Deep Plasma Proteomics at Scale with Proteograph™ Product Suite: A Performance Evaluation with Label-free and TMT Multiplexing Methods

Ryan W. Benz[1]*, *Xiaoju Zhang[1], Tianyu Wang[1], Mathew Ellenberger[1], Gabriel Castro[1], Ray Schmidt[1], Veder Garcia[1], Qiu Yang[1], Chenyi Zhao[1], Purvi Tandel[1], Michelle Dubuke[1], Ramón Díaz Peña[2], Alex Campos[2], and Khatereh Motamedchaboki[1]*

[1]Seer, Inc., Redwood City, CA 94065, USA, [2]Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA 92037, USA

The Proteograph Product Suite enables rapid sample preparation for reproducible, deep plasma proteomic analysis

Proteograph sample preparation provides high levels of reproducibility for both label-free and TMT workflows

Introduction

Human blood plasma is a widely accessible sample for assessing individual health status. However, the large dynamic range of circulating proteins combined with the diversity of proteoforms present in plasma have limited the comprehensive characterization of the plasma proteome in a high throughput manner. To address such challenges, current plasma proteomics workflows combine immunodepletion of high abundance proteins, peptide fractionation and sample multiplexing approaches such as tandem mass tags (TMT). Recent advancement in sample preparation (Seer's Proteograph™ Product Suite), coupled with improved mass spectrometry instrument sensitivity and speed, enable the quantification of thousands of proteins from plasma without compromising throughput or reproducibility, creating a unique opportunity to detect robust protein biomarkers for complex diseases[1]. Here we evaluate the performance of label-free Data Dependent Acquisition (DDA) and TMT multiplexing methods with a set of control plasma samples processed with Proteograph Product Suite for deep plasma proteomic analysis.

Proteograph Product Suite

Seer core technology provides unbiased, deep, and rapid proteomics at scale

From sample to peptides, ready for analysis on most LC-MS instruments with a variety of proteomics methods including; label-free and TMT workflows

Methods

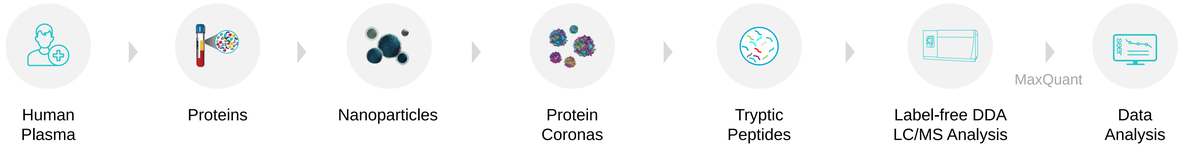
Label-Free DDA Analysis

A control pooled human plasma sample was processed with two different Proteograph automation instruments on two separate days each, resulting in a total of 4 batches (plates). Each batch contained 16 replicates of the control pooled plasma proteins enriched with 5 nanoparticles to produce 80 total wells of tryptic digested and desalted peptides for downstream LC-MS/MS analysis. Tryptic peptides were analyzed using a Thermo Fisher Scientific Orbitrap Exploris 480 Mass Spectrometer in a DDA mode with a 30-minute LC gradient (40 hours per batch of 16 plasma samples). LC-MS/MS data files were processed using MaxQuant, applying 1% FDR cutoff at the protein and peptide levels.

TMT Analysis

Two different control pooled human plasma samples were processed with Proteograph in 4 batches prepared on 4 different days. Tryptic peptides enriched with 5 nanoparticles were pooled together in one fraction and labeled with one of the TMTpro™ 16plex reagents followed by peptide fractionation (24 high pH RP fractions) and LC-MS/MS analysis on a Thermo Fisher Scientific Orbitrap Fusion Lumos Tribrid Mass Spectrometer and a FAIMS Pro Interface. LC-MS analysis were performed with a 48-hour workflow for 16 samples analysis, with 2-hr LC separation and 3 CV (compensation voltage) FAIMS peptide separation.

A Label-Free DDA Workflow



B TMTpro 16plex Workflow

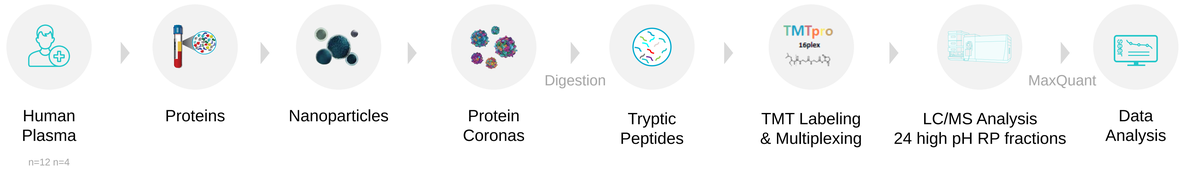


Figure 1. Label-free DDA and TMTpro 16plex workflows.

A) in the label-free DDA workflow, 16 replicates of a pooled human plasma sample were processed through the Proteograph sample prep workflow using a panel of five nanoparticles. After processing, a 96-well plate with digested peptides is ready for LCMS analysis. **B)** In the TMTpro 16-plex workflow, 4 aliquots of pooled human plasma sample (PC3) and two aliquots of pooled human plasma sample (PC5) were processed on 4 and 2 plates respectively on 4 different days on same Proteograph automation system (SP100)., Each of 4 PC3 processed peptides were aliquoted in 3 vials (total 12 samples) and each of 2 PC5 processed peptides were aliquoted in 2 vials (total 4samples), resulting in 16 samples for peptide labeling with Thermo Scientific TMTpro 16plex reagents. Labeled peptides were pooled into a single multiplexed pooled sample for LCMS analysis. Peptides were fractionated by high pH RP chromatographic separation into 24 fractions and each fraction was analyzed with a 3 CV FAIMS, MS2 method for TMT LCMS analysis.

Label-Free DDA Data & Results

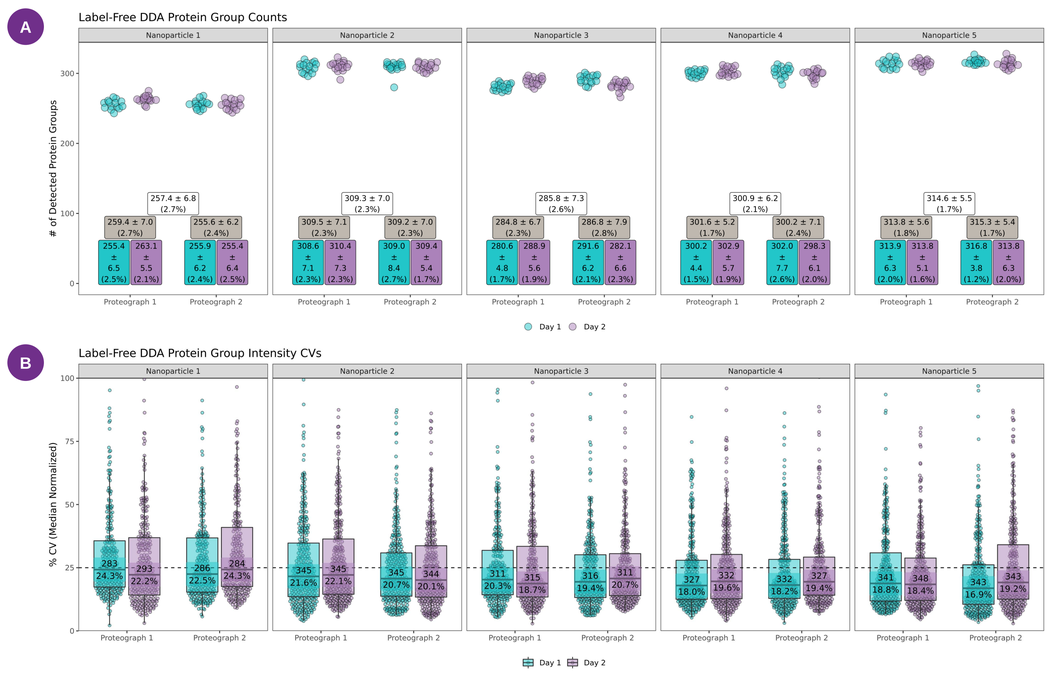
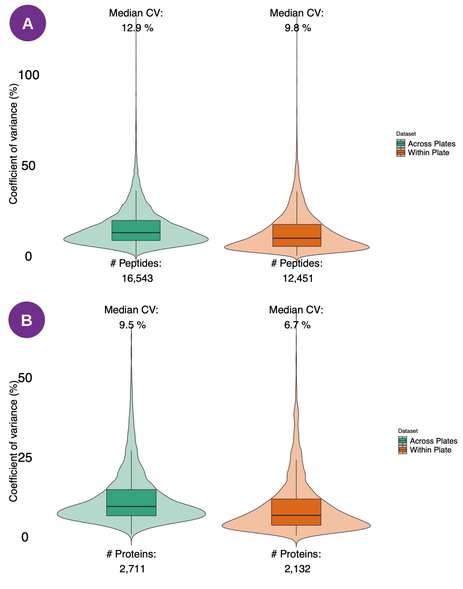


Figure 2. Number of detected protein groups and protein group intensity CVs across Proteograph instruments and day by nanoparticle.

A) Protein group counts, stratified by Proteograph instrument and day. Summary statistics are shown in the boxes (*mean ± std. dev, %CV*) where the teal and purple boxes show statistic within Proteograph instrument and day, gray boxes within Proteograph instrument across days, and white across Proteograph instruments and days. The protein group counts show excellent reproducibility with CVs below 3%. **B)** Protein group intensity CVs, broken out by Proteograph instrument and day. Summary statistics are shown within each boxplot box, with the upper number corresponding to the number of protein groups represented in the distributions (present in at least 3 replicates), and the lower number the median intensity CV. The median intensity CVs are all below 25%, and below 20% for most nanoparticles.

TMTpro 16plex Data & Results



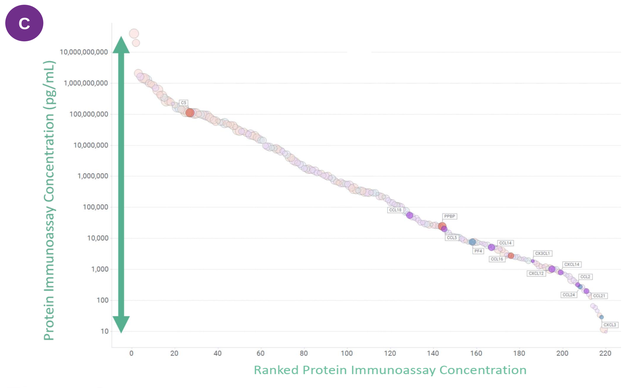


Figure 3.

A) Peptide and **B)** protein group intensity CVs computed for two different control pooled plasma samples within and between batches (plates)for the TMTpro 16plex runs. Peptide and protein group CVs are below 13% and 10%, respectively, across plates and a few points lower within plates showing an overall high degree of reproducibility across ~16,000 peptides and ~2700 proteins. **C)** Proteins from the TMT experiments were mapped to the Human Plasma Proteome Project (HPPP) protein database2 showing detection across the entire concentration range of the database spanning 9 orders of magnitude.

Conclusion

- The median CV (%) of the entire workflow including Proteograph sample preparation and mass spectrometry analysis is ~20% across 4 batches for reported label-free DDA workflow and < 10% for TMT workflow within & across batches.
- We have detected plasma proteins covering 9 orders of magnitude dynamic range reported in HPPP.
- The Proteograph Product Suite enables rapid, automated and reproducible sample preparation for deep plasma proteomic analysis for large cohort plasma proteomics studies.

References

[1] Blume et al. *Nat. Comm.* (2020)

[2] Schwenk, et al. *Journal of Proteome Research.* (2017)

Publications

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