Application note | 003046

Thermo Fisher SCIENTIFIC

Proteomics

Unveiling hidden protein depths: a high-throughput plasma proteomics workflow for enhanced biomarker discovery

Authors

Kevin Yang[1], Amarjeet Flora[2], Khatereh Motamedchaboki[1], Stephanie Samra[1], Bhavin Patel[2], and Amirmansoor Hakimi[1]

[1]Thermo Fisher Scientific, San Jose, CA, USA

[2]Thermo Fisher Scientific, Rockford, IL, USA

Keywords

Plasma, Orbitrap Astral mass spectrometer, blood, neat plasma, immuno-depletion, Vanquish Neo UHPLC, Seer SP100 automation instrument, Proteograph XT workflow

Goal

To develop robust plasma proteomics workflows with different sample preparation and various throughput methods for deep plasma proteomics and superior quantitation by a label-free data-independent acquisition (DIA) strategy on the latest Thermo Scientific™ Orbitrap™ Astral™ mass spectrometer

Introduction

Cancer continues to claim countless lives and holds the second-highest ranking in overall mortality worldwide. The late diagnosis of this disease frequently leads to a poor prognosis, emphasizing the crucial need for improved early detection tools. Present diagnostic tests either lack sensitivity or require invasive biopsies and are not always feasible or comfortable for patients.

One of the significant advantages of plasma proteomics is the ease of sample collection through routine blood draws. The analysis of blood plasma for cancer biomarkers presents a promising avenue for early detection, but the complexity of the associated workflows has hindered progress in this area.[1]

The analysis of plasma presents a substantial challenge due to the vast concentration range of proteins in blood plasma. Detecting low abundant proteins, which hold potential as biomarkers, proves challenging as they are often greatly overshadowed by high-abundant proteins. In fact, approximately 99% of the plasma proteome is composed of

thermo scientific

proteins such as albumin, globulins, and coagulants, effectively overshadowing potentially less abundant protein biomarkers. Therefore, there is a pressing need to develop methods that can accurately identify and quantify these low-abundant proteins in plasma samples.

Mass spectrometery-based proteomic analysis, in conjunction with advanced separation technology, is a leading method for studying low-abundant proteins due to its sensitivity, unambiguous peptide identification, and accurate quantitation. The cutting-edge Orbitrap Astral mass spectrometer offers unprecedented sensitivity and reproducibility for low-abundant proteins, which enables deeper plasma proteome coverage and facilitates the discovery of potential biomarkers.

Currently, several methods of sample preparation have been developed for plasma proteomics:

1. **Neat plasma workflow:** This method involves analyzing neat, undepleted plasma proteins. While it offers the highest proteomic analysis throughput, it provides shallow coverage of the plasma proteome as low-abundant proteins are masked due to the dominance of common, highly abundant plasma proteins.
2. **Immuno-depletion workflow:** This approach involves depleting highly abundant plasma proteins to provide a higher chance of lower abundant plasma protein identification and address the challenging dynamic range of the plasma proteome, effectively eliminating the highly abundant proteins. However, issues may arise with the throughput of depletion, the potential for non-specific protein removal, or potential high abundant biomarkers and the introduction of additional technical variation. To maintain throughput, the process needs to be automated, which adds to the cost per sample.
3. **Plasma protein enrichment:** This method involves using a panel of engineered nanoparticles for enrichment of the plasma proteome. To ensure efficient throughput and reproducibility of the assay, automation becomes essential when handling nanoparticle enrichment. However, it is important to note that automation also increases the cost per sample but is often required for large plasma proteomics studies. Techniques such as the use of the Seer Proteograph™ XT workflow (Seer, Inc.) allow for broad and deep coverage of the plasma proteome in a high-throughput and robust manner.[2]
4. **Fractionation schemes:** These methods separate digested peptides into less complex fractions, which can facilitate deeper protein identification. However, they also increase the time required for fractionation and analysis of each fraction.

Irrespective of the plasma sample preparation methods used, there is a pressing need for a standardized high-throughput liquid chromatography coupled to mass spectrometry (LC-MS) workflow for plasma proteomics, allowing larger studies with enhanced statistical power to see subtle health- or disease-related changes in the plasma proteome. This workflow should balance the depth of identification with the ability to handle large population cohorts. Given the requirement to analyze a high volume of samples with maximum coverage, LC-MS in data-independent analysis (DIA) mode has emerged as a clear choice for ensuring reproducible analyses.[2-4] The Orbitrap Astral mass spectrometer sets a benchmark for plasma proteomics, offering both deeper coverage and faster throughput due to its enhanced speed and sensitivity without compromising on assay reproducibility or quantitative performance.

Every aspect of the workflow needs to be optimized and robust to analyze hundreds to thousands of plasma samples in large-cohort studies. The Thermo Scientific™ Vanquish™ Neo UHPLC system delivers precise low flow rates and gradient formation to pair perfectly with the Orbitrap Astral mass spectrometer. Additionally, the Vanquish Neo UHPLC system provides vial bottom detection technology to ensure maximum sample recovery. The combination of the Orbitrap Astral mass spectrometer and the Vanquish Neo UHPLC system provides an optimal solution for high-throughput and reproducible plasma proteomics analysis.

For optimal separation of peptides from plasma proteins, we incorporated Thermo Scientific™ EASY-Spray™ HPLC columns into the workflow. Their fully integrated emitter minimizes dead volume, thus enhancing the detection of low-abundant proteins. Additionally, the user-friendly set-up, column-to-column consistency, and superior robustness of these columns contribute to maximized data reproducibility. In this study, we adhere to the high-throughput requirements of various translational research initiatives for plasma proteomics solutions (Figure 1). The innovative sample processing, in combination with the advanced Orbitrap Astral mass spectrometer, results in the most extensive plasma proteome coverage to date. Collectively, we have developed a range of sample preparation methods, each paired with a different LC-MS method with a different sample analysis throughput. This advancement will significantly contribute to translational research, providing the depth necessary for early disease detection through comprehensive high-throughput plasma proteome analysis.

Figure 1. Schematic diagram of high-throughput plasma proteomics analysis powered by the Orbitrap Astral mass spectrometer. An end-to-end workflow, from sample handling to data analysis, provides the analytical flexibility to meet the needs of translational researchers. This workflow facilitates faster throughput for large cohort studies or deeper coverage for more comprehensive proteomics insights.

Experimental

Common consumables

- Water with 0.1% formic acid (FA) (v/v), Optima™ LC-MS grade, Thermo Scientific™ (P/N LS118-500)
- 80% Acetonitrile (ACN), 20% water with 0.1% formic acid, Optima™ LC-MS, Fisher Chemical™ (P/N LS122500)
- Formic acid, 99.0+%, Optima™ LC-MS grade, Fisher Chemical™ (P/N A117-50)
- Thermo Scientific™ EASY-Spray™ PepMap™ column, 2 µm C18 150 µm x 15 cm, (P/N ES906)
- Thermo Scientific™ EASY-Spray™ PepMap™ Neo column, 2 µm C18 75 µm x 50 cm, (P/N ES75500PN)
- Thermo Scientific™ PepMap™ Neo trap cartridge, 5 µm C18 300 µm x 5 mm, (P/N 174500)
- Thermo Scientific™ High Select™ Depletion Spin Columns (P/N A36370)

Instrumentation

- Vanquish Neo UHPLC system
- Orbitrap Astral mass spectrometer
- Seer SP100 automation instrument

Samples

The neat plasma sample used in this experiment was from a pooled sample collected from multiple donors. The sample was prepared using the Thermo Scientific™ EasyPep™ Mini MS Sample Prep Kit (P/N A40006). Ensuing peptides were dried with a Thermo Scientific™ Savant™ SpeedVac™ concentrator and reconstituted in water with 0.1% FA and 5% ACN to a final concentration of 500 ng/µL. Depending on the experiment, 200 ng, 500 ng, or 1,500 ng of peptides were loaded on the column for LC-MS analysis.

For plasma protein enrichment with the Seer Proteograph XT Assay, 240 µL of plasma sample was mixed with each of the two nanoparticle (NP) wells included in the Proteograph XT Assay Kit. A one-hour incubation allowed high-affinity proteins to displace high-abundance proteins, resulting in a reproducible protein corona on each NP surface that is designed to probe the depth of the plasma proteome. A series of gentle washes removed non-specific and weakly bound proteins. The paramagnetic property of the NPs allowed for the accumulation of NPs with the protein corona after each wash step. This resulted in a highly specific and reproducible protein corona that contained the high-affinity protein binding partners selected by the NPs. Protein coronas were reduced, alkylated, and digested with Trypsin/Lys-C to generate peptides for LC-MS analysis. All steps were performed in a one-pot reaction directly on the NPs. The digested peptides were then desalted, and all detergents were removed using a mixed-media filter plate and positive pressure (MPE) system. Clean peptides were eluted in a high-organic buffer into a deep-well collection plate. The ensuing tryptic peptides were analyzed on the Orbitrap Astral mass spectrometer with methods with different sample loads and analytical throughput.

The top 14 abundant proteins in human plasma samples were depleted using depletion mini spin columns from the High Select Depletion Spin Columns (P/N A36370). Depleted protein samples were dried in a SpeedVac concentrator, followed by processing with the EasyPep Mini MS Sample Prep Kit and reconstitution

in sample loading buffer composed of water with 0.1% FA and 5% ACN to a final concentration of 500 ng/µL. Depending on the experiment, 200 ng, 500 ng, or 1500 ng of peptides were loaded on the column for LC-MS analysis.

LC-MS analysis

For all subsequent LC runs described later (neat, depleted, and NP-enriched plasma), mobile phase A = 0.1% FA in H_2O, and mobile phase B = 0.1% FA in 80% ACN using a column temperature of 55 °C and an autosampler temperature of 7 °C. All samples were run in replicates of three.

Source parameters, including spray voltage and ion transfer tube temperature, are tunable parameters and must be optimized for the individual setup. The details of the LC gradient, LC parameters, and mass spectrometer method are reported in Tables 1 and 2. Identical MS parameters were applied consistently across all throughputs. For gas phase fractionation, the Proteograph XT-enriched plasma samples were analyzed by 60 samples per day (SPD) methods with 100 *m/z* mass range spanning from *m/z* 380 up to 980 for a total of six windows.

Table 1. Liquid chromatography parameters

<table>
<tr><th colspan="4">24 SPD (50 cm EASY-Spray PepMap Neo column, P/N ES75500)</th></tr>
<tr><td rowspan="9">Gradient</td><td>Time</td><td>%B</td><td>Flow (µL/min)</td></tr>
<tr><td>0</td><td>8</td><td>0.5</td></tr>
<tr><td>2.5</td><td>8</td><td>0.5</td></tr>
<tr><td>3</td><td>8</td><td>0.25</td></tr>
<tr><td>37</td><td>28</td><td>0.25</td></tr>
<tr><td>48.5</td><td>35</td><td>0.25</td></tr>
<tr><td>48.9</td><td>98</td><td>0.25</td></tr>
<tr><td>49</td><td>98</td><td>0.35</td></tr>
<tr><td>54</td><td>98</td><td>0.35</td></tr>
<tr><td rowspan="5">LC parameters</td><td>Column temperature</td><td colspan="2">50 °C</td></tr>
<tr><td>Fast loading/equilibration</td><td colspan="2">Pressure Control</td></tr>
<tr><td>Pressure for loading/equilibration/wash</td><td colspan="2">Max Pressure</td></tr>
<tr><td>Equilibration factor</td><td colspan="2">2</td></tr>
<tr><td>Sampler temperature</td><td colspan="2">7 °C</td></tr>
<tr><th colspan="4">60 SPD (15 cm EASY-Spray PepMap column, P/N ES906)</th></tr>
<tr><td rowspan="9">Gradient</td><td>Time</td><td>%B</td><td>Flow (µL/min)</td></tr>
<tr><td>0</td><td>10</td><td>2</td></tr>
<tr><td>0.3</td><td>10</td><td>2</td></tr>
<tr><td>0.6</td><td>10</td><td>0.8</td></tr>
<tr><td>13.6</td><td>22.5</td><td>0.8</td></tr>
<tr><td>20.5</td><td>35</td><td>0.8</td></tr>
<tr><td>20.9</td><td>55</td><td>2</td></tr>
<tr><td>21.4</td><td>99</td><td>2</td></tr>
<tr><td>22.3</td><td>99</td><td>2</td></tr>
<tr><td rowspan="5">LC parameters</td><td>Column temperature</td><td colspan="2">50 °C</td></tr>
<tr><td>Fast loading/equilibration</td><td colspan="2">Pressure Control</td></tr>
<tr><td>Pressure loading/equilibration</td><td colspan="2">Max Pressure</td></tr>
<tr><td>Equilibration factor</td><td colspan="2">2</td></tr>
<tr><td>Sampler temperature</td><td colspan="2">7 °C</td></tr>
</table>

Table 1. Liquid chromatography parameters (continued)

100 SPD (15 cm EASY-Spray HPLC column, P/N ES906			
	Time	**%B**	**Flow (µL/min)**
Gradient	0	4	2.5
	0.2	8	2.5
	7.8	20	2
	11.4	35	2
	11.8	99	2
	12.5	99	2
LC parameters	Column temperature	50 °C	
	Fast loading/equilibration	Pressure Control	
	Pressure loading/equilibration	Max Pressure	
	Equilibration factor	2	
	Sampler temperature	7 °C	

Table 2. Mass spectrometer parameters

MS parameter		
MS1	Resolution	240K
	Scan range *(m/z)*	380–980
	AGC	500%
	Max-IT	5 ms
MS2	Resolution	15K
	Scan range *(m/z)*	145–2000
	Isolation window *(m/z)*	3
	Window placement optimization	on
	AGC	500%
	Max-IT	7 ms
	HCD	25%

Data analysis and post-processing

Acquired LC-MS data was processed by DIA-NN software (v1.8.1) or Thermo Scientific™ Proteome Discoverer™ software using the CHIMERYS™ intelligent search algorithm by MSAID™. The reference human proteome FASTA file was acquired from UniProt.

For Proteome Discoverer software, the CHIMERYS node was used with UniProt human FASTA files selected as the protein database of choice, and the INFERYS™ prediction model version 3 was used for the CHIMERYS search. Oxidation of methionine and carbamidomethylation of cysteine were selected as dynamic and static modifications, respectively. Trypsin was the enzyme of choice, with a maximum number of two missed cleavage sites per peptide.

For DIA-NN software, default settings were used for either direct DIA or library search. To generate a spectral library, the six gas phase fractionation runs were analyzed on DIA-NN software and a library of approximately 8,469 protein groups were generated. All the PSMs, peptides, and proteins were filtered at 1% FDR.

The resulting candidate tables and report files for data searched with either DIA-NN software or CHIMERYS in Proteome Discoverer software were exported to .csv or .tsv files. The ensuing tables were imported to Python™ or a spreadsheet for downstream data analysis and visualization.

Results and discussion

Neat plasma

The 60 SPD workflow offers a good balance of throughput and proteome coverage, given that no protein depletion or enrichment is required. However, due to abundant proteins being present, the depth of coverage is compromised compared to the other methods. With the sensitivity and higher throughput of the Orbitrap Astral mass spectrometer, the coverage of neat plasma for short LC gradients surpasses that of previous approaches. With the 60 SPD method, we identified on average 718 protein groups along with 7,763 peptide groups and a CV of approximately 6%, which demonstrates excellent quantitative precision (Figure 2).

Top 14 protein depletion

The employment of High Select Depletion Spin Columns presents an economical method to eliminate the most abundant proteins from human plasma samples. This depletion process enhances the detection and identification of low-abundant proteins of interest. In this study, we analyzed the depleted plasma using the three different throughput methods developed herein: high-throughput approaches with 60 and 100 SPD, and a max-identification method with 24 SPD on an Orbitrap Astral mass spectrometer.

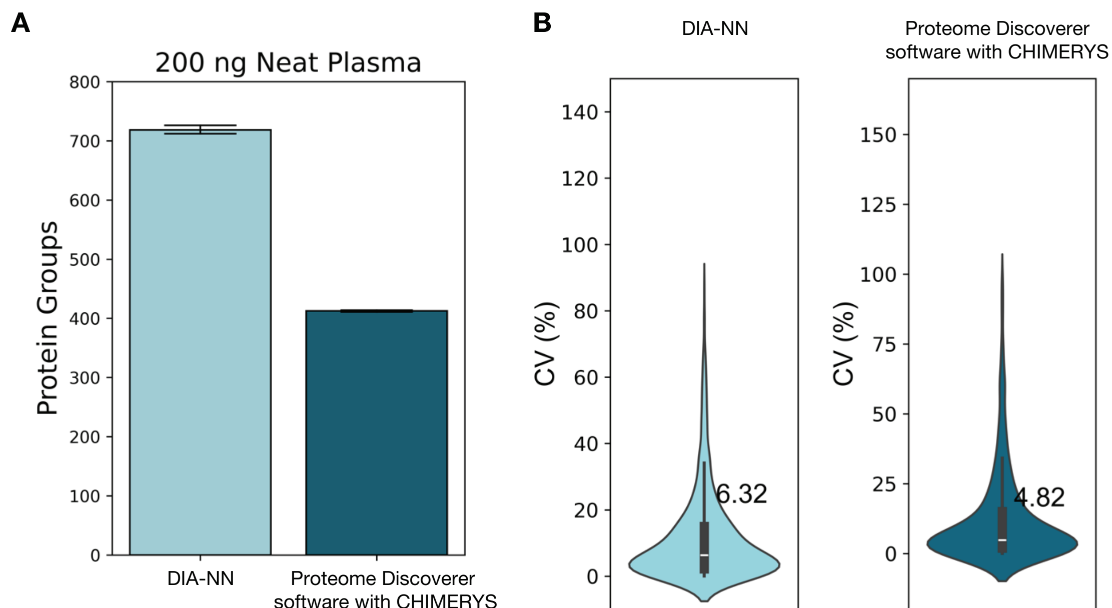

Figure 2. The Orbitrap Astral mass spectrometer improves plasma proteome coverage with a neat plasma digest workflow. (A) The number of protein groups identified from 200 ng of neat plasma digest. Samples were analyzed at a throughput of 60 SPD, and the ensuing raw files were processed by DIA-NN software or Proteome Discoverer software with CHIMERYS. (B) Coefficient of variation (CV) of protein groups from three individual runs.

With a daily throughput of 100 samples, we were able to identify on average 1,068 protein groups using a 500 ng sample load. The observed CV was around 10%, indicating extensive plasma proteome coverage and high quantitative precision from the depletion column. We further improved the plasma proteome coverage to 1,380 protein groups using the 60 SPD method (Figure 3).

To improve proteome coverage, we analyzed the depleted samples using a method that allowed us to process 24 SPD. This strategy led to the identification of 1,600 and 1,857 protein groups from the 500 ng and 1500 ng sample loads, respectively (Figure 3). Notably, the 24 SPD method with 1,500 ng load identified 20,439 peptides, nearly double the number of peptides as the 100 SPD method with 200 ng.

Seer Proteograph XT processed plasma

The two NP fractions were pooled into one sample in equal quantity and analyzed on an Orbitrap Astral mass spectrometer. At a throughput of 100 SPD, we identified on average 3,898 or 4,240 protein groups from 200 ng and 500 ng enriched plasma samples, respectively. Moreover, an excellent quantitative precision was exhibited with a CV of 6% (Figure 4). The results show that the Orbitrap Astral mass spectrometer allows high-throughput plasma proteomics analysis with unprecedented proteome coverage, fulfilling the need for ultra-fast biomarker screening.

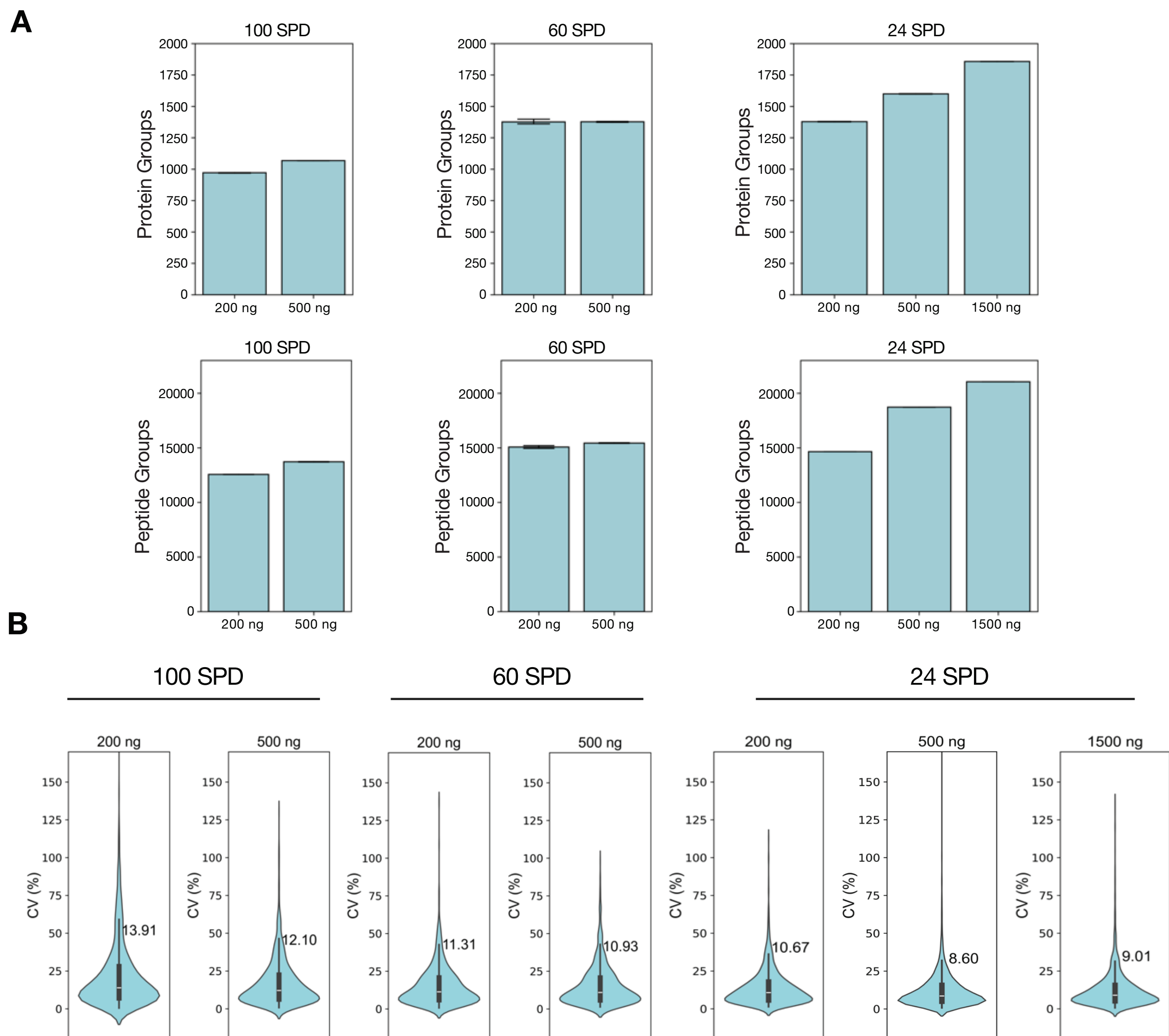

Figure 3. Top 14 protein depletion with mini columns improves plasma proteome coverage. (A) The number of protein groups (upper panel) and peptide groups (lower panel) from the top 14 mini column-depleted plasma samples at a throughput of 100, 60, or 24 SPD on the Orbitrap Astral mass spectrometer. (B) Violin plots showing the percent CV from three sample preparation replicates.

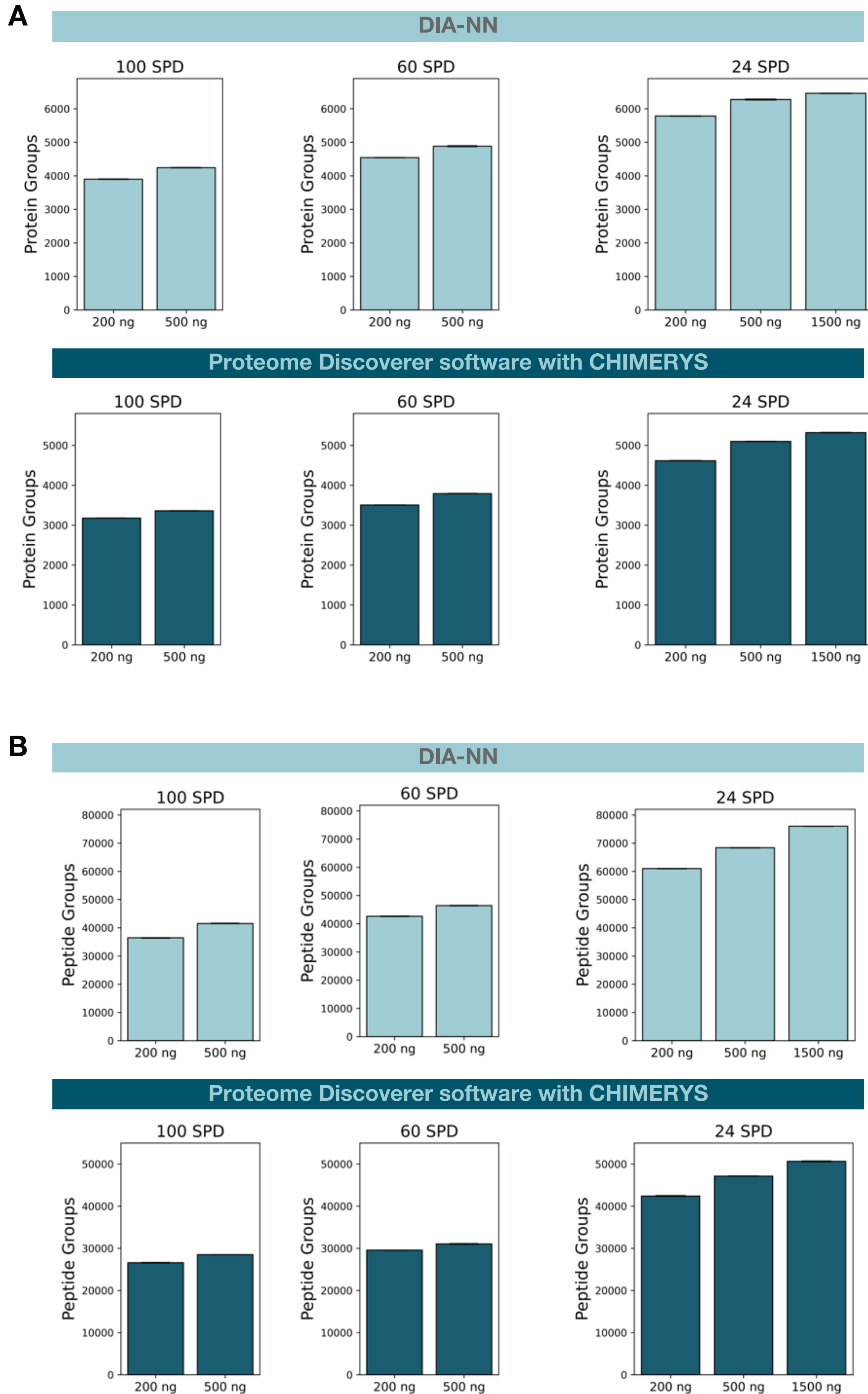

Figure 4. The Orbitrap Astral mass spectrometer coupled with the Seer Proteograph XT Assay enables deep plasma proteome coverage. The number of protein groups (A) and peptide groups (B) identified from different loads of enriched plasma. Samples were analyzed on the Orbitrap Astral mass spectrometer with 100 and 60 SPD high-throughput methods or a 24 SPD method to maximize identification.

The throughput of 60 SPD resulted in the quantification of approximately 4,881 protein groups. It is noteworthy that the CV % for these quantifications was 6%, indicating a relatively low level of variability (Figure 5). The 60 SPD method provides a good balance between high throughput and the depth of plasma proteome coverage, while maintaining excellent quantitative performance. These results demonstrate the effectiveness of this approach in protein quantification.

To enhance proteome coverage and facilitate a more comprehensive biomarker investigation, we expanded the workflow by utilizing a 50 cm EASY-Spray PepMap Neo column, allowing for a throughput of 24 SPD. Implementing this one-hour run-to-run method, we identified 6,273 protein groups from 500 ng of enriched plasma sample. Additionally, we observed a further increase in coverage, with a total of 6,457 protein groups identified from 1,500 ng of enriched plasma (Figure 4). These findings demonstrate that the Orbitrap Astral mass spectrometer enables the deepest proteome coverage from plasma samples to date. This enhanced coverage could provide novel insights by measuring previously hidden biomarkers.

Incorporating a library into the search frequently enhances proteome coverage, particularly when it comes to short gradient methods. To comprehend the advantages of library search in our high-throughput workflow, we created a spectral library using gas phase fractionation paired with the 60 SPD method. This combination resulted in the incorporation of spectra from 7,300 protein groups into our library, which further improved the identification by an additional 300 to 600 protein groups (Figure 6). Hence, a project-specific spectral library can be generated for even deeper proteome coverage in high-throughput plasma proteomics analysis conducted on the Orbitrap Astral mass spectrometer.

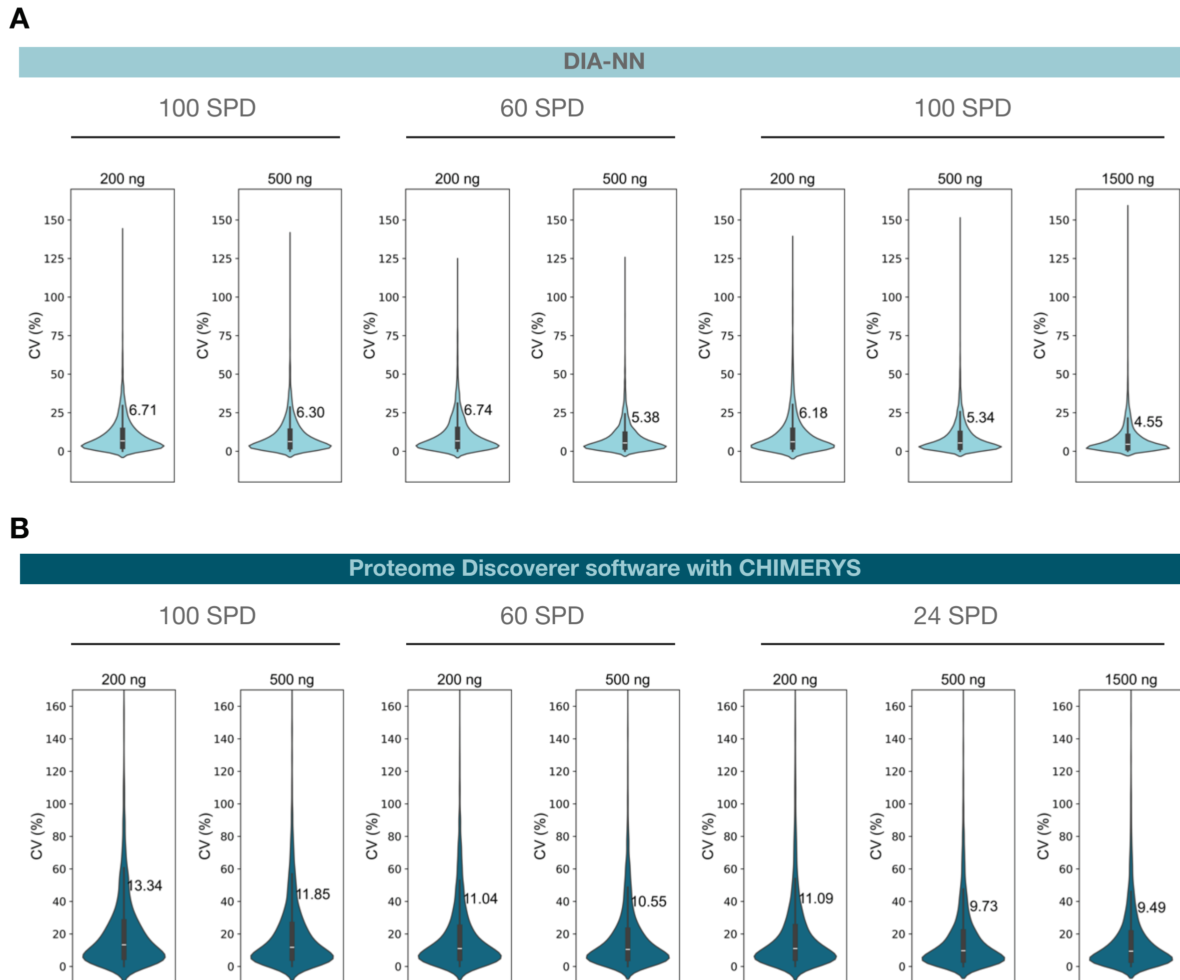

Figure 5. The Orbitrap Astral mass spectrometer coupled with the Seer Proteograph XT Assay enabling reproducible plasma protein identification. The coefficient of variance (CV) percentage of (A) protein groups and (B) peptide groups. Samples were analyzed on the Orbitrap Astral mass spectrometer with high-throughput methods, including 100 and 60 SPD or a 24 SPD method to maximize plasma protein identification.

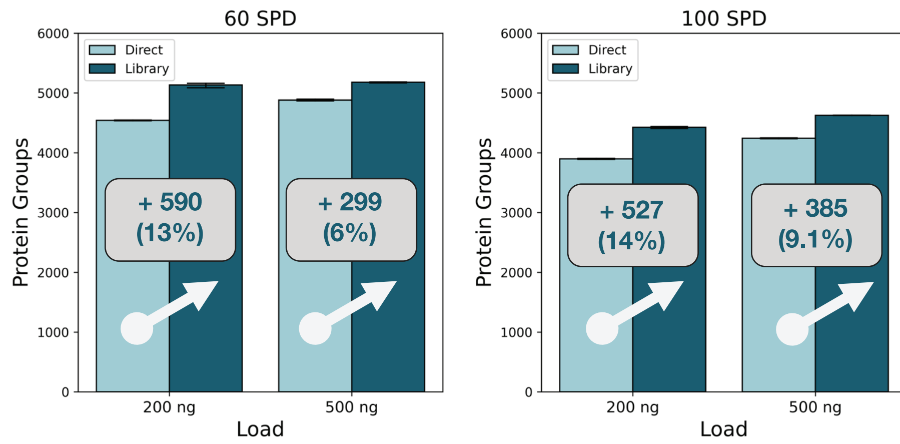

Figure 6. Utilizing a spectral library to improve proteome coverage for high-throughput workflows. Bar charts showing the number of protein groups from 100 or 60 SPD methods on the Orbitrap Astral mass spectrometer were searched with a spectral library generated by gas phase fractionation or through a library-free approach.

Enhance plasma proteome coverage through various sample preparation approaches

It's noteworthy that about 20–25% of proteins identified through the top 14 depletion workflow were not detected by the Seer Proteograph XT enrichment method, indicating this to be a cost-effective supplementary method to further enhance plasma proteome coverage if those proteins are biologically relevant in the study (Figure 7). In addition, approximately 120 protein groups were selectively detected from neat plasma (Figure 8) compared with the Seer Proteograph XT enrichment method. Altogether, our results indicate that the pairing of various sample preparation methods with the Orbitrap Astral mass spectrometer provides a potent tool for performing comprehensive, swift, and highly reproducible biomarker discovery with plasma proteomics.

Figure 7. Differential plasma protein identification with the Seer Proteograph XT and the top 14 depletion workflow. Venn diagrams showing the overlap between the Proteograph XT and the top 14 depletion methods from the 24 SPD method. Bar charts showing the KEGG pathway analysis from protein groups exclusively identified from the top 14 depletion method.

Thermo Fisher SCIENTIFIC

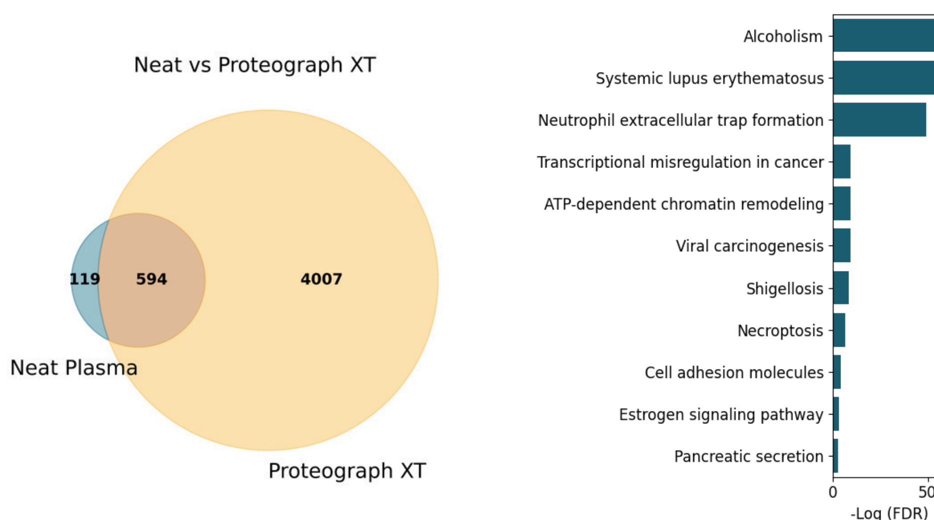

Figure 8. Differential plasma protein identification with the Seer Proteograph XT and near plasma digestion. Venn diagrams showing the overlap between the Proteograph XT and the neat plasma digest from the 60 SPD method. Bar charts showing the KEGG pathway analysis of protein groups exclusively identified from the neat plasma digest.

Conclusions

Neat plasma workflows, while easy to handle due to minimal sample preparation, suffer from dynamic range issues, which result in shallow coverage of the plasma proteome compared to enriched or depleted plasma workflows.

The use of High Select Depletion Spin Columns for plasma depletion has proven to significantly increase proteome coverage. Compared to neat plasma, proteome coverage is enhanced by 2- to 3-fold. This method provides an economical workflow and is a viable alternative to other depletion strategies used in plasma proteomics research.

The Seer Proteograph XT workflow coupled with the Orbitrap Astral mass spectrometer offers the most comprehensive depth of analysis, effectively balancing the need for high-throughput automated workflows with high quantitative precision. Taken together, these workflows with the Orbitrap Astral mass spectrometer enable new possibilities for biomarker discovery from large cohort plasma proteomics studies.

References

1. Deutsch E. et al., Advances and Utility of the Human Plasma Proteome. J. Proteome Res. 2021, 20, 12, 5241-5263.
2. Huang T. et al., Protein Coronas on Functionalized Nanoparticles Enable Quantitative and Precise Large-Scale Deep Plasma Proteomics. bioRxiv, 2023, Aug
3. Woo J. and Zhang Q. A streamlined high-throughput plasma proteomics platform for clinical proteomics with improved proteome coverage, reproducibility, and robustness. J. Am. Soc. Mass. Spectrom. 2023, Apr, 34, 4, 754-762.
4. Geyer P. et al., Plasma proteome profiling to assess human health and disease. Cell Syst. 2,3, 185-195

Learn more at **thermofisher.com/proteomics**

General Laboratory Equipment – Not For Diagnostic Procedures © Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. Seer and Proteograph are trademarks of Seer, Inc. MSAID, CHIMERYS, and INFERYS are trademarks of MSAID, GmbH. Python is a trademark of the Python Software Foundation. This information is presented as an example of the capabilities of Thermo Fisher Scientific products. It is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others. Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representative for details. **AN003046 0724S**

thermo scientific