Democratization of Metaproteome Analysis by Combining Fully Automated Sample Preparation and AI-driven Data Analysis

Amirmansoor Hakimi[1], Kevin L. Schauer[1], Runsheng Zheng[2], Alexander Boychenko[2], David Horn[1], Steven L. Reeber[1], Kristan Bahten[1], Woong Kim[1] and Daniel Lopez Ferrer[1]
[1]Thermo Fisher Scientific, San Jose, California, USA, [2]Thermo Fisher Scientific, Germering, Germany

ABSTRACT

Background
LCMS-based proteomics is a powerful tool for deep profiling of peptides and proteins in complex biological samples but when it comes to high throughput applications manual sample preparation workflow remains a major challenge in the field. Here we introduce the AccelerOme™ system, a fully automated sample preparation platform for proteomics samples that enables standardized, hands-off operation and provides robust workflows for label-free proteomics applications to reduce the sample preparation time, increase the reproducibility significantly and maximize the proteome coverage of metaproteome samples with high protein diversity.

Methods
Metaproteome standards were purchased from ZymoBIOMICS and prepared using the AccelerOme platform. Briefly, proteins were lysed, reduced, alkylated, and digested to peptides in the automated platform. 1µg of peptides were separated using an EASY-Spray™ PepMap™ Neo column on a Vanquish Neo UHPLC system coupled to an Orbitrap Eclipse Tribrid MS with/without FAIMS. The raw files were searched using Proteome Discoverer 3.0 software utilizing Sequest HT, INFERYS re-scoring algorithm and the CHIMERYS™ algorithm, a new intelligent search algorithm that rethinks the analysis of tandem mass spectra from the ground up.

Results
The automated AccelerOme sample preparation platform is an intuitive turnkey system (instrument, software, reagents) to provide robust and reproducible sample preparation for MS proteomics analysis associated with easy to use and limited user involvement and without any compromise in performance compared to traditional methods. We performed gas-phase fractionation using the FAIMS to gain dynamic range and improve proteome coverage. Using CHIMERYS alone, over 13200 proteins were identified in ZymoBIOMIC Microbial Community standard. Similar proteome coverage was observed in the ZymoBIOMIC Gut Microbiome standard dataset. Importantly, AccelerOme exhibited an extremely high digestion efficiency even for challenging samples. Zero peptide missed cleavages of 94% could be achieved for the metaproteome standards with 1-hour digestion. The sample preparation process is robust and reproducible and requires less than 20 min setup.

INTRODUCTION

Figure 1. Complete "end to end" proteomics workflow. The experimental design and other metrics such peptide amount information are carried out along the process to be consumed for improved results and reduce operator errors.

The AccelerOme platform includes Experiment Designer software, which guides the user through the experiment planning process to input sample names and assign study factors, values, and also provides an estimate of statistical power. It also benefits from an integrated touchscreen display with user interface for instrument control and operation through a graphical wizard. The liquid handling robot has the capacity to process up to 36 label-free samples, 33 Thermo Scientific™ TMT11plex isobaric labeling reagent samples or 32 Thermo Scientific™ TMTpro™ 16plex label reagent samples per session. The AccelerOme system has an automated and standardized workflows used to increase reproducibility and productivity which is part of an integrated workflow solution, from experiment design and sample preparation, to LC/MS analysis reducing training requirements while improving data quality.

Sample Preparation

The two metaproteome standards used in this study were purchased from Zymo Research Corporation called ZymoBIOMICS Microbial Community Standard (Catalog No. D6300) and ZymoBIOMICS Gut Microbiome Standard (Catalog No. D6331). Samples were prepared using glass beads and zirconium oxide beads before cleanup with Pierce™ Detergent Removal Spin Column. Dried samples were resuspended in 50ul of lysis buffer before being transferred into the AccelerOme 96 well plate.

Figure 2. Deep metaproteomic workflow including AccelerOme, Vanquish Neo UHPLC system and an Orbitrap Eclipse Tribrid mass spectrometer coupled with FAIMS Pro Interface.

LC Method

- Analytical Column: EASY-Spray™ PepMap™ Neo, 2 µm C18, 75 µm X 75 cm
- Mobile phases: [A] 0.1% FA in H_2O, [B] 0.1% FA in 80% ACN
- Flow rate: 250 nL/min
- Workflow: Direct injection
- Column temperature: 50˚C
- 1 µL injections (1µg on the column)
- Gradient: 5% to 28% B in 105 minutes and 28% to 40% B in 15 minutes.

MS Method

For the No-FAIMS experiment, a data dependent acquisition (DDA) method was used by having 240K resolution, AGC Target at 4.0e5, and Maximum IT set to Auto for the full scan. Intensity threshold was set to 5.0e3 and Dynamic Exclusion was set to 60 seconds. Data Dependent MS2 scan was set with 1 second cycle time, isolation window 0.7, Ion trap scan rate of Turbo with 1e5 AGC and 10ms maximum injection time. For the FAIMS dataset 3 CVs of -35V, -50V and -65V were used.

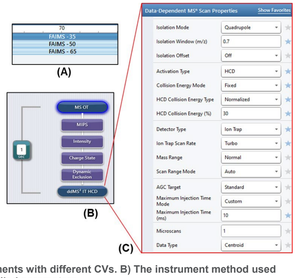


Figure 3. A) Three identical experiments with different CVs. B) The instrument method used for data acquisition. C) ddMS2 detailed parameters.

Data Processing

Proteome Discoverer 3.0 software with two different processing workflows was used for data processing. In one processing workflow (Figure 4 A), Precursor Detector with S/N 1.5 was used to handle chimeric spectra by identifying additional precursors within the isolation window of the precursor spectrum in the results. Sequest-HT was used to search the data with custom fasta files loaded for each standard and typical proteomics variable and static modifications were selected. Trypsin was the enzyme of choice with maximum number of 2 missed cleavage sites per peptide. The INFERYS Rescoring node was attached to the Sequest-HT used to improve confidence using a deep learning-based method as described in reference 3. Lower confidence spectrum researched by Sequest-HT with trypsin in a semi specific cleavage mode. In the second processing workflow (Figure 4 B) utilizing CHIMERYS, a Top N Peaks Filter is used to select top 20 of the most abundant peaks in every 100 Da window and send them to CHIMERYS for identification.

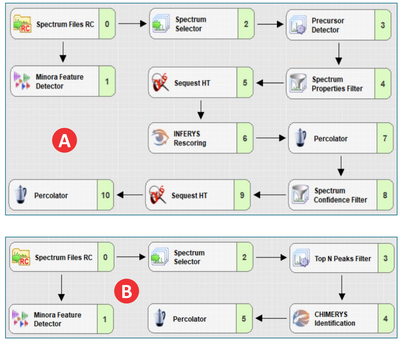


Figure 4. Processing workflows on Proteome Discoverer Software. The workflow with 2 Sequest-HT and INFERYS Rescoring node is shown in A and the CHIMERYS workflow is shown in B.

RESULTS

Identification

Using the Processing Workflow with CHIMERYS, 13,200 protein groups and over 76,000 peptides were identified from the Microbiome Community Standard and over 13,854 protein groups and 86,662 peptides from the Gut Microbiome Standard (Figure 5).

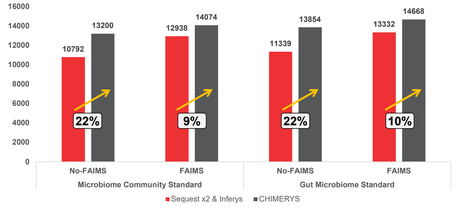


Figure 5. Number of proteins identified in each standard with and without FAIMS.

CHIMERYS Improvement Effect

When comparing the 2 different processing workflows, CHIMERYS shows improvement in identifying proteins with 22% higher in No-FAIMS groups and about 10% higher in FAIMS groups in both standards. The addition of the FAIMS Pro Interface to the workflow improved the protein identification by 6.6% and 6.3% in Microbiome Community Standard (14,704 proteins) and Gut Microbiome Standard (14,668 proteins), respectively (Figure 6).

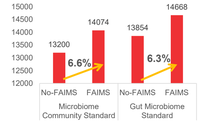


Figure 6. Number of proteins identified when FAIMS is used.

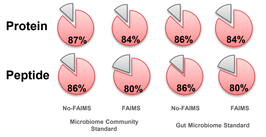


Figure 7. Percent quantified proteins and peptides

Quantification

Out of all the identified protein groups in both Microbiome Community Standard and Gut Microbiome Standard, on average, 85.2% were quantified accurately (Figure 7). Also, very high levels of quantification rate at the peptide level was observed with 83% (on average) of all the identified peptides were quantified.

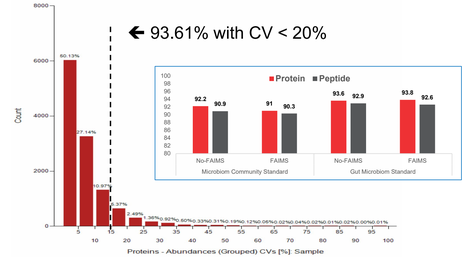


Figure 8. Quantitative reproducibility for triplicate injection. More than 93% of the proteins quantified with %CV of less than 20% in triplicates injections of Gut Microbiome Standard. Similar observation was made for another sample with and without FAIMS Pro Interface.

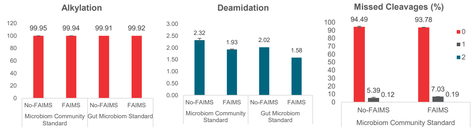
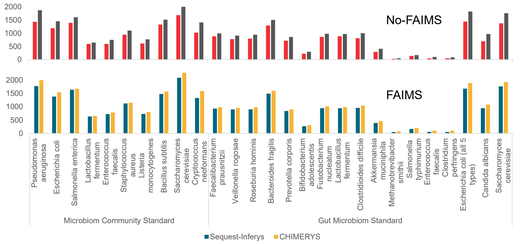


Figure 9. Different metrics including % missed cleavage, alkylation and deamidation, were evaluated. Very high levels of alkylation, and very low levels of deamidation were observed among all the standards prepared using the AccelerOme system Also, close to 95% of the peptides were identified with zero missed cleavages which indicates high digestion efficiency of the protocol. Overall these metrics show an excellent data quality and outstanding robustness of the LC-MS platform.

CHIMERYS Effect on Individual Species

Microbiome Community Standard and Gut Microbiome Standard are made from 10 and 17 different species, respectively. Looking deeper into the results and evaluate number of proteins identified in every species of each standard, a significant improvement was observed when CHIMERYS processing workflow was used in both No-FAIMS and FAIMS datasets (Figure 10).



CONCLUSIONS

- The AccelerOme platform offers simplified workflows with minimized user involvement and improved reproducibility through instrument functionality and automation. It helps increased efficiency and productivity through pre-built and validated sample preparation methods and reagents delivered in kit format, ensuring experiment democratization. The AccelerOme platform reduces total cost of sample prep by eliminating the need for dedicated personnel for mass spectrometry sample preparation.
- Analysis of metaproteome standards using a Vanquish Neo UHPLC system with an EASY-Spray PepMap Neo column coupled to an Orbitrap Eclipse Tribrid mass spectrometer ,enables great proteome coverage by identifying 13,200 and 13,854 proteins in Microbiome Community Standard and Gut Microbiome Standard, respectively when implementing CHIMERYS detection node in the processing workflow of Proteome Discoverer software.
- Addition of the FAIMS Pro Interface allows digging even deeper into the proteome by increasing the number of identified proteins by about 6.6% in Microbiome Community Standard and 6.3% in Gut Microbiome Standard.

REFERENCES

1. Wilmes, P. & Bond, P. L. Metaproteomics: studying functional gene expression in microbial ecosystems. Trends Microbiol. 14,92–97 (2006).
2. Verberkmoes, N. C. et al. Shotgun metaproteomics of the human distal gut microbiota. ISME J. 3, 179–189 (2009).
3. Zolg, D. P. et al. INFERYS rescoring: Boosting peptide identifications and scoring confidence of database search results. Rapid Communications in Mass Spectrometry, Special Issue, May 2021.

TRADEMARKS/LICENSING

PO66168-EN0422S