

A New Algorithm for Minimizing Chemical Noise in LC-MS: Matched Filtration with Experimental Noise Determination (MEND)

Victor P. Andreev, Tomas Rejtar, Hsuan-shen Chen,
Eugene V. Moscovets, Alex R. Ivanov and Barry L. Karger

Barnett Institute, Northeastern University, Boston, MA 02115

Noise in LC-MS is both random (white) and chemical (colored). In MALDI-MS, chemical noise results from matrix clusters, in ESI-MS from mobile phase impurities. Noise can cause either false negative or false positive identifications of sample components by masking or mimicking the signal. In addition, chemical noise can reduce mass accuracy by shifting the centroids of MS peaks. Chemical noise is more difficult to remove than white random noise because it has a pattern in the m/z domain similar to that of the signal. Our algorithm (MEND-matched filtration with experimental noise determination) is able to minimize both chemical noise and random noise by exploiting the differences in the patterns of noise and signal in the chromatographic time domain. By performing matched filtration in the chromatographic time domain MEND avoids distortion of peak shapes in m/z domain and by minimizing noise improves accuracy of peak centroids.

A typical LC-MS data set consisted of 3000 spectra with 130,000 m/z data points. First, MEND determined noise characteristics from "vacant" extracted ion chromatograms (EIC) that contained no chromatographic peaks. Noise characteristics were found to be m/z dependent (for both LC-MALDI-MS and LC-ESI-MS), so transfer functions for matched filtration were separately determined for a number of m/z regions (~200). Then, transfer functions were used for matched filtration of EICs for each m/z value. Chemical and random noise were substantially reduced with no distortion in MS peak shapes. Next, peak picking was performed based on the examination of denoised data both in chromatographic time and m/z domains.

MEND was used for denoising of data sets from LC-MALDI-MS (AB 4700 TOF/TOF) and LC-ESI-MS (AB Mariner) of tryptic digests of model protein mixtures and SCX fractions of yeast lysate. MEND minimized chemical and random noise both for LC-MALDI-MS and LC-ESI-MS and increased S/N by a factor of 5 - 8 relative to original spectra and by a factor of 2 - 3 relative to matched filtration with white noise assumption (Fig.1). By minimizing chemical noise, MEND improved the mass accuracy of peak centroids. Distances between first and second isotopes of the isotopic clusters in the denoised data set were shown to be within 1 ppm mass accuracy of the theoretically predicted value of 1.003. MEND was used for analysis of an LC-MALDI-MS data set of the tryptic digest of a cleavable ICAT labeled model mixture of 10 proteins. MEND found a higher number (by a factor of 1.3) of ICAT pairs relative to a program that only minimized white noise. This increase resulted from S/N improvement (detection of low abundant labeled peptides), from mass accuracy improvement (more precise determination of the distance between the light and heavy labeled peptides within the pair) and from removal of matrix related peaks masking the members of ICAT pairs (Fig. 2). As a result, 28 additional ICAT labeled peptides were identified. Good agreement (within 30%) between experimentally determined and expected abundances for each of 10 model proteins was found, even in the case of heavy/light ratio 10:1 and 1:10.

Advantages of MEND are: improvement in S/N due to minimization of both chemical and random noise, non-distortion of MS peaks due to denoising in chromatographic time domain, potential improvement in mass accuracy, minimization of peaks originating from matrix, increase in the number of selected ICAT pairs due to S/N and mass accuracy improvement, applicability of MEND principles to other hyphenated techniques, e.g. LC-MALDI-QTOF, CE-ESI-MS, CE-ESI-MALDI, LC(CE)-NMR, LC(CE)-diode array UV-VIS detection.

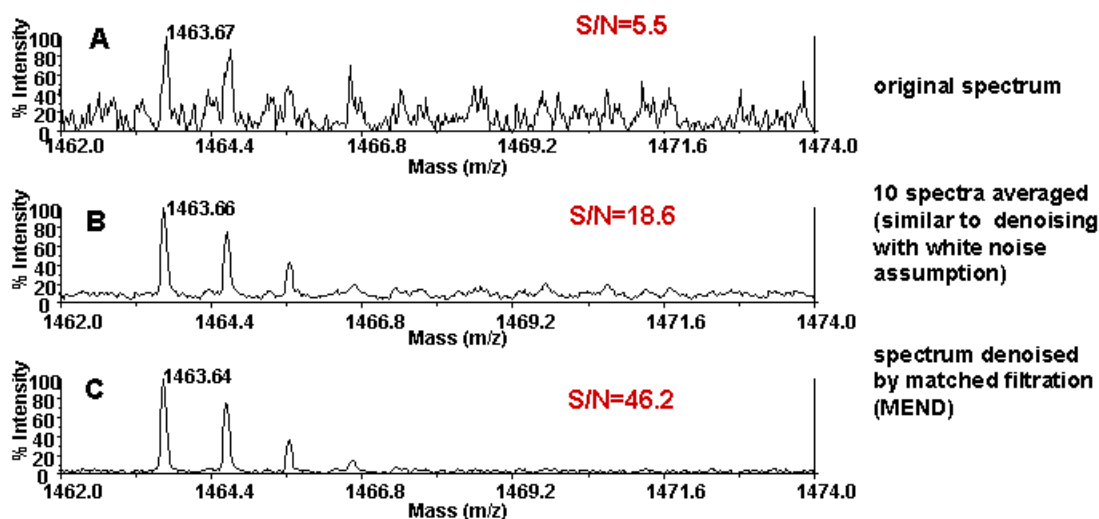


Fig 1. Chemical noise suppression by MEND in spectra from LC-MALDI-MS of complex mixture

Sample: SXC fraction of tryptic digest of yeast lysate.
Only limited spectral region is presented for clarity.

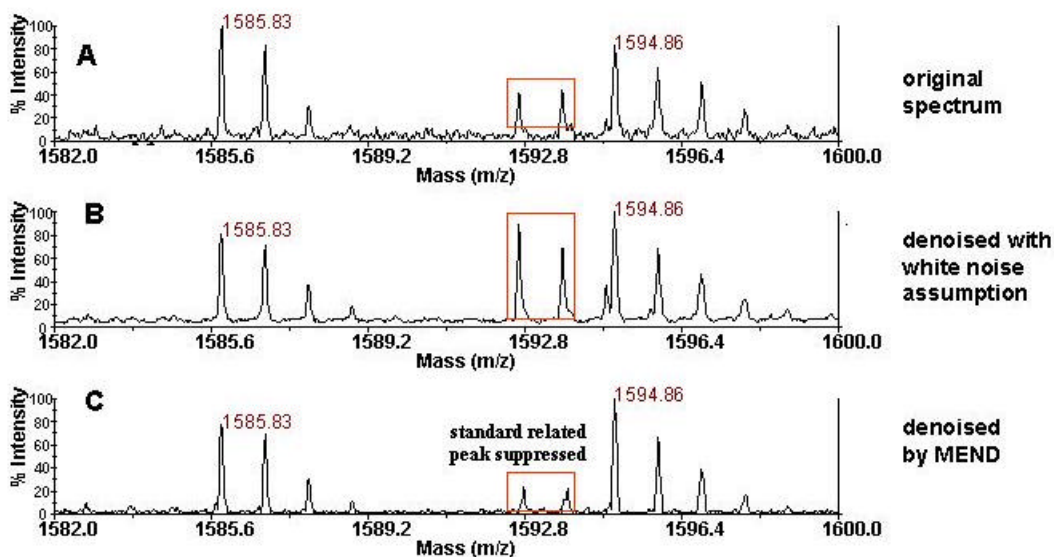


Fig. 2. ICAT pair selected by MEND due to suppression of sodium adduct of the internal standard.

Sample: tryptic digest of the mixture of 10 ICAT labeled proteins.
Spectral region of the ICAT pair of peptide SVIPSDGPSVACVK from human transferrin is presented. Light - $m/z=1585.83$, heavy - $m/z=1594.86$. Peak at $m/z=1592.8$ is a sodium adduct of internal standard peptide Glu1-Fibrinopeptide B ($m/z=1570.67$).