



# Digging for Biomarkers: ICPLQuant - A New Software Suite For Isotope-Labeling Based Proteomics

Brunner A<sup>1</sup>, Dosch D<sup>1</sup>, Keidel E<sup>1</sup>, Wright PC<sup>2</sup>, Kellermann J<sup>1</sup>, Lottspeich F<sup>1</sup>,

<sup>1</sup> Max Planck Institute of Biochemistry, Martinsried, Germany,

<sup>2</sup> University of Sheffield, Sheffield, UK

8<sup>th</sup> Siena Meeting 2008

## Overview

### ICPLQuant – the “one button solution” for proteomics data analysis from isotopic labeling experiments

One challenging problem in our days is the growing amount of data, generated by mass spectrometers (peaklists, spectra) and MS/MS search engines that try to identify proteins via their isotopic pattern. In case of the ICPL-technology it is important to automatically detect peptide pattern on MS level and filter out regulated precursors as possible biomarker candidates.

A new software package solving ICPL multiplex protein quantification and identification problems for MALDI TOF/TOF mass spectrometry, called ICPLQuant, was developed in our lab.

Protein and peptide result tables generated by the software can be saved in a multiplet database consisting of mass spectrometry data, information about the separation space (1D-LC fraction) and peptide and protein identification characteristics. Further, generated peptide and protein result tables can be easily used for cluster analysis or statistical scatterplots.

The functionality of our software is demonstrated with a Bovine Serum Albumin (BSA) spiked E. coli cell lysate in three different concentrations (ICPL triplex labeled).

## Introduction

A very popular approach to relatively quantify and identify proteins within a mixture is a 1D-polyacrylamide gel electrophoresis (1D-PAGE) separation of the proteins followed by enzymatic cleavage and liquid chromatography (LC) in combination with tandem mass spectrometry. The 1D-gel lane can be cut into about 20 gel slices. After enzymatic cleavage (e.g. with trypsin) of the proteins in each slice, the resulting peptide fragments are further separated using LC. Containing barely a few peptides, each LC fraction is automatically spotted onto a MALDI-target for a following MS-measurement (Fig. 2).

ICPLQuant only selects **regulated peptides** for MS/MS identification, which leads to a **time - and cost-efficient strategy** that is generally applicable to quantitative proteomics projects (Fig. 1).

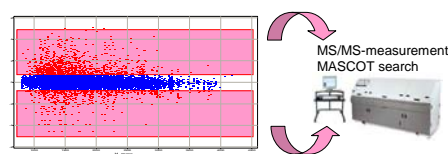


Fig. 1: Only regulated peptides will be chosen by ICPLQuant for further MS/MS analysis

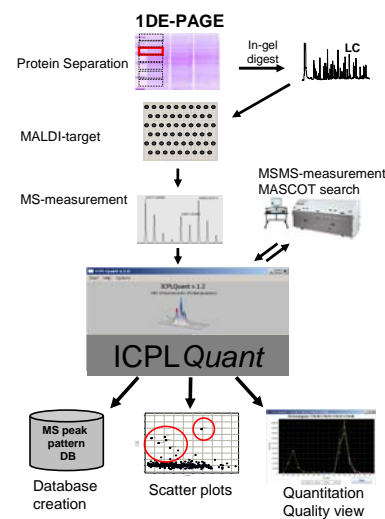


Fig. 2: Use of ICPLQuant in a proteomics experiment (connected with 1DE protein separation workflows)

## Methods

ICPLQuant consists of two logical units, **UNIT1** and **UNIT2**, and is able to **manage five bottlenecks** in an ICPL-labeled proteomic experiment:

- Project management, e.g. file and folder handling
- Quantification of ICPL-labeled peptides and proteins (**UNIT1**)
- Graphical view of quantification quality
- Automatic generation of precursor inclusion lists for targeted MSMS analysis (**UNIT1**)
- Embedding of MASCOT MS/MS results and creation of peptide and protein summary tables in excel format (**UNIT2**)

A basic workflow of the software is shown in Figure 3:

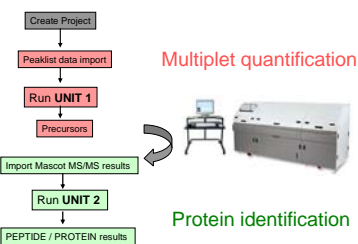


Fig. 3: Straightforward ICPLQuant workflow. Each UNIT is differently colored (UNIT1: red, UNIT2: green), indicating the different roles

## Results

For validation of ICPLQuant, three E. coli lysates (each 100 µg protein) and different amounts of Bovine Serum Albumin (at the ratio of 1 : 2 : 5) were isotopically labeled (<sup>12</sup>C, <sup>13</sup>C) and combined.

The pooled lysates were then separated on a 1D-SDS-Gel. One slice in the appropriate mass range of labeled BSA was excised and in-gel digested by trypsin. The extracted peptides were further separated into 192 fractions on a nano-RP-HPLC, and spotted directly onto a MALDI-TOF target. The fractions were then measured using MALDI-MS technique on an ABI 4700 Proteomics Analyzer.

ICPLQuant was used to detect triplet peak pattern in the generated LC-MS data, to extract precursors for a further MS/MS run and to merge MS/MS results with triplets.

Protein	observed mass (L)	expected mass (L)	Ratio (M/L)	Ratio (H/L)
BSA	922.5	922.51	1.7	5.37
BSA	1106.58	1106.57	2.26	4.58
BSA	1271.5	1271.51	2.14	5.56
BSA	1300.59	1300.59	2.15	5.05
BSA	1403.63	1403.65	1.78	3.84
			Mean: 2	4.9
			CV (%): 11.1	12.6

Fig. 4: ICPLQuant results for BSA

Identified and quantified peptide results were imported into a graphic program (e.g. Spotfire) to visualize the results. Clustering methods like hierarchical clustering or PCA can be applied to the generated data.

The cluster of all identified BSA peptides (clustered by regulation) can be seen in Figure 5:

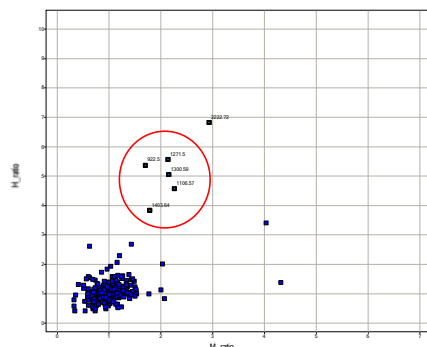


Fig. 5: Scatterplot of all MS/MS identified peptides. The plot shows the regulation of medium labeled peptides plotted against regulation of the heavy labeled peptides within one triplet

## Biomarker discovery

Within the present study we, for the first time, got the opportunity to monitor the development of a disease and thereby the changes in the protein expression pattern, using plasma of the same patient.

Five time points (Fig. 6), starting about three years before diagnosis of a colon tumor, were compared to each other by quantifying changes in the plasma proteome, using isotope coding protein labels (ICPL™, Serva), which allow the comparison of three samples in one experiment. In a preliminary study 16 patients were monitored.

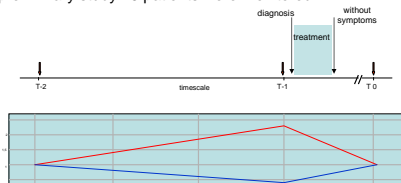


Fig. 6: Timescale for plasma samples within one patient. Blue and red profiles show expected concentration gradients of possible biomarker candidates

## Data mining strategy

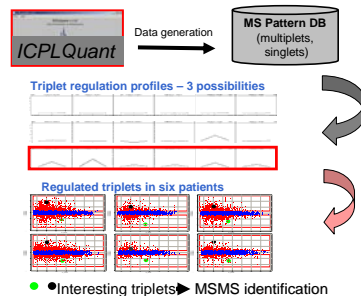


Fig. 7: Data mining strategy for biomarker discovery

Using the MS Pattern DB, it is now possible to specifically select **regulated peptides** that have the same regulation profiles in all patients (Fig. 7). Therefore it is sufficient to identify proteins in only one of the patients via MS/MS, which leads to an enormous time - and - cost reduction.

## Conclusion

The ICPLQuant software is able to handle a large number of 1D-LC/MS mass spectrometry data. The main features are:

- Data analysis adjusted to 1D-LC/MS gel slice handling and robust pattern search of ICPL-labeled peptides
- Exact quantification of ICPL-labeled peptides, accounting for chromatographic elution profiles and the isotopic effect
- Time-efficient precursor selection for MSMS peptide and protein identification.
- Exportable Protein/Peptide file formats (Excel, Text)
- ICPLQuant can be used for building up huge peak pattern databases. These databases comprise important statistical information about regulated proteins and peptides, which contributes to the detection of statistically significant biomarker candidates.
- Pattern recognition within datasets of a large number of experiments having the same regulation profile makes the software extremely useful for all proteomics studies using isotopic labeling.