Abundant Protein Depletion and Multiplexed Protein Quantitation of Human Plasma Samples

Sergei Snovida; Katherine Herting; Ramesh Ganapathy; Ryan Bomgarden; Barbara Kaboord; Chris Etienne; John Rogers; Thermo Fisher Scientific, 3747 N Meridian Rd, Rockford, IL 61101

ABSTRACT

Purpose: The large dynamic range in protein abundance of plasma samples is the main problem associated with plasma/serum-based biomarker discovery experiments, and depletion of abundant proteins is required to identify and measure changes in prognostic or diagnostic plasma proteins. We have optimized the production and immobilization of immunoaffinity ligands to develop new top2 and top12 abundant protein depletion resins. In addition, we comprehensively evaluated the specificity, efficiency, and reproducibility abundant protein depletion from human plasma samples. Finally, we demonstrated the use of these new depletion resins to identify differences in plasma from lung cancer patients using a multiplexed protein quantitation workflow.

Methods: Relative abundance of human plasma proteins was determined by ELISA and LC-MS. LC-MS samples were analyzed on Thermo Scientific Orbitrap Fusion™ Tribrid™ mass spectrometer and processed using Thermo Scientific™ Proteome Discoverer™ 2.2 and Skyline 3.6 (University of Washington) software.

Results: After optimization of the antibody ligand conjugation chemistries and resin blending protocols, we achieved >97-99% depletion efficiency of high abundance protein targets in human plasma samples. Assessment of depletion capacity of the new resins show high depletion efficiencies, post-depletion protein yields, and reproducibility over a defined range of sample amounts. The utility of the procedure was further demonstrated using a TMT reagent-based quantitative proteomics experiment which identified and quantified nearly 1100 proteins in normal and lung cancer patient plasma samples.

INTRODUCTION

Human plasma has a broad dynamic range in protein abundance with over 10^{12} difference in concentration between the lowest and highest abundant proteins.[1] In order to discover prognostic or diagnostic plasma biomarkers, it is necessary to remove the most abundant proteins. Reproducible removal of abundant proteins is of paramount importance when profiling biological samples using mass spectrometry.

We recently developed two new depletion resin formulations, Top2 (Albumin, Ig's) and Top12 (Albumin, Ig's, transferrin, fibrinogen, alpha-1 antitrypsin, apolipoprotein A1, alpha-2 macroglobulin, alpha-1 acid glycoprotein, haptoglobin), to be used for depletion of abundant human plasma proteins in a disposable spin column format. The new columns were designed to accommodate two sample sizes, 10µL and 100µL of plasma. All bind/wash/elute conditions were optimized to yield maximum depletion of the targeted proteins and maximum recovery of the non-targeted proteins; to minimize the number of handling steps and processing time; and to maintain excellent sample-to-sample reproducibility in depletion and recovery. In this study, we focused on assessing reproducibility in terms of (a) total protein material depletion, (b) specific depletion of targeted proteins, and (c) abundance of non-targeted proteins post-depletion.

MATERIALS AND METHODS

Normal and lung cancer patient plasma was purchased from a commercial source. For abundant protein depletion,10uL of human plasma were added to the column, which was then capped and vortexed gently back into suspension. The column was then placed on rotator to mix. After 30 minutes of incubation, the column was centrifuged at 1500 x g for 2 minutes and the flow through fraction was collected into a clean 2mL microcentrifuge tube. The Top2 and the Top12 flow through was in a final volume of 110uL and 210uL, respectively. Triplicate samples were generated for each Top2 and Top12 depletions, and protein recovery was measured using BCA assay (Thermo Fisher Scientific, P/N 23225) and compared to the total protein content of undepleted plasma.

The samples were dried and resuspended in 200 uL of 3% SDS, 100 mM HEPES (pH 8.5) containing 10 mM tris(2-carboxyethyl)phosphine (TCEP) and 20 mM chloroacetamide (CAA). The samples were reduced and alkylated by heating the mixture at 95 ºC for 5 minutes. Proteins were precipitated from acidified methanol, protein pellet was washed several times with neat methanol and dissolved in 200 uL of 25 mM HEPES (pH 8.5) by sonication in a water bath. MS-grade trypsin (Thermo Fisher Scientific, P/N 90058) was added in a 1:50 enzyme-to-substrate ratio and the samples were incubated at 40 ºC for 16 hours. After digestion, the samples were desalted and peptide recoveries were measured using Pierce™ Quantitative Colorimetric Peptide Assay (Thermo Fisher Scientific, P/N 23275) and were compared to the peptide content of the undepleted samples, which were prepared in parallel as controls.

The samples were analyzed by LC-MS by performing triplicate injections of ~750 ng (for each prepared sample replicate). All samples were run on a Thermo Scientific™ Orbitrap Fusion™ Tribrid™ mass spectrometer. Liquid chromatography was performed using Thermo Scientific™ Dionex™ Ultimate™ 3000 Nano LC system, utilizing a 50 cm C18 Thermo Scientific™ EASY-Spray™ column heated at 45 ºC. Raw data was processed using Thermo Scientific™ Proteome Discoverer™ 2.2 and Skyline 3.6 (University of Washington) software.

For the TMT reagent-based quantitative experiment, five biological replicates were drawn from a common pooled normal plasma and pooled lung cancer patient plasma each, for the total of two sets of five replicate samples. Each sample was prepared as described above. After peptide-level clean-up, each sample was labeled with a unique tag of a TMT10plex isobaric reagent set. After labeling, each sample was cleaned-up to remove excess TMT reagent and buffer salts using Pierce™ Peptide Desalting Spin Columns (P/N 89852) and peptide yields were quantified using Pierce™ Quantitative Colorimetric Peptide Assay (Thermo Fisher Scientific, P/N 23275). The samples were then mixed by volume (same volume of resuspended sample solutions were used), dried in a vacuum centrifuge, and subjected to high pH reversed-phase fractionation using Pierce™ High pH Reversed-Phase Peptide Fractionation Kit. Each fraction was analyzed by LC-MS as described above.

RESULTS

Efficiency and Reproducibility of Abundant Protein Depletion

After Top2 and Top12 depletion was performed on 10 uL of human plasma, the amount of protein material recovered was measured at the protein level using BCA assay and at the peptide level using Pierce™ Quantitative Colorimetric Peptide Assay. Percent depletion of total protein material was calculated and the protein/peptide level results were compared. The results show good agreement with low %CV within each replicate set (Figure 1). This information can be used to establish what fraction of the total recovered sample can be taken for LC-MS analysis to maximize the identification numbers without overloading the column.

Figure 1. Total abundant protein depletion determined at protein and peptide levels by BCA and quantitative peptide assays, respectively.

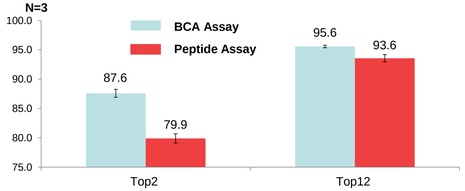
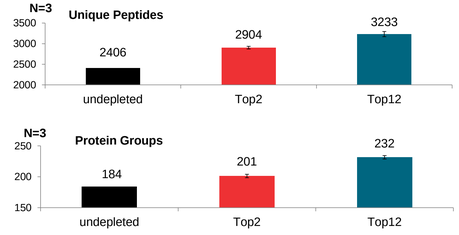


Figure 2. Effect of depletion on unique peptide and protein group identification numbers.



Efficiency and Reproducibility of Abundant Protein Depletion

LC-MS data from triplicate injections of 750 ng of each replicate sample was processed (high confidence peptides; minimum 2 unique peptides per protein) and identification numbers were compared (Figure 2). After Top2 and Top12 depletion, we obtained a 10% and 30% increase in protein group identification numbers relative to the undepleted samples, respectively.

Depletion efficiencies determined from ELISA experiments were in close agreement with the data obtained from label-free MS experiments using precursor-level peak areas of protein-representative peptides (Table 1). Total target depletion ranged from 97.4% to 100% for 99.9% for all abundant protein targets. Label-free MS data (not shown) produced CVs of <30% for abundances protein-specific peptides for both the depletion-targeted and untargeted recovered proteins.

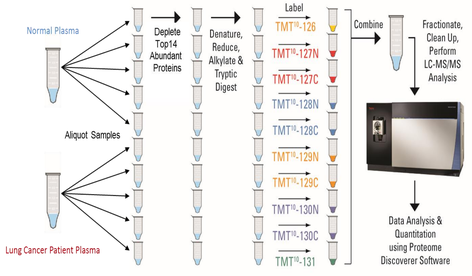
Table 1. Abundant protein depletion efficiencies measured by ELISA and label-free MS (expressed as % depletion) for selected proteins.

Protein	Top2 ELISA	Top2 MS	Top12 ELISA	Top12 MS
Albumin	99.2	99.6	99.6	99.9
IgG Kappa	99.1	99.6	99.9	99.9
IgG Lambda	99.0	99.2	99.9	99.8
Alpha-1 acid glycoprotein	N/A	N/A	99.3	99.9
Antitrypsin			100	99.9
Macroglobulin			99.9	99.8
Apolipoprotein			97.4	98.8
Fibrinogen			99.8	99.9
Haptoglobin			99.4	99.9
Transferrin			99.8	99.9

Multiplexed Quantitation of Plasma Protein Abundances

In order to identify and measure lower abundant plasma proteins, it is necessary to further reduce sample complexity using peptide fractionation in addition to abundant protein depletion. Analysis of multiple peptide fractions for multiple sample replicates significantly increases the mass spectrometry analysis time and decreases quantitative precision. To address these issues, we used TMT10plex reagents to label individual replicate samples from normal and lung cancer patient samples (Figure 3).[2]

Figure 3. Workflow for quantitative comparison of plasma proteomes with abundant protein depletion.



Samples for the TMT reagent-based quantitative experiments were prepared in parallel under identical conditions. After labeling and removal of the quenched reagent, yields of reconstituted samples were assessed using the colorimetric quantitative peptide assay. This analysis showed excellent reproducibility for the entire workflow, which eliminated the need for peptide normalization at the mixing stage prior to MS analysis (Figure 4). High reproducibility of protein quantitation was also observed in the TMT quantitation results with a majority of proteins quantified with CVs < 15% (Figure 5).

Figure 4. Reproducibility of TMT reagent-labeled peptide yields following abundant protein depletion, reduction/alkylation, digestion, clean-up, and labeling.

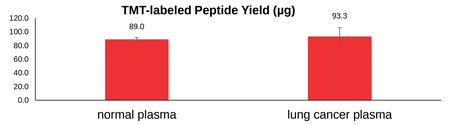


Figure 5. Distribution of %CVs of protein abundances determined using the TMT quantitation workflow.

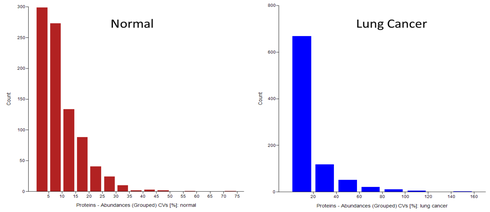
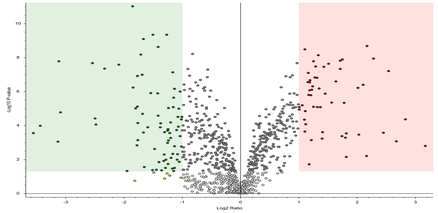


Figure 6. Volcano plot for the quantitative experiment data. Shaded areas represent a minimum two-fold change in relative protein abundance. Insets below showing examples of protein abundance changes represented by grouped abundance charts.

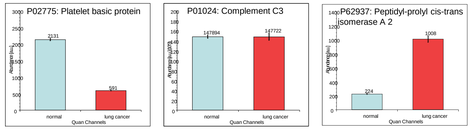


Determining Differences in Normal and Lung Cancer Plasma Proteome

A total of 14979 peptide groups and 1092 protein groups were identified from a single, three-hour gradient LC-MS analysis of the fractionated plasma peptide samples. Compared to a our label-free analysis, we were able to identify more than five times as many proteins with the TMT workflow which combined analysis combined with high pH fractionation. In addition, the precision of the quantitation with TMT reagent quantitation was much lower than our label free analysis.

Analysis of the searched data revealed clear significant differences between the two samples (Figure 6). In total, we found 56 proteins with > 2-fold increase in abundance in lung cancer patients compared to normal plasma and 72 proteins which were reduced by > 2-fold. Specific examples are shown in Figure 7. Platelet Basic Protein is an example of one protein that was found to be downregulated over 4-fold in lung cancer plasma samples. An example of a protein we found that was significantly upregulated in lung cancer plasma samples was Peptidyl-prolyl cis-trans isomerase A2.

Figure 7. Comparison of selected protein abundances in the plasma samples.



CONCLUSIONS

- Reproducible depletion of abundant proteins is obtained by using Pierce Top2 and Top12 abundant protein depletion spin columns.
- Abundant protein depletion from plasma allows for detection of more proteins in the sample enabling better detection and quantitation of relevant biomarkers.
- While reproducible depletion of the abundant proteins can be attained, MS-based label-free quantitation ultimately depends on the quality of sample preparation after the depletion, in terms of reproducibility of reduction/alkylation, digestion efficiency and peptide recovery.
- Abundant plasma protein depletion can be used effectively for quantitative comparison proteomics studies of plasma samples.

REFERENCES

1. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics. 2002 Nov;1(11):845-67.
2. Erickson B., Jedrychowski M., McAlister G., Everley R., Kunz R., Gygi S. Evaluating Multiplexed Quantitative Phosphopeptide Analysis on a Hybrid Quadrupole Mass Filter/Linear Ion Trap/Orbitrap Mass Spectrometer. 2015 Anal. Chem., 87:1241-1249.

TRADEMARKS/LICENSING

For Research Use Only. Not for use in diagnostic procedures.

Thermo Fisher SCIENTIFIC