BRUKER AT ASMS 2021 – ADVANCES IN 4D-PROTEOMICS™

Label-free quantification of oxidized peptides in eHAP cell lines via a high-throughput dia-PASEF workflow

Romano Hebeler [1], Renata Blatnik [1], Philipp Strohmidel [1],
Hamish Pegg [2], Joanna Kirkpatrick [2]

[1] Bruker Daltonics GmbH & Co. KG, Bremen, Germany
[2] The Francis Crick Institute, London, United Kingdom

Introduction

- dia-PASEF® (Meier et al., 2020; 1) merges the benefits of data-independent acquisition (DIA) with the advantages of ion mobility in proteomics experiments (Meier et al., 2018; 2). The ion mobility dimension improves the alignment of precursor and fragment spectra. In this study we applied a dia-PASEF workflow in combination with an Evosep One (Evosep, Bache et al., 2018, 3) chromatographic system for high-throughput analysis of eHAP cell line digests, while quantifying spiked-in, oxidized target peptides in Spectronaut™ (Biognosys, Bruderer et al., 2015, 4) on different gradient lengths.

1: https://pubmed.ncbi.nlm.nih.gov/33257825/
2: https://www.biorxiv.org/content/10.1101/336743v1.full
3: https://www.mcponline.org/content/17/11/2284
4: https://pubmed.ncbi.nlm.nih.gov/25724911/

Methods: chromatographic conditions and dda-PASEF method

- The PreOmics NHS iST Kit was used to digest eHAP cells and 15 oxidized target peptides were spiked in at concentration levels corresponding to 0, 25 and 100 fmol/µg.
- Chromatographic separation of tryptic peptides was done using an Evosep One (Evosep). Standardized methods from the Evosep for 30, 60, 100 and 200 Samples Per Day (SPD), see *Table 1*, were used for library generation with a dda-PASEF method. The same LC methods were used for the subsequent dia-PASEF measurements.
- Depending on the chromatographic method the loading on the Evotips™ was 20, 50 or 200 ng (1, 2.5 and 10 µl in 20 µl total volume of 0.1 % FA in water).
- The LC was coupled to a timsTOF Pro (Bruker) mass spectrometer via the Captive Spray source (Bruker) equipped with a Standard Captive Spray Emitter Tip (Bruker).The library generation was done with a dda-PASEF method optimized for short chromatographic separation times ("*DDA PASEF-short_gradient_0.5sec_cycletime.m*" see *Table 2* and *Figure 1*).

Throughput (samples/day)	Cycle Time (min)	Gradient Length (min)	Flow rate (µl/min)	Column dimensions (length/ID/bead size)	Sample load on column
200	7.2	5.6	2	4cm / 150µm / 3µm	20 ng
100	14.4	11.5	1.5	8cm / 100µm / 3µm	20 ng / 50 ng
60	24	21	1	8cm / 100µm / 3µm	50 ng / 200 ng
30	48	44	0.5	15cm / 100µm / 3µm	200 ng

Table 1: chromatographic conditions used on the Evosep One system

Source settings:	
Capillary	1.4 kV
Dry Gas	3.0 L/min
Dry Temperature	180 °C
MS settings:	
Mass Range	100 – 1,700 m/z
Ion Mobility Range	0.85 - 1.30 V-s/cm2
MS/MS settings:	
Precursor ions	Total Cycle Time: 0.5 sec
# of MS/MS Ramps	4 PASEF scan à 100 ms
Spectra Rate	> 100 Hz

Table 2: dda-PASEF method

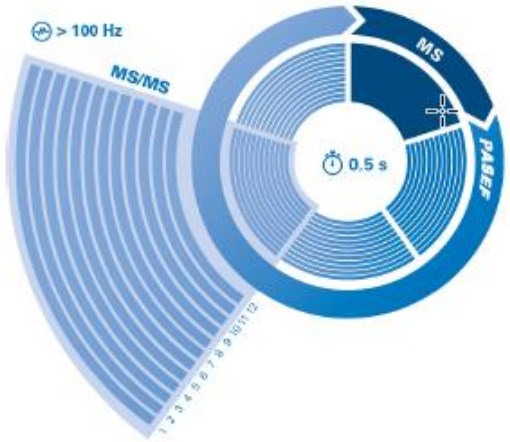

Figure 1: dda-PASEF cycle

Methods: dia-PASEF method and database search parameters

- Subsequent dia-PASEF measurements were done with an acquisition scheme covering the mass and mobility range of the multiply charged peptide ions in the heatmap, refer to *Table 3 and Figure 2*.
- All data processing steps were done using Spectronaut (Biognosys). Two different versions of the software package were used and results compared with each other:
  1. Version 14.3.200707 named as "v14" and
  2. Version 15.4.210913 named as "v15" in the following.
- For the library generation built-in standard settings were used. The database searches were carried out against two protein sequence databases (refer to *Table 4*):
  1. one containing the 15 spiked-in oxidized peptides and
  2. Human SwissProt (with a total of 36969 entries).
- For the matching of the dia-PASEF data against the libraries also build-in default processing parameters were used.
- For the directDIA™ searches the same database search settings were used.

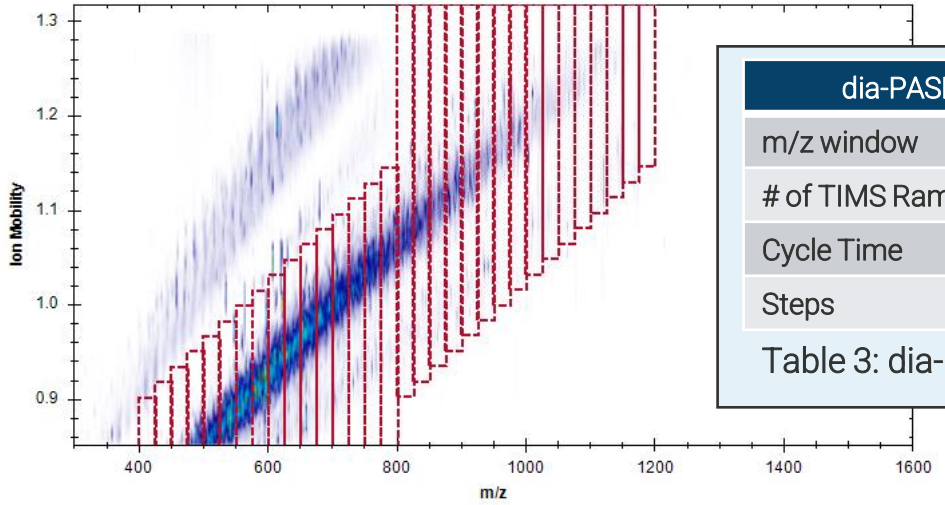

Figure 2: dia-PASEF acquisition scheme

dia-PASEF settings:	
m/z window	25 Da
# of TIMS Ramps	16
Cycle Time	1.8 sec
Steps	2

Table 3: dia-PASEF method

Database Search Settings	
Enzyme	Trypsin/P
Digest Type	Specific
Max Peptide Length	52
Min Peptide Length	5
Missed Cleavages	2
Max variable Modifications	5
Fixed Modifications	Cys Novartis
Variable Modifications	Acetyl (Protein N-term), Oxidation (K), Oxidation (M), Oxidation (P)

Table 4: Database Search Settings

Results: library generation with Spectronaut “v14” and “v15”

v14

Sample ID	Precursors	Modified Peptides	Peptides	Protein Groups	Proteins
200 spd	10,950	10,516	8,605	1,864	2,901
100 spd 1.0ul	16,098	15,471	13,447	2,707	4,250
100 spd 2.5ul	26,146	24,697	21,049	3,837	6,100
60 spd 2.5ul	32,246	30,380	25,576	4,264	6,709
60 spd 10ul	48,374	44,688	36,691	5,687	8,883
30 spd	53,451	50,258	39,947	6,029	9,244

Table 5a: Database Search Results in Spectronaut v14

v15

Sample ID	Precursors	Modified Peptides	Peptides	Protein Groups	Proteins
200 spd	10,999	10,533	9,383	1,967	3,051
100 spd 1.0ul	16,626	15,958	14,007	2,800	4,378
100 spd 2.5ul	26,955	25,435	21,951	3,887	6,191
60 spd 2.5ul	33,073	31,115	26,628	4,405	6,929
60 spd 10ul	49,486	45,749	38,339	5,811	9,029
30 spd	55,108	51,830	42,241	6,172	9,417

Table 5b: Database Search Results in Spectronaut v15

- The number of protein groups in the spectral libraries generated from dda-PASEF data in Spectronaut (v14 / v15) ranged from 1,864 / 1,967 in the 200 SPD runs to 6,029 / 6,172 in the 30 SPD runs. The corresponding number of identified peptides were 8,605 / 9,383 and 39,947 / 42,241, respectively.
- In general, the number of identified protein groups and peptides correlates quite well with the length of the gradient used and the sample load on the Evotips.
- There is a small average increase of 5 % for the Peptides and 3 % for the Protein Groups being identified comparing Spectronaut v15 to v14.
- The highest increase in detections was observed at the shortest separation time used 200 SPDs (first line in *Table 5a and 5b*): 9 % for the peptides and 6 % for the Protein Groups
- This improvement is a result from a much better support of the ion mobility enabled MS data generated by the timsTOF Pro system in the later Spectronaut version (v15).

Results: “DIA Analysis” with Spectronaut “v14” and “v15”

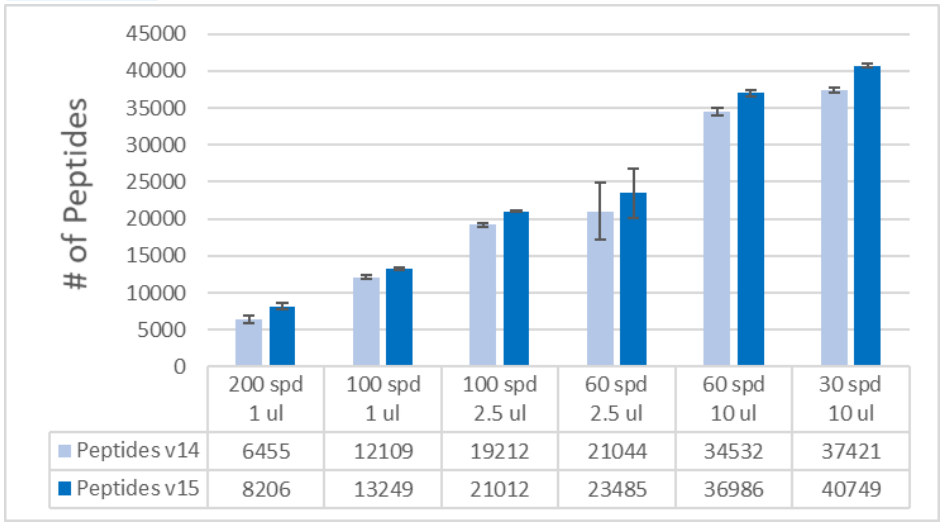

Figure 3a: identified Peptides in Spectronaut v14 vs v15

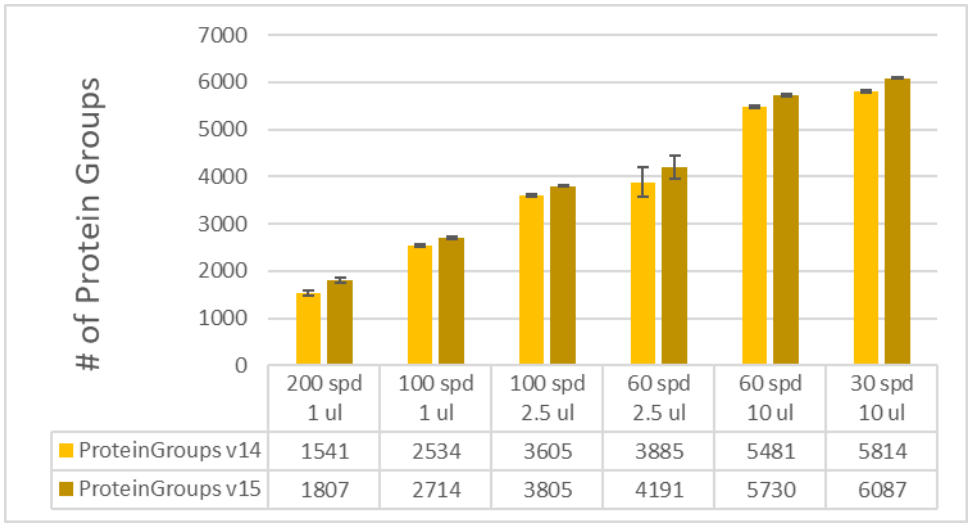

Figure 3b: identified Protein Groups in Spectronaut v14 vs. v15

dia-PASEF Experiment	Recovery	Median CV
200 spd 1 ul	98.7	24.3
100 spd 1 ul	99.6	23.8
100 spd 2.5 ul	99.7	19.6
60 spd 2.5 ul	99.6	23.6
60 spd 10 ul	99.9	15.0
30 spd 10 ul	99.8	20.8

Table 6: Library Recovery

- Again, the number of identified protein groups and peptides correlates quite well with the length of the gradient and the sample load used in the experiment done (*Figure 3a and 3b*).
- In the 30 SPD experiment on average 5,814 / **6,087** protein groups and 37,421 / **40,749** peptides were identified, whereas the fastest 200 SPD experiment resulted in 1,541 / **1,807** protein groups with 6,455 / **8,206** peptides being identified.
- The highest increase in identifications was observed at the shortest separation time (200 SPD runs): 27 % for the Peptides and 3 % for the Protein Groups. The average increase was 12 % for the Peptides and 1 % for the Protein Groups.
- The library recovery was generally quite high when matching the dia-PASEF data with same lengths of gradient against the dda-PASEF library (second row in *Table 6*, just v15 data shown).

Results: dia-PASEF match all against “deepest” library

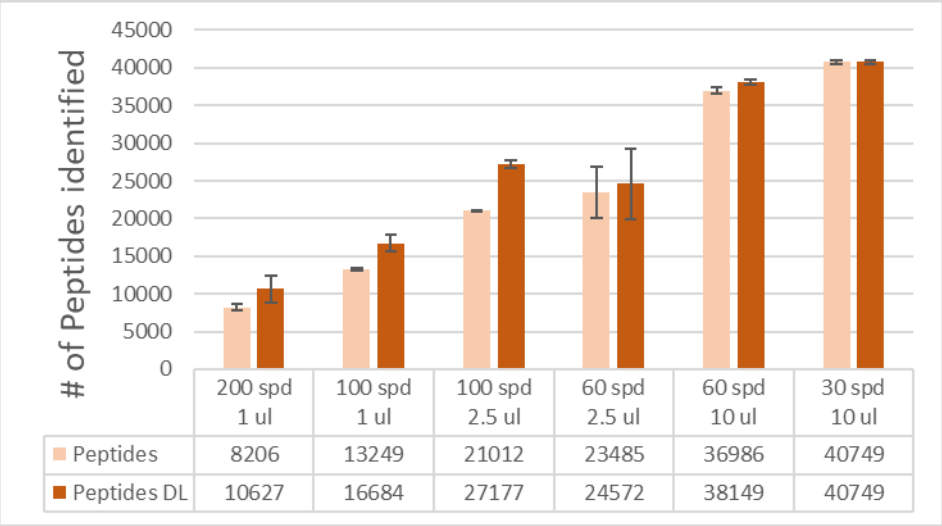

Figure 4a: identified Peptides in Spectronaut v14 vs v15

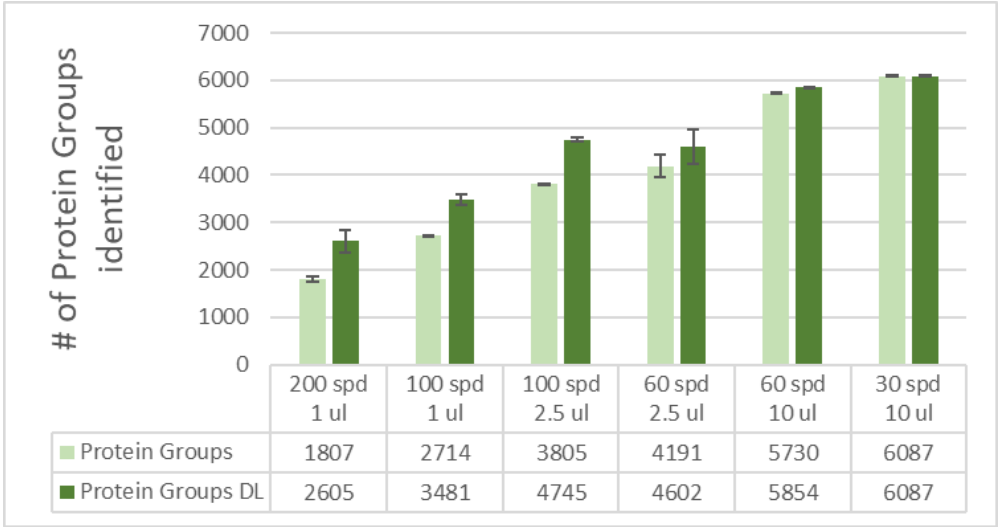

Figure 4b: identified Protein Groups in Spectronaut v14 vs. v15

dia-PASEF Experiment	Peptides increase %	Protein Groups increase %
200 spd 1 ul	30	8
100 spd 1 ul	26	5
100 spd 2.5 ul	29	3
60 spd 2.5 ul	5	2
60 spd 10 ul	3	0

Table 7: gain in percent

- To find the best balance between throughput and high identification rates, all shorter dia-PASEF runs were matched against the deepest library (named “DL” in *Figure 3*) which was generated from the 30 SPD dda-PASEF runs.
- There is a gain of 30 % in the number of identified Peptides and 8 % in the number of identified Protein Groups when matching the 200 SPD dia-PASEF runs against the 30 SPD dda-PASEF library. The increase raised significantly with the 100 SPD DIA runs (highlighted green in *Table 7*)
- The average gain is 15 % in the number of identified Peptides and 3 % in the number of identified Protein Groups.
- A library recovery between 44 % and 98 % was reached in the 200 and 60 SPD runs, respectively (*Table 8*). The recovery dropped at the 1 µl load of the 200 and 100 SPD runs (highlighted red in *Table 8*).

dia-PASEF Experiment	Recovery "DL"	Median CV "DL"
200 spd 1 ul	44	22.5
100 spd 1 ul	56	24.8
100 spd 2.5 ul	83	22.0
60 spd 2.5 ul	81	23.3
60 spd 10 ul	98	15.3

Table 8: Library Recovery

Results: directDia™

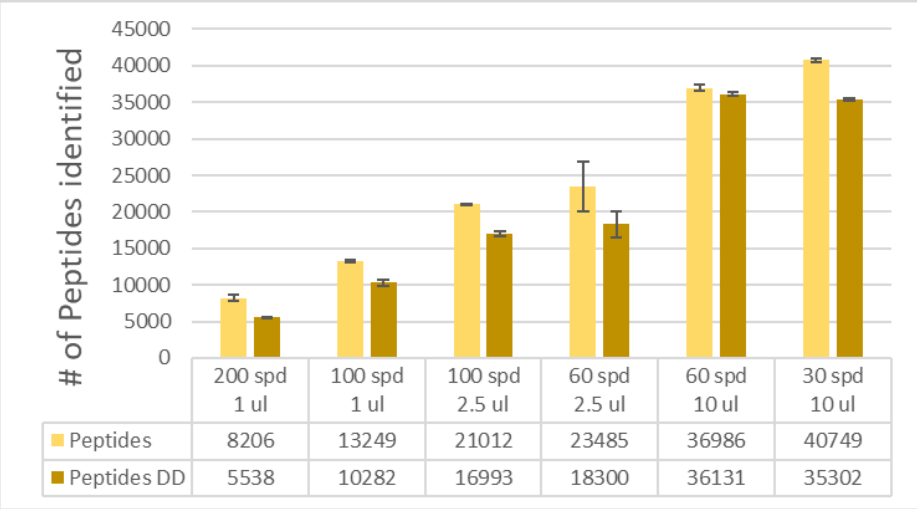

Figure 5a: identified Peptides "library" vs. "directDIA" (DD)

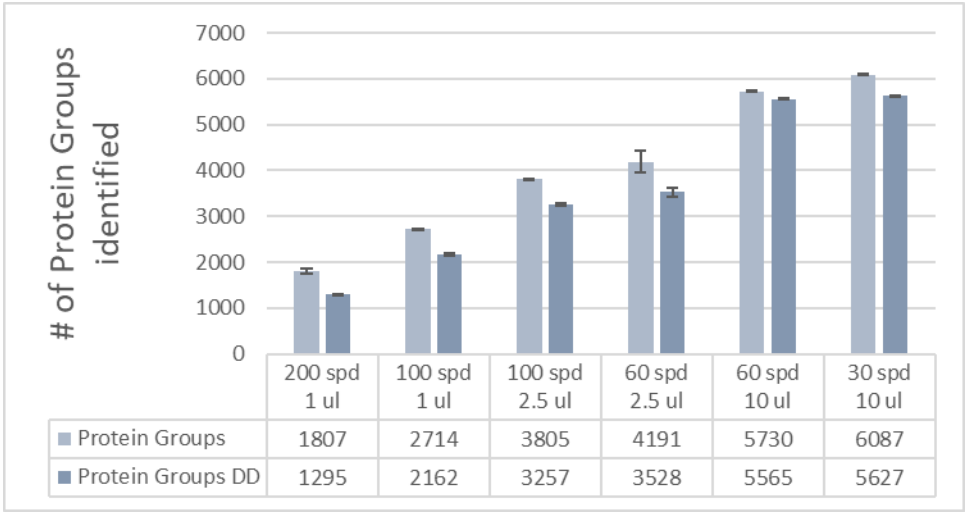

Figure 5b: identified Protein Groups "library" vs. "directDIA" (DD)

dia-PASEF Experiment	Peptides increase %	Protein Groups increase %
200 spd 1 ul	-33	-9
100 spd 1 ul	-22	-5
100 spd 2.5 ul	-19	-3
60 spd 2.5 ul	-22	-4
60 spd 10 ul	-2	0
30 spd 10 ul	-13	-1

Table 9: gain in percent

- To validate the potential of the Spectronaut software to directly perform a database search with dia-PASEF data, without the need to generate DDA-based libraries beforehand, all dataset were processed with the directDIA™ workflow in Spectronaut (named "DD" in *Figure 5*).
- In general the number of identified Peptides and Proteins is lower compared to a library based approach (*Figure 5a and b*).
- The difference is quite high for the faster chromatographic separation methods used (100 and 200 SPD, see *Table 9*).
- On average the number of identified Peptides is 18 % lower compared to a spectral library based approach, whereas the number of identified Protein Groups is 4 % lower.

Results: quantification of oxidized target peptides

Uniprot Accession	Sequence	Found
O00763	DVDEGLEAAERIGFP(+15.9949)LMIK	0
P31749	SGSP(+15.9949)SDNSGAEEMEVSLAK	1
P31749	LSP(+15.9949)PFKPQVTSETDTR	0
Q8TEP8	WHLSSLAP(+15.9949)PYVK	1
O00418	HMPDP(+15.9949)WAEFHLEDIATER	0
Q96AY3	ASP(+15.9949)AGAPLEDVVIER	1
Q16665	NPFSTQDTDLDLEMLAP(+15.9949)YIPMDDDFQLR	0
Q9Y4L1	VETGSEPGDTEPLELGGP(+15.9949)GAEPEQK	1
Q13438	VWGSP(+15.9949)GGEGTGDLDEFDF	1
Q13438	RVWGSP(+15.9949)GGEGTGDLDEFDF	1
P63151	IWDLNMENRP(+15.9949)VETYQVHEYLR	1
Q96T23	LEKPLP(+15.9949)ENEEK	0
P50454	KPAAAAAP(+15.9949)GTAEK	1
O43597	LLGSSFSSGP(+15.9949)VADGIIR	1
O43597	AQSGNGSQPLLQTP(+15.9949)R	1
	Total	**10**

Table 10: Spiked in oxidized peptides

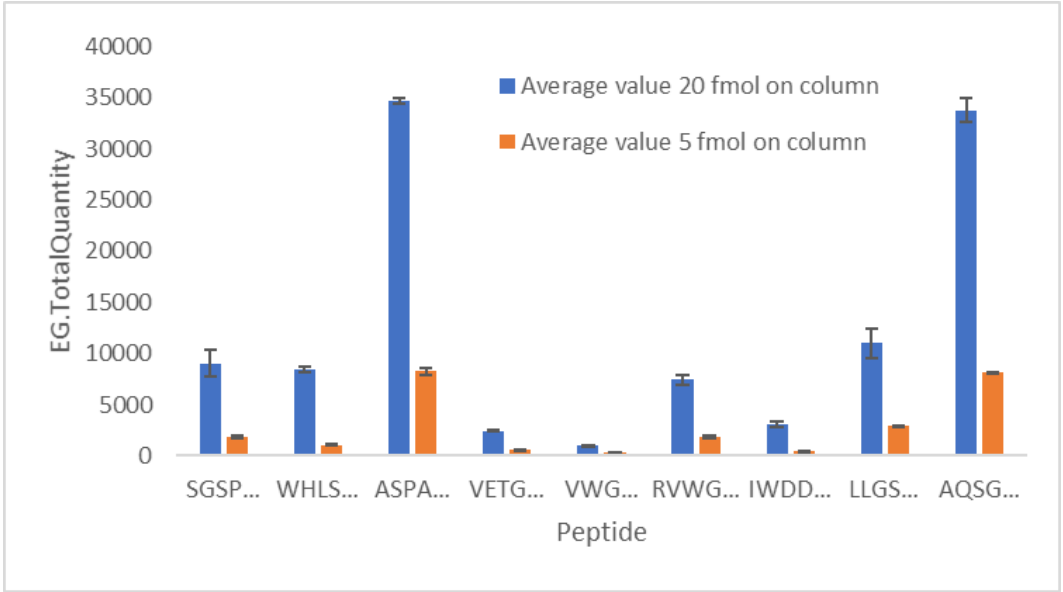

Figure 6: Quantification of oxidized peptides

- In the 30 SPD library based experiment analyzed with Spectronaut (v15) 10 out of 15 of the spiked-in oxidized target peptides were identified (*Table 10*).
- From this in total 9 could be successfully quantified, exhibiting the expected ratio of 4 ranging from 3.3 - 8.1 (*Figure 6*). The corresponding Standard Deviation ranges from 0.7 – 14.8 % (see error bars in *Figure 6*).
- The peptides were quantified in all triplicate injections done for 20 and 5 fmol on column spiked into 200 ng of background proteome (eHAP cell digest).

Summary and Conclusion

Summary:

- When comparing the two different software versions it is obvious that the newer one (v15) result in higher identification rates for both the library generation with dda-PASEF data and the subsequent dia-PASEF experiments.
- In general the increase was rather low on the Protein Group level, but significant on the Peptide level. The highest increase for the library generation was reached with the fastest separation method (200 SPD): +9 % Peptide identifications
- These numbers could be further improved by matching the shorter dia-PASEF runs against the deepest library generated with the 30 SPD method: here the best combination is 200 or 100 SPD dia-PASEF data matching against the “deepest” 30 SPD dda-PASEF library. This nicely demonstrates that very fast dia-PASEF experiments with 5.6 or 11 min. gradient length can benefit from longer 44 min. runtimes during library generation without the need for laborious fractionations of the sample.
- Finally, the directDia™ workflow in Spectronaut v15 showed the potential to even bypass the whole library generation step, at least for the longer runtimes being use in these experiments.
- In the exemplary chosen 30 SPD experiment accurate quantification was possible for most of the spiked-in target peptides in the presence of the complex tryptic digest of the eHAP cell line.

Conclusion:

- In this study the best balance between throughput and identification was found for the 100 and 200 SPD dia-PASEF runs matched against the deepest library generated from the 30 SPD dda-PASEF runs.