Designing a DIA Method

Written by Brian C. Searle. Email me at bsearle@systemsbiology.org or contact me @briansearle on Twitter to get answers to any questions you have.

- v1.2, updated on May 10, 2020
- v1.1, updated on March 28, 2018

Summary and further resources

Data independent acquisition (DIA) is mass spectrometry acquisition strategy that represents a series of compromises between comprehensive detection and selective quantitation where the goal is to detect as many peptides as possible, but in a way that maintains quantitative rigor. With modern mass spectrometers we typically can only scan at a maximum rate of 10-20 Hz without sacrificing spectrum quality. An additional constraint is that at least 6-8 measurements are required to describe a quantitative peak. DIA methods make several compromises to manage these constraints. In particular, instead of trying to monitor all peptides, the measurement range is typically limited to measure peptides in a restricted precursor window (e.g. 400-1000 m/z). In addition, precursor isolation windows are widened such that multiple peptides are usually isolated together and co-fragmented, resulting in higher interference.

This document is a quickstart guide to rapidly building a DIA method tailored for your instrument and experiment. It is focused on developing methods for Orbitrap-based instruments, but these concepts are generally applicable to other instrument configurations as well. A thorough investigation into the intuition behind these best practices is detailed in our paper in *Molecular and Cellular Proteomics*:

Pino LK, Just SC, MacCoss MJ, and Searle BC. (2020) Acquiring and Analyzing Data Independent Acquisition Proteomics Experiments without Spectrum Libraries. *Mol Cell Proteomics.* Apr 20. pii: mcp.P119.001913. https://doi.org/10.1074/mcp.P119.001913

A further DIA resource is a recorded talk on this subject for the Northeastern May Institute 2020 organized by Olga Vitek and Meena Choi. This talk is freely available at: https://www.youtube.com/watch?v=RidYXjvAk0s

Designing a windowing strategy

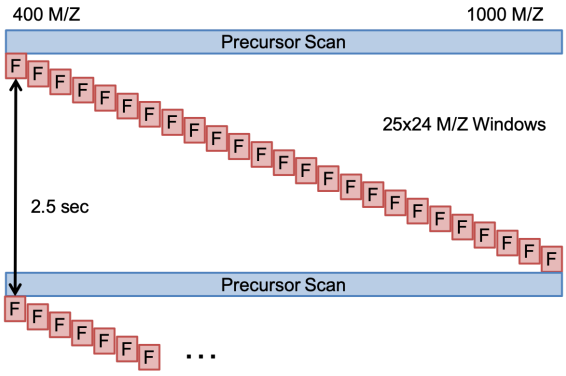
Summarizing Pino *et al*,[1] current Orbitrap-based instruments that are suitable for DIA fall into two categories:

10 Hz instruments: Thermo Fusion, QE, QE+, QE-HF

20 Hz instruments: Thermo Lumos, Eclipse, QE-HFX, Exploris

While most of these instruments can technically collect MS/MS faster than the above rates, sensitivity tends to be the important limiting factor. ToF instruments can scan much faster than

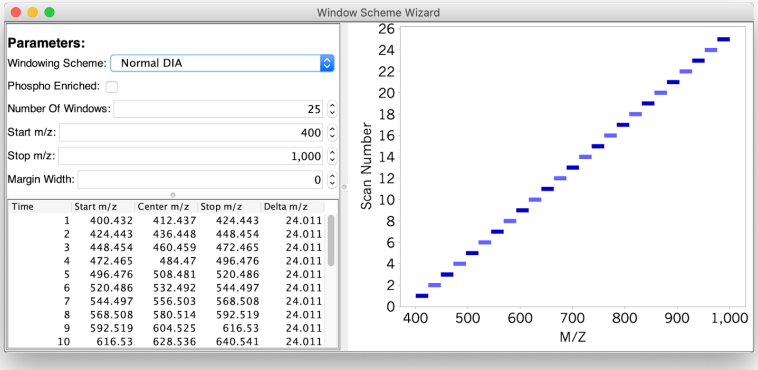
Orbitraps but require scan averaging to produce acceptable MS/MS. In general, we recommend considering these instruments as similar to 20 Hz Orbitrap instruments after scan averaging.



We structure windowing methods assuming 10 measurements across an average peak so that regions of the chromatogram with narrower peaks will not be underrepresented. We calculate the necessary cycle time and windowing scheme from the average peak width and maximum scan rate. Assuming peak widths of 25 seconds (typical for 300mm-400mm columns and 90 min gradients) and 10 Hz, we require cycle times of 2.5 seconds, resulting in 25 windows per cycle. This allows for 25x 24 m/z-wide precursor isolation windows to cover 600 m/z in windowing range (400-1000 m/z). However, if we use a faster scanning instrument that can achieve 20 Hz, we can use up to 50 windows per cycle. This allows for 50x 12 m/z-wide precursor isolation windows to cover the same 600 m/z in windowing range. Wider peptide peak widths produced by shorter columns (e.g. 150-200mm) or slower chromatography enable more scans to achieve 10 measurements per peak, allowing for narrower precursor isolation windows. Similarly, longer columns may require balancing with wider precursor isolation windows.

Choosing optimal precursor isolation windows

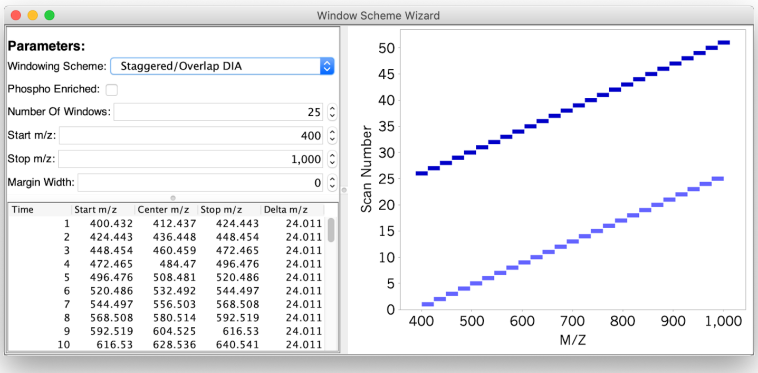
EncyclopeDIA[2] makes it easy to set up a DIA windowing scheme. After downloading and installing EncyclopeDIA (freely available at https://bitbucket.org/searleb/encyclopedia), navigate to the “Help/Window Scheme Wizard” menu option. This launches the following dialog:



From the windowing strategy calculations, enter the appropriate number of windows, and the desired start/stop m/z values. The table and graphic update automatically based on these selections. Here the default is designed for 10 Hz instruments with 25 windows measuring from 400 to 1000 m/z. Fractional optimized window placements put the window boundaries in regions between nominal m/z values where peptides are unlikely to exist. These window boundaries are designed for typical proteomes. However, some PTMs can adjust these calculations and if peptides are enriched for a certain type of amino acid (e.g. phospho) then these boundaries should be adjusted appropriately. Some older instruments (e.g. Thermo Fusion or QE) have quadrupole geometries that are not optimized for DIA. These instruments can benefit from small (<0.5 m/z) margins added to the window width. Newer instruments with segmented quadrupoles have relatively flat transmission efficiency across the precursor isolation window such that using optimized window placements sufficiently makes margins unnecessary.

Using staggered windows for Orbitrap instruments

For Orbitrap instruments, we recommend using “staggered” window schemes that offset every other cycle by 50%:

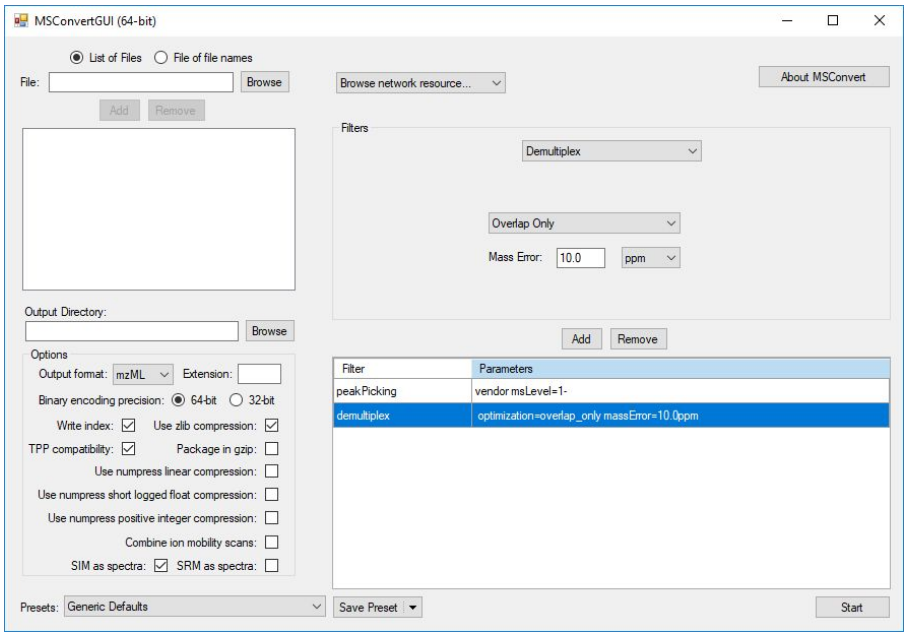


Data collected with this windowing scheme can be demultiplexed using the Proteowizard[3] GUI tool, which uses previous and next cycles to computationally separate each N-width MS/MS into two N/2-width MS/MS spectra.[4] A tutorial for how to perform this demultiplexing is described here:
https://www.mcponline.org/content/mcprot/suppl/2020/04/20/P119.001913.DC1/157767_2_supp_511351_q8sty4.pdf

Briefly, Proteowizard can be freely downloaded (from http://proteowizard.sourceforge.net/) to convert vendor-specific MS raw files into HUPO standard mzML interchange files that are vendor neutral. We recommend these settings for DIA window deconvolution on Orbitrap instruments using the command line:

```
msconvert.exe --zlib --64 --mzML --simAsSpectra --filter "peakPicking
true 1-" --filter "demultiplex optimization=overlap_only" *.raw
```

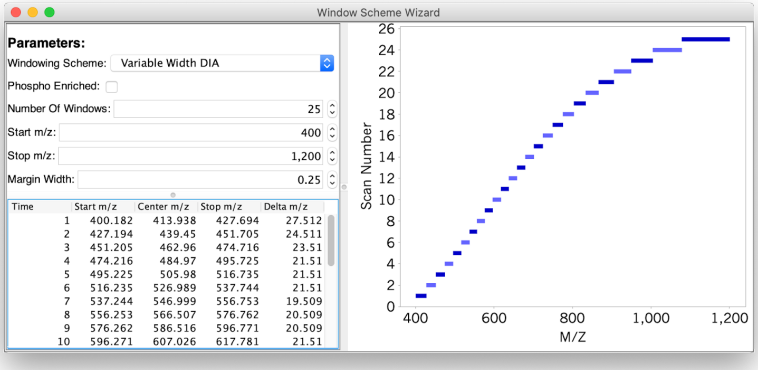
Alternatively, the MSConvert GUI can be used:



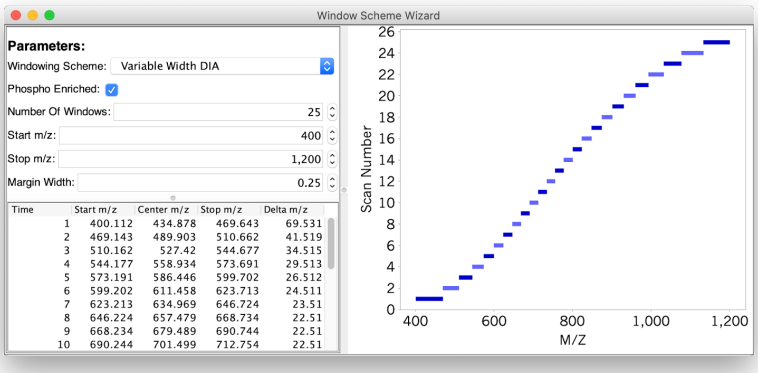
In both cases, the order of “filters” is important: the “peakPicking” filter must be listed first to enable vendor-library peakpicking. Note: if staggered windows are used, then it is important that margins are not added. Also note: make sure “SIM as spectra” is checked for tribrid instruments.

Using variable-width windows on ToF instruments

ToF instruments typically are able to scan faster than Orbitrap instruments, but naturally produce noiser MS/MS spectra because they collect true profile data. Consequently, measurements made by ToFs may not be deconvoluted as easily as those made by Orbitrap instruments. In this case, we normally recommend using variable-width windows and taking advantage of small margins and potentially extended precursor isolation ranges:

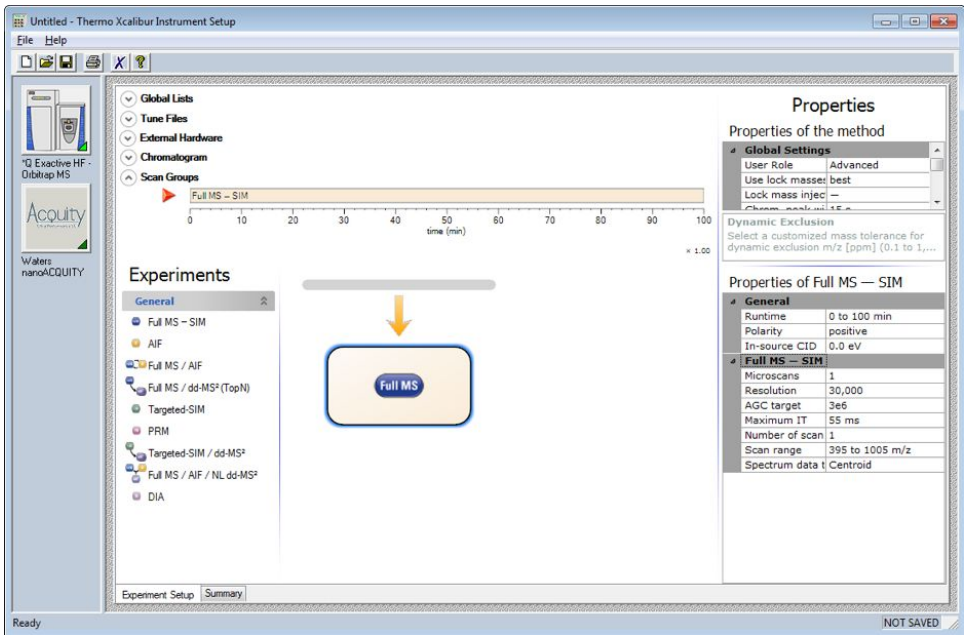


Note: as with optimized window placements, the optimal variable-width windows change when analyzing PTM-enriched samples such as samples of phosphopeptides:



Setting up a DIA instrument method

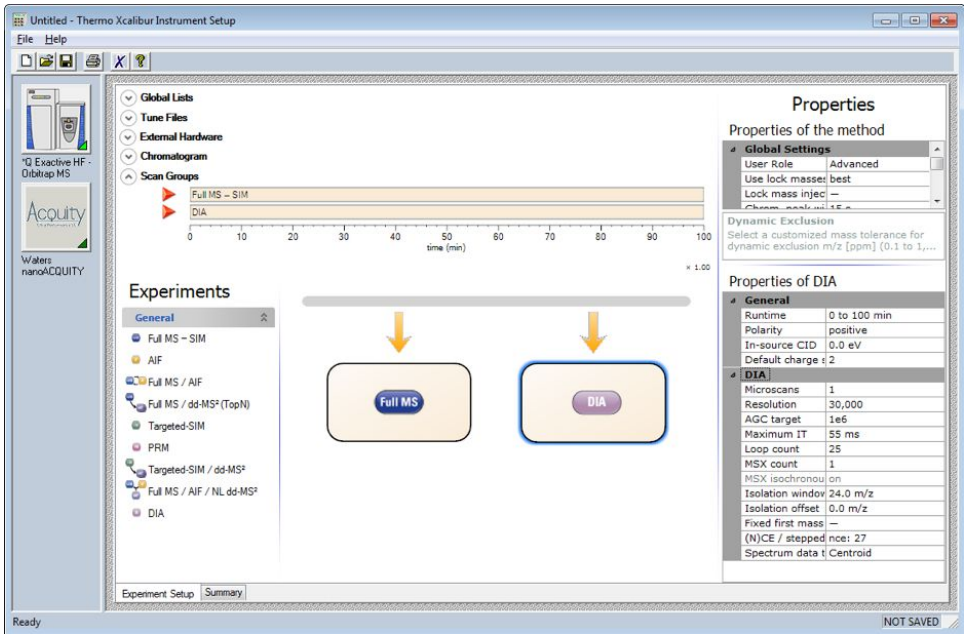
Here we use the Thermo QE-HF method editor, but many of these settings are transferable to other platforms. Starting with a new method, we recommend adding a precursor scan followed by DIA MS/MS scans:



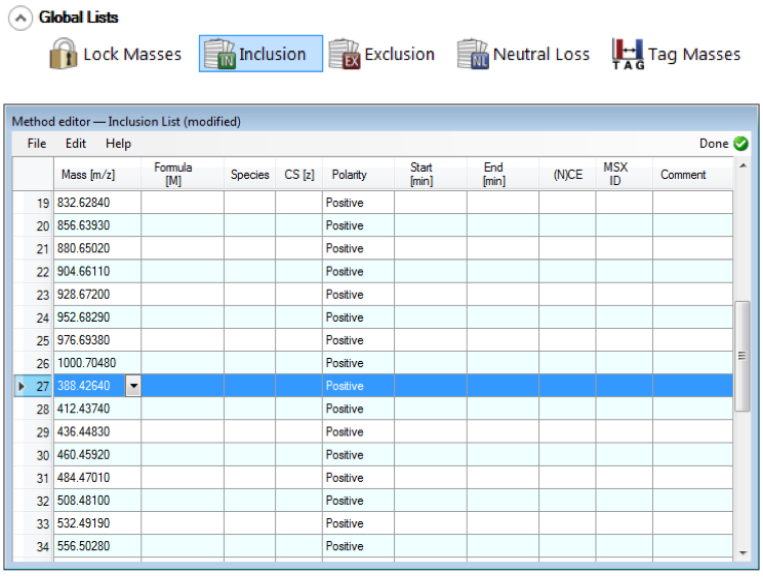
While precursor scans are not required for detection or quantitation, we find they can somewhat improve detection rates and are useful for troubleshooting experiments while acquiring data. We recommend setting the resolution and maximum ion injection time to a lower value than for a typical DDA experiment to limit the time wasted on collecting MS scans. Also, we recommend

only scanning through the precursor isolation range of the experiment. Here for a 400-1000 m/z experiment we use a precursor scan of 395-1005 m/z. Since these windows are not being used for quant, collecting MS survey scans in centroid mode helps to moderate file sizes.

For DIA scans, we set the default charge to 3 and a complementary collision energy setting (NCE) to what we typically use for DDA. In general we set automatic gain control (AGC) target intensity to a very high value so that every MS/MS scan is acquired to the maximum ion inject time (max IT) duration, ensuring regular scans time differences between cycles throughout the experiment. Again, collecting scans in centroid mode helps keep file sizes down.



Finally, we add the “center m/z” inclusion list from the Window Scheme Wizard:



	Mass [m/z]	Formula [M]	Species	CS [z]	Polarity	Start [min]	End [min]	(N)CE	MSX ID	Comment
19	832.62840				Positive					
20	856.63930				Positive					
21	880.65020				Positive					
22	904.66110				Positive					
23	928.67200				Positive					
24	952.68290				Positive					
25	976.69380				Positive					
26	1000.70480				Positive					
27	388.42640				Positive					
28	412.43740				Positive					
29	436.44830				Positive					
30	460.45920				Positive					
31	484.47010				Positive					
32	508.48100				Positive					
33	532.49190				Positive					
34	556.50280				Positive					

Collecting chromatogram libraries

Chromatogram libraries[2] are generated from a small collection of narrow precursor isolation window DIA experiments, rather than from DDA, as with typical spectrum libraries. To generate chromatogram libraries, we typically acquire between 4 and 6 gas-phase fractionated (GPF) injections, in a 4-6x 100 m/z configuration. These fractions are tiled to cover the entire precursor isolation window in our wide-window DIA experiments, for example 400-500 m/z, 500-600 m/z, 600-700 m/z, 700-800 m/z, 800-900 m/z, and 900-1000 m/z. Since the windows do not generally share peptides, retention time consistency between runs is of utmost importance. Consequently, we recommend collecting several conditioning runs of the same sample type on your column before collecting your chromatogram library. A great way to test chromatogram libraries on your instrument is to use the following acquisition order:

1. Typical DDA

2. Single-injection DIA (400-1000m/z)

3. Typical DDA

4. Single-injection DIA (400-1000m/z)

5. Typical DDA

6. Single-injection DIA (400-1000m/z)

7. GPF-DIA (400-500m/z)

8. GPF-DIA (500-600m/z)

9. GPF-DIA (600-700m/z)

10. GPF-DIA (700-800 m/z)

11. GPF-DIA (800-900 m/z)

12. GPF-DIA (900-1000 m/z)

This acquisition strategy allows you to collect triplicate DIA and DDA measurements of your proteome to compare for quantitative reproducibility, as well as 6x GPF-DIA experiments to build a chromatogram library. These GPF-DIA injections can be searched with library-free methods or with predicted spectrum libraries.[5]

Recommended starting settings

The following are starting conditions assuming a 25 second base-to-base peak width for peptides.[1] We recommend adjusting these settings following the rules above as you develop DIA methods specific to your instrument and experiment.

Recommended single-injection DIA (quantitative) acquisitions:

		QE/QE+	QE-HF	Fusion	QE-HFX/ Exploris	Lumos/ Eclipse
	Est. Scan Rate	10 Hz	10 Hz	10 Hz	20 Hz	20 Hz
MS1 Scans	**Range**	385-1015 m/z	385-1015 m/z	385-1015 m/z	390-1010 m/z	390-1010 m/z
	Resolution	35000	60000	60000 (Orbi)	60000	60000 (Orbi)
	Max IIT (ms)	60	60	60	60	60
	AGC Target	1e6	1e6	4e5	1e6	4e5
	Data Type	Centroid	Centroid	Centroid	Centroid	Centroid
MS2 Scans	**Windowing Scheme**	24 m/z staggered	24 m/z staggered	24 m/z staggered	12 m/z staggered	12 m/z staggered
	Resolution	17500	30000	30000 (Orbi)	15000	15000 (Orbi)
	Max IIT (ms)	60	60	60	20	20
	AGC Target	1e6	1e6	4e5	1e6	4e5
	Loop Count	25	25	25	50	50
	Default Charge	3	3	3	3	3
	(N)CE	27	27	33 (HCD)	27	33 (HCD)
	Data Type	Centroid	Centroid	Centroid	Centroid	Centroid

Recommended 6x injection GPF-DIA (chromatogram library) acquisitions:

		QE/QE+	QE-HF	Fusion	QE-HFX/ Exploris	Lumos/ Eclipse
	Est. Scan Rate	10 Hz	10 Hz	10 Hz	10 Hz	10 Hz
MS1 Scans	**Range**	395-505 m/z, 495-605 m/z, etc…	395-505 m/z, 495-605 m/z, etc…	395-505 m/z, 495-605 m/z, etc…	395-505 m/z, 495-605 m/z, etc…	395-505 m/z, 495-605 m/z, etc…
	Resolution	35000	60000	60000 (Orbi)	60000	60000 (Orbi)
	Max IIT (ms)	60	60	60	60	60
	AGC Target	1e6	1e6	4e5	1e6	4e5
	Data Type	Centroid	Centroid	Centroid	Centroid	Centroid
MS2 Scans	**Windowing Scheme**	6x GPF-DIA w/ 4 m/z staggered	6x GPF-DIA w/ 4 m/z staggered	6x GPF-DIA w/ 4 m/z staggered	6x GPF-DIA w/ 4 m/z staggered	6x GPF-DIA w/ 4 m/z staggered
	Resolution	17500	30000	30000 (Orbi)	30000	30000 (Orbi)
	Max IIT (ms)	60	60	60	60	60
	AGC Target	1e6	1e6	4e5	1e6	4e5
	Loop Count	25	25	25	25	25
	Default Charge	3	3	3	3	3
	(N)CE	27	27	33 (HCD)	27	33 (HCD)
	Data Type	Centroid	Centroid	Centroid	Centroid	Centroid

References:

1. Pino, L. K., Just, S. C., MacCoss, M. J. & Searle, B. C. Acquiring and Analyzing Data Independent Acquisition Proteomics Experiments without Spectrum Libraries. *Mol. Cell. Proteomics* (2020) doi:10.1074/mcp.P119.001913.
2. Searle, B. C. *et al.* Chromatogram libraries improve peptide detection and quantification by data independent acquisition mass spectrometry. *Nat. Commun.* **9**, 5128 (2018).
3. Chambers, M. C. *et al.* A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).
4. Amodei, D. *et al.* Improving Precursor Selectivity in Data-Independent Acquisition Using Overlapping Windows. *J. Am. Soc. Mass Spectrom.* **30**, 669–684 (2019).
5. Searle, B. C. *et al.* Generating high quality libraries for DIA MS with empirically corrected peptide predictions. *Nat. Commun.* **11**, 1548 (2020).