ASMS 2004

Towards Global Proteomics by Analysis of Exact Mass Retention Time Pairs: A proof-of-principle of the Quantitative Ion Mapping Technology

Scott J. Geromanos[1]; Keith Richardson[1]; Phillip Young[1]; Richard Denny[1]; Marc Gorenstein[1]; Guo-Zhong Li[1]; Timothy Riley[1]; Jeffrey Silva[1]; Gregory J. Opiteck[2]; Ashok R. Dongre[2]; and Stanley A. Hefta[2]
[1]Waters Corporation, Milford, MA; [2]Bristol-Myers Squibb, Pennington, NJ

Introduction

This poster illustrates the quantitative capabilities of the new Waters® Protein Expression System. The data presented were generated using commercially available tryptically digested human serum samples spiked with different concentrations of five exogenous internal reference proteins contained in the Waters Protein Expression System Standards. Data were acquired in alternate scanning low and elevated energy accurate mass LC/MS mode. Each spiked digest sample was analyzed in triplicate. The raw data was than processed through a suite of Protein Expression System Informatics tools to produce a list of unique Exact Mass Retention-Time (EMRT) components. The data are processed further to match and compare the intensity of like EMRT components from different sample injections. Accurate mass low and elevated information associated with EMRT components that illustrate a significant change in intensity may optionally be exported to facilitate their qualitative identification.

Sample Preparation

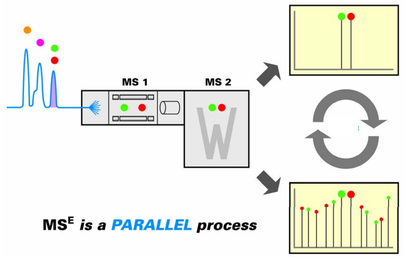
Human serum (Sigma source) was tryptically digested according to the procedure described in ***Poster TPY 458*** *(C. Dorschel et al., "Protocols to Assure Reproducible Quantitative and Qualitative Analysis of Tryptic Digests of Complex Protein Mixtures for Global Proteomic Experiments"). Five aliquots of this human serum digest were spiked with Waters Protein Expression System Standards, containing equimolar levels of Yeast Enolase and Alcohol Dehydrogenase, Rabbit Glycogen Phosphorylate, Bovine Serum Albumin and Hemoglobin tryptic digest.*

Data Collection

- *LC/MS System: Waters Protein Expression System comprised of the Waters CapLC® System with the Waters Micromass® Q-Tof Ultima API Mass Spectrometer equipped with a NanoLockSpray™ Source operated at 12,000 mass resolving power*
- *Column: Waters NanoEase™ Atlantis™ dC18 Column, 300 µm x 15 cm*
- *Mobile Phase: A = 1% Acetonitrile in Water, 0.1% Formic Acid, B = 80% Acetonitrile in Water, 0.1% Formic Acid*
- *Gradient: 6% to 40% B over 100 min. at 4.4 µL/min, followed by 10 min. rinse (99% B) and 20 min. re-equilibration at initial conditions*

*Each of the five spiked human serum (HS) digest samples was analyzed in triplicate. Injections were 5 microliters each and were made directly on-column. The total protein load for each injection was ~8.5 micrograms of Human Serum plus either 5.0, 1.0, 0.5, 0.25 or 0.1 picomoles of the spiked Protein Expression System Standards. Exact mass LC/MS data was collected using an alternating low (10eV) and elevated (28eV to 35eV) collision energy mode of acquisition such that one cycle of low and elevated collision energy data was acquired every 4.0 seconds (**Figure 1**). The NanoLockSpray source was switched every 10 seconds to obtain a reference scan of [Glu[1]]-Fibrinopeptide lockmass calibrant.*

Figure 1. *Data acquisition using the Waters Protein Expression System mass spectrometer*



Results and Discussion

All results illustrated in the Figures and Tables are based on data collected in the low collision energy portion of the low and elevated energy data acquisition cycle.

Figure 2 *panels A through O illustrate the base peak intensity (BPI) chromatograms of the triplicate analyses of each of the five spiked HS digest samples. Note that the appearance of all 15 BPI chromatograms is quite similar indicating that the spiked protein digest components represent only a minor fraction of the total peptide content of each HS digest mixture.*

Figure 3 *panels A through O illustrate the peak area integration of a non-lockmass corrected, selected ion chromatogram of an example spiked peptide (mass 978.49 eluting at approximately 77.5 minutes).*

Figure 4 *panels A through O illustrate the complexity of a single lock mass corrected spectrum extracted at the apex of each selected ion chromatogram.*

Figure 5 *panels A through O show an expanded mass range region of each of these spectra centered on m/z 978.49. The complexity of the spectra and the presence of significantly higher intensity ions indicate that the 978.49 EMRT is a minor component eluting at that point in the chromatogram. The software has extracted the 978.49 EMRT component with excellent reproducibility and its linearity in response across the various spiked injection levels indicates that ion suppression effects are minimal. The mass precision for all 15 measurements of the 978.49 EMRT is +/- 4.64 ppm.*

Figures 6 through 9 *compare data extracted from the fifteen spiked HS injections using the Protein Expression System Informatics Software.* ***Figure 6B*** *illustrates the log ratio of EMRT components from the 5.0 and 0.1 picomole spiked HS samples. The cloud of ions lying along the 45 degree axis represents EMRT peptides that were measured with an intensity ratio of approximately 1.0. These peptides are associated with HS peptides. The methodology employed allowed these ions to be measured over nearly 4 orders of magnitude in intensity. The cloud of ions lassoed in the upper left section of panel b illustrates the EMRT peptides associated with the spiked exogenous protein digests. This intensity ratio information is further illustrated in* ***Figure 6C*** *as a frequency plot. The intensity of the two EMRT groupings clearly illustrate the 1:1 ratio of the HS peptides and the 1:50 ratio of the spiked HS peptides.* ***Figure 6A*** *illustrates that the coefficient of variation of the intensity measurements made on the HS peptides was generally less than 20%.*

Figures 7 through 9 *compare 5.0 picomole vs. 0.5 picomole spiked samples, 0.50 versus 0.25 picomole spiked samples, and 0.25 versus 0.1 picomole spiked samples, respectively.*

A method for finding significant changes in peptide intensity is through the use of a paired Student t-Test. This test was employed to test the intensity of all the EMRT peptides that were observed to be in at least 2 out of 3 replicate injections in all of the spiked samples. This analysis produced a list of 63 EMRT peptides that had an intensity ratios between samples with a p-score lower than 0.01(data not shown). ***Figure 10*** *shows typical concentration plots for 6 randomly selected peptides. A quick perusal shows linearity across three orders of magnitude on peptides of varying ionization efficiencies. This list of peptides was searched against the SwissProt database using 10 ppm mass accuracy search criteria.* ***Table 2*** *illustrates that 44 of the 63 peptides (>67%) mapped back to the five exogenous proteins spiked into the HS digest samples.*

Table 2. *Peptide Mass Fingerprint Analysis of Student t-Test Results*

1. 16/63 matches (25%).
RABIT Name: Glycogen phosphorylase, muscle form (Myophosphor
Index: 463045 MW: 97290 Da pI: 6.8
Sequence
979.5407 979.5464 -5.8 835 843 1 (R)LIVAPDEKIF (L)
1117.5616 1117.5642 -2.3 415 425 0 (R)VAAAAPIQSIVDIR (L)
1177.5519 1177.5530 -0.90 471 479 0 (K)DFYELEPYHK (F)
1262.5911 1262.5945 -2.7 774 783 0 (K)VFADYTEYYK (C)
1278.5392 1278.5391 0.074 726 735 0 (R)GYNVAQEYYDR (I)
1386.7427 1386.7494 -4.8 413 427 1 (R)VAAAAPIQSIVDIRK (R)
1440.7295 1440.7375 -5.5 522 533 0 (K)LSYYIDDEAFR (D)
1442.6935 1442.6956 -1.5 279 290 0 (R)VLYPNDINFFEGK (E)
1550.7635 1550.7702 -4.3 508 520 0 (R)IGEETYSDLDGLK (K)
1566.7963 1566.7916 3.0 257 270 0 (K)DFNVIGCYIQAVLDR (N)
1580.8292 1580.8334 -2.0 741 754 0 (R)GIEGLSSGFISPK (Q)
1663.8410 1663.8300 6.6 2Met.ox 603 618 1 (R)
TVMGCIGKAAPGYHM
1840.9268 1840.9234 1.9 279 293 1 (R)VLYPNDINFFEGKEEIR (L)
1853.9938 1853.9948 0.56 492 507 0 (R)WLVCCNPDGAEIIAER (I)
1874.9087 1874.9071 0.87 372 387 0 (R)TCATYTNHTVFEAEAER (W)
1889.9995 1890.0085 -4.8 623 640 0 (K)LITAIGDVVNHDPYVGDR (L)

2. 10/63 matches (15%).
gi|30794280|ref|NP_851335.1| albumin [Bos taurus]
Index: 351651 MW: 69324 Da pI: 5.8
Sequence
1163.6200 1163.6312 -9.6 66 75 0 (K)LVNELTEFAK (T)
1291.5995 1291.6026 -2.4 300 309 0 (K)ECCDKPLLEK (S)
1479.7871 1479.7960 -6.0 421 433 0 (K)LGEYGFQNALIVR (Y)
1502.6156 1502.6143 0.85 375 388 0 (K)EYEATLEECCAK (D)
1554.6537 1554.6535 0.14 387 399 0 (K)DDPHACYSTVFDK (L)
1576.7693 1576.7681 0.75 139 151 0 (K)LKPDPNTLCDEFK (A)

3. 6/63 matches (9%).
YEAST Name: Enolase 1 (2-phosphoglycerate dehydratase) (2-ph
Index: 213440 MW: 46802 Da pI: 6.2
Sequence
807.4390 807.4365 3.1 179 185 0 (K)ITAEAUE (I)
1159.6121 1159.6111 0.82 186 195 0 (R)IGSEVYINKJK (S)
1373.6436 1373.6411 1.8 244 255 0 (K)DGICASSEFYK (D)
1416.7140 1416.7222 -5.8 16 28 0 (R)GNTPVTEVELTTK (D)
1578.7960 1578.8015 -3.5 89 103 0 (K)AVDDYLISDLTAANK (D)
1840.9268 1840.9227 2.2 33 50 0 (R)SIVPSGASTGIYHEALEMJR (D)

4. 6/63 matches (9%).
YEAST Name: Alcohol dehydrogenase I (7ADH-1)
Acc. #: 171025 Species: UNREADABLE Name: gi|171025|gb|AAA344
Index: 343993 MW: 36823 Da pI: 6.3
Sequence
1071.6357 1071.6414 -5.3 225 234 1 (K)IKDVGAVUK (A)
1136.5694 1136.5740 -4.1 9 18 0 (K)GVFPYESHQK (L)
1355.6079 1355.6088 -0.65 277 287 0 (K)CCSDVRNQVYK (S)
1386.7427 1386.7413 0.86 262 276 0 (R)IANGITIVSVGMHAGAK (C)
1447.7940 1447.8048 -7.5 320 332 0 (K)VVGLISTLPEIYEK (M)
2312.1337 2312.1482 -6.3 235 258 0 (K)ATDGGAHGVRVNVSVSEAAIEA
EAGVDTIIIVTAGEFQIGEEIRLR

35. 6/63 matches (11%).
gi|2781960|ref|NP_776342.1| hemoglobin, beta [Beta globin]
Acc. #: 122572 Species: BOVIN Name: Hemoglobin beta chain
Index: 521112 MW: 15955 Da pI: 7.0
Sequence
1328.7186 1328.7174 0.90 17 29 1 (K)VKVDEVGGEALGR (L)
1422.7242 1422.7269 -1.9 120 131 0 (K)ETPYVLQAFDFQK (V)
1448.6836 1448.6844 0.54 82 94 0 (K)GTFAALSEELHCDK (L)
2089.9682 2089.9541 6.8 40 58 0 (R)FFESFGDLSTADAVMNNPK (V)
1177.6812 1177.6812 1.21 3 15 0 (K)VVAGVANALAHR(S)
821.4074 821.4094 2.27 1 7 0 (L) MTAEEK(T)

Summary

- *The protein digestion protocols employed in this study provide quantitatively reproducible results even on complex protein mixtures as complex such as human serum.*
- *The analytical protocols employed in this study are capable of reproducibly measuring the intensity of peptides in complex protein digest mixtures over three to four orders of magnitude.*
- *Waters Protein Expression System Informatics is capable of extracting peptide accurate mass, chromatographic retention time, and intensity information from complex protein digest mixtures in a quantitatively reproducible manner.*
- *The analytical protocols employed in this study demonstrate that the combination of exact mass and chromatographic retention time can provide a very unique signature for each peptide contained in a complex protein digest mixture.*
- *Waters Protein Expression System Informatics uses peptide exact mass and retention time signature information to match and quantitatively compare the intensity of like identical peptides contained in protein digestion mixtures of a comparable control and experimental state.*

720000898EN May 2004