Balancing deep proteome coverage with limited sample amounts using Seer's Proteograph™ Product Suite

Shao-Yung Chen[1*], Lucy Williamson[1], Margaret Donovan[1], and Aaron S. Gajadhar[1]

[1]*Seer, Inc., Redwood City, CA 94065, USA*

Proteograph™ Product Suite enables rapid sample processing for reproducible, deep plasma proteomic analysis

Model organisms, like mice, rats, pigs, and monkeys, are utilized to unveil biological insights to human diseases. However, it is difficult to obtain 250 µL starting plasma volume required to perform the Proteograph™ Assay from small model animals. This in fact impacts the ability to conduct longitudinal sampling studies for model organisms, for example, collecting limited amounts of blood from the model organism without sacrificing the animals.

Herein, we explore the performance of the Proteograph Product Suite coupled with a state-of-the-art mass spectrometer with conventional capillary flow liquid chromatography workflows (nanoLC-MS/MS) when limited sample volumes are utilized. We evaluated Proteograph Product Suite SP100 Automation Instrument for sample processing of mouse serum with starting volumes of 250, 125, 50, 25, and 10 µL. We investigated the workflow performance using a ~30-minute LC methods coupled to Data Independent Acquisition strategies (DIA) MS and analyzed the data via our Proteograph Analysis Suite 2.1 platform, by evaluating depth of proteome coverage, dynamic range, peptide yield, and reproducibility.

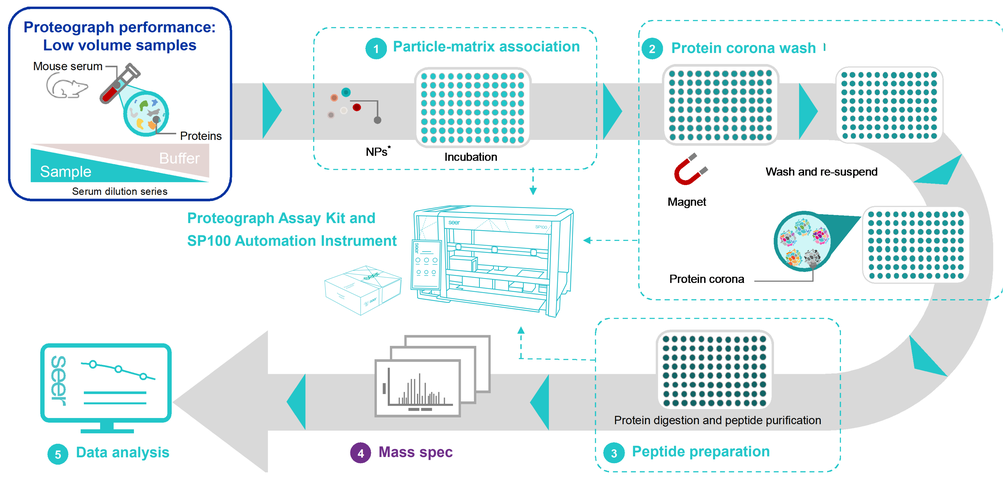

Figure 1. Schematic Proteograph Workflow for Sample Processing, MS Acquisition, and Data Analysis.
Triplicates of 250, 125, 50, 25, and 10 µL of mouse serum samples processed with the Proteograph and triplicates of traditional neat digestion workflow (10 µL) were processed in parallel. Data was acquired with Thermo Orbitrap Exploris 480 DIA LC-MS method, data was analyzed using Proteograph Analysis Suite 2.1.

Methods

Sample Preparation

The minimum sample volume required for each Proteograph assay is 250 µL. For samples with lower than this standard volume, samples were first mixed with an appropriate amount of Reconstitution Buffer A (Seer proprietary buffer) to yield a final volume of 250 µL. After samples are loaded onto the SP100 Automation Instrument, samples are ready to undergo protein corona formation and processing to purified peptides. To form the protein corona, Seer's proprietary nanoparticles are mixed with these diluted mouse serum samples in Reconstitution Buffer A and incubated at 37°C for 1 hr. Unbound proteins are removed, and downstream wash, reduction, alkylation, and protein digestion were all performed according to Seer's Proteograph Assay protocol (Figure 1).

Data Acquisition and Analysis

For Data-Independent Acquisition (DIA), 200 – 400 ng of peptides in 4 µL were reconstituted in a solution of 0.1% formic acid (FA) and 3% acetonitrile (ACN) spiked with 5 fmol/µL PepCalMix from SCIEX for constant mass MS injection between samples regardless of starting volume. Each sample was analyzed with a nanoLC system coupled with a Thermo Fisher Scientific Exploris 480 Mass Spectrometer. The mass spectrometer was operated in DIA mode using 10 m/z isolation windows from 380-1000 m/z and 3 s cycle time. MS1 scans were acquired at 60k resolution and MS2 at 30k resolution. DIA data was processed using ProteographTM Analysis Suite 2.1. Raw MS data was processed using the DIA-NN search engine (version 1.8.1) in library-free mode searching MS/MS spectra against an in silico generated spectral library of mouse protein entries (UP000000589_10090).. Different starting volume conditions (250, 125, 50, 25, and 10 µl) were searched individually to avoid boosting of identifications for lower volume conditions.

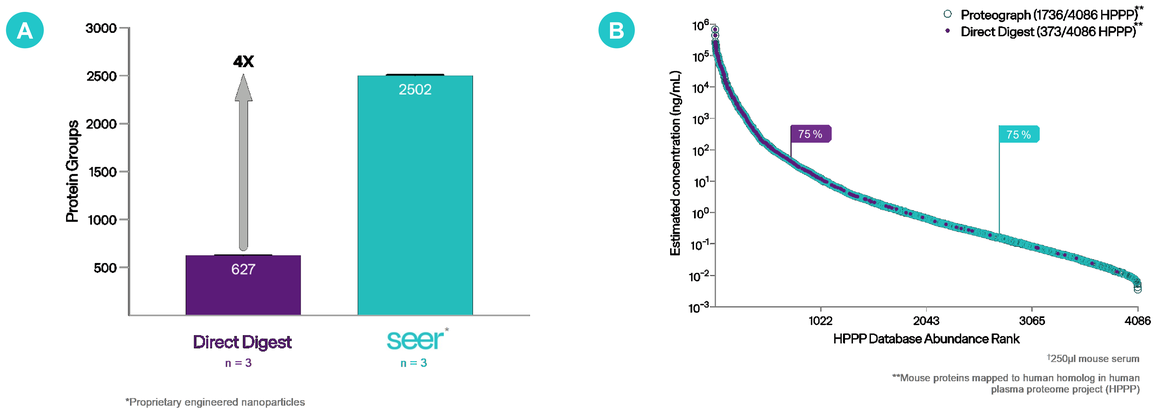

Figure 2. Performance of Proteograph vs. Traditional Direct Digestion Workflow Using Mouse Serum.
(A) Protein identifications for mouse serum samples using Proteograph with standard input volumes and direct digestion workflow, with error bars denoting standard deviation. (n= 3) **(B)** Waterfall plot mapped towards the human homolog in HPPP database showing the dynamic range and proteome coverage.

Proteograph™ workflow for limited-sample volume provides deep coverage and precision, outperforming traditional direct digest

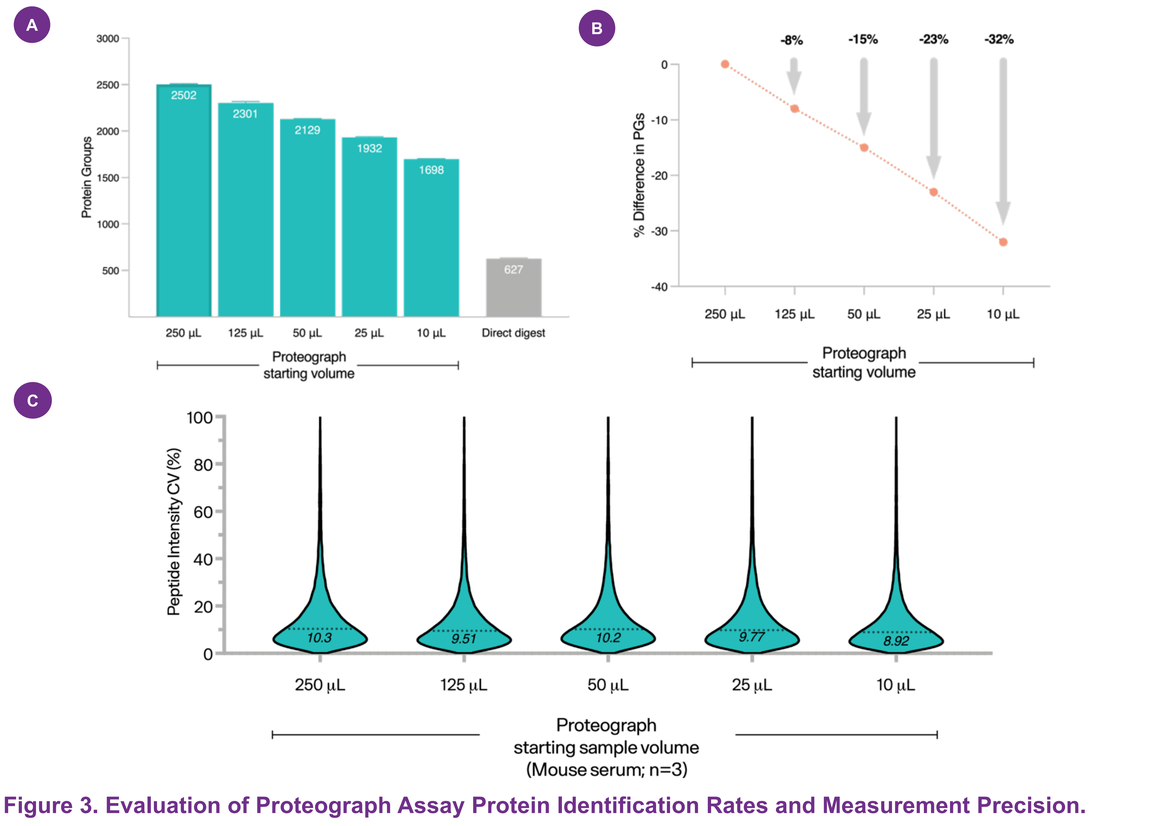

Figure 3. Evaluation of Proteograph Assay Protein Identification Rates and Measurement Precision.
A) Protein identifications for triplicates of 250, 125, 50, 25, and 10 µL of mouse serum samples with Proteograph and triplicates of traditional neat digestion workflow (10 µL). **(B)** Percentage decrease of protein identification rates (relative to standard Proteograph workflow starting volumes) when lower starting volume is used for Proteograph Assay, **(C)** Peptide intensity CV (%) plot across different Proteograph Assay within each starting sample volume (n=3) is shown, with dotted lines denoting median CV%.

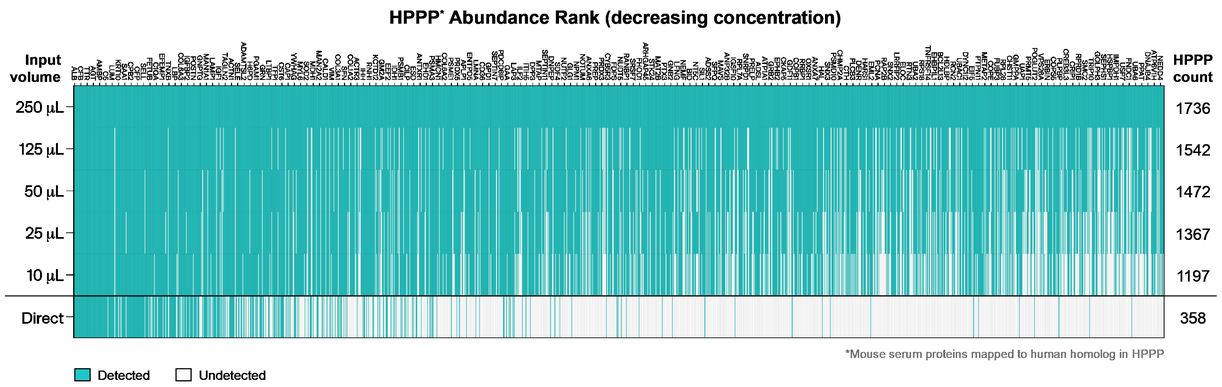

Figure 4. Evaluation of Protein Dynamic Range Coverage and Depth.
Protein identifications from 250, 125, 50, 25, 10 µL, sample volume processed with the Proteograph, and traditional direct digestion workflow were mapped towards the human homolog in the HPPP database. The protein concentrations are rank ordered in decreasing abundance from left to right. Comparison was made towards the protein groups detected in the 250 µL starting condition

Summary

In this study, we evaluated the performance of Proteograph Assay processing low volume samples (< 250 µL) using the standard Proteograph Assay protocol. Although peptide yield, depth, and protein identifications all decrease when volumes lower than 250 µL are utilized, the quantitative reproducibility is still consistent, and the median peptide intensity CV is below 10.5 % in all cases. However, the identified protein groups from all Proteograph Assays with lower starting volumes are still 2.5 X – 4 X higher than identification and depth of performance with traditional direct digestion workflow.

In summary, the Proteograph Product Suite with its current capabilities, is compatible with low volume sample studies, including model organism longitudinal sampling studies. In conclusion, this application describes a simple, straightforward dilution strategy to extend the utility of the Proteograph workflow for cases where only limited sample volumes are available, but still provides deep proteome coverage with highly reproducible quantification.

Conclusion

- Proteograph workflow performance across a range of low sample volumes outperforms direct digestion workflow (2.5 – 4X increase)
- At low, non-standard Proteograph Assay starting sample volumes, we observe a modest decrease in protein group identification (<15% drop at 50µl) relative to the standard input volume (250 µL) and higher coverage of low abundant proteins compared to direct digestion workflows
- Proteograph workflow provides deep proteome coverage for a variety of small model organisms with low volume of plasma/serum

Publications

Seer, Inc., Redwood City, CA 94065, USA | *schen@seer.bio