HUPO CONNECT 2020

A New Algorithm for FAIMS Data Analysis with Accurate In-depth Quantitative Profiling

Zia Rahman; Yandong Zhu; Zheng Chen; Jonathan R. Krieger. Baozhen Shan

Bioinformatics Solutions Inc., Waterloo, ON

EXPANDED ABSTRACT

Peptide identification by data-dependant mass spectrometry relies on the sequential selection of precursor ions as they elute from the chromatographic column. Complex samples represent a challenge due to the presence and occurrence of chimeric tandem mass spectra from co-eluting ions. Gas-phase ion fractionation opens new perspectives to overcome sample complexity. High-field asymmetric waveform ion mobility spectrometry (FAIMS) provides an important resource to combat this issue, however, quantitative analysis of FAIMS data remains a challenge. Here we present a new algorithm, including 4-dimensional feature detection and alignment for FAIMS data analysis. Preliminary testing indicates that algorithm enabled deep proteome coverage with high accuracy.

The algorithm contains four components:

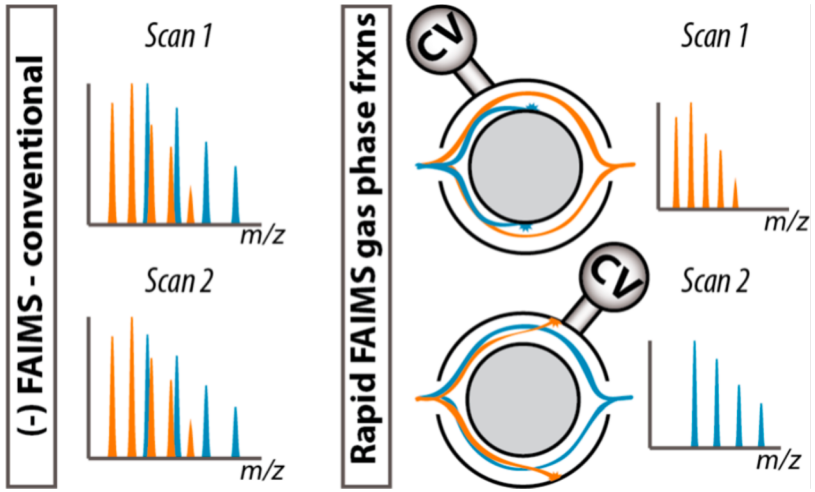
- FAIMS holds the promise of delivering gas-phase fractionation by stepping through multiple compensation voltages (CVs), deconvoluting overlapping peptide signals, and removing contaminants. Thus, peptide precursor features are detected by grouping each CV from the run together allowing for higher precursor resolution
- Based on precursor m/z, retention time, and CV, multiple features are associated with each MS/MS spectrum. This association further deconvolutes overlapping precursors for in-depth identification
- Label free quantification is achieved by performing 4D-feature alignment across different LC-MS/MS runs. Peptide intensity is calculated as summation of values in all CVs.
- TMT quantification is achieved by excluding chimeric spectra, and by also separating MS/MS spectra from distinct CV values resulting in higher accuracy quantification.

The algorithm was implemented in PEAKS XPro software platform and tested with multiple published available data sets. With one dataset(PXD009547), the new algorithm suggested that FAIMS aids in reducing the percentage of chimeric spectra from 39% to 25%. This reduction in chimerism leads to significant benefits in improving proteome coverage. A 36% gain in the number of protein identifications (2 unique peptides) compared with traditional LC-MS/MS analyses for the same number of injections. The reduction in complexity also has improving effects on quantification. With a second dataset, an increase of 1100 protein groups were quantified using FAIMS, with noticeable increases in feature detection and alignment correlations, and decreased in standard deviation across injection replicates. Together, this suggests that the new PEAKS algorithm provides accurate and sensitive quantitative analysis with FAIMS data.

Bioinformatics Solutions Inc
470 Weber St. North. Waterloo, ON
N2L 6J2, CANADA

www.bioinfor.com
@PEAKSProteomics
t. 1.855.885.8288
peaks@bioinfor.com

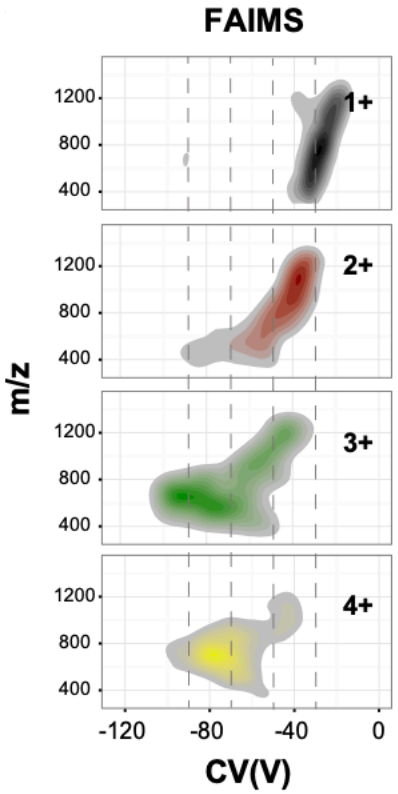
INTRODUCTION



Hebert, A. S., Prasad, S., Belford, M. W., Bailey, D. J., McAlister, G. C., Abbatiello, S. E., et al. (2018). Comprehensive Single-Shot Proteomics with FAIMS on a Hybrid Orbitrap Mass Spectrometer. *Analytical Chemistry*.

Principles of FAIMS

- FAIMS (high-Field Asymmetric waveform Ion Mobility Spectrometry) allows for the separation of ions based on their mobility in gas phase
- Ions get trapped by a carrier gas between two parallel electrodes. Adding a high frequency waveform (DV) to the inner electrode while the outer electrode remains grounded causes ions to be trapped



- A compensation voltage (CV) is added to the electrode to counteract the DV allowing ions to be transmitted through the FAIMS device into the mass spectrometer (MS)
- Different CVs will cause different ions to move through the FAIMS device into the MS.
- Ions of different charge states will distribute different as a function of CV
- This represents a second separation step to simplify the peptide mixture entering the mass spectrometer to improve detection

Cooper, H. J. (2016). To What Extent is FAIMS Beneficial in the Analysis of Proteins? *Journal of the American Society for Mass Spectrometry*.

Purves, R. W., & Guevremont, R. (1999). Electrospray Ionization High-Field Asymmetric Waveform Ion Mobility Spectrometry−Mass Spectrometry. *ACS Publications*

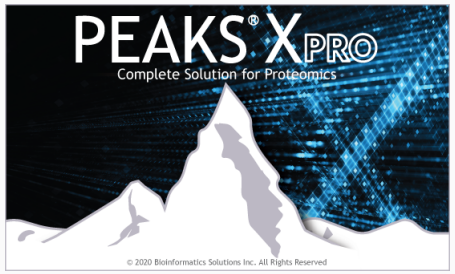
Pfammatter, S., Bonneil, E., McManus, F. P., Prasad, S., Bailey, D. J., Belford, M., et al. (2018). A Novel Differential Ion Mobility Device Expands the Depth of Proteome Coverage and the Sensitivity of Multiplex Proteomic Measurements. *Molecular & Cellular Proteomics : MCP*, *17*(10), 2051–2067.

Bioinformatics Solutions Inc
470 Weber St. North. Waterloo, ON
N2L 6J2, CANADA

PEAKS Complete Software for Proteomics | BSi | PEAKS Online Complete Multi-User Proteomics Solution

www.bioinfor.com
@PEAKSProteomics
t. 1.855.885.8288
peaks@bioinfor.com

PEAKS SOLUTIONS FOR 4D FAIMS-LC-MS/MS DATA



PEAKS® Online
Complete Multi-User Proteomics Solution

- PEAKS X PRO Studio and PEAKS Online are complete proteomics software solutions that now includes full support for FAIMS 4D Proteomics Projects, including:
  - FAIMS DDA
  - FAIMS DIA*
  - FAIMS SILAC
  - FAIMS TMT
  - Multiple CV values in a single run

*direct DIA in Studio, Library search in both Studio and Online/

Both versions if PEAKS employ a new algorithm specifically geared towards FAIMS data.

FAIMS Algorithm Feature	Application/Function
Individual CVs are grouped together	higher precursor resolution and better feature detection
Multiple overlapping feature associations per MS/MS spectrum	increased in-depth protein identification
Performs 4D feature alignment across different LC-MS/MS Runs	peptide intensity is calculated as a summation of values in all CVs for better and more accurate quantification
Specific to TMT, chimeric spectra are excluded and MS/MS spectra from distinct CV values are separated	higher quantification accuracy

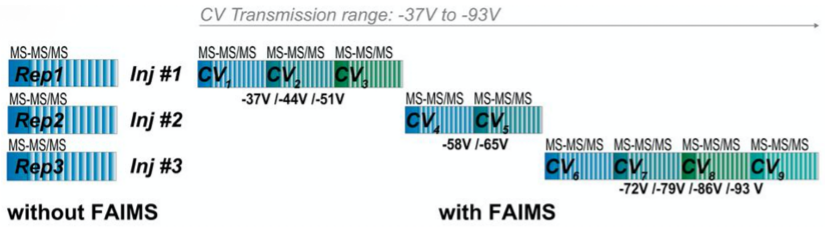
Bioinformatics Solutions Inc
470 Weber St. North. Waterloo, ON
N2L 6J2, CANADA

PEAKS® Complete Software for Proteomics
BSi
PEAKS® Online Complete Multi-User Proteomics Solution

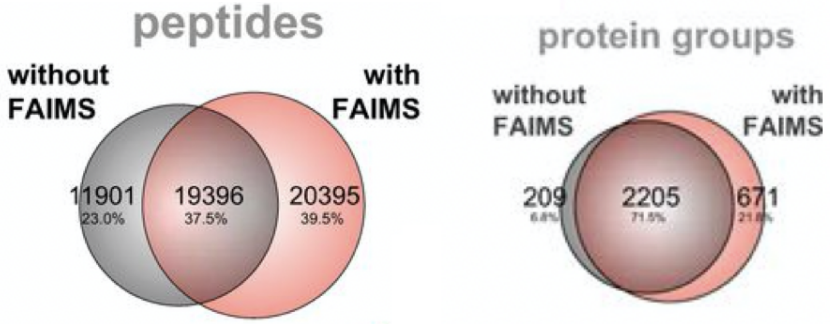
www.bioinfor.com
@PEAKSProteomics
t. 1.855.885.8288
peaks@bioinfor.com

IMPROVEMENTS IN ANALYSIS OF FAIMS vs. NON-FAIMS DATA

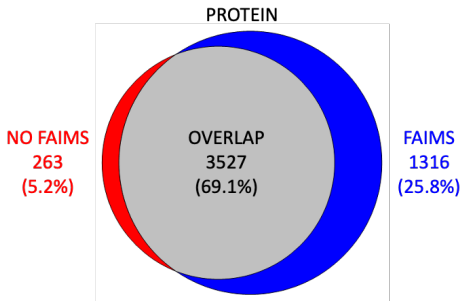
Using data from Pfammatter S. *et al*[1], which specifically compared non-FAIMS to FAIMS data and used PEAKS 8.5 for analysis, we sought to confirm and show how improvements in the algorithm have led to a marked increase in protein identifications

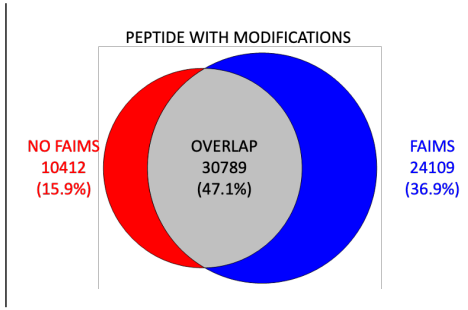


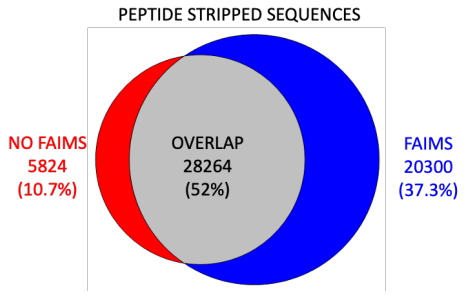
Experiment design by the study's authors show that each sample was either injected 3 times with no FAIMS or 3 times with various CVs

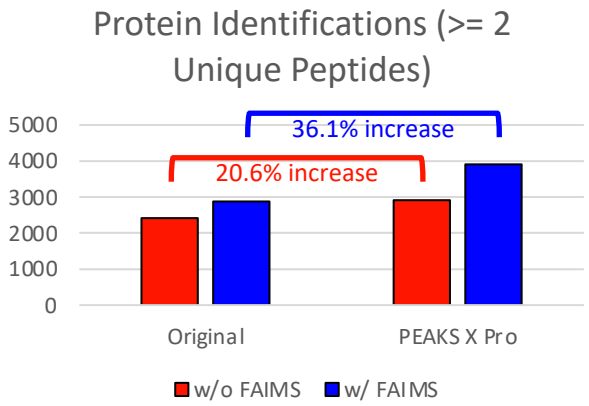


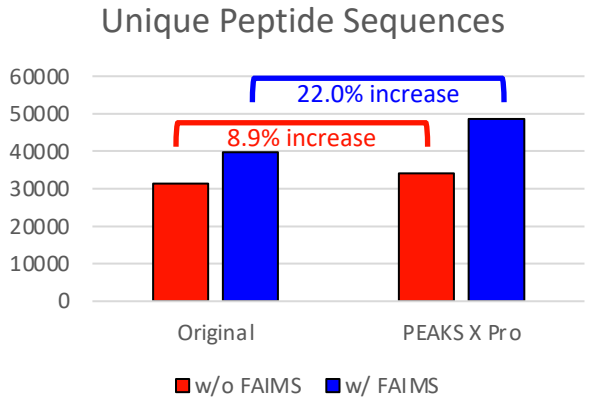
Peptide and protein group numbers from original publication (searched with PEAKS 8.5)











Re-analyzing this data set with the new PEAKS X PRO shows substantial gains in protein identifications for data generated both with and without FAIMS, but specifically the largest gains in sensitivity and identifications of both peptides and proteins comes with data generated using FAIMS.

[1]Pfammatter, S., Bonneil, E., McManus, F. P., Prasad, S., Bailey, D. J., Belford, M., et al. (2018). A Novel Differential Ion Mobility Device Expands the Depth of Proteome Coverage and the Sensitivity of Multiplex Proteomic Measurements. *Molecular & Cellular Proteomics : MCP*, *17*(10), 2051–2067.

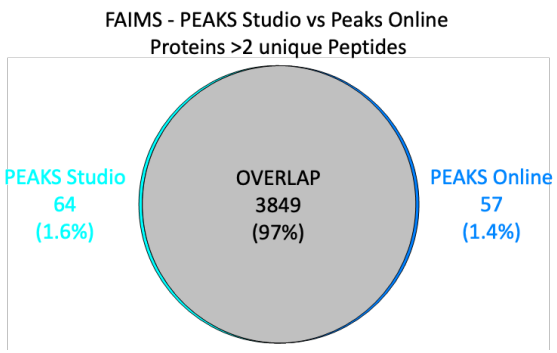
Bioinformatics Solutions Inc
470 Weber St. North. Waterloo, ON
N2L 6J2, CANADA

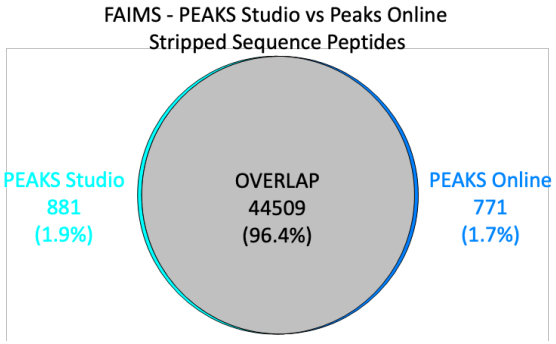
PEAKS Complete Software for Proteomics | BSi | PEAKS Online Complete Multi-User Proteomics Solution

www.bioinfor.com
@PEAKSProteomics
t. 1.855.885.8288
peaks@bioinfor.com

PEAKS STUDIO X PRO and PEAKS ONLINE offer Consistent Results

In order to assess whether there are differences in the identification of peptides and proteins between PEAKS Studio X Pro and PEAKS Online (v.1.4), we carried out a direct comparison using the FAIMS triplicate injections of HEK293T cells with multiple CV values run in DDA mode from the same dataset.

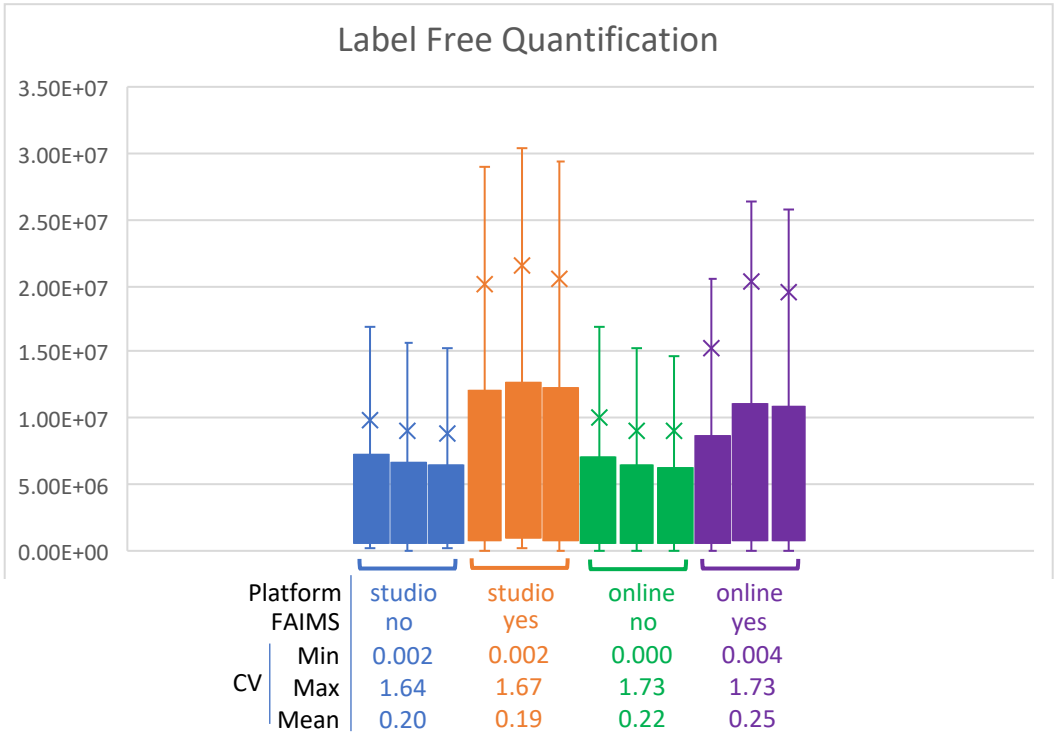




As illustrated in the Venn Diagrams, we observe >95% overlap between PEAKS Studio X Pro and PEAKS Online with DDA data.

CV and LFQ ANLYSIS

Further examining the same dataset, proteins identified in all replicate injections were quantified. CV values remain low both with and without FAIMS, across 3 replicates in each condition. However, quantification across the FAIMS data shows a broader distribution, likely due to several factors including more proteins quantified; more accurate quantification; the reduction in chimeric spectra while using FAIMS, and improvement of sensitivity including minimizing ion suppression. With the same dataset between PEAKS Studio X Pro and PEAKS Online, we obtain very similar results of label free quantification distribution



Bioinformatics Solutions Inc
470 Weber St. North. Waterloo, ON
N2L 6J2, CANADA

PEAKS Complete Software for Proteomics

BSi

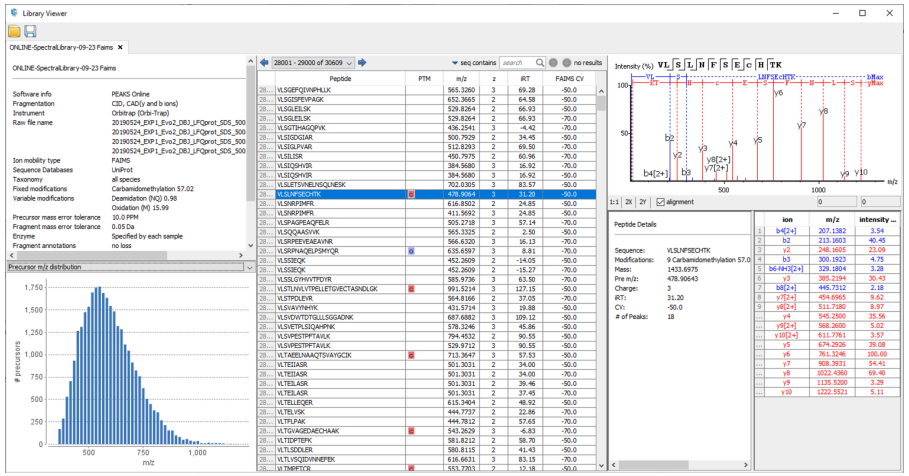
PEAKS Online Complete Multi-User Proteomics Solution

Min: 0.002
Max: 1.67
Mean: 0.19

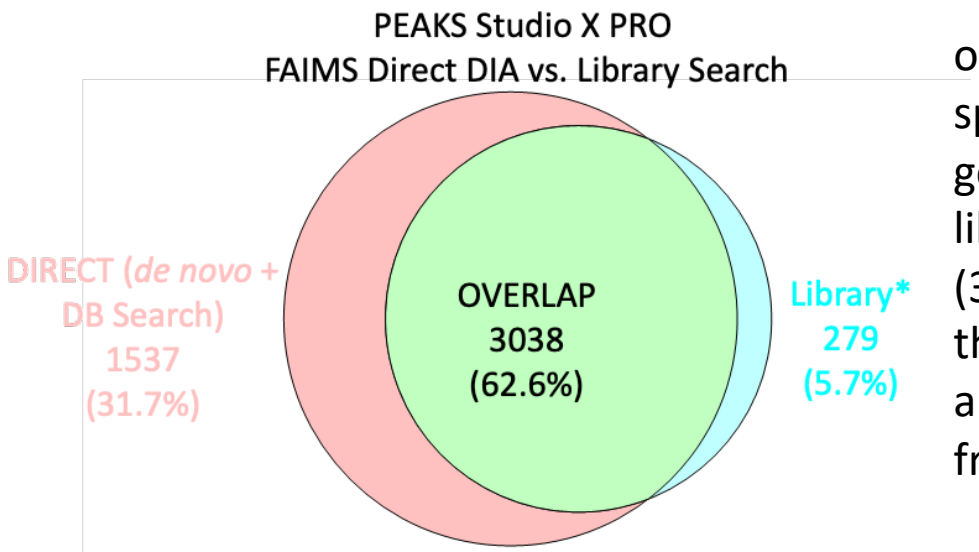
www.bioinfor.com
@PEAKSProteomics
t. 1.855.885.8288
peaks@bioinfor.com

PEAKS and FAIMS DIA DATA

To assess whether this trend holds true for DIA FAIMS data, we used a Hela digest run with various DIA windows and FAIMS settings of -50 and -70 CV. This data was searched against a library generated from a small DDA Hela FAIMS dataset. PEAKS Studio X Pro has the option to analyze FAIMS LC-MS/MS DIA data either by library search or by direct DIA analysis, which is based on our *de novo* algorithm.



New to PEAKS Studio X Pro is a library viewer that will allow the user to explore the library. PEAKS Studio X Pro accepts libraries generated with PEAKS Studio, PEAKS Online, OpenMS, and Spectronaut



* There is a limited amount of FAIMS data at the specified CV Values to generate a library with. This library used was quite small (30609 entries, explaining the disparity, or the large amount of identifications from the direct DIA Search.

A more extensive database would reduce this and provide more overlap. Regardless, this data shows that PEAKS Studio X Pro has the capability to identify and quantify proteins with and without a spectral library.

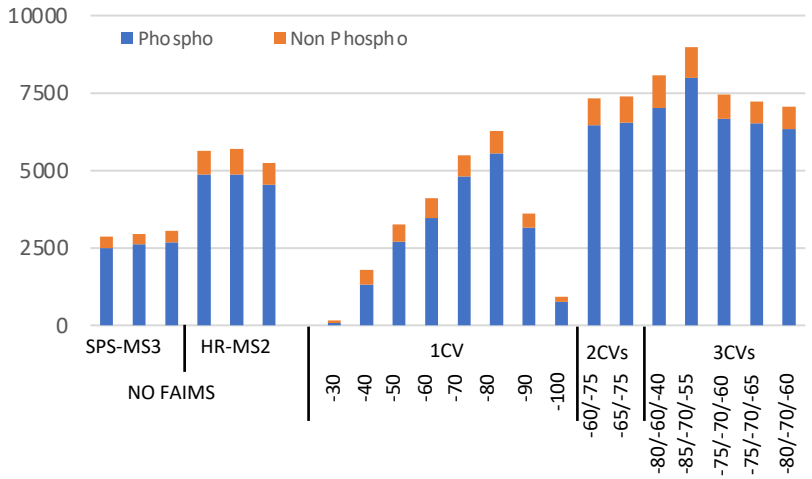
A manual search of randomly selected unique peptides from these 1537 proteins not identified via library search revealed that these peptides were not in fact in the spectral library.

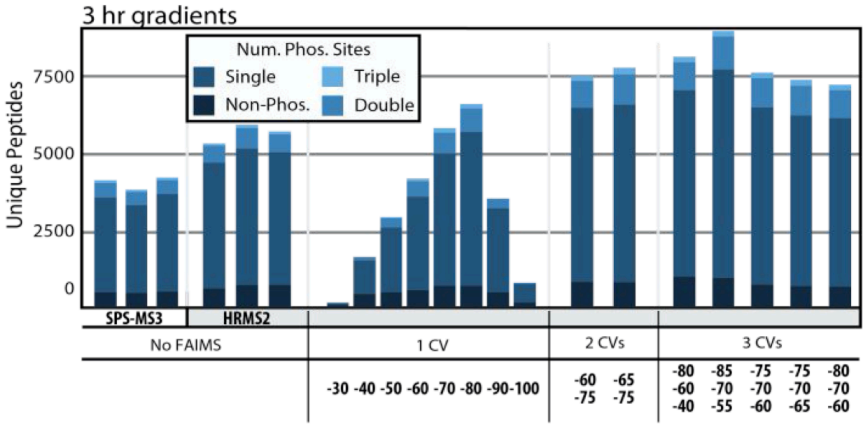
*Library data build with 4 DDA files (CV50 and CV70) from: Bekker-Jensen, D. B., Martínez-Val, A., Steigerwald, S., Rüther, P., Fort, K. L., Arrey, T. N., et al. (2020). A Compact Quadrupole-Orbitrap Mass Spectrometer with FAIMS Interface Improves Proteome Coverage in Short LC Gradients. *Molecular & Cellular Proteomics : MCP*, *19*(4), 716–729.

Bioinformatics Solutions Inc
470 Weber St. North. Waterloo, ON
N2L 6J2, CANADA

www.bioinfor.com
@PEAKSProteomics
t. 1.855.885.8288
peaks@bioinfor.com

TMT, Phosphorylation, and FAIMS with PEAKS





- With the clear beneficial advantages of FAIMS in reducing sample complexity and increasing identifications, its use and presence in the field will become more pronounced. Furthermore, due the increased throughput of isobaric labelling by Tandem Mass Tags, we have introduced the ability to do TMT quantification with the PEAKS line of products.
- In order to assess the ability of the PEAKS tools to identify and quantify phospho-peptides with isobaric tags, we took a published dataset from Schweppe *et ai*, and compared the numbers of unique peptide identifications across a number of conditions (Synchronous Precursor Selection MS3 (SPS-MS3) without FAIMS, High Resolution MS2 (HR-MS2) without FAIMS, or HR-MS2 with various FAIMS CVs as indicated.
- We observe results that benchmark extremely well against this published dataset.
- Isobaric labels can then be used for accurate quantification

Schweppe, D. K., Rusin, S. F., Gygi, S. P., & Paulo, J. A. (2019). Optimized Workflow for Multiplexed Phosphorylation Analysis of TMT-Labeled Peptides Using High-Field Asymmetric Waveform Ion Mobility Spectrometry. *Journal of Proteome Research*, *19*(1), 554–560.

Bioinformatics Solutions Inc
470 Weber St. North. Waterloo, ON
N2L 6J2, CANADA

PEAKS Complete Software for Proteomics
BSi
PEAKS Online Complete Multi-User Proteomics Solution

www.bioinfor.com
@PEAKSProteomics
t. 1.855.885.8288
peaks@bioinfor.com

Summary

- PEAKS Studio X Pro and PEAKS Online v1.4 can represent a new software solution for analyzing FAIMS Pro Data
- All types of FAIMS data can be analyzed, including:
  - DDA, DIA, TMT, SILAC
- Multiple CVs in same files can be used directly with PEAKS without the need for conversion or splitting data files
  - This also means that individual CV values do not need to be run independently of each other. CV switching on the FAIMS Pro device is relatively quick, and thus a combination approach has been shown to be the best
- PEAKS Studio X Pro can can directly identify peptides and proteins from DIA FAIMS experiments if a library is not available. This is done accurately using our well known *de novo* sequencing algorithm
- For a demonstration or a free trial, feel free to visit us or contact us (peaks@bioinfor.com)

Bioinformatics Solutions Inc
470 Weber St. North. Waterloo, ON
N2L 6J2, CANADA

PEAKS Complete Software for Proteomics | BSi | PEAKS Online Complete Multi-User Proteomics Solution

www.bioinfor.com
@PEAKSProteomics
t. 1.855.885.8288
peaks@bioinfor.com