples may be necessary to identify meaningful protein biomarkers associated with disease, thus dramatically increasing the number of analyses required per sample.

Dr. Pierre Thibault, a principal investigator at the Institute for Research in Immunology and Cancer (IRIC) and the director of the Proteomics Core Facility, stresses the difficulty of biomarker discovery. "When using a non-targeted shotgun proteomics approach, an overwhelming amount of MS/MS data can be acquired; out of that data, only a small subset will represent proteins of interest. This significant challenge [is like searching for] the proverbial differen-

tially expressed needle in a widely diverse proteomic haystack. Indeed, proteins of interest might only represent 5% of the overall population, and appropriate strategies are required to successfully identify these candidates."

Dr. Thibault's research program consists of developing a reproducible LC/MS proteomics platform for applications in cancer and immunology, and is heavily focused on characterizing low-level amounts of proteins and potential biomarkers in complex cell extracts.

Using the Agilent HPLC-Chip/6210 TOF LC/MS enables reproducible low-level analysis of multiple samples with particular focus on protein profiling for biomarker discovery (Figure 1). The Agilent HPLC-Chip system provides high resolution to increase the number of peptides found and proteins identified.The Agilent 6210 TOF LC/MS is an electrospray ionization (ESI) TOF mass spectrometer with both high sensitivity and highly accurate mass capabilities necessary for this profiling approach to reveal potential biomarkers. The powerful combination of the HPLCChip/ MS system with the 6210 TOF allows dependable reproducibility of sample analysis.

Dr. Thibault favors the Agilent HPLC-Chip/6210 TOF LC/MS for his research. "The novel microfluidics approach with the HPLC-Chip/TOF system eliminates all those uncertain issues with traditional nano-separation and nanoelectrospray. Dead volumes, dispersion issues, and sample losses are minimized because the pre-columns and columns are integrated into a chip format, allowing us to mine complex proteomics samples reliably and efficiently. This approach has also proven to be very reproducible and we have a more consistent platform day-in, day-out. We have lost the fear of wondering if we will get the same performance when we run our samples; there is now a level of reliability in our comparative sample studies."

One key study in Dr. Thibault's laboratory is the identification of protein expression changes in monocyte cells upon chemical stimulation with tumor promoting agents. "We used the TOF profiling strategy to first reveal those potential proteins that change in abundance and could be involved in early signaling events between control cells and PMA (phorbol 12-myristate 13- acetate)-treated monocytes. As in most cases, the majority of proteins do not change but we were able to find those that were both up- and
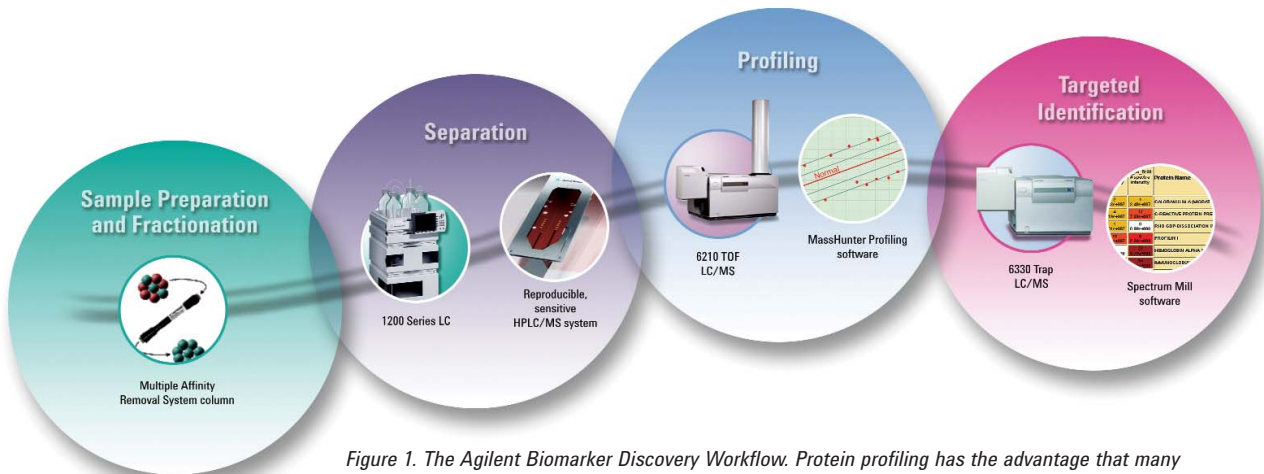

Figure 1. The Agilent Biomarker Discovery Workflow. Protein profiling has the advantage that many sample sets (e.g., control versus disease) can be screened in MS mode first to identify putative biomarkers that show statistically meaningful changes in expression across sample sets.

down-regulated (Figure 2). For this profiling approach, it was essential that we have highly accurate mass capability, within 5 ppm. The Agilent 6210 TOF LC/MS delivers this capability routinely without the need to invest in more expensive systems."
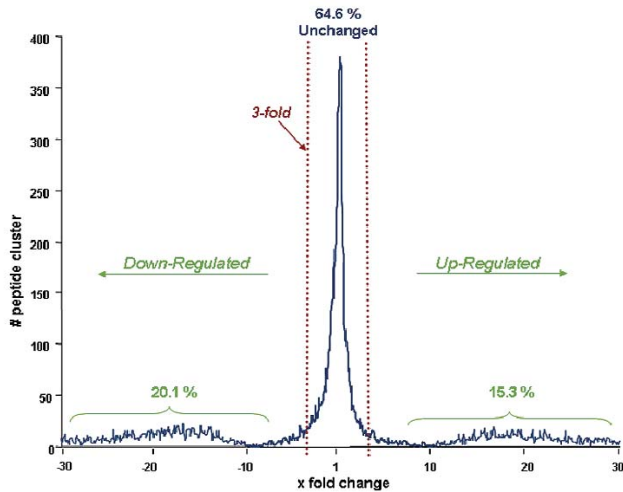

Figure 2. Expression profiling of peptide clusters for PMA-treated monocytes showing those that are up- and down-regulated.

"Once we have a list of peptides that are differentially expressed, we can target these for identification by MS/MS on our ion trap mass spectrometer (Figure 3). This profile-directed approach means we are not wasting any instrument time on identifying peptides that are not likely to be of significance in our study, and we can concentrate on those we're targeting with high sensitivity. The HPLC-Chip system is totally transferable between our ESI-TOF MS and our ion trap MS, so we can use essentially the same methods and HPLC separations for both our profiling and identification strategies."
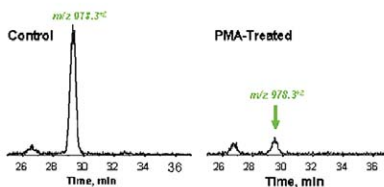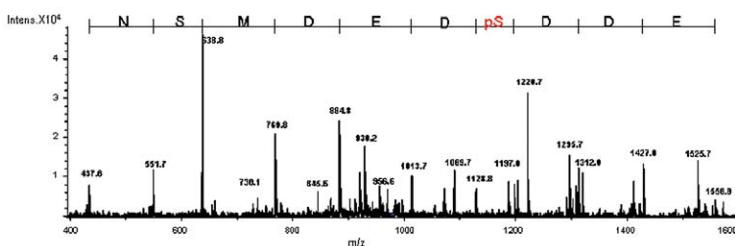
Biomarker discovery strategies essentially utilize the same strategy as described by Dr. Thibault, where control samples are compared with disease or drug-treated samples. With regard to Dr. Thibault's own biomarker studies, he says "The studies with these cultured cells exhibit a lower level of biological variation. With our other samples from sources such as tissues or blood, there is much more biological variation which means that the number of replicates must increase in order to statistically determine which proteins are truly differentially expressed. When analyzing a large number of samples, reproducibility is the key, from the HPLC separation to mass spectrometry analysis."

"By using the HPLC-Chip system with ESI-TOF mass spectrometry, we have a system that routinely and reproducibly analyzes low-level proteins in complex mixtures. This enables us to profile many samples for differential expression analysis and then target differentially-expressed peptides for further identification. This profile-directed approach allows us to identify those lowlevel proteins in the 5% subset of the population often missed with regular shotgun proteomics techniques."

Dr. Pierre Thibault, Principal Investigator

In reviewing the overall benefits of the Agilent HPLC-Chip system with ESI-TOF mass spectrometry, Dr. Thibault emphasizes, "By using this novel chip-based LC/MS we can routinely run more samples for our comparative studies with confidence in reproducibility. This system enables us to concentrate our efforts on our experimental design, sample handling, and results—rather than on the LC/MS system—to further our research."

Figure 3.
a. Profiling by ESI-TOF reveals peptides that are up- and down-regulated
b. Targeted MS/MS of these peptides leads to their identification

About The Institute for Research in Immunology and Cancer

The Institute for Research in Immunology and Cancer is focused in advancing health research in the workings of the immune system and cancer. It is also dedicated to training investigators in the development as well as generation of therapies to stop the suffering and premature death caused by cancer. www.iric.ca

Related publications

1. M. Ghitun, E. Bonneil, M.-H. Fortier, H. Yin, K. Killeen, P. Thibault, "Integrated Microfluidic Devices with Enhanced Separation Performance; Application to Phosphoproteome Analyses of Differentiated Cell Model Systems", *J. Sep. Sc.*, in press.

2. M.-H. Fortier, E. Bonneil, P. Goodley, P. Thibault, "An integrated microfluidic device for mass spectrometry-based proteomics and its application to biomarker discovery programs", *Anal. Chem.*, 77, 1631–1640 (2005).

The Effect of MS/MS Fragment Ion Mass Accuracy on Peptide Identification in Shotgun Proteomics

David M. Horn, Christine A. Miller, and Bryan D. Miller, Agilent Technologies, Inc., Santa Clara, California, U.S.A.

Abstract

It is widely accepted that high precursor ion mass measurement accuracy significantly improves confidence in MS/MS database search results throughthe reduction of false positives. In contrast, the impact of high fragment ion mass accuracy has been largely ignored. The Agilent Accurate Mass Q-TOF, a recent advancement in MS instrumentation, now enables high-throughput accurate mass measurement of both precursor and fragment ion data. This study examines the effects of precursor and fragment ion mass accuracy on protein identification using the Accurate Mass Q-TOF. Results demonstrate that fragment ion mass accuracy is as important (if not more important) than precursor ion mass accuracy in reducing false positive identification rates in MS/MS database searches, and plays a critical role in improving the throughput of proteomics research projects.

Introduction

Perhaps the most fundamental goal of proteomics research is the identification and characterization of proteins, particularly in the context of cellular expression. The shotgun technology of tandem mass spectrometry paired with advanced liquid chromatography has emerged as the standard technique for high-throughput protein identification (Liu, J. *et al.*, 2007; VerBerkmoes, N.C. *et al.*, 2002). To perform a typical shotgun experiment, a sample is first fractionated and the resulting mixture of proteins is then digested into peptides by an enzyme such as trypsin. The peptide mixture is then separated by high performance liquid chromatography (HPLC) and subsequently analyzed by mass spectrometry to determine the mass/charge (m/z) ratio of each peptide. Peptides of interest are selected for further fragmentation in a collision cell to produce tandem (MS/MS) mass spectra. This resultant data is then searched against a protein database to identify the peptide sequences and further infer the protein content of the sample.

It is well known that high mass accuracy measurement of precursor ions significantly improves confidence in MS/MS protein database search results. However, the impact of high mass accuracy measurement of fragment ions has been largely ignored. Up until recently, MS instruments that possessed sufficient throughput capability to capture both precursor and fragment ion data with high mass accuracy were not readily available. The advent of the Agilent Accurate-Mass Q-TOF LC/MS provides the sensitivity, speed, and accurate mass capabilities to enable the assessment of precursor and fragment ion mass tolerance as the maximum mass deviation, or MMD (Zubarev, R. and Mann, M., 2007). The Accurate-Mass Q-TOF LC/MS quickly captures information from complex samples with attomole-level sensitivity, a wide in-scan dynamic range that covers 3.5 orders of magnitude, and a fast acquisition rate (20 spectra/sec) at up to 15,000 resolving power.

Here, we apply the Q-TOF LC/MS system and demonstrate that fragment ion mass accuracy can be at least as significant as precursor ion mass accuracy in reducing false positive identification rates in MS/MS protein database searches. In addition, we compare false positive rates generated by the mass tolerance parameters particular to other instrument types used for protein identification research.

Figure 1. The experimental workflow consisting of a 3100 OFFGEL Fractionator, a 1200 HPLC-Chip/6510 Q-TOF LC/MS, and Spectrum Mill software for data analysis. As part of this study, the performance of the Q-TOF was compared with the 6330 Ion Trap mass spectrometer.

Materials and Methods

The experimental workflow for this analysis used a 3100 OFFGEL Fractionator to fractionate the HeLa sample, a 1200 HPLC-Chip/6510 Q-TOF LC/MS to acquire the data, and Spectrum Mill Protein Identification software to perform analysis of the collected data and assess the MMD for HeLa cells (Figure 1).

HeLa Cell Culture

HeLa S3 cells were grown in Ham's F12 medium (F12K) with 2 mM L-glutamine and 1.5 g/L sodium bicarbonate supplemented with 10% fetal bovine serum until 90% confluency. Cells were washed in PBS, lysed in hypotonic solution containing protease inhibitors cocktail, then homogenized with a Dounce homogenizer. Soluble HeLa lysate was prepared by centrifugation (1 hr at 16,000 rpm at 4°C). An aliquot (~300 µg total protein) of the supernatant was digested with trypsin using a 2,2,2-trifluoroethanol (TFE) based protocol for solubilization and denaturation. The digest was dried to remove the TFE and ammonium bicarbonate.

OFFGEL Fractionation

The trypsinized lysate was fractionated into 23 fractions in the pH range from 3 to 10. The current was limited to 50 µA and fractionation was stopped after 50 kVh (about 24 hr).

LC/MS Analysis

Approximately 5% of each fraction (5 µL) was analyzed using a microfluidics based HPLC-Chip connected to an Agilent 6510 Quadrupole Time-of-Flight (Q-TOF) mass spectrometer or an Agilent 6330 Ion Trap mass spectrometer.

HPLC-Chip: Protein ID chip with 150 x 0.075 mm analytical column and 160 nL enrichment column. Sample load: 5 µL of fractions obtained from OFFGEL. Flow: 300 nL/min analytical pump, 4 µL/min loading pump. Mobile phases A: 0.1% formic acid (FA), B: 90% acetonitrile, 0.1% FA. Gradients: 3%B to 6%B at 2 min, then 10% B at 10 min, 30%B at 65 min, 40%B at 75 min, 80%B at 80 min until 85 min, then 3%B at 85.1 min. Stop time: 90 min and post time 10 min.

Ion Trap MS Conditions: Drying gas: 4 L/min, 300°C; Skim 1: 30 V; Capillary exit: 75 V; Trap Drive: 85; Averages: 1; ICC: On; Max. Accumulation time: 50 ms; Smart Target: 500,000; MS Scan range: 300-1800; Ultra Scan; MS/MS Scan range: 100-2000; Ultra Scan.

Q-TOF MS Conditions: Drying gas: 4 L/min, 300°C; Skimmer: 65 V; Fragmentor: 175 V; Collision energy: slope 3.7 V, offset 2.5 V; MS Scan range and rate: 300-2000 at 3 Hz; MS/MS scan range and rate: 50-3000 at 3Hz. AutoMS/MS: 8 precursors, active exclusion on with 1 repeat and release after 0.17 min. Preferred charge state: 2, 3, >3, unknown.

Database Searches

Protein database searches were performed with Spectrum Mill Protein Identification software. All searches used the IPI Human Version 3.28 database with trypsin specificity, 2 missed cleavages, 50% minimum scored peak intensity, and dynamic peak thresholding. The "forward" database search of the ion trap data was performed using 2.5 Da precursor and 0.7 Da fragment ion tolerance while the Q-TOF search used 10 ppm precursor and 40 ppm fragment ion tolerance. Protein identifications were validated manually. A reversed version of the IPI Human database was created using a Perl script. Searches against this reversed database were performed using a variety of different precursor and product ion mass tolerances. The false positive rates for a given Spectrum Mill score and MMD regime were estimated by comparison of the distribution of Spectrum Mill scores for the forward Q-TOF search and reversed database search results. All results were exported to Microsoft Excel and database search scores were divided into 0.5 unit bins and plotted.

Results and Discussion

Selection of MMD values for Q-TOF data

As a first step in this study, it was necessary to determine the optimal mass tolerances of the 6510 Q-TOF to use for the subsequent protein database search. Using various precursor and product ion mass tolerances, a series of forward database searches were performed. The search results that produced the largest number of validated protein identifications used 10 ppm precursor ion and 40 ppm fragment ion tolerance.

Figure 2 shows the distribution of precursor and fragment ion mass errors for the validated peptides. These results indicate that fragment ion mass tolerance displays a wider range than that of the precursor due to significantly reduced signal-tonoise ratios. However, the higher abundance fragments demonstrate similarly high mass accuracy and improvements to the MS/MS search algorithm could take this into account to improve the results.
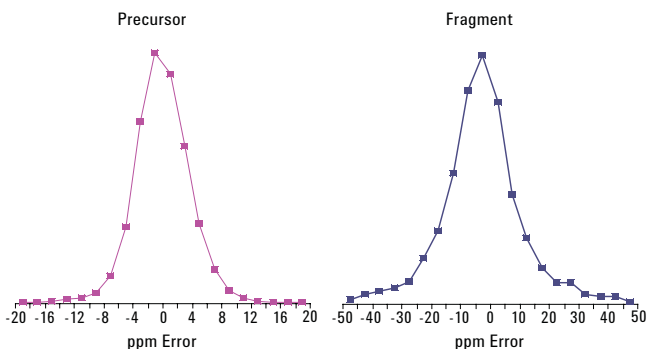

Figure 2. Mass measurement deviation distributions for precursor (3.9 rms ppm) and fragment ions (13.8 rms ppm) using the 6510 Q-TOF.

Comparison of ion trap and Q-TOF results

Having established the optimal mass tolerance values for the Q-TOF, we next compared the performance of the Q-TOF with the 6330 Ion Trap mass spectrometer, an instrument that is commonly used for protein identification research. Both instruments analyzed aliquots of the same HeLa sample. Table 1 displays the results generated from both instruments. In the ion trap, there were a number of very high scoring random protein hits. Further, the number of spectra with "random" peptide identifications was more than 10-fold for the ion trap compared to the Q-TOF system. This is presumably partly due to the higher precursor MMD of the ion trap, but also potentially due to the effect of high fragment ion tolerance as well. Also, the total number of MS/MS acquired on the Q-TOF was lower by 90,000 spectra, largely because the high resolution of the Q-TOF better distinguished real peptidesignal from background noise. Finally, a greater number of validated peptide identifications were found in the Q-TOF data.

Table 1. Comparison between 6330 Ion Trap and 6510 Q-TOF search results and hits.

	Ion Trap	Q-TOF
Total number spectra acquired	355,136	363,692
MS/MS acquired	271,507	186,049
Number of extracted files (.pkl)	92,081	129,752
Number of search result files (.spo)	72,028	20,871
Number of validated spectra	10,618	16,072
Number of validated unique peptides	4151	6253
Number of protein IDs	873	994
Number of unvalidated spectra	61,410	4799

This combination of factors led to a significantly higher identification yield for the Q-TOF compared to the ion trap. Besides the higher mass accuracy, other factors that might have contributed to the high number of valid spectra generated by the Q-TOF include the ability to assign charge state for the fragment ions and the general cleanliness of the Q-TOF MS/MS data, especially when compared to the ion trap data for MS/MS acquired and the number of extracted files. Regardless, the reduction in matches to random proteins exhibited by the data obtained from the Q-TOF enables a vast improvement in confidence related to peptide identification.

18

Precursor and fragment ion mass tolerance

Based on the results of these preliminary studies indicating the significance of mass accuracy on the ability to identify peptides, we next proceeded to determine the contributing effects of precursor and fragment ion mass tolerances on confident protein identification. Past research has indicated that high stringency in the precursor ion tolerance can dramatically reduce the rate of false positive identifications for a given search (Liu, T. *et al.*, 2007). To demonstrate this effect, a series of reversed database searches were performed holding the fragment ion mass tolerance constant at 40 ppm (the established threshold for the Q-TOF data obtained in thisstudy) while varying the precursor ion tolerance. It is known that reversed database searches are a good method for deter-mining the false positive rate, because searching a reversed database would generate a number of matches that would be known to be false (Elias, J.E. and Gygi, S.P., 2007). As precursor ion tolerance decreased below the fragment ion tolerance of 40 ppm, the false positive rate decreased as expected (Figure 3). Interestingly, as the precursor ion tolerance was increased above the fragment ion tolerance, there was little change from 50 to 200 ppm and only a

slight increase with 500 and 1000 ppm. The increase in the number of false positives for 500 and 1000 ppm may be attributed to matches of peptides with 1 Da higher and lower nominal mass. These results suggested that while precursor ion mass tolerance most certainly does affect the false positive rate in database searching, the effect of a low fragment ion mass tolerance is also an important factor in limiting the false positive rate.

To determine the significance of low fragment ion mass tolerance on false positive rates, a series of reversed database searches were performed holding the precursor mass tolerance at 10 ppm and varying the fragment mass tolerance (Figure 4). In contrast to results obtained when the fragment mass tolerance was held constant, the results of this experiment showed that the number of false positives continued to increase as the fragment ion mass tolerance increased. Importantly, the number of false positive matches decreased quite dramatically as the fragment ion mass tolerance was decreased, indicating that high precursor ion mass accuracy is not sufficient to mitigate the effect of large fragment mass measurement errors.

Figure 4. Spectrum Mill score distributions of reversed database hits for variable fragment tolerance at constant 10 ppm precursor tolerance using the 6510 Q-TOF.
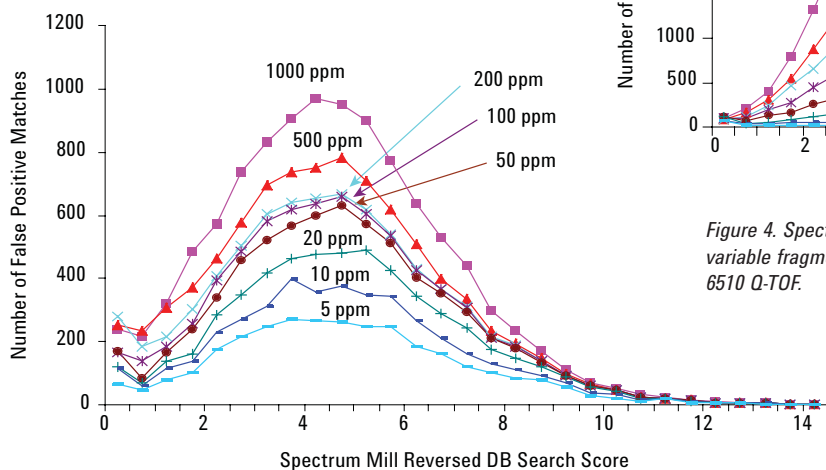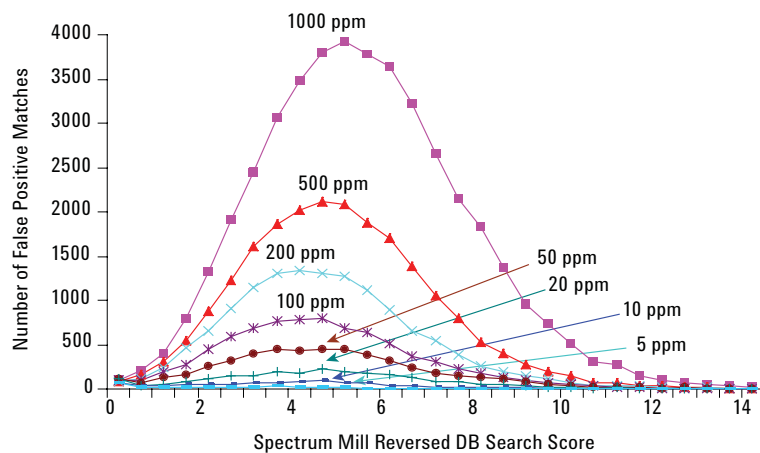

Figure 3. Spectrum Mill score distributions of reversed database hits for variable precursor tolerance at constant 40 ppm fragment tolerance using the 6510 Q-TOF.

19

We next examined the effects of mass accuracy on database search efficacy more closely by performing reversed database searches with 6510 Q-TOF data using different sets of precursor and fragment mass accuracy settings that represent three different MS instruments—an ion trap, a high-resolution hybrid linear ion trap, and the Q-TOF itself (Figure 5). As expected, when the precursor ion mass tolerance was improved from 1000 ppm (ion trap) to 5 ppm (high-resolution hybrid linear ion trap), the distribution of false positive matches decreased proportionally as well. However, when the fragment tolerance was lowered from 1000 ppm (ion trap & high-resolution hybrid linear ion trap) to 40 ppm (Q-TOF), this led to a substantial decrease in the score distribution of random matches. These results indicate that the decrease in false positive identification was attributable to not just the precursor ion mass tolerance but also just as strongly to the product ion mass tolerance.

In the final part of this study, we examined how both the precursor and fragment ion mass tolerances could ultimately affect confidence in protein identification. Using the Q-TOF data generated at 10 ppm precursor MMD and 40 ppm fragment MMD, the reverse score distributions and the validated protein identifications from the forward database search were compared (Figure 6). A false positive rate was estimated for each scoring bin by comparing the number of random matches in the reversed database search to the number of confidently identified peptide matches. The results indicated that high mass accuracy led to a downward shift in the score distribution of the reversed database matches relative to the forward search. As a result, the Spectrum Mill score required for a particular confidence level was lowered, leading to an increase in the number of confident hits. In fact, the number of confident matches increased up to 6.5-fold with the use of both highly accurate precursor and fragment ion mass measurements on the 6510 Q-TOF (Table 2).
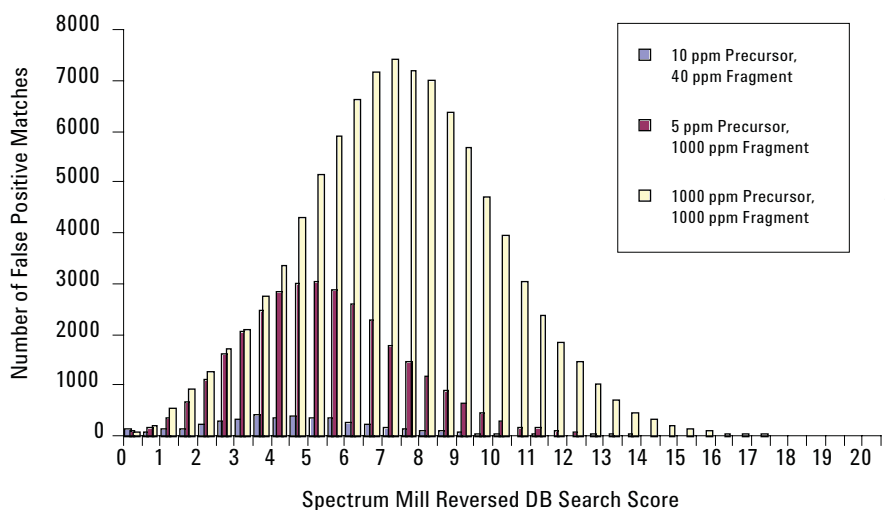

Figure 5. Spectrum Mill score distributions of reversed database hits using common precursor and fragment mass accuracy parameters on the 6510 Q-TOF.
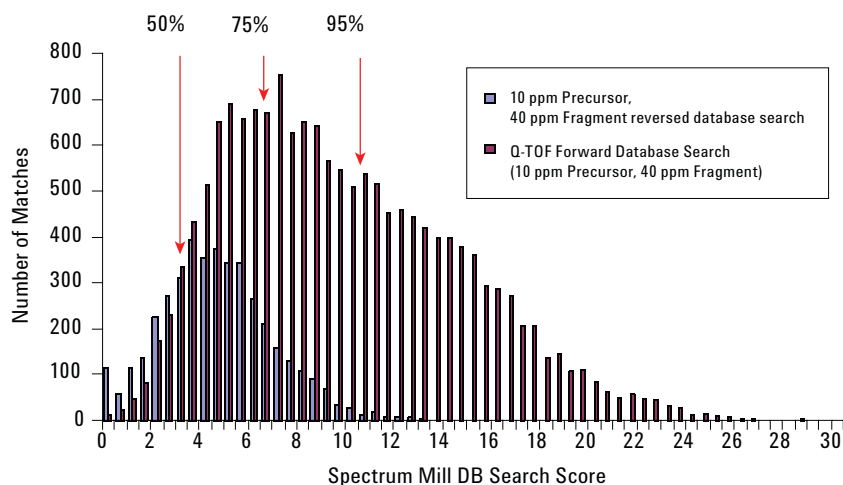

Figure 6. Calculation of peptide identification confidence by comparison of forward and reversed search results of 6510 Q-TOF data. The Spectrum Mill score required for a particular confidence level (denoted at the top of the chart) was lowered, leading to an increase in the number of confident hits.

Table 2. The number of confident hits for three different combinations of precursor and fragment mass accuracy settings on the 6510 Q-TOF.

ppm Tolerance		Number of Valid Peptide Matches with % Confidence			Number of Valid Protein Matches with % Confidence (minimum of 2 or more unique peptides/protein)		
Precursor	Fragment	>95%	>75%	>50%	>95%	>75%	>50%
1000	1000	954	2220	3361	102	213	282
5	1000	4181	6590	7645	338	478	524
10	40	6590	11552	15509	478	728	829

Table 2 also shows the number of protein matches with 2 or more unique peptide identifications with at least 95%, 75%, or 50% confidence. Using both high-accuracy precursor and fragment ion mass measurements, the 6510 Q-TOF demonstrated roughly a 5-fold improvement in the number of proteins identified with confidence (95% confidence interval). As a result, a much larger proportion of sample identifications can be automatically validated, greatly increasing the automated throughput of proteomics experiments.

Conclusions

The results of this study indicate that fragment ion mass accuracy is an important component to accurate protein identification. Overall, fragment ion mass accuracy was found to render a significant impact on false positive rates. When the fragment tolerance was lowered from 1000 ppm to 40 ppm, a substantial decrease in the score distribution of random matches in database searches was realized. The positive effect of low fragment mass tolerance has at least two explanations. First, large fragments generated by MS/MS of a peptide will exhibit a similar distribution of masses for a given nominal mass as the precursor. Thus, lowering the fragment tolerance will have a similar effect as lowering the precursor tolerance for these fragments, with the result of limiting the distribution of potential database matches. Second, a wider fragment mass tolerance will greatly increase the chance that random peaks in the MS/MS spectrum will be matched against random entries in the database. Decreasing the fragment tolerance will reduce the chance of such random matches.

These data demonstrate that high fragment ion mass accuracy can significantly increase valid protein identification by reducing the number of false positives when searching large databases for proteins, and consequently decreasing the score required for a high-confidence match. Other protein database search engines, such as Mascot, that can accommodate accurate mass data would be expected to yield similar results. In addition, these data demonstrate that the Agilent Accurate Mass Q-TOF LC/MS provides high mass accuracy for both the precursor and product ions, while generating MS/MS spectra at a rate that is well matched to complex sample LC/MS, unlike other high-resolution hybrid linear ion trap instruments.

References

1. Elias, J.E., S.P. Gygi (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods*, 4(3):207–14.

2. Liu, J., A.W. Bell, J.J. Bergeron, C.M. Yanofsky, B. Carrillo, C.E. Beaudrie, R.E. Kearney (2007) Methods for peptide identification by spectral comparison. *Proteome Sci*, 5:3.

3. Liu, T., M.E. Belov, N. Jaitly, W.J. Qian, R.D. Smith (2007) Accurate Mass Measurements in Proteomics. *Chem. Rev.*, 107:3621–3653.

4. VerBerkmoes, N.C., J.L. Bundy, J.L. Hause, K.G. Asano, J. Razumovskaya, F. Larimer, R.L. Hettich and J.L. Stephenson (2002) Integrating top-down and bottom-up mass spectrometric approaches for proteomic analysis of Shewanella oneidensis. *J of Proteom Res*, 1(3):239–252.

5. Zubarev, R., and M. Mann (2007) On the proper use of mass accuracy in proteomics. *Mol Cell Proteomics*, 6(3):377–81.

Data Showing the Feasibility of Confirming Putative Biomarkers Using Multiple Reaction Monitoring on the Agilent 6410 Triple Quadrupole LC/MS/MS System

Ning Tang, Hongfeng Yin, and Keith Waddell, Agilent Technologies, Inc., Santa Cara, California, U.S.A.

Abstract

Confirmation of putative protein biomarkers in complex biological samples requires an instrumental method that is both highly selective and sensitive. This work shows how a triple quadrupole mass spectrometer, in combination with a microfluidic HPLC-Chip, provides extremely specific biomarker confirmation at the low-attomole level.

Introduction

Biomarker discovery using mass spectrometry is now well established, so it is logical that mass spectrometry should also be used for the confirmation of candidate biomarkers. Multiple reaction monitoring (MRM) on a triple quadrupole (QQQ) mass spectrometer (MS) provides superior sensitivity and selectivity for targeted compounds in complex samples. MRM also offers high precision and rapid MS cycle time, which makes it an ideal technology for validating biomarkers in a high-throughput fashion.

In this study, we tested the MRM methodology by spiking standard protein digests into human plasma digest. To enable detection of low-level proteins, the most abundant proteins were removed by immunodepletion using the Agilent Multiple Affinity Removal System. We spiked various concentrations that represented the wide dynamic range that would be observed in that biological fluid.

The Agilent Spectrum Mill software was used to accelerate setup of the MRM method for the standard proteins. Spectrum Mill features the Peptide Selector tool, which was used to predict the optimal tryptic peptides and their MS/MS productions. These predicted results were then compared with experimental results from digests of the standard proteins and the lists of MRM transitions were created. The spiked plasma samples were analyzed by a robust and reproducible nanoflow LC/MS/MS using the Agilent HPLC-Chip interfaced to the high-performance Agilent 6410 Triple Quadrupole LC/MS/MS. The results showed excellent sensitivity, with on-column detection limits in the tens of attomoles.

Experimental

Sample preparation

A human plasma sample from a patient with rheumatoid arthritis (Genomics Collaborative) was used for this study. The sample was depleted of 14 highly abundant proteins with an Agilent Human 14 (HU-14) Multiple Affinity Removal Column. The depleted proteins include albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin, fibrinogen, alpha2-macroglobulin, alpha1-acid glycoprotein, IgM, apolipoprotein Al, apolipoprotein All, complement C3, and transthyretin. The standard Agilent protocol was used. After depletion, the sample was buffer-exchanged into an ammonium bicarbonate solution, reduced, alkylated with iodoacetamide (IAA), and digested with trypsin under denaturing conditions.

Luciferase (firefly), peroxidase (horseradish) and myoglobin were purchased from Sigma-Aldrich (St. Louis, MO) and digested as above. These standard protein digests were then spiked into aliquots of the immunodepleted plasma digest.

Standards for external calibration curves were prepared by spiking various amounts of luciferase or peroxidase digests into the depleted plasma digest. Additionally, the following test samples were prepared from these proteins:

- Sample A contained 10 fmol of luciferase spiked into 1 µg of immunodepleted human plasma digest.
- Sample B contained 500 amol luciferase spiked into 1 µg of immunodepleted human plasma digest.

Myoglobin standards were prepared for acquisition of quantitation curves with internal standards. Various amounts of myoglobin digest were spiked into the immunodepleted human plasma digest. An internal standard consisting of an isotopically labeled absolute quantitation (AQUA) peptide was spiked at 1 fmol into all of the myoglobin samples.

LC/MS/MS analysis

All samples were analyzed using the Agilent 6410 Triple Quadrupole LC/MS/MS connected to the Agilent HPLC-Chip Cube interface and Agilent 1100 Series LC modules.

HPLC-Chip

- Protein ID chip with 150 x 0.075-mm analytical column and 160-nL enrichment column
- Sample load: 1 µg of immunodepleted human plasma digest spiked with various amounts of standard protein digests
- Flow: 300 nL/min analytical pump, 2 µL/min loading pump
- Mobile phases: A = 0.1% formic acid (FA); B = 90% acetonitrile, 0.1% FA
- Gradients: 2% B to 8% B at 1 min, then 12% B at 5 min, 30% B at 40 min, 40% B at 45 min, 80% B at 50 min until 55 min, then 2% B at 55.1 min
- Stop time: 60 min
- Post time: 10 min

QQQ MS conditions

- Drying gas: 6 L/min, 325 °C
- Fragmentor: 120 V
- Collision energy: optimized for each MRM transition
- Dwell time: 200 ms
- Delta EMV: 700 V
- No. of transitions: 4 for luciferase study, 8 for myoglobin AQUA peptide study

Results and discussion

Peptide quantitation with external standards

In order to quantitate protein biomarkers, the MRM mode of the Agilent 6410 QQQ LC/MS/MS system was used to detect peptides. The MassHunter Quantitative software was then used to generate calibration curves. Various amounts of luciferase and peroxidase were spiked into aliquots of immunodepleted human plasma; the resulting samples were then analyzed. In an MRM experiment with a QQQ MS, Q1 (the first quadrupole) is set to pass a desired precursor ion, Q2 is used as a collision cell to fragment that precursor ion, and Q3 is set to monitor a specific fragment ion.

Using Peptide Selector to choose peptides and MRM transitions from proteins

For this study, the Peptide Selector module within the Agilent Spectrum Mill software was used to help choose MRM transitions that would produce sensitive analyses. Peptide Selector first performs an in silico digestion of the proteins, based on sequences that are manually copied into the software or are retrieved by an accession number from a database. The simulated digestion provides a list of component peptides. The Peptide Selector then filters the peptides based on user input such as the number of missed cleavage sites, the mass range and the specific amino acid composition. The software lists precursor and product ions for these peptides. It also highlights the best product ions for MRM experiments, based on the cleavages that are expected to produce intense signals (i.e., N-terminal side of Pro and C-terminal side of Asp and Glu). The software can also search any user selected database to determine the uniqueness of any peptide transitions that are being considered. Figure 1 shows an example Peptide Selector output for luciferase.

Figure 1. The Peptide Selector module in Agilent Spectrum Mill software assisted in the selection of MRM transitions that would produce the lowest detection limits.

The luciferase and peroxidase results from Peptide Selector were subsequently compared with Q-TOF results from the protein tryptic digests (see Figure 2 for Q-TOF results for peroxidase peptide MGNITPLTGTQGQIR) and the lists of MRM transitions were then created. The collision cell on the Agilent Q-TOF is identical to that of the QQQ. As shown in Table 1, data was acquired for multiple peptides from each protein and two transitions (two fragments) from each peptide.
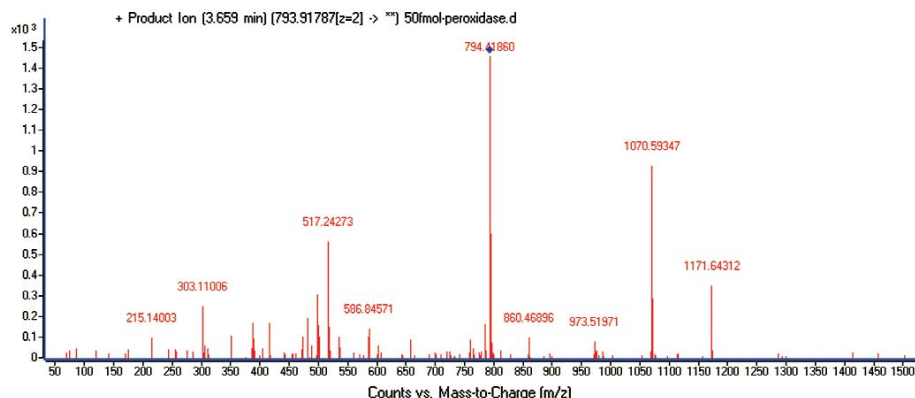

Figure 2. MS/MS spectrum of peptide MGNITPLTGTQGQIR with precursor mass at 793.9196. The major fragments are as labeled. Y10 and Y11 ions are chosen as the transitions for Q3. The m/z 1070.6 ion was successfully predicted by Peptide Selector.

Table 1. MRM transitions for peptides from luciferase and peroxidase.

Proteins	Peptides	Q1 Precursor m/z	z	Q3 Transition 1	Q3 Transition 2
Luciferase	FVDEVPK	417.2	2	587.3	686.4
	EIVDYVNSQVVNHK	548.6	3	651.8	594.3
	QHVALLMNSSGSTGLPK	580.9	3	748.9	659.4
Peroxidase	DTIVNELR	480.3	2	630.4	531.3
	SSDLVALSGGHTFGK	492.5	3	651.3	537.3
	MGNITPLTGTQGQIR	793.7	2	1070.6	1171.9

Linearity of external calibration curves

External calibration curves were plotted based on results of the MRM data acquisition experiments. The calibration curves for luciferase (Figure 3) and perox-idase (Figure 4) showed excellent linearity, with correlation coefficients of better than 0.998. In addition, these data demonstrated the wide dynamic range capability and sensitivity of the QQQ system, showing a range of four orders of magnitude. Figure 5 shows MRM chromatogram data for one peptide of luciferase at two transitions. The chromatograms are fairly clean and both the MRM taken for quantitation (bottom trace) and the MRM used as a qualifier show the peptide at a retention time of 18.4 minutes. These results are particularly impressive given the complexity of the plasma matrix, the low-attomole levels of some of the calibration standards, and the fact that internal standards were not used for this analysis.
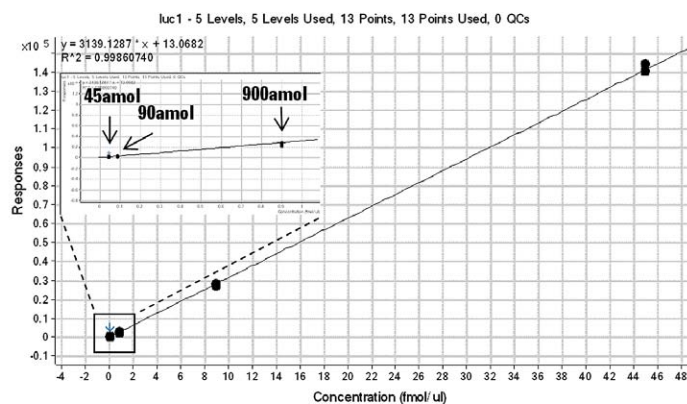

Figure 3. External quantitation curve of luciferase peptide FVDEVPK from 45 amol to 45 fmol.
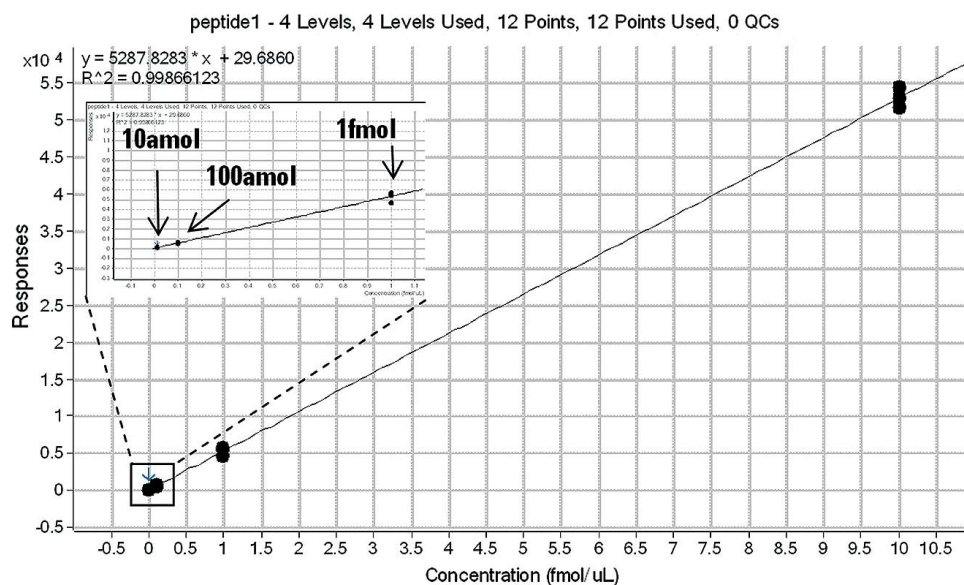
Figure 4. External quantitation curve of peroxidase peptide DTIVNELR from 10 amol to 10 fmol.
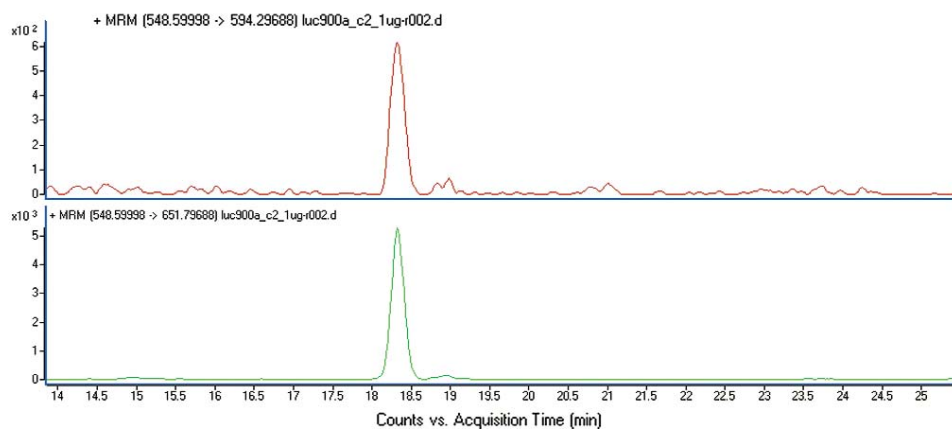

Figure 5. MRM chromatograms for peptide EIVDYVNSQVVNHK from luciferase. The top trace shows a transition of 548.6 to 594.3 m/z. The bottom trace shows a transition of 548.6 to 651.8 m/z. (900 attomoles of depleted serum injected.)

The calibration curve shown in Figure 3 was then used to quantitate luciferase in synthetic Samples A and B, with the results shown in Table 2. The results showed good agreement between the spiked concentration and the amount calculated based on the calibration curve. It should be noted that these results were obtained without the use of internal standards.

Table 2. Quantitation of luciferase in Samples A and B using external standard curve.

Sample	Spiked concentration (fmol/µL)	Calculated concentration (fmol/µL)	n	% RSD
A	10	13.9	3	3.4
B	0.5	0.55	3	7.2

25

Peptide quantitation with internal standards

Use of AQUA peptides as internal standards

As a next step to confirm the viability of MRM for protein biomarker validation, myoglobin was used as the test protein to evaluate MRM quantitation using internal standards. The internal standards were AQUA peptides – tryptic peptides of interest where a single amino acid is completely labeled with ^{13}C and ^{15}N. Table 3 shows the MRM transitions that were used to measure peak areas for the myoglobin peptides LFTGHPETLEK and ALELFR and their labeled counterparts.

Figure 6 illustrates excellent repeatability of peak areas for the MRM analysis of labeled ALELFR at the 40-amol level. These chromatograms show that the combination of the HPLC-Chip and the 6410 QQQ delivered outstanding precision, even for peptides at low-attomole levels.
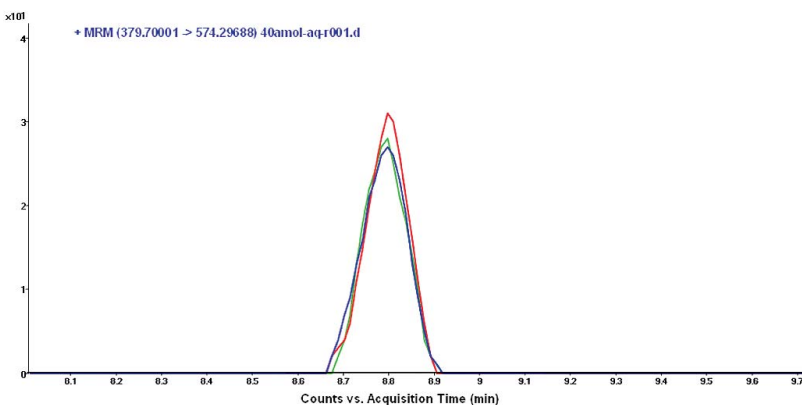
Table 3. Peptides and transitions for myoglobin quantitation.

Peptide	Precursor (Q1)	Ion type[†]	Fragment (Q3)
LFTGHPETLEK	636.3	Quant	716.4
		Qual	1011.6
LFTGHPETLEK*	640.3	Quant	724.4
		Qual	1019.9
ALELFR	374.7	Quant	564.3
		Qual	435.3
ALELFR*	379.7	Quant	574.3
		Qual	445.3

* = Isotopically labeled AQUA peptide

† : Quant/Qual = quantitation ion/qualifier ion

Figure 6. Detection limit and reproducibility of peptide ALELFR* at 40 amol.
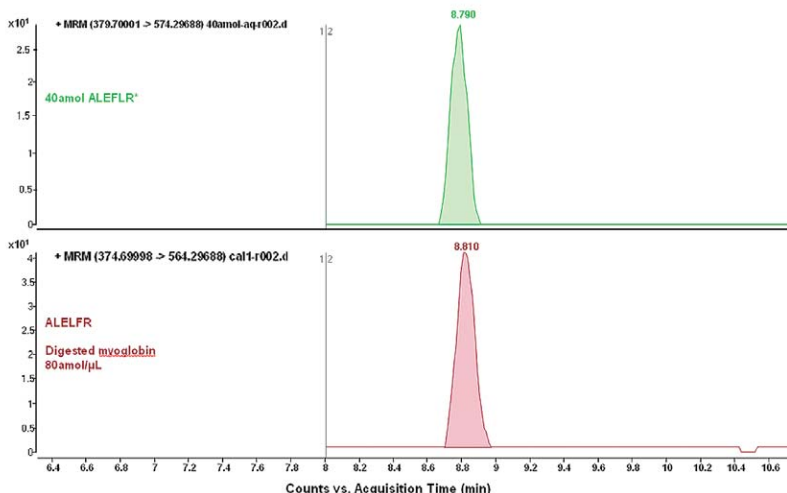
Figure 7. Single-point quantitation of ALEFR with labeled internal standard.

The chromatograms in Figure 6 also illustrate the superb selectivity of QQQ MRM for this analysis. Although the labeled ALELFR peptide was present in a complex mixture of plasma proteins, the chromatograms show clean baselines and an absence of interferences. With the MRM technique, the ions are detected only when a specific precursor ion gives rise to a specific fragment ion. This combination of signature ions has a low probability of occurrence for molecules other than the analyte of interest. With the flat baseline, the signal-to-noise is excellent even for the low-level peptides that are typically of interest in biomarker confirmation.

With the AQUA peptides, one typically quantitates peptides via peak ratios of the native peptides to the labeled internal standards, as shown in Figure 7. In the present study, the ALELFR peak from 80 amol digested myoglobin was calculated to be 63 amol using the ALEFLR* AQUA peptide at the 40-amol level. One possible explanation for this result would be the incomplete digestion of myoglobin, which would generate fewer of the targeted peptides.

Linearity of calibration curve with internal standard

To further evaluate the QQQ MRM approach for protein biomarker confirmation, a calibration curve was constructed for the peptide ALELFR, using the isotopically labeled analog as an internal standard (Figure 8). The linearity was excellent, showing the feasibility of this technique for reliable quantitation.
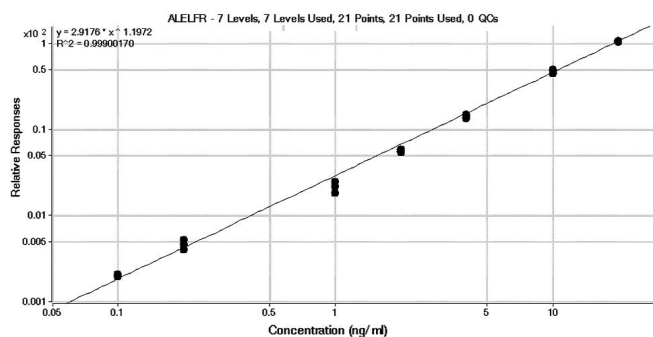

Figure 8. Calibration curve for ALELFR using ALELFR* as internal standard.

Conclusions

These experiments demonstrate the viability of triple-quad-based MRM for high-throughput validation of protein biomarkers. The Agilent HPLC-Chip interfaced to the Agilent 6410 Triple Quadrupole enabled robust, easy-to-use, sensitive, and reproducible nanoflow MRM analysis for peptide quantitation. While biomarker confirmation is often performed with antibody-based assays, the MRM approach is a reliable alternative that can save months of development time and considerable cost when the necessary antibodies are not commercially available.

For sensitive biomarker confirmation, it is extremely important to pick the best ions to monitor in Q1 and Q3 of the QQQ MS. The Peptide Selector in the Agilent Spectrum Mill software plays a valuable role in this task. This software performs theoretical digests of proteins and generates a list of peptides that are suitable for sensitive MS analysis, along with precursor and product ions for MRM experiments.

Used in MRM mode, the Agilent 6410 Triple Quadrupole LC/MS/MS system delivered excellent selectivity and limits of detection between 10 and 100 amol for standard protein digests spiked into immunodepleted human plasma digests. These low detection limits demonstrate that the MRM technique produces the sensitivity that is needed for real-world biomarker confirmation in complex biological samples.

High-Throughput Protein Quantitation Using Multiple Reaction Monitoring

Ning Tang, Christine Miller, Joe Roark, Norton Kitagawa, and Keith Waddell, Agilent Technologies, Inc., Santa Cara, California, U.S.A.

Abstract

Quantitative proteomics using multiple-reaction monitoring (MRM) has emerged as an important methodology for biomarker validation. Multiple-reaction monitoring on a triple quadruple (QQQ) mass spectrometer provides superior sensitivity and selectivity for targeted peptides in a complex sample. MRM also offers high precision in quantitation and a fast scan speed, which makes it an ideal technology for validating biomarkers in a high-throughput fashion. This application note demonstrates the entire workflow for biomarker validation offered by Agilent Technologies using multiple reaction monitoring together with statistical analysis software. Thousands of peptide transitions were generated using Agilent Q-TOF and Spectrum Mill MRM selector. The transition list was imported into MassHunter acquisition software equipped with dynamic MRM functionality and monitored in a single LC/MS run. Finally, MRM quantitation results were analyzed using Mass Profiler Professional for principal component analysis (PCA), hierarchical clustering, and ANOVA analysis.

Introduction

Peptide quantitation using multiple reaction monitoring (MRM) has been established as an important methodology for biomarker validation. Quantitative proteomics can require high-throughput as often hundreds of target peptides need to be monitored in each sample and thousands of biological samples may need to be analyzed. The dynamic MRM algorithm allows the system to acquire transition ion data only during the retention window when each peptide is eluting. This reduces the number of concurrent ion transitions and therefore improves quantitation and sensitivity. In this study, peroxidase was spiked at different concentrations into human plasma to demonstrate the entire workflow from biomarker discovery to validation. Reproducibility of peak abundances and retention time at nanoflow range were studied with 443, 2,000 and 3,293 ion transitions using dynamic MRM method on a nanoflow LC/MS system.

Experimental

Sample preparation

Human plasma sample was purchased from Sigma (St. Louis, MO). The sample was depleted of 14 highly abundant proteins using a Hu-14 immunoaffinity column (Agilent) following the standard protocol. After depletion, the sample was buffer-exchanged into an ammonium bicarbonate solution, reduced, alkylated (IAA) and digested with trypsin under denaturing conditions. Horseradish peroxidase was purchased from Sigma (St. Louis, MO), reduced, alkylated and digested with trypsin. Peroxidase digest was spiked at 500 amol (A) or 5 fmol (B) per 0.5 µg human plasma digest.

LC/MS analysis

With electrospray LC/MS, smaller chromatographic elution volumes result in enhanced peak height and thus greater sensitivity. A unique microfluidic chip, the HPLC-Chip, has been developed for nanoflow LC/MS that integrates sample enrichment and separation columns, microvalve connections, and the nanospray tip on a biocompatible polyimide chip. Interfacing this device to a triple quadrupole mass spectrometer provides easy-to-use, highly reliable, and highly sensitive LC/MS analysis. For this work, the HPLC-Chip was interfaced to an Agilent 6410 Triple Quadrupole (QQQ) LC/MS and an Agilent 6520 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) LC/MS.

HPLC-Chip Parameters:

Chip and columns: Protein ID chip with 150 x 0.075 mm analytical column and 40 nL enrichment column.

Sample load: 0.5 µg of human serum digest spiked with different amounts of peroxidase digest.

Injection volume: 1 µL.

Flow: 300 nL/min analytical pump, 3 µL/min loading pump.

Mobile phases: A. 0.1% formic acid (FA), B. 90% acetonitrile (ACN), 0.1% FA.

Gradients: 3% B at 0 min,10% B at 3 min, 12% B at 8 min, 30% B at 42 min, 45% B at 45 min, 70% B at50 min, 90% B at 90 55 min, then 3% B at 55.1 min.

Stop time: 60 min.

Post time: 10 min.

MS Conditions:

Drying gas: 5 L/min, 325°C

Collision energy: slope 3.6, offset -4.8

Capillary voltage: 1,800 V.

Software

The Q-TOF data was searched against the SwissProt database using Agilent Spectrum Mill MS Proteomics Workbench. A new tool within Spectrum Mill, MRM Selector, was used to directly generate the dynamic MRM methods based on results from the Q-TOF database search results. After LC/MS analysis, the results were analyzed using MassHunter Quantitative Analysis software. From this, quant batch report XML files were imported into Mass Profiler Professional software, a chemometrics software package designed specifically for mass spectrometry data. Spiked-in peptide features were analyzed in the context of human serum peptides via principal components analysis (PCA). Additionally, a naïve hierarchical clustering analysis was performed.

Results and Discussion

Step 1: Q-TOF

The biomarker validation workflow is illustrated in Figure 1. In the first step, the samples were run on HPLC-Chip/ Q-TOF in data-dependent MS/MS mode. The HPLC-Chip provided excellent reproducibility as shown by the overlaid base peak chromatograms (BPC) of 5 replicate injections (Figure 2).
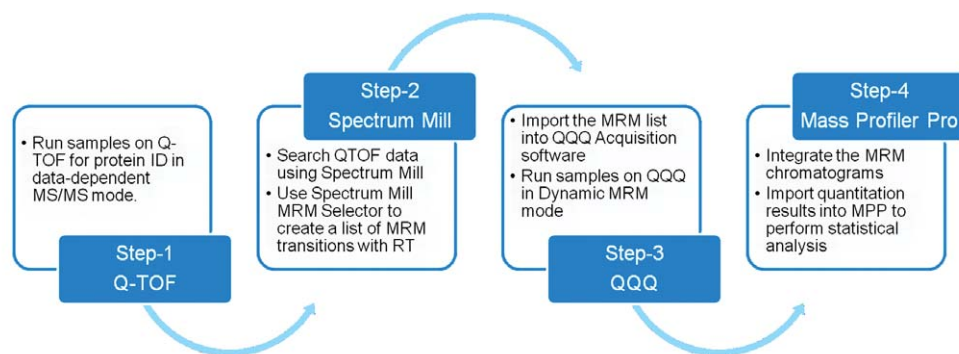

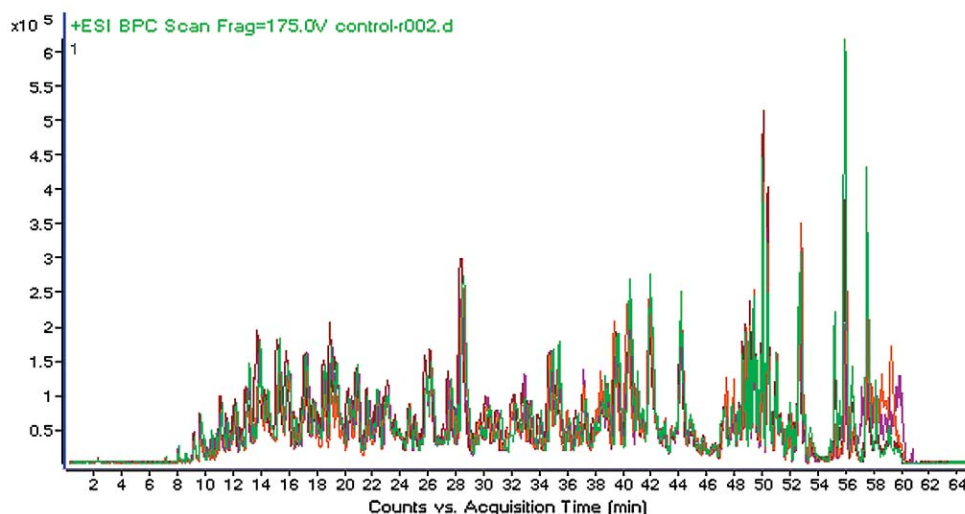Figure 1. Biomarker validation workflow.

Figure 2. Overlaid BPCs from 5 replicate runs of depleted human plasma tryptic digest on HPLC-Chip/ Q-TOF. The samples were analyzed in data-dependent MS/MS mode for protein identification.

Step 2: Spectrum Mill (Figure 3)

Q-TOF data was searched using Spectrum Mill. A dynamic MRM list was generated using MRM Selector based on the validated peptide hits. MRM Selector is a utility tool in Spectrum Mill workbench that allows the user to select the MRM transitions from the experimental MS/MS data. The user can input several parameters (listed below) to filter the ion transitions to be monitored on QQQ. The MRM Selector results contain protein accession number and peptide sequence, ion transition values, retention time (RT), peak width, collision energy, and fragmentor values. The saved list can be pasted directly into the QQQ acquisition software.

MRM Selector parameters:

- Number of peptides per protein
- Number of product ions per peptide, choice of above precursor and y-ions
- Peptide score and %SPI
- Required AA and disallowed AA
- Peptide pI
- Protein accession number

Figure 3. MRM Selector generates dynamic MRM methods from discovery Q-TOF data.

Step 3: QQQ (Figures 4–6)

The dynamic MRM lists containing hundreds to thousands of ion transitions were imported into a QQQ acquisition method. The cycle time was set to obtain at least 15 data points across the peaks.

In order to assess the reproducibility of the MS and RT with increasing number of ion transitions, four dynamic MRM experiments were set up using MRM Selector with 443, 2,000 and 3,293 ion transitions in each method. The retention time (RT) window for monitoring the ion transitions was also varied between 1–2 min resulting in different minimum dwell time and maximum number of concurrent MRMs. The RSD of MS response and RT from 12 peroxidase peptide transitions were calculated and are listed in Table 1. The %RSD of the MS responses was below 5% while the RSDs of RT were less than 0.04 min for the 60-minute run.
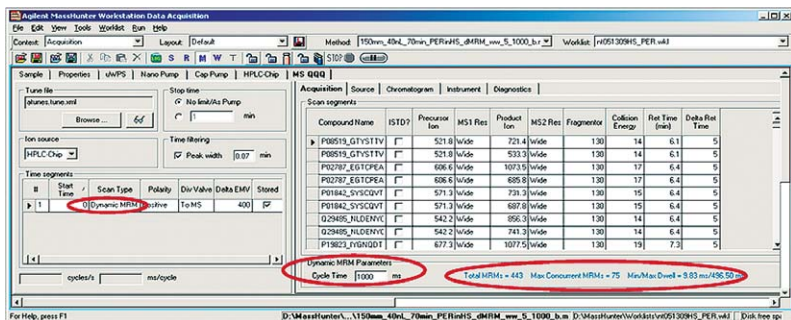

Figure 4. Example of dynamic MRM method.
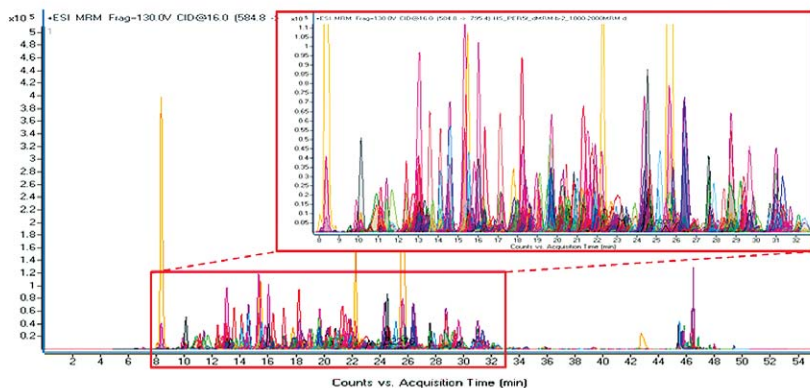
Figure 5. Overlaid 2,000 MRM chromatograms acquired in a single run using dynamic MRM.
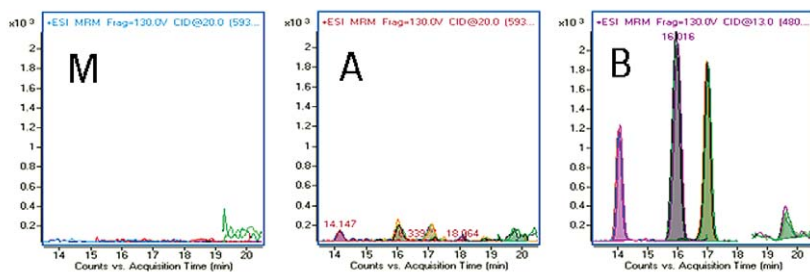
Figure 6. Overlaid MRM chromatograms of peroxidase peptides showed excellent reproducibility of retention time and MS response. M: human plasma matrix; A: 500 amol peroxidase spiked into human plasma; B: 5 fmol peroxidase spiked into human plasma.

Table 1. Reproducibility of MS response and RT.

# MRM	RT window (min)	Cycle time (ms)	Min. dwell (ms)	Max. # concurrent MRM	% RSD Area	RSD RT (min)
443	2	1,000	16.5	50	2.5	0.038
443	1	1,000	29.83	30	3.2	0.016
2,000	2	1,000	2.75	160	4.5	0.030
3,293	1	1,050	2.18	185	4.7	0.025

Step 4: Mass Profiler Professional

Mass Profiler Professional can easily classify, compare, and analyze sample groups using a combination of powerful statistical and mathematical models to analyze complex MS data sets. Figure 7 shows the abundance of of 443 ion transitions across all samples (three B samples, three A samples and two controls). Each line represents one ion transition and the color represents the relative intensity of the ion transition (red being high, gray being low). The ion transitions for the four peptides from peroxidase are highlighted in green. The mean of the 443 abundances is displayed (black) to show that the peptides from plasma did not vary from sample to sample.

Principal component analysis (PCA) can be used to find differences in groups in either an unsupervised fashion (classify without knowing the group assignments) or a supervised fashion (must classify groups before analysis). As shown in Figure 8, both levels of peroxidase-spiked plasma were clearly differentiated from the control plasma.
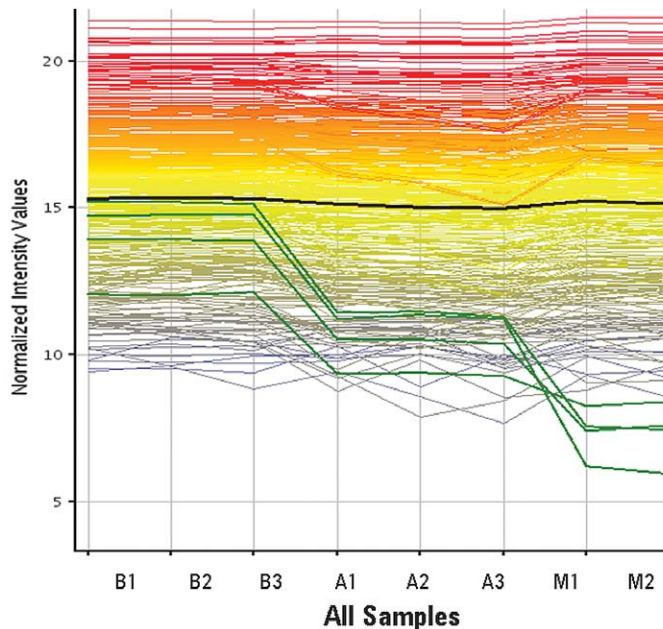

Figure 7. Profile of abundances across all samples for 443 ion transitions. Four peptides from peroxidase were highlighted in green. The mean of the 443 abundances is displayed (black) to show that the peptides from plasma did not vary from sample to sample.
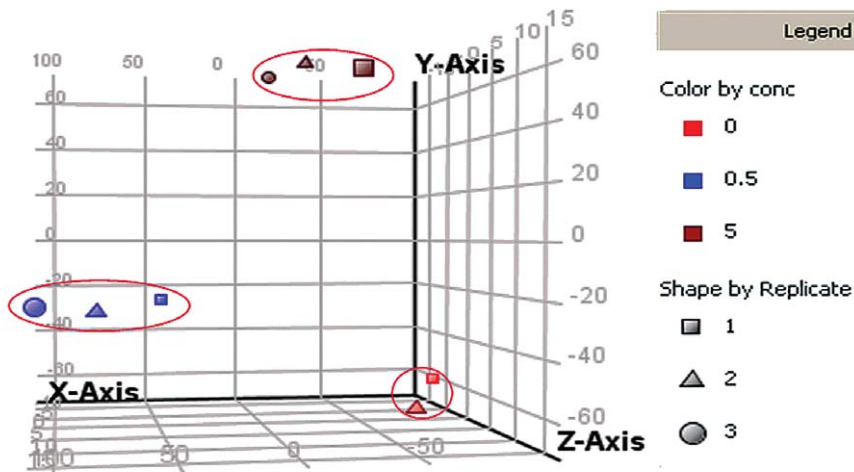

Figure 8. PCA analysis of different samples. Samples at different peroxidase concentrations were correctly grouped together.

32

Clustering analysis by sample groups organizes the relationships based on the similarity of entities' abundance profiles. The tree diagram produced by hierarchical clustering in Mass Profiler Professional reveals the relationships between mass entities in one dimension and between samples in the other dimension. In Figure 9, the technical replicates for the different peroxidase levels were correctly grouped using hierarchical clustering.
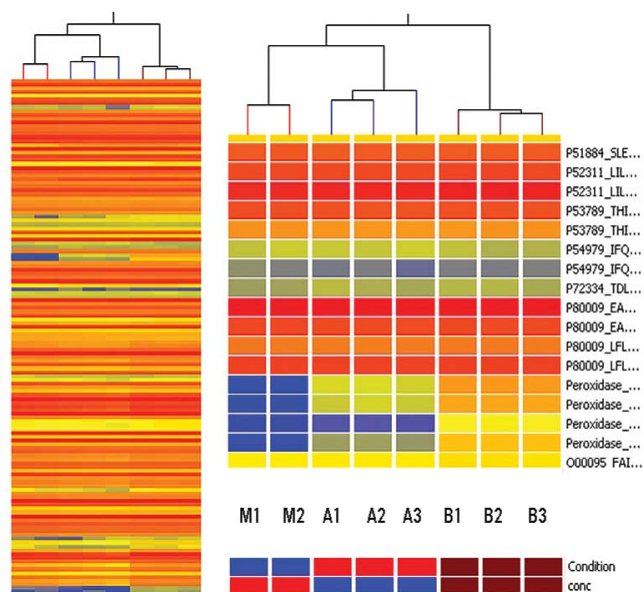

Figure 9. Hierarchical clustering successfully clustered samples at different peroxidase concentrations. A condition was generated with peroxidase concentration color-coded on the tree branches as in Figure 8, along with the peptide features labeled on each row. The heat map is colored from blue to red, where blue is low abundance and red is high abundance. The full view of all the features is on the left. The zoom view is on the right.

Table 2. Analysis of Variance. A one-way ANOVA on concentration was performed on the peptide abundances. Benjamini-Hochberg multiple testing correction was applied. Additionally, a filter for fold change ≥ 5.0 was applied to the list. The four peroxidase peptides each had a corrected p-value of 0.0.

Compound	Corrected p-value	p-value	Retention time
Peroxidase_DTI	0	0	15.952
Peroxidase_GFP	0	0	16.999
Peroxidase_YYV	0	0	14.034
Peroxidase_SSD	0	0	19.595

Conclusions

A complete biomarker workflow from Q-TOF discovery to QQQ validation has been demonstrated using a spike-in study with human plasma. The protein database search results from Q-TOF discovery data was used to create dynamic MRM methods via MRM Selector, a new tool being developed in Spectrum Mill software. These dynamic MRM methods allowed hundreds to thousands of peptides to be monitored in a single LC/MS run, while still producing excellent RSDs on MS abundance and retention time. Mass Profiler Professional software allowed chemometric analysis of the quantitative results for confirmation of the significant differences between samples. Excellent reproducibility of the HPLC-Chip/MS system is the key element for high-throughput and sensitive analysis of biomolecules.

This material was presented in a poster at the 57th ASMS Conference on Mass Spectrometry in 2009. Research presented in posters at scientific conferences may include results from instruments or products that are not yet commercially available.

Quantitation of Protein Phosphorylation Using Multiple Reaction Monitoring

Ning Tang, Christine A. Miller and, Keith Waddell, Agilent Technologies, Inc., Santa Clara, California, U.S.A.

Abstract

This application note demonstrates quantitation of percentage of phosphorylation of two sites on a single peptide using the HPLC-Chip/QQQ system. ^{13}C-labeled peptides from the extracellular signal-regulated kinase (ERK) were synthesized and MRM assays were developed to quantify each peptide. The standard curve of each ^{13}C-labeled peptide was established. The amount of each phosphopeptide from active ERK spiked into human plasma was measured and calculated in reference to the ^{13}C-labeled standards.

Introduction

Peptide quantitation using multiple-reaction monitoring (MRM) has emerged as an important methodology for biomarker validation. MRM on a triple quadruple mass spectrometer provides superior sensitivity and selectivity for targeted compounds in a complex sample. MRM also offers high precision in quantitation and fast scan speed, which makes it an ideal technology for validating biomarkers in a high-throughput fashion.

One area of great interest is the quantitation of protein phosphorylation. Reversible protein phosphorylation plays an important role in cell signaling pathways and the percentage of phosphorylation is often very important to signal transduction. The ERK pathway is essential to trans-mitting signals from many extracellular agents to regulate cellular processes such as proliferation, differentiation and cell cycle progression. Signaling via the ERK cascade is mediated by sequential phosphorylation and activation of protein kinases in the different tiers of the cascade. The recombinant human ERK1 containing an N-terminal GST tag was expressed in *E. coli* and activated with MEK1.

Up to 7 phosphorylation sites were identified from the activated ERK1. The two key regulatory phosphorylation sites have been identified as the neighboring t202 and y204.[1] Phosphorylation of these sites, located upstream of the conserved protein kinase subdomain VIII region impli-cated in substrate binding, might function by relieving inhibition of substrate access to the active sites of the kinase domain. The peptides that contain either no phosphorylation site, a single phosphorylation site, or both phosphorylation sites have been identified. Hence, the degree of phosphorylation of these two sites is very important to the overall function of ERK1 in the signal transduction pathway. Therefore, in this study, we developed an MRM-based assay to quantify the percentage of phosphorylation by reference to ^{13}C-labeled peptides.

Experimental

Sample Preparation

Four peptides labeled at the valine residue with ^{13}C were purchased. The 4 peptides were as follows: 1) without phosphorylation, 2) with single phosphorylation at t202 or 3) with single phosphorylation at y204 and 4) double phosphorylation at both t202 and y204.

Active MAP kinase ERK1 was purchased from Millipore (Temecula, CA). It was denatured, reduced, alkylated and digested using trypsin. ^{13}C-labeled peptide standards were purchased from Cell Signaling Technology (Danvers, MA). The peptides were labeled at the valine residue. Human plasma was purchased from Sigma (St. Louis, MO). The plasma was depleted of six highly abundant proteins using an Hu-6 immunoaffinity column (Agilent) following the standard protocol. After depletion, the sample was buffer-exchanged into an ammonium bicarbonate solution, then reduced, alkylated (IAA), and digested with trypsin under denaturing conditions. Internal calibration curves were acquired using different amounts of ^{13}C-labeled peptides spiked into 1 µg of human plasma digest. ERK digest (35 fmol/µg plasma protein) was spiked into the samples as the internal standard.

LC/MS Analysis

An Agilent 6410 Triple Quadrupole (QQQ) LC/MS was connected to the HPLC-Chip interface (see Figure 1).

HPLC-Chip Conditions:

Protein ID chip with 150 x 0.075 mm analytical column and 160 nL enrichment column.

Sample load: 1 µg of human plasma digest spiked with different amounts of standard protein digests

Injection volume: 1 µL

Flow: 300 nL/min analytical pump, 3 µL/min loading pump

Mobile phases A: 0.1% formic acid (FA) in water, B: 90% acetonitrile, 0.1% FA

Gradients: 2% B to 42% B at 20 min, 90% B at 23–25 min, then 2% B at 25.1 min

Stop time: 35 min

Triple Quadrupole MS Conditions:

Drying gas: 6 L/min, 325°C

Fragmentor: 100 V

Collision energy: optimized for each transition

Dwell time: 70 ms

Delta EMV: 400 V

Results and discussion

Previous work[2] demonstrated that peptide quantitation using the HPLC-Chip Agilent 6410 Triple Quadrupole LC/MS can be achieved down to 10 amol. In this study, we took on the challenge of quantitating the degree of phosphorylation of two sites on a peptide from ERK protein (amino acid residues 190–208).

The sequence and the product ions selected for monitoring the synthetic peptides in the MRM assay are illustrated in Figure 2 and Table 1. The selection of the transitions was critical because of the similarity of the sequences.

Figure 1. HPLC-Chip/QQQ system for protein quantitation.
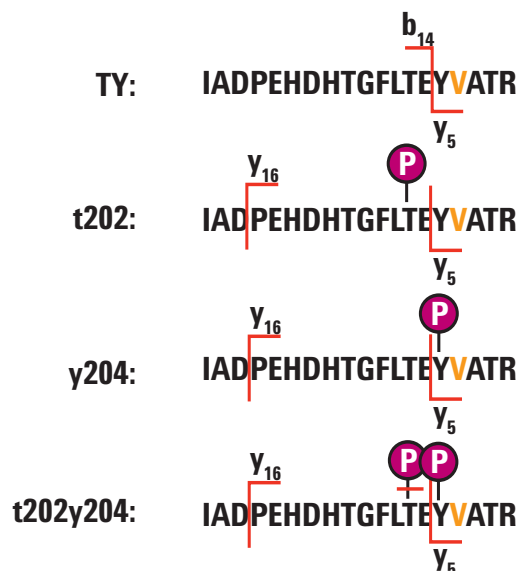
Figure 2. Sequence and product ions selected for monitoring the four ^{13}C-labeled peptides. Each peptide was named by the phosphorylation state at amino acids 202 and 204. Phosphotyrosine was labeled as lower case y and phosphothreonine was labeled as lower case t. V represents ^{13}C-labeled valine.

Table 1. The charge states, type of fragment and *m/z* values of the precursor ions and product ions monitored for each ^{13}C-labeled peptide standard. Bold indicates transition shared between two peptides.

	Precursor ion	Product ions	
TY	545.3 [M+3H] $^{3+}$	615.3 y_5	782.5 $b_{14}{}^{2+}$
t202	**753.3** [M+2H] $^{2+}$	615.3 y_5	**979.9** $y_{16}{}^{2+}$
y204	**753.3** [M+2H] $^{2+}$	**979.9** $y_{16}{}^{2+}$	695.3 y_5
t202y204	780.0 [M+2H] $^{2+}$	647.6 $y_{16}{}^{3+}$-H_3PO_4	695.3 y_5

35

Chromatographic separation of the four peptide standards

The transition from 753.3 –> 979.9 is the same for both the t202 and y204 peptide, thus it was important to have complete chromatographic separation. The four ^{13}C-labeled peptide standards were mixed at equal molar ratios and 100 fmol was injected on the HPLC-Chip. A gradient method was developed to separate the four peptides (Figure 3).
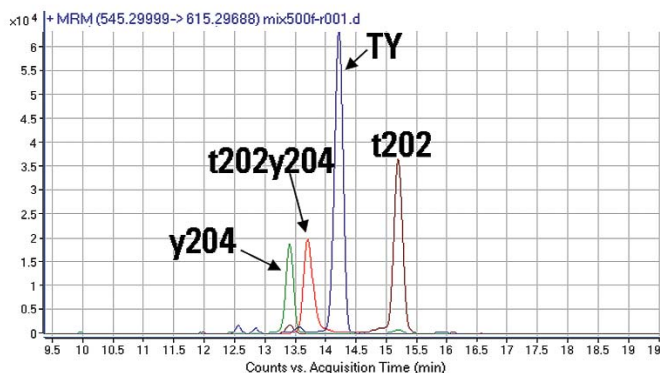

Figure 3. MRM chromatogram of four ^{13}C-labeled peptides showed complete chromatographic separation.

QC of incoming phosphopeptide standards

We also performed quality control of the incoming ^{13}C-labeled peptides using the MRM assays developed. The four ^{13}C-labeled peptide standards were analyzed individually (500 fmol on-column). The non-phosphory-lated peptide (TY) was pure and does not contain phosphorylation residues (Figure 4). The two singly phosphorylated peptides (t202 and y204) contained a small percentage of non-phosphorylated peptide (TY). The doubly phosphorylated peptide (t202y204) contained both the singly phosphorylated peptide (t202) and the non-phosphorylated peptide (TY). However, overall the percentage of impurities was relatively low (Table 2). The purity of each synthesized peptide was above 97.6%.
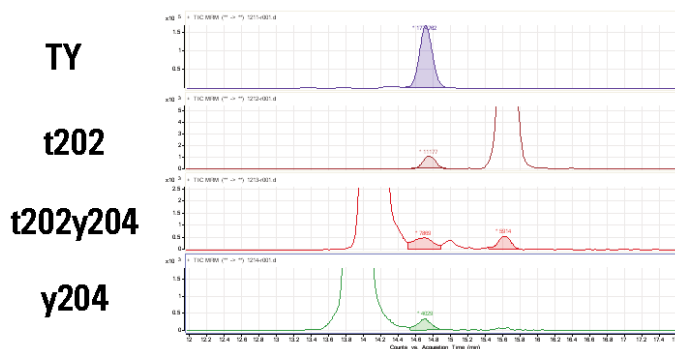

Figure 4. The four ^{13}C-labeled peptide standards were injected individually at 500 fmol on-column. Trace amounts of incomplete phosphorylation were observed. The percentage of impurities for each synthetic peptide was analyzed by MRM and calculated in Table 2.

Table 2. The percentage of each component in the synthetic peptides was calculated by the ratio of MRM peak area. All four peptides are greater than 97% pure.

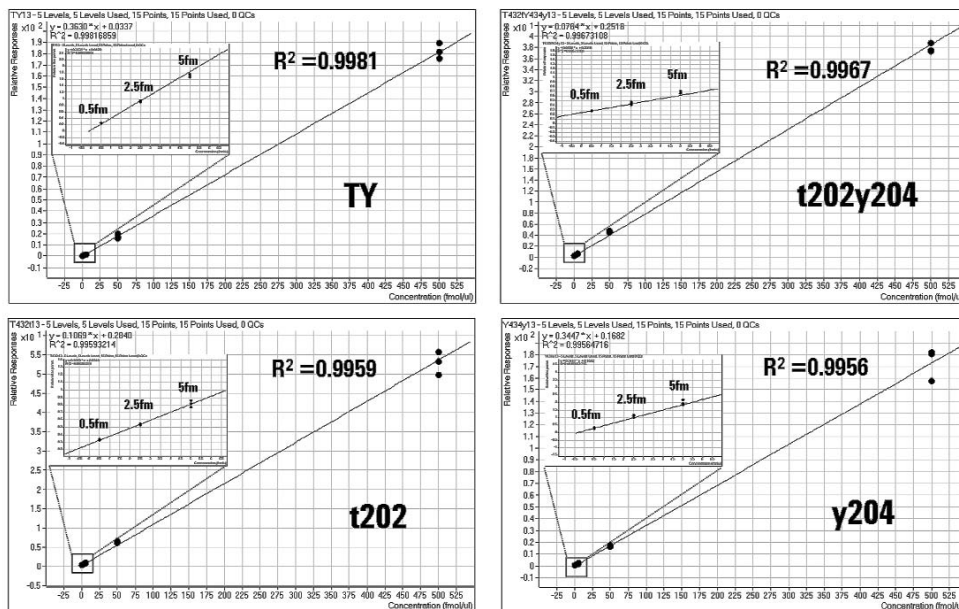| Peptides	Percentage of components			
	TY	t202	y204	t202y204
TY	100%	–	–	–
t202	0.6%	99.4%	–	–
y204	0.2%	–	99.8%	–
t202y204	0.4%	2%	–	97.6%

Figure 5. The calibration curves of the four ^{13}C-labeled peptides spiked into human plasma.

Quantitation curves for the four ^{13}C-labeled peptide standards

For each LC/MS analysis, 1 μg of human plasma digest was injected. The four calibration curves all had excellent linearity with correlation coefficients better than 0.995 (Figure 5). Calibration curves were acquired using human plasma digest spiked with both ^{13}C-labeled peptides and ERK1 tryptic digest. Limit of quantitation (LOQ) was determined to be 500 amol and limit of detection (LOD) was 50 amol.

The degree of phosphorylation at t202 and y204 in active ERK1 protein

The degree of phosphorylation can be measured by the relative percentage of each peptide as shown in Figure 2. The ratio between the abundance of unlabeled peptide from ERK tryptic digest and the labeled peptide in each sample provided the absolute amount of unlabeled peptide. The percentage for each phosphorylation form was derived from these results and listed in Table 3. The relative standard deviation was below 15%.

Table 3. Relative amount of four different peptides detected in active ERK1. RSDs of all measurements are below 15%.

	% Molar ratio	RSD (n=9)
TY	20%	0.13
t202	25%	0.15
y204	21%	0.12
t202y204	34%	0.08

Conclusions

- An MRM-based method to quantitate the different phosphorylation states on a single peptide was developed. A critical part of the method development was the chromatographic separation of the target peptides.

- Because of the sequence similarity between these peptides, the selection of transitions was critical. Chromatographic separation of the peptides allowed the same transition to be used for different peptides.

- The microfluidic-based HPLC-Chip has been demonstrated to be sensitive, reproducible and capable of fast separation, especially in combination with a triple quadrupole MS/MS.

References

1. D. L. Charest, G. Mordret, K. W. Harder, F. Jirik, S. L. Pelech (1993) Molecular cloning, expression, and characterization of the human mitogen-activated protein kinase p44erk1. *Mol. Cell. Biol.*, 13:4679–4690.

2. N. Tang, C. Miller, K. Waddell (2007) Approaches for the Confirmation of Putative Biomarkers Using Multiple Reaction Monitoring. HUPO Conference Proceedings.

This material was presented in a poster at the 56th ASMS Conference on Mass Spectrometry in 2008. Research presented in posters at scientific conferences may include results from instruments or products that are not yet commercially available.

Citations

A list of papers from peer-reviewed proteomics publications using Agilent LC/MS systems.

Q-TOF LC/MS

1. Quantitative proteomic profiling of prostate cancer reveals a role for miR-128 in prostate cancer.

Khan AP, Poisson LM, Bhat V, Fermin D, Zhao R, Kalyana-Sundaram S, Michailidis G, Nesvizhskii AI, Omenn GS, Chinnaiyan AM, Sreekumar A.
Mol Cell Proteomics. 2009 Nov 9. [Epub ahead of print]

2. Prioritization of Candidate Protein Biomarkers from an In Vitro Model System of Breast Tumor Progression Towards Clinical Verification.

Lau TY, Power KA, Dijon S, de Gardelle I, McDonnell S, Duffy M, Pennington S, Gallagher WM.
J Proteome Res.. [Epub ahead of print]

3. A depletion strategy for improved detection of human proteins from urine.

Kushnir MM, Mrozinski P, Rockwood AL, Crockett DK
J Biomol Tech. 2009 Apr;20(2):101-8

4. Elevated cleavage of human immunoglobulin gamma molecules containing a lambda light chain mediated by iron and histidine.

Ouellette D, Alessandri L, Piparia R, Aikhoje A, Chin A, Radziejewski C, Correia I.
Anal Biochem. 2009 Jun 15;389(2):107-17. Epub 2009 Mar 24

5. Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of Arabidopsis cell culture peroxisomes.

Eubel H, Meyer EH, Taylor NL, Bussell JD, O'Toole N, Heazlewood JL, Castleden I, Small ID, Smith SM, Millar AH.
Plant Physiol. 2008 Dec; 148(4):1809-29. Epub 2008 Oct 17.

6. Structure and evolution of a novel dimeric enzyme from a clinically important bacterial pathogen.

Burgess BR, Dobson RC, Bailey MF, Atkinson SC, Griffin MD, Jameson GB, Parker MW, Gerrard JA, Perugini MA.
J Biol Chem. 2008 Oct 10;283(41):27598-603. Epub 2008 Aug 5

7. Amyloid-beta-anti-amyloid-beta complex structure reveals an extended conformation in the immunodominant B-cell epitope.

Miles LA, Wun KS, Crespi GA, Fodero-Tavoletti MT, Galatis D, Bagley CJ, Beyreuther K, Masters CL, Cappai R, McKinstry WJ, Barnham KJ, Parker MW.
J Mol Biol. 2008 Mar 14;377(1):181-92. Epub 2008 Jan 30.

8. AMP-activated protein kinase subunit interactions: beta1:gamma1 association requires beta1 Thr-263 and Tyr-267.

Iseli TJ, Oakhill JS, Bailey MF, Wee S, Walter M, van Denderen BJ, Castelli LA, Katsis F, Witters LA, Stapleton D, Macaulay SL, Michell BJ, Kemp BE.
J Biol Chem. 2008 Feb 22;283(8):4799-807. Epub 2007 Dec 13.

TOF LC/MS

1.Bovine milk glycome.

Tao N, DePeters EJ, Freeman S, German JB, Grimm R, Lebrilla CB.
J Dairy Sci. 2008 Oct;91(10):3768-78.

2. Molecular mass analysis of antibodies by on-line SEC-MS.

Brady LJ, Valliere-Douglass J, Martinez T, Balland A.
J Am Soc Mass Spectrom. 2008 Apr;19(4):502-9. Epub 2008 Jan 4.

3. Daily variations in oligosaccharides of human milk determined by microfluidic chips and mass spectrometry.

Niñonuevo MR, Perkins PD, Francis J, Lamotte LM, LoCascio RG, Freeman SL, Mills DA, German JB, Grimm R, Lebrilla CB.
J Agric Food Chem. 2008 Jan 23;56(2):618-26. Epub 2007 Dec 19.

4. Glycoprofiling of bifidobacterial consumption of human milk oligosaccharides demonstrates strain specific, preferential consumption of small chain glycans secreted in early human lactation.

LoCascio RG, Ninonuevo MR, Freeman SL, Sela DA, Grimm R, Lebrilla CB, Mills DA,German JB.
J Agric Food Chem. 2007 Oct 31;55(22):8914-9. Epub 2007 Oct 5.

Triple Quadrupole LC/MS

1. Identification of cows treated with recombinant bovine somatotropin.

Le Breton MH, Rochereau-Roulet S, Chéreau S, Pinel G, Delatour T, Le Bizec B.
J Agric Food Chem. 2010 Jan 27;58(2):729-33.

2. HPLC-CHIP coupled to a triple quadrupole mass spectrometer for carbonic anhydrase II quantification in human serum.

Callipo L, Foglia P, Gubbiotti R, Samperi R, Laganà A.
Anal Bioanal Chem. 2009 Jun;394(3):811-20. Epub 2009 Mar 24.

Agilent Literature

Please contact your Sales Representative to order the following:

Brochures

5989-7761EN: Faster, Easier, Flexible Proteomics Solutions

5990-4052EN: Agilent Accurate-Mass 6500 Series Q-TOF and 6200 Series TOF LC/MS Systems: The clearly better choice in TOF and Q-TOF LC/MS

5990-4788EN: Agilent 6430 Triple Quadrupole LC/MS System: Ideal Quantitative LC/MS for UHPLC with Dynamic MRM and Fast Polarity Switching, plus Unsurpassed Sensitivity with Chip LC

5989-9897EN: Agilent 1200 Series HPLC-Chip/MS System – Easy, reliable and completely integrated microfluidic nanospray LC/MS

5990-4164EN: Agilent Mass Profiler Professional Software – MS data: Don't just process it. Understand it.

5990-3567EN: Agilent Proteomics Reagents for LC/MS

Technical Overviews

5990-3595EN: New Dynamic MRM Mode Improves Data Quality and Triple Quad Quantification in Complex Analyses

5990-4499EN: Better detection of trace components: in-spectrum dynamic range of Agilent 6540 UHD Q-TOF exceeds that of Thermo LTQ Orbitrap XL

5990-4337EN: Agilent 6540 UHD Q-TOF produces more accurate isotope ratios than Thermo LTQ Orbitrap XL

5990-4507EN: Superior resolution of Agilent 6540 UHD Q-TOF over Thermo LTQ Orbitrap XL for fast UHPLC applications

5990-3168EN: Achieving Unmatched Speed and Confidence for Complex Sample Analyses: Agilent True High-Definition TOF Technology

5989-7408EN: Ion optics innovations for increased sensitivity in hybrid MS systems

Application Notes

5990-4098EN: Facilitating Phosphopeptide Analysis Using the Agilent HPLC Phosphochip

5989-3499EN: Comparison of Different Approaches for the Label-Free Relative Quantitation of Proteins

5989-5083EN: A Highly Accurate Mass Profiling Approach to Protein Biomarker Discovery Using HPLC-Chip/MS-Enabled ESI-TOF MS

5989-5084EN: The Agilent HPLC-Chip/6210 TOF LC/MS Enables Highly Accurate Profiling of Peptide Maps for Differential Expression Studies

Agilent Technologies