Label-free quantitation of protein mixtures using Zeno SWATH data-independent acquisition (DIA)

Randy Arnold[1], Takeshi Shibata[2], Ushio Takeda[2], Haruka Kumabe[3], Takeshi Masuda[4], Sumio Ohtsuki[3] and Patrick Pribil[5]
[1]SCIEX, USA; [2]SCIEX, Japan; [3]Kumamoto University, Japan; [4]Keio University, Japan; [5]SCIEX, Canada

ABSTRACT

This work describes the use of DIA for the label-free quantitation of trypsin-digested proteins from multiple organisms mixed at different ratios. Using nanoflow LC and variable-window Zeno SWATH DIA on a SCIEX ZenoTOF 7600 system, a high depth of protein detection and a high degree of quantitative fidelity are demonstrated. More than 10,000 total protein groups were quantified in the mixtures, with the ratios of protein abundances for between the mixtures closely matching the expected theoretical ratios.

INTRODUCTION

The detection and quantitation of proteins are the cornerstone of proteomics and are of great interest to researchers in a variety of fields and applications. The ability to quantify changes in protein abundance, particularly in complex matrices, is key to clinical and disease biomarker discovery. The main mass spectrometric techniques used for protein profiling consist of either data-dependent acquisition (DDA) or DIA methods. With DDA methods, sample labeling/tagging strategies are often used, whereby samples are modified with either isobaric or non-isobaric labels, followed by MS and/or MS/MS quantitation. Labeling strategies, while having advantages such as the ability to multiplex samples, also suffer from drawbacks such as dynamic range compression or added complexity of the MS1 space. In addition, the costs of labeling strategies can be prohibitive[1]. By comparison, label-free quantitation strategies are attractive because of the more simplified and cost-effective nature of the workflow. DIA methodologies have become increasingly popular and are often used for label-free quantitation of peptides and proteins because of their sensitivity, simplicity, high accuracy and reproducibility.

MATERIALS AND METHODS

Sample preparation
Commercial tryptic digest of yeast protein extract were purchased from Promega. Tryptic digests of human and *E. coli* were prepared from human cultured cells (KMS-12PE) and DH5α, respectively, using a phase transfer surfactant method[2]. Digests were diluted in water containing 0.1% formic acid and mixed at the indicated ratios (Sample A – 65% human, 30% yeast, 5% *E. coli*; Sample B – 65% human, 15% yeast, 20% *E. coli*). The final concentration of total combined protein in both samples was set to 200 ng/µL.

Chromatography
Nanoflow LC analysis was carried out using a Waters M-Class LC system. Direct-inject chromatographic separation was done using a Waters nanoEase M/Z Peptide CSH (75 µm x 150 mm, 130 Å, 1.7 µm particle size) column. Mobile phase A was water with 0.1% formic acid and mobile phase B was acetonitrile with 0.1% formic acid. The flow rate used was 300 nL/min and 2.5 µL injections (i.e. 500 ng on column) were done. The chromatographic gradient profile used was a 70-minute gradient and a 120-minute total runtime, as described in Table 1.

Mass spectrometry
Analyses were done with a SCIEX ZenoTOF 7600 system, using the OptiFlow interface and OptiFlow Turbo V ion source with a nanoflow probe for nanoflow LC. A nebulizing gas setting of 10 psi and curtain gas setting of 25 psi were used. The ionspray voltage was set to 3000 V and a source temperature of 200º C was used. Zeno SWATH DIA experiments consisted of a TOF MS scan from 400 – 1250 amu with an accumulation time of 100 ms, followed by 100 variable-width SWATH DIA windows spanning the Q1 mass range from 400 – 1000 amu. MS/MS scans were done with a mass range of 100 – 1500 amu and accumulation times of 25 ms, with using dynamic collision energy and Zeno trap pulsing either on or off.

Data processing
All Zeno SWATH DIA data were processed using DIA-NN software, version 1.8.1[3]. Library-free searches were performed using a combined FASTA comprising human, yeast and *E. coli* protein sequences from UniProt (combined library containing 31,293 proteins and 29,926 genes)[4]. Default processing settings were used, with match between runs (MBR) checked. The number of protein groups and precursors identified and quantified in the datasets were determined from the resulting output pg.matrix.tsv and pr.matrix.tsv files, respectively.

RESULTS

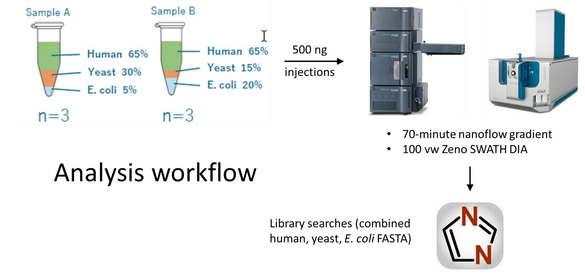
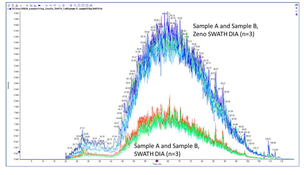


Figure 1. Overview of the label-free quantitation workflow used in this work. Two sets of mixtures were prepared from human, yeast and *E. coli* tryptic digests, with different ratios of each digest. Samples were analyzed using nanoflow chromatography and Zeno SWATH DIA on a SCIEX ZenoTOF 7600 system, with a 100 variable-width window SWATH DIA method (with or without Zeno trap pulsing in MS/MS). Data were processed with a library-free approach using DIA-NN software, with a FASTA file combining the known protein sequences for the 3 organisms.

Time (min)	Flow rate (nL/min)	% A	% B
0	300	98	2
34	300	98	2
104	300	65	35
105	300	20	80
109	300	20	80
110	300	98	2
120	300	98	2

Table 1. LC gradient profile used for the nanoflow chromatography analysis of these samples.

Figure 2. Reproducibility of the LC-MS injections and the sensitivity gains from Zeno trap pulsing in MS/MS with SWATH DIA. Overlays of the total ion chromatograms (TIC's) for replicate injections of each sample (3 each), using a 100 variable-width window SWATH DIA method either with or without Zeno trap pulsing for MS/MS. Zeno trap pulsing in MS/MS on the ZenoTOF 7600 system typically improves fragment ion signal intensities of tryptic peptides between 6- and 10-fold[5], as evidenced by the increase in overall TIC signal intensities.



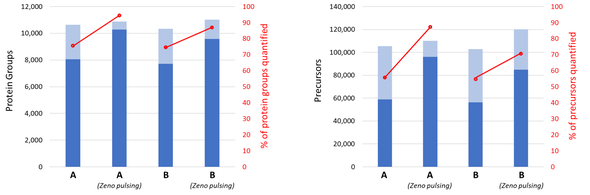


Figure 3. Protein groups and precursors identified and quantified in both samples. SWATH DIA were processed with DIA-NN software using a library-free approach, against the combined FASTA sequences for human, yeast and *E. coli*. The numbers of protein groups (left) and precursors (right) identified and quantified are shown. Light bars represent the total numbers of protein groups and precursors detected at 1% FDR, while the numbers of protein groups and precursors quantified with coefficients of variation (CV's) of 20% or better are shown by the dark-shaded bars. The percentage of protein groups or precursors quantified out of the total numbers detected are shown in red. Values were extracted from the DIA-NN software output files for the protein groups (pg.matrix.tsv) and precursors (pr.matrix.tsv), respectively. Experiments were done either with or without Zeno trap pulsing for SWATH DIA MS/MS. It is evident that the added sensitivity of Zeno trap pulsing improves detection and quantitation rates in both samples.

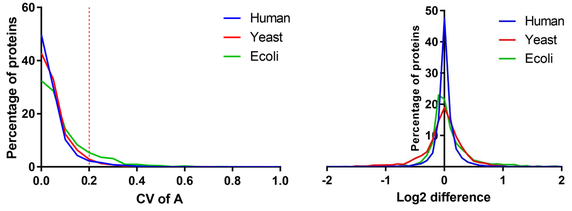


Figure 4. Quantitative statistics of the label-free analysis. (Left) Distribution of the CV's for protein groups identified in Sample A, showing the majority have CV's below 20%. (Right) Distribution of Log2 difference of the A/B ratio from the average for each species, Log2 difference = log2 (A/B of each protein for that species divided by the average A/B for that species).

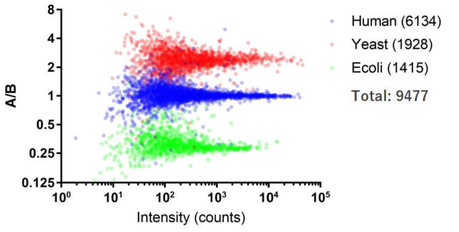


Figure 5. Quantitative ratios of proteins detected for the different organisms in the 2 mixtures from Zeno SWATH DIA. The list of detected proteins from the DIA-NN software output file was parsed for each organism and the ratios of protein abundances between the 2 samples were calculated. The ratios were plotted as a function of protein intensity. The results show a very close match to the expected values for human (1:1), yeast (2:1) and *E.coli* (1:4), owing to the high degree of reproducibility and accuracy of Zeno SWATH DIA.

CONCLUSIONS

- More than 10,000 total protein groups were detected and quantified in the mixtures of human, yeast and *E.coli* tryptic digests using Zeno SWATH DIA.
- Zeno trap pulsing improved the sensitivity of MS/MS and resulted in significant gains in protein groups and precursors quantified in the mixtures.
- The majority of protein groups and precursors detected in the mixtures were quantifiable, particularly when using Zeno SWATH DIA.
- The ratios of the protein abundances for each organism between the 2 mixtures closely matched the expected theoretical ratios (1:1 for human, 2:1 for yeast and 1:4 for *E.coli*), thereby highlighting the power of Zeno SWATH DIA for label-free proteomics quantitation.

REFERENCES

1 Rozanova, S et al. (2021) Quantitative mass spectrometry-based proteomics: an overview. Methods Mol. Biol. 2228, 85-116.
2 Masuda, T et al. (2008) Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. J. Proteome Res. 7, 731-730.
3 Demichev, V et al. (2020) DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. Nature Methods 17, 41–44.
4 Going library-free for protein identification using Zeno SWATH DIA and in silico-generated spectral libraries. SCIEX technical note, RUO-MKT-02-14675-A.
5 Zeno MS/MS with microflow chromatography powers the Zeno SWATH data-independent analysis (DIA) workflow to quantify more proteins, SCIEX technical note, RUO-MKT-02-14668-A.

TRADEMARKS/LICENSING

To receive a copy of this poster:
- Scan the code with your phone camera
- Complete the form