thermoscientific

# Enhancing the Throughput and Robustness of DIA Using a 150 µm EASY-Spray Column

**Joshua J. Nicklay[1], Kerry M. Hassell[1], and Aran Paulus[2],**
**[1]Thermo Fisher Scientific, 265 Davidson Ave Suite 101 Somerset, NJ 08873, [2]Thermo Fisher Scientific 355 River Oaks Parkway San Jose, CA 95134**

## ABSTRACT

**Purpose:** To demonstrate the benefits of running Data Independent Analysis (DIA) by capillary flow LC using a 150 µm ID EASY-Spray column on a Thermo Scientific™ Q Exactive™ HF mass spectrometer relative to nanoflow LC for proteomics.

**Methods:** A complex, three proteome mixture of yeast, *E. coli* and human cell digests was spiked with the Promega 6x5 peptide mixture and the Biognosys iRT peptide calibrants. Peptide spectral libraries were generated by Data Dependent Analysis (DDA) using nanoflow LC. DIA was run using various gradient lengths and sample amounts at both nanoflow rates and capillary flow rates. Library recovery and sensitivity were compared between the two flow regimes.

**Results:** Running shorter gradients by capillary flow DIA greatly increases the number of proteins and peptides identified per hour relative to nanoflow DIA. Sensitivity as measured by 6 heavy labeled peptides spanning five orders of magnitude is minimally reduced by capillary flow DIA.

## INTRODUCTION

Data-independent analysis (DIA) on Thermo Scientific™ Orbitrap™ mass spectrometers is uniquely suited for large scale differential analysis of proteomic samples due to its ability to rapidly identify and accurately quantify all ions within an LC-MS analysis. Assuming a sufficiently large spectral library, DIA is capable of deep proteomic characterization with minimal method development. To achieve the highest level of quantitative certainty, DIA requires replicate analyses, which can be extremely time-consuming when using traditional nanoliter flow rates. The throughput, efficiency and overall robustness of DIA can be greatly increased by moving to slightly larger bore columns, shorter gradients and higher flow rates. Here we compare throughput, depth of characterization, and overall sensitivity of DIA between nanoliter flow rates (300 nL/min) and capillary flow rates (2 µL/min).

## MATERIALS AND METHODS

**Sample Preparation**

For DDA, 600 ng/µL of yeast protein digest (Promega), 600 ng/µL of *E.coli* protein digest (Waters) and 900 ng/µL of HeLa cell protein digest (Thermo Fisher Scientific) were each spiked with iRT peptides (Biognosys) at the recommended 1:10 ratio and analyzed separately.

For DIA, 20 µL of 2 µg/µL yeast protein digest was combined with 20 µL of 2 µg/µL *E.coli* protein digest and 40 µL of 1 µg/µL HeLa cell protein digest. iRT peptides (Biognosys) were spiked in at the recommended 1:10 dilution. 50 µL of the 6x5 peptide mix (Promega) was also added such that the reference peptides spanned five orders of magnitude at a final concentration of 170 fmol/µL to 17 amol/µL. Final protein concentration was 825 ng/µL (275 ng/µL per proteome).

**Test Method(s)**

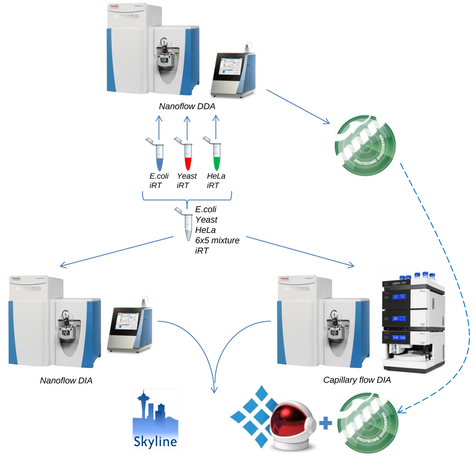
See figure 1 for workflow overview. For spectral library generation, each proteome (yeast, *E.coli*, HeLa) was separately analyzed in triplicate by DDA on the Q Exactive HF MS by nanoflow LC using four hour gradients using the Thermo Scientific™ EASY-nLC™ 1200 system. Specifically, 1.2 µg of yeast, 1.2 µg of E.coli and 900 ng of HeLa were individually injected onto a 75 µm i.d x 50 cm Thermo Scientific™ Acclaim™ PepMap™ column packed with 2 µm, 100 Å C18. The gradient ramped from 2% B to 20% B over 200 min, then from 20%B to 32%B in 40 min ((A): 0.1% FA in water, [B]: 0.1% FA in 80/20 ACN/water). The Q Exactive HF instrument was run by a top 10 data dependent MS/MS method where the full scan from *m/z* 400-1600 was taken at a resolving power of 60,000 with an AGC target of 1e6. MS/MS were taken at a resolving power of 15,000 with an AGC target of 1e5 and a maximum inject time of 40 ms. Quadrupole isolation was set to 1.5 *m/z*, intensity threshold was set to 5.0e4, and a dynamic exclusion of 30 s was used.

For DIA, the three proteome mixture was analyzed by nanoflow using the aforementioned 75 µm x 50 cm EASY-Spray column using both one⁺ hour and two⁺⁺ hour gradients (2-25%B in 50⁺/100⁺⁺ min, 25-40%B in 10⁺/20⁺⁺ min) and 1 µL injections (825 ng total protein on column) in triplicate. The same mixture was analyzed by capillary flow using a prototype 150 um x 15 cm Thermo Scientific™ Acclaim™ PepMap™ column packed with 2.2 µm C18 using both one⁺ hour and 30⁺⁺ min gradients (2-25%B in 25'50⁺ min, 25-40%B in 5'/10⁺⁺ min) using the Thermo Scientific™ Ultimate™ 3000 RSLCnano LC system equipped with a capillary flow selector. Various injection volumes were tested. The DIA method on the Q Exactive HF MS consisted of a full scan acquired at 120,000 resolving power from m/z 400-1000 and an AGC target of 3e6, followed by 15 DIA scans using 20 *m/z* isolation windows covering m/z 400-700 at a resolving power of 30,000 using an AGC target of 1e6, followed by a second full scan from *m/z* 400-1000, followed by another 15 DIA scans using 20 *m/z* isolation windows covering *m/z* 700-1000.

**Data Analysis**

DDA files were searched in Thermo Scientific™ Proteome Discoverer™ 2.1 software using SEQUEST HT and Percolator to filter to a 1% FDR. DIA files were analyzed in Spectronaut 10 to determine the number of identified proteins and peptides. DIA files were also analyzed in Skyline 3.6.1.10690 to determine limits of detection based on the spiked in 6x5 peptide mixture. All DIA quantitation was performed at the MS1 level.

**Figure 1. Experimental workflow. Spectral libraries were generated by Proteome Discoverer software using 4 hour nanoflow DDA runs on the individual proteome samples. DIA runs were acquired on the combined, three proteome mixture (with the 6x5 peptide mixture spiked in) by both nanoflow and capillary flow analyses using various gradient lengths. DIA runs were analyzed by Spectronaut and Skyline.**




## RESULTS

**Data Dependent Analysis**

The results for the data dependent spectral library generation are shown in table 2. In total, the library consisted of 10,847 proteins and 89,859 peptides.

**Table 1. Final spectral library peptide and protein group numbers. 1.2 µg of yeast, 1.2 µg of E.coli and 900 ng of HeLa were separately analyzed in triplicate by nanoflow DDA using 4 hour gradients.**

| | Rep. 1 Proteins/Peptides | Rep. 2 Proteins/Peptides | Rep. 3 Proteins/Peptides | Total Proteins | Total Peptides |
|---|---|---|---|---|---|
| ***E.coli*** | 1,552/8,685 | 1,581/8,904 | 1,372/7,266 | **1,700** | **10,514** |
| **Yeast** | 3,426/24,223 | 3,486/25,110 | 3,561/25,825 | **3,819** | **33,814** |
| **HeLa** | 5,049/34,659 | 5,089/34,748 | 5,139/35,458 | **5,711** | **46,567** |
| **Combined** | | | | **10,847** | **89,859** |

**Data Independent Analysis – Library Recovery**

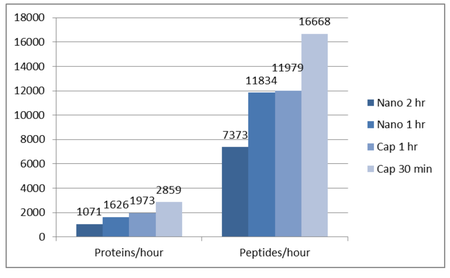
As expected, DIA using longer nanoflow LC gradients identified a greater number proteins and peptides relative to capillary flow LC. However, throughput is greatly reduced when running nanoflow LC due to the extra time spent in loading and column equilibration. Protein and peptide identifications for each experimental condition as reported by Spectronaut are shown in table 2.

**Table 2. Protein groups and peptide recovery from DIA for the various flow regimes and injection volumes relative to total analysis time across triplicate runs. Total protein concentration was 825 ng/µL. *Proteins and peptides reported by Spectronaut as "sparse" groups, meaning they did not need to be identified in all replicates.**

| Flow Regime | Injection Volume (µL) | Gradient Duration | Total Analysis Time | Total Proteins* | Total Peptides* |
|---|---|---|---|---|---|
| **Nano** | 1 | 2 hour | **8.4 hours** | **9,594** | **69,821** |
| **Nano** | 1 | 1 hour | **5.9 hours** | **8,999** | **61,937** |
| **Capillary** | 1 | 1 hour | **4.1 hours** | **8,090** | **49,112** |
| **Capillary** | 2 | 1 hour | **4.1 hours** | **8,088** | **49,538** |
| **Capillary** | 4 | 1 hour | **4.1 hours** | **8,132** | **49,911** |
| **Capillary** | 1 | 30 min | **2.6 hours** | **7,433** | **43,336** |
| **Capillary** | 2 | 30 min | **2.6 hours** | **7,462** | **42,882** |
| **Capillary** | 4 | 30 min | **2.6 hours** | **7,639** | **44,911** |

Despite the reduction in identified proteins and peptides when running DIA by capillary flow, the clear benefit is that efficiency is greatly increased. Comparing the total analysis time for the nanoflow 1 hour gradient experiment to the capillary flow 30 min gradient experiment, a 2.3x reduction in instrument time only results in a 15% reduction in identified proteins, This gain in efficiency is also depicted in figure 2.

**Figure 2. DIA protein and peptide recovery per hour for 1 hour and 2 hour nanoflow gradients as well as 1 hour and 30 min capillary flow gradients for identical injection volumes. Total triplicate analysis time was considered.**




**Data Independent Analysis – Sensitivity**

Sensitivity of each assay was measured by the use of six heavy labeled peptides each spiked in at five different concentrations spanning five orders of magnitude, each 10x apart in concentration. For the same gradient duration and same amount on column DIA by nanoflow showed lower LOQ's for four of the six peptides; for two of the six peptides, nanoflow showed similar LOQ's as capillary flow, as shown in table 3.

**Table 3. Sensitivity of nanoflow DIA and capillary flow DIA with 1 µL injection volumes as measured by the 6x5 peptide dynamic range mixture. All quantitation was performed at the MS1 level by summing the areas for three isotopes of each parent peptide. LOQ was defined as the lowest concentration with a %CV < 20%.**

| | Nano 1hour 1 µL injection LOQ | Nano 1hour 1 µL injection %CV @ LOQ | Capillary 1 hour 1 µL injection LOQ | Capillary 1 hour 1 µL injection %CV @ LOQ |
|---|---|---|---|---|
| **VTSGSTSTSR** | 170 amol/µL | 20.16% | 1.7 fmol/µL | 0.54% |
| **LASVSVSR** | 1.7 fmol/µL | 16.05% | 1.7 fmol/µL | 13.21% |
| **YVYVADVAAK** | 1.7 fmol/µL | 8.69% | 1.7 fmol/µL | 6.52% |
| **VVGGLVALR** | 170 amol/µL | 6.31% | 1.7 fmol/µL | 13.45% |
| **LLSLGAGEFK** | 170 amol/µL | 13.26% | 1.7 fmol/µL | 17.17% |
| **LGFTDLFSK** | 170 amol/µL | 10.67% | 1.7 fmol/µL | 6.24% |

As injection volume is increased with capillary flow from 1 µL to 4 µL, LOQ's for three of the six peptides (LASVSVR, YVYVADVAAK and LGFTDLFSK) are reduced from 17 fmol/µL to 170 amol/µL.

**Table 3. Sensitivity of nanoflow DIA and capillary flow DIA with 4 µL injection volumes as measured by the 6x5 peptide dynamic range mixture. All quantitation was performed at the MS1 level by summing the areas for three isotopes of each parent peptide. LOQ was defined as the lowest concentration with a %CV < 20%.**

| | Nano 1hour 1 µL injection LOQ | Nano 1hour 1 µL injection %CV @ LOQ | Capillary 1 hour 4 µL injection LOQ | Capillary 1 hour 4 µL injection %CV @ LOQ |
|---|---|---|---|---|
| **VTSGSTSTSR** | 170 amol/µL | 20.16% | 1.7 fmol/µL | 0.54% |
| **LASVSVSR** | 1.7 fmol/µL | 16.05% | 170 amol/µL | 10.59% |
| **YVYVADVAAK** | 1.7 fmol/µL | 8.69% | 170 amol/µL | 15.2% |
| **VVGGLVALR** | 170 amol/µL | 6.31% | 1.7 fmol/µL | 13.45% |
| **LLSLGAGEFK** | 170 amol/µL | 13.26% | 1.7 fmol/µL | 17.17% |
| **LGFTDLFSK** | 170 amol/µL | 10.67% | 170 amol/µL | 10.01% |

## CONCLUSIONS

DIA by capillary flow LC using a 150 µm I.D. EASY-Spray column on a Q Exactive HF MS was demonstrated to greatly reduce analysis time relative to nanoflow LC while maintaining a significant number of protein identifications. Low level analytes can be measured at LOQ's similar to nanoflow by adjusting injection volume to account for the inherent loss of sensitivity when moving from nanoflow to capillary flow.

- With triplicate 30 minute gradients yielding a total analysis time of 2.6 hours per sample, over 7,600 proteins and nearly 45,000 peptides can be identified and quantified by DIA using capillary flow LC.
- With identical injection volumes, DIA by nanoflow LC was shown to be more sensitive. By increasing the injection volume of the capillary flow LC by 4x, similar sensitivities to nanoflow were demonstrated.

## REFERENCES

1. MacLean, Brendan, et al Bioinformatics. 2010 Apr 1; 26(7): 966–968.

## ACKNOWLEDGEMENTS

We would like to thank Sue Abbatiello for providing the Skyline document for the 6x5 peptide mixture.

## TRADEMARKS/LICENSING