The Development and Application of an Automated MALDI-TOF Mass Spectrometric Method for the Reproducible Analysis of Large Data Sets

Anne McLachlan[1,2], Michael Borchers[2], Patrick A. Limbach[2], Prakash Velayutham[1], Michael Wagner[1]

[1]Cincinnati Children's Hospital Medical Centre, Cincinnati OH;
[2]University of Cincinnati, Cincinnati OH

Understanding the sources and magnitude of both technical (experimental) and biological (naturally occurring) variability is essential to the assessment of the statistical significance of differential expression protein patterns and for the accurate interpretation of the experimental results. A seemingly large difference in mean expression levels in two groups can be meaningless if the variability of the measurements, as determined for example by biological replicates, is of the same order of magnitude. Experimental precision is important in all analytical analyses. The determination of technical reproducibility of a method enables the number of replicates per sample required for the method to be considered reproducible, to be determined. The number of biological replicates considered for analysis should be large enough to obtain reasonable estimates for biological variability and therefore enable a valid interpretation of observed differences in sample populations.

A protocol has been developed to automate the collection of reproducible protein profiling data from large sample sets using matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry. This protocol has been applied to the analysis of proteins that are differentially expressed between challenged and controlled samples of biological fluids.

Proteomic profiling using magnetic bead technology enables the capture of different subsets of proteins/peptides

from complex biological samples. A magnetic bead-based hydrophobic interaction chromatography resin (Bruker Daltonics, Billerica MA) was used to enrich specific subsets of proteins in the biological fluid samples. Each of the samples was spotted on 5 locations on an Anchorchip target plate (Bruker Daltonics, Billerica MA). An automated MALDI-TOF mass spectrometric method was used to collect data in the linear positive mode over a mass range of 1kDa to 70kDa. The method was developed to collect data from a 384-600 µm Anchorchip plate using a series of 5 spiral rasters across each Anchorchip location, giving a total of 25 replicates per sample.

MALDI-TOF mass spectrometry is particularly suited to the analysis of biological samples as it enables the analysis of large thermally-labile compounds. It can be used to analyse heterogeneous samples and is a soft ionisation technique yielding little to no observable fragmentation. The automation of data acquisition enabled experimentation with larger data sets than was possible using manual acquisition. These larger data sets were used to test the technical reproducibility of the method, which showed good reproducibility with both a sample spotted on five locations on the Anchorchip target plate and across each anchor.

The MALDI-TOF mass spectra obtained from the purification of the analyte solutions using beads of two hydrophobicities and the analysis of data using custom bioinformatics software showed that each bead type extracted different subpopulations of proteins. Sequential extraction using two bead types provided little additional purification of the analyte solutions. The mass spectral analysis of unbound material from the initial bead extraction enabled optimum bead loading capacity to be established.

Software written in-house (Excel Visual Basic for Applications) and interfaced to the MALDI-TOF mass spectrometer enabled the tabulation of information related to experimental details and sample annotation, as well as mapping the sample locations on the Anchorchip target

plate. The inputs to this macro were MALDI spot locations on the plate along with the spot name (using a template worksheet) and sample annotation information. The macro generates two different output files. One of the output files can be fed directly as input to the MALDI-TOF instrument and the other file is used for analysing the mass spectra generated using the instrument. Custom bioinformatic software was used to compare these protein profiles and to develop dendograms of potential over and under expressed proteins.

Proteins found in the bronchoalveolar lavage (BAL) samples from mice were analysed using this protocol. The technical reproducibility of the protocol was established using pooled BAL samples from mice, and the biological variability was tested using BAL samples from 96 genetically identical mice. The samples analysed were obtained from non-exposed mice and mice exposed to the pulmonary irritant acrolein. Acrolein (2-propenal) is a potent respiratory tract irritant found in tobacco smoke and photochemical smog that can reproduce the lesions associated with Chronic Obstructive Pulmonary Disease (COPD) in experimental animals including epithelial damage, airspace enlargement, mucus hypersecretion and macrophage accumulation.

The mice lungs were lavaged with Hanks' balanced salt solution, the cells were removed by centrifugation and the supernatant retained. A magnetic bead-based C3-hydrophobic interaction chromatography resin (Bruker Daltonics, Billerica MA) was used to enrich specific subsets of proteins in the BAL samples.

The mass spectra of the pooled lung lavages from the exposed and non-exposed mice clearly showed differences in their protein profiles. The preliminary analysis of a subset of 12 of the 96 genetically identical mice showed one peak region that was differentially expressed between the exposed and the non-exposed mice. Bioinformatic analysis, however, of the MALDI-TOF data of the BAL of all 96 mice did not confirm the preliminary data, and further work is required to investigate these observations.