thermo scientific

# A Benchmarking Workflow for High-Throughput DIA Label-Free Quantification using a Novel High-Resolution Accurate Mass Platform

**Anna Pashkova[1], Julia Kraegenbring[1], Pedro Navarro[1], Tabiwang N. Arrey[1], Eduard Denisov[1], Hamish Stewart[1], Jeff Op de Beeck[2], Eugen Damoc[1]** (1) Thermo Fisher Scientific GmbH, Hanna-Kunath-Str. 11, 28199 Bremen, Germany; (2) Thermo Fisher Scientific MSS, Gent, Belgium

## ABSTRACT


- Purpose: Evaluate the performance of the novel Thermo Scientific™ Orbitrap™ Astral™ mass spectrometer for Label-Free Quantitation (LFQ) applications under a variety of conditions, such as sample loads, gradient lengths, and MS acquisition parameters.
- Methods: LFQ with a micropillar array-based column for highly reproducible separation of peptide digests coupled to an Orbitrap Astral mass spectrometer operated in data-independent acquisition (DIA) mode with narrow isolation windows (2Th), providing data acquisition at scan rates close to 200 Hz.
- Results: Very high quantitative accuracy and precision were demonstrated for 3-proteome digest mixture loads from 50 to 500ng. The dynamic range for quantitation spanned 5 orders of magnitude, which translated into reliable quantitation of proteins within the entire range of expression (from 100s to 100,000s of copies per cell).


## INTRODUCTION

Bottom-up proteomics has proven to be the most suitable technology for the high-throughput analysis of very complex biological samples, such as cell lysates or blood. As the obtained data become more and more employed in biomedical research, the challenge of analyzing the smallest sample input amounts in the shortest time remains. Data analysis software must keep up with the fast data acquisition and increasing size of raw data files. To meet all these challenges, a novel mass spectrometer has been developed by combining three mass analyzers: a quadrupole mass analyzer for precursor ion selection, a Thermo Scientific™ Orbitrap™ mass analyzer to acquire high dynamic range HRAM spectra, and the novel Thermo Scientific™ Astral™ mass analyzer to acquire high sensitivity, high dynamic range HRAM spectra at a rate of up to 200 Hz.
An integrated workflow for label-free quantitative proteomic studies, based on the Orbitrap Astral mass spectrometer, is presented. Two samples containing *E.coli*, HeLa (human) and *S.cerevisiae* (yeast) digests mixed in known ratios, were analyzed in DIA mode, and the results are presented in this poster.

## MATERIALS AND METHODS

Sample: Mix A: E.coli 36%, Human 46%, Yeast 18%;
Mix B: E.coli 18%, Human 46%, Yeast 36% (3 replicates/vial for both).
For the data presented in Fig.1-5 and 8-10:
- LC: Thermo Scientific™ Vanquish™ Neo UHPLC, Direct Injection mode, 500ng sample load, Thermo Scientific™ EASY-Spray NG Source, Thermo Scientific™ µPAC™ Neo 50 cm, 20 min gradient;
- Orbitrap Astral MS Acquisition method: OTMS1 240K, max inj. time 3ms, 200Hz Astral DIA, 2Th isolation window, max. inj. time 3ms, AGC 500%, NCE25.

For the data presented in Fig.6-7:
- LC: Vanquish Neo UHPLC, Trap and Elute mode, 50 to 500ng sample load, Thermo Scientific™ EASY-Spray™ NG Source, Thermo Scientific™ EASY-Spray™ PepMap™ 15cm x 150 µm; SPD/total run time: 180/8 min, 100/14.4 min, 60/24 min;
- Orbitrap Astral MS Acquisition method: OTMS1 240K, max. inj. time 3ms, 200Hz Astral DIA, 2Th isolation window, max. inj. time 3ms, AGC 500%, NCE25.

Data Processing:
- Biognosys Spectronaut™ 17.4, directDIA, normalization by Human FASTA, Q≤0.01 (1% FDR)
- Thermo Scientific™ Proteome Discoverer ™ 3.1 software with CHIMERYS™ 2.5.15 intelligent search algorithm, 1% FDR

## REFERENCES

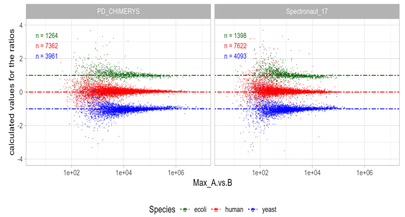
1. Ghaemmaghami, S., Huh, WK., Bower, K. *et al.* Global analysis of protein expression in yeast. *Nature* **425,** 737–741 (2003). https://doi.org/10.1038/nature02046

## TRADEMARKS/LICENSING



## RESULTS

**Figure 1. Orbitrap Astral MS detected log2 ratios (y-axis) distributed by the abundance (x-axis) for the 3-proteome mix samples A and B, processed with Proteome Discoverer 3.1 software using CHIMERYS and Spectronaut 17**




| | Species | Expected Log2 Ratio | Median Detected Log2 Ratio |
|---|---|---|---|
| CHIMERYS | ECOLI | 1 | 1.019 |
| CHIMERYS | HUMAN | 0 | 0.037 |
| CHIMERYS | YEAST | -1 | -1.023 |
| SN17 | ECOLI | 1 | 0.998 |
| SN17 | HUMAN | 0 | 0.025 |
| SN17 | YEAST | -1 | -1.028 |

Distribution of average log2 ratios of intensities (highest detected abundance among both conditions) of the identified (1% FDR) human, yeast, and *E.coli* proteins between the two 3-proteome mixtures, A and B, are presented in Fig.1. Data processed with the two search strategies, Spectronaut 17 and Proteome Discoverer 3.1 software with CHIMERYS, demonstrate very high correlation. The numbers of identified proteins are shown in the matching colors.

**Figure 2. Number of Proteins (1%FDR) quantified in 500 ng LFQ samples**

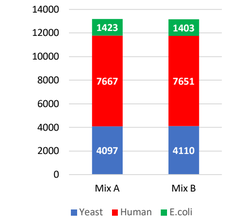
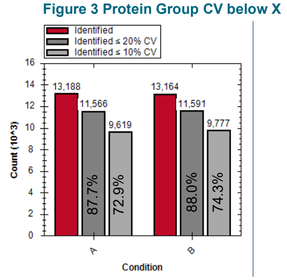



**Figure 3 Protein Group CV below X**




Over 13,000 proteins were identified and quantified across the 3 proteomes from 500ng of the mixed samples A and B, from triplicate runs (Fig.2). In addition to the deep proteome coverage, 87% of identified proteins had a CV less than 20% and 72% of identified proteins had a CV less than 10% (Fig.3), with a median CV 4.7% (Fig.4).

**Figure 4. Protein CV distribution per condition**

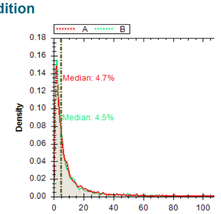
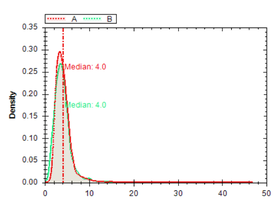
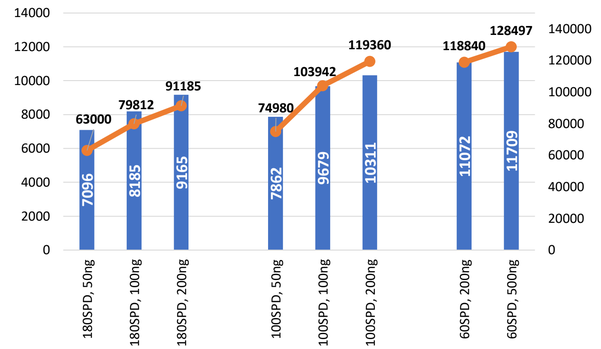


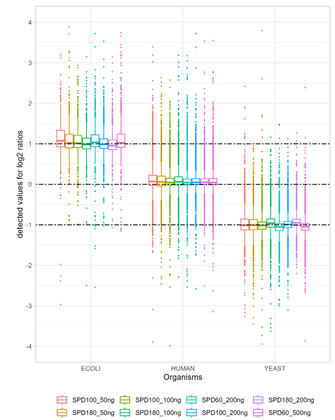

**Figure 5. Datapoints per peak distribution**




According to the Spectronaut 17 data, such a high quantitation accuracy was obtained with a median value of 4 datapoints per peak (Spectronaut defines peak width as 1.7* FWHM) (Fig.5). Increasing the DIA window size from 2Th to 3 and 4Th raised the number of SN datapoints to 5 and 6 per peak, yet this did not bring an improvement to the average quantitation accuracy (data not shown).

**Figure 6. Number of proteins and peptides identified and quantified with Orbitrap Astral MS for different sample loads and gradient lengths**




In Fig. 6, the unsurpassed quantitation performance of the novel Orbitrap Astral MS is demonstrated for a 3-proteome mixture with sample loads from 50 to 500 ng, for 3 different high throughput methods (180, 100 and 60 SPD). Quantitation accuracy was equally high throughout all studied conditions, showing little dependency on the sample load or throughput (Fig. 7).

**Figure 7. Quantitation accuracy of Orbitrap Astral MS LFQ data acquired for different sample loads and different throughput/gradient lengths**

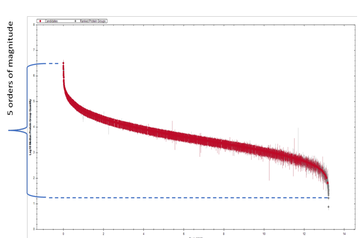
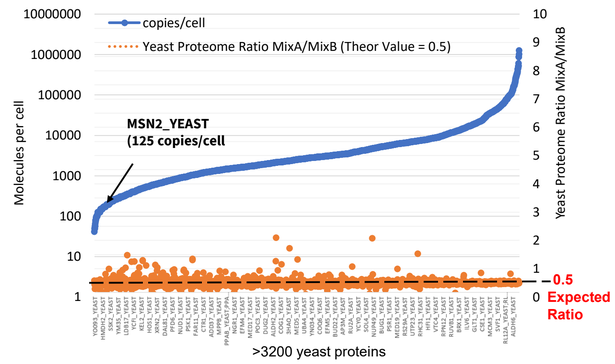



**Figure 8. Ranked Protein Groups**

Dynamic range for signal intensities for the identified proteins spans over 5 orders of magnitude. The ranked protein groups curve shown in this figure was obtained with 500ng load and 20 min gradient. For the 50ng load, the dynamic range is still close to 5 orders. With such a depth, proteins with expression levels as low as 100 copies per cell can be reliably quantified (Fig.9-10).

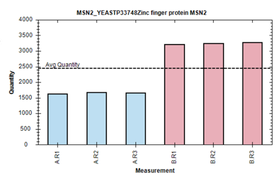
**Figure 9. LFQ accuracy for Yeast proteome over the protein expression scale**




For the Yeast proteome, the abundance of proteins ranges from fewer than 50 to more than 10(6) molecules per cell [1], [Figure 9]. With Orbitrap Astral MS, nearly the entire Yeast proteome could be reliably quantified using a 20-minute LC gradient and 500ng load of the mixed 3-proteome samples containing 18% or 36% Yeast. High quantification accuracy was observed throughout the entire expression range.

**Figure 10. An example of label-free quantitation for a low-abundant protein MSN2_YEAST (Zinc finger protein MSN2)**

Zinc finger protein MSN2 is a positive transcriptional factor that acts as a component of the stress responsive system. It recognizes and binds to the stress response element (STRE), which is involved in the response to various forms of stress (heat, oxidative, osmotic, etc.). MSN2 is present with 125 molecules/cell in log phase SD medium [1]. This protein was quantified with high accuracy and precision in all 6 runs using a 20-minute LC gradient with 500ng total sample containing either 18% or 36% yeast.




## CONCLUSIONS

The novel Orbitrap Astral mass spectrometer ensures ultra-fast 200 Hz data acquisition rates while delivering high sensitivity and dynamic range, spanning at least 5 orders of magnitude. With such performance, proteins with expression levels from 100s to 100,000s of copies per cell can be reliably quantified, opening new frontiers for life sciences research. For LFQ applications, accurate and precise quantitation using DIA is achieved with 6 datapoints per peak (mean value). Run-to-run precision of quantitation was high with less than 1% error, and quantitation based on single-peptide IDs was found to be as reliable as for 2 and more peptides. Equally high quantitation accuracy was demonstrated for sample loads from 50 to 500ng and LC run times from 8 to 24 minutes, showing little dependency on the sample load. Data processed with Spectronaut 17 and Proteome Discoverer 3.1 software with CHIMERYS demonstrated very high correlation.
PO002303-Anna-Pashkova-en

ThermoFisher SCIENTIFIC