Development of a Robust and Reproducible Global Plasma Proteome Profiling Workflow for Clinical Research

Scott Peterman[1], Emily Chen[1], David Sarracino[1], Amol Prakash[2], and Ken Miller[1]; [1]Thermo Fisher Scientific PMSC, Cambridge, MA, USA, [2]Optys Technologies, Shrewsbury, MA, USA

ABSTRACT

Purpose: Increase the routine plasma proteome profiling capabilities while maintaining 1 hour sample injection cycles and maximizing system robustness.

Methods: The developed method modified the chromatographic separation by incorporating analytical flow UHPLC pumps, a PLRP-S trapping column, and 50-cm analytical column to maximize loading and peak capacities.

Results: The alternative chromatographic method enabled significantly greater amounts of plasma digest to be evaluated reproducibly while mitigating adverse chromatographic and mass spectral performance. Over 350 proteins were routinely quantified using a 52-minute method.

INTRODUCTION

Plasma proteome profiling presents unique challenges to developing comprehensive, high-throughput analytical methods due to the highly abundant proteins that, while small in number, constitutes ca. 85% of the protein weight. Efforts continue to develop highly reproducible depletion routines in order to increase profiling the remaining plasma proteome, but depletion introduces perturbations that may inhibit accurate profiling of the original proteome as well as insert additional sample preparation steps and costs to the workflow. Performing global proteome analysis on non-depleted plasma typically results in routine quantitation for 250-320 of the most abundant proteins[1] loading 1 µg of digested plasma on column using capillary and nanoflow chromatographic separation.[2] Typically the 1 µg loading amount is limited by capacity of the capillary column.

The approach we have taken is to leverage the benefits of analytical flow rates driven by UHPLC pumps to maximize loading capacity as well as peak capacity to increase proteome coverage while maintaining 60 minute sample-to-sample analysis time. The chromatographic separation method utilized PLRP-S trapping columns and two divert valves to maximize workflow performance while maintaining instrument robustness and extend the lifetime of the analytical column.

MATERIALS AND METHODS

Sample Preparation

A single donor provided 16 different blood draws on successive days from which all samples were created. The blood tubes were prepared by resting for 1 hour, then separating the plasma using centrifugation (2000 RCM for 30 minutes) and the plasma was extracted and placed into a vial. Seven to 10 different 20 µL aliquots were extracted per vial and placed into separate 96-well plate wells for digestion for the ruggedness test. A second sample was created by pooling 30 µL aliquots from each tube, reduced, alkylated and digested using trypsin for the load study.

Test Method(s)

Two sets of experiments were performed to evaluate the workflow. The first was performed to test the loading capacity by increasing the plasma digest amounts in 2-fold increments from 7 to 220 µg on column. The pooled digest sample was used for all loading levels. All other aspects of the experimental method were kept constant. The second experiment utilizes "biological" and technical replicates from the 16 different whole blood draws. Each replicate was spiked with beta galactosidase (Sigma Aldrich) prior to digestion and the Thermo Scientific™ Pierce™ Peptide Retention Time Calibration (PRTC) Kit post-digestion.

Data Analysis

A Thermo Scientific™ Vanquish™ Horizon UHPLC system equipped with an extra divert valve in the column oven compartment was used for sample injection and chromatographic separations. A 1.4% minute^{-1} gradient consisting of a standard binary solvent system flowing at 200 µL minute^{-1} was used with a PLRP-S trapping column (Agilent Technologies) and two Thermo Scientific™ Acclaim™ 120 columns linked in series with dimensions of 500 x 2.1 mm. The trapping and analytical columns were heated to 60 ºC.

All data were acquired using a Thermo Scientific™ Orbitrap Fusion™ Lumos™ Tribrid™ mass spectrometer utilizing a DDA method with MS acquisition at a resolving power of 120,000 (@m/z 200). A loop count of 10 was used for all triggered tandem mass spectral events acquired using precursor isolation widths of 8 Da, resolution setting of 35,000, HCD collision energy setting of 35, max ion fill time of 50 msec, and dynamic exclusion width of 4 Da (centered on trigged m/z value). The 4 Da window was excluded for 6 seconds.

Data Processing

All data were processed using spectral matching using a peptide-based fraction library. A pooled plasma sample was digested, and fractionated into 48 reverse-phase high pH fractions. Each

Data Processing (continued)

fraction was analyzed using standard DDA conditions and the same chromatographic methods to create an accurate mass/accurate retention time library. The spectral matching routine used a retention time window of ±2 minutes, precursor isotopic and product ion distribution correlation coefficient tolerance of 0.9 and 0.6, respectively. All proteins and peptides were considered high quality if the measured %CV was ≤ 20%.

RESULTS

Plasma Digest Load Study

The primary goal of the experiment was to evaluate higher chromatographic flow rates, UHPLC and alternative trapping column chemistries to maximize loading and peak capacities, maintaining a hour sample injection cycle time, and significantly increasing the number of proteins and peptides commonly reported using nano- and capillary-flow chromatography. By increasing the flow rate (and sample dilution) the loading capacity has to be maximized to increase dynamic range for proteins and peptides. A stock set of plasma was used to evaluate the loading effects and a large set of samples created from a set of whole blood draws was used to evaluate the reproducibility and ruggedness.

Figure 1. Comparative analysis of protein (1A) and peptide (1B) detection and quantitation as a function of loading amount on column. The results were calculated based on the average number of proteins and peptides from the technical replicates (shown in blue with error bars) and resulting from match between runs (red). The percent values listed above each level reflects reproducibility measured across the individual replicates. All raw files were processed within a single workbook, therefore peptides detected, sampled and confidently matched at higher loading values, could still be processed by precursor quantitation. However, protein identification had to contain at least one peptide confidently matched by retention time, precursor, and product ion matching. The overlap across technical replicates is used to evaluate the mass spectral detection method to not only identify increased numbers of proteins and peptides but reproducibly sample peptides run-to-run. Clearly, increasing the loading amounts results in an increase in protein and peptide detection without compromising.

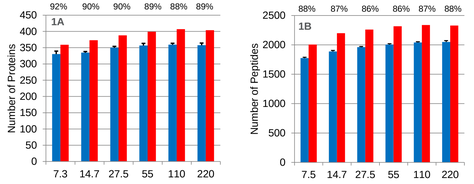


Figure 2. Evaluation of potential retention time perturbation from increasing the digest loading amounts on column. The set of ca. 2000 peptides was used to compare the retention time stability between lowest and highest loading amounts, relative to the 27.5 µg loading amount. Typically, increasing loading amounts caused peptide retention times to migrate earlier, even with trapping columns comprised of packing materials that have similar retentive strength. Incorporating the PLRP-S trap, which has a significantly lower retentive strength than the Acclaim 120 column, allowed for peptides eluted off of the trapping column to refocus onto the analytical columns to minimize the loading effects.

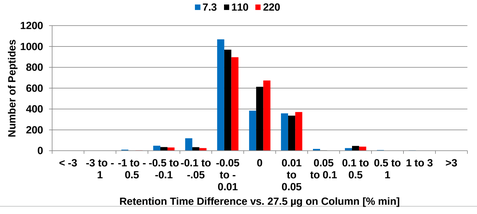


Figure 3. Comparative analysis of chromatographic peak widths as a function of loading amounts. All detected and quantified peptides measured across the entire load study were evaluated for the comparison. The chromatographic peak widths measured for the loading extremes were compared against the measured peak widths from the 27.5 µg loading amounts. The results show very little negative loading effects.

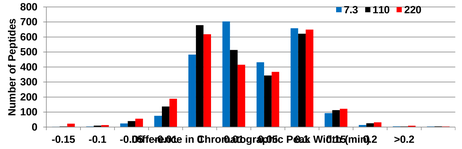


Figure 4. Evaluation of the peptides not sampled using the 7.3 µg loading amount on column compared to the higher loading amounts. The presented distribution is based on the peptide AUC measured based on the integrated precursor isotopic signal. The peptides were identified using the "match-between-runs" routine.

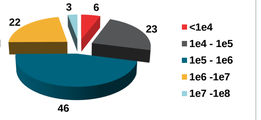
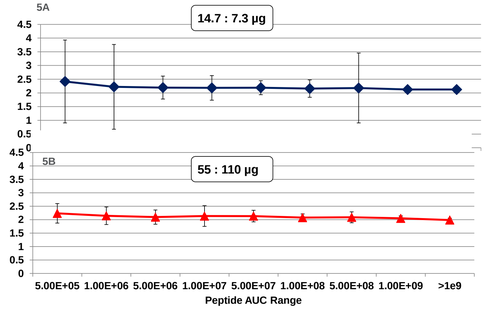


Figure 5. Comparative analysis of peptide AUC ratios between the two lowest (Figure 5A) load levels and two of the highest levels (Figure 5B). All peptides per ratio were arranged based on measured AUC values and standard deviations.



RESULTS

Stability Analysis Across Larger Number of Samples

The plasma load study was used to evaluate the performance of the trapping and analytical column to minimize loading effects as well as determine the increased detection and quantitation. While the increased plasma proteome coverage is advantageous, injecting 220 µg of plasma digest on the column and introducing it into the mass spectrometer does require periodic interruption of data acquisition for maintenance and cleaning. A final loading amount of 90 µg was used for to test the robustness and reproducibility of the method across larger numbers of plasma digest samples.

Figure 6. Evaluation of the entire workflow robustness and reproducibility based on protein level analysis. Each blood tube generated 8-10 replicates and the green portion represent proteins that were measured across all replicates to have ≤ 20% while the yellow portion represents proteins that were routinely identified based on sampled peptides, but did not meet the quantitative thresholds. A total of 490 proteins were quantified using match between runs and averaged ca. 460 proteins (94%) per blood tube. An additional 310 proteins were identified using match between runs and averaged ca. 136 proteins per tube.

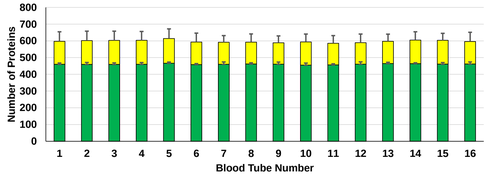


Figure 7. Similar reproducibility analysis of peptides across all samples. A total of 2402 peptides were quantified using match between runs compared to an average of 2222 peptides per tube (93%). An additional 1212 peptides were identified across all runs and an average of 553 per tube.

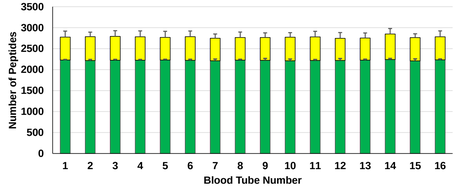


Figure 8. Comparative analysis of protein variance from grouped replicates. The samples originating from the grouped blood tubes are shown in the legend. The inset shows the measured protein AUC distribution loading 90 µg of plasma digest per sample. The AUC range for quantified proteins was measured to be in excess of 5 orders of magnitude.

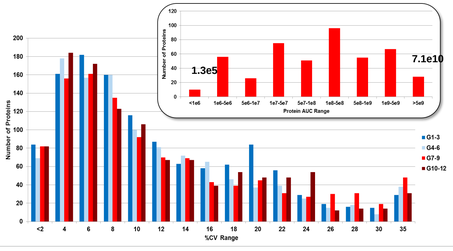


Figure 9. Distribution analysis of the measured peptide retention time across the entire study determined with and without alignment.

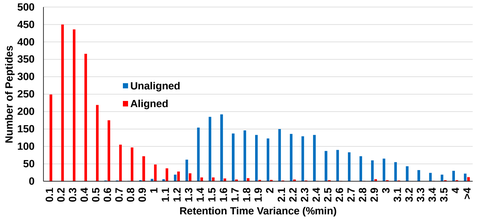
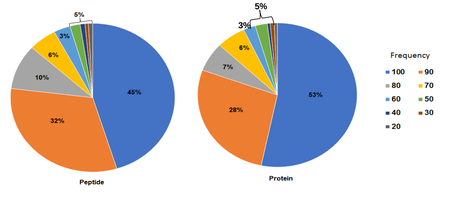


Figure 10. Frequency determination for peptides and proteins measured across the 169 plasma samples analyzed in the study. Frequency was based on confident peptide sampling per individual RAW file resulting from DDA triggering and matching against the spectral library. The protein frequency was based on how many RAW files attributed peptides were successfully matched.



CONCLUSIONS

The presentation focused on developing an experimental workflow to increase global plasma proteome profiling while maintaining high-throughput capabilities by utilizing UHPLC pumps, wide-bore columns, and alternative trapping columns. The benefits include:

- Increased loading capacity using the PLRP-S trapping column and additional divert valve.
- Maintain high peak capacity using trapping/analytical column configuration.
- Modifying comprehensive data acquisition strategies to maximize peptide sequencing efficiency.

REFERENCES (if necessary)

1. Geyer, P. E. et al. Cell Systems, 2016, 2 (3), pp. 185-195
2. Geyer, P. E. et al. Mol. Systems Biology, 2017, 13(942), pp. 1-15.

TRADEMARKS/LICENSING

ThermoFisher SCIENTIFIC