HUPO CONFERENCE 2024

High-throughput plasma proteomics with cap-flow LC separation and dia-PASEF-MS/MS detection

Bernd Müller[1], Andreas Schmidt[1], Claudia Martelli[2], Raphael Schuster[1], Markus Lubeck[1]

[1] Bruker Daltonics GmbH & Co. KG, Bremen, Germany,
[2] Bruker Switzerland AG, Fällanden, Switzerland

Introduction

Blood is one of the least invasive biopsies and a valuable specimen for clinical research. Transferring proteomic analysis from cell lines to plasma samples faces three major challenges: First, the dynamic range of protein abundances in plasma covers more than 10 orders of magnitude with only 20 proteins contribution to >95% of the protein mass. Secondly, the typical study size in the range of hundreds of individuals is significantly larger than the that of a well-defined laboratory experiment with low sample variability which requires only few biological replicates. Finally, additional variability is introduced by plasma sampling due to the anti-coagulant and the centrifugation conditions for blood cell removal. Here we present data for 3 different plasma preparations and 2 LC-MS analysis conditions. Whereas nanoLC-MS analysis covers a higher percentage of the plasma proteomes, capLC-MS analysis multiplies the sample throughput due to low turnover times.

Plasma analysis with nano-RP HPLC-DIA-MS/MS on the TimsTOF Ultra 2

Nano-LC-MS analysis of neat plasma samples taken from 17 healthy donors prepared from plasma pheresis with citrate as anticoagulant as well as centrifugation with EDTA with single centrifugation at 1200xg and double centrifugation (1. 1200xg, 2. 2000xg) demonstrates consistency of results within a sample group, but significant differences between plasma preparations. Most peptide and protein identifications were obtained from single-spun EDTA plasma (EDTA_1spun, in average ~7500 peptides and 1000 protein IDs), citrate plasma from plasma pheresis (citrate_PP) and double-spun EDTA plasma (EDTA_2spun) were significantly lower (fig 1 AB). The three datasets show similar dynamic range and a high overlap on protein IDs (fig 1 CD). Enriched plasma preparations lead to an 3x higher number of protein IDs and almost 4x more peptide IDs for EDTA_1spun, however , EDTA_2spun had only 50% increase in protein IDs. These two datasets overlapped however protein abundance was in a similar range.

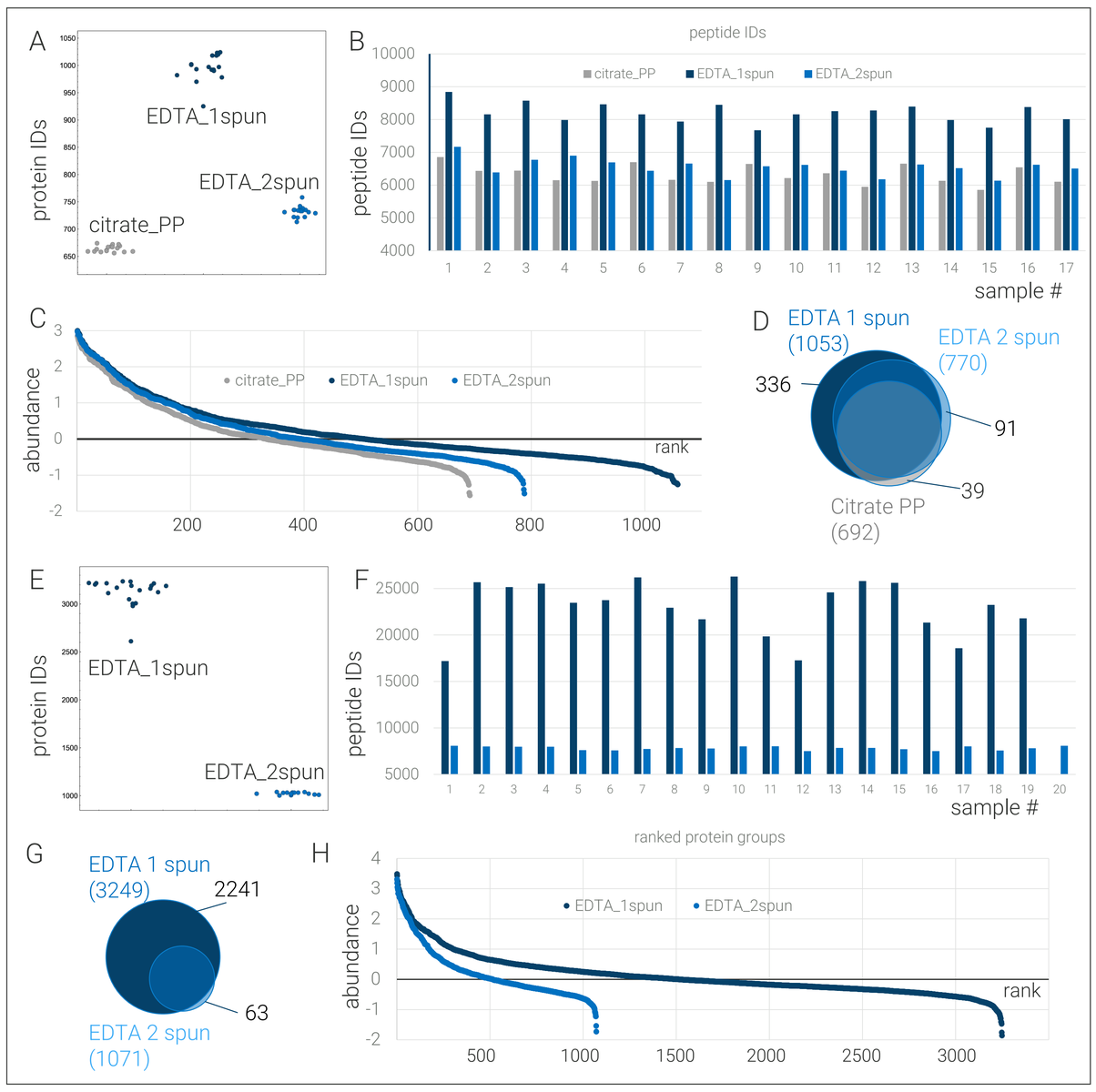
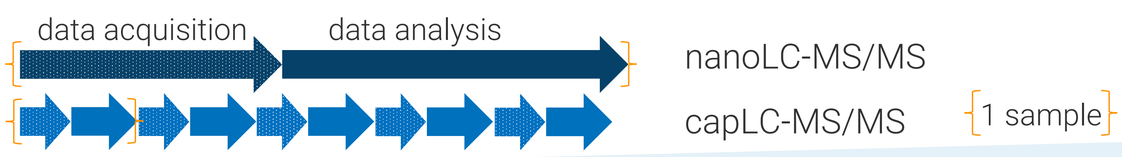


Fig. 1 plasma analysis with nanoLC-dia-PASEF-MS/MS
A) Protein IDs for 17 plasma samples/preparation condition for neat plasma digest with the iST-BCT kit. B) Peptide IDs for individual plasma samples of the Bruker cohort. C) Ranked average protein abundances in the three plasma preparation cohorts. Protein abundances were log10 transformed and the average of each sample was subtracted before averaging the protein abundance over all samples. D) Overlap of protein IDs of the three neat plasma datasets. E) Protein IDs for 19/20 plasma samples prepared with Enrich-iST kit from EDTA plasma sample with 1 (1spun) and 2 (2spun) centrifugation steps. F) Peptide IDs for individual plasma samples of the Bruker cohort. G) Overlap of protein IDs of the two Enrich-iST plasma datasets. H) Ranked average protein abundances in the two enriched plasma datasets.

Summary

Controlled sampling is one major factor for reproducibility and coverage for proteomics-based clinical research. Whereas mild sampling increases protein identifications, more stringent conditions lead to better comparability. Additional sample processing by enrichment steps improve the proteome coverage of plasma datasets without requiring additional time. Nano-LC provides the deepest coverage of the plasma proteomes, however, has a limited sample throughput. Cap-flow LC provides a good alternative that requires only about twofold higher injection amounts and can increase the sample throughput up to 5-fold.



Methods

Plasma samples from 20 healthy donors were prepared with the iST and ENRICH-iST sample preparation kits (PreOmics); peptide content for each sample was determined by UV-absorption. The TimsTOF Ultra 2 mass spectrometer was coupled to a nanoElute LC system equipped with a 25 cm Aurora reversed-phase LC column (IonOpticks). A 22 min gradient from 4% to 35% ACN was employed for separation of ~200 ng of peptides/sample. Data were acquired with a dia-PASEF acquisition program using the ICC 2.0 feature and covering precursors in the range from 400 to 1000 m/z. For capillary flow LC-MS analysis was conducted on a M-Class HPLC system (Waters) operated in HyStar and equipped with a BrukerMAX column (10cmx150um, C18 reversed-phase 1.5um) and operated in direct injection mode at a flow rate of 2.5 ul/min. The dia-PASEF method was optimized for the 6 min separation program (5% to 36% ACN in 3.2 min), by reducing the ion accumulation and TIMS ramp time to 45 ms. Fast-flow experiments were conduced with 500ng peptides/sample. Data were analyzed in Spectronaut 19 in directDIA+ mode against a human Swissprot database.

Plasma analysis with cap-RP HPLC-DIA-MS/MS on the TimsTOF HT

Implementing fast capillary-flow separation required optimization of the dia-PASEF methods (e.g. TIMS ramp time of 45 ms) to obtain sufficient datapoints. In comparison to nanoLC, capillary-flow separation of only 2.5-fold more injected peptides had a lower coverage of the plasma proteome at a more than 5-fold increased sample throughput (fig. 2A). Datasets of 3x 19 donors with 3 sampling conditions showed a high overlap of data, and a similar dynamic range (fig. 2BC). EDTA-plasma samples prepared with the ENRICH-iST kit recovered more than 1200 protein IDs from single spun plasma and ~600 protein IDs from double spun plasma using 6 min separation programs with reproducible peptide detection and with a dynamic range of more than 4 orders of magnitude (fig. 2 DEF). Plasma samples prepared with the ENRICH plus kit exceeded more than 2200 protein IDs using cap-flow LC-MS/MS (fig. 2G). Both neat plasma analysis and enriched plasma demonstrated high reproducibility in protein quantitation and datapoints/peak (fig. 2 H-J).

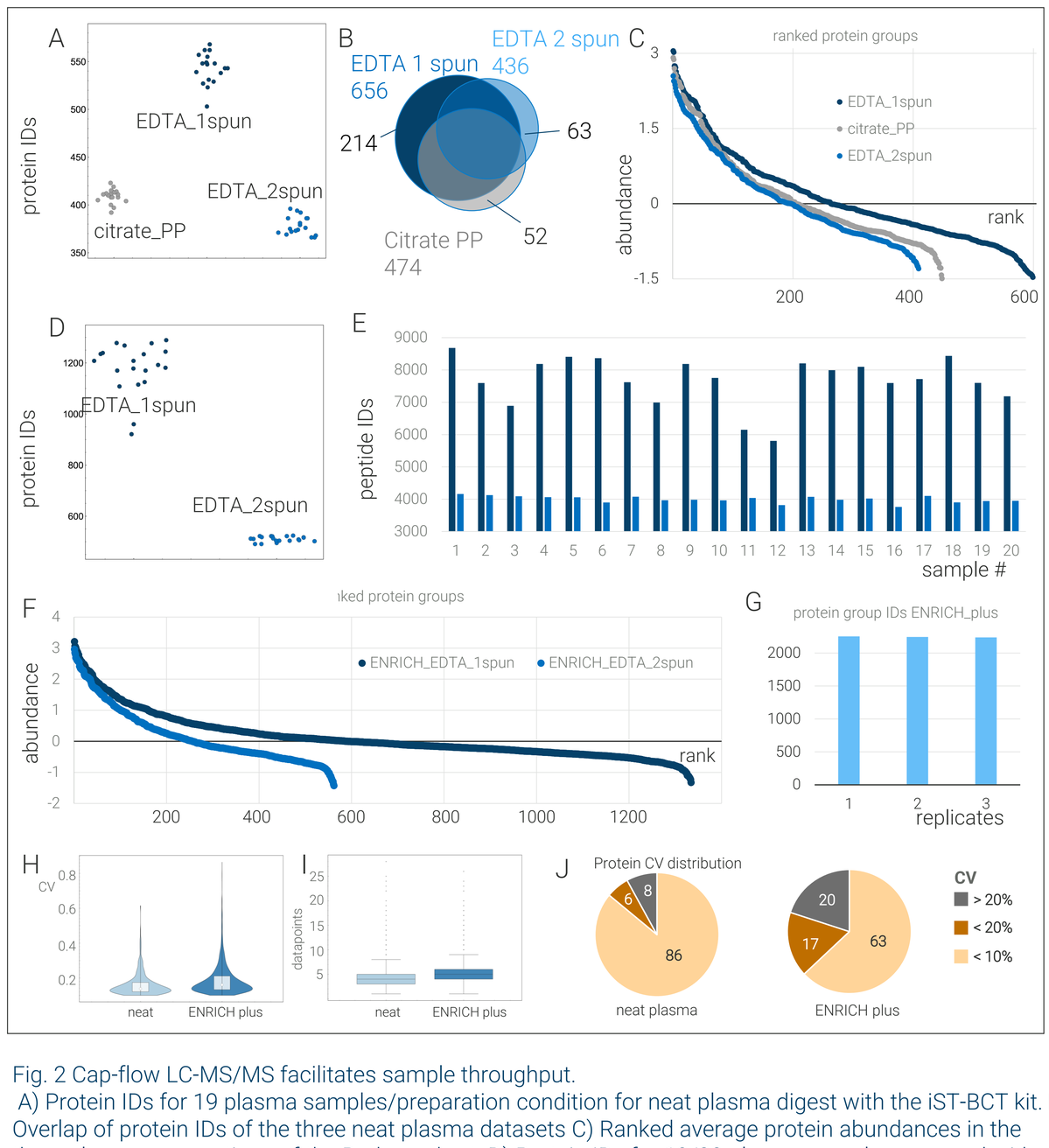


Fig. 2 Cap-flow LC-MS/MS facilitates sample throughput.
A) Protein IDs for 19 plasma samples/preparation condition for neat plasma digest with the iST-BCT kit. B) Overlap of protein IDs of the three neat plasma datasets C) Ranked average protein abundances in the three plasma preparations of the Bruker cohort. D) Protein IDs for 19/20 plasma samples prepared with Enrich-iST kit from EDTA plasma sample with 1 (1spun) and 2 (2spun) centrifugation steps. E) Peptide IDs for individual enriched EDTA plasma samples. F) Ranked average protein abundances in enriched plasma datasets. G) Protein identifications in replicates of ENRICH plus plasma analysis with cap-LC. H) CVs in neat and ENRICH plus plasma samples. I) Number of datapoints in for high-throughput plasma analysis. J) distribution of CVs in cap-LC proteomics datasets of neat and ENRICH plus plasma preparations.

Conclusion

- Controlled sampling is a key requirement for successful plasma analysis and clinical proteomics
- Deeper proteome coverage and more significant proteins are observed with ENRICH-iST at high reproducibility (CV ~20%)
- Capillary-flow chromatography facilitates sample throughput by a factor of 5-fold at similar injection amounts and high reproducibility
- dia-PASEF technology is suited for high sample throughput in clinical studies