**Data-independent Acquisition for the Orbitrap Q Exactive HF. A tutorial.**

Léon Reubsaet*[1,2], Michael J. Sweredoski[1], Annie Moradian[1]

1) Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, 1200 E. California Blvd. Pasadena, CA 91125, USA

2) Department of Pharmaceutical Chemistry, School of Pharmacy, University of Oslo, Sem Sælands vei 3, N-0316 Oslo, Norway

*) Corresponding author. Email: leonr@farmasi.uio.no

***Abstract***

Data-independent acquisition (DIA) is a powerful mass spectrometric technique to perform both protein identification and quantification of complex protein samples. Setting up DIA methods on Orbitrap analyzers requires a thorough overview over the actions the Orbitrap mass spectrometers carry out. This tutorial is written with the intention to give an overview over the important parameters to consider as well as which measurements to carry out to get the most out of your DIA method when setting it up on an Orbitrap mass analyzer. Instead of giving the optimal DIA settings, all steps in the construction and optimization of the DIA method are shown and discussed in a way that allow tailored DIA methods. Key steps are: building the spectral library after sample fractionation, decide upon the number of data points per chromatographic peak, determine scan times of each mass spectrometric step, construct various DIA methods using these data and evaluate their performance. This proposed DIA method development strategy was tested on digested lysates from Pseudomonas Aeruginosa and compared to conventional DDA analysis to put the DIA results in a perspective.

***Keywords:*** Data-independent acquisition, method development strategy, actual scan times, orbitrap mass analyzer

***Introduction***

Since it was coined first in 1997[1], the field of proteomics has evolved very rapidly, aided by fast developments in mass spectrometry[2]. The readily accessible high-resolution mass spectrometers like orbitrap and time-of-flight (TOF) boosted proteomics as a diverse discipline with many varying application areas. The ability to trigger automated MS/MS analysis of tryptic peptides detected in complex samples made it possible to generate and analyze large amounts of qualitative and

quantitative information of proteins reasonably fast [3-5]. Within a short time this data dependent acquisition (DDA) became a standard strategy in the field. It generally is based on the selection for MS/MS fragmentation of the top N most abundant peptides in a full scan high resolution single mass spectrum using relative narrow isolation widths for each fragmented *m/z*. These selected *m/z* values are thus excluded for fragmentation during a given time-frame (typically 15-30 sec depending on the chromatographic peak width) to avoid redundancy. With respect to optimal settings for DDA, the impact of essential parameters like resolution, AGC, maximum ion injection time and number of selected top ions on sensitivity and selectivity were investigated and discussed earlier [6, 7]. In contrast to DDA, data-independent acquisition (DIA) carries out fragmentation using wider isolation widths (typically 10-50 Da) covering all *m/z* values within a certain mass range. In this way all ions within this mass range will be fragmented generating more signal for a single peptide giving the advantage of more reliable relative quantifications compared to DDA [8]. It is obvious that fragmentation spectra are far more complex than those obtained by DDA and therefor DIA relies on high resolution mass spectrometry and often a spectral library (either in-house generated or available from other sources). DIA has been described for the first time more than ten years ago [9, 10] and the number of citations has increased substantially over the years. Within DIA, the term SWATH (Sequential Window Acquisition of all Theoretical fragment-ion spectra) was introduced[11] for analyses of sequential (minimally overlapping) isolation windows within a predefined mass range on certain TOF mass spectrometers. DIA can however also be carried out on orbitrap mass analyzers: several groups have evaluated the influence of cycle time, scan time, maximum ion injection time, resolution, isolation widths on the DIA performance of the Q Exactive[12, 13]. Alternative approaches like wide SIM DIA (using MS1 for quantification) and untargeted DIA (allowing for search outside the spectral library) were investigated through comparison with “classic” DIA by Koopmans et al. [14]. There is an overwhelming number of possible combinations of mass spectrometric settings and as well as experimental parameters to investigate and to construct an optimal DIA method. This tutorial proposes the order of parameter optimization for DIA analyses on a Q Exactive orbitrap mass spectrometer with respect to both label free quantitation and number of identifications. The settings found after this optimization might be copied to other laboratories. This report is not meant as a comparative study between DDA and DIA, however, DDA results are included for comprehensive purposes.

***Experimental***

Unless otherwise stated, chemicals used were of analytical grade.

*Biological samples*

Pseudomonas Aeruginosa (strain PA14) lysates (provided by Newman Lab at Caltech) were digested using a modified Filter-Assisted Sample Preparation (FASP) as described earlier [15].

*Basic pH fractionation*

The complex biological samples were fractionated using the Pierce Thermo high-pH fractionation kit (ThermoFisher Scientific, Rockford, IL, USA). A pooled digested Pseudomonas Aeruginosa lysate was produced by mixing 75 µL of 4 biological samples. This resulted in 300 µL pooled sample with a protein concentration of 0.1 µg/µL. This sample was subjected to fractionation in its whole. Fractionation was carried out as described according to the manufacturer with the exception that instead of 0.1% trifluoroacetic acid (TFA), 0.2 % formic acid (FA) was used to re-dissolve the samples. A total of 11 fractions were collected: the flow through when the 300 µL sample is applied, the consecutive water wash (300 µL) and the eluates (300 µL) using 5 %, 7.5 %, 10 %, 12.5 % 15 %, 17.5 %, 20 % and 50 % acetonitrile (MeCN) in 0.1 % triethylamine (TEA) followed by a final wash where 300 µL 100% MeCN was used.

After fractionation the fractions were dried down using SpeedVac. The dried fractions were reconstituted in 10 µL 0.2% FA containing 1: 10 diluted iRT peptides (according to the manufacturer's protocol - Biognosys, Zürich, Switzerland). Approximately 8 µL of each reconstituted sample was transferred to a sample vial.

*Chromatographic conditions - standard 120 minute run*

Gradient elution was performed using Proxeon easy-nLC 1200 (Thermo Scientific, San Jose, CA, USA). A 50 µm I.D x 20 cm column with a 10 µm electrospray tip (PicoFrit ™ from New Objective, Woburn, MA, USA) in-house packed with ReproSil-Pur C18-AQ 1.9 µm (Dr. Maisch GmbH, Ammerbuch, Germany) was used for separation. The flow was set to 220 nL/min. Mobile phase A consisted of 2 % MeCN in 0.2 % FA, mobile phase B consisted of 80 % MeCN in 0.2 % FA. The following gradient program was carried out: 2–6% Solvent B (7.5 min), 6-25% B (82.5 min), and 25-40% B (30 min) and to 100% B (9 min). Re-equilibration of the column (>15 column volumes) using starting conditions was performed for 30 min before injection of the sample.

*Chromatographic conditions - runs < 120 minutes.*

For the DDA and DIA runs with run times 60, 30 and 15 minutes, the chromatographic conditions in the essential part of the gradient were changed proportionally. Table 1 shows the time points and the percentage of mobile phase B for all run times.

Table 1: Gradient runs for the four different tested run lengths.

| *Percentage mobile phase B (%)* | *120 min run (min)* | *60 min run (min)* | *30 min run (min)* | *15 min run (min)* |
|---|---|---|---|---|
| *2* | *0* | *0* | *0* | *0* |
| *6* | *7.5* | *3.75* | *1.97* | *1* |
| *25* | *90* | *45* | *22.5* | *11.25* |
| *40* | *120* | *60* | *30* | *15* |
| *100* | *121* | *61* | *31* | *16* |
| *100* | *130* | *70* | *40* | *25* |

*Mass spectrometric detection*

Mass spectrometric detection was carried out on a Q Exactive HF orbitrap mass analyzer (Thermo Scientific, San Jose, CA, USA). The system was operated either in the DDA mode or in the DIA mode. Xcalibur version 4.0.27.13 and Tune version 2.7 Build 2659 were used to generate instrumental methods as well as to operate the mass spectrometer.

*Mass spectrometric conditions - DDA*

For the DDA runs, the optimal FullMS/dd-MS2 (top 12) method described by Kelstrup et al. [6] was used. In brief: The MS1 had a resolution of 60000 and a scan range between 400 and 1650 *m/z*. The maximum ion injection time was 15 ms with an AGC target of 3e6. The MS/MS had a resolution of 30000, a maximum injection time of 45 ms with an AGC target of 1e5. The isolation width was 1.2 *m/z* and the nce was set to 28. DDA was triggered at an intensity threshold of 1e5. The dynamic exclusion time was set to 45 sec. Lockmass was 445.12000

*Mass spectrometric conditions - DIA*

For the DIA runs, each cycle was built up as follows: one full MS scan at resolution 60000 (the resolution in the full MS was kept constant, scan range between 400 and 1650 *m/z*) followed by narrow isolation widths covering *m/z* 400 - 900 and a wide isolation window covering *m/z* 900 - 1650. Number of isolation windows, isolation widths, maximum ion injection times and resolutions

were varied, the actual values are given in the *"Results and Discussion"* section. The AGC target was 3e6, NCE was set to 28. Most of the methods last for 120 minutes, unless otherwise stated.

*LC-MS/MS methods for the determination of the actual cycle times*

The LC is set to isocratic flow (200 nL/min) @ 100% mobile phase B. There will be no sample injection in this method, so the MS signal obtained will only be from the LC background. This is to create an ion poor situation.

Each cycle in the MS method consists of one full MS scan over a *m/z* range 400-1650 followed by 10 relative narrow isolation windows (*m/z* range 400 - 900, isolation width *m/z* 50.5) and one relative wide isolation window (*m/z* range 900-1650, isolation width *m/z* 751) each isolation width overlapping 0.5 *m/z.* The AGC target was for this particular experiment set to $3x10^6$ to allow the maximum injection time to be reached. In other experiments it is set this high to increase the dynamic range[12, 13].

Using this MS method, determination of the actual scan times is carried out stepwise: first scan times of MS1 are evaluated (keeping the resolution in MS2 constant at 15 000). After this scan times in the MS2 are evaluated (keeping the resolution in the MS1 constant at 60 000).

Each method is acquired for one minute.

*Data analysis*

For spectral library purposes and DDA experiments MaxQuant was used.

Raw data were analyzed with MaxQuant (version 1.6.1.0) [16, 17] against a Pseudomonas Aeruginosa database (downloaded from UniProt on Dec 9th 2016, containing > 5500 sequences) and a contaminant database (245 entries). A decoy database was constructed by MaxQuant on-the-fly to determine the false discovery rate (FDR). Trypsin ([KR][^P]) was specified as the proteolytic enzyme with up to two missed cleavages. Carboxyamidomethyl modification of cysteine (57.0215 Da) was specified as a fixed modification. Variable modifications included oxidation of methionine (15.9949 Da), and protein N-terminus acetylations (42.0106 Da). Peptides were searched with a tolerance of 4.5 ppm MS1 in MaxQuant and an MS/MS tolerance of 20 ppm (FT) and a peptide and protein FDR of 1%.

DIA experiments were analyzed using Spectronaut (v. 11 Biognosys AG). The project specific library was built according to the instructions given in the User Manual of Spectronaut v.11. In short, the MaxQuant search results from each of the DDA runs of the fractionated sample are loaded in the "Load Experiment form" in Spectronaut. It is recommended to link the raw-files to the MaxQuant results. Spectronaut then builds the project specific library that consisted of 49213

precursors. After this, each DIA experiment was loaded ("Load Raw Data…") and the project specific library was assigned to these DIA runs. The spectral libraries specify the peptide search space and are required in many DIA analyses. The spectral libraries provide actual observed fragment ion intensities, rather than basic in-silico predictions common to DDA database searches, which help to provide highly sensitive and accurate peptide identifications. Specifically, the project specific library generated from the DDA runs were used for "6) Evaluate the DIA performance at various resolutions and various settings" and "8) Comparison with DDA and evaluation of the DIA method length". Dynamic mass tolerance at the MS1 and MS2 levels was employed. The XIC RT Extraction Window was set to Dynamic with a correction factor of 1. Calibration mode was set to automatic with non-linear iRT calibration and precision iRT enabled. Decoys were generated using the scrambled method and a dynamic limit (default settings). P-value estimation was performed using a kernel density estimator. Interference correction was enabled and with no proteotypcicity filter. Major grouping was by Protein-Group Id and minor grouping was by stripped sequence. Major group quantity was mean peptide quantity. Major group top N was enabled with a minimum of 1 and maximum of 3. Minor group quantity was mean precursor quantity. Minor group top N was enabled with a minimum of 1 and maximum of 3. Quantity MS-Level was MS2 and quantity type was area. Qvalue was used for data filtering. Cross run normalization was enabled with Qvalue sparse row selection and local normalization. Default labeling type was label free with no profiling strategy and unify peptide peaks no enabled. Protein inference workflow was set to automatic.

## ***Results and discussion***

### *Choice of strategy*

As mentioned earlier, there is an overwhelming number of combinations of settings which can be varied for the optimization of a DIA method. The following sections describe the order and nature of the steps to go through to design a balanced DIA method. It generally consists of:

1 - sample fractionation for maximal proteome depth
2 - building a project specific spectral library
3 - making the decision of number of points per chromatographic peak
4 - measurement of the actual scan time each MS step takes
5 - construction of the number of isolation width's, isolation width and isolation times
6 - evaluate the DIA performance at various resolutions and various settings
7 - decide upon the settings optimal for its purpose

Step 1 and 2 create the space in which the DIA searches will be carried out. These steps can be omitted by using public available spectral libraries.
Step 3 and 4 give the information needed to model the cycle times.
Step 5 and 6 are the actual DIA method generation and evaluation of important settings. After this the most optimal settings can be used for further DIA analyses (step 7)

All these steps were, in our case, carried out using a complex lysate from the Pseudomonas Aeruginosa PA14 strain as a model.

figure 1 shows a graphic overview over the steps to follow to develop a satisfactory DIA method.

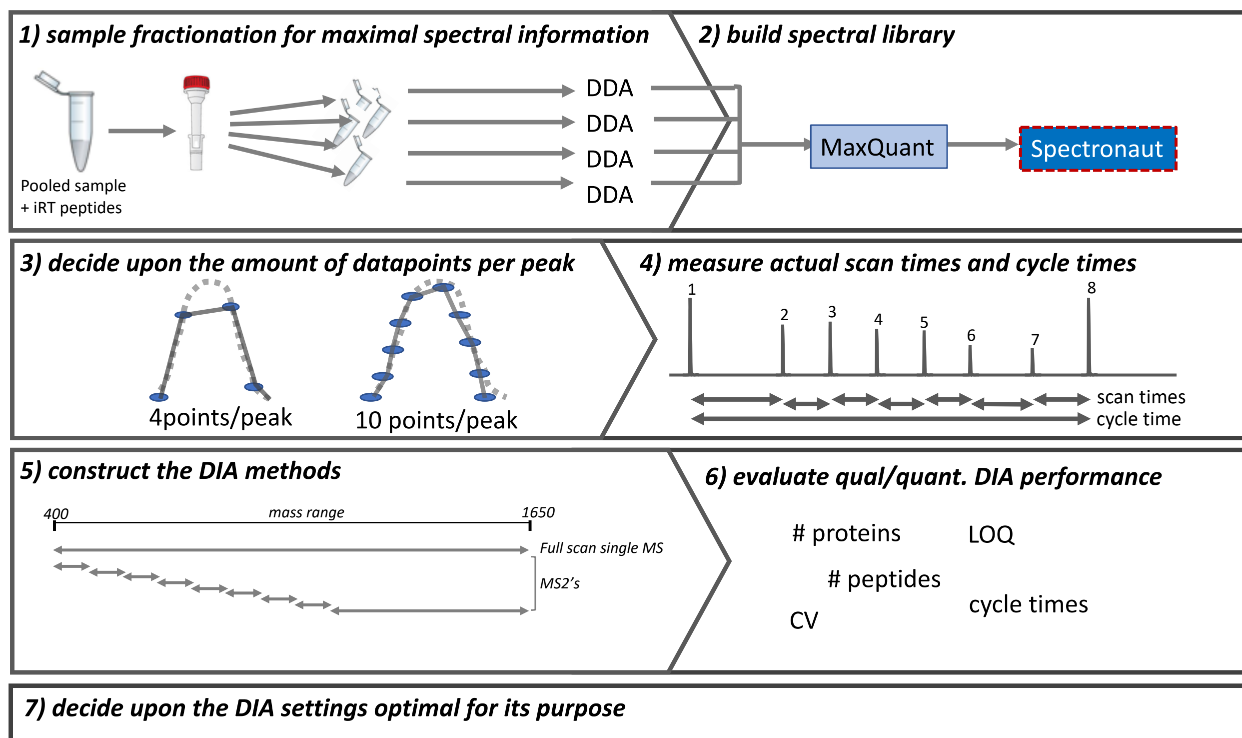



***figure 1***: *schematic overview over all the steps, and the order, to develop DIA methods*

*1) Sample fractionation for maximal proteome depth*

The post-proteolytic sample fractionation prior to spectral library generation was carried out to maximize the depth of proteome coverage. Although sample fractionation itself can be subject to choice of retention mechanism (SCX, basic pH fractionation, HILIC, etc.) and optimization

(experimental parameters, number of fractions, etc.) it was decided not to include these in this report. Here a commercial basic pH fractionation with its accompanying protocol was used to reveal as much spectral information as possible. A previously optimized DDA method described by Kelstrup et al. [6] was used to generate the MS/MS spectra for the spectral library.

Pseudomonas Aeruginosa digested lysate fractionation

As a model for spectral library generation digested lysates from Pseudomonas Aeruginosa (PA14) were injected as a pooled sample as well as 9 injections from the fractions of a pooled sample.

In the supplementary information the number of identified peptides and proteins for various injections are shown. Comparing the information generated from a single injection of a pooled sample with the information generated from the combined 9 fraction injections it is obvious that many more peptides were identified in the combined 9 fraction injections (40735 peptides vs 24756 peptides, possible contaminants and reversed database subtracted). Subsequently many more proteins were identified too (3755 proteins vs 2718 proteins, possible contaminants and reversed database subtracted, at least two unique peptides). The number of entries for the predicted number of proteins for Pseudomonas Aeruginosa PA14 from UniProt is 5888. Given this number, approx. 64% of the predicted proteome is identified in the combined fractionated sample.

*2) Create the library through DDA*

To analyze the collected DIA data, a spectral library is often used to identify and quantify the peptides present in the sample. One common method for generating a project specific library is to pool aliquots of the digested samples and then fractionate this pooled sample prior to DDA analysis. This will lead to a spectral library that is large enough to contain most peptides that could be observed in the DIA analyses while minimizing the number of peptides in the library that are unlikely to be in the samples. Since we used Spectronaut for our spectral library generation and subsequent DIA analysis, iRT peptides were added to the fractionated digested lysate for retention time alignment.

We used both MaxQuant and Spectronaut to build a sample specific library. Raw data collected from each of the 9 injected fractions was searched by MaxQuant (v. 1.6.1.0) against the UniProt P. aeruginosa PA14 reference proteome and a contaminant database. Default settings were used. These results were then loaded into Spectronaut (v 11) to create the sample specific spectral library, again with default settings. The final library contained 49213 precursors, 40570

modified peptides, 36451 peptides, 4062 proteins (with 285 single hits), and 282956 fragment ions.

*3) Decide upon the number of points per chromatographic peak / MS cycle time*

One of the goals of the DIA experiments is to quantify the peptides. This is achieved by integrating the area under the chromatographic peaks. In order to get a reliable quantification, analytical variation must be minimized as much as possible. One factor that can affect analytical variation is the sampling frequency across the chromatographic peak. For a fixed peak width, more frequent sampling, and thus more points per peak, leads to a better approximation of the true chromatographic peak and therefore lower analytical measurement variation. In the typical DIA configuration, each peptide chromatographic peak is sampled once per total scan cycle. In the proteomics community approx. 8 points per peak is considered to be satisfactory[13] while in the chromatography community between 12-15 points per peak are considered to be the standard[18]. In general, the more points per peak, the lower the analytical variation will be. On the other hand, increasing the amount of points per peak (assuming constant peak width) requires shorter cycle times (the time between consecutive data points represents a whole cycle of MS actions) and potentially to less identifications and lower sensitivity. When deciding on number of points per peak, the tradeoffs between analytical variation and depth of coverage should be considered in the context of each specific project. In this report we chose 10 points per chromatographic peak. The next step then is to find out the experimental chromatographic peak width. This was done on the files used to create the spectral library. From the “evidence” file after MaxQuant analysis of these injections, the peak width (named “retention length” in the MaxQuant output) of all the peptides is given. From this a distribution can be made. In our case, as well in most cases, the retention widths vary to a large extent due to the varying chromatographic behavior of the peptides as well as the variation in concentration of the peptides. It will thus be challenging to decide upon a retention width which suits for all peaks. We therefore used the median retention width as the guiding experimental parameter. For the tryptic peptides generated from the Pseudomonas Aeruginosa PA14 strain, this median was 20.5 sec thus leading to a target total cycle time to be set at 2 sec.

*4) Measure the actual cycle time considering all steps.*

As the total cycle time is set to 2 sec, the MS settings need to be designed such that all MS scans will be performed within this time frame. For the DIA method that means that both a full scan in

the single MS mode and all fragmentation MS2 scans of all ions in the anticipated mass range need to be carried out within these 2 seconds.

An advantage with a trapping mass spectrometer is that it simultaneously can carry out the transient (performing the actual scan) and many of the ion pre-treatments (like fragmentation) for the next scan (see figure 2). In the orbitrap, the transient time is a value which is determined by the setting of the resolution: the higher the resolution the longer the transient time. This transient time is the shortest time the mass spectrometer needs to perform a scan. In addition to the transient time, a certain amount of time is needed to allow ions to leave and enter the orbitrap. This additional time is referred to as *overhead A* in figure 2. Besides transient and overhead A (which are sequential actions) the MS needs to inject, eject, collect, fragment and transport ions. The time the MS needs for all these actions is referred to as *overhead time B (see figure 2).* This overhead B varies with different settings. Overhead B is carried out parallel to the transient and before overhead A. The time needed for all these actions in overhead B.

The contribution of the *overhead time A* for a full MS scan is rather low (approx. 10 ms) [12], especially when the resolution is set to 60 000 or higher. Initial experiments (using the method described under "*LC-MS/MS methods for the determination of the actual cycle times* ") showed that the contribution of the overhead time can be considerable for the MS/MS steps. To exemplify this: an MS/MS scan at 15 000 resolution with a single HCD fragmentation and a measured ion injection time of 5 ms, the MS/MS scan takes 41 ms. This is, compared to the anticipated transient time of 32 ms, an increase of 29%.

Fragmentation in the HCD using stepped collision energy allows optimal fragmentation for compounds with various optimal collision energies. One should however be aware of the fact that when a stepped HCD collision instead of a single HCD collision is carried out, this increase in overhead time becomes considerably more: with a maximum injection time of for example 22 ms (resolution 15 000) and a stepped collision energy (NCE) of 25.5, 27 and 30, the MS/MS scan time was measured to be 72 ms (an increase of 125% compared to the anticipated 32 ms). At a maximum injection time of 54 ms (resolution 30 000) and 118 ms (resolution 60 000) this was measured to be 101 ms (increase 58 %) and 167 ms (increase 30%) respectively. The mass spectrometer uses the injection time as set (for each of the three energy levels it uses 1/3 of the ion injection time). However, it will use additional time to perform collision, collection and transportation of ions as this needs to be carried out for each energy level. This this will take longer time than when single HCD collision is carried out [12, 13].

Taking into account that the cycle time mainly consists of time required to perform MS/MS scans, the overhead can increase the cycle time greatly, leading to less data points per chromatographic peak and thus unwanted higher CV. It is therefore necessary to take into account the contribution of these overhead times producing a realistic overall cycle time when calculating the number of DIA scans to be carried out per cycle. Additionally, we observed that when the isolation band width exceeds 500 *m/z*, the scan time of the scan preceding this wide isolation window as well as the scan time of the wide isolation window itself is much longer than that of the isolation windows where the band width is below 100 *m/z*.

Taking all the above into account, it is recommended to determine the scan times and overhead times experimentally. These data can then be used to construct the MS cycle such that it will take the expected time (will be discussed below). In this section a proposed set-up is presented to perform the experimental determination of both scan- and overhead times within 2-3 hours. In our DIA set-up, a cycle consists of 1 full MS scan and a certain number of MS/MS scans with relative narrow isolation width and 1 MS/MS scan with a relative wide isolation window (>500 *m/z*).

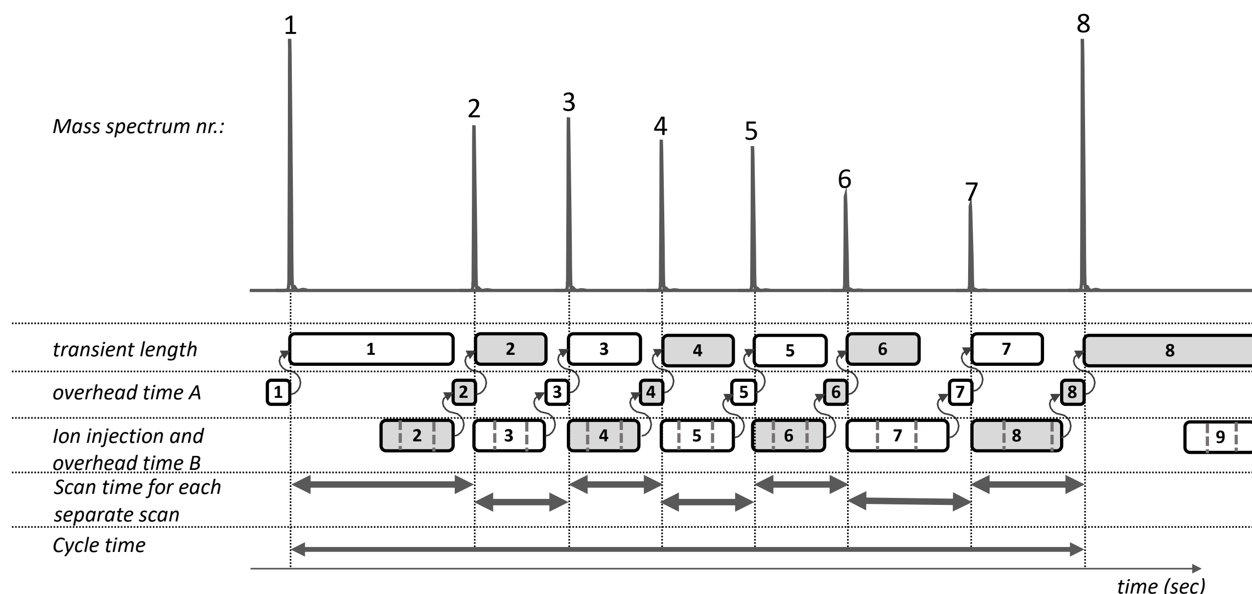



***Figure 2:*** *Chromatographic representation of a full cycle from a full scan single MS (scan 1) to a next full scan single MS (scan 8). Thus, the cycle in this example consists of a full scan single MS (scan 1) and six sequential isolation windows (scans 2-7). Besides the transient length (performance of the actual scan), the overhead times A and B determine the length of the scan. Overhead time A is time needed for ions to leave and enter the orbitrap. Ion injection and overhead time B is the time needed for ion injection and ion isolation in the quadrupole, ion transport and/or fragmentation in the HCD. Each number belongs to a certain set of ions going through all separate steps. It is assumed that the time point of each scan given in the chromatogram is the time at the start of the transient. The scan times are the time difference between the scans: the full scan single MS time is the difference in time between scan 2 and scan*

*1. The average scan time for the narrow isolation window is calculated from scan 2 - 6. The time needed around the wide isolation window is the difference between scan 8 (start of the next cycle) and scan 6.*

Using the LC-MS/MS method described in the experimental section (*LC-MS/MS methods for the determination of the actual cycle times*) the scan times are measured as described in figure 2.
First the combined time for the transient and overhead A in the full scan single MS (MS1) is measured. This will be a sum of the “transient length 1” and “overhead A 2” in figure 2.
Table 2a shows the measured scan times at various resolutions for the MS1. The settings for the MS2 are kept constant (R = 15 000). In all cases the maximum ion injection time was set substantially lower than the transient. This to ensure that the ion injection time will not be rate determining. The data from table 2a also revealed that the transient times were slightly different from the values normally reported: plotting measured scan times against the resolution gave a perfect straight line ($R^2$=1) with an offset of 8.6375 ms (this is the calculated overhead A). From this the actual transients were calculated. supplementary 1 gives more detailed information on how these actual transients were calculated. These actual transients are not essential in constructing the DIA method as proposed.

*Table 2a: measured scan times for full scan single MS at various resolutions.*

| Resolution | Max ion injection time* (ms) | Measured ion injection time (ms) | Measured scan time (ms)** | Actual transient times*** |
|---|---|---|---|---|
| 15000 | 5 | < 5 | 39.6 | 31.0 |
| 30000 | 14 | < 5 | 70.5 | 61.9 |
| 60000 | 78 | < 5 | 132.6 | 124.0 |
| 120000 | 206 | < 5 | 256.2 | 247.6 |
| 240000 | 462 | < 5 | 504 | 495.4 |

** The maximum ion injection time is set to 50 ms lower than the stated transient time [19] with the exception of the resolution 15000, which was set to 5 ms. In all cases, the maximum ion injection time was not the rate determining factor. All full scan single MS time measurements were carried out with a resolution of 15000 in the consecutive MS/MS scans.*

***Combined transient time and overhead time A.*

****Calculated using an overhead time A of 8.6 ms (see supplementary 2), assuming that transient times double each with each doubling of the resolution.*

After the scan times for the MS1 are determined, the scan times for all the MS2 scans need to be determined. In contrast to the MS1, the sum of the ion injection time and the actions in overhead B easily becomes longer than the transient, especially when a low resolution is chosen as a setting. This implies that while performing a scan it is of no use to have ion injection times much shorter than the combined times of the transient and overhead A as this will result in potentially under-collecting ions. It is therefore advised to determine the optimal maximum ion injection time. This is defined as the transient time minus the overhead time B (see figure 3B). As it can be seen from figure 3, overhead B occurs both before and after the ion injection. This representation is chosen as some of the actions in overhead B occur before and some occur after the ion injection.

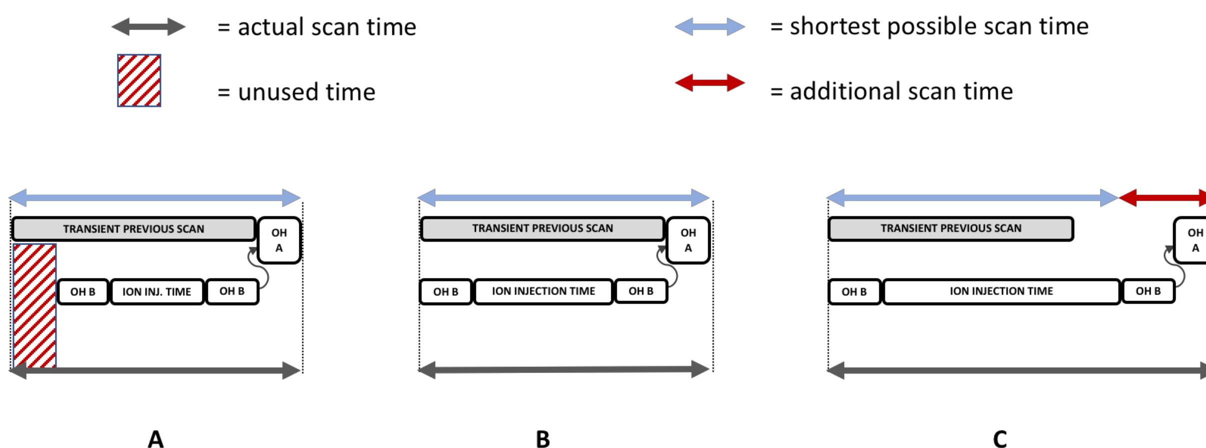



***figure 3****: Close-up of figure 2. A) maximum ion injection time and overhead B are shorter than the transient causing under-sampling of the orbitrap (red area). B) optimal maximum ion injection time allows optimal parallelization of transient scan acquisition and ion injection without leading to longer scan times. C) sum of the maximum ion injection time and the overhead B is longer than the transient, leading to longer scan times.*

To determine the optimal maximum ion injection time for the narrow isolation windows in the MS/MS, we first calculated the overhead time by setting the maximum injection time substantially higher than the transient and then subtracted the transient time from the measured scan time (situation figure 3C, data in table 2b). we then set the maximum injection time substantially lower than the transient to measure the minimum scan time (situation figure 3A). We then subtracted the calculated overhead from the minimum scan time to determine the optimal maximum ion injection time (see table 2c). It is important to notice that these measurements will be done in an

ion limited situation. In other words, the maximum ion injection time is reached before the AGC target (3e6 in our case) is reached.

Table 2b: Measurement of the scan time when ion injection time exceeds the transient time.

| | Resolution | maximum ion injection time (ms)* | Measured scan time narrow isolation width (ms) | Calculated overhead (ms)** |
|---|---|---|---|---|
| ion injection >> transient time | 15000 | 50 | 82 | 32 |
| | 30000 | 80 | 112 | 32 |
| | 60000 | 150 | 181 | 31 |

** The maximum ion injection time is set longer than the transient time. with the exception of the resolution 15000, which was set to 5 ms. In all cases, the maximum ion injection time was not the rate determining factor (see figure 3C).*

*** The overhead time is calculated by subtracting the maximum ion injection time from the measured scan time. Overhead time here is the sum of overhead A and overhead B. For both the calculation of the optimal maximum ion injection time and the overhead time it is important to notice that the measurements need to be done in an ion limiting situation i.e. that the maximum ion injection time is used in its full.*

Table 2c: Measurement of the minimal scan time and calculation of optimal maximum ion injection time.

| | Resolution | maximum ion injection time (ms)* | Measured scan time narrow isolation width (ms) | Calculated overhead (ms) (from Table 2b) | Optimal maximum ion injection time (ms)** |
|---|---|---|---|---|---|
| ion injection << transient time | 15000 | 5 | 41 | 32 | 9 |
| | 30000 | 14 | 73 | 32 | 41 |
| | 60000 | 78 | 136.8 | 31 | 106 |

** The maximum ion injection time is set to 50 ms lower than the transient times mentioned in the instrument manual with the exception of the resolution 15000, which was set to 5 ms. In all cases, the maximum ion injection time was not the rate determining factor (see figure 3A).*

*** The optimal maximum ion injection time is calculated by subtracting the calculated overhead from the measured scan time where the transient length is the rate limiting factor (situation in figure 3B)*

It is not recommended to use the "auto" setting for the maximum ion injection time. It gives the illusion that the most optimal ion injection time is chosen. This is not the case since it sets the maximum ion injection time at a preset "free ion time" for a particular resolution: 22 ms at a resolution of 15000, 54 ms at 30000, 118 ms at 60000 and so on. Since the real overhead is longer than 10 ms, it will logically lead to a situation as shown in figure 3C: the overhead and ion injection time together is longer than the transient time.

The time used around the isolation window with the wide isolation is longer than for narrow isolation windows (Table 2d). This is due to the fact that the isolation in the quadrupole becomes rate determining when it needs to sweep over a wide range: this occurs both for the isolation of ions prior to the wide isolation and isolation of ions prior to the full scan MS resulting in longer measured scan times. Supplementary 3 shows that the quadrupole selection time increases discontinuously when the isolation width of a window increases, therefore wide isolation widths should preferably be avoided.

Table 2d: Measured scan time around the wide isolation width.

| | Resolution | Set maximum ion injection time (ms) | Measured scan time wide isolation width (ms) |
|---|---|---|---|
| ion injection << transient time | 15000 | 5 | 187.6 |
| | 30000 | 14 | 249 |
| | 60000 | 78 | 438.6 |

*5) Construct the isolation windows at the chosen conditions.*

Construction of the number of isolation windows and their isolation width is now rather straightforward using the data from table 2 and figure 2. As an example, a method is constructed with a resolution of 60000 in the full scan MS, a resolution of 15000 and the optimal maximum injection time (9 ms) in the MS/MS (narrow isolation windows in the *m/z* range 400 to 900, one wide isolation window from *m/z* 900 to 1650 - isolation width) and 2 sec cycle time.

Since 132.6 ms is used by the full scan MS (at a resolution of 60000 - table 2a) and 187.6 ms is used for the wide isolation window (that is the total time of the scan preceding this isolation window and the scan time of the isolation window itself at resolution 15000 - table 2d), 1679.8 ms is left for narrow isolation windows, which take 41 ms each (From table 2c, assuming optimal maximum ion injection times as shown in figure 3b). This results in 42 narrow isolation windows with an isolation width of 12.4 *m/z*.

In our experiments we constructed several DIA methods with the purpose to investigate which MS2 settings were optimal. Table 3 shows the investigated settings. The 10 tested methods can be classified as methods where the ion injection time is optimal (i.e. the sum of the overhead time B and the ion injection time does not exceed the sum of the transient time and overhead time A - see figure 3B) or as methods where the ion injection time exceeds the transient (as in figure 3C). It was decided not to explore methods where the maximum ion injection time was considerably lower than the optimal maximum ion injection time as these methods were to expect to “under-collect” ions (situation figure 3A). In all cases the resolution of the full scan single MS was set to 60 000.

Steps 4 and 5 can be carried out using an alternative strategy where the time of scan events are not measured directly (using the chromatographic representation as shown in figure 2) but rather the time of each whole cycle. This alternative approach is shortly explained in supplementary 4.

*6) evaluate the DIA performance at various resolutions and various settings*

After the various DIA methods have been run, a three way evaluation is to be performed

a) evaluation of the actual cycle times to check the methods performance in relation to the settings

b) qualitative performance by means of number of peptide and protein identifications

c) quantitative performance by means of CV's and percentage of peptides and proteins above the LOQ.

*Table 3: Measured number of points per peak, measured cycle time and percentage of scans which reach maximum ion injection time (acquisition sequence order see Supplementary 5)*

| Method settings for the Q Exactive HF | | | Measured values | | |
|---|---|---|---|---|---|
| Resolution | #isolation windows | Max IT | Points per peak | % Max IT reached MS2 (n=3) | M.cycle time (sec) n=3 |
| 15 000* | 43 | 9 | 9 | 99.8 | 2.07 |
| 30 000 * | 24 | 41 | 9-10 | 97.0 | 2.14 |
| 60 000* | 12 | 106 | 10 | 82.1 | 2.18 |
| 120 000*^ | 6 | 325 | 5-6 | 56.3 | 2.71 |
| 15 000** | 24 | 41 | 7 | 96.3 | 1.99 |
| 15 000** | 12 | 106 | 12-13 | 81.5 | 1.78 |
| 15 000** | 6 | 325 | 12 | 44.0 | 1.54 |
| 30 000** | 12 | 106 | 11 | 80.8 | 1.89 |
| 30 000** | 6 | 325 | 9 | 46.4 | 1.68 |
| 60 000** | 6 | 325 | 8 | 47.9 | 1.94 |

** methods using a resolution of 60000 in the full scan and 15000, 30000 or 60000 in the MS/MS designed with the number of isolation windows needed with an optimal ion injection time (situation Figure 3B)*

*** methods using a resolution of 60000 in the full scan and 15000, 30000 or 60000 in the MS/MS designed with the number of isolation windows needed with a longer ion injection time (situation Figure 3C, this situation is the case when operating in the "auto" setting for the maximal ion injection time)*

*^ at 120000 the number of isolation windows as well as their isolation widths are estimated. This was necessary since cycle time measurements for the 120000 resolution in MS2 as performed in table 2 were not consistent.*

*Evaluation of the actual cycle time*

Table 3 shows the various DIA settings evaluated. In addition to the various MS2 resolutions with optimal scan times as shown in table 2, DIA settings with ion injection times longer than the "optimal maximum ion injection time" were also explored. The rationale behind this is that longer ion injection times will lead to more ions in the orbitrap, which in turn might lead to better LOQ's in ion poor parts of the chromatogram. From the table 3 it can be seen that the measured cycle time for 15000, 30000 and 60000 are close to the calculated 2 sec. Decreasing the number of isolation windows at a certain resolution leads to shorter cycle times, mainly because in some of the isolation window scans the AGC target is reached before the max ion injection time has elapsed. This can be seen as an advantage with respect to reliable quantitation of the chromatographic peaks since more points per peak were obtained. The relative long measured cycle times for the 120 000 resolution in the MS2 are at this point not understood.

*Evaluation of identifications*

Table 4B shows that at a resolution of 15000 with 24 isolation windows and a resolution of 60000 with 12 isolation windows, the number of protein identifications exceeds 2950. Slightly suboptimal results were achieved with a resolution of 30000 and 24 isolation windows and a resolution of 120000 with 6 isolation windows, both of which resulted in slightly fewer than 2950 protein group identifications. All other settings were considered to be non-optimal with respect to protein group identifications. Logically, the number of peptide identifications measured follows the same trend (see table 4A). For the 30000 and the 60000 resolution an increasing number of isolation windows seem to improve the number of identifications slightly. The same goes for the 15000 resolution, except for the 43 isolation windows. At this number of isolation windows, the number of identifications drops dramatically. This might be caused by severe underfilling of the orbitrap at this setting. Independent of the resolution, the percentage of scans that reached the maximum ion injection time is strongly dependent on the number of isolation windows. The methods using 43 and 24 isolation windows have maximum ion injection times which are so short, that most of the scans are performed in a situation where AGC target has not been reached (percentage of scans that reached the maximum ion injection time > 95%, see table 3). Decreasing the number

of isolation windows to 12 or lower improves the percentage of scans that reaches the maximum ion injection time.

***Table 4:*** *Average number of peptide (A) and protein (B) identifications. All identifications are averaged from Spectronaut, (DIA) (n=3). One-way ANOVA analysis gave a p-value of $4.31*10^{-17}$ for the peptide identifications and a p-value of $2.74*10^{-26}$ for the protein identifications.*

**A**

| Resolution | # Isolation windows: 43 | 24 | 12 | 6 |
|---|---|---|---|---|
| 15000 | 12297 | 21388 | 14965 | 13417 |
| 30000 | | 19654 | 17497 | 18109 |
| 60000 | | | 20051 | 20329 |
| 120000 | | | | 21013 |

**B**

| Resolution | # Isolation windows: 43 | 24 | 12 | 6 |
|---|---|---|---|---|
| 15000 | 2093 | 3025 | 2468 | 2095 |
| 30000 | | 2937 | 2759 | 2597 |
| 60000 | | | 2992 | 2829 |
| 120000 | | | | 2911 |

The steep decrease in protein identifications going from 24 to 6 isolation windows at 15000 is probably caused by the increased complexity of the fragmentation spectrum at this resolution: the resolution becomes the limiting factor in the identification. This is supported by the fact that the number of identifications (approx. 2900) done at 60000 and 120000 using 12 and/or 6 isolation windows are of similar size.

*Evaluation of LOQ and CV*

Since there are no standards included in these analyses, it will be impossible to evaluate the absolute value for the LOQ for certain peptides or proteins. The evaluation of the DIA methods are based on the FDA's definition that the method's CV should not exceed 20%. In other words, when the CV of a certain peptide or protein exceeds 20%, it can only be identified by the presence of its fragment ions but not be quantitated.

***Table 5:*** *Percentage of peptides (A) and proteins (B) quantified with a CV < 20 % (number of peptides or proteins with a CV<20% / number of identified peptides or proteins in full profile). The higher the number, the darker the color. C) median CV of protein quantifications, the lower the number, the darker the color. Data from Spectronaut analysis. (n=3)*

**A**

| Resolutic | # isolation windows: 43 | 24 | 12 | 6 |
|---|---|---|---|---|
| 15000 | 83.6 | 83.8 | 85.0 | 81.8 |
| 30000 | | 83.7 | 85.6 | 81.8 |
| 60000 | | | 86.0 | 83.6 |
| 120000 | | | | 81.8 |

**B**

| Resolutic | # isolation windows: 43 | 24 | 12 | 6 |
|---|---|---|---|---|
| 15000 | 80.5 | 78.9 | 83.4 | 79.6 |
| 30000 | | 79.9 | 83.3 | 79.3 |
| 60000 | | | 83.5 | 79.1 |
| 120000 | | | | 76.8 |

**C**

| Resolutic | # isolation windows: 43 | 24 | 12 | 6 |
|---|---|---|---|---|
| 15000 | 5.2 | 6.0 | 4.3 | 5.6 |
| 30000 | | 4.9 | 4.7 | 5.6 |
| 60000 | | | 4.6 | 6.4 |
| 120000 | | | | 7.8 |

The best performance was obtained with 12 isolation windows: regardless of the resolution used in MS2, approximately 83 % of all signals can be quantified with sufficient confidence (defined with CV < 20%) (see table 5A). In all the other settings, not more than 80% or less of all the signals can be quantified (with a value above the LOQ). This is similar for the protein group quantitation. The share of protein groups quantified with sufficient confidence is best, up to 85-86%, for the 12 isolation window measurements, regardless of the resolution (see table 5B). Additionally, none of the settings gave lower than 80 % protein groups quantified with sufficient confidence.

The median CV did not vary that much either, most of the median CVs were around 5%, except for the runs with a resolution of 120 000 (table 5C). The CVs were most likely so low because the observed number of data points per chromatographic peaks (5-6) was significantly lower than the target number of points per peak (10). (see table 3). This might explain why the 20% CV data and the median CV was lower at this resolution compared to the runs with lower resolutions.

*7) Determine the settings optimal for its purpose*

All in all, within the tested DIA settings on the given sample, both MS2 resolution 15000 with 24 isolation windows and MS2 resolution 60000 with 12 isolation windows gave similar results with respect to the number of identifications. The 60000 12 isolation window seems slightly better when it comes to quantitation reliability.

Both methods can thus be seen as optimal DIA methods for our instrument setup and can be used as such.

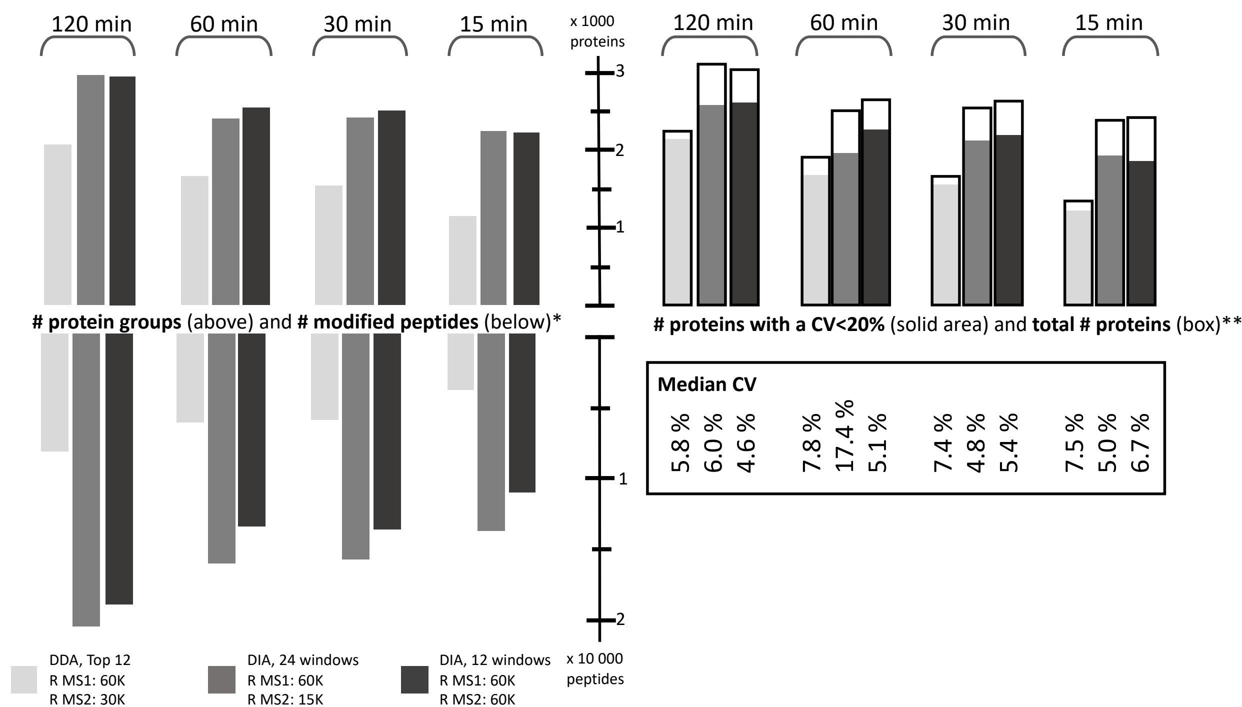
The following section uses both methods in the comparison with the conventional DDA, to evaluate the potential of shorter gradients in order to get shorter analysis times.

*8) Comparison with DDA and evaluation of the DIA method length.*

This tutorial has taken an optimal DDA method [6] as the basis for the construction and evaluation of several DIA methods with a gradient time of 120 min. How does an optimal DIA method now relate to this DDA method? Additionally, what is the effect of shortened gradient times on the DIA and DDA methods? In other words, how short can an analysis be performed without compromising too much on the quality of the identifications and the quantifications. Shorter gradients with steeper slopes resulted in sharper chromatographic peaks and thus less points per peak as the same cycle time (2 sec).

Both the 24 isolation windows at 15000 resolution and the 12 isolation windows at 60000 resolution as well as the initial DDA method were compressed in their gradient length. Besides the 120 min runs, also 60, 30 and 15 minutes gradients were carried out.
Figure 4 shows the number of identified peptides, proteins, the median CV of each tested condition as well as the number of proteins which are quantified with a CV<20%.

***Figure 4:*** *Number of identifications and quantitative performance (by means of the CV) for both DDA and DIA at various resolutions and various gradient lengths (acquisition sequence order see Supplementary 5): * number of quantifiable "modified peptides" or "protein groups" in both MaxQuant (for the DDA) and Spectronaut (for the DIA). MaxQuant was carried without "match between runs" and with an LFQ minimal ratio count of 1. Only peptides/proteins which had LFQ intensities > 0 in all three runs are shown. For the Spectronaut the full profile numbers are shown. Both reversed database and possible contaminants were removed from the list of quantifiable peptides/proteins.*
*** Spectronaut calculates the CV of a quantifiable protein when it has an intensity > 0 in at least 2 of the 3 runs. The same principle is also applied on the data obtained from MaxQuant. The total number of proteins with a calculated CV is shown in brackets.*
*Settings described under * and ** are used to get as comparable Spectronaut and MaxQuant data as possible.*




*Direct performance comparison between 120 min DDA and 120 min DIA*

The direct comparison between the original DDA method and the optimal DIA settings using a 120 minute run shows (Figure 4) that the DIA outperforms the DDA with respect to number of identifications and number of reliably quantitated proteins: the number of protein group identifications in the DIA is 43% higher than the DDA, additionally, 22-23% more proteins are quantified with a CV lower than 20% in the DIA compared to the DDA.

*Effect of shorter gradients on the number of identifications and quantification.*

Figure 4 also shows that the DIA runs perform better with respect to number of identifications and the quantitation compared to the DDA which was run with a similar gradient. Interestingly, the DIA runs performed with a 30 minutes gradient were comparable in number of protein group identifications and quantitation to the DDA run performed with a 120 minutes gradient; thus demonstrating the DIA's potential for either deeper proteome coverage or higher throughput, depending on the user's needs.

*9) Wrapping up: Is it worth the effort to develop and run tailormade DIA methods?*

As in all scientific cases, there is no straightforward answer to this. It depends to a large extend on the aspect of the analysis which the researcher considers as the most important. This can be

the aspect of time, depth of proteome coverage, interest in reliable identifications or the quantitative power of the DIA method. Usually it is a mix of these.

However, setting up a tailor-made DIA method can be time consuming. In addition to the development of the method, the construction of in-house spectral libraries is also time consuming, since in many cases fractionation and a fair amount of DDA analyses need to be carried out. However, the time lost in development can be partially offset since spectral library searches are in general faster due to the restricted search space of a library that they have, and due to the utilization of the known intensities of the fragment ions, higher specificity and sensitivity. The disadvantage is that only peptide ions present in the library can be identified.[20].

Construction of in-house spectral libraries can be circumvented when open access spectral libraries are used [21]. In this case one needs to be aware of the fact that the quality of such a spectral library will differ from the in-house build spectral library. This might have an impact on the qualitative data. On the other hand, Collins et al.[22] showed highly reproducible DIA results between laboratories (all using the same equipment, methods and spectral library) indicating that optimized DIA methods could be transferred to other laboratories when similar settings, libraries and equipment is used.

Once the optimal DIA method is in place, its use can provide significant time savings compared to DDA. For example, the 30 minute DIA runs gave qualitative and quantitative comparable results to 120 minute DDA runs. The consideration of time depends strongly on the number of samples to be run: it may not be worth the effort to set-up a DIA for just a few runs, while it would gain tremendous amount of time when the same type of samples were run routinely.

Seen from a protein identification and / quantification point of view, DIA seems to perform better than the DDA: more than double the number of peptides were identified, which led to an increase of approx. 25% in protein identifications. Median CVs were lower for the DIA and the number of proteins groups quantified with a CV lower than 20 % was higher. Although it is an easy choice from this perspective, it again will come down to the trade off in time: how much does the aspect of protein identification and / quantification weigh in order to justify the time use in constructing the best DIA method.

It should be stressed that the strategy development in this tutorial was carried out given a certain framework.

- We chose to operate with 10 data points per chromatographic peak, which in our chromatographic system resulted in approx. 2 sec per cycle. Other cycle time settings will result in a different DIA settings to test, but it will not affect the strategy to follow for the development of the method.

- The settings which were evaluated were restricted to 43, 24, 12 and 6 isolation windows. These numbers of isolation windows followed from the calculation of the optimal maximum ion injection time. Although we choose to use these number of isolation windows in settings where the maximum ion injection time was longer than the optimal one, other numbers of isolation windows could have been investigated too. This would of course increase the time needed for evaluation and method development.
- It is also recommended to tailor the DIA method for a certain type and complexity of sample. Samples which are less complex as well as samples which are more complex might give different optimal DIA settings than the ones presented for the Pseudomonas Aeruginosa.
- Comparing data from Spectronaut and MaxQuant is challenging. This probably holds for all comparisons of seemingly similar data from different proteomic platforms. Thorough consideration should be given to how a proteomic platform generates the numbers. E.g. is the number of protein identifications in one proteomic platform generated the same way as in the other?

The strategy for developing a DIA method (steps 3 – 5) was also tested using a Fusion Tribrid mass spectrometer to investigate if this strategy can be used on a mass spectrometer with a different geometry. We showed that this approach is also necessary to be carried out on the Fusion Tribrid mass spectrometer in order to determine the best DIA method setting (Supplementary 3). Due to the differences between the Q Exactive HF and the Fusion Tribrid mass spectrometers the measured scan times vary when similar settings on both instruments were used despite similar orbitrap transient times for the same resolution. Optimal ion injection times can be longer in the Fusion Tribrid mass spectrometer as it uses less overhead time allowing a larger portion of the ions to be accumulated. Despite the differences between the instruments, the strategy of setting up the DIA method worked also for the Fusion Tribrid mass spectrometer. Evaluation of the performance of several DIA methods was further only carried out on the Q Exactive HF mass spectrometer. This evaluation must also be performed for the Fusion Tribrid mass spectrometer if the optimal DIA settings need to be determined.

***Conclusion***

We have described in this tutorial a general approach to develop and evaluate DIA methods on orbitrap mass spectrometers. The whole strategy is tested using a digested lysate of

Pseudomonas Aeruginosa on an Q Exactive HF mass spectrometer, but can also be carried out on a Fusion Tribrid mass spectrometer.

Key elements in this development are spectral library building, measurement of the actual scan times and total cycle time to ensure the correct pre-decided number of data points per chromatographic peaks. Following this it allows construction of the DIA method in terms of the number of isolation windows and their width. To obtain the best DIA settings evaluation of optimal ion injections times and ion injection times longer than optimal need to be carried out using median CV, number of peptides and proteins quantified with a CV < 20%, number of identified peptides and number of identified proteins as response variables. Within the chromatographic time frame of 120 minutes, DIA performs better than the DDA. Furthermore, shorting the chromatographic gradient length down to 30 minutes for the DIA runs gave similar qualitative and quantitative results as DDA runs with a chromatographic gradient length of 120 minutes. However, all in all the use of DIA is very time demanding when only few samples need to be analyzed.

### *Supporting Information*

The following supporting information is available free of charge at ACS website http://pubs.acs.org

Supplementary 1: Calculation of the actual overhead time A and actual transient time.

Supplementary 2: DIA on the Fusion Tribrid Mass Spectrometer.

Supplementary 3: Quadrupole selection time vs. isolation width.

Supplementary 4: Description of the model-free method for calculating number of SWATHs and maximum injection times

Supplementary 5: Acquisition sequence of the experiments

### *Acknowledgement*

The authors wish to thank Prof. Dianna Newman and David Basta (California Institute of Technology) for preparation and supplying the Pseudomonas Aeruginosa (strain PA14) digested lysates. The authors also thank Roxana Eggleston-Rangel from the Proteome Exploration Laboratory (Beckman Institute at the California Institute of Technology) for her technical assistance during the nano LC-MS/MS analyses. The Proteome Exploration Laboratory is supported by, the Beckman Institute, and NIH 1S10OD02001301.

***References***


1. James, P., Protein identification in the post-genome era: the rapid rise of proteomics. *Quarterly Reviews of Biophysics* **1997,** *30* (4), 279-331.
2. Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003,** *422*, 198.
3. Link, A. J.; Hays, L. G.; Carmack, E. B.; Yates, J. R., Identifying the major proteome components of Haemophilus influenzae type-strain NCTC 8143. *Electrophoresis* **1997,** *18* (8), 1314-1334.
4. Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R.; Iii, Direct analysis of protein complexes using mass spectrometry. *Nature Biotechnology* **1999,** *17*, 676.
5. Yates, J. R.; Eng, J. K.; McCormack, A. L.; Schieltz, D., Method to Correlate Tandem Mass Spectra of Modified Peptides to Amino Acid Sequences in the Protein Database. *Analytical Chemistry* **1995,** *67* (8), 1426-1436.
6. Kelstrup, C. D.; Young, C.; Lavallee, R.; Nielsen, M. L.; Olsen, J. V., Optimized Fast and Sensitive Acquisition Methods for Shotgun Proteomics on a Quadrupole Orbitrap Mass Spectrometer. *Journal of Proteome Research* **2012,** *11* (6), 3487-3497.
7. Kalli, A.; Smith, G. T.; Sweredoski, M. J.; Hess, S., Evaluation and Optimization of Mass Spectrometric Settings during Data-dependent Acquisition Mode: Focus on LTQ-Orbitrap Mass Analyzers. *Journal of Proteome Research* **2013,** *12* (7), 3071-3086.
8. Doerr, A., DIA mass spectrometry. *Nature Methods* **2014,** *12*, 35.
9. Venable, J. D.; Dong, M.-Q.; Wohlschlegel, J.; Dillin, A.; Yates Iii, J. R., Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nature Methods* **2004,** *1*, 39.
10. Silva, J. C.; Gorenstein, M. V.; Li, G.-Z.; Vissers, J. P. C.; Geromanos, S. J., Absolute Quantification of Proteins by LCMSE : A Virtue of Parallel ms Acquisition. *Molecular & Cellular Proteomics* **2006,** *5* (1), 144-156.
11. Gillet, L. C.; Navarro, P.; Tate, S.; Röst, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R., Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Molecular & Cellular Proteomics* **2012,** *11* (6), O111.016717.
12. Kelstrup, C. D.; Bekker-Jensen, D. B.; Arrey, T. N.; Hogrebe, A.; Harder, A.; Olsen, J. V., Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics. *Journal of Proteome Research* **2018,** *17* (1), 727-738.
13. Bruderer, R.; Bernhardt, O. M.; Gandhi, T.; Xuan, Y.; Sondermann, J.; Schmidt, M.; Gomez-Varela, D.; Reiter, L., Optimization of Experimental Parameters in Data-Independent Mass Spectrometry Significantly Increases Depth and Reproducibility of Results. *Molecular & Cellular Proteomics* **2017**, 2296-2309.
14. Koopmans, F.; Ho, J. T. C.; Smit, A. B.; Li, K. W., Comparative Analyses of Data Independent Acquisition Mass Spectrometric Approaches: DIA, WiSIM-DIA, and Untargeted DIA. *Proteomics* **2018,** *18* (1), 1700304.
15. Moradian, A.; Porras-Yakushi, T. R.; Sweredoski, M. J.; Hess, S., A Targeted MRM Approach for Tempo-Spatial Proteomics Analyses. In *Proteomics in Systems Biology: Methods and Protocols*, Reinders, J., Ed. Springer New York: New York, NY, 2016; pp 75-85.
16. Tyanova, S.; Temu, T.; Cox, J., The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **2016,** *11*, 2301.

17. Cox, J.; Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **2008,** *26*, 1367.
18. Lundanes, E.; Reubsaet, L.; Greibrokk, T., *Chromatography: Basic Principles, Sample Preparations and Related Methods*. Wiley: 2013; p 224.
19. Scheltema, R. A.; Hauschild, J.-P.; Lange, O.; Hornburg, D.; Denisov, E.; Damoc, E.; Kuehn, A.; Makarov, A.; Mann, M., The Q Exactive HF, a Benchtop Mass Spectrometer with a Pre-filter, High-performance Quadrupole and an Ultra-high-field Orbitrap Analyzer. *Molecular & Cellular Proteomics* **2014,** *13* (12), 3698-3708.
20. Deutsch, E. W., Tandem Mass Spectrometry Spectral Libraries and Library Searching. In *Data Mining in Proteomics: From Standards to Applications*, Hamacher, M.; Eisenacher, M.; Stephan, C., Eds. Humana Press: Totowa, NJ, 2011; pp 225-232.
21. Muntel, J.; Xuan, Y.; Berger, S. T.; Reiter, L.; Bachur, R.; Kentsis, A.; Steen, H., Advancing Urinary Protein Biomarker Discovery by Data-Independent Acquisition on a Quadrupole-Orbitrap Mass Spectrometer. *Journal of Proteome Research* **2015,** *14* (11), 4752-4762.
22. Collins, B. C.; Hunter, C. L.; Liu, Y.; Schilling, B.; Rosenberger, G.; Bader, S. L.; Chan, D. W.; Gibson, B. W.; Gingras, A.-C.; Held, J. M.; Hirayama-Kurogi, M.; Hou, G.; Krisp, C.; Larsen, B.; Lin, L.; Liu, S.; Molloy, M. P.; Moritz, R. L.; Ohtsuki, S.; Schlapbach, R.; Selevsek, N.; Thomas, S. N.; Tzeng, S.-C.; Zhang, H.; Aebersold, R., Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nature Communications* **2017,** *8* (1), 291.

For TOC only: