

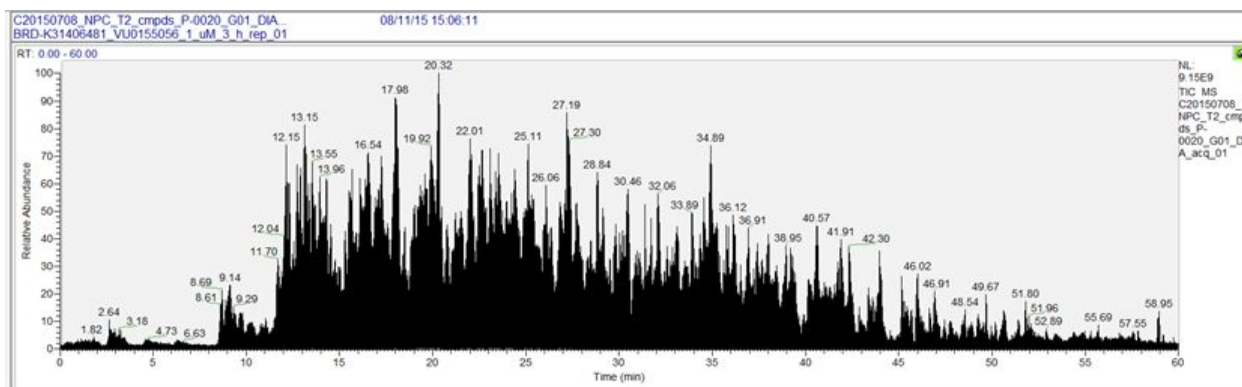
P100 Data Acquisition Guidelines

The purpose of this data sheet is to give general guidelines and tips for acquiring DIA data for P100. This is not an exhaustive database of everything that can go wrong, just the basics.

DIA data acquisition

On a newly conditioned column with the instrument running well you can expect the TIC of a P100 DIA run to look like the one below. The file size is expected to be between 2.5-3.5Gb.

TIC intensity ideally high e9's, does shape look normal? Keep an eye out for sparse MS2 spectra, TIC only reflects the MS1, file size low could be bad MS2's



Over the course of running the plate, the overall file size will decline somewhat due to column aging. **This is normal.** You should be worried if the file sizes start to drop drastically and continues across multiple sets of replicates. Sometimes a drug treatment will kill the cells and result in very little signal for its triplicates but this isn't indicative of a problem unless multiple triplicate sets have low signal.

If the file size drops below **1Gb**, you should check the MS quality by running Jurkats and will likely need to change the column. This can also be indicative of the MS being dirty and it may need to be cleaned.

A good rule to monitor how the column and MS are doing is to run Jurkats once per quadrant and search them as soon as they're finished running. In general, a good Jurkat peptide count number for Copernicus is between ~34,000 and 37,000, but good DIA data can still be acquired with peptide counts as low as ~28,000. Use good judgment.

Summary:

Ideal File Size: **1.5-3.5Gb**

Inject Volume: **3uL**

Current DIA method:

C:\Xcalibur\methods\LINCS\CurrentDIA\P100_DIA_11amu_Overlap_22amuwindow_60min_50msIT_400-1000_27loopcount

DDA data acquisition

Data dependent acquisitions of P100 samples are needed for:

- The eventual creation of spectral libraries
- Phosphopeptide Enrichment Quality Metrics

DDA runs should be done at the beginning of running a plate and back to back with the sample's DIA run (DIA followed by DDA). Historically, we've collected DDA runs for 2uL injections of B01, G01, A12, H12, and D06 in order to assess if there are plate effects if an enrichment fails.

The TIC should look like the same as the DIA run (the method uses the same gradient) and the file size is expected to be approximately 0.5Gb.

Summary:

Ideal File Size: **0.5Gb**

Inject Volume: **2uL**

Current DDA method:

C:\Xcalibur\methods\LINCS\LINCS_60min_DDA_2uL_centroid_USEforSAMPLES

Searching DDA data using SpectrumMill

Browse to \\musketeer\msdataSMLINCS
Create a new folder for the current P100 LINCS plate
Create subfolders named B01, G01, A12, H12, and D06
Place copies of the appropriate DDA files into their respective folders

Using internet explorer, open SpectrumMill server **dunlop**:

<http://dunlop/millhome.htm>

Open 'Workflows' and scroll to 'Jenn\QE_PhosphoProteome_Human'

The parameters used in this workflow can be found below

Select the directories that you just created on musketeer and press 'Execute'

SpectrumMill will search the data according to the parameters specified. This should take 1-2 hours for 5 files.

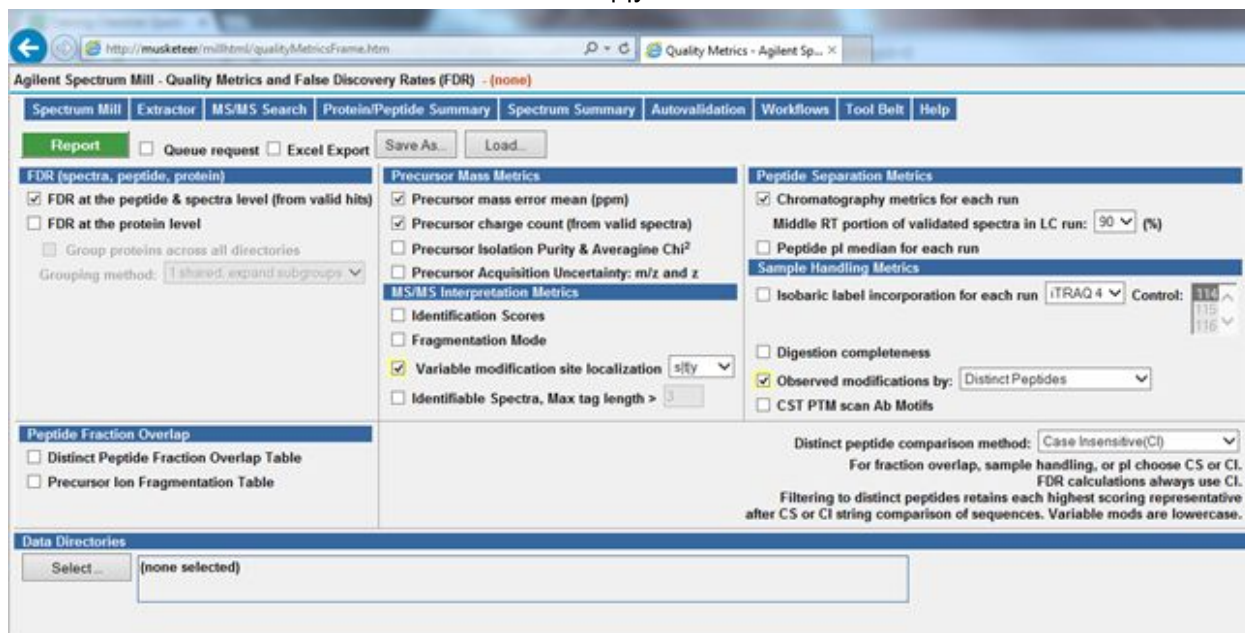
In SpectrumMill, return to the home page and browse to 'Quality Metrics and FDR'

Select the data directories that you just searched

Check the following parameters to return relevant quality metrics:

FDR at the peptide & spectra level (from valid hits)

Precursor mass error mean (ppm)
 Precursor charge count (from valid spectra)
 Chromatography metrics for each run, Middle RT portion of the validated spectra in LC run: 90%
 Observed modifications by: Distinct Peptides
 Distinct peptide comparison method: Case Sensitive
 Variable modification site localization: s|t|y



Press 'Report' to return quality metrics

For a quick look at how well the phosphopeptide enrichment performed, look at the following metrics:

s|t|y Sites spectra (%): Ideal is above 90%, average result is 85% but >70% is usually acceptable

s|t|y Sites spectra (#): Average result is around 8,000

Distinct Peps CI Total (#): Average result is around 5,000 to 7,000 peptides identified

Another way to assess enrichment efficacy is to make a SpectrumMill report of phosphopeptide only intensities and another with intensities of all peptides ID'd in the run. The percentage of total intensity is generally >95% phosphopeptides. Briefly, to do this:

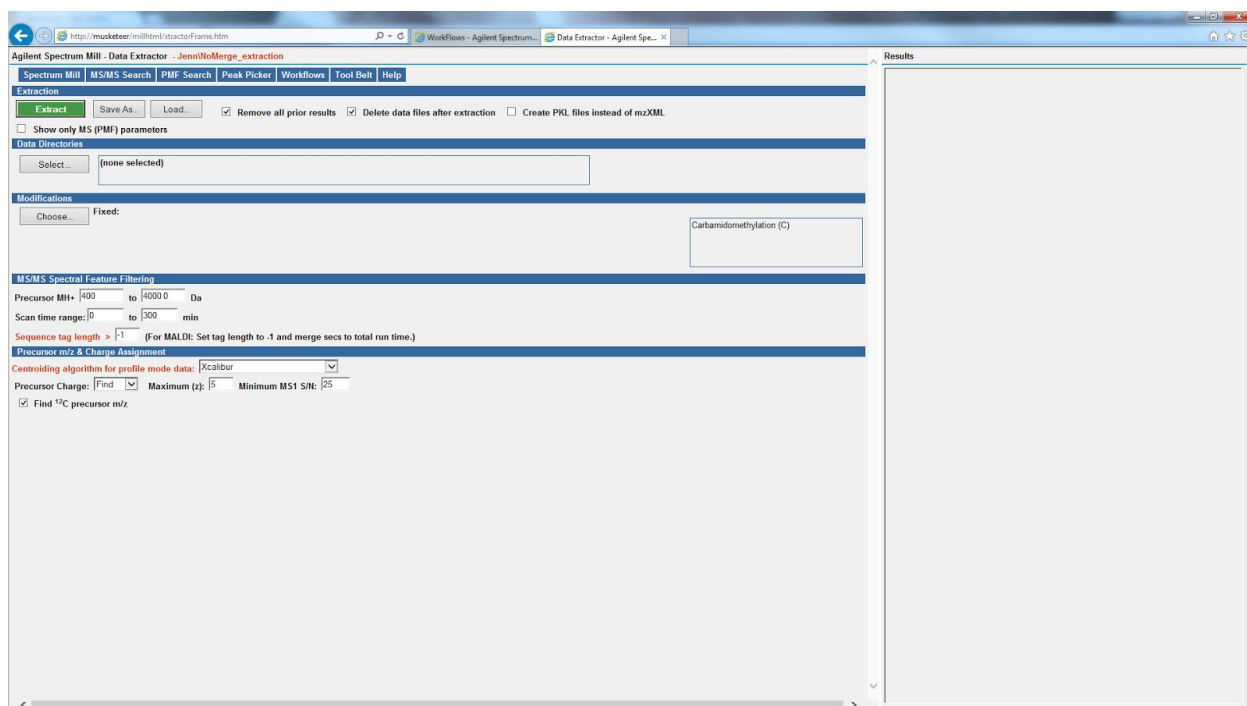
Create a SM peptide report of all peptides ID'd

Create a SM peptide report of only s|t|y peptides ID'd

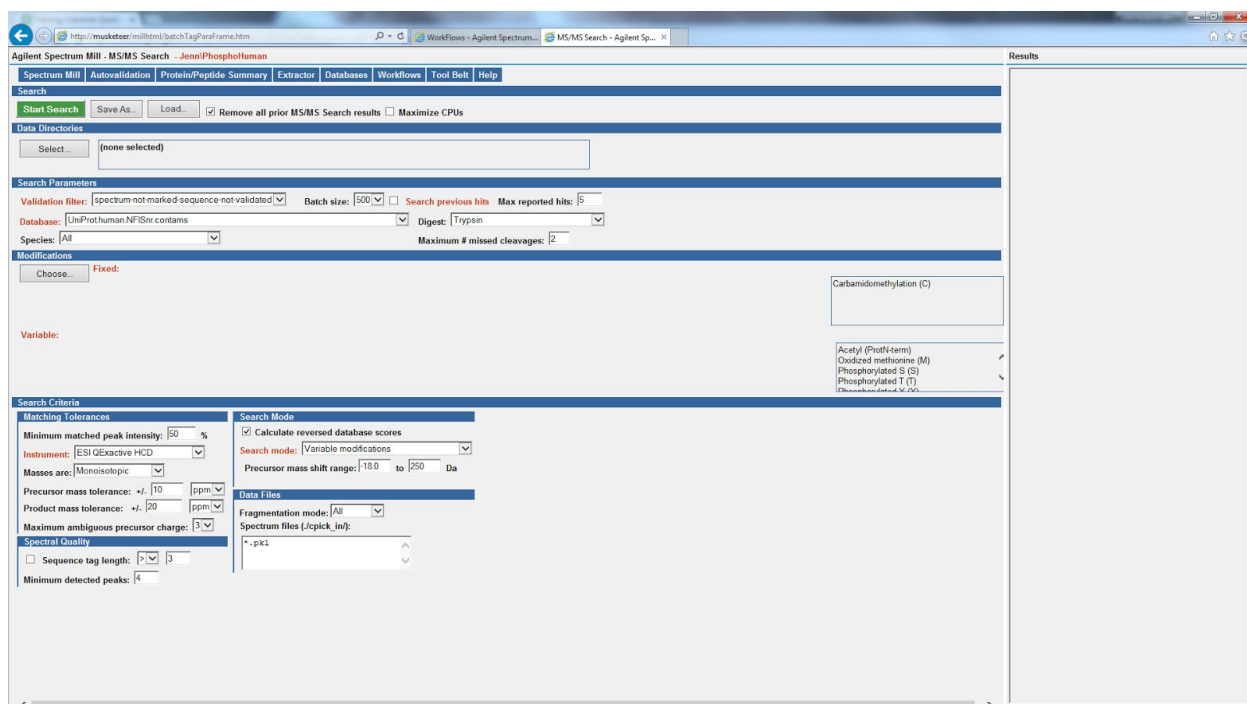
Sum the intensity of both reports

Divide phosphopeptide intensity by total intensity to get % phosphopeptides

DDA Search Workflow Parameters: IN THIS ORDER



The screenshot shows the 'Data Extractor' window of the Agilent Spectrum Mill software. The interface includes a menu bar with 'Spectrum Mill', 'MS/MS Search', 'PMF Search', 'Peak Picker', 'Workflows', 'Tool Belt', and 'Help'. Below the menu is an 'Extraction' section with buttons for 'Extract', 'Save As...', and 'Load...', and checkboxes for 'Remove all prior results', 'Delete data files after extraction', and 'Create PKL files instead of mzXML'. A 'Data Directories' section has a 'Select...' button and a text box containing '(none selected)'. The 'Modifications' section has a 'Choose...' button and a 'Fixed:' field containing 'Carbamidomethylation (C)'. The 'MS/MS Spectral Feature Filtering' section includes input fields for 'Precursor MH+' (400 to 4000.0 Da), 'Scan time range' (0 to 300 min), and 'Sequence tag length' (> -1). The 'Precursor m/z & Charge Assignment' section has a 'Centroiding algorithm for profile mode data' dropdown set to 'Xcalibur', 'Precursor Charge' set to 'Find', 'Maximum (z)' set to 5, and 'Minimum MS1 S/N' set to 25. A checkbox for 'Find ¹³C precursor m/z' is checked.



The screenshot shows the 'MS/MS Search' window of the Agilent Spectrum Mill software. The interface includes a menu bar with 'Spectrum Mill', 'AutovValidation', 'Protein/Peptide Summary', 'Extractor', 'Databases', 'Workflows', 'Tool Belt', and 'Help'. Below the menu is a 'Search' section with buttons for 'Start Search', 'Save As...', and 'Load...', and checkboxes for 'Remove all prior MS/MS Search results' and 'Maximize CPUs'. A 'Data Directories' section has a 'Select...' button and a text box containing '(none selected)'. The 'Search Parameters' section includes a 'Validation filter' dropdown set to 'Spectrum not marked-sequence not validated', 'Batch size' set to 500, 'Search previous hits' checkbox, and 'Max reported hits' set to 5. The 'Database' is set to 'UniProtHuman.NFISv8.contains', 'Digest' is set to 'Trypsin', and 'Species' is set to 'All'. The 'Maximum # missed cleavages' is set to 2. The 'Modifications' section has a 'Choose...' button and a 'Fixed:' field containing 'Carbamidomethylation (C)'. A 'Variable:' field contains a list of modifications: 'Acetyl (ProN-term)', 'Oxidized methionine (M)', 'Phosphorylated S (S)', 'Phosphorylated T (T)', and 'Phosphorylated Y (Y)'. The 'Search Criteria' section is divided into 'Matching Tolerances' and 'Search Mode'. 'Matching Tolerances' includes 'Minimum matched peak intensity' (50%), 'Instrument' (ESI QExactive HCD), 'Masses are' (Monoisotopic), 'Precursor mass tolerance' (+/- 10 ppm), 'Product mass tolerance' (+/- 20 ppm), and 'Maximum ambiguous precursor charge' (3). 'Search Mode' includes 'Calculate reversed database scores' (checked), 'Search mode' (Variable modifications), and 'Precursor mass shift range' (180 to 250 Da). The 'Data Files' section includes 'Fragmentation mode' (All) and 'Spectrum files (.cpick_in):' with a text box containing '*_pk1'.

Autovalidation of MS/MS Spectra - Agilent Spectrum Mill - Internet Explorer

Agilent Spectrum Mill - MS/MS Autovalidation - Karl\peptide_autoPpm_1_0

Help

Validate Files Queue request Undo Last Clear All Save As... Load...

Data Directories

Select...

(none selected)

Fragmentation mode: All

Search result files: *.spo

Validation Parameters

Strategy: Fixed thresholds Auto thresholds Auto thresholds - discriminant

Mode: Peptide Auto determine using score, delta R1-R2 thresholds to reach a target FDR

Optimize score & R1-R2 score thresholds with max FDR: 1.0 % across each: LC run Directory

Precursor charge range: 2 to 4 Min Sequence Length: 6

Required AAs: any Disallowed AAs: none

Filtering	Automatic variable range for each run	Fixed range for all runs
<input type="radio"/> None (ppm)	<input checked="" type="radio"/> Auto precursor mass error	<input type="radio"/> Fixed precursor mass error Low -1.0 High 30.0 ppm
<input checked="" type="radio"/> None (SC/pl)	<input type="radio"/> Auto SCX Solution Charge, pH3	<input type="radio"/> Fixed Solution Charge Low -2 High 6
	<input type="radio"/> Auto OGE/IEF peptide pl	<input type="radio"/> Fixed peptide pl Low 3.0 High 10.0

Agilent Spectrum Mill - Protein/Peptide Summary - Jana/peptide_labelfree

Spectrum Mill | Summary Settings | Autovalidation | MRM Selector | MS/MS Search | Spectrum Summary | Build TIC | Workflows | Tool Belt | Help

Summarize Results for Review

Summarize Save As... Load...

Queue request Excel AMRT export

Mode: Peptide

Filter to distinct peptides: Case insensitive

Data directories: Select...

(none selected)

Search result files:
*.apo

Search result files exclude:

Validation and Sorting

Filter results by: valid

Validation preset: none

Sort peptides by: Score

Filter peptides by: Score: > 0 % SP: > 30

Required AAs: any

Disallowed AAs: none

Accession #'s:

Review Fields

<input checked="" type="checkbox"/> Filename	<input checked="" type="checkbox"/> Sequence	<input type="checkbox"/> by map
<input checked="" type="checkbox"/> Score	<input checked="" type="checkbox"/> Rev Sequence	<input type="checkbox"/> Rank 2
<input checked="" type="checkbox"/> FDR (Discriminant)	<input checked="" type="checkbox"/> VML sequence	<input type="checkbox"/> Reporter Ratios: [TFOAQ 4]
<input checked="" type="checkbox"/> Fwd-Rev score	<input type="checkbox"/> Prec Av Ch ²	<input type="checkbox"/> Intensities
<input checked="" type="checkbox"/> Rank 1.2 score	<input checked="" type="checkbox"/> Ret time, width	Control: 117
<input checked="" type="checkbox"/> SPI (%)	<input checked="" type="checkbox"/> Precursor m/z	<input checked="" type="checkbox"/> MH ⁺
<input type="checkbox"/> Unmatched ions	<input checked="" type="checkbox"/> Delta mass	<input type="checkbox"/> Pepp pl
<input checked="" type="checkbox"/> Var mod sites	<input checked="" type="checkbox"/> Protein MW	<input type="checkbox"/> Prot pl
<input checked="" type="checkbox"/> VML score [rt]	<input checked="" type="checkbox"/> Species	<input checked="" type="checkbox"/> Modification names
<input checked="" type="checkbox"/> Solution charge	<input checked="" type="checkbox"/> Accession #	<input checked="" type="checkbox"/> Xterm <input type="checkbox"/> Caerm <input type="checkbox"/> Cysteines
<input checked="" type="checkbox"/> Start AA position	<input checked="" type="checkbox"/> Protein name	<input checked="" type="checkbox"/> Fragmentation mode
		<input checked="" type="checkbox"/> Max tag length <input type="checkbox"/> Longest tag
		<input checked="" type="checkbox"/> # Backbone Cleavages
		<input type="checkbox"/> # by pairs
		<input type="checkbox"/> Category