

Software-supported Process Optimization for Magnetic Beads Purification of Glycoproteins

● ABRF 2008, V1-S

Peter Hufnagel¹;
Ulrike Schweiger-Hufnagel¹;
Arndt Asperger¹; Katrin Sparbier³;
Irina Kessler³; Markus Kostrzewa³;
Stefan Weise²; and Detlev Suckau¹
¹ Bruker Daltonik GmbH, Bremen, Germany
² Protagen AG, Dortmund, Germany
³ Bruker Daltonik GmbH, Leipzig, Germany

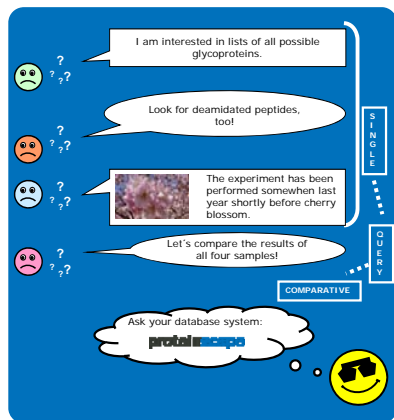
Introduction

Comprehensive proteomic analyses deal with large data sets comprising a range from very few up to thousands of MS and MS/MS spectra. Considering multiple measurements deriving from method optimization, quality control and long-term studies, this number even multiplies and necessitates a structured overview and summary about the different results. In the presented approach the analysis of human glycosylated serum proteins captured by means of different functionalized magnetic beads was performed. For the queries a software package featuring proteomics-specific queries for mass spectrometric data was used. The queries allowed investigating specific aspects with a focus on different sample preparations, certain peptides, or protein specific attributes including biological properties.

Fig. 1: "Simple" queries are applied to the data of each separation separately. Then, the query results are compared with respect to identified glycoproteins recovered from the specific type of magnetic beads.

Methods

20 µl of human serum was incubated with four kinds of glycospecific magnetic beads (ConA, LCA, WGA, BA, respectively). The beads were washed, and after elution the proteins were tryptically digested. A second magnetic bead purification was applied to enrich the glyco-peptides. The eluate was submitted to RP HPLC (Pepmap C18, LCPackings) and spotted onto a PAC target for MALDI TOF analysis on an Ultraflex III (Bruker). For each of the four kinds of magnetic beads, protein identification and result processing was done on the bioinformatics platform ProteinScope.



Results

In ProteinScope the database searches were performed for each of the four samples. All proteins fulfilling the 5 % False Positive Rate criterion were subjected to queries which were defined according to Fig. 1. Glycoproteins were searched for following two strategies:

- (1) potential glycoproteins with a NX[ST] motif in the protein sequence
- (2) proven glycoproteins, where peptides with the NX[ST] motif have been identified by MS, and which are deamidated due to de-glycosylation by PNGase F treatment.

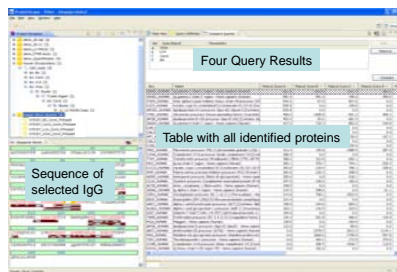
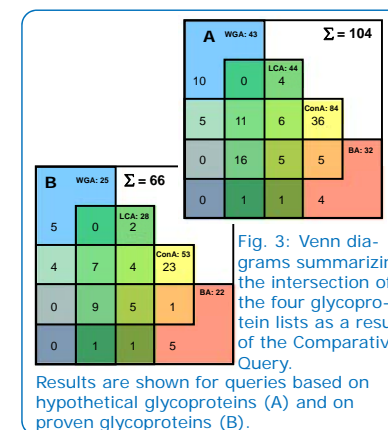


Fig. 2: ProteinScope 2 view with human glyco project open in the Navigation Tree. The table shows the query-based comparison of the glycoprotein lists of all four separations.

From the Comparative Queries a common protein list resulted. Originally, 135 proteins have been identified from all samples together. 104 of them contain the NX[ST] motif in the protein sequence, and 66 are glycoproteins confirmed by MS (PNGaseF-deamidation, Fig. 2). The intersections of the protein lists are summarized in the Venn diagrams (Fig. 3).

It becomes obvious, e.g. that the ConA sample shows up the highest number of exclusively identified proteins (23) and the highest total number of identified proteins (53).



Summary

The optimization of a complex separation workflow is often a multidimensional task. Many parameters might have to be varied. In the end, a huge number of datasets is generated and must be compared in various ways. Even the rather simple experiment presented here shows that a database-driven software platform makes life much easier. It keeps track of all data and allows the setup of simple, relevant queries.

Conclusions

- Comparative Queries allow for
 - quick and simple extraction of tailored and concise information and
 - give an excellent overview about large data amounts.
- ➔ They are an ideal tool for method optimization.