Maximizing Protein Profiling Using HRAM MS and DDA Methods: Sweeping Through the Proteome Using Wide-Window DDA Parameters

Scott M. Peterman[1], David Sarracino[1], Amol Prakash[2], Bradley Hart[1], and Shen Luan[1] [1]Thermo Fisher Scientific BRIMS, Cambridge, MA [2]Optys Technologies Inc. Boston, MA

ABSTRACT

Purpose: Create an LC-MS workflow to perform global proteome profiling that enables the greatest depth, breadth, and reproducibility while satisfying clinical research requirements that the resulting method is easy to set up, amenable to UHPLC separations, and facilitates retrospective data mining.

Methods: Develop a mass spectrometry-based acquisition method that merges concepts of DDA/DE and DIA to maximize Orbitrap-based ion accumulation and detection. The method decouples the qualitative and quantitative data streams and employs spectral matching routines for robust, automated data processing routines.

Results: Optimization of the wide-window DDA/DE acquisition method resulted in 23% more peptides that were confidently detected and quantified as compared to DIA methods.

INTRODUCTION

Data acquisition strategies for clinical research must maximize proteome coverage and quantitative rigor while addressing high-throughput requirements. Recently, UHPLC separation technics were introduced to maximize peak capacity with shorter gradients to address throughput, but the narrow chromatographic peak widths reduce global data acquisition cycle times retarding exhaustive precursor sampling using data independent acquisition.[1-2] To address this, we utilize wide-window data dependent acquisition (wDDA)[3] and automated data processing via spectral matching to perform exhaustive and reproducible global profiling on complex samples. Employing data dependent acquisition (DDA) methods decouples qualitative and quantitative data acquisition on a Thermo Scientific™ Orbitrap™ mass spectrometer, enabling simultaneous detection and quantification of full scan MS signal and acquisition of richer product ion spectra throughout the entire peak width.

MATERIALS AND METHODS

Sample Preparation

All experiments were performed on tryptic digestions of peripheral blood mononuclear cells (PBMCs) and plasma, obtained from a single donor and prepared using similar tryptic digestion conditions. A total of 20 µg of PBMC digest was repetitively injected for all initial experiments testing various DDA settings, and a series of samples comprised of a constant amount of PBMC digest was spiked with increasing amounts of plasma digest. Both PBMC and plasma digests were separately fractionated off-line for spectral library creation using normal DDA routines.

Test Method(s)

All experiments were performed using a Thermo Scientific™ Q Exactive™ Plus mass spectrometer performing two different set of experiments. The first set acquired replicate measurements in an experimental matrix changing the precursor isolation window, AGC setting, and MS2 resolution setting. DDA isolation ranges used were 0.5, 1, 1.5, 2, 3, 5, 10, 15, 25, and 50 Da) and AGC settings were 5e4, 1e5, and 5e5. Max ion fill times for narrow precursor isolation settings were 50 msec (0.5 through 3 Da) and increased to 100 msec for 5 through 50 Da). Resolution settings were 17,500 for shorted max ion fill times and 35,000 for longer fill times. Static full scan MS data acquisition settings were used for all of the first portion of the research, including resolution setting of 70,000, AGC setting of 5e6, and max ion fill time of 50 msec. The overall loop count and cycle times were set based on the average peak width of 10 seconds equating to 12 DDA events per MS for 0.5 to 3 Da settings and a loop count of 8 for longer max ion fill times and wider precursor isolation windows.

The second set of experiments used a static, wide-window DDA method and a DIA method. The wide-DDA method used resolution settings of 70,000 and 35,000, AGC settings of 5e6 and 1e5, and max ion fill times of 50 and 100 msec for MS and MS/MS, respectively. The wide DDA method used a loop count of 10. The DIA method used resolution settings of 70,000 and 17,500, AGC settings of 3e6 and 1e6, and max ion fill times of 50 msec for both precursor and product ion acquisition. The DIA settings used 15 MS2 events with 30 Da precursor isolation windows covering a precursor m/z range of 420-900 Da.

Chromatographic separations were performed using a Thermo Scientific™ Vanquish™ UHPLC pump with a flow rate of 200 µL/min, a standard binary solvent system, and a 250 x 2.1 mm Thermo Scientific™ Acclaim™ 120 (2.2µm particle size) column. The total analysis time was 30 minutes, which included loading and diverting at 2% B for 1 min, stepping the concentration of B to 7%, and initiating a linear gradient over 25 minutes (1.2%/min) before rinsing the column at 500 µL/min and re-equilibrating prior to the next injection.

Data Analysis

Sequencing matching was performed on the 94 DDA fractions for PBMC and plasma digests using Thermo Scientific™ Proteome Discoverer™ 2.1 software. A Uniprot databased as indexed with a static alkylation modification on Cys residues, differential modifications on Met (ox), Ser/Thr (phosphorylation), and Asp (deamination). The sequence match results from the fractions were imported into the Pinnacle software (Optys Technologies Corp.) creating spectral libraries. The Pinnacle software was used to perform automated global proteome profiling by spectral matching for forward and decoy hit rate.

MATERIALS AND METHODS (cont.)

analysis. A mass tolerance of 10 ppm was used for both precursor and product ion XIC analysis prior to determining scoring attributes and area under the curve (AUC) values. Scoring attributes used to stratify peptides were based on product ion distribution analysis (dot product), precursor isotopic distribution scoring, chromatographic peak shape, and %CV across the technical replicates. Specific thresholds for the highest confidence levels (green) were (in the order list above): 0.6, 0.9, 15% peak shape variance, 20%. The scoring thresholds used for medium confidence levels (yellow) were: 0.6, 0.8, 20%, 25% and a peptide specific FDR threshold of 5%. Decoy hit rate analysis was performed using a library created from scrambling peptide sequences. The identical scoring and stratification metrics used for the forward analysis were maintained for the decoy hit rate analysis. A second round of DIA data processing for the plasma-spiked PBMC sample analysis utilized more stringent scoring criteria to reduce the overall decoy hit rate, changing the peptide specific FDR to 2.5% and product ion dot product coefficient from 0.6 to 0.7.

RESULTS

Figure 1. Representative analysis using wide DDA windows showing a narrow mass range from a full scan HRAM MS spectrum. The solid line centers on the targeted precursor m/z value, and the resulting ions isolated during the DDA event prior to simultaneous dissociation. The targeted precursor (m/z 636) has an intensity of 7e5.

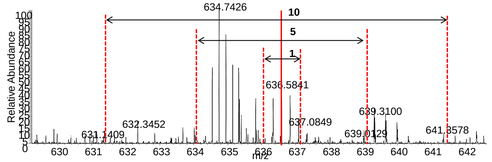

Figure 2 shows the MS/MS spectra as a function of precursor isolation width for the peptide TVWDWELMNDIKPIWQRPSK (+4). Clearly the wider the precursor isolation, the greater the complexity. To increase spectral quality, the AGC and max ion fill times are increased to handle the larger number of charge density and spectral matching results as processed in the Pinnacle software. The dot product correlation coefficient is listed for each spectrum. The spectral quality is maintained as defined by the dot-product coefficient, resulting in a confident match despite the increased number of precursors simultaneously sampled.

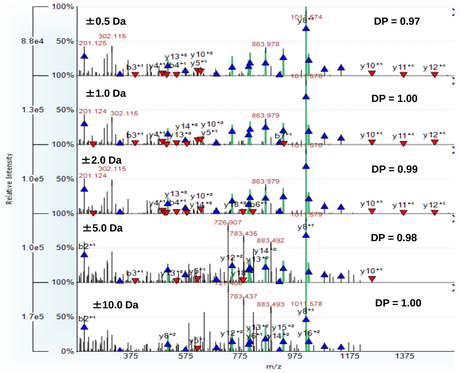

Figure 3. Comparative number of MS/MS spectra acquired per DDA/DE acquisition parameters. The settings used for wider DDA (high AGC setting, longer max fill time, and smaller loop count) reduced the number of MS/MS spectra acquired significantly at narrower precursor isolation settings. As the isolation window widths increase, precursor charge density was raised, reducing the required fill times needed to reach the AGC setting, and increasing the number of MS/MS spectra acquired.

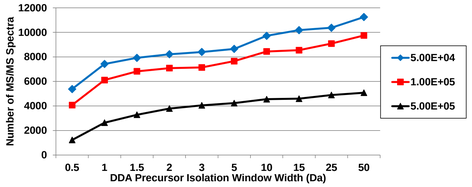

Figure 4. Comparative analysis of matched peptides as a function of wDDA parameters for the same data using sequence and spectral matching routines.

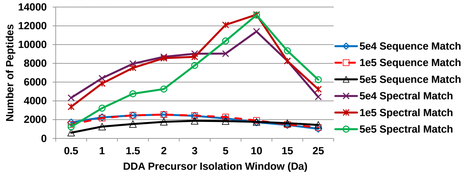

Table 1. The results shown above demonstrate that spectral matching leverages the increased product ion content resulting from wDDA methods, much the same as for DIA. Likewise, the size and quality of the spectral library plays a primary role in confidently extracting data through spectral matching. Table 1 lists the number of true positive hits as well as the number of decoy hits and decoy hit rate (DHR) using the same data, spectral library, and scoring thresholds. The DHR percentage is listed with the "green" hits and the total DHR with the "yellow" hits. The table lists the comparative results from the optimal DDA settings for each AGC target value.

	True Positive Matches			Decoy Hit Rates		
	Green	Yellow	Missed	Green (%)	Yellow (%)	Missed
5e4 2 Da						
3 msf files	2740	665	83	0	0	3488
6 msf files	2916	755	332	0	0	4002
9 msf files	2969	804	505	0	4 (0.11)	4278
94 fractions	4880	3180	58619	1 (0.02)	3 (0.05)	66675
1e5 5 Da						
3 msf files	2618	568	301	0	0	3488
6 msf files	2837	729	437	0	6 (0.17)	3997
9 msf files	2943	790	545	0	2 (0.05)	4276
94 fractions	6068	4927	55684	3 (0.05)	61 (0.58)	66615
5e5 10 Da						
3 msf files	2319	457	284	1 (0.04)	14 (0.54)	3473
6 msf files	2573	536	894	2 (0.08)	28 (0.96)	3960
9 msf files	2676	569	1033	3 (0.11)	21 (0.75)	4254
94 fractions	6340	4642	55697	18 (0.28)	202 (2.00)	66459

Figure 5. Correlation of sampling frequency to DDA/DE settings per peptide across three technical replicates. The matched MS/MS spectra were comprised of all charge states and met product ion distribution overlapping through dot product correlation coefficient analysis.

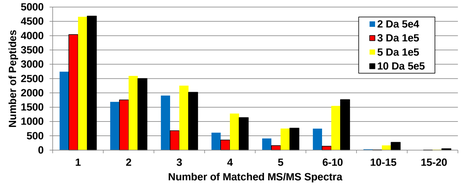

Figure 6. Evaluation of the spectral matching quality as a function of MS/MS acquisition parameters using the difference in calculated dot product correlation coefficients. The primary comparison is between the wider windows and default DDA parameters as well as between the two wider window settings. Omitted from the histogram were the values showing now discernable coefficients. Only peptides identified with correlation coefficients greater than 0.6 were considered.

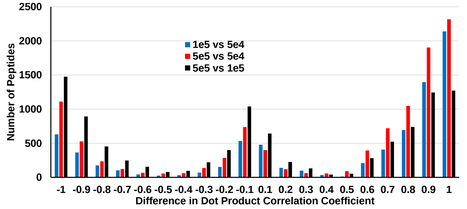

Table 2. List of peptide counts with reproducibility assessment based on MS1 and MS/MS data processing using 94 fraction spectral library. The columns present the comparison for all peptides successfully matched across technical replicates. The discriminating scoring aspects differentiating peptides scored as "green" and "yellow" are either precursor isotopic distribution overlap (0.9) and/or %CV for MS1 AUC reproducibility of 20%.

	94 Fraction Library		
	All 3 Reps	2/3 Replicates	1/3 Replicates
5e4 2 Da			
MS1 (Green)	4719	122	39
MS1 (Yellow)	1994	826	334
MS2 (Green)	3060	889	934
MS2 (Yellow)	373	922	1785
1e5 5 Da			
MS1 (Green)	5688	273	107
MS1 (Yellow)	2859	1219	763
MS2 (Green)	3890	1168	1009
MS2 (Yellow)	828	1447	2417
5e5 10 Da			
MS1 (Green)	6040	261	40
MS1 (Yellow)	2573	1547	734
MS2 (Green)	4106	1253	602
MS2 (Yellow)	676	1287	2133

Table 3. Spectral matching results for different amounts of plasma spiked into PBMC. The two methods compared were DIA and wDDA/DE with 5 Da precursor isolation and 1e5 AGC setting. The initial scoring was consistent with the described methods and the true hits and decoy hits and rates. The second round of DIA processing used dot product correlation coefficient of 0.7 to reduce the decoy hit rate.

	True Positive Matches			Decoy Hit Rates		
	Green	Yellow	Missed	Green (%)	Yellow (%)	Missed
DIA						
50 ng	2664	337	2448	220 (7.6)	76 (9.0)	1070
100 ng	2928	350	3580	278 (8.7)	96 (10.2)	1737
250 ng	2971	324	3813	345 (10.4)	129 (12.6)	1891
500 ng	2752	345	4327	361 (11.6)	124 (13.5)	2190
DIA*						
50 ng	1861	135	3453	80 (4.1)	20 (4.8)	1266
100 ng	1989	163	4706	122 (5.8)	37 (6.9)	1952
250 ng	2034	180	1894	166 (7.5)	61(9.3)	2138
500 ng	2752	345	4327	169 (5.8)	60 (6.9)	2446
1e5 5 Da						
50 ng	2255	173	465	11 (0.5)	35 (1.9)	113
100 ng	2418	200	542	20 (0.8	38 (2.2)	95
250 ng	2459	225	583	29 (1.2)	42 (2.6)	105
500 ng	2629	255	749	28 (1.1)	38 (2.2)	148

CONCLUSIONS

The presented research demonstrated a robust DDA/DE experimental method for exhaustive and confident global proteome profiling. The method leverages wide window precursor isolation coupled with increased trapping efficiencies of the Orbitrap mass spectrometer resulting in reproducible chimeric spectral acquisition and quantitation. Spectral matching routines were used for retrospective data processing to maximize proteome coverage and provides significant advantages for clinical research due to:

- Decoupling quantitative and qualitative data enables sensitive and selective profiling using UHPLC separations
- Increasing the AGC target setting and max ion fill times increased proteome coverage while minimizing decoy hit rates
- Wide window DDA resulted in greater proteome quantitation than standard DIA when factoring in decoy hit rates

For Research Use Only. Not for use in diagnostic procedures.

REFERENCES

1. Venable, J. D., et al. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra, Nature Methods, 2004, 1 (1), 39-45
2. Gillet, L. C., et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis, Mol. & Cell. Proteomics, 2012, 11(6): 0111.016717.DOI: 10.1074/mcp.0111.016717
3. Scigocki, D., et al. Application note #1, Physikron, http://www.mswil.com/images/Physikron%20application%20note%20n1_Final_Version.pdf

TRADEMARKS/LICENSING