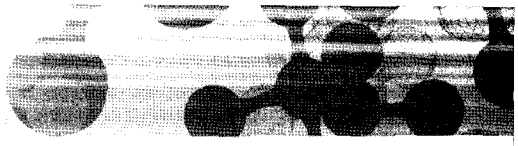

S4-1

Monoisotopic mass filtering and refinement of tandem mass spectrometric data: A step toward label-free approach for quantitative plasma proteomics

Sang-Won Lee

Department of Chemistry, Korea University, Seoul 136-701, Korea

There is tremendous interest in the use of liquid chromatography/mass spectrometry (LC/MS) as a technique to quantify protein abundances in complex proteome samples, such as plasma proteome. While quantitative proteomics holds considerable promise for elucidation of basic biology and for clinical biomarker discovery, the necessary proteomic technologies, including techniques of separation, mass spectrometry, and informatics, to achieve accurate relative/absolute quantitation are still far from their completion. Especially, the methods of treating the LC/MS data to obtain accurate peptide identification and quantitation across multiple LC/MS runs have been difficult to develop due to many technical challenges.

Here we describe new algorithms to filter and refine the tandem mass spectrometric data and to accomplish high-throughput and highly accurate "label-free"quantitative proteomics. By utilizing already developed analysis tools of mass spectrometric data, such as ICR2LS of PNNL, these algorithms generate lists of monoisotopic masses observed in LC/MS experiments, clusters the monoisotopic masses based on their masses and retention times, compare the masses with the ones in MS/MS data (i.e. in DTA files), and finally filter the MS/MS data based on the presence of corresponding monoisotopic masses and refine the precursor ion masses. This method provides not only decrease in database search time by efficiently filtering out "garbage" data, which consist of about 40% of original DTA files, but also improvement in mass measurement accuracies (average mass measurement accuracy of c.a. 4 ppm) of peptide identification. This mass spectrometric data processing method was applied to the analyses of glycoprotein enriched serum samples and is demonstrated to be a viable method for "lable-free" quantitative proteomics.