Robust Identification of Low Abundant Proteins Using a Precipitation Workflow on the ProTrap XG

Kévin Ly[1], Victoria A Miller[2], Sara Lahsaee Little[2], Jessica L Nickerson[3], Alan A Doucette[3], Jean-François Noël[1], Jean-Phillippe Couture[1], Hugo Gagnon[1]

[1] PhenoSwitch Bioscience, Sherbrooke QC Canada, [2] Allumiqs (Proteoform Scientific), Halifax NS Canada, [3]Dalhousie University, Halifax, NS Canada

ABSTRACT

Introduction
In biomarker discovery, conclusions are dependent on the quality of the data, which is contingent on the integrity of sample preparation. Efficient recovery of all proteins, regardless of abundance, is necessary to accurately represent the proteome. The ProTrap XG is a filter-based sample preparation cartridge that simplifies the traditional precipitation workflow.
Sample preparation conditions for the ProTrap XG were optimized to deliver highly efficient and robust identification of low abundance proteins from complex systems.

Methods
The effect of 2% SDS on the precipitation and recovery of 3 standard proteins, spanning 4 orders of magnitude: Beta-galactosidase (50 µg), cytochrome c (50 ng) and enolase-1 (5 ng) were combined in the presence of 2% SDS. 50 ug of the mixture was subjected to acetone precipitation followed by resolubilization and trypsin digestion in the ProTrap XG. A control (no SDS) was also digested in solution.
Bovine liver was extracted in 1% SDS and 50 µg protein was processed in the ProTrap XG.
Human plasma was diluted 100-fold, extracted in 1% SDS and processed in the ProTrap XG using acetone/MeOH precipitation or directly processed. Proteins were digested with trypsin and analysed on LC-MS/MS in data independent acquisition mode

Preliminary Data
Bottom-up analysis of the 3 protein mix on a Thermo Orbitrap Fusion Lumos with easy nLC identified all 3 proteins with equivalent sequence coverage in the SDS-based workflow relative to the control. Moreover, the SDS workflow yielded higher peptide abundance for Beta-galactosidase (4x) and cytochrome c (2x) vs. the control.
Bottom-up MS analysis of the bovine liver samples yielded on average ~8000 proteins and ~30,000 peptides. Several low abundance proteins, such as transcription factors *(e.g.* NFκB) were identified. Further data analysis is ongoing.
Bottom-up analysis of the plasma samples on a TripleTOF 6600 (Sciex) coupled to a Micro LC-200 (Eksigent) using SWATH/DIA in various acquisition strategies allowed relative quantification of more than 400 proteins and more than 10,000 peptide. The effect of protein precipitation and acquisition strategies on the quantitative dynamic range of protein abundance will be presented.

Novel Aspect
A detergent-based workflow for biological sample processing in the ProTrap XG improves the MS identification of low abundance proteins.

EXPERIMENTAL

Figure 1: β-galactosidase (50 µg), cytochrome C (50 ng), and enolase-1 (5 ng) were prepared in the presence or absence of 2% SDS in 50 mM Tris, 50 mM NaCl. The SDS was removed from the SDS-containing sample using a workflow process similar to illustrated to the left. Resulting peptides were desalted with SPE, dried and forwarded to SPARC (Toronto Canada) for MS analysis.

Figure 2 and 3. Bovine liver tissue was extracted with either 5% SDS in 50 mM Tris pH 8, or in 1% SDS in 50 mM Tris pH 8, 50 mM NaCl. Cleared protein samples were diluted to 0.5 mg/mL under 3 conditions: extracted at 5% SDS & precipitated at 5% SDS, extracted at 5% SDS & precipitated at 1% SDS or extracted at 1% SDS & precipitated at 1% SDS. Samples were processed on the ProTrap XG following a workflow similar to illustrated here. Resulting peptides were desalted with reversed-phase SPE, dried and forwarded to Bioinformatics Solutions Inc (Waterloo ON Canada) for MS analysis.

The residual SDS , normalized to starting protein concentration, was quantified using the methylene blue active substances (MBAS) assay.

Figure 4 Human plasma (40 µg) was processed as illustrated to the left, using the Bottom-Up Sample Preparation Kit and ProTrap XG Filtration Cartridges. Samples were digested with Trypsin/Lys-C (Promega V5072). Resulting peptides were desalted, dried and underwent LC-MS analysis at PhenoSwitch Bioscience.

FIGURE 2. BOVINE LIVER – PRECIPITATION CONDITIONS

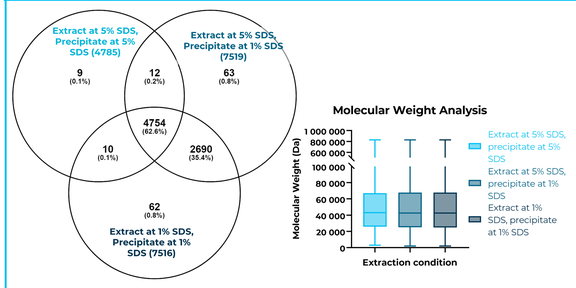

Figure 2. Three replicates at each precipitation condition were prepared and forwarded to Bioinformatics Solutions Inc in Waterloo, Canada. The samples were solubilized in loading solvent and 150 ng was injected onto a timsTOF Pro. The resulting peptides were searched at Bioinformatics Solutions Inc using PEAKS Online v1.6 against a bovine protein database. The SDS level was reduced in all cases, with samples precipitated at 1%, regardless of extraction condition resulted in more consistent sample SDS levels
Proteins found in common for each extraction/precipitation set were compared across the groups. Variability in the number of peptides found and the proteins identified for group 1- extracted and precipitated at 5% SDS was noted. Analysis of the molecular weights of the identified proteins showed a slight bias against low molecular weight proteins, when the sample was precipitated at 5% SDS, likely due to protein precipitation variability. Proteins with a molecular weight as low as 2kDa were identified across all three conditions.

TABLE 1. CELL AND PLASMA PRECIPITATION

Amount of starting material	1µg	8-10µg	40µg
Sample type	bovine oocytes	mouse CSF	human plasma
LC-MS/MS (uHPLC SWATH) length	45 min	120 min	30 min
Number of proteins quantified	1638	1284	405
Number of peptides quantified	13977	15136	10379
Average protein %CV	24%	27%	17%

Table 1: The precipitation and Digestion workflow described in the Sample Precipitation Workflow In The ProTrap XG With Bottom-up Sample Kit section was applied to different sample types with a range of initial protein contents. Samples were digested with Trypsin/Lys-C (Promega V5072). Resulting peptides were desalted, dried and underwent LC-MS analysis at PhenoSwitch Bioscience. A range of gradient lengths were used. The number of proteins and peptide quantitated was in line with expectations for the gradient length and sample type.

FIGURE 4. DYNAMIC RANGE OF PLASMA PROTEINS

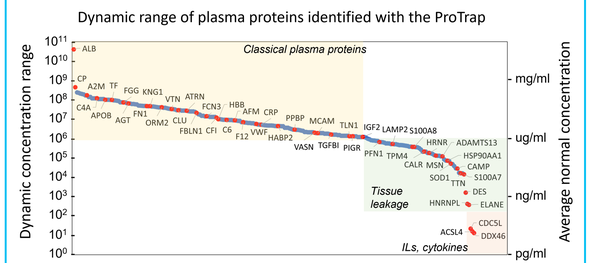

Figure 4. Graphical representation of the range of proteins identified in the plasma samples described in Table 1. The undepleted plasma was analyzed using a relatively short gradient. Proteins representing a wide dynamic range, into the ng/mL range, were identified. Average normal concentrations taken from www.proteinatlas.org.

SAMPLE PRECIPITATION WORKFLOW IN THE PROTRAP XG WITH BOTTOM-UP SAMPLE KIT

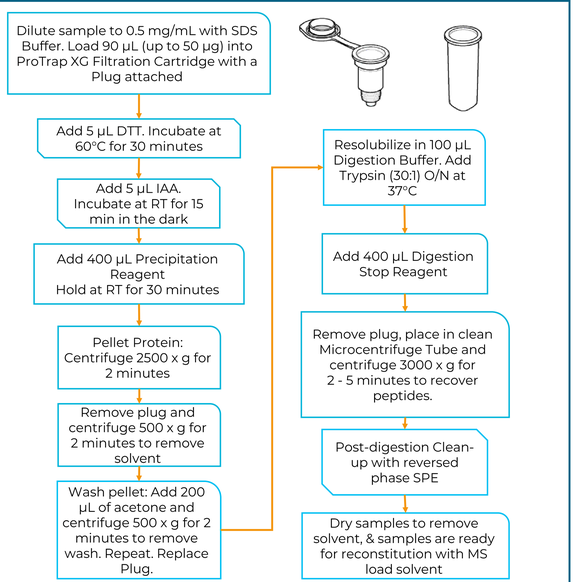

FIGURE 1. DYNAMIC RANGE

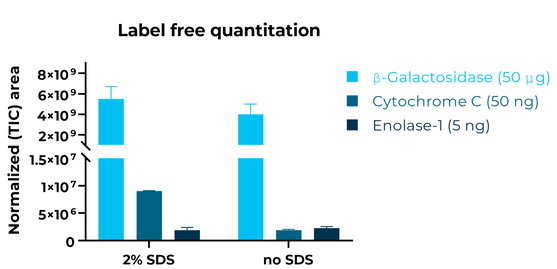

Figure 1. 150 ng of the sample was loaded. on a Orbitrap™ Fusion Lumos Tribrid mass spectrometer using an EASY-nLC. The resulting peptides were searched at BSI using PEAKS Online v1.6 against a combined database of the 3 species represented (e. coli, bovine and yeast). All three proteins were detected in approximately the proportion present in the mix, and good coverage of each protein was observed. Cytochrome C revealed a larger area in the sample precipitated in the presence of SDS indicating the possibility that the denaturing conditions of the detergent, and the contributions of the precipitation process, result in a more thorough digestion. Further replicates are planned, using a variety of proteins to determine if this phenomena is specific to this protein group. Additional experiments to cover a broader dynamic range are in process.

FIGURE 3. LOW ABUNDANCE BOVINE PROTEINS

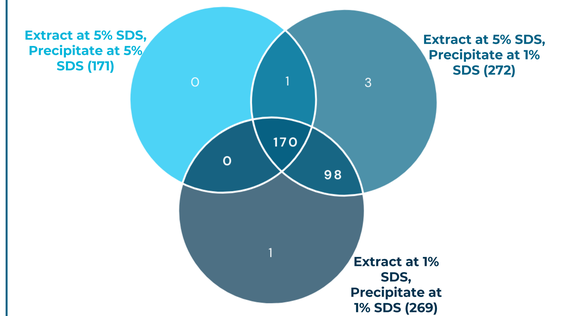

Figure 3. Proteins identified in common in each of the 3 replicates for each extraction/precipitation combination were compared against proteins listed in the abundance database PaxDb version 4.2 (https://pax-db.org/), noting that the database covered 54% of the predicted proteome size at the time of the search. Proteins assigned a normalized abundance of 1 or 2 ppb were classified as low abundance. Extraction and precipitation at 5% SDS were less likely to efficiently precipitate these low abundant proteins than extraction at 5% SDS followed by precipitation at 1% SDS or extraction and precipitation at 1% SDS.

https://pax-db.org/ Wang M, Herrmann CJ, Simonovic M, Szklarczyk D, von Mering C. Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. Proteomics. 15(18):3163-8 (2015).

CONCLUSIONS

- The ProTrap XG device can process a range of sample types and initial sample amounts, from 1 µg to 50 µg with good reproducibility.
- Precipitation is a viable method to capture low abundant proteins (initial concentration range: ng/mL and below).
- Biological samples processed using a detergent-based workflow and the ProTrap XG results in improved identification of low abundance proteins.

ACKNOWLEDGEMENTS

The authors wish to thank SPARC BioCentre (Molecular Analysis) (Toronto, ON Canada and Bioinformatics Solutions Inc (Waterloo, ON, Canada) for assistance in mass spectrometry analysis.

Interested in more information about this poster? Scan the QR code to get in touch by email after the conference.

DALHOUSIE UNIVERSITY

PhenoSwitch Bioscience
A Division of Allumiqs Corporation

Proteoform Scientific is now allumiqs