On the Advantage of Denoising and Peak Picking by MEND in LC-MALDI-QqTOF Analysis

Werner Ens [1], Victor Andreev [2], Oleg Krokhin [1], Tomas Rejtar [2], Hsuan-shen Chen [2], Eugene Moskovets [2], Kenneth G. Standing[1], Barry L. Karger [2]

[1] University of Manitoba, Winnipeg, Canada; [2] Northeastern University, Boston, MA 021

Introduction 1

The MEND (matched filtration with experimental noise determination) algorithm for denoising of LC-MS data and selecting precursors for MS/MS analysis has been recently implemented for LC-MALDI-TOF-TOF analysis [1]. The main advantage of MEND is the ability to minimize chemical noise in LC-MS by exploiting the difference between the patterns of noise and true signal in the chromatographic time domain. By performing filtration in the chromatographic time domain, MEND avoids distortion of peak shapes in the *m/z* domain and by minimization of noise improves accuracy of peak centroids.

In this work, MEND was used in **LC-MALDI-QqTOF** analysis for minimization of chemical noise and selection of MS/MS precursors, leading to a significant increase in the number of identifications of low abundance peptides.

MEND Algorithm 2

A typical LC-MS data set consisted of 400 spectra with 120,000 m/z data points. First, MEND determined noise characteristics from "vacant" extracted ion chromatograms (EIC) that contained no chromatographic peaks. These characteristics were then used to generate transfer functions to minimize noise in EICs for each m/z value using matched filtration. Noise characteristics were found to be *m/z* dependent, so transfer functions for matched filtration were separately determined for about 200 m/z regions. Matched filtration substantially reduced chemical and random noise with no distortion in MS peak shapes. Next, peak picking was performed based on the examination of denoised data both in the chromatographic time and m/z domains.

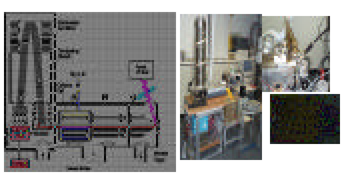
Flow Chart 3

- Analyze LC-MS data in chromatographic time dimension to select vacant EIC.
- Determine noise characteristics of LC-MS data by analyzing noise in vacant EIC.
- Determine transfer functions based on noise characteristics.
- Reduce noise in LC-MS data by performing matched filtration of EIC with transfer functions.
- Analyze noise-reduced LC-MS data in chromatographic time dimension to detect peak candidates for m/z values in mass spectral dimension.
- Calculate respective peak score values for each peak candidate.
- Analyze noise-reduced LC-MS data in mass spectral dimension to detect peaks by comparing peak score values.
- Deisotoping/deadducting

Experimental 4

A single SCX fraction of a tryptic digest of a yeast lysate was separated by RPLC and deposited on the MALDI plate as a series of 400 spots, each corresponding to 5 seconds of LC separation. MS and MS/MS analyses were performed with the Manitoba/Sciex prototype QqTOF mass spectrometer [2]. In this instrument, ions are produced by irradiation of the sample with photon pulses from a 20 Hz nitrogen laser (VCL 337ND, Spectra-Physics, Mountain View, CA) with 300 mJ energy per pulse. Orthogonal injection of ions from the quadrupole into the TOF section normally produces a mass resolving power of ~10,000 FWHM, and accuracy within a few mDa in the TOF spectra, in both the MS and MS/MS modes. Denoising of LC-MS data set and selection of precursors for MS/MS analysis were performed with MEND software. The acquired MS/MS spectra were submitted for database search and identification to SONAR (www.proteome.ca).

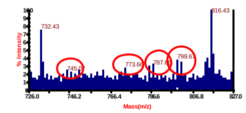
Manitoba MALDI QqTOF spectrometer 5



Results 6

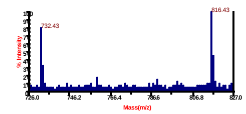
The goal of the experiment was to determine if MEND could help in the identification of low abundance peptides in the presence of high abundance ones. From a total of 2262 precursors selected by MEND, 90 were chosen with the lowest values of S/N (all with S/N<10) and submitted to MS/MS analysis. About 80% of these low S/N peaks provided good MS/MS spectra and 23 peptides (26% of those submitted to MS/MS) were identified by database search with SONAR. In order to evaluate the peak picking ability of MEND, the same LC-MS data was analyzed with the "M/Z" software that is typically used with the LC-MALDI-QqTOF Manitoba/Sciex instrument. The S/N threshold was set to the value 2.5. Nevertheless "M/Z" picked only 25 out of the 90 low S/N peaks. Out of these 25, only 5 were identified by SONAR. Thus, application of denoising and peak picking MEND software enabled a significant (23 versus 5) increase in the number of identified low abundance peptides.

Example of MEND Denoising of QqTOF Data 7



original spectrum

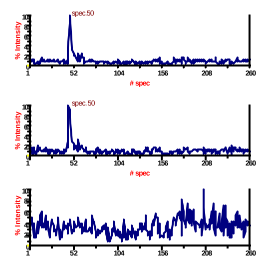
I_{max}= 932 counts



spectrum denoised by MEND

Circled MS peaks correspond to chemical noise and are minimized by MEND. For EICs see next panels.

Extracted Ion Chromatograms of Real and False Peaks 8

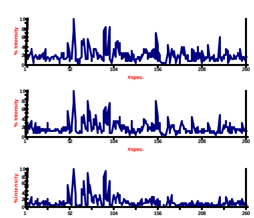


EIC for m/z= 816.43 Real peak

EIC for m/z= 732.43 Real peak

EIC for m/z= 745.07 Chemical noise filtered out by MEND

9



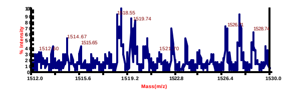
EIC for m/z= 773.66

EIC for m/z= 787.66

EIC for m/z= 799.67

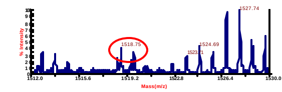
Chemical noise patterns are very similar for 3 EICs. Corresponding MS peaks are filtered out by MEND.

Low S/N Peak Selected and Identified from MEND 10



original spectrum

I_{max}= 81 counts
S/N = 3.7



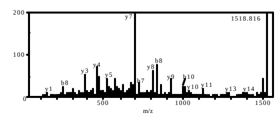
spectrum denoised by MEND

Chemical noise reduced. Overlapping peaks distinctly separated. Mass accuracy of the centroid improved.

As a result :

precursor m/z=1518.75 identified by SONAR with significant score

11



MS/MS spectrum of low abundance peptide selected by MEND

Residue	b ions		y ions	
	[illegible]	[illegible]	[illegible]	[illegible]
Glu	130.050	-	1518.817	1518.816
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Ser	330.167	-	1276.690	1276.710
[illegible]	[illegible]	[illegible]	[illegible]	
Ser	488.236	488.244	1118.621	1118.596
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Asn	715.363	715.354	918.505	918.516
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Pro	925.499	-	691.378	691.378
[illegible]	[illegible]	[illegible]	[illegible]	
Ala	1067.574	-	523.288	523.283
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Ser	1225.643	-	381.214	381.220
[illegible]	[illegible]		[illegible]	[illegible]
Lys	1500.806	-	147.113	147.110

Table: MS/MS identification of 1518.816Da peptide ELSASLNLPAAASFK from Saccharomyces cerevisiae gi6323768.

Low Abundance Peptides Selected by MEND and Identified by SONAR 12

m/z	Peptide sequence, identified protein	S/N	"M/Z" software
[illegible]	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	
1041.59	NLIQLISLK gi1586982	4.7	-
1646.79	FYDSNYPEFPVLR gi29420847	5.8	-
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	
1020.48	SDDNIESIK gi999493	7.9	-
1152.55	QVYEELESR gi6321650	8.0	-
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	
2289.05	DYSEELSELSSNGEISQDLR gi6323768	8.6	+
1027.54	QVLIRPSSK gi6322271	8.9	-
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
1366.77	VLDTGGPISVPVGR gi114575	9.7	-

Conclusions: 13

- MEND minimized both chemical and random noise in LC-MALDI-QqTOF data.
- MEND significantly (by a factor of 4.5) increased the number of identified low abundance peptides relative to standard software.

References:

1. Andreev V. P.; Rejtar T.; Chen H.-S.; Moskovets E. V.; Ivanov A. R.; Karger B. L.; Anal. Chem. 2003; **75**: 6314-6326.

2. Loboda A.V.; Krutchinsky A.N.; Bromirski M.; Ens W.; Standing K.G.; Rapid Commun. Mass Spectrom. 2000; **14**:1047-1057.