Protein profiling and identification in complex biological samples using LC-MALDI

by Dr S. Brand, Dr S Hahner and Dr R Ketterlinus

Offline-coupling of liquid chromatography (LC) to MALDI-mass spectrometry (LC-MALDI) is an method for the analysis of complex biological samples that is becoming more and more widely used. The increasing popularity of this technique is due to several reasons. Firstly, the approach is relatively straightforward and secondly, once fractions are spotted on MALDI-targets there are no time constraints on carrying out further analyses. This means that intelligent data-dependent MS/MS analysis can be carried out on the basis of prior MS and MS/MS analysis, resulting in significantly greater protein identification success rates. This article describes the LC-MALDI technique and illustrates its power through an example analysis involving the complex lysate of *Corynebacterium glutamicum* ATCC13032. 350 non-redundant proteins were clearly identified in the lysate.

Over the last few years mass spectrometry has developed at an incredible rate, not only in terms of the technology itself and its flexibility, but also in the diversity of the fields of scientific research to which it can be applied. The ability to combine data from different mass spectrometric techniques such as MALDI-TOF/TOF and ESI MS/MS offers the ultimate opportunity for in-depth exploration of the proteome. Advanced proteomic applications frequently generate highly complex data sets. For example, the proteomics-based identification of biomarkers requires the MS analysis and comparison of hundreds of clinical sample profiles, in combination with a MS/MS identification and validation of the differentiated components. Such applications require advanced software tools not only to process these complex data sets, but also to extract the maximum information from the combination of MS data from different MS instruments.

The prior preparation of complex biological samples such as serum, urine or cerebro-spinal fluid is also necessary for the success of any clinical research project aimed at identifying biomarkers. Such samples require pre-fractionation to simplify the MS profiles and also to remove species which could interfere with MALDI analysis. Without pre-fractionation vital sample information is lost due to ion suppression during the desorption/ionisation process, with the result that there is only a partial representation of the proteins and peptides.

2D-gel based proteomics

Complex mixtures of proteins can be separated using two-dimensional gel electrophoresis. The proteins are separated initially on the basis of their isoelectric point, and are then further fractionated according to molecular weight. Protein spots of interest are then cut from the gel and digested; the resulting peptides are extracted and prepared for mass spectrometric analysis. This entire process can be automated using the PROTEINEER system from Bruker Daltonics [1] which allows the in-depth analysis of samples even in terms of post-translational modifications (PTMs) and polymorphisms. Qualitative and quantitative information can be generated from 2D gel experiments using image analysis software such as Decodon Delta2D or Definiens Proteomweaver.

LC-based proteomics

LC-MALDI offers a promising alternative and complementary approach to the analysis of complex biological samples [Figure 1]. LC MALDI couples an upfront liquid chromatographic separation with analysis on a MALDI-TOF/TOF mass spectrometer. In a recent comparative study [2] a complex sample was separated on a 1D gel from

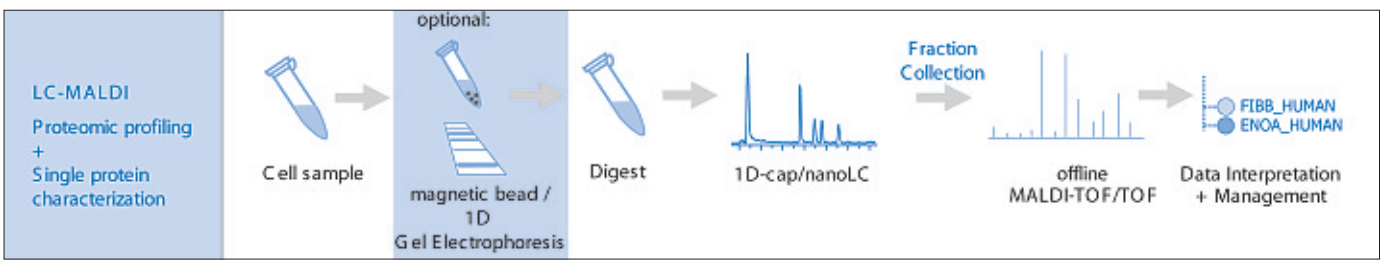

Figure 1. The LC-MALDI workflow. Reduction of sample complexity prior to MALDI-MS analysis can be performed by enzymatic digestion of a sample followed by liquid chromatography. Alternatively, the sample can be prefractionated with CLINPROT magnetic beads or by 1D-gel electrophoresis followed by enzymatic digestion. After LC separation, the fractions are collected on a MALDI-target and analysed by MALDI-TOF/TOF, and the data evaluated.

which a single gel band was digested and analysed by either MALDI or LC-MALDI. Direct MALDI analysis of the sample identified three proteins, all with poor sequence coverage. However, LC-MALDI analysis provided substantially higher sample information content as evidenced by the identification of seven proteins with dramatically improved sequence coverage, thus increasing confidence in the identified proteins.

Sample fractionation and archiving of whole LC-runs

LC separation reduces sample complexity, thus in turn reducing ion suppression in the MALDI desorption/ionisation process. Fractionation of the LC-run directly onto MALDI targets makes these fractions available for MS and MS/MS analysis "offline" from the LC-run. In contrast to "online" LC-ESI MS experiments that require instant analysis of the sample, the spotting of the sample on to the MALDI targets enables the analysis to be carried out at any suitable time. Moreover, using disposable prespotted targets such as the Prespotted AnchorChip from Bruker Daltonics, fractionated samples can be stored for several weeks for re-analysis or reference over the course of a proteomics project.

LC-MALDI workflow

Complex biological samples such as serum, plasma, tissue lysates or cell culture supernatants are digested enzymatically prior to 1D capillary or nano LC separation. After fraction collection, offline MALDI-TOF/TOF analysis is performed. Sample complexity can be further reduced by running 1D SDS PAGE, followed by enzymatic digestion of individual gel bands [2] prior to LC-MALDI analysis. A more elegant reduction of sample complexity can be achieved by using CLINPROT magnetic beads (Bruker Daltonics). These use dedicated surface functionalities such as reverse phase, cation exchange, Cu-IMAC or Protein G-immobilised antibodies for the enrichment of peptides and proteins. The eluate from CLINPROT beads can be analysed directly by MALDI MS [3]. For a more in-depth analysis, the eluate can be separated by LC prior to MALDI MS/MS analysis [4]. The PROTEINEER fc fraction collector performs the MALDI target preparation in MTP format for up to 1536 fractions on four MALDI targets, respectively. Transponder-mediated sample tracking and integration into the ProteinScape database system ensure fully automatic tracking and acquisition of spectra.

MALDI MS creates peptide mass fingerprints (PMF) in TOF mode from every fraction. WARP-LC, the intelligent core software module of the overall PROTEINEER-LC approach, determines the chromatographic compounds of the LC runs and selects fractions based on signal intensity and spectrum complexity from which MS/MS spectra will be acquired. This provides maximum qualitative and quantitative information content while minimising sample consumption and analysis time by avoiding the acquisition of redundant data.

In a final step the data are visualised, interpreted, validated and organised using the ProteinScape database system and the WARP-LC and BioTools software for protein data interpretation.

Data evaluation and validation

In a manner analogous to the representation of protein spots on a two-dimensional gel, the WARP-LC SurveyViewer provides a comprehensive overview of the peaks of all fractions as a 2D density plot, and shows their retention time, m/z value, and intensity in MS [Figure 2]. The SurveyViewer provides a rapid method for validating the quality of the LC run in terms of chromatographic peak width, calibration, contaminating species and the elution of a range of peptides during the entire run. MS and the associated MS/MS data can also be viewed simultaneously in the LC Survey Viewer.

Protein identification in complex samples

The power of the techniques can be judged from the following experiment in which a digested sample of 8 µg of a *Corynebacterium glutamicum* ATCC13032 lysate was separated on an Agilent 1100 Cap-LC using a 90 min gradient (2% - 40% acetonitrile in 0.1% TFA), and spotted in fractions of 15 sec on Prespotted AnchorChips using the PROTEINEER fc fraction collector. MALDI-TOF analysis enabled the identification of approximately 4100 compounds for subsequent LIFT-MS/MS measurement on an ultraflex II TOF/TOF mass spectrometer. Figure 2 shows a 2D density plot of all MS spectra generated

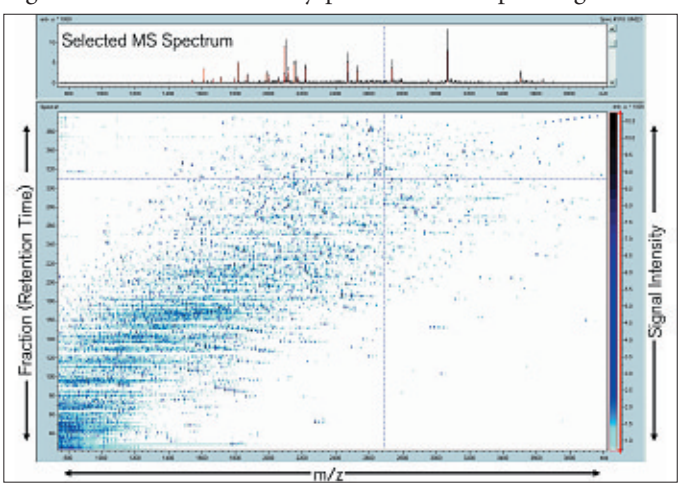

Figure 2. A 2D density plot generated by the LC SurveyViewer of a lysate from Corynebacterium glutamicum ATCC13032. The X axis represents m/Z, the Y axis the fraction number, and the fraction size is 15 sec. The colour indicates the signal intensity. In the top a MS spectrum is displayed, selected by the cross hairs in the 2D plot.

thorugh the WARP-LC SurveyViewer. The WARP-LC ProteinBrowser summarises all identified proteins and associated peptides from the database search results. Redundancies in the protein lists from the search engines can be removed by the ProteinExtractor tool.

In this experiment, 350 non-redundant proteins of *Corynebacterium glutamicum* ATCC13032 could be identified significantly by a MASCOT database search applying the MudPIT scoring algorithm. These 350 proteins cover 1675 of the 4100 analysed compounds. Figure 3 shows the MS annotation of the identified homocysteine methyltransferase and LIFT-MS/MS spectra of two peptides. In addition, the MASCOT score obtained from the fragment analysis is also indicated. Finally, ProteinScape gathers all data and archives all project-relevant parameters and results. With the ProteinScape Data Management system and the Prespotted AnchorChips, further automated searches are possible several weeks after the initial experiment.

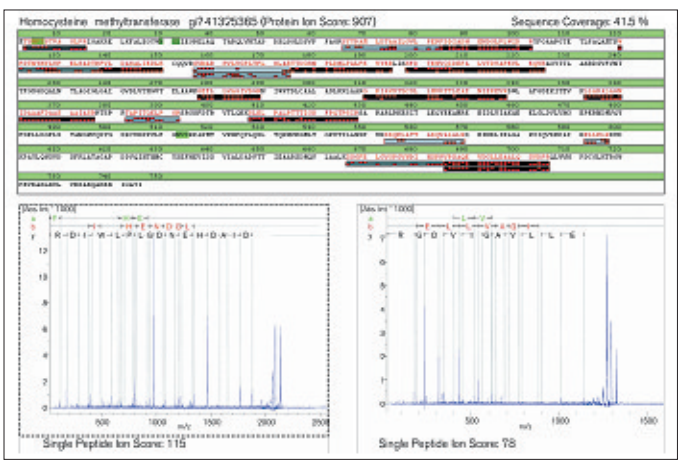

Figure 3. MS annotation of the protein homocysteine methyltransferase identified by LC-MALDI analysis in a complex lysate of Corynebacterium glutamicum ATCC13032. The grey bars represent an identified peptide, the red boxes inside the bars display the detected b- and y-ions. Additionally, the annotated LIFT-MS/MS spectra of two peptides of this protein are shown.

Biomarker detection and identification

Sample preparation with the CLINPROT magnetic bead-based system permits both the direct acquisition of linear profile spectra using the eluate from the magnetic beads for proteome profiling experiments [5] as well as further analysis by LC-MALDI from the same sample preparation. Dedicated software tools such as ClinProTools enable easy evaluation of peptide profile fingerprints for biomarker detection and validation [6]. In LC-MALDI based profiling experiments, the LC Survey Viewer 2.0 can compare two LC-runs, thus allowing excellent resolution and sensitivity for the detection of low-abundance biomarkers. The combination of LC-MALDI fractions increases the complexity of a profile spectrum by a factor of four compared to a linear profile spectrum [4], thus providing enhanced detection of low abundance putative biomarkers.

Modular and open architecture

The above experiments show how LC-MALDI technology can rapidly identify proteins and peptides, thus providing a fast and convenient way for identifying putative biomarkers that are essential for pushing back the boundaries of clinical proteomics research. LC-MALDI is a very powerful technology for protein characterisation and the profiling of complex proteomes. The modular structure of the PROTEINEER-LC system allows individual modification of workflows to suit individual and future requirements. Quantification can be achieved using stable isotopic labelling experiments (SILE) such as ICPL or ICAT, which are fully supported by the system. For smart in-depth analysis an ion trap or an API-TOF can easily be added to the system to create an intelligent feed-back driven combined LC-ESI-MALDI system [7] with all the advantages of further validation possibilities and higher detection power. Furthermore, the system allows the import of all major mass spectrometric data formats and simplifies data sharing by allowing data export in mzXML file format and in the open data format mzData. With the GelViewer Pack&Go tool, the ProteinScape system allows datasets from complete gels to be exported for subsequent browsing using a standalone viewer.

References

1. Wattenberg *et al.* Automatically identifying proteins from 2D gels using the PROTEINEER line; Bruker Daltonics Application Note 2004; MT-76.
2. Asperger A *et al.* LC-MALDI - a powerful tool for high confidence protein identification from protein mixtures; Bruker Daltonics Application Note 2004; MT-71.
3. Elssner T *et al.* Sample preparation from MALDI-TOF MS using CLINPROT magnetic beads. Bruker Daltonics Technical Note 2005; 013.
4. Niemeyer D *et al.* Higher sample information content by combining magnetic bead based fractionation and LC-MALDI-TOF/TOF on complex biological samples; DGMS Poster Presentation 2005; P73.
5. Villanueva J *et al.* Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. Anal. Chem 2004; 76: 1560-1570.
6. Ketterlinus R *et al.* Fishing for biomarkers: analyzing mass spectrometry data with the new ClinProTools software. BioTechniques 2005; 38: S37-S40.
7. Baessmann C *et al.* Coupling LC with on-line ESI and off-line MALDI-MS/MS for maximum information readout from protein samples. LCGC 2005; The Applications Book.

The authors

Sven Brand Ph.D., Stephanie Hahner Ph.D. and Ralf Ketterlinus Ph.D.
Bruker Daltonik GmbH,
Bremen, Germany
Tel.: +49 241 2205 0
E-mail: rak@bdal.de

DrugPlus Info 30530
or☑ Hot Line at www.drugplusinternational.com