1 General Information .................................................................................................................. 8
  1.1 Scope of Spectronaut™ Pulsar X Software ................................................................. 8
  1.2 Spectronaut™ Pulsar X Release Features ................................................................. 8
  1.3 Computer System Requirements ............................................................................... 9
  1.4 Post installation recommendations ......................................................................... 10
  1.5 Supported Mass Spectrometers ............................................................................. 11
  1.6 Supported Data Acquisition Methods .................................................................... 11
  1.7 iRT Kit .................................................................................................................. 12
2 Getting Started .................................................................................................................. 13
  2.1 Getting Spectronaut™ Pulsar X .............................................................................. 13
    2.1.1 Spectronaut Pulsar X Activation .................................................................... 13
  2.2 Demo Data ............................................................................................................. 13
3 Spectronaut™ Pulsar X Usage .......................................................................................... 14
  3.1 Structure of Spectronaut™ Pulsar X ..................................................................... 14
    3.1.1 Layout ........................................................................................................ 14
    3.1.2 Tips for a Better Experience ....................................................................... 15
  3.2 Before starting ......................................................................................................... 16
  3.3 Library Perspective ................................................................................................... 16
    3.3.1 Library Generation from Pulsar .................................................................... 17
    3.3.2 Library Generation from External Search Engines ........................................ 22
      3.3.2.1 Spectral Library Generation from BGS Generic Format ...................... 24
    3.3.3 Importing an External Library ..................................................................... 26
      3.3.3.1 Library Columns ................................................................................ 28
      3.3.3.2 Modification Parsing ......................................................................... 31
    3.3.4 Merging Libraries ......................................................................................... 32
    3.3.5 Library Overview .......................................................................................... 33
    3.3.6 Making a Labeled or a Spike-in Library ....................................................... 34
    3.3.7 Generate a Quality Control Kit from a Library .......................................... 35
3.4 Analysis Perspective........................................................................................................35

3.4.1 Setting up and Running an Analysis ........................................................................35

3.4.1.1 Performing a DIA Library-based Analysis .........................................................36

3.4.1.2 Performing a directDIA™ Analysis ..................................................................39

3.4.1.3 Configure Conditions .......................................................................................40

3.4.1.4 Quantitative workflows supported in Spectronaut Pulsar X .........................42

3.4.2 Reviewing your analysis .......................................................................................43

3.4.2.1 Analysis Perspective Plots ..............................................................................44

3.4.2.2 Tree Filtering .................................................................................................45

3.4.2.3 Experiment Tab Options ................................................................................46

3.4.2.4 Manual Analysis Refinement ..........................................................................46

3.4.2.5 Library Refinement ........................................................................................50

3.5 Post Analysis Perspective .....................................................................................51

3.5.1 Analysis Overview ............................................................................................52

3.5.2 Scoring Histograms ..........................................................................................53

3.5.3 Analysis Details ................................................................................................53

3.5.4 Differential Abundance ....................................................................................53

3.5.4.1 Candidates table ...........................................................................................53

3.5.4.2 GO Enrichment .............................................................................................54

3.5.4.3 GO Clustering ................................................................................................56

3.5.4.4 Differential Abundance Plots ........................................................................57

3.6 Report Perspective ..................................................................................................58

3.6.1 Report Schemas ................................................................................................59

3.6.2 Normal Report ...................................................................................................60

3.6.3 Run Pivot Report ................................................................................................60

3.7 Quality Control Perspective ...................................................................................61

3.7.1 QC Panels ..........................................................................................................62

3.8 Pipeline Perspective ................................................................................................63
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.8.1</td>
<td>Pipeline Reporting Settings</td>
<td>64</td>
</tr>
<tr>
<td>3.9</td>
<td>Databases Perspective</td>
<td>64</td>
</tr>
<tr>
<td>3.9.1</td>
<td>Table Import</td>
<td>65</td>
</tr>
<tr>
<td>3.9.2</td>
<td>Modifications</td>
<td>65</td>
</tr>
<tr>
<td>3.9.2.1</td>
<td>Importing Modifications from Search Engine</td>
<td>65</td>
</tr>
<tr>
<td>3.9.2.2</td>
<td>Creating custom modifications</td>
<td>66</td>
</tr>
<tr>
<td>3.9.3</td>
<td>Protein Databases</td>
<td>66</td>
</tr>
<tr>
<td>3.9.4</td>
<td>GO Databases</td>
<td>67</td>
</tr>
<tr>
<td>3.9.4.1</td>
<td>Gene Ontologies</td>
<td>68</td>
</tr>
<tr>
<td>3.9.4.2</td>
<td>Gene Annotations</td>
<td>68</td>
</tr>
<tr>
<td>3.9.5</td>
<td>Digest Rules</td>
<td>68</td>
</tr>
<tr>
<td>3.10</td>
<td>Settings Perspective</td>
<td>70</td>
</tr>
<tr>
<td>3.10.1</td>
<td>DIA Analysis Settings</td>
<td>71</td>
</tr>
<tr>
<td>3.10.2</td>
<td>Pulsar Search Settings</td>
<td>71</td>
</tr>
<tr>
<td>3.10.3</td>
<td>directDIA™ Settings</td>
<td>71</td>
</tr>
<tr>
<td>3.10.4</td>
<td>Library Generation Settings</td>
<td>71</td>
</tr>
<tr>
<td>3.10.5</td>
<td>Global Settings</td>
<td>71</td>
</tr>
<tr>
<td>3.10.5.1</td>
<td>General</td>
<td>72</td>
</tr>
<tr>
<td>3.10.5.2</td>
<td>Directories</td>
<td>72</td>
</tr>
<tr>
<td>3.10.5.3</td>
<td>Plotting</td>
<td>73</td>
</tr>
<tr>
<td>3.10.5.4</td>
<td>Reporting</td>
<td>73</td>
</tr>
<tr>
<td>3.10.6</td>
<td>Spectronaut Pulsar X Command Line Mode</td>
<td>73</td>
</tr>
<tr>
<td>4</td>
<td>HTRMS Converter</td>
<td>75</td>
</tr>
<tr>
<td>4.1</td>
<td>File Conversion</td>
<td>76</td>
</tr>
<tr>
<td>4.2</td>
<td>Folder Conversion</td>
<td>76</td>
</tr>
<tr>
<td>5</td>
<td>References</td>
<td>77</td>
</tr>
<tr>
<td>6</td>
<td>Appendixes</td>
<td>79</td>
</tr>
<tr>
<td>6.1</td>
<td>Appendix 1. DIA Analysis Settings</td>
<td>79</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>6.2</td>
<td>Appendix 2. Pulsar Search Settings</td>
<td>87</td>
</tr>
<tr>
<td>6.3</td>
<td>Appendix 3. directDIA™ Settings</td>
<td>88</td>
</tr>
<tr>
<td>6.4</td>
<td>Appendix 4. Library Generation Settings</td>
<td>93</td>
</tr>
<tr>
<td>6.5</td>
<td>Appendix 5. Analysis Perspective Plots</td>
<td>96</td>
</tr>
<tr>
<td>6.5.1</td>
<td>Run Node Plots</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>iRT Calibration Chart</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>XIC Extraction Width Chart</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>TIC Chromatogram &amp; TIC Overlay Chromatogram</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Analysis Log</td>
<td>98</td>
</tr>
<tr>
<td>6.5.2</td>
<td>Peptide and Fragment Plots</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>MS2 XIC</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>MS2 XIC Sum</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>MS2 Intensity Correlation</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>MS1 Isotope Envelope XIC</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>MS2 Isotope Envelope Correlation</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>MS2 XIC Alignment and MS1 XIC Alignment</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>XIC graph</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>iRT XIC Sum Overlay</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>MS2 Intensity Alignment</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Cross Run RT Accuracy</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>MS1 Spectrum at Apex</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>MS2 Spectrum at Apex</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>Peptide Data Match (PDM) Plot</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>Protein Coverage</td>
<td>111</td>
</tr>
<tr>
<td>6.6</td>
<td>Appendix 6. Experiment Tab Options</td>
<td>112</td>
</tr>
<tr>
<td>6.7</td>
<td>Appendix 7. Post Analysis Perspective Plots</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>Data Completeness</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>Coefficients of Variations</td>
<td>116</td>
</tr>
</tbody>
</table>
CVs Below X ................................................................. 117
Normalization .................................................................. 118
Heatmap ......................................................................... 119
Volcano Plot .................................................................... 120
Sample Correlation Plot .................................................... 121

6.8 Appendix 8. Most Relevant Report Headers ...................... 122
1 General Information

1.1 Scope of Spectronaut™ Pulsar X Software

Spectronaut™ Pulsar X is a commercial software package aimed at analyzing data independent acquisition (DIA) experiments. Spectronaut Pulsar X can quantitatively profile 100s to several 1000s of proteins in one experiment. Large experiments with several conditions and replicates consisting of up to 1000s of LC-MS runs can be analyzed.

Spectronaut Pulsar X can analyze DIA data without the use of a retention time calibration kit. However, the addition of the iRT Kit is highly recommended as it ensures calibration on difficult matrices and allows for detailed quality control read-outs.

1.2 Spectronaut™ Pulsar X Release Features

- Support for Hybrid Library workflows
- Introduction of Search Archives for (Box 2)
- Introduction of Source specific iRT calibration (Box 3)
- Improved library generation from DIA data
- Improved directDIA™ workflow: 50% faster; up to 10% more identifications
- Improved Protein-FDR estimation
- Introduction of peptide level FDR in library generation from Pulsar
- Introduction of GO Clustering in Post Analysis
- Improved support for Waters
- Improved Analysis Perspective
- Improved Post Analysis Perspective
- Up to 10% faster saving and loading experiment
- Improved Spike-in workflow
- Improved support for Host Cell Proteome (HCP) workflow
1.3 Computer System Requirements

Spectronaut™ Pulsar X is only available for Windows operating systems. Command line operation is also supported (see section 3.10.6). The minimum and recommended system specifications are:

**Minimum**

- Operating System: Windows 7, x64
- CPU: Intel® Core™ CPU, 2.7 GHz (quad-core) or similar
- Hard drive: 200 GB free space
- Memory: 16 GB
- Software: .NET 4.7 or higher

**Recommended**

- Operating System: Windows 10 or higher, x64
- CPU: Intel Core i7 4770, 3.4 GHz (octa-core) or similar
- Hard drive: 500 GB free space, solid state drive (SSD)
- Memory: 64 GB or more
- Software: .NET 4.7 or higher

The recommended amount of total system memory can be estimated using the following equation:

\[
RAM_{GB} = 4 + \frac{0.85 \times (n + 5000) \times r}{1024^2}
\]

where \( n \) is the number of precursors in the library and \( r \) the number of runs in the experiment. According to this, 32 GB of RAM should be sufficient to process 500 runs with a library containing 50000 precursors.
1.4 Post installation recommendations

Performance improvements after the installation:

1. **Directories:** Spectronaut™ Pulsar X will set all directories in the C: drive by default. However, it is likely that the C: drive has a limited storage capacity. Thus, we **strongly recommend** changing the Temporary Directory and the Local Search Archives directory to a local destination with enough free memory. To do that, go to the Settings Perspective → Global → Directories (Figure 1).

![Figure 1](image)

Figure 1. Change the default location for temporary files and Search Archives.

2. **Network drives and virtual machines:** We strongly recommend running Spectronaut Pulsar X local, i.e., having the resources (especially run files and temporary directories) on a local drive. The use of virtual machines for Spectronaut Pulsar X is feasible but not advised. A failure in the connection to any network drive can cause Spectronaut Pulsar X to abort the process due to third-party library dependencies.

3. **CPUs:** Spectronaut Pulsar X is designed to perform highly resource intensive tasks, especially when running searches with Pulsar. For resource management purposes, you can set a maximum number of CPUs Spectronaut Pulsar X may use. To do that, go to the Settings Perspective → Global → General → CPU Affinity, and uncheck one or more of the CPUs (Figure 2). Be aware this will prolong the analysis time.
1.5 Supported Mass Spectrometers

Spectronaut™ Pulsar X supports mass spectrometers from Thermo Scientific™, Sciex, Bruker and Waters. The specific supported models are:

- Thermo Scientific™ Q Exactive™ Series
- Thermo Scientific™ Orbitrap Fusion™ Series
- Sciex TripleTOF® Series
- Bruker impact II™
- Waters Xevo® G2-XS QToF

1.6 Supported Data Acquisition Methods

Spectronaut™ Pulsar X analyzes a large variety of DIA data. Minimum requirements are a reversed phase chromatography with either a linear or nonlinear gradient that spans at least 10-35% acetonitrile. Methods acquiring only MS2 scans are supported, as well as methods with both, MS1 and MS2 scans. The cycle time of the DIA method should be in the range of 2-3 seconds depending on the peak width of the chromatography used. MS1 as well as MS2 precursor ranges can be segmented. The MS2 scans should cover at least 500-900 m/z of precursor range. Gas phase fractionation is supported. More specifically Spectronaut Pulsar X supports HRM™, WiSIM-DIA, AIF, SWATH™ (Gillet et
al. 2012) and SWATH™ 2.0. Fractionated DIA experiments are not recommended (Box 6). Multiplexed DIA is not supported.

In case you experience technical problems with the software or if you have feature suggestions please contact support@biognosys.com.

1.7 iRT Kit

To enable fully automated and sensitive signal processing with Spectronaut™ Pulsar X, we developed the iRT Kit, that can be spiked into each sample before measurement. The iRT Kit contains a mixture of synthetic peptides which are non-naturally occurring. Based on these peptides, Spectronaut Pulsar X calibrates important parameters for peak detection (iRT, mass) and can store quality control information available in the QC perspective (see section 3.7). The integration of Spectronaut Pulsar X with the iRT Kit generates better results:

- **Smart in-run parameter calibration**: guarantees quick and robust analysis
- **Increases sensitivity and specificity**: by accurate iRT determination (Escher et al. 2012; Bruderer et al. 2016) and advanced handling of interfering signals
- **Higher quantitative accuracy**: by supporting optimized ion current extraction
- **Automated quality control**: by monitoring LC and MS performance

*The use of the iRT Kit is highly recommended when performing any analysis with Spectronaut Pulsar X.*

For more information about the iRT Kit, please see [here](#).
2 Getting Started

2.1 Getting Spectronaut™ Pulsar X

Spectronaut™ Pulsar X software licenses can be purchased on our webpage. We also provide licenses for a trial period upon request on our webpage. After requesting a license, you will get an email with:

1. A link to download the installer
2. A key to activate your copy.

If this is not the case, please contact us at support@biognosys.com.

**Important:** license keys are computer-bound. If you need to install Spectronaut Pulsar X on more than one computer, please contact us at support@biognosys.com.

2.1.1 Spectronaut Pulsar X Activation

When you install and start Spectronaut Pulsar X for the first time, you will be asked to activate your software by pasting your license key into the Spectronaut Pulsar X activation dialogue. If your computer has access to the internet, activation will be automatic. If your Spectronaut Pulsar X computer does not have an internet connection, or the connection is blocked by a firewall, you can also activate your software offline.

The respective instructions will appear after a few seconds if online activation was not successful. Save the registration information file on your computer and send this file to support@biognosys.com. In general, you will receive a license file within one or two working days. To activate Spectronaut Pulsar X using a license file, click on the "Browse License File..." button in the Spectronaut Pulsar X Activation dialogue. **Your license period will start upon activation.**

2.2 Demo Data

In section 3, Spectronaut™ Pulsar X Usage, we will guide you through the software perspective by perspective. The examples shown for the classic DIA analysis (section 3.4.1.1) are generated with the demo data available for downloading here. Please note this demo data was intentionally prepared to be as small as possible for demoing purposes. Most DIA experiments will require larger storage space and more resources to be analyzed.
3 Spectronaut™ Pulsar X Usage

3.1 Structure of Spectronaut™ Pulsar X

3.1.1 Layout

Spectronaut™ Pulsar X is structured in different levels (Figure 3). The highest level is the Perspective. Within each perspective, you can often find several pages separated into tabs. The layout of each page is normally structured into a left menu (tree) containing elements (nodes) and a right panel containing information related to the selected nodes (plots, reports and summaries).

Figure 3. Spectronaut Pulsar X general layout structure.
3.1.2  Tips for a Better Experience

1. Spectronaut Pulsar X is full of informative tool-tips throughout the software (Figure 4). They will appear as you hover over many of the elements.

2. There are many nonobvious functionalities by right-clicking an element: experiment tab, plots, nodes, etc.

![Figure 4. Spectronaut Pulsar X contains nonobvious tip and menus when you hover over some elements or right-click on them.](image)

3. It is possible the software will display warning messages like this:

![Warning Message](image)

Warnings are sometimes just informational, and do not require action. Error messages during library generation or data processing are shown in red. If an error occurs, please send the error log to support@biognosys.com. The error logs can be found by clicking "Show Error Logs" in the About Perspective.
3.2 Before starting

Make sure you have everything you need ready before starting your analysis in Spectronaut™ Pulsar X. There are two main types of quantitative analyses you can run in Spectronaut Pulsar X: classic, library-based DIA analysis or library-free directDIA™ analysis. Table 1 shows which resources are required to perform each of the pipelines.

Table 1. Input resources for each DIA approach supported in Spectronaut Pulsar X.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Classic DIA analysis (library-based)</th>
<th>directDIA (library-free)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIA run files</td>
<td>required</td>
<td>required a</td>
</tr>
<tr>
<td>Library b</td>
<td>required</td>
<td>not applicable</td>
</tr>
<tr>
<td>Protein database (FASTA)</td>
<td>recommended</td>
<td>required</td>
</tr>
<tr>
<td>Gene Ontology annotation</td>
<td>optional</td>
<td>optional</td>
</tr>
</tbody>
</table>

a directDIA is currently supported only for Thermo Scientific™ and Sciex run files
b Libraries can be generated in the Library Perspective of Spectronaut Pulsar X (section 3.3)

3.3 Library Perspective

The main tasks you can perform in the Library Perspective of Spectronaut™ Pulsar X are:

1. Generating a library with Pulsar, Biognosys’ proprietary search engine (section 3.3.1).
2. Generating a library using search results from external search engines.
3. Importing an external library

Guidelines on how to generate the data for an optimal library can be found in Box 1.
3.3.1 Library Generation from Pulsar

Pulsar is Biognosys’ proprietary search engine, integrated into Spectronaut Pulsar X for library generation. Pulsar can search data-dependent acquisition (DDA), data-independent acquisition (DIA) and parallel reaction monitoring (PRM) data. Both centroid and profile mode data can be processed. Pulsar is designed to be fast and scale as the number of runs to be analyzed increases. Further, Pulsar is able to identify co-fragmented peptides in multiple search rounds by subtracting previously identified fragment ions from the spectra. False identifications are controlled by a false discovery rate (FDR) estimation at three levels: peptide-spectrum match (PSM), peptide, and protein group.

The specific vendors and acquisition modes supported by Pulsar are:

- Thermo Scientific™ (DDA, DIA and PRM)
- Sciex (DDA and DIA/SWATH™)
- Bruker (DDA)
- Waters (DDA)

Figure 5. Library Generation with Pulsar. In the Library Perspective, Spectral Library tab click "Generate Library from Pulsar…". Follow the wizard to complete the process.
Every time a search is performed, Spectronaut Pulsar X will save the results (PSMs) for each run as a Search Archive (Box 2). These Search Archives can be used to generate libraries without the need to search the runs from scratch. Run files and Search Archives can be combined conveniently to generate new libraries. Refer to Table 2 to see a summary of the resources you will need in each of these cases.

To generate a library from Pulsar, go to the Library Perspective and click on "Generate Library from Pulsar..." in the bottom left corner. A wizard will appear to help you set up the experiment (Figure 5). A schematic view showing the wizard steps you will encounter, depending on your input resources, is shown in Table 3. The sequential steps are described below:

1. **Set up Library Generation from Pulsar.**
   - Choose an experiment name
   - Click on "Add Runs from File..." or "Add Runs from Folder..." and select the runs from which you want to create the library. You can mix runs acquired in different modes. *If you want to generate a library from Search Archives only, skip this step.*

<table>
<thead>
<tr>
<th>Library based on</th>
<th>FASTA file</th>
<th>Gene Ontology information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run file(s) only</td>
<td>required</td>
<td>optional</td>
</tr>
<tr>
<td>Search Archive(s) only</td>
<td>not applicable</td>
<td>optional</td>
</tr>
<tr>
<td>Run file(s) and Search Archive(s) combined</td>
<td>required</td>
<td>optional</td>
</tr>
</tbody>
</table>

2. **Choose Fasta File(s)** by clicking "Fasta File...". Protein databases can be assigned on a run basis. Multiple protein databases can be selected. FASTA files can be added at this step by clicking "Import..." in the bottom left corner. *If you are doing a library from Search Archives only, skip this step.*
Box 1. Library Guidelines

To generate a spectral library, typically shotgun runs of your samples of interest are acquired, searched with a database search engine and the search results are then condensed into a spectral library. To maximize the coverage, we recommend measuring pools of samples that have been biochemically fractionated (e.g. using high pH reverse phase fractionation). Technical LC-MS/MS replicates are still recommended due to the semi-stochastic nature of shotgun proteomics. The optimal number of pools, fractions and replicates depends on the experimental setup and the complexity of the samples. However, overly large spectral libraries where only a small percentage can be recovered from the data might negatively influence the sensitivity of your analysis.

We strongly recommend generating the library in Spectronaut™ Pulsar X. Although for most common samples the iRT Kit is not strictly required, we do recommend spiking the iRT Kit into samples aimed at library generation. Spectronaut™ Pulsar X will take care of calculating iRTs for all peptides identified even if the iRT Kit was not used.

Alternatively, spectral libraries can be imported. If an external library is used, please consider the following recommendations to ensure the best performance:

- Keep peptide precursors non-redundant in the library
- Do not keep more than 15 fragment ions per peptide precursor
- Precursors with less than 3 fragment ions should not be considered
- Spike the iRT Kit into your sample for shotgun analysis
- Search your shotgun files with the iRT peptide FASTA file together with your FASTA file. The iRT peptide FASTA fusion file can be downloaded here.
- Check whether the iRT peptides were identified in your search. If they are not this could be an indication of a low-quality measurement. The iRT peptides are relatively high in concentration and most, if not all, should be easily detected.
- Generate a spectral library using SpectraST (Lam et al., 2007) or similar.
- iRTs can be calculated from retention times (RT) in the database search using a linear regression on iRT and RT values of the iRT peptides \( iRT_{\text{pep}} = f(\text{RT}_{\text{pep}}) \). This linear transformation can then be applied to all peptides identified in the database search. The calculation can be performed in Excel, R or similar.

3. **Choose Pulsar Search Settings** by clicking "Search settings..." (for detailed explanations about each setting, see Appendix 2. Pulsar Search Settings (section 6.2)).

4. Choose either the default schema which can be modified on the fly or a previously saved settings-schema. Schemas can be assigned at either the experiment or run...
level. Only one settings-schema can be set per run. When nothing is selected for a run, default settings will be applied. *If you are doing a library from Search Archives only, skip this step.*

5. Next, you can add Search Archives (for more information see Box 2) to your library. Search Archives prevent you from having to research run files if you have already searched them in the past.

6. In Spectronaut Pulsar X you can generate your library containing Gene Ontology (GO) annotation information. To select a GO annotation (*.goa) file at this point, you need to have the file already imported in the Databases Perspective. Learn how to do so in section 3.9.4.2.

---

**Box 2. Search Archives**

Search Archive is a new concept introduced with Spectronaut™ Pulsar X.

Before Search Archives, already searched run files had to be searched again from scratch to include them in a library with other runs and maintain control of the FDR. This resulted in a great amount of time and computational resources which had to be reinvested.

With Search Archives, every time a library is generated using Pulsar, the result relevant to this Pulsar search is saved as a Search Archive, and will appear in the Search Archive page of the Library Perspective. Search Archives contain the information from a search before applying any FDR filter. This allows several Search Archives to be combined together, or Search Archives with runs files, to generate libraries with a proper, library-wide control of the FDR.

There are two types of Search Archives:

1. Search Archives from run files (complete Search Archives). These archives contain all the information related to the Pulsar search of each particular run file.
2. Search Archives from other archives or from combinations of archives (meta-Search Archives). These archives contain the metainformation of:
   - A re-utilized complete archive
   - A combination of archives

When a library is generated combining archives and run files, a complete Search Archive will be stored for each run file and a meta-Search Archive will be done containing the information of the combined analysis.

When meta-Search Archives are used, the process will be directed to the corresponding complete Search Archives to retrieve the relevant information.
7. The next wizard page contains experiment-wide settings for library generation, such as PSM, peptide and protein FDR thresholds (for a detailed explanation about each setting, see Appendix 4. Library Generation Settings (section 6.4)).

8. The last page shows an overview of the whole experiment set-up. Clicking "Finish" will start the experiment. Using the "View Live Log…" it is possible to follow the progress of the experiment. As soon as the library is generated, it will appear in the library tree. Libraries with FASTA files assigned are marked with a blue protein icon.

Table 3. Schematic view of the wizard steps required during library generation depending on the input resources

<table>
<thead>
<tr>
<th>Library generation step</th>
<th>Run files only</th>
<th>Run files and Search Archives combined</th>
<th>Search Archives only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose an experiment name</td>
<td>required</td>
<td>required</td>
<td>required</td>
</tr>
<tr>
<td>Add runs</td>
<td>required</td>
<td>required</td>
<td>not applicable</td>
</tr>
<tr>
<td>Choose Fasta Files</td>
<td>required</td>
<td>required</td>
<td>not applicable</td>
</tr>
<tr>
<td>Choose Search Settings</td>
<td>required (with default)</td>
<td>required (with default)</td>
<td>not applicable</td>
</tr>
<tr>
<td>Choose Search Archives</td>
<td>not applicable</td>
<td>required</td>
<td>required</td>
</tr>
<tr>
<td>Specify Gene Ontology annotation</td>
<td>optional</td>
<td>optional</td>
<td>optional</td>
</tr>
<tr>
<td>Choose library settings</td>
<td>required (with default)</td>
<td>required (with default)</td>
<td>required (with default)</td>
</tr>
</tbody>
</table>
3.3.2 Library Generation from External Search Engines

To generate a library from external search engine results, you will need:

1. The search result files or folders → required
2. The run files from which the search was done → required

Spectronaut Pulsar X supports search results from:

- MaxQuant (Cox et al. 2011)
- Proteome Discoverer™
- ProteinPilot™
- Mascot™

Table 4 summarizes the type of files or folders needed for each search engine, and whether some actions are required for correct integration of the post-translational modification (PTM) annotations (read note b in Table 4).

In addition to the specific result formats above, Spectronaut Pulsar X also supports results in mzIdentML format (containing fragment ion information). Finally, any search results can be reformatted into the Biognosys (BGS) Generic Format.

To generate a library from search results:

1. Go to the Library Perspective → Spectral Library and click on "Generate Spectral Library from..." in the bottom left corner (Figure 6). Choose your search engine.

2. Navigate to the files or folders containing the search results (see Table 4). Spectronaut Pulsar X will try to map the run files automatically (see Box 4). If it fails to do so, you will have to manually link the files by clicking "Assign Shotgun Files..."

3. Choose your library settings in the Library Settings panel or run under default settings. (for a detailed explanation of each setting, see Appendix 4. Library Generation Settings (section 6.4).

4. Choose a FASTA file in the FASTA File tab. If your FASTA file is not yet in the tree, you can add it at this point by clicking "Import..." in the bottom left corner.
5. Select your Gene Ontology annotation information in the Gene Annotation panel. You should have your file previously loaded into the Databases Perspective (to learn how to do this, go to section 3.9.4). By clicking “Load”, Spectronaut Pulsar X will perform the library generation. Your new library will automatically appear in the Library Perspective upon completion.

**Box 3. Source-specific iRT calibration**

Source-specific iRT calibration is a new feature introduced with Spectronaut™ Pulsar X.

In the past, the iRT value of a given peptide in a spectral library was summarized by taking a median across all runs where the peptide was identified. However, when building a spectral library from chromatographically heterogenous data, this can lead to a loss in iRT-precision. A **good example** of this case would be a situation in which you generate a hybrid library using DDA files from public repositories together with your quantitative DIA files.

To improve the targeted extraction of such data, we have introduced the concept of source-specific iRT calibration. Spectronaut Pulsar X will generate libraries containing as many iRT values as there are different sources in the dataset. When using this library on a quantitative dataset, Spectronaut Pulsar X will use the iRT calibration from the best source available for each assay.

By using source-specific iRT calibration, you will be able to keep the iRT-precision of project specific data while benefiting from the depth of a large resource dataset.

**Source-specific iRT calibration and Search Archives**

Spectronaut Pulsar X will create different iRT sources in the libraries in the following cases:

1. Different acquisition methods: DDA and DIA
2. Different calibration type: linear, fragmented (non-linear), *in silico* (no iRT peptides present)
3. Different library inputs: Search Archive(s) or run files. One source will be created per Search Archive. You can create archives with iRT heterogeneity in mind so that you can conveniently reuse them in the future.
Table 4. Supported search engines and information required by Spectronaut Pulsar X when generating a library from search results

<table>
<thead>
<tr>
<th>Search engine</th>
<th>Search result</th>
<th>Peptide modifications</th>
<th>Custom</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MaxQuant</strong></td>
<td>Folder with result files</td>
<td>Included</td>
<td>Imported (*.xml file)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>from MaxQuant installation folder (\bin\conf\modifications.xml)</td>
</tr>
<tr>
<td><strong>Proteome Discoverer</strong></td>
<td>*.msf for PD 1.4</td>
<td></td>
<td>Included with the search results</td>
</tr>
<tr>
<td></td>
<td>*.pdResult for PD &gt; 2.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Protein Pilot</strong></td>
<td>MS Excel with the suffix &quot;_FDR&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mascot</strong></td>
<td>*.dat file</td>
<td>Download the latest Unimod XML database themselves from <a href="http://www.unimod.org/downloads.html">www.unimod.org/downloads.html</a></td>
<td>Add manually (see 3.9.2.2)</td>
</tr>
</tbody>
</table>

*Detailed information on how to setup a MaxQuant search for spectral library generation can be found in the Biognosys Science Hub at www.biognosys.com/science-hub.*

*These defaults apply only to upgrades from older versions of Spectronaut. If Spectronaut Pulsar X is your first Spectronaut version, no action is required concerning default modifications.*

### 3.3.2.1 Spectral Library Generation from BGS Generic Format

Spectronaut Pulsar X supports generating spectral libraries from the minimalistic BGS Generic Format. This allows end-users to use their favorite search engine with the aid of a basic script which would convert their search result into BGS Generic format. This is a plain-text format where each row represents a PSM. Table 5 shows the information required in this file.
Box 4. Mapping run files to search results

Spectronaut Pulsar X will try to map the run files automatically by name matching. First, it will look in your Shotgun Raw Repository (Settings Perspective → Global → Directories). If unsuccessful, it will look in the search results location. If the automatic mapping fails, you will see a red cross.

If this is the case, you will have to manually map the runs. Click on “Assign Shotgun Files…” to find the missing runs. You can either navigate to a common directory or browse for your runs individually.

After the runs have been found, the red cross will change into a green tick mark.

Figure 6. Library Generation from external search engines. In the Library Perspective, Spectral Library tab click “Generate Library from…” and choose your search engine. Load your search results and assign your shotgun run files (Box 4). Follow the wizard to complete the process.
Table 5. BGS generic format search results required information

<table>
<thead>
<tr>
<th>Header</th>
<th>Information</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw File</td>
<td>The name of the DDA file in which this PSM was found. This column is used to map the DDA file to the PSM.</td>
<td>required</td>
</tr>
<tr>
<td>Stripped Sequence</td>
<td>The stripped sequence of the peptide that was found by the search engine for this PSM.</td>
<td>required</td>
</tr>
<tr>
<td>Precursor Charge</td>
<td>The charge that was associated with this PSM by the search engine.</td>
<td>required</td>
</tr>
<tr>
<td>Labeled Sequence</td>
<td>The sequence with encoded modification and label information. The only requirement is that the modification/label information should be enclosed in [modification/label] or (modification/label) brackets.</td>
<td>required</td>
</tr>
<tr>
<td>Retention Time</td>
<td>The retention time of the PSM. If available, retention time at apex intensity should be used.</td>
<td>required</td>
</tr>
<tr>
<td>Scan Number</td>
<td>The scan number of the PSM.</td>
<td>required</td>
</tr>
<tr>
<td>Scan Event</td>
<td>The scan event of the PSM. *This is only relevant for Sciex <em>.wiff files.</em></td>
<td>required</td>
</tr>
<tr>
<td>MS1 Intensity</td>
<td>The intensity of the PSM as reported by the search engine.</td>
<td>recommended</td>
</tr>
<tr>
<td>Protein Group Id</td>
<td>The protein group assigned by the search engine for this peptide. It is not necessary if using protein inference in Spectronaut Pulsar X.</td>
<td>recommended</td>
</tr>
</tbody>
</table>

3.3.3 Importing an External Library

To import an external library into Spectronaut Pulsar X, click on "Import Spectral Library…" in the bottom left corner of the Spectral Libraries tab in the Library Perspective (Figure 7).

There are two ways of importing a library into Spectronaut Pulsar X:

1. Importing a *.kit library (Biognosys' library format). In this case, no further action is required, and the library will be loaded automatically into the Library tree.
2. Importing a compatible spreadsheet as a plain text, separated value format (*.txt, *.csv, *.tsv, *.xls). Headers defining your columns are mandatory in these files.

The "Import Spectral Library" dialogue (Figure 7) will try to auto-detect column names. If there are new column names, Spectronaut Pulsar X will ask you whether or not you want to store them as a recognized synonym for this column. This allows Spectronaut Pulsar X to automatically select these columns the next time you load a spectral library with a similar format (you can remove the user-defined column synonyms in the Databases Perspective → Table Import).

Figure 7. Importing an external library. The "Import Spectral Library" dialog only applies to formats different than *.kit. You can refine your library using the lower panel tabs.

The import function also allows you to refine your library. In the Library Settings panel, you can choose several options to be applied to your library (for details, see Appendix 4. Library Generation Settings, in section 6.4). For example, you can perform protein inference again. To do this, go to the FASTA File panel and choose your protein database. You can also add Gene Ontology annotation information using the Gene Annotations panel.
The minimum requirements for a library are two columns: precursor ion \( m/z \) and fragment ion \( m/z \). Although an analysis would work with only these two minimally required columns, it is advised that the information listed in Table 6 is also included to achieve the best possible results.

### 3.3.3.1 Library Columns

A Spectronaut Pulsar X library is similar to a typical MRM/SRM transition list. Refer to Table 6 to know what information a library should contain.

<table>
<thead>
<tr>
<th>Header</th>
<th>Requirement</th>
<th>Refers to</th>
</tr>
</thead>
<tbody>
<tr>
<td>PrecursorMz</td>
<td>Required</td>
<td>The \textit{in silico} calculated ( m/z ) of the peptide precursor ion. Do not round this number.</td>
</tr>
<tr>
<td>FragmentMz</td>
<td>Required</td>
<td>The \textit{in silico} calculated ( m/z ) of the peptide fragment ion. Do not round this number.</td>
</tr>
<tr>
<td>iRT</td>
<td>Highly recommended</td>
<td>The peptide retention time in the reverse phase chromatography converted into iRT space (Escher \textit{et al.} 2012). If accurate iRTs are provided, the analysis will speed up significantly, and the quality of your results will increase (sensitivity, specificity, accuracy). If no iRTs are available, Spectronaut Pulsar X will predict an iRT for each peptide. This is less accurate than an empirically determined iRT value. To derive iRTs for your peptides, we recommend spiking the Biognosys' iRT Kit into your shotgun runs. If the Library Perspective of Spectronaut Pulsar X is used, iRT values will be automatically determined for your library.</td>
</tr>
<tr>
<td>Header</td>
<td>Requirement</td>
<td>Refers to</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>RelativeFragmentIntensity</td>
<td>Highly</td>
<td>The relative peptide fragment ion intensity expressed as a percentage of the most intense fragment ion. Please report the raw number without the percentage sign. This information improves the limit of detection.</td>
</tr>
<tr>
<td>StrippedSequence</td>
<td>Recommended</td>
<td>The stripped amino acid sequence of the peptide excluding any modifications. Please only use the single letter code for the 20 standard proteinogenic amino acids. This information is used for labeling and scoring your fragment ions in Spectronaut Pulsar X. Further, it is included in the automatically generated unique ID for your precursor if necessary.</td>
</tr>
<tr>
<td>PrecursorCharge</td>
<td>Recommended</td>
<td>The peptide precursor ion charge. This information is used to label your precursors in Spectronaut Pulsar X and to automatically generate a unique ID for your precursor if necessary.</td>
</tr>
<tr>
<td>FragmentType</td>
<td>Recommended</td>
<td>The peptide fragment ion type. Usually, this is &quot;y&quot; or &quot;b&quot;. This information is used for labeling and scoring your fragment ions in Spectronaut Pulsar X.</td>
</tr>
<tr>
<td>FragmentNumber</td>
<td>Recommended</td>
<td>The peptide fragment ion number. This number should be between 1 and the length of your peptide in amino acids minus one. This information is used for labeling and scoring your fragment ions in Spectronaut Pulsar X.</td>
</tr>
<tr>
<td>FragmentCharge</td>
<td>Recommended</td>
<td>The peptide fragment ion charge formatted as a number. This information is used for labeling and scoring your fragment ions in Spectronaut Pulsar X.</td>
</tr>
<tr>
<td>Header</td>
<td>Requirement</td>
<td>Refers to</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>FragmentLossType</strong></td>
<td>Recommended</td>
<td>Allows you to specify the fragment loss-type (e.g. NH₃ or H₂O). This is mainly used to label your fragment ions within plots.</td>
</tr>
<tr>
<td><strong>ExcludeFromQuantification</strong></td>
<td><strong>If applicable</strong></td>
<td>This column can be specified optionally and is meant for annotating fragment ions that should only contribute to identification but not to quantification (TRUE or FALSE). TRUE means that Spectronaut Pulsar X will never consider this fragment for quantification and FALSE if Spectronaut Pulsar X will always use it for quantification. If you leave the column empty, the automatic interference correction will decide whether or not to use it (the latter is the default if the column is not specified).</td>
</tr>
<tr>
<td><strong>ModifiedSequence</strong></td>
<td>optional</td>
<td>In the event that your peptide is modified use this column to specify the amino acid sequence including modifications. The modified sequence should be constant for one unique precursor. This information is used to label your precursors in Spectronaut Pulsar X and automatically generate a unique ID if necessary. Spectronaut Pulsar X will try to parse and map modifications from the provided sequences to the internal modification database. This field does not contain any label specific modifications (see LabeledSequence).</td>
</tr>
</tbody>
</table>
**Header** | **Requirement** | **Refers to**
---|---|---
LabeledSequence | *If applicable* | Similar to modified sequence. This sequence will specify the respective label for each channel in a labeled experiment. It will also contain any modifications. An example SILAClabeled peptide could look like this: _NAYVC [+57] WTLK_ for the light channel and _NAYVC [+57] WTLK [+8]_ for the heavy channel of the same peptide.

ProteinId | *optional* | The ID of the protein, the peptide is derived from. This information is used to label your peptides. Spectronaut Pulsar X provides filtering capabilities in the Analysis Perspective including filtering for the protein ID.

UserGroup | *optional* | A column for additional user information that can be used for grouping and filtering.

To view an example of a library, see our online material for download here or export a library from the Library Perspective in *.xls format.

### 3.3.3.2 Modification Parsing

Once the library is imported, Spectronaut Pulsar X will try to parse all values imported from the "ModifiedSequence" and the "LabeledSequence" columns to assign modification specifications to them. This allows Spectronaut Pulsar X to have greater control over decoy generation. If possible, Spectronaut Pulsar X will automatically assign known modifications from its internal database. If a certain modification is unknown, you will be prompted to assign the modification specification from the database to the new keyword (Figure 8). The only parsing requirement for external modification definitions is the modification tag which is specified within round or square brackets. Spectronaut Pulsar X will not parse modifications specified as single letter special amino acids (like ‘B’ for carbamidomethyl cysteine or ‘O’ for oxidized methionine). You can remove previously assigned parsing synonyms in the Modifications page of the Databases Perspective (see section 3.9.2).
### 3.3.4 Merging Libraries

Although Spectronaut Pulsar X allows for the merging of libraries, this can lead to an inflation of the protein FDR. Instead, we strongly advise to generate a new library from a search which includes all relevant LC-MS/MS runs, and so controlling the protein FDR at this step. Alternatively, you can use Pulsar Search Archives (see section 3.3.1 and Box 2).

If the above is not possible, two or more spectral libraries can be merged in the Library Perspective. To do so, simply select the spectral libraries you would like to merge while holding the Ctrl key and then, right-click to open the context menu and select the "Merge" option. This will open a setup window similar to when generating a spectral library from a database search (see section 3.3.2 and Figure 6).

![Assign Modifications](image)

**Figure 8.** Modification assignment during import of an external library. Add the synonym from the database by double-clicking or dragging it to the unassigned modifications.
Please note that if the spectral libraries used for merging have different types of protein annotations, protein counts in the merged library will be inflated (as the same protein could be counted twice). This will not happen if the libraries were generated in Spectronaut Pulsar X performing protein inference using the implemented IDPicker algorithm (Zhang et al. 2007).

### 3.3.5 Library Overview

Spectronaut Pulsar X provides several different plots for obtaining an overview of your spectral library. You can access these plots by clicking on the spectral library node in the tree and then selecting an appropriate plot in the right panel (Figure 9).

![Figure 9. Spectral library overview. Several plots can be selected from the drop-down menus. In this example, the top plot shows the Library Summary, while the bottom plot shows the missed cleavage overview.](image-url)
3.3.6 Making a Labeled or a Spike-in Library

By right-clicking on a spectral library in the Library Perspective, you can attach heavy labels to an existing library. Doing so will open the label editor form where you can select which labels should be applied to the existing library by double-clicking (Figure 10). The selected library will be stripped of any pre-existing labels. The selected workflow will be included in the library to define how these new peptides will be treated during analysis. For detailed information about the supported workflows see section 0.

Figure 10. Right-clicking on a Library opens a context menu with several options, like generating a labeled library or enable library for QC. In the figure the is an example of applying a SILAC label to an existing library. The two isobaric modifications Arg6 and Lys4 are selected to be applied as label to all applicable peptides.
3.3.7 Generate a Quality Control Kit from a Library

When right-clicking on a spectral library in the Library Perspective you have the option to generate a new QC kit using this library (Enable QC, Figure 10). This will select 250 highly abundant peptides from the spectral library which will be added as a QC kit to the quality control perspective. The selection of peptides can also be altered manually within the dialogue. These peptides can then be tracked for quality control purposes within the quality control perspective whenever the corresponding library was used.

3.4 Analysis Perspective

Spectronaut™ Pulsar X starts up in the Analysis Perspective (Review Perspective in older versions). This perspective allows you to:

1. Set up a DIA or a directDIA™ analysis
2. Review your data at any level through many of the useful plots, reports, and filters
3. Refine your quantification by modifying integration boundaries, fragment ions used for quantification, defining interferences, etc.

3.4.1 Setting up and Running an Analysis

Setting up your DIA analysis is straightforward thanks to the set-up wizards in Spectronaut Pulsar X.

Before starting, see Table 1 to make sure you have everything you need. After completing the wizard and click "Finish", Spectronaut Pulsar X will switch back to the Analysis Perspective and start the analysis.

First, in every run, Spectronaut Pulsar X will perform the a basic, linear iRT calibration using the iRT Kit peptides; then, Precision iRT calibration will be applied using a stored set of endogenous iRT peptides. Browsing your data is possible a few seconds after the initial calibration process is finished. Once the analysis has finished the number of unique precursors, peptides and protein groups identified for the Qvalue cutoff defined in the settings (default 0.01, equivalent of an FDR cutoff of 1%) will be shown at the bottom right (Figure 15).
Box 5. Spectronaut Pulsar X plots: how to get the most of them

Spectronaut Pulsar X provides, across all its perspectives, a comprehensive number of plots of many types to show you all the relevant details about your analysis, from MS data acquisition to post-analysis results. Most plots in Spectronaut Pulsar X are interactive and customizable at some extend. For example: zoom in on a plot by selecting the area you want to enlarge (find back to default scale by right-clicking on the plot). Drag or navigate a plot horizontally by Ctrl +click and drag.

By right-clicking on a plot, you will find a context menu with a large list of functionalities (see figure below):

![Plot with context menu]

In this example, you can show or hide the legend, save the data used for the plot, choose the unit you want to show numbers for (protein, peptides), among many others.

3.4.1.1 Performing a DIA Library-based Analysis

Most examples shown in this section are generated with the demo data available here. Please, download this data to reproduce the results. Alternatively, you can generate your own DIA data to test Spectronaut Pulsar X. To check which resources you will need in order to perform a DIA library-based analysis, refer to Table 1.
To start a library-based DIA analysis, go to the Analysis Perspective (add a new experiment tab if needed) and click on "Set up a DIA Analysis from..." in the bottom left corner. This will let you navigate to your run files or folders. Once you have chosen your DIA data, a wizard will start guiding you through the set-up (Figure 11):

1. **Set up DIA Analysis.**
   - Give a name to your analysis.
   - Select the whole experiment or individual runs, then click on "Assign Spectral Library". Different libraries can be assigned to different runs. More than one library can be assigned to a single run.

![Figure 11. Starting a library-based DIA analysis. After selecting your run files, a wizard will guide you through the process. You will be prompted to assign one or more libraries, the FASTA files, the GO annotation file, and to select your analysis settings. Finally, you will see a summary of the analysis set-up (Figure 12).](image)

2. **Choose Spectral Library.** Select the library from the Recently Used list, From File, or From Library Perspective. If the library is chosen from file, this action is similar to importing an external library, described in section 3.3.3. Click "Load" to add the library to your analysis.
3. In the next page, you select your analysis settings schema. Use one of the schemas available or modify one on the fly. These settings will define many important aspects of the analysis, such as FDR thresholds, quantification preferences, how to filter your data, among others. The BGS Factory Settings (default) schema is a good starting point for most projects. Find a detailed explanation of the analysis settings in Appendix 1. DIA Analysis Settings (section 6.1).

4. Choose a protein database (FASTA file) if you want Spectronaut Pulsar X to perform protein inference. Spectronaut Pulsar X performs protein inference according to the IDPicker algorithm (Zhang et al. 2007). Refer to Appendix 1. DIA Analysis Settings (section 6.1) for more details about this option.

5. Specify your experimental set-up (conditions, replicates) so Spectronaut Pulsar X can test for differential abundance and perform other Post Analysis processing steps. See section 3.4.1.3 for more information about how the conditions editor works.

Figure 12. Summary of the analysis set-up. Click “Finish” to proceed with the calculations.
6. Choose a Gene Ontology (GO) annotation if you want Spectronaut Pulsar X to give you extra biological insight into your experiment. This includes GO term enrichment and GO clustering.

7. Before clicking "Finish", a summary of your analysis set-up will be shown (Figure 12).

### 3.4.1.2 Performing a directDIA™ Analysis

Spectronaut Pulsar X enables directDIA™, Biognosys' library-free DIA workflow. This novel analysis allows you to directly search DIA files using nothing but a FASTA file for identification. Currently, directDIA is only supported for Thermo Scientific™ Orbitrap, and Sciex TripleTOF data. To know the resources required to perform a directDIA analysis, refer to Table 1.

To start a new directDIA analysis, go to the Analysis Perspective (add a new experiment tab if needed) and click “Set up a directDIA Analysis…” in the bottom-left corner. This will prompt you to navigate to your run files. Once you have chosen your DIA data, a wizard will start guiding you through the set-up (Figure 13):

1. **Set up directDIA Analysis.** Give your analysis a proper name.

2. Choose one or more FASTA files. The searches will be done against these databases.

3. Choose the search and analysis settings for your experiment. Use one of the schemas available or modify one on the fly. These settings will define many important aspects of the analysis, such as digestion rules, FDR thresholds, quantification preferences, how to filter your data, etc. The BGS Factory Settings (default) schema is fitting most projects. Find a detailed explanation of the directDIA settings see Appendix 3. directDIA™ Settings (section 6.3).

4. Specify your experimental set-up (conditions, replicates) so Spectronaut Pulsar X can test for differential abundance and perform other Post Analysis processing steps. See section 3.4.1.3 for more information on how the conditions editor works.

5. Choose the Gene Ontology (GO) annotation if you want Spectronaut Pulsar X to provide extra biological insight for your experiment, including GO term enrichment and GO clustering.
Note that a directDIA analysis includes a database search of DIA data by Pulsar. This step is resource intensive and can be time consuming. The algorithm followed by Pulsar to search DIA data is based on DIA-Umpire (Tsou et al. 2015).

Figure 13. Starting a directDIA analysis. After selecting your run files, a wizard will guide you through the process. You will be prompted to assign the FASTA files, the GO annotation file, and to select your analysis settings. Finally, you will see a summary of the analysis set-up before clicking “Finish”.

### 3.4.1.3 Configure Conditions

To let Spectronaut Pulsar X perform the differential abundance tests and other condition-wise metrics, you need to specify your experimental set-up during the configuration of your analysis. Spectronaut Pulsar X will ask you to annotate your runs and specify to which condition, biological replicate, and fraction (if applicable), they belong (see Box 6 to learn more about fractionation in Spectronaut Pulsar X). Each condition in Spectronaut Pulsar X will get a color assigned during the set-up which will be used for post-analysis plot labelling. The Conditions Set up panel contains several columns (Figure 14):
• The "Label" column is used for plotting purposes.

• You can specify which condition should be considered as the reference. "Is Reference" is used only for plotting purposes (when right-clicking on one of the profile plots under Candidates there is the option to choose "Normalize by Reference". This will divide every condition by the defined reference condition).

• You also have the option to enter quantity correction factors for each sample. Spectronaut Pulsar X will multiply the final quantities by this factor. An application example would be if quantities should be expressed per initial sample volume (for instance in plasma).

![Figure 14. Conditions Setup panel during DIA Analysis set-up. You can manually adjust your conditions on the panel or Import Conditions Setup from a text file.](image)

Unless actively disabled in the Analysis Settings, Spectronaut Pulsar X will perform a pairwise comparison (Student's t-test) of all conditions specified in the Condition Setup panel. The results are reported in the Post Analysis Perspective.

There are several ways to introduce the annotation information into the Condition Setup panel:

1. If you maintain a file-name structure which is self-annotating, you can define a parsing rule to automatically parse the conditions and replicates from it (Settings → Global → General → File Name Parsing Schema; see section 3.10.5.1).

2. The Condition Setup table is editable: you can directly write in any of the fields to enter your information (Figure 14). The table will recognize your changes and adapt them to the rest of the fields automatically. Be aware that the conditions editor is space and case sensitive.
3. Import your annotation from an external text file. The easiest way to do this is by exporting the current set-up, modify it and save it as a text file to import it.

### Box 6. Fractionation in Spectronaut Pulsar X

We do not recommend sample fractionation in DIA analyses. While for DDA, sample fractionation results in significantly higher coverage, the gains in the case of DIA are not that clear. In general, increasing the coverage in a DIA analysis is achieved by building a better spectral library.

In addition, one of the main features of DIA datasets is the low CVs and the high reproducibility. The process of sample fractionation introduces variability, sometimes notably high, which renders in a dataset of lower quality.

Although not recommended, Spectronaut Pulsar X supports sample fractionation. If you have your samples fractionated, you need to annotate this properly in the Conditions Setup. This will allow Spectronaut Pulsar X to perform normalization fraction-wise. Furthermore, libraries might have to be optimized in case of fractionation. If a peptide is not expected to be in a fraction, ideally it should not be targeted in that fraction. The reason for this is shown in the following table:

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Identified in F1R1</th>
<th>Identified in F1R2</th>
<th>Identified in F2R1</th>
<th>Identified in F2R2</th>
<th>Full Profile</th>
<th>Sparse Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>TRUE</td>
<td>TRUE</td>
<td>FALSE</td>
<td>FALSE</td>
<td>FALSE</td>
<td>TRUE</td>
</tr>
<tr>
<td>B</td>
<td>TRUE</td>
<td>FALSE</td>
<td>Not Targeted</td>
<td>Not Targeted</td>
<td>FALSE</td>
<td>TRUE</td>
</tr>
<tr>
<td>C</td>
<td>TRUE</td>
<td>TRUE</td>
<td>Not Targeted</td>
<td>Not Targeted</td>
<td>TRUE</td>
<td>TRUE</td>
</tr>
</tbody>
</table>

#### 3.4.1.4 Quantitative workflows supported in Spectronaut Pulsar X

Spectronaut Pulsar X not only performs label-free quantification. Labeled workflows are also supported, and specific scoring methods are developed for each approach.

- **Label-free**: Default workflow for all channel experiments. Peak detection, scoring and identification are applied as usual.
- **Labeled**: Peak detection and scoring will be applied to all channels. Quantification in Post Analysis will be performed on the light to heavy ratio.
• Spike-in: Peak detection will be performed on only the reference (heavy) channel. Scoring and identification will be performed on the target (light) channel. The heavy channel is expected to be easily detectable and considered a peak-picking aid in this experiment. Quantification in Post Analysis will be performed on the target to reference ratio.
• Inverted spike-in: Similar to spike-in but the light channel is considered the reference.

For more details about how to set these workflows, see Appendix 1. DIA Analysis Settings (section 6.1).

3.4.2 Reviewing your analysis

An analysis is shown in the Analysis Perspective organized as a data tree on the left and plots, reports and summaries on the right side (Figure 15). By default, the hierarchy of the data tree is:

> Run
  > Precursor window
    > Elution group
      > Precursor
        > Fragment ions

You can change this structure by right-clicking on the experiment tab → Group by, then select one of the options (more about this functions in Appendix 6. Experiment Tab Options, section 6.6).

In this data tree, you can right-click on any of the elements to find new functions which can be applied to that element (Figure 19).

The right side is divided into two panels (Upper and Lower panels), so you can display two different plots at the same time. These plots will change based on what is selected in the data tree. Different plots are applicable to different levels. To know which plots are available for each level, see Appendix 5. Analysis Perspective Plots (section 6.5).
Figure 15. Analysis view in the Analysis Perspective. Data tree is on the left side. Plots and summaries are on the right side. Right-clicking on the tab will open the Experiment Tab Options menu. A summary of the number of identifications is shown in green at the bottom.

### 3.4.2.1 Analysis Perspective Plots

One of the main tools Spectronaut Pulsar X provides for analysis reviewing in the Analysis Perspective is the comprehensive set of plots and reports available for the different levels of the data: Run, precursor and fragment ion.

- **Run level plots:** information about the calibration status, DIA method used, TICC, run meta information and cross run performance
- **Precursor and fragment level plots:** XIC chromatograms, score-centric plots and cross-run profile visualizations. The latter ones are only available in multi-run experiments and disabled for experiments containing only one run or peptides that are only targeted in one run.

Please, visit the Appendix 5. Analysis Perspective Plots (section 6.5). to find an example and a description of each plot.

To learn some tips about how to use the plots in Spectronaut Pulsar X, see Box 5.
3.4.2.2 Tree Filtering

Using the filtering system implemented in Spectronaut Pulsar X, one can apply one or several filters to the data tree. These filters only influence what is shown in the Analysis Perspective but not, for instance, the post-analysis perspective. Select a filter from the drop-down menu (Figure 16) and set the filter criteria. The filter is now checked within the drop-down menu. To combine filters, select a different filter and define the value that should be applied. A precursor has to apply to all selected filters in order to be shown in the review tree. By default, the precursor and protein Qvalue filters are set to what was chosen in the settings (default is 0.01).

*Note: sometimes is nonobvious that a filter is applied. Make sure you check the filter list before reviewing your analysis further.*

You can specify a custom criterion with the "User Group" filter. This value can be set during the library import by selecting a specific column as "User Group".

![Figure 16. Filtering the tree in the Analysis Perspective. Check the box for a filter and give the corresponding value. An example of filtering for a peptide sequence is shown on this figure.](image-url)
3.4.2.3 Experiment Tab Options

Right-clicking on the experiment tab in the Analysis Perspective opens a context menu with many functionalities that can be applied to the experiment (Figure 15). To see the full details for these options, refer to Appendix 6. Experiment Tab Options (section 6.6). Some of the most relevant ones are:

1. **Save and Save as:** Spectronaut Pulsar X will not save the analysis automatically. To save an analysis you will have to manually save it. You can save your analysis with or without ion traces (XICs, Figure 17):
   a. **With ion traces (FULL):** the file generated (*.sne file) will be larger, but Spectronaut Pulsar X will not require the run files to be available when you load your saved analysis again.
   b. **Without ion traces (XICs):** The *.sne file will be smaller, and you can map the run files after loading it.

![Figure 17. Saving your *.sne file with or without ion traces (XICs).](image)

2. **Group by:** change the structure of the data tree. The main level will still be "Run". See details in Appendix 6. Experiment Tab Options (section 6.6).

3. **Settings:** this option allows you to review and change many of the analysis settings without having to run the analysis again. For instance, you can change the FDR cutoffs, quantification settings, FASTA file for protein inference, conditions set-up, and many more (for details see Appendix 6. Experiment Tab Options, section 6.6).

3.4.2.4 Manual Analysis Refinement

Spectronaut Pulsar X will do high-throughput, automatic peak picking of your data. However, you can review and modify the assignment of peaks to precursors if you are not satisfied with some of them (see Box 7 on how to optimize the manual reviewing
process). For this purpose, the Analysis Perspective provides a number of aspects you can interact with and modify. The most relevant are:

- **Refine elution group integration boundaries.** Select an elution group in the data tree and set the right-side plots to, for example, MS2 XIC. You will see the ion chromatogram corresponding to the selected elution group within two green lines (Figure 18). You can manually slide these lines on both sides of the integrated peak, to set different boundaries. A new Qvalue should be calculated. A hand icon will appear next to the elution group in the data tree, denoting it was manually modified.

![Figure 18: Reviewing Spectronaut Pulsar X peak picking. Integration boundaries can be modified by dragging them. The precursor will be marked as manually modified.](image)

- **Manually select a different peak in the XIC.** Select an elution group in the data tree and set the right-side plots to, for example, MS2. Hover over the peak you want to assign and click when the cursor changes into a hand. The integration will be transferred, and a new Qvalue should be calculated. Similar to the action above, the precursor will be marked as manually modified in the data tree.
• **Manually accept or reject an elution group:** If you right-click on an elution group, you can manually accept or reject it. The icon next to the precursor will change to denote it has been manually modified (Figure 19).

• **Manually define an interfering fragment ion:** in the data tree, when you expand the precursor, you can see the fragments present in the library for that particular precursor. The ones used for quantification will have a blue icon, while the ones detected as interferences have a grey icon (Figure 20). You can define interferences manually by right-clicking on the fragment ion and unchecking the "Used for Quantification" option.

![Figure 19. Reviewing Spectronaut Pulsar X peak picking. Precursors can be manually accepted or rejected by right-clicking and choosing the option. The precursor will be marked as manually accepted or rejected.](image-url)
Box 7. Tips to optimize manual reviewing of your data (UI responsiveness)

If you need to manual review and actively navigate through your analysis in the Analysis Perspective, you might find some processes to be a bit slow and the software not as responsive as desirable. There are several things you can check in order to make the process as fast as possible:

1. Have your run files locally. Having your run files on a network drive is not recommended, and can significantly slow down the processes.
2. Convert the run files to HTRMS files (see section 4) before running your analysis
3. If you saved your *.sne file without XICs, re-extract your XICs (see section 3.4.2.3).
4. Group your data tree by precursor window

Happy reviewing!

Figure 20. Manually define interfering fragment ions or manually accept for quantitation fragments defined as interferences by Spectronaut Pulsar X.
3.4.2.5 Library Refinement

Spectronaut Pulsar X allows you to refine the fragment ion selection of your libraries. You can remove fragments that show interferences, specifically select fragments to cover interesting modification sites, or add additional fragments that were not detected in the original DDA analysis but are nicely visible in DIA.

To perform library refinement, the peptide assay must originate from a library generated in the Library Perspective. Right-click on the elution group node and select "Refine Fragment Selection". A new dialog will appear that shows you the selected peak in detail (Figure 21). The list on the left will show you all the fragments of this peptide that are present in the library, and which of those are currently selected. Additionally, a list of theoretical fragments is generated. In order to change the selection of theoretical fragments, right-click on the "Theoretical Fragments" node and select "Set Fragment Filter". This will allow you to expand the set of theoretical fragments so that it will contain different ion types, as well as common loss types.

![Fragment Selection Dialog](image)

Figure 21. Refine your library while reviewing your analysis. You can, for example, use different ion series (z-ions) by adding them into the fragment tree and selecting them. A preview of how the XIC looks with the current selected ions is shown on the right. You can also look at how individual ions XICs appear by selecting individual ions.

On the right, you can see a live preview of how the fragment selection will affect both the XIC and the match between the predicted and measured fragmentation pattern. Please
note that theoretical fragments will not contain a predicted intensity. To add or remove a
fragment from the spectral library, simply check or uncheck the corresponding fragments.

After the new fragment selection is complete, click "Apply Selection" in the bottom-right
corner to mark this peptide for refinement. Please note that the refinement is not
performed immediately. In order to effectively change the peptide in your experiment, as
well as in the spectral library, right-click on the experiment tab in the Analysis Perspective
and click "Commit Library Changes…".

This extra step is necessary since the update will require a re-extraction of this peptide
from all currently loaded runs. Spectronaut Pulsar X will, therefore, perform this
operation in batches once all manual fragment selection is done, and not for each peptide
individually. All changes made to a peptide’s fragment selection will only take effect once
the library changes are committed.

After clicking "Commit Library Changes…" another popup window will appear asking you
for a version name. Spectronaut Pulsar X features a version control for spectral libraries
that allows you to switch between different versions of a given library. This way, any
changes to the original library can be reverted. In order to change the version of a
spectral library, go to the Library Perspective, right-click on the respective library node
and select "Set Selected Version". A small window featuring a drop-down list will allow
you to select which version of this library to use.

3.5 Post Analysis Perspective

The Post Analysis Perspective in Spectronaut™ Pulsar X reports information that is less
raw data focused, as compared to the Analysis Perspective. It shows summary
information about identification, quantification, and results of the differential abundance
test, hierarchical clustering and GO terms enrichment and clustering (Figure 22).
3.5.1 Analysis Overview

Here you will find some experiment-wide information that will give you a rough idea about dataset characteristics. Under Overview, you will see a summary with the number of proteins and peptides identified (by conditions), miss-cleavages, library recovery, and other metrics that can help you judge the experimental outcome. To support this overview, you will find several plots related to the number of identifications, the data completeness, the coefficient of variation, and the normalization. On each of these plots, you can change many settings by using the right mouse click option. To see the full details of each plot, see Appendix 7. Post Analysis Perspective Plots (section 6.7). Learn more about plots in Spectronaut Pulsar X in Box 5.
3.5.2 Scoring Histograms

Under this node, you will find plots related to the behavior of the target and the decoy distribution estimation. This behavior defines the discriminant scores (Cscores), Qvalues, and sensitivity on the precursor level. Scoring histograms are shown for each workflow in the current experiment depending on your experimental set-up. Find all the details about these plots in Appendix 7. Post Analysis Perspective Plots (section 6.7).

3.5.3 Analysis Details

The binned identification plot in the analysis details shows the number of identifications across conditions and are binned according to three variables: iRT, intensity and m/z. This provides valuable feedback on the performance of the measurements for different conditions according to technical criteria such as liquid chromatography and mass spectrometer performance. The binned coefficients of variation show CVs within conditions and are binned according to the same three variables: iRT, intensity and m/z. See these plots in Appendix 7. Post Analysis Perspective Plots (section 6.7).

3.5.4 Differential Abundance

3.5.4.1 Candidates table

The results of the differential abundance testing will show up under this node. The Candidates node shows a table with the results, annotated by pair-wise comparison (Figure 22):

- The direction and the percentage of change are noted by color and color intensity, respectively; the level of significance is noted by the size of the circle.
- The fold changes are expressed as log transformed ratios of averaged replicates (AVG Log2 Ratio).
- The table is, by default, filtered by a multiple testing corrected Qvalue of 0.05 and an absolute log2 ratio of 0.58. You can change these filters to your preferred cutoffs. The filters applied to this table will automatically apply to the volcano plot as well.
- You can add and hide columns in this table by right-clicking on any of the headers and selecting Column Chooser. For example, you may want to add the pvalue column.
It is possible to search any character in the table with the Search field at the bottom of the table.

The candidates table can be exported as an excel file by clicking on "Export Table..." at the bottom.

In addition to the table, the candidates will be shown as plots on the right side. These plots can be customized in several ways by right-clicking on it and choosing your preferred options.

### 3.5.4.2 GO Enrichment

Under the Differential Abundance node, you will also find the results from the Gene Ontology (GO) term enrichment and the GO term clustering. If you added GO annotation to your experiment, either within the library or during the analysis set-up, Spectronaut Pulsar X will perform a GO term enrichment test.

Spectronaut Pulsar X comes with the human GO annotation implemented. If you are working with a different organism, you can download your relevant annotation from [http://geneontology.org/page/download-annotations](http://geneontology.org/page/download-annotations) or any other source, and import it into Spectronaut Pulsar X via the Databases Perspective (Databases Perspective → GO Databases → Import Gene Annotation). Your annotation is now ready to be appended to a library during the library generation, or to be selected during the analysis set-up.

The term enrichment test will check whether there are biological processes, functions or cell compartments over or under-represented within the candidate set. In other words, it will highlight processes, functions or compartments affected by the experimental conditions.

During the analysis, the first step is to determine how frequently a GO term occurs in the background proteome, i.e., all proteins identified throughout the whole experiment. Based on this information, this term is expected to be found a certain number of times in a random set of a given size. If the GO term occurs more frequently in your candidate set than expected, it is considered as significantly overrepresented; if it occurs less frequently, the term is considered as significantly underrepresented (Mi et al. 2013). The level of significance is given by a P-value. Spectronaut Pulsar X will perform two multiple testing correction methods to this test: Bonferroni (Dunn 1961) and Benjamini-Hochberg
(Benjamini and Hochberg 1995), for which the corresponding corrected P-values are also displayed.

If you change the candidate set, i.e., you apply a different filter in the Candidates table, the enrichment must be recalculated.

The result of the enrichment test will be shown as an interactive table where you can group the results according to a column or filter according to any of the features (Figure 23). Similar to the Candidates table, you can easily search within the table with the search field at the bottom. The term enrichment table can be exported by clicking "Export Table..." below the table panel.

Figure 23. GO term enrichment result. Similar to the Candidates table, you can apply filters. You can also group the results by any of the column headers by dragging them to the Group by field. If you want to do GO clustering on manually selected terms, use the first column of this table.
3.5.4.3 GO Clustering

GO clustering is a step further towards reducing the complexity of the differential abundance test results into an easier to interpret picture. If your GO term enrichment seems too convoluted, GO clustering will group related terms by similarity. The result is a shorter list showing groups of GO terms. GO clustering works based on the REVIGO algorithm (Supek et al. 2011). The semantic similarity of two terms is calculated based on their position/relation in the Gene Ontology graph.

Spectronaut Pulsar X will perform a GO term clustering on a subset of the terms from the enrichment analysis. This subset can be defined in two ways:

1. Manually selecting them in the GO Enrichment node, by activating the check box in the first column (Figure 23).
2. Filtering from the GO enrichment node by:
   - Namespace: biological process, molecular function, subcellular compartment
   - Type of representation: over, under, or both
   - Number of terms to cluster, ranked by Pvalue
   - Fold change
   - Number of proteins per term

In both cases, you will have to specify:

- The pair-wise comparison you want look at
- The term similarity cutoff you want to apply (being 0 lowest and 1 highest)
- Whether you want to apply a p-value cutoff and which one (corrected or not): If two terms are clustered, one of the terms will be chosen as cluster representative. The other will be discarded. Usually, the more specialized term "wins" since it carries more meaning (like ER lumen over cell). However, if you use p-values, the term with the lower p-value will win, even if it is a very general term.

When you are ready, click on "Start Clustering…"

The results show a list of GO terms (cluster representatives) on the left, and the dispensed terms that were clustered underneath them. The Dispensability score shows at what similarity cutoff the GO term would be clustered under another term.

As usual, you can export the table of results as an excel sheet by clicking on "Export Table…" at the bottom.
3.5.4.4 Differential Abundance Plots

Under the Differential Abundance node, several plots related to the significance test are also generated. Please, find detailed information on each of these plots in Appendix 7. Post Analysis Perspective Plots (section 6.7). The most relevant are the Heatmap and the Volcano Plot:

1. The Heatmap will be clustered row and column wise according to the Post Analysis settings. The raw data of the Heatmap can be exported via right-click on the plot (Figure 24).

![Heatmap with clustering in both rows and columns. The heatmap is built using the set of confidently identified datapoints. By right-clicking on the plot, you can perform a number of actions, such as save the Image As, export the data matrix or modify the scaling.](image)

2. The Volcano Plot shows the results of the differential abundance test by plotting the peptides or proteins fold change against the significance level. The candidates will appear in red on the plot (Figure 25).
Figure 25. The Volcano Plot shows the candidates in red. This plot is updated when you modify the Candidates table. By right-clicking, you can choose several actions, like deactivate the legend or annotate the candidates or not.

### 3.6 Report Perspective

Spectronaut™ Pulsar X has a very powerful reporting strategy. In the Report Perspective, you can design and customize your report to contain any information you may need about the analysis. Report schemas can be saved to be reused. You can also change the column names to fit your needs (Figure 26).

The Report Perspective contains four panels, from left to right:

1. The Schema tree: all different report building schemas. If you save a custom one, it will appear here.
2. Column chooser: all possible reportable elements with check boxes to add or remove them. Below this panel, there is a search field to help you navigate through the different fields.

3. Filters applied to the report.

4. Report preview: this one will only be visible if there is an analysis loaded. It will show a preview of how your report will look like. This is very useful when you are modifying a schema. When you are happy with your report structure, you can export it by clicking on "Export Report…" in the bottom left corner to be able to see the whole matrix.

Figure 26. Report Perspective. The figure shows the process of customizing a Normal Report schema and exporting the data. Detailed explanation of the headers can be found by hover over them of in Appendix 8. Most Relevant Report Headers

3.6.1 Report Schemas

Spectronaut Pulsar X includes several preconfigured reporting schemas that may fit most frequent needs. If you want to design your own, you can use one of the included as a base to build your preferred report.
Within the report schemas, there are two main formats you can export your data into: Normal Report and Run Pivot Report. Find detailed information about each format below.

### 3.6.2 Normal Report

In a Normal Report (long format), you will find each reported event in a single row (Figure 26). A Normal Report will usually have many more rows than a Run Pivot Report. This format is the one allowing for the most comprehensive report of your data. To build your Normal Report, add or remove columns from the Columns panel by checking or unchecking them (Figure 26). The Columns are organized by levels, from more general (Experiment) to more specific (Fragment):

- Experiment
- Run
  - Protein Group
  - Peptide
    - Elution Group
    - Fragment Group
    - Fragment

Within each of these levels, the columns are again organized by categories (e.g., identification, quantification, scoring, etc.). The whole Columns tree is quite comprehensive, and expanding/collapsing categories when looking for a column can be cumbersome: to make the search for columns easier, there is a search filed at the bottom of the Columns panel where you can type what you are looking for (Figure 26). Finally, to know which information a header contains, hover over it and you will see a text box popping up with a description.

To see a detailed description of some of the most relevant columns, see Appendix 8. Most Relevant Report Headers

### 3.6.3 Run Pivot Report

In a Run Pivot Report (wide format), each run (sample) will be a header column. You can choose which element you want to be rows in the Columns panel under Row Labels (e.g., stripped peptide sequence) and which value you want in the cells under Cell Values (e.g., quantitative value, Figure 27). If you choose more than one Row Label or
Cell Value, the table will multiply its length column-wise. This report will probably have fewer rows than a Normal Report.

Figure 27. Run Pivot Report. This report is in wide format contains one column per run (sample).

3.7 Quality Control Perspective

The quality control perspective of Spectronaut™ Pulsar X is based on the peptides in the iRT Kit. Chromatography, mass spectrometer performance and analysis can be monitored over time using several performance indicators. Every successful analysis is stored in the quality control perspective (Figure 28). Spectronaut Pulsar X automatically detects various instruments and will create a separate quality control history for each of them. If you have more than one instrument of the same type it might be useful to rename them manually. Additional folder structures can be made according to the established quality control testing performed in a specific laboratory.
Figure 28. QC Perspective. Runs in which the QC panel is detected are saved in the History tree. You can monitor instrument performance with help of many plots related to several aspects of the experiment, from LC-MS to Spectronaut Pulsar X analysis.

Only as many runs as specified in Settings → Global → General → QC Plot History Length are shown in the plots.

3.7.1 QC Panels

In addition to the iRT Kit, sample specific QC panels can be created in the Library Perspective. Right-click on a spectral library and select "Enable QC". Whenever this spectral library is used in an analysis, a respective QC file is written for each of the runs included in this analysis. When selecting this novel QC panel in the QC perspective, all the corresponding QC files (runs analyzed with the library enable for QC) will be available for QC monitoring.
3.8 Pipeline Perspective

The Pipeline Perspective is used to batch process library-based DIA analyses using predefined settings. Spectronaut™ Pulsar X works most efficiently when several experiments are processed sequentially rather than in parallel (because of disc IO). If you are not interested in manual evaluation of your peaks, the pipeline perspective might be your preferred choice. The set-up of an experiment works similarly to the set-up in the Analysis Perspective. The setup Analyses will be added to the Pipeline Queue. Clicking "RunPipeline" will start to process the queued experiments sequentially. Spectronaut Pulsar X will automatically generate the report according to the settings in the chosen schema (Figure 29).

Figure 29. Pipeline Perspective. Queue DIA analyses so Spectronaut Pulsar X can process them sequentially. Experiment files, reports, plots and summaries will be generated and stored according to the DIA Analysis Reporting Settings.
3.8.1 Pipeline Reporting Settings

Here you can choose which reports should be written, whether run-based or experiment-based analysis should be performed, whether scoring histograms should be reported and if the whole experiment should be saved to an *.sne file (Figure 30).

Figure 30. DIA Analysis Reporting Settings. Define which and were the results from the Pipeline Perspective should be saved.

3.9 Databases Perspective

The Databases Perspective allows you to store and manage information that you will need to use when setting up analyses. This includes protein databases, gene annotations, peptide modifications, etc.
3.9.1 Table Import

Spectronaut™ Pulsar X can remember column names in user spectral libraries. Once you import a new library format into Spectronaut Pulsar X, it will ask you whether it should store novel synonyms for column header. You can manage those synonyms in the column recognition settings tab.

3.9.2 Modifications

The Spectronaut Pulsar X modifications panel provides a central location to manage modifications. Modifications contained here are used to automatically detect modified sequences in a search engine output when creating a spectral library from an external search engine. They are also used to detect modified sequences encoded in an external spectral library by means of synonyms. We recommend that modifications in a spectral library are mapped to Spectronaut Pulsar X modifications for optimal downstream analysis (see section 3.3.3.2 and Figure 8). If a library is generated using Spectronaut Pulsar X, this is taken care of automatically.

Spectronaut Pulsar X comes with a database of default modifications for all search engines. If you use special modifications, please import the corresponding modifications file into Spectronaut Pulsar X.

3.9.2.1 Importing Modifications from Search Engine

To import non-default modifications into Spectronaut Pulsar X, you can batch import (see also Table 4):

- For MaxQuant, using the modifications.xml file in its bin folder
- For ProteinPilot, using Unified Modification Catalog.xlsx, located in the ProteinPilot/Help folder in the Program Files
- For Proteome Discoverer, no action is required
- For Mascot, non-default modifications have to be created as custom modifications, see below.

When possible, Spectronaut Pulsar X will merge identical modification from multiple sources and save only the necessary search engine specific mapping information. However, if it is not able to unambiguously merge two or more modifications, you will be asked to resolve any conflicts at the time of import. You can tell if a modification has been mapped to multiple search engines by looking at the "Mapped to" data grid in the panel.
3.9.2.2 Creating custom modifications

It is also possible to specify a new modification. This action has two main applications:

1. Incorporate modifications for Mascot searches which are not in the Unimod database (non-default ones).
2. Add a new label to generate a labeled library

To create a new modification, click "new" in the bottom left corner, give a name to your modification and click "OK" (Figure 31). Edit your new modification as desired, and click "Save" in the bottom left corner. You can also modify an existing modification by clicking "Save As…".

Figure 31. Adding a new modification to the database.

3.9.3 Protein Databases

This section of the settings perspective lets you import and manage your protein databases. Spectronaut Pulsar X uses protein databases (FASTA files) to make searches for library generation with Pulsar and to do protein inference. The protein databases contain all of the sequences, as well as Meta information extracted from the
FASTA protein headers, using the specified parsing rule. Spectronaut Pulsar X already contains the UniProt parsing rule, but you can add a new rule by clicking "New Rule" in the Protein Databases page or during an importation (Figure 32).

In order to import a new proteome database from FASTA click on "Import..." in the bottom left corner" (Figure 32). While importing a new protein database from FASTA, Spectronaut Pulsar X will try to find the appropriate parsing rule for this file format from the already specified rules. Should there be no matching parsing rule found, you will be asked to specify a new one. Once your new protein database is imported, it will be available in the Databases tree for setting-up Analysis.

3.9.4 GO Databases

Similar to the Protein Databases, gene ontology (GO) Databases in Spectronaut Pulsar X are used to further annotate your data. This annotation will be used for calculating term enrichment and give further biological insight into the differential
abundance results. The GO databases section manages two different data structures: Gene Ontologies and gene annotations. Find more details about each below.

### 3.9.4.1 Gene Ontologies

These allow you to import complex gene ontology structures in the shape of graphs. These structures are used for hierarchical grouping of functions, components and processes. Currently, Spectronaut Pulsar X supports the *.obo file format from the GO Consortium. The go-basic.obo is already part of the Spectronaut Pulsar X installation. Information from a gene ontology tree can only be used in combination with an organism specific gene annotation file.

### 3.9.4.2 Gene Annotations

The gene annotation file functions as a link between the protein identifier (Uniprot accession number) and the GO tree. In its most basic form, the gene annotation file must feature two columns:

1. The protein id that specifies the UniProt accession number
2. GO-ID that specifies the unique identifier of each GO term. In case you don't have the GO-ID information, this column will have to be substituted by two: the unique name of the term and its corresponding namespace (function, component or process).

Using this format, Spectronaut Pulsar X will connect the protein entries of your analysis via the GO-ID with the respective entries in the Gene Ontology to annotate your data further.

The official GO Consortium annotation file (*.gaf) is recommended but you can also specify a custom annotation file.

To import a new gene annotation file into Spectronaut Pulsar X, go to the GO Databases page of the Databases Perspective. Click "Import Gene Annotation..." and navigate to your *.gaf file. The GO annotation will automatically appear in your Gene Annotations tree.

### 3.9.5 Digest Rules

This tool lets you define the rule to in silico digest your proteins from the protein database(s). Digest rules are applied whenever you do a Pulsar search (in library
generation or in directDIA™). The most frequent rules are already included in Spectronaut Pulsar X, such as Trypsin, Trypsin/P, and LysC.

To design your own rule, you have to click on an existing one, modify it and click on "Save as..." in the bottom left corner (Figure 33). The rules are defined by which sites are cleaved by the enzyme. In the Digest Rule page, you will see a 20 x 20 matrix containing all possible combinations of amino acids. Select the combination where your enzyme cleaves (Figure 33). At the bottom, you will see a preview of how a sequence will look like after being cleaved following your digest rule. You can also include a description.

Figure 33. Define a new Cleavage Rule. The Cleavage Rule editor will allow the generation of new cleavages rules in a very friendly manner.
3.10 Settings Perspective

The Settings Perspective of Spectronaut™ Pulsar X is meant to define custom settings schemas for any of the processes performed by the software. In this perspective, you will see a tab corresponding to each of these processes: DIA Analysis, Pulsar Search, directDIA™ and Library Generation (Figure 34). In addition, you can alter global settings of Spectronaut Pulsar X in the Global page (see below).

Detailed information regarding each setting option can be obtained by hovering the mouse over the label of a specific settings variable (Figure 34).

Make your own setting schema by modifying one of the predefined ones. Go throughout the nodes and edit the corresponding settings. Once you are done with the customization, click “Save as…” in the bottom left corner to give a name to your schema, and click “OK” (Figure 34). Your new schema will appear in the tree and it will be available to be chosen during the set-up of your next analysis.

See the Appendixes for detailed information about the numerous settings within each process.

Figure 34. Make a custom schema for your analysis. The new schemas will be available during the subsequent analysis set-ups.
3.10.1 DIA Analysis Settings

The DIA Analysis Settings define the details of how Spectronaut Pulsar X should analyze the data, from DIA targeted data extraction to Post-Analysis calculations. These settings will specify important metrics, such as FDR cutoffs, decoy set generation to estimate scores, quantification settings, workflow to be used (label-free, labeled, spike-in), among many others. Find details of each setting in Appendix 1. DIA Analysis Settings (section 6.1).

3.10.2 Pulsar Search Settings

These settings define how Pulsar should create the search-space when performing a search. You can specify the expected peptide characteristics (enzyme used, length, modifications, among others). Find details of each setting in

3.10.3 directDIA™ Settings

The directDIA™ Settings combine configurations related to the Pulsar search and to the quantitative analysis. They define how Pulsar will generate the search-space and how the identification (FDR cutoffs) and the quantification will be performed. Find details of each setting in Appendix 3. directDIA™ Settings (section 6.3).

3.10.4 Library Generation Settings

This set of settings defines the Library Generation process, both from Pulsar or from an external search engine. Metrics such as MS1 and MS2 tolerances, FDR cutoffs for identification confidence, peptide-based filters for your library, among others. Find details of each setting in Appendix 4. Library Generation Settings (section 6.4).

3.10.5 Global Settings

The "Global" settings tab in the Settings perspective will allow you to change parameters that can be considered analysis unspecific. Here you will find options regarding plotting, working directories, as well as some general settings.
3.10.5.1 General

This section contains settings options that allow you to modify the default behavior of Spectronaut Pulsar X. For more information about these options use the tool-tip hover for each individual entry.

One important aspect of this settings is the File Name Parsing Strategy to let Spectronaut Pulsar X read information directly from the run file name. One of the most relevant uses of this function is the Conditions Setup annotations. Parsing of the file names requires the file annotation to be separated by the "_" character. By defining the meaning of the different blocks in the file name, Spectronaut Pulsar X is able to obtain this information automatically (Figure 35).

![Parsing Rule Editor in the General Settings of the Global Settings page. Setting this rule properly will let Spectronaut Pulsar X read annotation information directly from the run file name.](image)

3.10.5.2 Directories

Here you can setup the different storage paths for data managed by Spectronaut Pulsar X. Should you have a central storage location for all your DDA, you
can specify this location here. This will allow Spectronaut Pulsar X to automatically map the correct shotgun acquisitions during the setup of the library generation pipeline (see Box 4). Please note that all changes within the "Directories" section will require a restart of Spectronaut Pulsar X in order to take effect. See some recommendation in section 1.4.

3.10.5.3 Plotting

The plotting section allows you to customize the look and feel of most of the plotting options used in Spectronaut Pulsar X. You can specify whether XIC plots should show the integration boundaries, as well as the expected elution time. Additionally, you can also apply smoothing to your plots. For more information about these options use the tool tip hover for each individual entry.

3.10.5.4 Reporting

Here you can specify two things:

1. Where to locate the results of the analysis performed via the Pipeline Perspective.
2. The default name of the files exported from Spectronaut Pulsar X.

3.10.6 Spectronaut Pulsar X Command Line Mode

In addition to the visual pipeline mode, Spectronaut Pulsar X is also capable of running the pipeline from command line. To run Spectronaut Pulsar X in command line mode you simply call the Spectronaut.exe file using the following parameter.

-\r adds a raw file to the experiment
-\d adds all raw files of a specified directory to the experiment
-\a assigns a spectral library to every run in the experiment
-\ar assigns a spectral library to the last run added to this experiment
-\s [OPTIONAL] selects a settings schema (the schema must already exist)
  If this schema is not present, the default settings are used
-\o [OPTIONAL] specifies an output directory for the reports
  If this parameter is not present, reports will be generated in %Appdata%/Spectronaut/Results
-\n [OPTIONAL] specifies a name for this experiment
-\f [OPTIONAL] A regex file filter to select only certain runs while using the directory-add methods (-\d). Default parameter is ".*\(\.htms |\.raw |\.wiff\)$"
-con  [OPTIONAL] The path to condition setup file as exported via the Spectronaut Pulsar X Conditions Setup.
-lg   Starts the library generation

An example:

Spectronaut.exe  -d "C:\data\My Experiment"  -a "C:\data\My Experiment\library.txt"
-s "my_settings"  -o "C:\data\My Experiment\Results"  -n "My Experiment 1"  -f ".*\.wiff"

If you encounter problems with the automatic parsing of your spectral library, please first try to load the spectral library using Spectronaut Pulsar X’s graphical user interface and make sure that all necessary columns are recognized automatically.
4 HTRMS Converter

Together with Spectronaut™ Pulsar X installation, a complementary software called HTRMS Converter will be also installed.

The HTRMS Converter converts DIA run files into a Biognosys compatible format called HTRMS. These files are pre-processed and optimized to be analyzed in Spectronaut Pulsar X. Converting run files into HTRMS files is very useful if you need to analyze the same files several times. The overall analysis time will be significantly reduced.

The HTRMS converter is free to use and can be run on multiple computers without the requirement of a license key.

Figure 36. The new HTRMS converter with multiple tasks added to the task-list and monitoring a local directory for new MS/MS files.
4.1 File Conversion

In order to select one or more files from your hard-drive to be converted, click on the drop-down arrow in the Converter perspective and select "Add Files…". After selecting one or more MS/MS files from the hard-drive an input form will show that allows you to specify the conversion parameter. Click on "OK" to add all selected files to the main task list in the Converter perspective.

You can add more tasks to the list at any moment.

4.2 Folder Conversion

Using the drop-down arrow and selecting "Add Folder…" will ask you to specify an input folder. The folder conversion will automatically convert all valid MS/MS files within the target folder that meet some basic filter criterion.

Using the folder conversion, you now have access to the "Batch Conversion" settings, which allows you to specify filter criteria like vendor or file age. You can also specify monitoring of the folder in order to automatically convert every new file that is added to the input folder.
5 References


Tsou CC, Avtonomov D, Larsen B, Tucholska M, Choi H, Gingras AC, and Nesvizhskii

6 Appendixes

6.1 Appendix 1. DIA Analysis Settings

Data Extraction

Under default settings (Dynamic), Spectronaut™ Pulsar X will calculate the ideal mass tolerances for data extraction and scoring based on its extensive mass calibration. However, you can also specify your preferred tolerances for both MS1 and MS2 levels. You can choose among:

- Dynamic (default): determined by Spectronaut Pulsar X based on the extensive mass calibration. You can set a correction factor (default is no correction = 1)
- Relative: set a relative mass tolerance in ppm of the target ion mass
- Static: set a fix mass tolerance in Thomson

XIC Extraction

XIC RT Extraction Window defines whether iRT should be used to predict the elution of a peptide:

- Dynamic: Spectronaut Pulsar X will determine the ideal extraction window dynamically depending on iRT calibration and gradient stability. Sections of the gradient that show higher variability during the calibration step will automatically be extracted using wider windows. A correction factor of 2.0 would mean that you want to use 2 times the window that Spectronaut Pulsar X suggests. The default settings are recommended.
- Static: Spectronaut Pulsar X will use a fixed width (in min).
- Full: Spectronaut Pulsar X will use the full gradient width to find the target.
## Calibration

<table>
<thead>
<tr>
<th>Allow source specific iRT Calibration</th>
<th>Enable or disable source-specific iRT Calibration (Box 3).</th>
</tr>
</thead>
</table>
| Calibration Mode (only relevant when using HTRMS files) | None: keep current stored calibration.  
Automatic: Spectronaut Pulsar X will decide whether a recalibration is needed or not.  
Force: will recalibrate HTRMS files. |
| Precision iRT | Spectronaut Pulsar X will use a larger set of calibration peptides, if available, to perform an extensive calibration and improve iRT precision and accuracy.  
**iRT ↔ RT Regression Type:**  
Linear iRT calibration: will do a linear iRT calibration  
Non-linear iRT calibration: this allows to calibrate non-linear gradients, but also to correct for possible variations of a linear gradient. If Spectronaut Pulsar X does not find enough data points, it will automatically change to linear calibration |
| Calibration Carry-Over (recommended only for HCP workflows) | Spectronaut Pulsar X will take the calibration from the best performing run in the experiment and will apply this calibration to the rest of the runs. Consider using a higher XIC extraction width factor. |
## Identification

<table>
<thead>
<tr>
<th>Exclude Duplicate Assays</th>
<th>Spectronaut Pulsar X will keep the best performing assay only if a peptide is duplicated in the libraries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generate Decoys</td>
<td>If unchecked, decoys have to be provided in the library for Spectronaut Pulsar X to estimate Qvalues</td>
</tr>
<tr>
<td></td>
<td>Decoy Method: defines how to generate the decoys. For details, please use the text hovers in the software.</td>
</tr>
<tr>
<td></td>
<td>Decoy Limit Strategy: set the maximum number of decoys to be generated:</td>
</tr>
<tr>
<td></td>
<td>None: generates decoys for the whole dataset</td>
</tr>
<tr>
<td></td>
<td>Dynamic: set a fraction of the targets for decoy generation</td>
</tr>
<tr>
<td></td>
<td>Static: choose a fix number of decoys</td>
</tr>
<tr>
<td>Machine Learning</td>
<td>Per run: calculate Cscores and Qvalues run based</td>
</tr>
<tr>
<td></td>
<td>Across Experiment: makes an experiment-wise Cscore space, but can compromise the sensitivity.</td>
</tr>
<tr>
<td>Precursor Qvalue Cutoff and Protein Qvalue Cutoff</td>
<td>Choose your Qvalue (FDR) cutoff on precursor and protein level. Only those passing the cutoff will be considered identified and use for other subsequent processes.</td>
</tr>
<tr>
<td>Pvalue Estimator</td>
<td>Specify how you preferred the null distribution to be estimated to calculate the Pvalues: Kernel Density or Normal Distribution Estimator</td>
</tr>
<tr>
<td>Single Hit Definition</td>
<td>Define what should be considered a single hit: stripped sequence, modified sequence, or precursor</td>
</tr>
<tr>
<td>Exclude Single Hit Proteins</td>
<td>Discard protein groups identified with only one peptide hit (as defined above)</td>
</tr>
</tbody>
</table>
## Quantification

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interference Correction</td>
<td>Exclude fragment ions detected as interferences across all runs (Bilbao et al. 2015). If checked, set a minimum number of features to keep at the MS2 and MS1 level so Spectronaut Pulsar X can still do quantification.</td>
</tr>
<tr>
<td>Proteotypicity Filter (only if protein Inference is selected)</td>
<td>Choose whether you want to quantify only based on non-shared peptides, either at the level of protein (very stringent) or at the level of protein group.</td>
</tr>
<tr>
<td>Major (Protein) Grouping</td>
<td>Specify what should be considered as a protein (Protein Group Id or Gene Id).</td>
</tr>
<tr>
<td>Minor (peptide) Grouping</td>
<td>Specify what should be considered as a peptide (Stripped Sequence, Modified Sequence or Precursor).</td>
</tr>
<tr>
<td>Major group quantity</td>
<td>Specify how you want the minor groups to be used to calculate the major group quantities.</td>
</tr>
<tr>
<td>Major Group Top N</td>
<td>Use a specific range of the best minor group elements to calculate the major group quantities.</td>
</tr>
<tr>
<td>Minor group quantity</td>
<td>Specify how you want the precursors to be used to calculate the minor group quantities.</td>
</tr>
<tr>
<td>Minor Group Top N</td>
<td>Use a specific range of the best precursors to calculate the minor group quantities.</td>
</tr>
<tr>
<td>Quantity MS-Level</td>
<td>Choose which MS level you want to use to perform quantification: MS1 or MS2.</td>
</tr>
<tr>
<td>Quantity type</td>
<td>Decide which feature of the peaks should be used for quantification: area under the curve within integration boundaries or apex peak height.</td>
</tr>
<tr>
<td>Data Filtering</td>
<td>Decide how to apply the Qvalue filter on the precursors in an experiment-wide manner to quantify protein groups:</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>• Qvalue (default): only those precursors passing the Qvalue cut-offs will be reported (considered as quantified) and, accordingly, used for statistical testing of differential abundance. This filter is the only one producing a data matrix containing missing values, tagged as Filtered or NaN.</td>
<td></td>
</tr>
<tr>
<td>• Qvalue sparse: if a peptide precursor was identified passing the cut-off in at least one of the samples, it will be reported for all the samples. In the samples where it was not identified below the significance threshold (default ≤ 0.01) the best picked signal will be reported. This value can correspond to the real signal or to noise, and can be considered similar to an imputation. Qvalue sparse filtering is less stringent than Qvalue filtering.</td>
<td></td>
</tr>
<tr>
<td>• Qvalue percentile: this is a modified version of the Qvalue sparse in which you define in how many of your samples the peptide precursor needs to pass the Qvalue threshold. For instance, if you set a 50th percentile cutoff (0.5), the peptide precursor needs to pass the Qvalue in 50% or more of your samples to be reported</td>
<td></td>
</tr>
<tr>
<td>• Qvalue complete: the peptide precursor needs to pass the Qvalue threshold in all the samples to be reported. This is the most stringent filter and produces the smallest data matrix.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cross Run Normalization</th>
<th>Apply label free normalization to the whole dataset. Choose among:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Local Normalization, based on the Local Regression Normalization described by Callister et al. 2006.</td>
<td></td>
</tr>
<tr>
<td>• Global Normalization (by median or mean).</td>
<td></td>
</tr>
</tbody>
</table>

Row Selection: choose which full profiles to use for normalization (sparse, complete or percentile).
**Workflow**

Here you can specify if you are running a label-free analysis or a different kind of quantification.

<table>
<thead>
<tr>
<th>Multi-Channel Definition</th>
<th>Workflow</th>
</tr>
</thead>
<tbody>
<tr>
<td>From Library Annotation: will take the definition annotated in the library. Any precursors containing only one channel specification will automatically be treated as label-free. You can set what is the fallback option if the annotation detection fails.</td>
<td></td>
</tr>
<tr>
<td>Label-free: peak detection, scoring and identification are applied in all channels</td>
<td></td>
</tr>
<tr>
<td>Labeled: peak detection and scoring will be applied to all channels. Quantification in post-analysis will be performed on the light to heavy ratio.</td>
<td></td>
</tr>
<tr>
<td>Spike-in: peak detection will be performed on only the reference (heavy) channel. Scoring and identification will be performed on the target (light) channel. The heavy channel is expected to be easily detectable and considered a peak-picking aid in this experiment. Quantification in post-analysis will be performed on the target to reference ratio.</td>
<td></td>
</tr>
<tr>
<td>Inverted: change which channel should be considered as reference</td>
<td></td>
</tr>
<tr>
<td>Reference-based Identification: not only peak detection but also identification will be based on the reference channel</td>
<td></td>
</tr>
</tbody>
</table>
Profiling Strategy

The profiling workflow allows the user to carry over the measured iRT of peptides that could be identified (Qvalue ≤ 0.01) in certain runs to fix integration boundaries in runs where an identification could not be achieved:

Template Correlation Profiling: takes the best peptide signal in all runs as a template to find low abundant signals in the rest of the runs.

iRT profiling: takes the best peptide signal in all runs as a template and translates the empirical iRT to the integration boundaries of the low abundant signals in the rest of the runs.

Profiling Row Selection: choose a row-wise Qvalue threshold for to choose which rows to profile (Minimum, Average or none).

Profiling Target Selection: specify which precursors should be readjusted (non-identified or automatic).

Unify Peptide Peaks

Unify the peak picking across different charge states of the same modified peptide based on the highest scoring instance.

Protein Inference

Spectronaut Pulsar X is able to perform protein inference using the IDPicker algorithm (Zhang et al. 2007). Protein grouping will be well defined and protein group counts will be comparable across search engines and spectral libraries. Spectronaut Pulsar X also checks which peptides are proteotypic. The options are:

- Automatic: when using a Spectronaut library and all information is available, Spectronaut Pulsar X will re-calculate protein groups based on all identified peptides.
- From Search Engine: will keep the inference from the provided library.
- From protein-db matching: you can overwrite the existing grouping by choosing a database and digest rules.
### Post-Analysis

<table>
<thead>
<tr>
<th>Post-Analysis</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculate Explained TIC</td>
<td>Choose if and how the relative TIC is explained.</td>
</tr>
<tr>
<td>Calculate Sample Correlation Matrix</td>
<td>Choose whether you want to calculate this matrix. For large experiments it can be very time-consuming.</td>
</tr>
<tr>
<td>Differential Abundance Grouping</td>
<td>Select what the biological unit you want your results to be based on as defined in the quantification settings: Major (proteins) or Minor (peptides) Group. Smallest quantity unit: the unit to perform t-tests on.</td>
</tr>
<tr>
<td>Differential Abundance Testing</td>
<td>Choose whether you want Spectronaut Pulsar X to perform testing (t-test) or not.</td>
</tr>
<tr>
<td>Run Clustering</td>
<td>Choose whether you want to cluster your samples and potential candidates, and how.</td>
</tr>
</tbody>
</table>

### Reporting

These settings are only relevant when running analyses from the Pipeline Perspective. Specify if you want your analysis to be saved (with or without ion traces), which reports should be generated and saved, etc.
6.2 Appendix 2. Pulsar Search Settings

Configure the conditions for Pulsar to perform the search:

**Sequences**

Specify settings related to the peptide sequences.

<table>
<thead>
<tr>
<th>Digest Type</th>
<th>Specific: both N- and C-terminus follow the specified digest rules</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Semi-specific: only of the termini follows the specified digest rules</td>
</tr>
<tr>
<td></td>
<td>Unspecific: no digest rules</td>
</tr>
<tr>
<td>Enzyme/Cleavage Rules</td>
<td>Proteases used to <em>in silico</em> digest the proteins from the protein database(s). Defined in Databases → Cleavage Rules</td>
</tr>
<tr>
<td>Maximum Peptide Length</td>
<td>Maximum number of amino acids allowed for a peptide</td>
</tr>
<tr>
<td>Minimum Peptide Length</td>
<td>Minimum number of amino acids allowed for a peptide</td>
</tr>
<tr>
<td>Missed Cleavages</td>
<td>How many consecutive cleavage sites the protease could miss</td>
</tr>
<tr>
<td>Special AAs for Decoy Generation</td>
<td>Amino acids that, for decoy generation are swapped with their preceding amino acid in order to generate decoys that have the same peptide length distribution but slightly different masses than their target versions</td>
</tr>
<tr>
<td>Toggle N-terminal M</td>
<td>Pre-processing of the protein database by removing the N-terminal M (when there is one)</td>
</tr>
</tbody>
</table>
### Applied modifications

<table>
<thead>
<tr>
<th>Maximum Variable Modifications</th>
<th>Maximum number of variable modifications allowed to happen in one peptide at the same time. The higher this number, the more possible combinations, and the bigger the search space and the longer the calculation time</th>
</tr>
</thead>
</table>
| Select Modifications             | Fixed: the amino acid always contains the modification  
Variable: the amino acid might or might not be modified. The more variable modifications, the bigger the search space and the longer the calculation time |

#### 6.3 Appendix 3. directDIA™ Settings

The directDIA™ Settings are similar to a combination of DIA Analysis Settings, Pulsar Search Settings and Library Generation Settings. Find the details below.

### Sequences

Specify settings related to the peptide sequences.

| Digest Type   | Specific: both N- and C-terminus follow the specified digest rules  
Semi-specific: only of the termini follows the specified digest rules  
Unspecific: no digest rules |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme/Cleavage Rules</td>
<td>Proteases used to <em>in silico</em> digest the proteins from the protein database(s). Defined in Databases → Cleavage Rules</td>
</tr>
<tr>
<td>Maximum Peptide Length</td>
<td>Maximum number of amino acids allowed for a peptide</td>
</tr>
<tr>
<td>Minimum Peptide Length</td>
<td>Minimum number of amino acids allowed for a peptide</td>
</tr>
<tr>
<td>Missed Cleavages</td>
<td>How many consecutive cleavage sites the protease could miss</td>
</tr>
</tbody>
</table>
Special AAs for Decoy Generation

Amino acids that, for decoy generation are swapped with their preceding amino acid in order to generate decoys that have the same peptide length distribution but slightly different masses than their target versions.

Toggle N-terminal M

Pre-processing of the protein database by removing the N-terminal M (when there is one).

Applied modifications

| Maximum Variable Modifications | Maximum number of variable modifications allowed to happen in one peptide at the same time. The higher this number, the more possible combinations, and the bigger the search space and the longer the calculation time. |
| Select Modifications | Fixed: the amino acid always contains the modification Variable: the amino acid might or might not be modified. The more variable modifications, the bigger the search space and the longer the calculation time |

Identification

<p>| Machine Learning | Per run: calculate Cscores and Qvalues run based Across Experiment: makes a experiment-wise Cscore space, but can compromise the sensitivity. |
| Precursor Qvalue Cutoff and Protein Qvalue Cutoff | Choose your Qvalue (FDR) cutoff on precursor and protein level. Only those passing the cutoff will be considered identified and use for other subsequent processes. |
| Pvalue Estimator | Specify how you preferred the null distribution to be estimated to calculate the Pvalues: Kernel Density or Normal Distribution Estimator |</p>
<table>
<thead>
<tr>
<th>Single Hit Definition</th>
<th>Define what should be considered a single hit: stripped sequence, modified sequence, or precursor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclude Single Hit Proteins</td>
<td>Discard protein groups identified with only one peptide hit (as defined above)</td>
</tr>
</tbody>
</table>

**Quantification**

<table>
<thead>
<tr>
<th>Interference Correction</th>
<th>Exclude fragment ions detected as interferences across all runs (Bilbao et al. 2015). If checked, set a minimum number of features to keep at the MS2 and MS1 level so Spectronaut Pulsar X can still do quantification.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteotypicity Filter (only if protein Inference is selected)</td>
<td>Choose whether you want to quantify only based on non-shared peptides, either at the level of protein (very stringent) or at the level of protein group.</td>
</tr>
<tr>
<td>Major (Protein) Grouping</td>
<td>Specify what should be considered as a protein (Protein Group Id or Gene Id).</td>
</tr>
<tr>
<td>Minor (peptide) Grouping</td>
<td>Specify what should be considered as a peptide (Stripped Sequence, Modified Sequence or Precursor).</td>
</tr>
<tr>
<td>Major group quantity</td>
<td>Specify how you want the minor groups to be used to calculate the major group quantities.</td>
</tr>
<tr>
<td>Major Group Top N</td>
<td>Use a specific range of the best minor group elements to calculate the major group quantities.</td>
</tr>
<tr>
<td>Minor group quantity</td>
<td>Specify how you want the precursors to be used to calculate the minor group quantities.</td>
</tr>
<tr>
<td>Minor Group Top N</td>
<td>Use a specific range of the best precursors to calculate the minor group quantities.</td>
</tr>
<tr>
<td>Quantity MS-Level</td>
<td>Choose which MS level you want to use to perform quantification: MS1 or MS2.</td>
</tr>
<tr>
<td><strong>Quantity type</strong></td>
<td>Decide which feature of the peaks should be used for quantification: area under the curve within integration boundaries or apex peak height.</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| **Data Filtering** | Decide how to apply the Qvalue filter on the precursors in an experiment-wide manner to quantify protein groups:  
  • Qvalue (default): only those precursors passing the Qvalue cut-offs will be reported (considered as quantified) and, accordingly, used for statistical testing of differential abundance. This filter is the only one producing a data matrix containing missing values, tagged as Filtered or NaN.  
  • Qvalue sparse: if a peptide precursor was identified passing the cut-off in at least one of the samples, it will be reported for all the samples. In the samples where it was not identified below the significance threshold (default ≤ 0.01) the best picked signal will be reported. This value can correspond to the real signal or to noise, and can be considered similar to an imputation. Qvalue sparse filtering is less stringent than Qvalue filtering.  
  • Qvalue percentile: this is a modified version of the Qvalue sparse in which you define in how many of your samples the peptide precursor needs to pass the Qvalue threshold. For instance, if you set a 50th percentile cutoff (0.5), the peptide precursor needs to pass the Qvalue in 50% or more of your samples to be reported  
  • Qvalue complete: the peptide precursor needs to pass the Qvalue threshold in all the samples to be reported. This is the most stringent filter and produces the smallest data matrix. |
| **Cross Run Normalization** | Apply label free normalization to the whole dataset. Choose among:  
  • Local Normalization, based on the Local Regression Normalization described by Callister et al. 2006.  
  • Global Normalization (by median or mean).  
Row Selection: choose which full profiles to use for normalization (sparse, complete or percentile). |
| **Modifications** | Filter peptides according to modifications. Find more details by hovering over the option in the software |
Amino Acids | Filter peptides containing specified amino acids
---|---
Fragment Ions | Filter peptides not fulfilling the conditions specified regarding fragment ions. Find more details by hovering over the option in the software

**Workflow**

Here you can specify if you are running a label-free analysis or a different kind of quantification

| Profiling Strategy | The profiling workflow allows the user to carry over the measured iRT of peptides that could be identified (Q-value ≤ 0.01) in certain runs to fix integration boundaries in runs where an identification could not be achieved. Usually it is recommended to only profile peptides that were identified at least once over all runs. To determine which peptides within a profile require alignment, Spectronaut Pulsar X performs an outlier detection based on retention time shifts and differences in fragmentation patterns of peptides across runs. Further, the user has the option to only consider peptides for alignment that were not identified. The default setting for "profiling row selection" is "Minimum Q-value Row Selection" of 0.01. This is equivalent to a sparse profile filter which selects any profile that was identified with a Q-value cutoff of 0.01 in at least one of the runs.
| Unify Peptide Peaks | Unify the peak picking across different charge states of the same modified peptide based on the highest scoring instance.

**Post-Analysis**

| Calculate Explained TIC | Choose if and how the relative TIC is explained.
| Calculate Sample Correlation Matrix | Choose whether you want to calculate this matrix. For large experiments it can be very time-consuming.
Differential Abundance Grouping

Select what the biological unit you want your results to be based on as defined in the quantification settings: Major (proteins) or Minor (peptides) Group.

Smallest quantity unit: the unit to perform t-tests on.

Differential Abundance Testing

Choose whether you want Spectronaut Pulsar X to perform testing (t-test) or not.

Run Clustering

Choose whether you want to cluster your samples and potential candidates, and how.

6.4 Appendix 4. Library Generation Settings

Configure the library generation settings. Most settings are described below. For further information, there are also helpful text hovers directly in the software.

Tolerances

Spectronaut™ Pulsar X will, by default, calculate the ideal mass tolerances to generate the library. Spectronaut Pulsar X performs two calibration searches: based on the first-pass calibration (rough calibration), the ideal tolerance for the second-pass calibration is defined; based on the second-pass calibration (finer calibration), the ideal tolerance for the main search is defined. Spectronaut Pulsar X will do this under default settings (Dynamic).

However, Spectronaut Pulsar X allows you to set your preferred tolerances for the different MS instruments (Thermo Ion Trap, Thermo Orbitrap, TOF). Hence, for both the calibration search (second-pass, finer calibration), and the main search, you can define your tolerances:

- Dynamic: determined by Spectronaut Pulsar X based on the precedent search (default). You can set a correction factor for MS1 and MS2 levels (default is no correction = 1)
- Relative: set a relative mass tolerance in ppm for both MS1 and MS2 levels
- Static: set a fix mass tolerance in Thomson for both MS1 and MS2 levels
Protein Inference

You can have this option activated or deactivated. If you let Spectronaut Pulsar X do your proteins inference, you can refine your sequence settings further in these settings:

<table>
<thead>
<tr>
<th>Enzyme/Cleavage Rules</th>
<th>Proteases used to \textit{in silico} digest the proteins from the protein database(s). Defined in Databases $\rightarrow$ Cleavage Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digest Type</td>
<td>Specific: both N- and C-terminus follow the specified digest rules</td>
</tr>
<tr>
<td></td>
<td>Semi-specific: only of the termini follows the specified digest rules</td>
</tr>
<tr>
<td></td>
<td>Unspecific: no digest rules</td>
</tr>
<tr>
<td>Toggle N-terminal M</td>
<td>Pre-processing of the protein database by removing the N-terminal M (when there is one)</td>
</tr>
</tbody>
</table>

Spectral Library Filters

You have a number of options to filter the search engine results for library generation. There are filters at the level of fragment ion and at the level of precursor. The filters are quite self-explanatory. Please, use the hover text-tools if you need more information. Find below some of the most relevant.

<table>
<thead>
<tr>
<th>Modifications</th>
<th>Filter peptides according to modifications. Find more details by hovering over the option in the software</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino Acids</td>
<td>Filter peptides containing specified amino acids</td>
</tr>
<tr>
<td>Fragment Ions</td>
<td>Filter peptides not fulfilling the conditions specified regarding fragment ions. Find more details by hovering over the option in the software</td>
</tr>
</tbody>
</table>
Workflow

In-Silico Generate Missing Channels: if generating a library for an isotopically labeled sample, you can activate this option to *in-silico* generate an assay for a missing label pair. For example, if you have a sample of only spiked-in heavy aqua peptides, it will make a light/heavy library where the label-free channel is generated *in-silico* based on the heavy channel and with a consistent fragment ion selection. This option works with n-channel search results as long as you are working with isotopic labels.

iRT Calibration

Set your preferences for iRT calibration:

<table>
<thead>
<tr>
<th>Minimum Rsquare</th>
<th>Choose how strict you want to be to accept the iRT calibration of your data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use Stripped Sequence to Identify Reference Peptides</td>
<td>Ignore modifications when identifying iRT peptides for calibration</td>
</tr>
</tbody>
</table>
6.5 Appendix 5. Analysis Perspective Plots

6.5.1 Run Node Plots

iRT Calibration Chart

This chart shows the status of run calibration. Initially, only showing the peptides of the used calibration kit. Spectronaut™ Pulsar X now also supports non-linear gradients using a refined calibration based on the initial calibration and detailed information in the user library.

![iRT Calibration Chart](image)

Figure 37. iRT Calibration Chart showing the non-linear transformation from library iRT to actual predicted retention times. The chart is shown after extended non-linear calibration has been performed. The extended calibration allows you to correct even small gradient fine-structure fluctuations in order to get the most accurate retention time prediction for your library.
**XIC Extraction Width Chart**

Based on extensive calibration using the respective calibration Kit and information about the spectral library used, Spectronaut Pulsar X can determine the optimal XIC extraction width for your experiment. This extraction width is calculated dynamically over the whole gradient and allows Spectronaut Pulsar X to automatically adapt to areas with lower retention time prediction accuracy. The window width can be influenced with a correction factor in the Settings perspective (Settings → Analysis → Peak Detection → Correction Factor).

Figure 38. The XIC Extraction Width plot gives insight into your gradient stability and the overall accuracy of your library’s iRT values. The blue and the orange lines show the window selections as suggested by Spectronaut Pulsar X (blue) and as set by the user (orange → correction factor in the settings). These extraction window widths (y-axis) change over time (x-axis) based on gradient stability and iRT accuracy. The red dot's show your libraries iRT accuracy assuming a linear iRT to RT transformation. The green dots show the iRT accuracy using the extended non-linear iRT to RT transformation. The later one is used for the actual analysis.
**TIC Chromatogram & TIC Overlay Chromatogram**

The TIC Chromatogram plots show your run’s total ion current chromatogram. The TICC overlay combines the information of all runs of your experiment in one plot for better insight into instrument stability and amount of sample injected.

![Figure 39](#)

Figure 39. The TIC Overlay chromatogram shows the total ion current of all runs in your experiment in a single plot, giving insight into instrument stability and amount of sample injected.

**Analysis Log**

The analysis log contains all the information pertaining to the analysis of your whole experiment. In the event of errors one can consult the analysis log for detailed information of what went wrong.

![Figure 40](#)

Figure 40. The analysis log with detailed information about the analysis processes in Spectronaut Pulsar X.
6.5.2 Peptide and Fragment Plots

**MS2 XIC**

The default plot on Elution Group (EG or peptide), Fragment Group (FG or peptide precursor) and Fragment (F or fragment ion) level. This plot shows the extracted ion current chromatogram of the selected peptide. The plot contains the XICs for all fragments present in the library. Additionally, the expected retention time is marked (black dotted line) and the currently selected peak for this peptide (green area). On the Elution Group and Fragment Group level the coloring of the fragments is based on its expected relative intensity ranging from red (expected most intense fragment) to blue (expected least intense fragment). You can directly change integration boundaries in the plot.

On fragment level, this plot only shows the selected fragment in color and all others in gray.

There are a number of options available upon right-clicking in the plot such as switching the y-axis to log scale, toggling accept/reject the peak, changing x-axis scale to iRT, showing normalized intensities and showing the XIC chromatogram for the whole gradient.

![Figure 41. XIC chromatogram for the peptide MPEMNIK++. The color coding of the fragments indicates an overall good correlation to the expected fragment intensities. The dotted blue line indicates a potential interference that was detected by Spectronaut Pulsar X and automatically removed for relative quantitation.](image)
**MS2 XIC Sum**

The XIC Sum chromatogram chart shows the selected peptides quantitative information. The XIC shown is the sum of all fragment XICs that qualified for quantitation. All fragments that were excluded due to interfering signals are not used to calculate the sum XIC chromatogram.

![MS2 XIC Sum chromatogram](image)

Figure 42. The MS2 XIC Sum plot of the peptide MPEMNIK++. The interference highlighted in the MS2 XIC plot (Figure 41) has already been removed from the sum XIC chromatogram.
**MS2 Intensity Correlation**

This plot shows detailed information about the correlation of predicted and measured relative fragment ion intensity. The predicted values in red correspond to the relative intensities provided by the spectral library. The black lines correspond to the relative measured intensity of each fragment ion. Fragment ions with potential interferences are displayed as dotted lines.

![MS2 Intensity Correlation Plot](image)

Figure 43. The MS2 Intensity Correlation plot for a given peptide precursor. The plot indicates a very good correlation between the expected relative intensities (red) and the measured intensities of the library fragments.
**MS1 Isotope Envelope XIC**

This plot shows the monoisotopic precursor plus its most abundant isotopic forms as an XIC chromatogram. The XIC chromatograms on MS1 and MS2 level are expected to have identical apex retention times and elution shapes. As with the MS2 XIC chromatogram plot, the MS1 XIC chromatogram is also color coded to reflect the predicted relative intensities. A color coding from red (highest) to blue (lowest) indicates high correlation with the predicted abundance.

Figure 44. The MS1 Isotope Envelope XIC chromatogram for the precursor IILDLISESPIK++ as extracted by Spectronaut Pulsar X. The color pattern again indicates a high correlation to the expected relative abundances.
**MS2 Isotope Envelope Correlation**

Similar to the MS2 Intensity Correlation plot, this plot highlights the correlation between the expected and the observed fragment isotope patterns. A high correlation between the measured (black) and the predicted (red) abundances signals high confidence in the identification and quantification.

![MS2 Isotope Envelope Correlation](image)

Figure 45. The envelope correlation plot for two fragments with very high correlation with respect to the predicted abundance. The measured intensities (black) are almost perfect mirror images of the predicted (red) intensities.

**MS2 XIC Alignment and MS1 XIC Alignment**

The MS2 XIC Alignment and MS1 XIC Alignment plots allows you to visualize the extracted ion chromatograms of a single peptide across all of your runs. You can directly change integration boundaries in the plot.

![MS2 XIC Alignment and MS1 XIC Alignment](image)

Figure 46. The MS2 XIC Alignment of the peptide IILDLISERPIK++ across 4 runs. The x-axis is automatically changed to iRT to reduce chromatographic variance. The axis can be changed to retention time by right-clicking on the plot and un-selecting the “Use iRT Scale” option.
**XIC graph**

This plot combines the MS2 XIC and the MS1 Isotope Envelope XIC together and across all your runs. This is a nice way of looking how well MS2 and MS1 peaks correlates and how reproducibly they behave across runs.

Figure 47. The XIC graph for one peptide over the six runs files analyzed. The upper row shows the MS2 XIC. The dashed blue line indicates a potential interference. The lower row shows the MS1 Isotope Envelope XIC.
**iRT XIC Sum Overlay**

This plot allows you to show all sum XIC chromatograms of your selected peptide from all runs as an overlay plot. The x-axis scale is by default in iRT but on right-click can be changed to actual retention time.

![iRT XIC Sum Overlay Chart](image)

Figure 48. The iRT XIC Sum Overlay chart for the peptide IILDLISESPIK++. The 4 XICs correspond to the sum XIC of one peptide in the 4 different runs loaded for this experiment. The title of the plot additionally shows the peptides coefficient of variation, in this case 4.5%.
**MS2 Intensity Alignment**

Similar to the MS2 XIC Alignment, this plot gives detailed information about the signal stability for one peptide across several runs. The different bars show the relative abundance of each fragment ion across multiple runs. Using this plot, one can quickly identify an inconsistent signal by the change in the color pattern. Right-click on the plot and un-selecting "Normalize" to show intensities on absolute scale.

![Image of MS2 Intensity Alignment]

Figure 49. The MS2 Intensity alignment for a peptide containing 8 fragment ions. Each fragment ions relative intensity compared to the total peak height is indicated using a differently colored bar. The peptide was targeted in 24 runs. An inconsistent signal can be easily identified due to the sudden change in the color pattern.
**Cross Run RT Accuracy**

Similar to the MS2 Intensity Alignment chart, the Cross Run RT Accuracy plot allows one to quickly validate the peak picking across several LC-MS runs. The x-axis shows colored bars that correspond to the peptide in different runs. On the y-axis one can see the retention time in iRT.

The height of each box corresponds to the peak width at the start and the end iRT according to the y-axis. The line through the middle of the box shows the apex retention time in iRT while the blue colored boxes in the back show the total XIC extraction width. The colors of the bars again correspond to the relative intensities of the measured fragment ions. The bar with the green background is the currently selected node. The black, dotted, horizontal line corresponds to the expected retention time in iRT. You can hop to any other node by clicking on the colored bar.

**Figure 50.** The Cross Run RT Accuracy plot for a peptide measured in 4 different runs. The multi-colored bars correspond to the detected peak with the colors encoding the relative fragment intensities and the upper and lower boundary of the bar corresponding to the peaks start and end retention time in iRT.
**MS1 Spectrum at Apex**

This plot shows the MS1 signal at apex retention time for a given peptide. The chart automatically zooms in on the isotopic envelope and labels the different peaks accordingly. The red bars indicate the expected relative intensities of the different isotopic forms of the precursor.

Figure 51. MS1 spectrum at apex for the peptide IILDLISESPIK++ showing the monoisotopic precursor plus its first 3 isotopic forms.
**MS2 Spectrum at Apex**

Similar to the MS1 Spectrum at Apex, this chart shows the full recorded MS2 spectrum corresponding to a given peptide's LC-peak apex. By default, the plot only highlights the fragments as provided by the library. Additionally, other fragments can be calculated and annotated by right-clicking on the plot and selecting "Show All Theoretical Fragments". Furthermore, you can open the "Set Fragment Filter" window to open up the range of fragments and show, for instance, a-series ions.

![Figure 52](image.png)

Figure 52. MS2 Spectrum at Apex for the peptide ILLDLISESPIK++. The option "Show All Theoretical Fragments" was turned on and a-ions were selected in the fragment filter editor.
**Peptide Data Match (PDM) Plot**

The PDM plot is newly introduced with Spectronaut Pulsar X. This plot shows the MS features for a precursor and how well it matches the expected values. The main area of the plot (on the right side) shows the full recorded MS2 spectrum corresponding to the peptide LC-peak apex. Highlighted, you can see expected fragments for that peptide. At the bottom of this plot, the mass errors of each of the highlighted fragments are shown. On the left side, the MS1 spectrum at apex is depicted (similar to the MS1 Spectrum at Apex plot described before). Finally, at the top-right side, you can choose among several elements to show on the MS1 and MS2 plots. Some of the options are:

- Label the fragment peaks with the predicted ion name or with the m/z of the peak
- Show the mass error in ppm or in Th
- Choose which fragments to highlight on the MS2 spectrum: the ones used for the assay (Library Fragments), the ones detected in the library but not selected (Extended Library Fragments), or All Theoretical. You can also filter by fragment length, by type and by neutral loss.

![Figure 53. Peptide Data Match (PDM) plot for the precursor SHGQDYLVGNK+++. At the top-right panel you can select what to highlight on the MS1 and MS2 spectrum plots. In this case, all fragments in the library fulfilling the filter criteria where shown. In blue you can see the b-ions: in red, the y-ions. The mass errors, in ppm, are shown for each fragment. The MS1 spectrum is shown at the bottom-left.](image-url)
**Protein Coverage**

Spectronaut Pulsar X gives you a detailed overview of a protein's coverage within your analysis. The protein coverage plot shows you all the peptides of a protein that were targeted within your current experiment. The color coding by default indicates the confidence level for each peptide. This can be toggled on right-click to show the proteotypicity status for each peptide. This option is only available if protein inference was enabled during the DIA analysis.

Figure 54. Protein coverage plot for protein P31327. The protein sequence is covered by 77% by the peptides of the selected spectral library. You can choose to show the proteotypicity by right-clicking and changing the coloring schema.
### 6.6 Appendix 6. Experiment Tab Options

In the Analysis Perspective, right-click on the experiment tab. A context menu will open with several functionalities to apply to the analysis (Figure 15):

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add runs and Remove runs</td>
<td>Add or remove runs from the analysis. For the changes to take effect, you need to refresh the Post-Analysis.</td>
</tr>
<tr>
<td>Map missing runs</td>
<td>If Spectronaut™ Pulsar X lost the link with the run files, you can map them back. If your analysis contains XICs, this is not needed.</td>
</tr>
<tr>
<td>Recalculate Qvalues</td>
<td>If you manually change the peak picking, by selecting a different peak of changing the integration boundaries, the Qvalues need to be re-calculated.</td>
</tr>
<tr>
<td>Refresh Post Analysis</td>
<td>If you do any modification on the analysis, such as manually modify a peak or add new runs, the Post-Analysis has to be refreshed.</td>
</tr>
<tr>
<td>Re-extract all XICs</td>
<td>If you saved your experiment without XICs, it is recommended that you re-extract them from the run files for a better performance.</td>
</tr>
<tr>
<td>Export Experiment Settings</td>
<td>Export a report with the settings you used to run your analysis</td>
</tr>
<tr>
<td>Order Run by</td>
<td>Choose a criteria to order your runs</td>
</tr>
<tr>
<td>Group by</td>
<td>You can group your data tree under different criteria. Default and recommended for better performance is by precursor window.</td>
</tr>
<tr>
<td></td>
<td>If you have several workflows in your analysis (e.g., label-free and spike-in), Group by Workflow is very useful.</td>
</tr>
<tr>
<td>Reset All Peaks</td>
<td>Revert the manually modified peak picking back to the automatic one. Qvalues need to be recalculated</td>
</tr>
<tr>
<td>Commit Library Changes</td>
<td>If you refined your library in the Analysis Perspective, the changes will not make effect until your commit the changes (see section 3.4.2.5).</td>
</tr>
<tr>
<td>Save as and Save</td>
<td>Spectronaut Pulsar X will not save your analysis automatically. You need to actively save your experiment as. Choose to save with or without XICs (see section 3.4.2.3).</td>
</tr>
</tbody>
</table>
Export All… (Ctrl + R)  With this function, you can batch export many reports of your choice relevant to your analysis.

<table>
<thead>
<tr>
<th>Settings</th>
<th>This tool allows to explore and <strong>change</strong> many settings of your analysis:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Identification, Quantification, Protein Inference and Post Analysis.</td>
</tr>
<tr>
<td></td>
<td>• Different FASTA files can be loaded here, as well as GO annotation files.</td>
</tr>
<tr>
<td></td>
<td>• You can also change the Conditions set-up</td>
</tr>
<tr>
<td></td>
<td>• Finally, in the Summary tab, you can change the name of your analysis</td>
</tr>
<tr>
<td></td>
<td>Changing these settings will let you recalculate your analysis, which is significantly less expensive than running it again from scratch.</td>
</tr>
</tbody>
</table>
6.7 Appendix 7. Post Analysis Perspective Plots

**Data Completeness**

In the data completeness plot the "Cumulative Sparse Profiles" describes the cumulative number of identifications when looking at 1 to n runs, where n corresponds to the complete experiment. The "Cumulative Full Profiles" corresponds to identifications that were consistent across 1 to n runs, i.e. identified in all the runs currently looked at. The Coefficients of Variation plot shows the distribution of CVs in the experiment. On right-click you can change the basis of quantitation between precursor, peptide, protein and protein group.

Figure 55. The upper plot under the "Data Completeness" node shows you the difference between cumulative sparse and full profiles. The green bars represent the growth of cumulative precursor identification. So for example, if you only consider the first 3 runs \( (x = 3) \) 26878 precursors where identified at least once in these 3 runs. The blue bars represent the decline of full profiles. For example, after the first 3 runs, only 22747 precursors where identified in all 3 previous runs. Adding more runs, the green bars can only ever go up or remain constant while the blue bars can only ever go down or remain constant. Like with many other plots you can change the context to display the data completeness on precursor, peptide, protein-group or protein level.
Figure 56. The lower plot under the "Data Completeness" node shows the histogram of missing values. In this plot, about 16500 precursors did not have any missing values and were identified in every run (this is equivalent to the last blue bar in Figure 55) while about 2200 precursors had one missing value and so forth. The last bar represents the number of single-hit precursors. These are the precursors that were only identified in one single run (hence they were only identified in 1 out of 18 runs in this case). Like with many other plots you can change the context to show you the missing values on precursor, peptide, protein-group or protein level.
**Coefficients of Variations**

Figure 57. The upper plot under the "Coefficients of Variation" node shows the %CV distribution for all conditions in your experiment. You can also change the context by right-clicking on the plot to show you the CVs for precursor, peptide, protein-group and protein quantities or to show the CV distribution across all conditions.

Figure 58. Similar to Figure 57, this plot shows the distribution of CVs for each condition. For experiments with large numbers of conditions, this view might be preferable. Like with the previous figures, you can change the context of the plot by right-clicking on it to select either precursor, peptide, protein-group or protein scope.
Figure 59. The CVs below X plot shows the number of precursors that were below either 20% or 10% CVs. The red bar shows the number of all identifications for each condition regardless of quantitative precision. The bar for ≤ 20% CVs also includes all counts from ≤ 10% CVs. Like with the previous figures, you can change the context of the plot by right-clicking on it to select either precursor, peptide, protein-group or protein scope.
Normalization

In the normalization you can see boxplots of responses for the individual runs before and after normalization.

Figure 60. The "Normalization" node shows details about the normalization status of your experiment. The left side shows boxplots of precursor responses before normalization for each run. The right side shows boxplots of the same precursor responses after normalization.
Figure 61. Heatmap showing the clustering of 9 runs from 3 conditions. Runs within the same condition cluster nicely as illustrated by the condition based color code in the bottom of the heatmap and the x-axis dendrogram.
Volcano Plot

Figure 62. Volcano plot showing the potential candidates of an experiment containing 3 conditions to each 3 replicates. By default, the filters are set to $\geq 1.5$ absolute fold change and $\leq 0.05$ Qvalue.
Sample Correlation Plot

Figure 63. The sample correlation matrix shows correlation of precursor quantities between all samples. The rows and columns of the matrix are ordered by condition and replicate annotation. A high correlation between technical replicates of the same sample is to be expected while low correlation between different samples might indicate biological variance. By default, the coloring range is set from 0.75 to 1.0 but can be changed via right-click option.
6.8 Appendix 8. Most Relevant Report Headers

Many headers have a text hover tool directly in the software. If you don’t find the information you are looking for, do not hesitate to contact us at support@biognosys.com.

Protein Group (PG) headers

Headers related to Protein Group (PG) as defined in the settings. Most headers related to PGs are self-explanatory. Here are the most relevant and some which are not too obvious.

| PG.ProteinGroups | One or several protein groups separated with a "|". Protein ids within protein groups are separated with a ";". The protein groups can either originate from the Spectronaut Pulsar X IDPicker protein grouping or from the search engine used to generate the spectral library. |
|-------------------|--------------------------------------------------------------------------------------------------|
| PG.Quantity       | The quantitative value as defined in the settings.                                              |
| PG.Qvalue         | The Qvalue (FDR) for that PG. Qvalues for PGs are experiment-wise.                             |
| PG.RunEvidenceCount | Number of hits (precursors) found for that PG in that run.                                    |
| PG.IsSingleHit    | True or False. It tells you whether the PG was found with only one hit, as defined in the settings. |
| PG.ProteinAccessions | The protein accessions in the same order as the protein ids in "PG.ProteinGroups". The value corresponds to what was defined by the parsing rule. If all the values are identical a unique value is reported |
| PG.Genes          | The genes in the same order as the protein ids in "PG.ProteinGroups". The value corresponds to what was defined by the parsing rule. If all the values are identical a unique value is reported |
| PG.ProteinDescriptions | The protein descriptions in the same order as the protein ids in "PG.ProteinGroups". The value corresponds to what was defined by the parsing rule. If all the values are identical a unique value is reported |
### PG.UniProtIds

The UniProt ids in the same order as the protein ids in "PG.ProteinGroups". The value corresponds to what was defined by the parsing rule. If all the values are identical a unique value is reported.

### PG.ProteinNames

The protein names in the same order as the protein ids in "PG.ProteinGroups". The value corresponds to what was defined by the parsing rule. If all the values are identical a unique value is reported.

### PG.FastaHeaders

The FASTA headers in the same order as the protein ids in "PG.ProteinGroups". The value corresponds to what was defined by the parsing rule. If all the values are identical a unique value is reported.

---

**Peptide headers**

Headers related to Peptides (PEP) as defined in the settings. Many headers related to Peptides are self-explanatory. Here are the most relevant and some which are not too obvious.

<table>
<thead>
<tr>
<th><strong>PEP.GroupingKey</strong> and <strong>PEP.GroupingKeyType</strong></th>
<th>Tells you which element is considered as peptide (defined in the settings). Default is stripped sequence.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PEP.IsProteinGroupSpecific</strong></td>
<td>True or False. Tells you whether the peptide only belongs to one Protein Group.</td>
</tr>
<tr>
<td><strong>PEP.RunEvidenceCount</strong></td>
<td>Number of hits (precursors) found for that peptide in that run.</td>
</tr>
<tr>
<td><strong>PEP.Quantity</strong></td>
<td>The quantitative value for that peptide as defined in the settings.</td>
</tr>
<tr>
<td><strong>PEP.UsedForProteinGroupQuantity</strong></td>
<td>True or False. Tells you whether this peptide was used to calculate PG quantities, as defined in the settings.</td>
</tr>
</tbody>
</table>
**Elution Group (EG) headers**

Headers related to Elution Groups (EG). Many headers related to EG are self-explanatory. Here are the most relevant and some which are not too obvious.

<table>
<thead>
<tr>
<th>EG.PrecursorId</th>
<th>Unique Id for the precursor: [modified sequence] plus [charge]</th>
</tr>
</thead>
<tbody>
<tr>
<td>EG.Library</td>
<td>The library this assay is from</td>
</tr>
<tr>
<td>EG.Workflow</td>
<td>The workflow this EG was used in. It can be LABEL_FREE, SPIKE_IN or LABEL</td>
</tr>
<tr>
<td>EG.IsUserPeak</td>
<td>True or False. Tells whether the EG was manually integrated</td>
</tr>
<tr>
<td>EG.Identified</td>
<td>True or False. The EG has to pass the precursor and protein Qvalue cutoff to be considered identified</td>
</tr>
<tr>
<td>EG.Qvalue</td>
<td>The Qvalue (FDR) of the EG.</td>
</tr>
<tr>
<td>EG.TotalQuantity</td>
<td>The quantitative value for that EG as defined in the settings.</td>
</tr>
<tr>
<td>EG.Label</td>
<td>A label for a peptide precursor compiled of the protein id, stripped sequence and the precursor ion charge. The label is not necessarily unique and therefore not intended for downstream peptide profiling across runs</td>
</tr>
<tr>
<td>EG.HeightCV</td>
<td>In a multi-run experiment, this value reports the CV of the peptide based on the summed apex height of the detected peak</td>
</tr>
<tr>
<td>EG.AreaCV</td>
<td>In a multi-run experiment, this value reports the CV of the peptide based on the summed area under the curve of the detected peak</td>
</tr>
<tr>
<td>EG.iRTEmpirical</td>
<td>The iRT (Escher et al. 2012) as determined in this specific analytical run</td>
</tr>
<tr>
<td>EG.MeanApexRT</td>
<td>The average retention time of the peak apexes across all fragment ions of this peptide</td>
</tr>
<tr>
<td>EG.MeanTailingFactor</td>
<td>The average tailing factor of the elution group across all the fragment ions determined at the FWHM</td>
</tr>
<tr>
<td><strong>EG.DeltaRT</strong></td>
<td>The difference between the predicted and empirical iRT. A measure of the reproducibility of chromatography</td>
</tr>
<tr>
<td><strong>EG.IsVerified</strong></td>
<td>Reports manual accepting or rejecting of peaks assigned in the Analysis Perspective</td>
</tr>
<tr>
<td><strong>EG.Cscore</strong></td>
<td>The Spectronaut Pulsar X identification score, which is based on an advanced mProphet (Reiter et al. 2011) scoring. A high score indicates high quality identifications</td>
</tr>
<tr>
<td><strong>EG.IsUserPeak</strong></td>
<td>Specifies whether the peak was integrated automatically or manually in the Analysis Perspective</td>
</tr>
<tr>
<td><strong>EG.AllProteinAccessions</strong></td>
<td>All protein accessions this peptide points to. This field is only reported when the Spectronaut Pulsar X protein inference was used. It represents the input to the IDPicker protein grouping algorithm</td>
</tr>
</tbody>
</table>

**Fragment Group (FG) headers**

Headers related to Fragment Group (FG). FG is only relevant in labeled and spike-in workflows. Two FGs belong to one EG. The FG id corresponds to the EG id plus the isotopic labelling.

Many headers related to FG are self-explanatory. Here are the most relevant and some which are not too obvious.

| **FG.Id** | A unique ID of the peptide precursor. Corresponds to the EG.Id if the experiment is label free |
| **FG.Label** | A label of the fragment ion group. The label is not necessarily unique and therefore not intended for structuring data |
| **FG.Charge** | The charge state of the peptide precursor |
| **FG.PrecMz** | The peptide precursor m/z |
| **FG.PrecWindowNumber** | The precursor window number in which the precursor was measured |
FG.TotalPeakArea | The summed up peak area of all fragment ions for the corresponding peptide precursor
---|---
FG.TotalPeakHeight | The summed up apex peak height of all fragment ions for the corresponding peptide precursor
FG.Quantity | The quantitative value for that FG as defined in the settings.

**Fragment (F) headers**

Headers related to Fragment ions. If you choose fragment ion level information in your report, you will have one row per fragment ion. This can make the report considerably large.

Most headers related to Fragments are self-explanatory. Here are the most relevant and some which are not too obvious.

<table>
<thead>
<tr>
<th>F.InterferenceScore</th>
<th>Fragment ion interference score as determined by Spectronaut Pulsar X</th>
</tr>
</thead>
<tbody>
<tr>
<td>F.PossibleInterference</td>
<td>True or False. Tells you whether an ion is a probable interference.</td>
</tr>
<tr>
<td>F.ExcludeFromQuantification</td>
<td>True or False. An ion can be excluded via the library or if it is considered an interference.</td>
</tr>
<tr>
<td>F.NormalizePeakArea</td>
<td>The quantitative value calculated as the area under the curve</td>
</tr>
</tbody>
</table>