## Quantitative Analysis of Complex Peptide Mixtures Using FTMS and Differential Mass Spectrometry

### Fanyu Meng

Molecular Profiling Proteomics, Merck Research Laboratories, Merck and Co. Inc., Rahway, New Jersey, USA

Matthew C. Wiener, Jeffrey R. Sachs, Chrissina Burns, and Priyanka Verma

Applied Computer Science and Mathematics, Merck Research Laboratories, Merck and Co. Inc., Rahway, New Jersey, USA

### Cloud P. Paweletz, Matthew T. Mazur, Ekaterina G. Deyanova, Nathan A. Yates, and Ronald C. Hendrickson

Molecular Profiling Proteomics, Merck Research Laboratories, Merck and Co. Inc., Rahway, New Jersey, USA

Label-free LC-MS profiling is a powerful quantitative proteomic method to study relative peptide abundances between two or more biological samples. Here we demonstrate the use of a previously described comparative LC-MS method, differential mass spectrometry (dMS), to analyze high-resolution Fourier transform mass spectrometry (FTMS) data for detection and quantification of known peptide differences between two sets of complex mixtures. Six standard peptides were spiked into a processed plasma background at fixed ratios from 1.25:1 to 4:1 to make two sets of samples. The resulting mixtures were analyzed by microcapillary LC-FTMS and dMS. dMS successfully identified five out of the six peptides as statistically significant differences ( $p \le 0.005$ ). In this experiment, the smallest fold change reliably detected by our method was 1.5:1, and the errors of estimated ratios of concentrations were less than 20% for peptides spiked at 1.5:1 to 4:1. We conclude that LC-FTMS coupled with dMS is a useful label-free quantitative MS method that can be used to detect subtle yet statistically significant peptide differences in complex protein mixtures, including plasma samples. (J Am Soc Mass Spectrom 2007, 18, 226–233) © 2007 American Society for Mass Spectrometry

n recent years, quantitative proteomics has been used in a wide range of applications, including Lbiomarker discovery, the analysis of cellular dynamics, and biological pathway studies [1–3]. Many quantitative proteomics techniques use isotopic labeling of proteins, including metabolic labeling [4-6] or chemical labeling strategies [7–9], before mass spectrometry (MS) analysis. Methods that involve direct analyses of complex protein samples without isotope labeling (labelfree MS profiling methods) have also had some success [10–14]. Label-free methods do not require sample mixing or metabolic labeling, meaning that diverse samples, including human tissues or body fluids, can be analyzed. Lastly, label- free methods permit direct comparison of multiple samples across multiple conditions [13], enabling complex experimental designs. These characteristics make label-free mass spectrometry methods appropriate for large scale biomarker studies, e.g., plasma profiling experiments [15].

In label-free LC-MS experiments, a key step is to analyze raw LC-MS data files, and to determine the relative quantities in different samples of the observed ion species. In 2004, we reported a method termed differential mass spectrometry (dMS) [14], which can identify statistically significant differences between two or more LC-MS datasets. This method was demonstrated using standard protein digests; a 2-fold change in peptide abundance was found in data collected by an ion trap instrument (LCQ) [14]. In this application note, we describe a well-controlled experiment to evaluate the ability of dMS to analyze Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS, or FTMS) datasets. Six standard peptides were differentially spiked into processed rat plasma samples at known fixed ratios, and the resulting mixtures were profiled by LC-FTMS to measure the m/z, retention time, and ion intensity of species present in the complex

Published online October 25, 2006

Address reprint requests to Dr. R. C. Hendrickson, Molecular Profiling Proteomics, Merck Research Laboratories, Merck and Co. Inc., 126 E. Lincoln Ave., RY800-B300, Rahway, NJ 07065, USA. E-mail: ronald\_hendrickson@merck.com

Peptide (Sigma #)	Peptide name	Peptide sequence	Concentration (fmol/uL)ª	Targeted ratio (B:A)	
A9525	Angiotensin-II human	DRVYIHPF	14	2.00:1	
O3632	Osteocalcin fragments	GAPVPYPDPLEPR	13	1.50:1	
G3774	Thymopoietin II fragment 29-41	GEQRKDVYVQLYL	33	1.25:1	
C1680	Chromostatin-20	SDEDSDGDRPQASPGLGPGP	22	2.00:1	
C5900	Bovine b-casomorphin	YPFPGPI	75	1.25:1	
C6446	Chromogranin A fragment 324-337	WSKMDQLAKELTAE	12	4.00:1	

Table 1. Six peptides were spiked into a complex plasma digest at different ratios

<sup>a</sup>Concentration of each peptide for Mixture A.

mixture. dMS was used to automatically detect these peptide changes, integrate their selected ion chromatographic (SIC) peaks, and calculate their relative abundance ratios. Based on this experiment, we determined figures of merit for this label-free LC-MS profiling method, including the minimum fold change detectable and the accuracy of the relative ratio calculations.

### Experimental

### *Preparation of Mixture A (Low Concentration) and Mixture B (High Concentration)*

Rat plasma sample was purchased from Charles River Laboratory (Wilmington, MA) and N-linked glycopeptides were enriched using the method reported by Zhang@et@al.@[16,@17].@The@resulting@glycopeptide-enriched plasma pool was used as the constant peptide background in this study. Six commercially available standard peptides (Sigma Aldrich, St. Louis, MO) were spiked into the plasma digests to make Mixture A (lower concentration) and Mixture B (higher concentration)@s@hown@n@Table@.@To@ninimize@ny@ariability of sample spiking, three technical replicates were made for Mixture A and Mixture B (designated as samples A1, A2, A3 and B1, B2, B3).

### LC-FTMS Analysis of Complex Peptide Mixtures

An Agilent HP1100 capillary pump (Palo Alto, CA), a Famos autoinjector (Sunnyvale, CA), and a hybrid linear ion trap-FTMS Thermo Electron, Bremen, Germany) were used in this study.

Samples that were equivalent to ~0.1 uL of plasma were loaded onto a trap column (100 um i.d., 2.5 cm, New Objective, Woburn, MA) packed with ProteoPepII C<sub>18</sub> media through the autoinjector. After washing by 100% Solvent A (0.1 M acetic acid in H<sub>2</sub>O) for 5 min at 3 uL/min, the peptides were eluted out through a spraying column (100 um i.d., packed in-house with POROS R2 media) using a 50-min gradient from 100% Solvent A to 90% Solvent B (0.1M acetic acid in 90% acetonitrile and 10% H<sub>2</sub>O) at 1uL/min. Both columns were then equilibrated by 100% Solvent A before next injection. Samples from different conditions were interleaved (e.g., A1, B1, A2, B2, A3...) to average out any systematic chromatographic and mass spectrometric performance changes over time. Each sample was analyzed four times, generating a total of 24 high-resolution LC-FTMS data files. Key parameters for the mass spectrometer were: AGC = 1E + 6; maximum injection time 1.0 s; resolving power = 50,000; and one FTMS full scan, three data-dependant ion trap MS/MS, and one ion trap full scan were acquired sequentially (the total cycle is ~0.9 to 1 s). A nanospray source with no sheath gas was used, and the spraying voltage was 3.0 kV.

A commercial data management system, Xcalibur, was used to acquire and analyze LC-FTMS data. This software employs an automatic "background reduction" algorithm during data acquisition, which reduces most of the background data points before data storage. In this report, signal to background ratios for spiked peptides were all estimated in the following fashion. The signal is calculated by integrating selected ion chromatographic peak area (observed accurate mass,  $\pm 5$  ppm mass range). To estimate the background level, we averaged SIC peak areas for five selected mass ranges (0.5 *m*/*z* each) that appeared to be noise at the corresponding elution time.

### dMS Analysis and Ratio Determination

Differential mass spectrometry was performed as previously@eported@14].@Measured@time,@n/z,@ntensity) points were regularized onto a standard (time, m/z) grid,@vith@pacing@f 3 s in time@nd@.01@n/z units@14]. A *t*-test was performed to produce a single *p* value describing the extent to which mean intensities differ between the two conditions at each elution time and m/z. Full scan FT data from each of the 12 data files acquired for Mixtures A and B were used to determine the *p* values. Because real differences are expected to persist in (elution) time, only sufficiently long runs of statistically significant pointwise differences (here, p <0.005) were considered to be differentially expressed features (as defined by the relevant m/z and time range). Here, we required persistence over at least 18 s (seven points including both ends), allowing for a single, less significantly different point within that range. In this study, signals with m/z between 400 and 1500 and retention time between 13 and 50 min, were analyzed;



**Figure 1.** Base peak chromatograms of four LC-FTMS runs for sample A3. Base peak m/z and retention time are labeled for five selected peaks. These data were acquired at different times over a 39 h period.

this range included most of the peptidic ion species. LTQ data were analyzed by dMS in a similar fashion, except that p < 0.05 and an m/z bin width of 1.0 were used.

A peak alignment algorithm was used to compensate for peak retention time shift (retention time shifts of up to 1.5 min were allowed). Briefly, at each binned m/z, we searched the selected ion chromatograms for peaks existing in all samples within a condition. Because chromatography can change slightly from sample to sample, the common peaks will probably not all occur at exactly the same time. For example, a peak at m/z800.05 might appear at 26.9 min in the reference sample (which we choose arbitrarily), and at 27.5 min in another sample. This gives rise to a shift that will align the peak with the corresponding peak in a reference sample. In the example above, the time 27.5 would need to be shifted back 0.6 min. We assume that the actual shift required depends only on elution time in a sample and, therefore, is the same for all peaks. Nonetheless, due to estimation issues, different peaks may indicate somewhat different shifts. Smoothing the set of shifts arising from different peaks gives a "consensus" function to align a particular sample to the reference sample for the appropriate condition. Similarly, peaks existing in the mean data for each condition were used to align

the reference samples for different conditions. Applying the within-condition and then the between-condition alignments put all samples into a common time frame for analysis.

Because of the high resolving power of FTMS, each peptide can give rise to features representing multiple isotopes at multiple charge states. Based on the mass (as determined by observed charge and m/z) and peak elution time, features believed to arise from the same peptide were automatically grouped. In other controlled experiments, we found that eliminating groups containing only one feature (that is, requiring that at least two isotopes of a differentially expressed peptide to be found for it to be considered for further analysis) substantially reduced the relative frequency of false positives. Therefore, we accepted only groups with at least two features as positive peptide differences in this experiment.

For each feature, a relative ratio was automatically determined by dMS using mean SIC peak area from each condition. First, dMS finds the full peak surrounding the area of statistically significant difference based on the mean SIC from the higher-abundance condition. dMS then extracts the peak areas for both conditions and calculates the relative ratio.



**Figure 2.** Fifteen full scan FT mass spectra summed over retention time range of 20.65 min to 21.01min. Only peaks between m/z 450–605 are shown. In Total, 35 isotopic distributions can be observed in this region. Three low-intensity isotopic distributions around m/z 575 are shown in the inset (note change in scale); triangles, squares, and circles represent different sets of isotopic distribution.

### **Results and Discussion**

# Reproducible LC-FTMS Analysis of Complex Peptide Mixtures

The processed plasma background samples used in this study contain thousands of different ion species. Four base-peak chromatograms for sample A3 are shown in Figure Cl. CWe Cused Cone-dimensional Ccapillary CLC Cto separate this mixture into ~50 chromatographic peaks, each@ontaining@many@oeluting@on@pecies.@Figure@ shows a small region of a typical averaged FTMS mass spectrum (20.65 to 21.01 min, 15 scans, *m*/*z* 450 to 600 is shown), which contains 35 isotopic distributions. Three isotopic distributions can be observed between m/z573.0@nd@*n*/z 576.5@Figure@,@nset).@A@loubly-charged ion with m/z of 575.3134 (circle), and a triply-charged ion measured at m/z of 575.2690 (square) are readily resolved by FTMS, and another doubly-charged ion with m/z of 573.8317 (triangle) is also observed in this 3.5 m/z window. It is not uncommon to observe multiple components with similar m/z coeluting when analyzing complex mixtures like plasma. Importantly, the high@esolving@ower@f@TMS@18]@nakes@t@ossible@o simultaneously detect and quantify these three individual peptides rather than mistakenly integrate the signals as a single species.

Key experimental requirements in our label-free LC-FTMS method are reproducible chromatography and mass spectrometry performances. We checked the chromatographic reproducibility of our LC-FTMS platform by monitoring the retention time shift of selected peaks. Figure©1©shows©base-peak©chromatograms©for©four LC-MS analyses of sample A3 acquired at different time points over the span of 39 h, and the *m*/z values and retention time for five randomly selected abundant peaks are labeled. As shown, the overall base peak chromatograms are very similar, but the retention times for these five peaks still have shift of ~0.2 min. Based on our experience, larger experiments that involve more LC-MS analyses and run over multiple weeks can give a wider range of retention time shift using the nanoflow LC systems we employ (e.g., up to 1.2 min for ~200 samples@19].@We@herefore@ind@t@accessary@o@erform retention time correction (i.e., alignment) before performing dMS analysis. We have found that retention time shift of less than 1.5 min can be compensated for by our alignment algorithm. We have found that retention@ime@hifts@are@sually@ess@han@.5@nin.

Further, we investigated the reproducibility of MS signal intensity by calculating the coefficients of variance (CVs) for observed ion species. For example, an endogenous peptide from the rodent plasma matrix has a triply charged ion with m/z of 551.2734 (mono) and an observed retention time of 25.0 min. MS/MS analysis reveals the peptide as a haptoglobin peptide, R.ATDLKDWVQETM#AK.N (data not shown). SICs for the monoisotopic peak of this peptide (3 + , m/z)551.2734, ±5 ppm) for all 24 LC-FTMS runs were manually integrated using peak integration software in Xcalibur, and the CV for peak areas was calculated as 17.5%. Additionally, we calculated CVs for 5000 selected features (excluding those that we believe are from the spikes) using the automatic peak integration tool in dMS, and the mean and median CVs were 27 and 25%, respectively. Peak intensity variability is inherent in label-free LC-MS experiments, and this variability can arise from multiple sources, including biochemical sample processing (enzymatic digestion, reductive al-



**Figure 3.** (a) Selected ion chromatogram for a feature detected by dMS, which corresponds to the monoisotopic peak of Angiotensin-II peptide (2+ ); solid black line, averaged SIC for B samples; solid red line, averaged SIC for A samples; dashed red and black lines, one standard error about the mean for the A and B samples, respectively. The FTMS scans in the region between dotted lines were statistically different (*t*-test, p < 0.005) as detected by dMS. (b) SIC for individual samples. B samples are shown in the upper panel, and A samples are shown in the lower panel. Red portions are the region of detected significant difference.

kylation, etc.) and LC-MS profiling (e.g., variation in sample injection or electrospray condition). Had we compared data from an individual LC-MS acquisition for each condition (Mixture A and Mixture B) and used a subtraction method to look for differences in abundance, almost all of the observed ion species would have been found to have different abundances. The majority of these differences could be considered false positives because they arise due to variability in peak intensity measurements rather than from the differentially-spiked@peptides@in@this@experiment.@dMS@[14] minimizes these false positives because it considers only peptide changes that are statistically significant between two or more conditions. In other experiments, biological variability among different subjects as well as technical variability in sample preparation and MS measurement must be taken into consideration in the statistical analysis of the MS signals.

### Detection of Differentially Spiked Peptides by dMS

To identify the differentially spiked peptides, we performed dMS analysis (see the Experimental section) and detected 70 features as statistically significant differences between Mixture A (N = 12 samples) and B (N =12 samples) at a pointwise *p*-value of  $\leq 0.005$ . These features fall into 16 groups, each containing features believed to arise from a single peptide. An exemplary feature for the monoisotopic peak of Angiotensin-II (2+) is@hown@n@igure@a,@where@ed@nd@lack@olid@ines show the averaged SIC for A samples and B samples, respectively. For the mass range between *m*/*z* of 523.76 and 523.78, signal intensities in the time bins from 20.3 to@20.9@min@Figure@a,@between@lue@dotted@vertical lines) were statistically different between Mixtures A and B ( $p \le 0.005$ ). Therefore, this signal was considered a differentially-expressed feature, and the SICs of this feature@or@ach@ndividual@ample@re@hown@n@igure 3b.@or@he@Angiotensin-II@eptide,@wo@dditional@eatures from 2+ charge state (m/z 524.28 at 20.5 min and m/z 524.78 at 20.5 min) and one from 1+ charge state (m/z 1046.54 at 20.5 min) were found to be differentially expressed between the two mixtures, and all these features@vere@rouped@ogether@Table@).

Overall, dMS detected five of the six spiked peptides but missed the thymopoietin II fragment peptide, which was@spiked@in@at@a@target@ratio@of@1.25:1@(Table@2). Fourteen of the 16 detected groups of features can be attributed to spiked standards (as they were observed in a separate LC-FTMS analysis of a sample with only the spikes). These fourteen groups can be assigned as spiked@eptides@five@roups,@s@hown@n@able@),@heir adduction forms (such as oxidation or sodium adduction, six groups), or other unknown impurity ion species in the spikes (three groups). The other two groups were from plasma peptide background and they were not expected to be different. These two groups were less statistically significant than the others (ranked 14th and 16th among the 16 groups). In dMS analyses performed with a less stringent pointwise *p*-value threshold of 0.05, allowing more results to be classified as significantly different, features corresponding to the five spiked peptides remain at the top of the list, with relatively low *p* values. A manual check of the raw data in Xcalibur showed that thymopoietin II was actually present with ratio of the mean intensities nearly equal to 1:1, so it is not surprising that it was not detected even using the less stringent point wise *p* value.

A power analysis showed that with the sample size (N = 12) and p value used here, and assuming constant

m/z	Charge	Mean area (B) <sup>b</sup>	Mean area (A) <sup>b</sup>	Signal to background ratio <sup>c</sup>	Ratio (FT) <sup>d</sup>	Ratio (LT) <sup>e</sup>	Ratio (FT, Xcal) <sup>f</sup>	Relative error, FT (%) <sup>g</sup>
Chromostatin (2.00:1) <sup>a</sup>					2.26+/-0.10		2.30	13.0
651.96	3	5.07E+05	2.33E+05	28	2.18	2.32	2.26	
652.29	3	4.85E+05	2.17E+05	26	2.24			
652.63	3	2.36E+05	1.13E+05	15	2.08			
652.96	3	1.04E+05	4.67E+04	5	2.22			
653.29	3	2.94E+04	1.00E+04	_	_			
977.43	2	3.70E+05	1.58E+05	37	2.34	2.29	2.33	
977.93	2	3.49E+05	1.45E+05	36	2.42			
978.43	2	1.78E+05	7.73E+04	18	2.30			
978.94	2	6.90E+04	3.02E+04	6	2.29			
979.44	2	2.13E+04	8.62E+03	—	_			
Angiotensin-II (2.00:1)					1.91+/-0.04		1.94	-4.5
523.78	2	8.64E+05	4.64E+05	28	1.86	1.87	1.94	
524.28	2	5.85E+05	3.06E+05	15	1.91			
524.78	2	1.88E+05	9.67E+04	4	1.95			
1046.54	1	1.98E+04	9.38E+03	—	_			
Chromog	ranin A (4.	00:1)			3.31+/-0.16		3.62	-17.3
550.62	3	7.97E+05	2.34E+05	44	3.40	3.74	3.70	
550.95	3	9.17E+05	2.70E+05	38	3.39			
551.28	3	4.59E+05	1.35E+05	18	3.41			
551.62	3	1.42E+05	4.09E+04	5	3.47			
551.95	3	3.53E+04	9.33E+03	—	_			
825.43	2	9.19E+04	2.87E+04	8	3.21	3.79	3.54	
825.92	2	5.86E+04	1.78E+04	6	3.29			
826.43	2	4.62E+04	1.53E+04	2	3.01			
826.91	2	1.54E+04	5.07E+03	—	—			
Osteocalo	in (1.50:1)				1.51+/-0.04		1.60	0.7
469.91	3	8.11E+04	5.13E+04	4	1.58	_	1.71	
470.25	3	7.56E+04	4.93E+04	3	1.53			
704.38	2	3.40E+05	2.29E+05	33	1.49	_	1.49	
704.87	2	3.51E+05	2.36E+05	27	1.49			
705.37	2	1.48E+05	1.00E+05	10	1.48			
705.88	2	4.22E+04	2.79E+04	2	1.51			
Bovine b casomorphin (1.25:1)					1.64+/-0.03		1.66	31.2
790.41	1	3.69E+05	2.28E+05	36	1.62	1.76	1.66	
791.41	1	1.42E+05	8.73E+04	16	1.63			
792.42	1	5.01E+04	3.00E+04	3	1.67			

Table 2. Five spiked peptides detected by dMS and their relative ratios

<sup>a</sup>Spiked peptide name; numbers in the parenthesis are their respective theoretical ratio.

<sup>b</sup>Mean peak areas are determined automatically by dMS using mean SIC peak across each condition (12 A or B samples in this study), in arbitrary units.

 $^\circ$ Signal to background ratio for selected ion chromatographic peaks for A sample (see Experimental section for details).

<sup>d</sup>Ratio determined by dMS (Mean Area B divided by Mean Area A) for each detected feature using FTMS data. Averaged ratio across all isotopes for each peptide and the standard deviation of the measurements are displayed in bold.

<sup>e</sup>Ratio determined by dMS using LTQ data. Only four peptides were found by dMS as significant differences.

fRatio is manually determined using Xcalibur software on monoisotopic peaks in FTMS data, and averaged ratios are showed in bold (see text for details).

<sup>9</sup>Relatively error of the averaged ratio for each peptide as determined by dMS, e.g., error for chromostatin is calculated as (2.26-2.00)/2.00 = 13%, where 2.26 is the experimental ratio, and 2.00 is the theoretical (intended) ratio.

CV of 25% regardless of signal intensity, we would expect to detect a 1.5-fold difference about 90% of the time, and a 1.25-fold difference only about 20% of the time. This is consistent with what we observed in this study. Taken together, we estimate that the minimum fold change that can be reliably detected with our current experimental and computational techniques is 1.5:1. For experiments in which it is desired to detect smaller fold changes, we will probably have to increase the sample size (increase N), improve our measurement reproducibility (i.e., lower the CV), or use a less stringent pointwise p value (accepting the risk of more false positives).

### *Relative Quantitation of Differentially Spiked Peptides*

For each detected feature, dMS automatically integrated the@nean@IC@eaks@or@lixtures@@nd@@Table@).@Ve found that only SIC peaks above certain abundance level can be clearly defined and accurately integrated. In our LC-FTMS system, this minimum signal requirement is mean peak area of ~150,000 arbitrary units. Taking this into consideration, relative abundance ratios were determined for all isotopes of the five detected peptides that meet the minimum signal level requirement@Table@).

For the five detected peptide differences, peak area ratios were consistent for different isotopes and charge statesCofCeachCpeptideC(TableC2).CForCexample,Cthe calculated ratios for different isotopes of Chromostatin peptide ranged from 2.08 to 2.42, and the overall unweighted averaged ratio was 2.26, with standard deviation of 0.10 (CV 5%). In terms of the absolute value, the bovine b-casomorphin peptide (target ratio 1.25:1) had the biggest relative error at 31%, while all the other peptides had relative errors of less than 20%. Along with the fact that thymopoietin II fragment, also spiked at 1.25:1, was not detected as differentially expressed, and the error of 31% for bovine b-casomorphin suggests that the minimum detectable fold change in our LC-FTMS and dMS platform was 1.5:1. Overall, our method can measure relative ratios with less than  $\sim$ 20% errors for these small and subtle (1.5:1 to 4:1) peptide differences in the complex matrix.

For comparison purposes, we also analyzed the acquired low-resolution LTQ data. Six charge states arising from four of the spiked peptides were found as statistically significant differences. As in the FT data, the expected results, along with a few adduction forms, were the highest ranked results. The minimum detectable fold change was higher using LTQ data than using FT data: osteocalcin (targeted ratio of 1.50:1) was not found by dMS using the LTQ data. For the spiked-in peptides that were detected, the ratios calculated using LTQ data were similar to those calculated using FT data (column@Ratio@T"@n@Table@).@However,@he@FTMS data shows isotopic distributions and charge state information. In more complex mixtures, we would expect a larger number of overlapping distributions like the ones@n@igure@,@nset,@where@ising@he@igher-resolution data would be necessary for accurate quantitation. This study was designed to test our ability to detect and accurately measure small differences with ratios ranging from 1.25:1 to 4:1. In this case, it is reasonable to assume comparable or linear response for both the lower and higher abundant species. However, it is important to note that when quantifying large differences with fold changes greater that 10:1, the mass spectrometer response may deviate from linearity and introduce systematic errors into the measured ratio.

### *Errors of Ratio Measurement Introduced by dMS Peak Integration*

Many factors can contribute to errors in ratio determination, including sample preparation (such as pipetting errors), LC/MS profiling, and dMS peak integration. We focus our attention mainly on the error introduced by dMS peak integration algorithm. To evaluate these errors, we manually integrated SIC peak areas for all monoisotopic peaks of peptide spikes. The relative abundance@atios@vere@alculated@nd@isted@ndTable@ (Column "Ratio Xcal"). These ratios represent the ratio raw LC-MS data could achieve, and comparison between@veraged@atio@or@ach@peptide@Table@,Column "Ratio Xcal," bold) and its ratio determined by dMS revealed discrepancies from as low as 1% up to 10% (i.e., for chromogranin A). This suggests that up to 10% ratio determination error can come from the automated peak integration of dMS. These results suggest that improvements can be achieved by optimizing the automatic chromatographic peak area integration in dMS (work in progress).

### Conclusions

We have demonstrated that automatic analysis of LC-FTMS data by dMS is a powerful tool to analyze complex protein mixtures and detect small peptide changes between two sets of complex samples. dMS considers not only differences between conditions, but also accounts for variability within each condition, so that statistically significant differences can be efficiently found and studied. The minimum detectable fold change using our current experimental and computational systems is  $\sim$ 1.5:1, and the error of ratio measurements for these subtle changes (1.5:1 to 4:1) at relatively low signal strength (signal to background ratio 2-44) was less than 20%. Although we designed this experiment to test the ability of dMS to analyze relatively small signals, similar minimum fold change detectable and quantification accuracy have been observed for ion species with high signal levels (signal to background of up@o@50)@20].@herefore,@his@nethod@s@pplicable@or abundant signals as well as less abundant ones, which is important for analyses of complex protein samples with large dynamic ranges, e.g., plasma samples.

Unlike many other quantitative proteomic analysis methods (9, 00], @MS (Finds @ignals @hat@re@tatistically) different without relying on the identification of their amino acid sequences. In our laboratory, we find this workflow particularly useful for experiments designed to detect quantitative changes in biological systems because it separates protein identification efforts from the detection of differences. Indeed, in some cases, we have performed experiments where we detect thousands of peptide ions in the mass spectrometer but observed no significant quantitative change between conditions. In these cases we can quickly conclude that our comparative LC-MS profiling does not currently have sufficient sensitivity to detect the quantitative changes that may be present in the experiment, our platform variability is too large to detect the fold change that is present, or there are no statistically significant quantitative changes in the biological experiment. In such cases, we no longer need to perform protein ID, and efforts are then focused to redesign the experiment or explore an alternate experimental system or hypothesis. For features found to be differentially expressed, tandem mass spectra for features can be acquired in a targeted fashion by running the samples again. Because only selected peptides are targeted, the MS/MS data acquisition and data analysis efficiency will be improved. Many relatively "slower" fragmentation techniques@uch@s@lectron@apture@lissociation@ECD)@21] and@nfrared@nultiphoton@lissociation@IRMPD)@22]@an be considered, and multiple MS/MS scans can be averaged. These techniques can improve MS/MS data quality, which may lead to more confident and efficient protein identifications.

Lastly, we used peptides in this study, but it is expected that these techniques would be equally applicable to small molecules such as human metabolites. This label-free relative quantitation method could play an important role in many proteomics, metabolomics, and drug metabolism studies.

### Acknowledgments

The authors thank Dr. Haihong Zhou for helpful comments on the manuscript and Drs. Alan Sachs and Stephen Friend for their support.

### References

- 1. Aebersold, R.; Mann, M. Mass Spectrometry-Based Proteomics. *Nature* 2003, 422(6928), 198–207.
- Ideker, T.; Thorsson, V.; Ranish, J. A.; Christmas, R.; Buhler, J.; Eng, J. K.; Bumgarner, R.; Goodlett, D. R.; Aebersold, R.; Hood, L. Integrated Genomic and Proteomic Analyses of a Systematically Perturbed Metabolic Network. *Science* 2001, 292(5518), 929–934.
   Kratchmarova, I.; Blagoev, B.; Haack-Sorensen, M.; Kassen, M.; Mann,
- Kratchmarova, I.; Blagoev, B.; Haack-Sorensen, M.; Kassem, M.; Mann, M. Mechanism of Divergent Growth Factor Effects in Mesenchymal Stem Cell Differentiation. *Science* 2005, 308(5727), 1472–1477.
- Ballif, B. A.; Roux, P. P.; Gerber, S. A.; MacKeigan, J. P.; Blenis, J.; Gygi, S. P. Quantitative Phosphorylation Profiling of the ERK/p90 Ribosomal S6 Kinase-Signaling Cassette and Its Targets, the Tuberous Sclerosis Tumor Suppressors. *Proc. Natl. Acad. Sci. U.S.A.* 2005, 102(3), 667–672.
- S. Krijgsveld, J.; Ketting, R. F.; Mahmoudi, T.; Johansen, J.; Artal-Sanz, M.; Verrijzer, C. P.; Plasterk, R. H.; Heck, A. J. Metabolic Labeling of C. elegans and D. melanogaster for Quantitative Proteomics. Nat. Biotechnol. 2003, 21(8), 927–931.
- Wu, C. C.; MacCoss, M. J.; Howell, K. E.; Matthews, D. E.; Yates, J. R. III. Metabolic Labeling of Mammalian Organisms with Stable Isotopes for Quantitative Proteomic Analysis. *Anal. Chem.* 2004, 76(17), 4951–4959.
- Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags. *Nat. Biotechnol.* 1999, 17(10), 994–999.
- Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D.

J. Multiplexed Protein Quantitation in *Saccharomyces cerevisiae* Using Amine-Reactive Isobaric Tagging Reagents. *Mol. Cell. Proteom.* 2004, 3(12), 1154–1169.

- 9. Yao, X.; Freas, A.; Ramirez, J.; Demirev, P. A.; Fenselau, C. Proteolytic 180 Labeling for Comparative Proteomics: Model Studies with Two Serotypes of Adenovirus. *Anal. Chem.* **2001**, *73*(13), 2836–2842.
- Bondarenko, P. V.; Chelius, D.; Shaler, T. A. Identification and Relative Quantitation@f@rotein@Mixtures@y@nzymatic@jgestion@ollowed@y Capillary Reversed-Phase Liquid Chromatography-Tandem Mass Spectrometry. *Anal. Chem.* 2002, 74(18), 4741–4749.
   Chelius, D.; Bondarenko, P. V. Quantitative Profiling of Proteins in
- Chelius, D.; Bondarenko, P. V. Quantitative Profiling of Proteins in Complex Mixtures Using Liquid Chromatography and Mass Spectrometry. J. Proteome Res. 2002, 1(4), 317–323.
- Old, W. M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K. G.; Mendoza, A.; Sevinsky, J. R.; Resing, K. A.; Ahn, N. G. Comparison of Label-Free Methods for Quantifying Human Proteins by Shotgun Proteomics. *Mol. Cell. Proteom.* 2005, 4(10), 1487–1502.
- Wang, W.; Zhou, H.; Lin, H.; Roy, S.; Shaler, T. A.; Hill, L. R.; Norton, S.; Kumar, P.; Anderle, M.; Becker, C. H. Quantification of Proteins and Metabolites by Mass Spectrometry Without Isotopic Labeling or Spiked Standards. *Anal. Chem.* 2003, 75(18), 4818–4826.
- Wiener, M. C.; Sachs, J. R.; Deyanova, E. G.; Yates, N. A. Differential Mass Spectrometry: A Label-Free LC-MS Method for Finding Significant Differences in Complex Peptide and Protein Mixtures. *Anal. Chem.* 2004, 76(20), 6085–6096.
- Anderson, N. L.; Anderson, N. G. The Human Plasma Proteome: History, Character, and Diagnostic Prospects. *Mol. Cell. Proteom.* 2002, 1(11), 845–867.
- Zhang, H.; Li, X. J.; Martin, D. B.; Aebersold, R. Identification and Quantification of N-Linked Glycoproteins Using Hydrazide Chemistry, Stable Isotope Labeling, and Mass Spectrometry. *Nat. Biotechnol.* 2003, 21(6), 660–666.
- Zhang, H.; Yi, E. C.; Li, X. J.; Mallick, P.; Kelly-Spratt, K. S.; Masselon, C. D.; Camp, D. G.; Smith, R. D.; Kemp, C. J.; Aebersold, R. High Throughput Quantitative Analysis of Serum Proteins Using Glycopeptide Capture and Liquid Chromatography Mass Spectrometry. *Mol. Cell. Proteom.* 2005, 4(2), 144–155.
- Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* 1998, 17(1), 1–35.
- Deyanova E. G.; Zhao, X.; Meng, F.; Mazur, M. T.; Paweletz, C.; Lee, A. Y. H.; Mistry, J.; Wiener, M. C.; Sachs, J. R.; Yates, N. A.; Hendrickson, R. C. Analysis of an FTMS Based Platform for Plasma Profiling and Comparative Proteomics. *Proceedings of the 53rd ASMS Conference on Mass Spectrometry and Allied Topics*; San Antonio, TX, June 2005.
- 20. Meng, F.; Wiener, M. C.; Sachs, J. R.; Zhao, X.; Paweletz, C.; Lee, A. Y. H.; Deyanova, E. G.; Mistry, J.; Mazur, M. T.; Yates, N. A.; Hendrickson, R. C. Quantitative Protein Profiling Using FTMS and Differential Mass Spectrometry. *Proceedings of the 53rd ASMS Conference* on Mass Spectrometry and Applied Topics: San Antonio. TX, June 2005.
- on Mass Spectrometry and Applied Topics; San Antonio, TX, June 2005.
   Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. Anal. Chem. 1994, 66(18), 2809–2815.
- Biomolecule Sequencing. Anal. Chem. 1994, 66(18), 2809–2815.
  Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations. Anal. Chem. 2000, 72(3), 563–573.