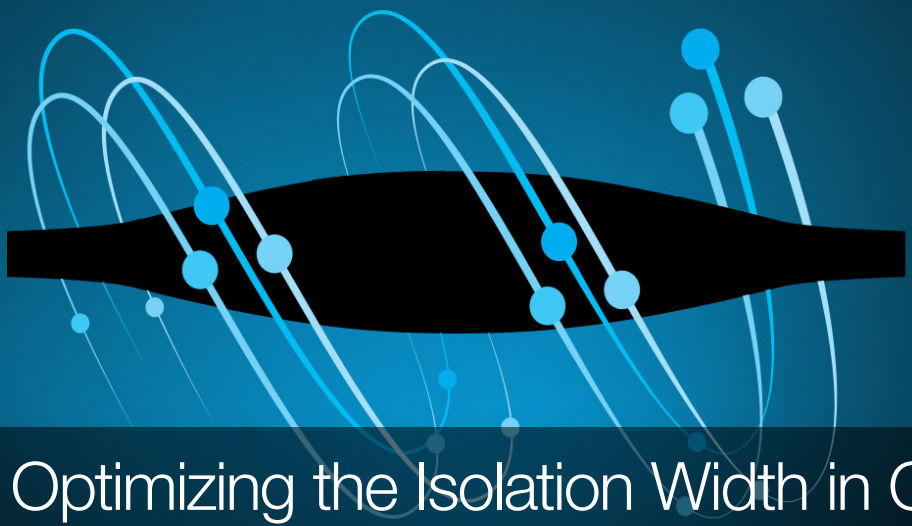
thermo scientific

POSTER NOTE 65545



Optimizing the Isolation Width in Orbitrap Instruments to Maximize the Number of Label-Free Quantified Peptides and Proteins

Authors

Carmen Paschke[1]; David Horn[2]; Tabiwang N. Arrey[1]; Rosa Rakownikow Jersie-Christensen[1]; Martin Zeller[1]; Romain Huguet[2]; Pedro Navarro[1], Bernard Delanghe[1];

[1]Thermo Fisher Scientific, Bremen, Germany; [2]Thermo Fisher Scientific, San Jose, CA

ABSTRACT

Purpose: Maximizing the number of quantified proteins in Label Free Quantification experiments.

Methods: Optimizing isolation width and using improved identification methods.

Results: Optimized method produces 50% more protein groups.

INTRODUCTION

In recent years, the performance of the filtering quadrupole in Orbitrap instruments has been improved. Isolation widths down to 0.4 Th are available. For reporter based quantification, smaller widths should be used to avoid interference from the co-isolated peptides. For Label Free Quantification (LFQ) that might not be the case. Here we will investigate the optimal isolation width for maximizing the number of identified and quantified proteins and peptides, together with the best possible accuracy and precision in LFQ experiments on Thermo Scientific™ Orbitrap Eclipse™ Tribrid™ MS and Thermo Scientific™ Orbitrap Exploris™ 480 mass spectrometers.

ThermoFisher SCIENTIFIC

MATERIALS AND METHODS

Sample Preparation and Data Acquisition

For the LFQ experiments different amount of yeast protein digest (Promega) was mixed in a constant HeLa protein digest (Pierce) background (200ng) to obtain a ratio of 2, 5 and 10. The samples were loaded onto Thermo Scientific™ EASY-Spray™ PepMap™ RSLC C18, 25 cm C18 column and separated with a 60 min gradient (8-30 % B [80% ACN in 0.1% FA] in 60 min, 5min to 50 %, another 5 min to 90 %B, 8 min at 90 % B). The eluting peptides were analyzed on the Orbitrap Exploris 480 mass spectrometer. The system was operated in a data dependent mode, selecting as many precursors as possible in 1 second cycle time. Isolation widths from 0.4 Th to 4 Th were used for MS2.

For the LFQ data using High-Field Asymmetric Waveform Ion Mobility (FAIMS), the Thermo Scientific™ FAIMS Pro™ interface was connected to an Orbitrap Eclipse Tribrid mass spectrometer. A similar sample and chromatography was used as described above. The isolation width was set to 3 Th, MS2 fragments were measured in the Orbitrap analyzerr.

Figure 1. Processing workflow

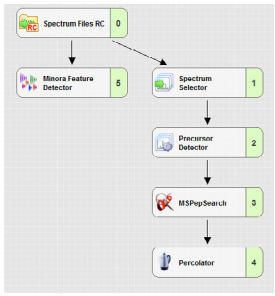
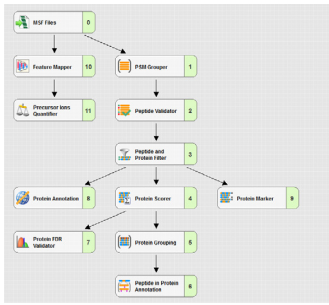


Figure 2. Consensus workflow



Data Analysis

Data analysis was performed using a beta version of Thermo Scientific™ Proteome Discoverer™ 2.4 software. To analyze the chimeric spectra, a new node, Precursor Detector, was developed. This node detects all precursor masses that are co-isolated with the targeted precursor and consequently fragmented together to produce the chimeric MS2 spectrum. Each precursor was then searched with that MS2 spectrum either using SEQUEST HT or MSPepSearch search engines. FASTA files used for SEQUEST HT: Homo sapiens (SwissProt TaxID=9606) (v2017-10-25) and Saccharomyces cerevisiae (SwissProt TaxID=4932_and_subtaxonomies) (v2017-10-25).

The Human and Yeast spectral libraies used with MSPepSearch were predicted using PROSIT (2) with the following settings: sequence length 7 to 30, precursor charge 2 and 3, up to 2 missed cleavages, up to 2 M(ox) per peptide.

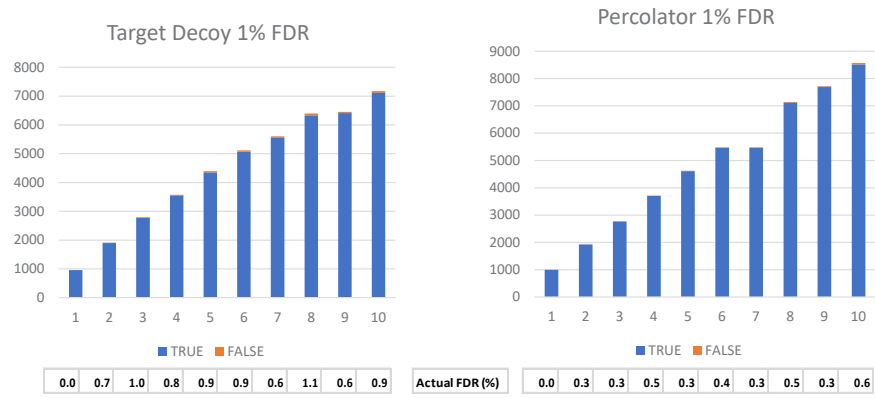
All data were filtered at 1% FDR on psm, peptide and protein level.

RESULTS

Validation of the SEQUEST HT results

To validate the identifications of the chimeric spectra, artificial chimeric spectra were created in silico by combining the mass peaks from known synthetic peptides from the Proteome Tools project (1). One to up to 10 of those spectra were merged into one and 1000 of each of those spectra were searched using SEQUEST HT. The validation was performed with a classical target decoy approach or using Percolator. With the known peptide sequences in each spectrum the real FDR was calculated by dividing the false positive identifications by all identifications (see figure 3). As expected, Percolator outperformsg the Target Decoy approach in number of identified peptides. The identification rate decreases with increasing multiplexing rate but is nevertheless superior to 85% using Percolator with FDR well below 1 %.

Figure 3. Number of identified multiplexed spectra versus multiplexing rate for SEQUEST HT



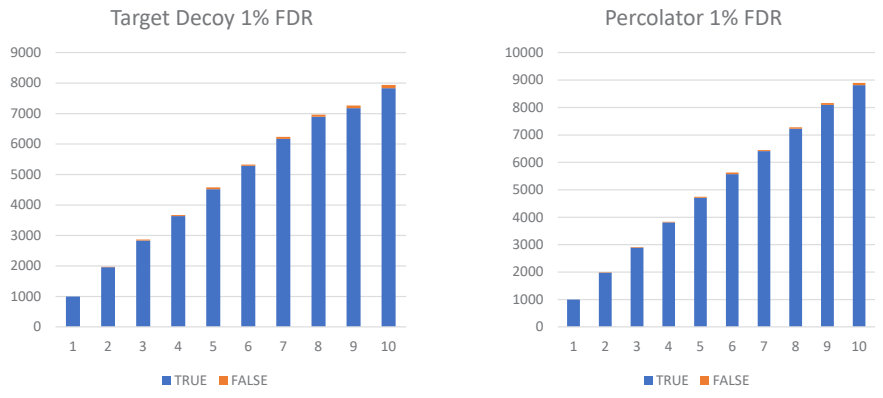
0.0	0.7	1.0	0.8	0.9	0.9	0.6	1.1	0.6	0.9	Actual FDR (%)
0.0	0.3	0.3	0.5	0.3	0.4	0.3	0.5	0.3	0.6	

The x axis displays the number of multiplexed spectra per spectrum. SEQUEST HT shows excellent recovery rates, from 96% to 85% for 1 up to 10 multiplexed MS2 spectra with real FDR below 1%. As expected, Percolator outperforms Target Decoy FDR calculation.

Validation of the MSPepSearch results

The same approach as for SEQUEST HT (see above) has been used for MSPepSearch .

Figure 4. Number of identified multiplexed spectra versus multiplexing rate for MSPepSearch



0.0	0.8	1.0	0.9	0.9	0.9	1.0	1.0	0.8	0.9	Actual FDR (%)
0.0	0.8	0.9	0.5	0.8	1.1	0.6	0.7	0.8	0.9	

MSPepSearch showsexcellent recovery rates, from 96% to 88% for 1 up to 10 multiplexed MS2 spectra with real FDR below 1%. As expected, Percolator also outperforms Target Decoy FDR calculation.

Optimal isolation width for maximal peptide and protein identification

Figure 5. Identified Protein Groups at 1% FDR for the different isolation widths

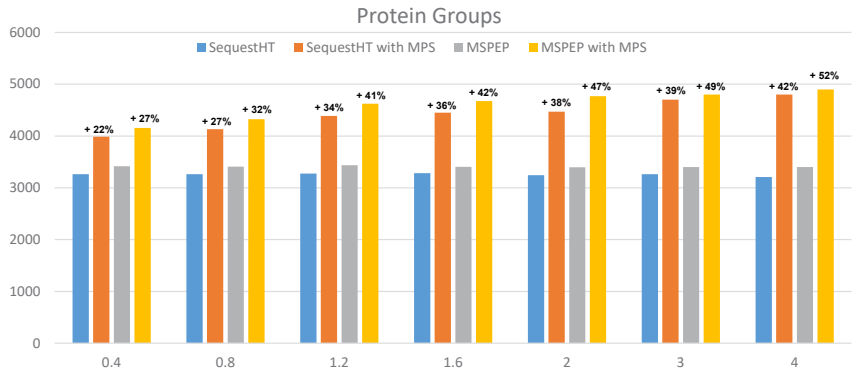
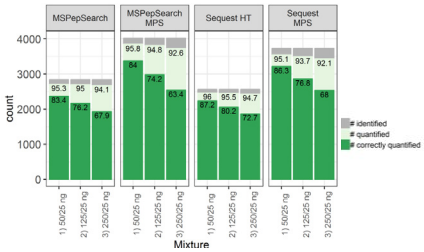


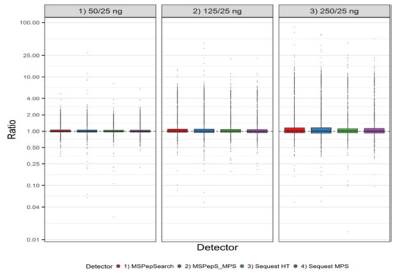
Figure 5 shows the number of identified protein groups in a mixture 200ng Hela + 25ng Yeast digest using different insolation width. Peptide identifications was done using SequestHT and MSPepsearch with and without multiple peptide search (MPS). Around 50 % more protein groups are identified with the chimeric search maximizing around 3 to 4 Th. Other samples are showing similar increases (data not shown).
Precision and accuracy of the LFQ with and without MPS are shown in figures 6-9.

Figure 6. Number of quantified Human protein groups at isolation width 2 Th for the different identification methods



Number of quantified and correctly (< 20% of expected ratio) quantified protein groups. MSPepSearch with MPS outperforms the rest

Figure 7. Precision and accuracy of the quantified Human protein groups for the different identification methods at isolation width 2 Th



Accuracy of the ratios is excellent for all the isolation widths, however the precision is improving for larger isolation widths

Figure 8. Precision and accuracy of the quantified Yeast protein groups for the different identification methods at isolation width 0.4 Th

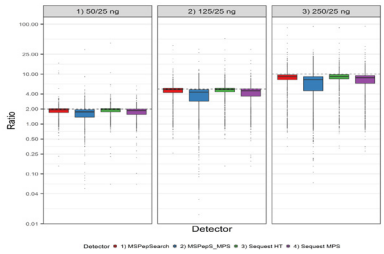
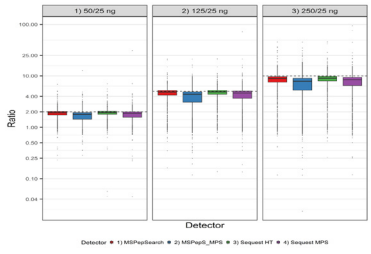


Figure 9. Precision and accuracy of the quantified Yeast protein groups for the different identification methods at isolation width 2 Th



Precision increases with greater isolation width; this is due to the fact that more peptides are identified. The accuracy slightly decreseswith MPS and increased isolation width, mainly due to the fact that the majority of additional proteins are of lower abundance (see Figures 10. and 11.)

All the samples measured with the different isolation widths have been analysed with SEQUEST HT and MSPepSearch, without and with searching for chimeric of multiple precursors (MPS). All validation was performed with Percolator. The number of identified protein groups increases with increasing isolation width, maximizing around 3 to 4 Th. MSPepSearch with MPS, in all cases, produces the most psms, peptide groups and protein groups (all at 1% FDR) increasing the number of protein groups by more then 50 % (figure 5.). Precision increases with increasing isolation width. This is due to the act that more peptides are identified (Figure 8. and 9.). MPS is slightly decreasing the precision due to the fact that more low abundant peptides and proteins are detected (Figure 10. and 11.). Accuracy for high abundant proteins is excellent for all data (Figure 7.) For lower abundant proteins (additional proteins identified with MPS) this is slightly worse (Figure 8., 9., 10. and 11.).

Figure 10. Protein ratio distribution at isolation width 0.4 Th for the different identification methods

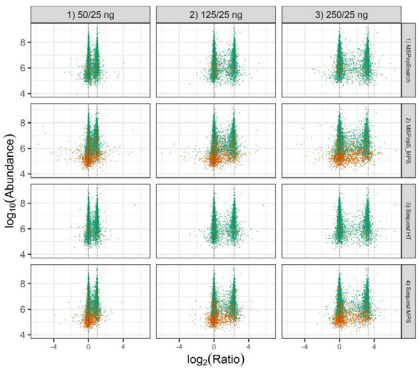
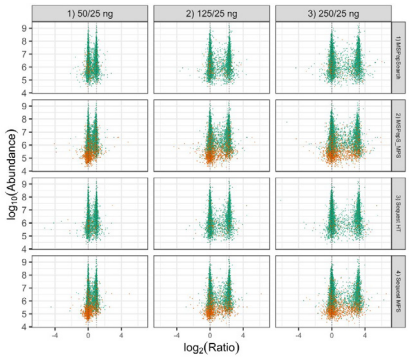


Figure 11. Protein ratio distribution at isolation width 2 Th for the different identification methods

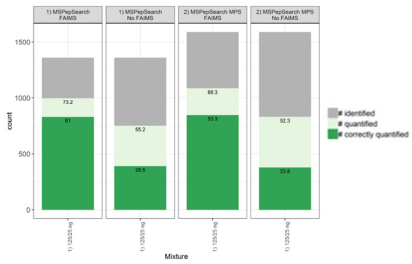


MPS is identifying and quantifying more low abundant peptides and proteins. Common between with and without MPS are marked in green, unique ones in organge.

Using FAIMS

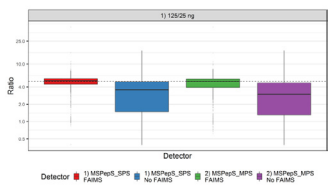
Adding the FAIMS Pro interface in front of the Orbitrap Eclipse Tribrid MS dramatically improves the precision (figures 12 and 13) and accuracy of the quantified peptides and proteins. In addition, that also increases the number of Identified proteins (Figure 12. and Figure 13.)

Figure 12. Number of quantified Yeast protein groups at isolation width 3 Th for the different identification methods



FAIMS improves the number of identified and quantified protein groups.

Figure 13. Precision and accuracy of the quantified Yeast protein groups for the different identification methods at isolation width 3 Th and ratio of 5

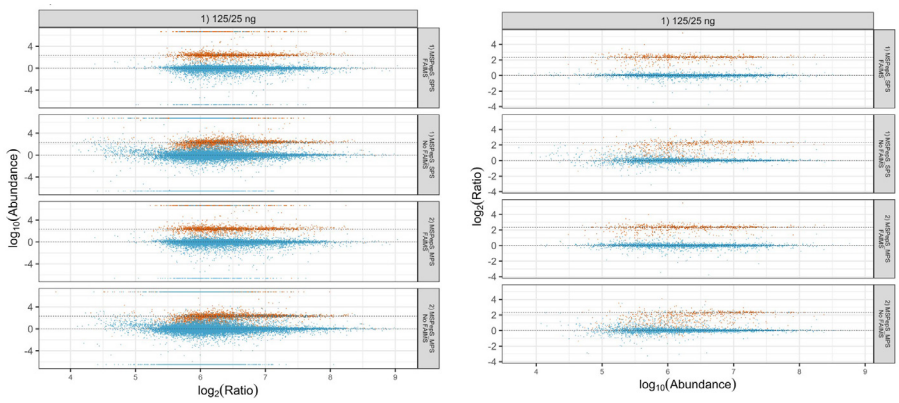


Using FAIMS improves greatly the accuracy and precision, especially for lower abundant proteins.

thermo scientific

Figure 14. Peptide ratio distribution at isolation width 3 Th for the different identification methods

Figure 15. Protein ratio distribution at isolation width 3 Th for the different identification methods



Adding FAIMS in front of the Orbitrap Eclipse Tribrid improves the accuracy and precision of the peptide and protein ratios .

CONCLUSIONS

- Optimal isolation width for Label Free Quantification is around 3 to 4 Th.
- Up to 50% more protein groups can be identified searching for multiple peptides per spectrum using SEQUEST HT as well as MSPepSearch.
- Using FAIMS dramatically increases the accuracy and precision of quantified protein groups. FAIMS also increases the number of quantified protein groups

REFERENCES

1. D. Zolg et al, Nature Methods, 14, 259 (2017): “Building ProteomeTools based on a complete synthetic human proteome”.
2. Gessulat & Schmidt et al. Nature Methods (2019): “ Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning”.

ACKNOWLEDGEMENTS

We would like to thank Dr Mathias Wilhelm and Siegfried Gessulat from the Technical University of Munich for supplying the PROSIT predicted spectral libraries for Human and Yeast.

Find out more at **thermofisher.com/Orbitrap**

ThermoFisher SCIENTIFIC