



# Byos<sup>®</sup> User's Manual

December 2025

5.12

Protein Metrics LLC, Boston, Massachusetts, USA

# Contents

Introduction.....	1
Byos Default System Workflows .....	1
Windows Support .....	3
System Requirements .....	3
Data Flow.....	3
Byos Menus .....	4
File .....	4
Server .....	4
Help .....	5
Project Creation .....	6
Project Creation Actions .....	7
Launch Workflow .....	7
Save Workflow .....	7
Save as Portable Workflow.....	7
Workflow Properties.....	7
Project Creation Tabs .....	8
Samples .....	8
Sequences .....	14
Processing nodes .....	22
Customizing System Workflows.....	22
Intact, Reduced, ADC, and icIEF-MS Workflows .....	22
Sequences and masses Tab .....	22
Sample-protein input Tab .....	30
Processing nodes .....	30
PTM, PTM (in-silico), PTM-DIA, HCP, S-S, SVA (C57) – Specific, SVA (C58) – Specific, HotSpot, System Suitability, Oxidative Footprinting, and MAM New Peak Detection .....	34
Processing nodes .....	34
HDX Workflow .....	55
Comparison Chromatography, Comparison Chromatography (in-silico), Reference Chromatography, Reference Chromatography (in-silico) Workflows .....	55
Processing nodes .....	56
Multi-Protein Quantitation, Multi-Protein Identification, Multi-Protein Quantitation .....	58
Applications.....	58
Multi-Protein Quantitation .....	58
Processing nodes .....	60
Multi-Protein Identification .....	64
Processing Nodes .....	65
Multi-Protein Quantitation Data-Independent Acquisition (DIA) .....	68

Processing nodes .....	68
Multi-Protein Preview .....	72
Processing nodes .....	73
Licensing Requirements .....	76
Supported Data .....	76
Released Glycans (N-Linked, O-Linked, and IgG) Workflows .....	76
Glycan Table .....	77
Glycan Options .....	78
Processing nodes .....	79
De novo Sequencing Workflow .....	80
Processing nodes .....	80
Oligonucleotide Analysis .....	83
MOBILion High Resolution Ion-Mobility (HRIM) Analysis .....	83
Charge Variant Analysis with Reconstruction .....	84
Byos Preview Workflow .....	91
My Workflows .....	91
Saving a Customized Workflow or Portable Workflow .....	91
Save as Portable Workflow .....	92
Populate "My Workflows" .....	92
Create Project to Generate an Analysis and Report .....	92
Byos UNIFI API Integration .....	93
Customized Workflows .....	94
Appendix .....	94
Byos: Advanced Commands .....	94

# Introduction

## Byos Default System Workflows

This manual will focus on the Byos<sup>®</sup> software program, which is based on the legacy products Byonic<sup>™</sup>, Byologic<sup>®</sup>, Byomap<sup>™</sup>, Intact Mass<sup>™</sup>, and Supernovo<sup>™</sup>. Byos is now the default application to open \*.byrslt, \*.blgc, \*.ntms, and \*.bmap project files and standalone apps (excluding Preview and Byonic) are no longer supported. Byos currently includes several default System Workflows, as shown in the figure below.

System Workflows


 <b>Intact</b> Identification, characterization, and relative quantification of large molecules	 <b>Native Intact</b> Identification, characterization, and relative quantification of non-denatured molecules	 <b>Reduced</b> Identification, characterization, and relative quantification of reduced large molecules	 <b>Isotope Resolved</b> Identification, characterization, and relative quantification of smaller molecules with i...	 <b>ADC</b> Drug loading and mass confirmation with multi-sample comparison	 <b>Intact Reconstruction</b> Intact Protein Analysis combining peptide level analysis for mass reconstruction	 <b>icIEF-MS</b> Annotation of icIEF peak ID with MS characterization	 <b>Oligo</b> Identification, annotation of Oligonucleotide molecules
 <b>Digested Oligonucleotides</b> Enzymatically Digested Oligonucleotides	 <b>PTM</b> Identification and quantification of modifications	 <b>PTM (pure in-silico)</b> Identification and quantification of modifications	 <b>PTM-DIA</b> Identification and quantification of modifications from DIA data	 <b>HCP</b> Identification and quantification of unknown/unexpected host cell proteins	 <b>S-S</b> Identification and quantification of disulfide bonded, trisulfide bonded, and cross-linked peptides	 <b>SVA (C57) - Specific</b> Identification, validation, and quantification of sequence variants	 <b>SVA (C58) - Specific</b> Identification, validation, and quantification of sequence variants
 <b>Hotspot</b> Identification and quantification of hotspot modifications	 <b>System Suitability</b> Assessment and verification of LC-MS/MS system	 <b>Oxidative Footprinting</b> Identification and quantification of sites of oxidation	 <b>HDX</b> Analyze hydrogen-deuterium exchange experiment	 <b>MAM New Peak Detection</b> Identification of reference peaks and new peak detection based on LC-...	 <b>Multi-Protein Quantitation</b> Identification and Quantification of Protein Mixtures	 <b>Multi-Protein Identification</b> Proteomics ID, characterization and annotation of PTMs	 <b>Multi-Protein Preview</b> Preview MS2 Data Quality and Search Parameters
 <b>Comparison Chromatography</b> Quantitative alignment and comparison of many chromatograms against ...	 <b>Comparison Chromatography (in-silico)</b> Quantitative alignment and comparison of many...	 <b>Reference Chromatography</b> Annotation of peak IDs for a TIC, LC-UV, or cIEF trace based on MS or MS/MS I...	 <b>Reference Chromatography (in-silico)</b> Annotation of peak IDs for a TIC, LC-UV, or cIEF tra...	 <b>Charge Variant Analysis with Reconstruction</b> Charge variant analysis combining peptide level analysis for charge reco...	 <b>Released Glycan (N-linked, neg mode)</b> Identification and quantification of released N-linked glycans	 <b>Released Glycan (O-linked, neg mode)</b> Identification and quantification of released O-linked glycans	 <b>Released Glycan (IgG, neg mode)</b> Identification and quantification of released glycans
 <b>de novo sequencing</b> de novo sequencing of monoclonal antibodies	 <b>HRIM Intact</b> Identification, characterization, and relative quantification of large molecules with MQ...	 <b>HRIM Peptide</b> Identification and quantification of protein modifications with MOBILion HRIM-MS	 <b>HRIM Glycan</b> Identification and quantification of released N/O-linked glycans with MOBILion HRIM-MS	 <b>Byonic</b> Byonic only - Search and identification of modifications	 <b>Preview</b> Preview for Byonic Identification of Peptides and PTMs		

Figure 1: The Byos default System Workflows included with installation.



In this manual, the default System Workflows are separated into eleven sections:

1. **Intact Analysis** – Intact, Native Intact, Reduced, ADC, Intact Reconstruction, and icIEF-MS workflows
2. **Peptide Analysis** – PTM, PTM (in-silico), PTM-DIA, HCP, S-S, SVA (C57) – Specific, SVA (C58) – Specific, HotSpot, System Suitability, Oxidative Footprinting, and MAM New Peak Detection
3. **HDX Analysis** – HDX workflow
4. **Chromatogram Analysis** – Comparison Chromatography, Comparison Chromatography (in-silico), Reference Chromatography, and Reference Chromatography (in-silico) workflows
5. **Multi-Protein Analysis** – Multi-Protein Quantitation, Multi-Protein Identification, and Multi-Protein Preview workflows
6. **Released Glycan Analysis** – Released Glycan (N-linked), Released Glycan (O-linked), and Released Glycan (IgG) workflows
7. **De novo Sequencing Analysis** – De novo Sequencing workflow
8. **Oligonucleotide Analysis** – Oligo and Digested Oligonucleotides workflow
9. **MOBILion High-Resolution Ion-Mobility (HRIM) Analysis** – HRIM Intact, HRIM Peptide and HRIM Glycan workflows
10. **Charge Variant Analysis with Reconstruction**
11. **Preview-like Analysis**

Note that this document outlines project creation for individual workflows, but more in-depth documentation for most analysis types and their associated Byos workflows can be found in the installer.

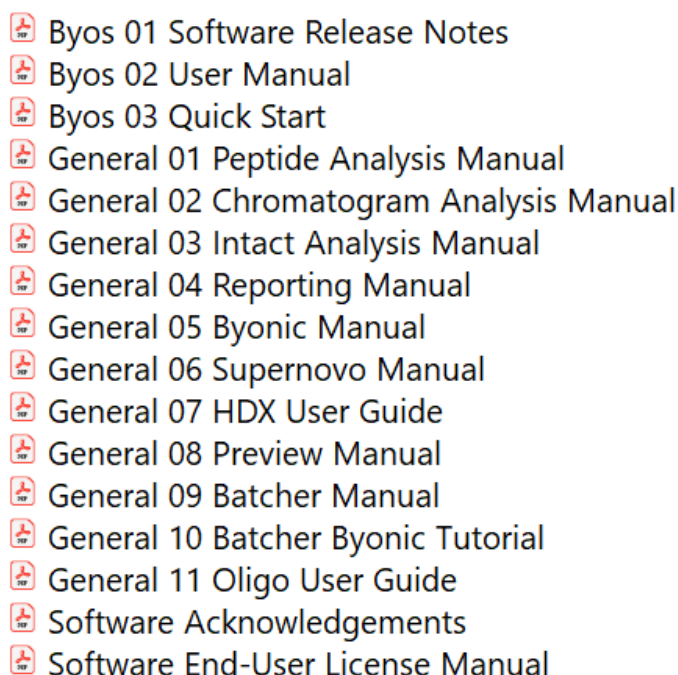
- 
- Byos 01 Software Release Notes
  - Byos 02 User Manual
  - Byos 03 Quick Start
  - General 01 Peptide Analysis Manual
  - General 02 Chromatogram Analysis Manual
  - General 03 Intact Analysis Manual
  - General 04 Reporting Manual
  - General 05 Byonic Manual
  - General 06 Supernovo Manual
  - General 07 HDX User Guide
  - General 08 Preview Manual
  - General 09 Batcher Manual
  - General 10 Batcher Byonic Tutorial
  - General 11 Oligo User Guide
  - Software Acknowledgements
  - Software End-User License Manual

Figure 2: Manuals and guides available in the installer

## Windows Support

Byos Desktop is currently supported on Windows 10 and Windows 11. **Note that for Windows 10, you must have version 1809 or later.**

## System Requirements

### Recommended PC:

- Windows 10/11 64-bit
- 32 GB RAM
- 1TB disk space (Solid State SSD)
- Recent version of Intel Core i7 or i9 / AMD Ryzen 7 or 9 (with AVX support)
- Oracle JRE or OpenJDK
- C++ compiler version 16 or higher

### Recommend PC for *high performance computing* (e.g. 32+ cores)

- Windows Server 2022 or Windows 10/11
- 64 GB RAM
- 2 TB disk space (Solid State SSD)
- Xeon CPU(s) (at least 16 physical cores) (with AVX support)
- Oracle JRE or OpenJDK
- C++ compiler version 16 or higher

Note: Installation of Byos on a server within Virtual Machines is not a supported configuration. Any technical issues experienced using Byos in a virtual machine will be the responsibility of the end user to resolve. Note the terms of the license to confirm compliance with any such setup.

## Data Flow

The user may select any default workflow from the System Workflows section or customized workflow from the **My Workflows** section. Default parameters for each workflow, including a report template, have been set in the Sequence and masses and Processing nodes tabs.

To create a project, the user will click on a workflow icon, drag and drop their raw data file(s) and a FASTA file, and click **Create Project**. A completed analysis and optimized report will then be available. This data flow is represented in [Figure 2](#) below.

A simplified diagram shows the data flow:

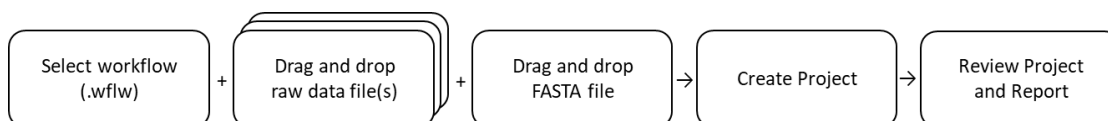


Figure 3: Simplified diagram of the Byos data flow.

# Byos Menus

## File

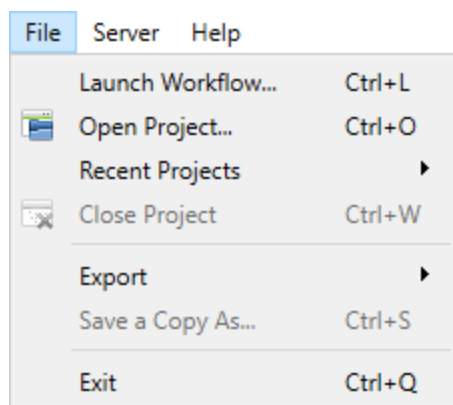


Figure 4: File menu options.

Menu	Sub Menu	Action and Notes
File	Launch Workflow	Launch a workflow to begin the creation of a project and its corresponding report. The user will be directed to the directory of default “System Workflows” available within the installer at: C:\Program Files\ProteinMetrics\PMI-Suite\Base\presets\Byos\workflows.
	Open Project	Open a Byos project that has been previously processed.
	Recent Projects	Select from recently created Byos projects.
	Close Project	Close the currently open Byos project.
	Export	Use this functionality to generate a Multi-Document Report. This feature allows the user to combine multiple Byos projects into one comprehensive report. Refer to the Reporting section for additional details.
	Save a Copy As	Save a copy of the currently open Byos project.
	Exit	Exit the Byos program.

## Server

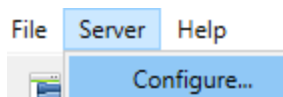


Figure 5: Server menu

Menu	Sub Menu	Action and Notes
Server	Configure	Configure allows Byos integration with third party servers.

Currently **Server > Configure** is supported for the Waters UNIFI server. See the **PMI UNIFI Integration User Guide** for instructions on how to configure UNIFI integration.

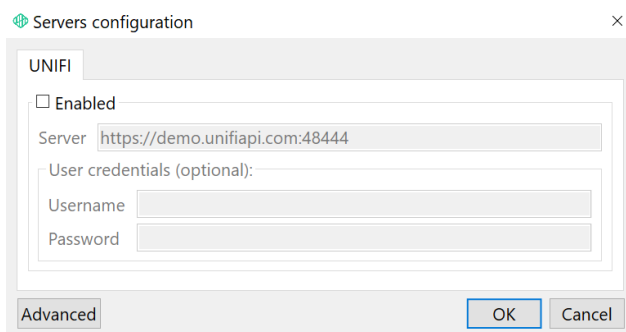


Figure 6: UNIFI Server configuration options

## Help

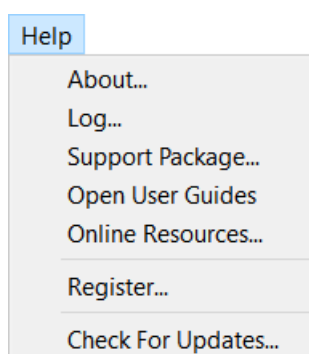
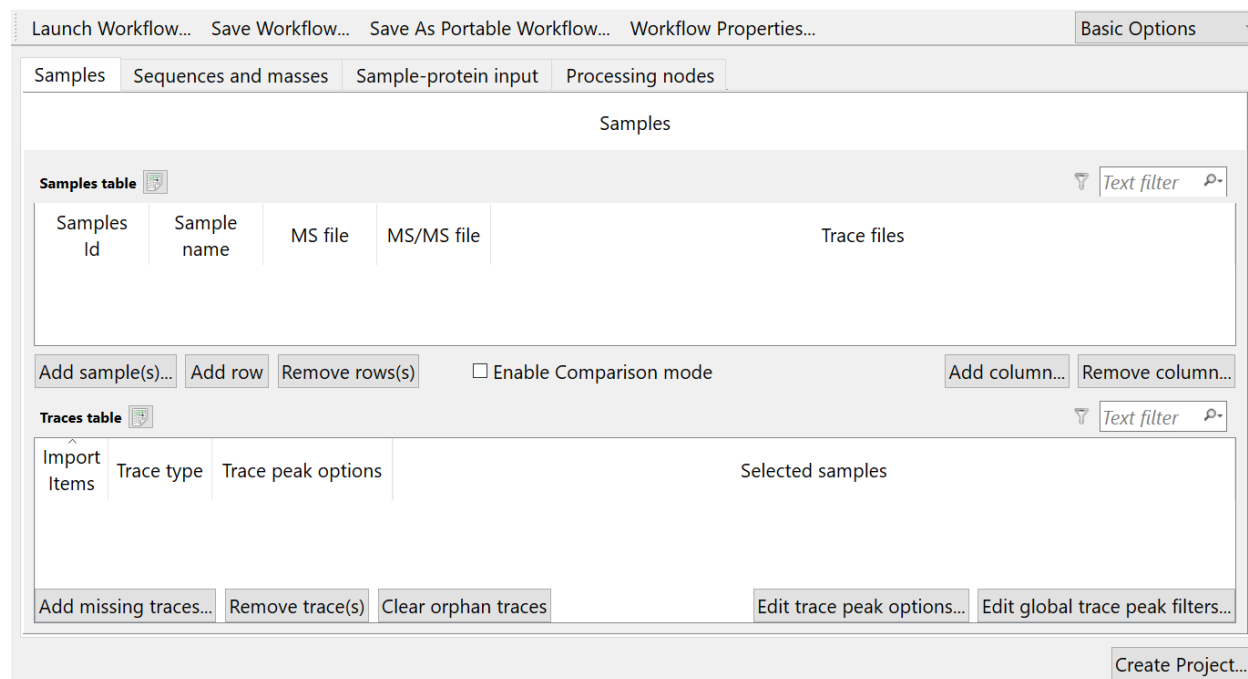


Figure 7: Help menu options.

Menu	Sub Menu	Actions and Notes
Help	About	Details on the version number, date, program path, copyright information, and patent details.
	Log	Displays the log file. The user is provided the options to “Open log folder”, “Copy All”, or “Clear”.
	Support Package	Collects all details necessary to assist a Support request. This file can be shared directly with <a href="mailto:support@proteinmetrics.com">support@proteinmetrics.com</a> .
	Open User Guides	Directs the user to the Manuals, QSGs, Release Notes, Licensing Details, FAQs, etc. included with the installation package. These documents are for Byos along with the Protein Metrics suite of legacy products.
	Online Resources	Directs the user to the online resources available at <a href="https://www.proteinmetrics.com/secure-resources/">https://www.proteinmetrics.com/secure-resources/</a> . If Support is up to date, the password will be automatically entered and the user granted access to all application notes, videos, posters, etc.
	Register	Directs the user to the License Activation window. Please refer to the licensing instructions pdf provided with purchase.
	Check For Updates	Directs the user to upgrade if a newer version is available.

## Project Creation




When a workflow is launched, the **Project Creation** window is activated. This is where the user directs Byos to create the project and corresponding report. The windows are shown below in Figure 8, Figure 9, Figure 10, **Error! Reference source not found.** and options detailed in the following Project Creation Actions section.



Launch Workflow... Save Workflow... Save As Portable Workflow... Workflow Properties... Basic Options




Samples Sequences and masses Sample-protein input Processing nodes

Samples

Samples table   Text filter 

Samples Id	Sample name	MS file	MS/MS file	Trace files
------------	-------------	---------	------------	-------------

Add sample(s)... Add row Remove rows(s) ☐ Enable Comparison mode Add column... Remove column...

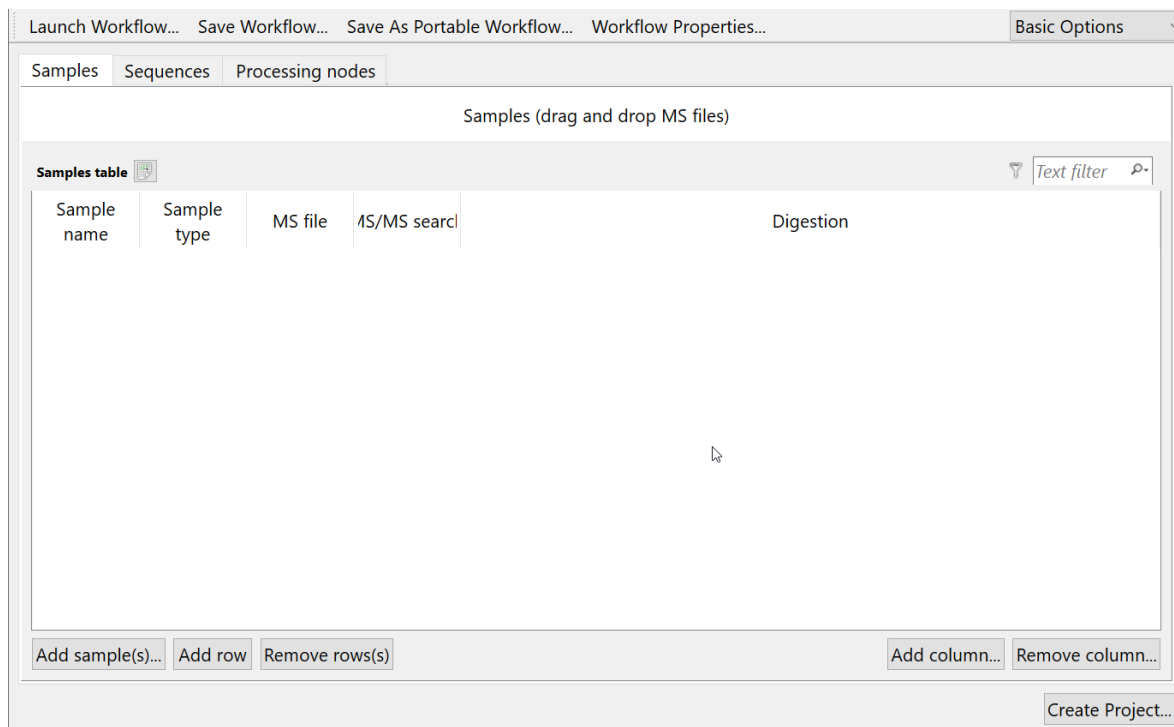
Traces table   Text filter 

Import Items	Trace type	Trace peak options	Selected samples
--------------	------------	--------------------	------------------

Add missing traces... Remove trace(s) Clear orphan traces Edit trace peak options... Edit global trace peak filters...

Create Project...




Figure 8: Intact Analysis project creation window.



Launch Workflow... Save Workflow... Save As Portable Workflow... Workflow Properties... Basic Options

Samples Sequences Processing nodes

Samples (drag and drop MS files)

Samples table   Text filter 

Sample name	Sample type	MS file	MS/MS search	Digestion
-------------	-------------	---------	--------------	-----------

Add sample(s)... Add row Remove rows(s) Add column... Remove column...

Create Project...

Figure 9: Peptide Analysis project creation window.

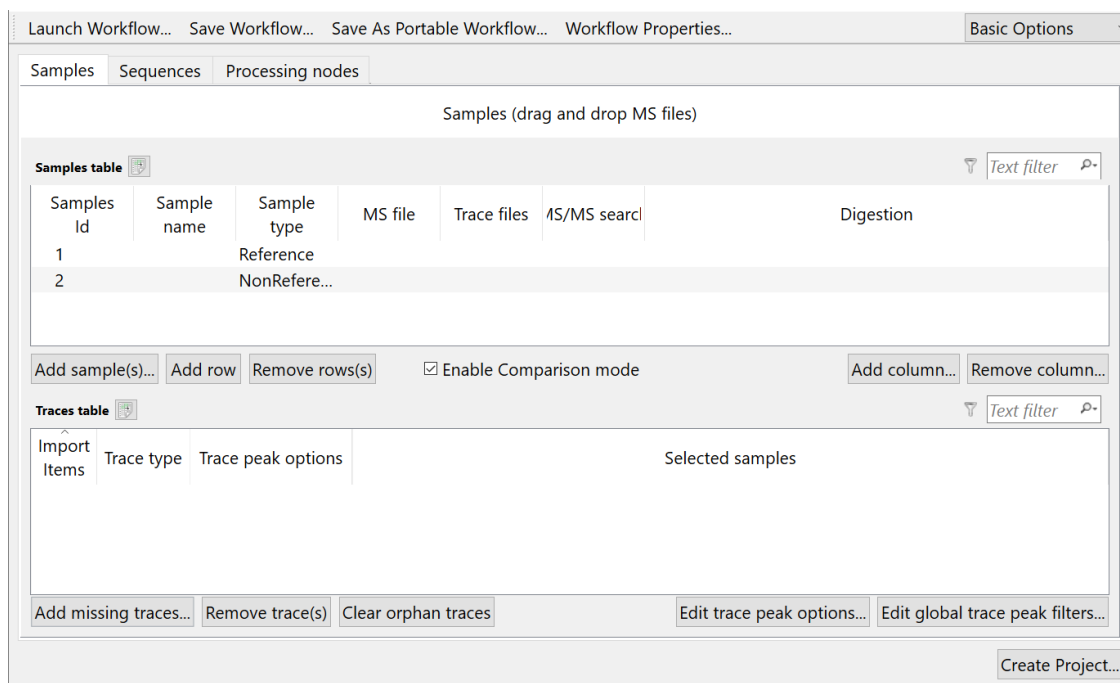


Figure 10: Chromatogram Analysis project creation window.

## Project Creation Actions

### Launch Workflow

This enables the user to open a different workflow from the one currently open.

### Save Workflow

This enables the user to save the current workflow (\*.wflw) file. This can be saved within a portable workflow (\*.wflwp) folder or as a stand-alone .wflw file. It can be saved within the “My Workflows” folder, detailed in the section [Saving a Customized Workflow or Portable Workflow](#).

### Save as Portable Workflow

This enables the user to move a workflow and all related components necessary for processing to another system. This process creates a folder that includes the workflow (.wflw), report template (.rptc), layout or filter (.ini), databases (.txt), etc. files. The user can then save changes to any \*.wflwp component within this folder. It can be saved within the “My Workflows” folder, detailed in the **My Workflows** section.

### Workflow Properties

This enables the user to edit the icon, name, and description of the workflow currently open. The icons are available to select from using a drop-down menu, as shown below:

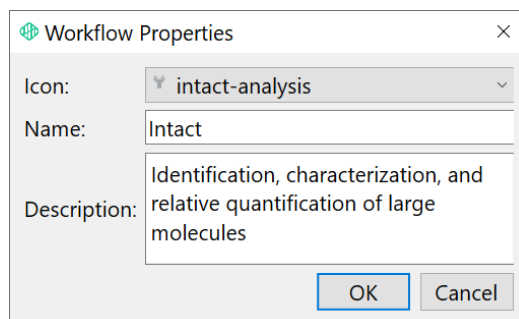


Figure 11: Workflow Properties – Icon, Name, and Description.

Please note each workflow has an icon, representative of the type of project being created. Byos also includes a “default” icon, that can be used if preferred. This is shown below:

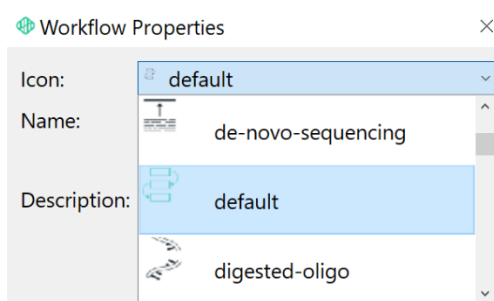


Figure 12: A default icon can also be assigned to the workflow.

## Project Creation Tabs

### Samples

In the Samples tab, the user drags and drops mass spectrometry raw data files to be processed within the project. The process is the same for all workflows. The PTM workflow is shown below:

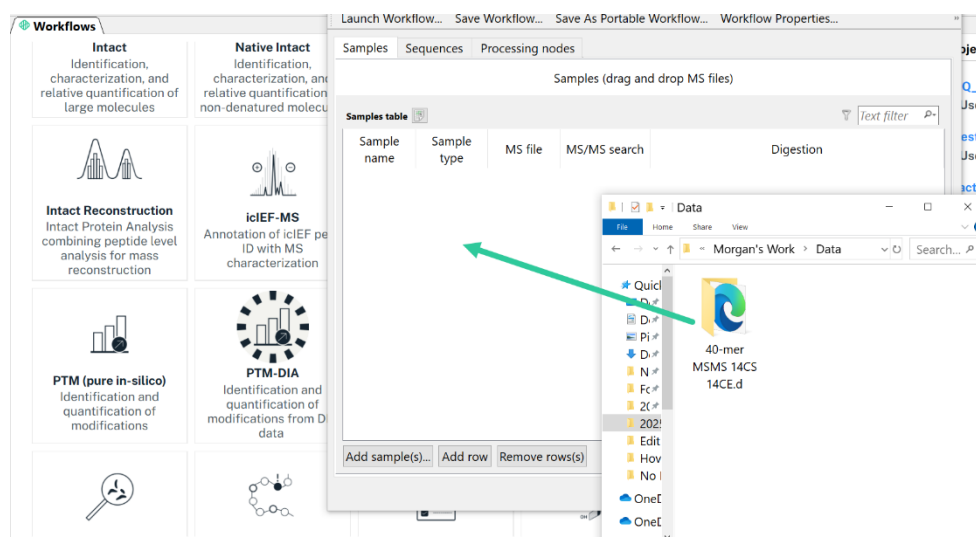


Figure 13: Drag and drop the raw data file(s) to be processed into the Samples tab.

The user can also use the **Add row** button to manually add a sample and then direct to the raw data file. **Remove sample(s)** will remove the highlighted line.

The user can use **Add column** to add an additional column to the default settings available. These include the options shown below:

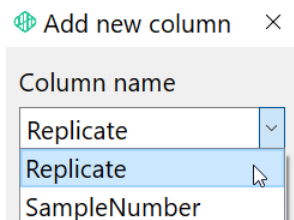
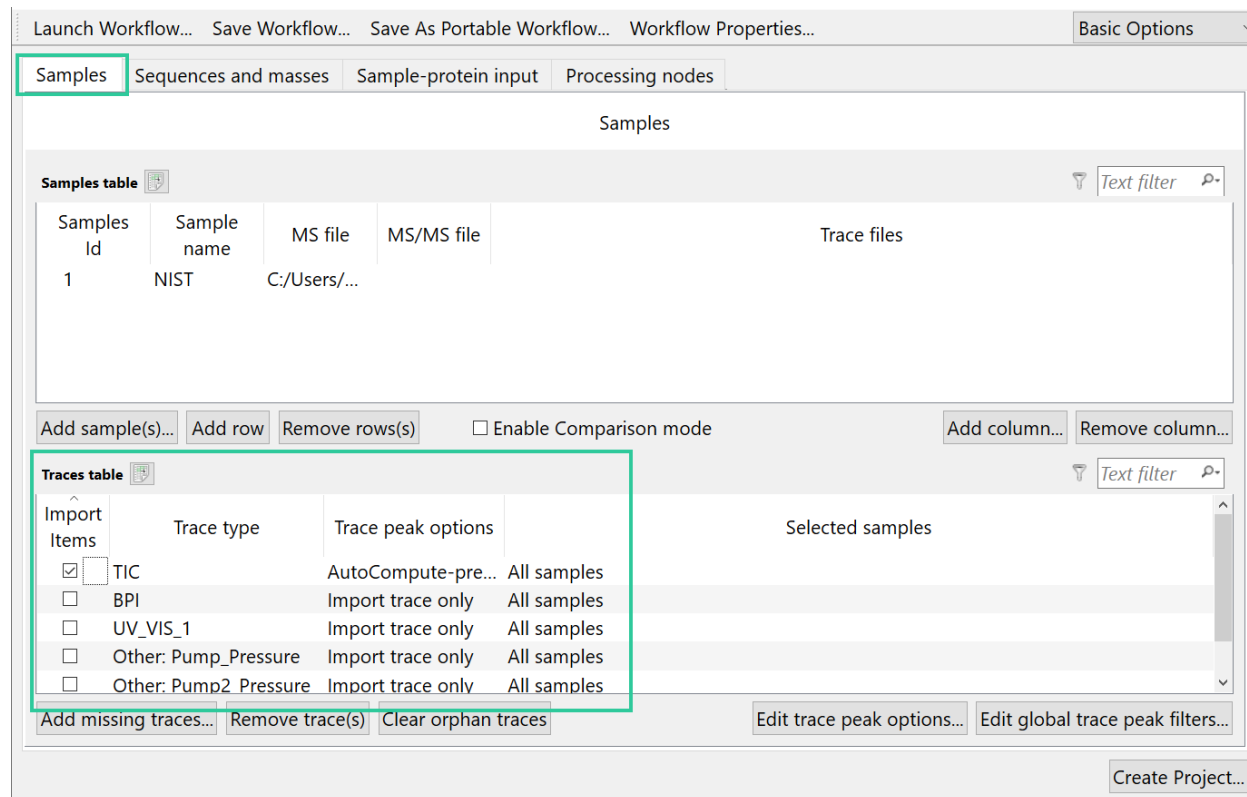


Figure 14: Add new column option.

The user can also create any custom names by typing in this dialog. The user can remove any column using the **Remove column** button.

For Intact Analysis and Chromatogram Analysis workflows, the MS file populates both the **Samples table** and the new **Traces table** at the bottom.



Launch Workflow... Save Workflow... Save As Portable Workflow... Workflow Properties... Basic Options

Samples Sequences and masses Sample-protein input Processing nodes

Samples

Samples table

Samples Id	Sample name	MS file	MS/MS file	Trace files
1	NIST	C:/Users/...		


Add sample(s)... Add row Remove rows(s) ☐ Enable Comparison mode Add column... Remove column...

Traces table

Import Items	Trace type	Trace peak options	Selected samples
<input checked="" type="checkbox"/> TIC		AutoCompute-pre...	All samples
<input type="checkbox"/> BPI		Import trace only	All samples
<input type="checkbox"/> UV_VIS_1		Import trace only	All samples
<input type="checkbox"/> Other: Pump_Pressure		Import trace only	All samples
<input type="checkbox"/> Other: Pump2 Pressure		Import trace only	All samples

Add missing traces... Remove trace(s) Clear orphan traces Edit trace peak options... Edit global trace peak filters... Create Project...

Figure 15: Samples tab with populated Traces table

Selected (not checked) traces can be removed by clicking **Remove trace(s)**. Removed traces can be restored by clicking **Add missing traces**, check the trace to restore and click **OK**. To associate specific traces to the samples, double-click the text in the Selected samples column, then click the  button to open the Edit samples selection dialog:



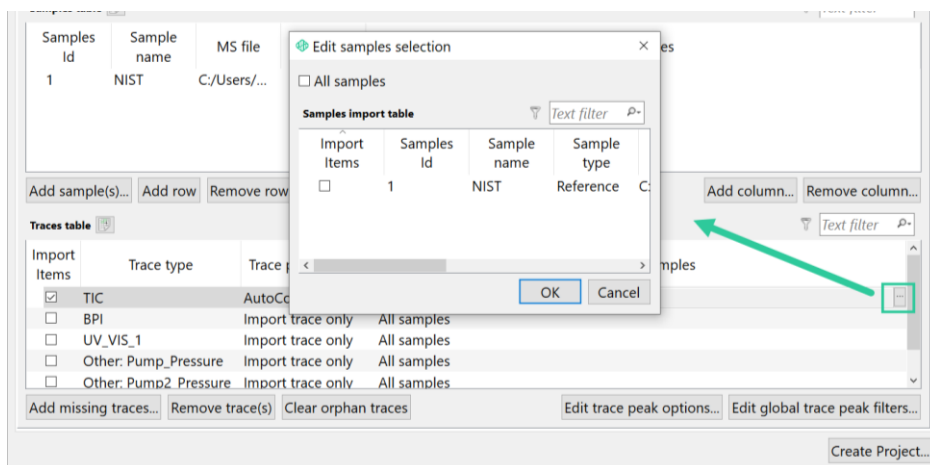


Figure 16: Edit samples selection dialog to associate traces with sample names

To associate the selected trace with individual samples, uncheck **All samples**, check one sample, and click **OK** (only **All samples** or a single sample Id can be checked).

Individual traces can be imported from sample file types supported in the MS file cell, as well as \*.csv, \*.txt files, and IntaBio \*.itb files. To import a trace from a file, drag the file into the **Traces files** column in the Samples table. Alternatively, double-click in the Samples table row under the Traces file column, click **...** and browse to the sample file. Click **Open**. The trace file is added to the sample name and the associated traces are added to the Traces table:

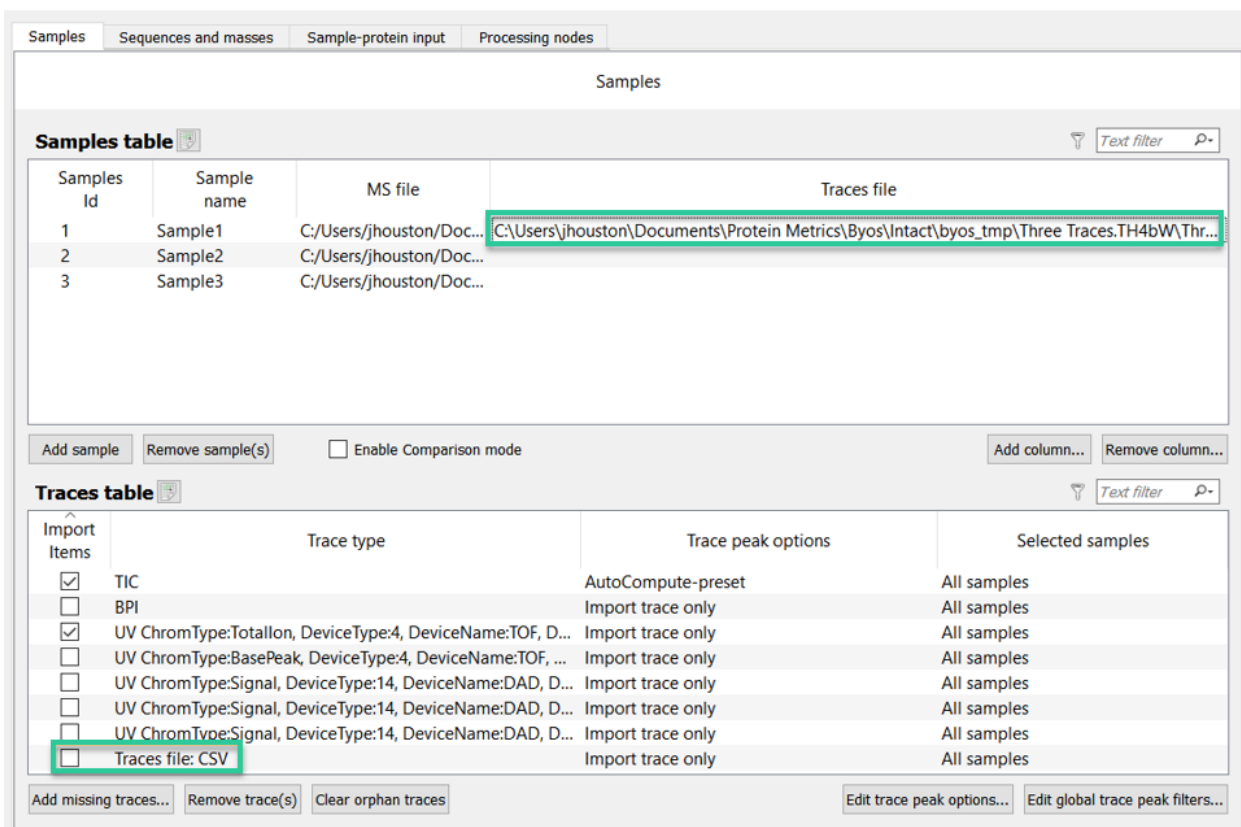


Figure 17: Traces imported from a \*.csv file

To control how the trace peaks are processed, click **Edit trace peak options** at the bottom:

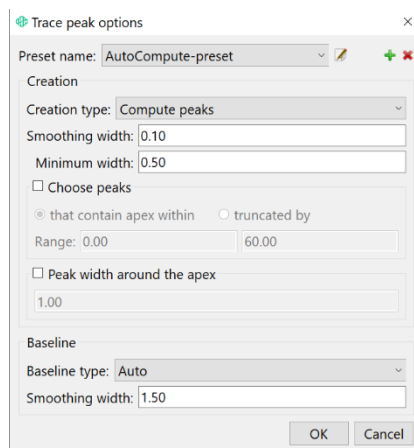


Figure 18: Trace peak options

At the bottom of the **Trace peak options presets** dialog are the **Baseline** parameters. The baseline type can be set to **Auto** or **Flat**. Baseline smoothing width is also set here. Larger smoothing factors will merge trace peaks while smaller values will split peaks.

The **Trace peak options presets** dialog contains five default peak processing presets:

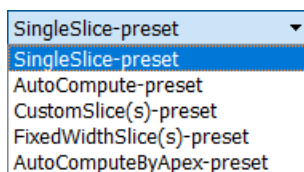


Figure 19: Default trace peak presets

Select a processing preset and edit the peak parameters. To create a custom preset, click the **+** button, add a preset name, and click **OK**. Next, select a creation type and edit those peak parameters. The new preset is added to the list. Click the **✎** button to edit a selected preset name. Click the **✖** button to delete a selected preset.

The default and custom peak processing presets can use one of four **Creation types**:

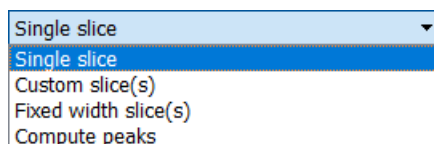
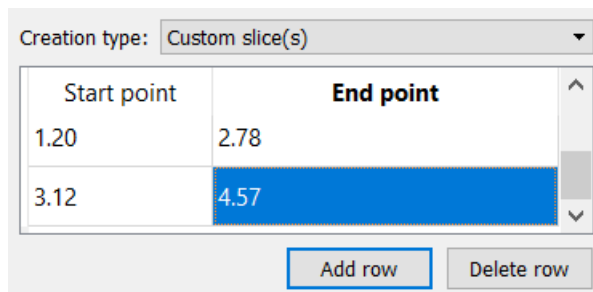


Figure 20: Trace peak Creation types

These Creation types use the following peak processing methods:

- **Single Slice** processes the entire chromatogram as single trace peak.
- **Custom slice(s)** divides trace peaks by user defined start and end points:



Creation type: Custom slice(s)

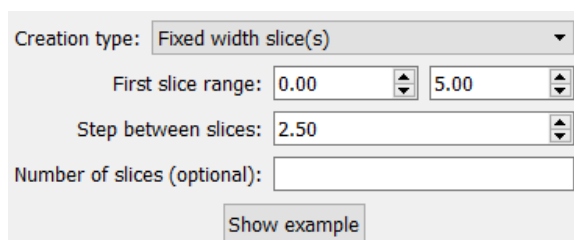
Start point	End point
1.20	2.78
3.12	4.57

Add row Delete row

Figure 21: Trace peak Custom slice(s) option

Click **Add row** and enter start and end points for each slice. To remove a slice, select it and click **Delete row**.

- **Fixed width slice(s)** divides trace peaks by regular intervals defined by the user:



Creation type: Fixed width slice(s)

First slice range: 0.00 5.00

Step between slices: 2.50

Number of slices (optional):

Show example

Figure 22: Trace peak Fixed width slice(s) option

- o **First slice range** sets the start and end points for the first slice. This also sets the width for all slices.
- o **Step between slices** is the delta that sets the start of subsequent fixed width slices by adding the step value to the start of the preceding slices. For example, the values above define the first slice as 0 - 5, the second slice as 2.5 – 7.5, etc., through the end of the chromatogram.
- o **Number of slices** sets the maximum number of defined slices. If blank, the last slice is determined by the end of the chromatogram x-axis range.
- o **Show example** displays an example fixed-width slice configuration with definitions and explanations:

Usage example

The slice depends on:

1. Input parameters
  - First slice range:** Sets the width for all slices as well as the start point and end point for the first slice.
  - Step between slices:** Add this value to the start point of each slice to get the next slice, for e.g. to get the second slice, we add 3 to the start point of the first slice. So the second slice starts at 4, the third slice starts at 7 (4+3) etc.
  - Number of slices (optional):** If specified, the number of slices determines the maximum end point of fixed width slices as detailed in Table 2.
2. Other parameters
  - Chromatogram range:** Actual start and end point of the chromatogram determined by the MS file.
  - Range of interest:** (set by Edit global elution peak filters) narrows the chromatogram range by dynamically filtering out peaks not overlapping with this range.

Assume that the chromatogram range is 1-11.

Close

Usage example

First slice range 1 3  
Step between slices 3

**Table 1**

Slice #	Start point	End point
1	1	3
2	4	6
3	7	9
4	10	12

Note that the last slice may extend beyond the chromatogram range. This does not affect computations (such as elution peak area) and will allow for consistent reporting across samples.

**Table 2**

Number of slices specified by user	Actual number of slices created	End point of last slice
None	4	12
3	3	9
7	4	12

Close

Figure 23: Fixed width slice(s) Usage example

**Compute peaks** automatically determines the trace peaks based on their minimum width, after smoothing:

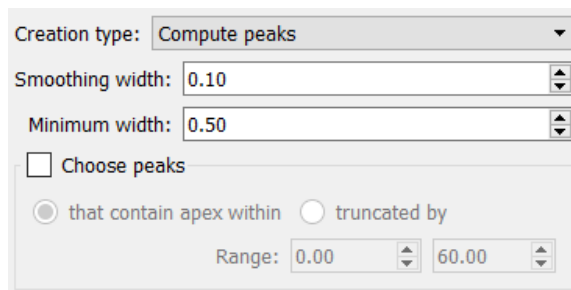


Figure 24: Trace peak Compute peaks option

- o Choose peaks that contain apex within, when selected, includes the detected peaks with apices in the specified x-axis range.
- o Choose peaks truncated by, when selected, ensures all peaks within the specified range will be chosen. Additionally, the first and last peak will be trimmed to fit within the specified range.

Individual trace filters are replaced with a single peak filter across all traces. To set peak property filters for all the traces, click **Edit Global trace peak filters**:

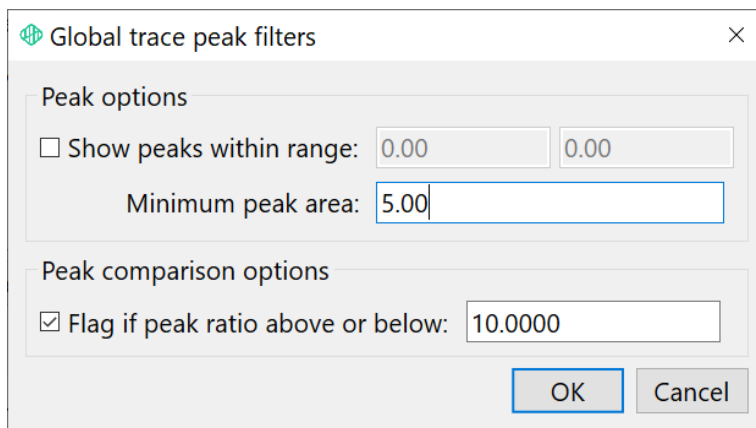


Figure 25: Global trace peak filters

Check **Show peaks within range** to manually set start and end points. **Minimum peak area** and a peak ratio threshold flag can be set in this dialog.

The Traces table and trace peak settings are backwards-compatible. Traces associated with samples in projects from software versions before 3.8 will correctly populate the Traces table in the current version. Trace peak settings and presets from versions before 3.8 will correctly migrate into the current version. However, Intact and Chromatogram Analysis projects created in versions 3.8 and above will not correctly open in older versions of Byos.

## Sequences

The Sequences tab is where the user enters protein sequences to be investigated. FASTA files can be dragged and dropped directly into the Sequences tab or click **Browse for FASTA file** and navigate to and select the file. The user will then be prompted to select the **Protein candidates** (heavy chain, light chain, etc.) to be considered for processing. The PTM workflow is shown below:

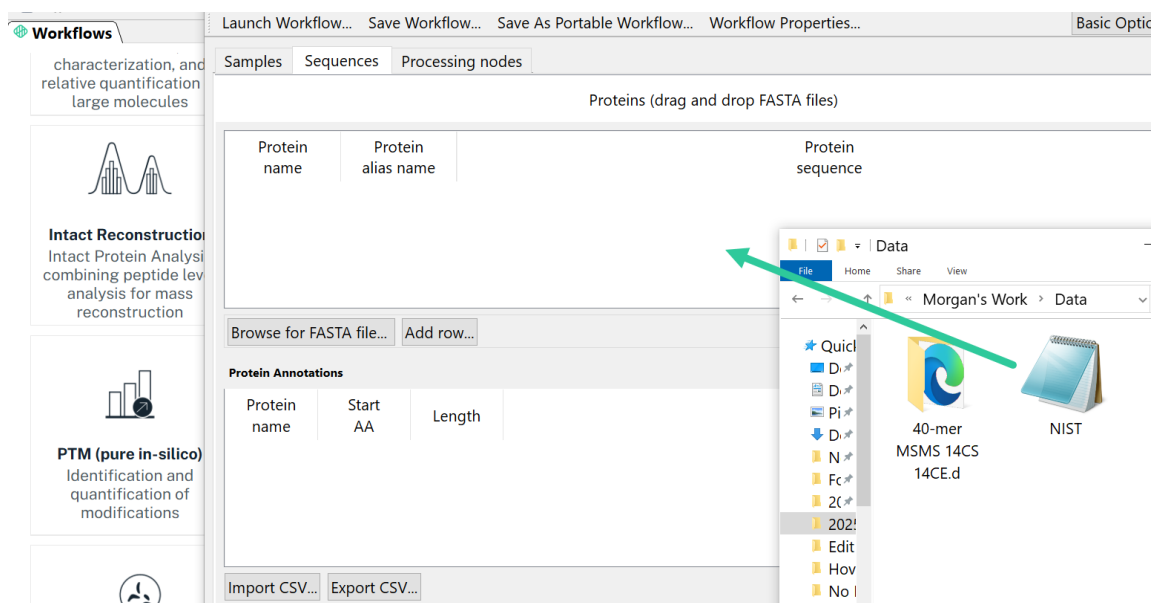
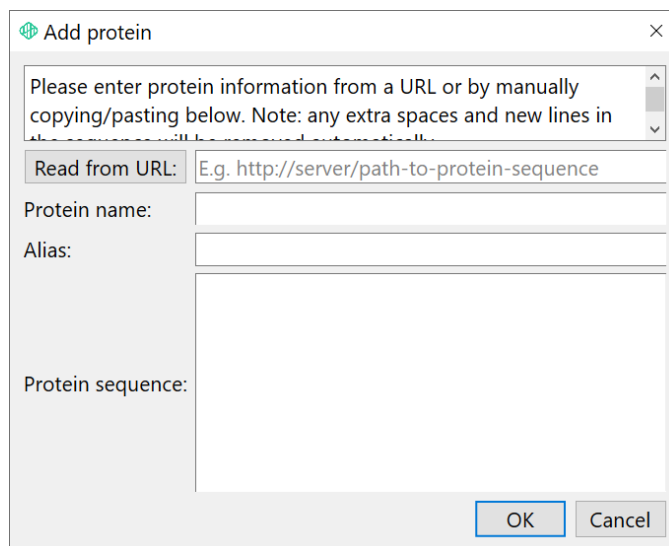


Figure 26: Drag and drop the FASTA file to be processed into the Sequences tab.

Alternatively, click **Add row** to enter a protein name and sequence directly:



The 'Add protein' dialog box is shown. It has a title bar with a close button. The main text says: 'Please enter protein information from a URL or by manually copying/pasting below. Note: any extra spaces and new lines in the sequence will be removed automatically.' Below this are four input fields: 'Read from URL:' (with a placeholder 'E.g. http://server/path-to-protein-sequence'), 'Protein name:', 'Alias:', and 'Protein sequence:'. The 'Protein sequence:' field is a large text area. At the bottom right are 'OK' and 'Cancel' buttons.

Figure 27: Add protein manually

Copy and paste the protein name and sequence or enter a website URL to load published protein data.

NOTE: In the HCP workflow, the user must instead point to a protein database. The user can either drag and drop as shown above or direct Byos to a database file. This is done by clicking in the space to activate the light blue "...". The user is then prompted to select a file, as shown:

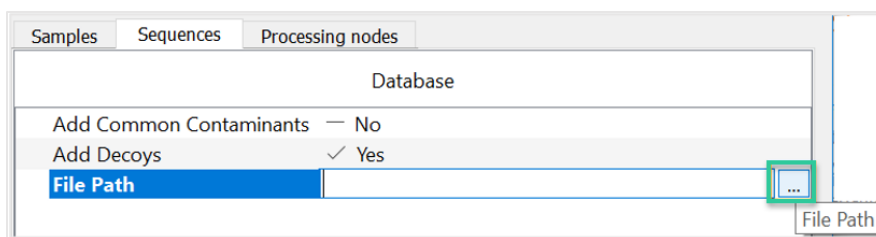


Figure 28: Click to activate the file path prompt.

NOTE: In the **System Suitability** workflow, the Sequences tab is populated by default to include the Pierce RT standard. This is shown below:

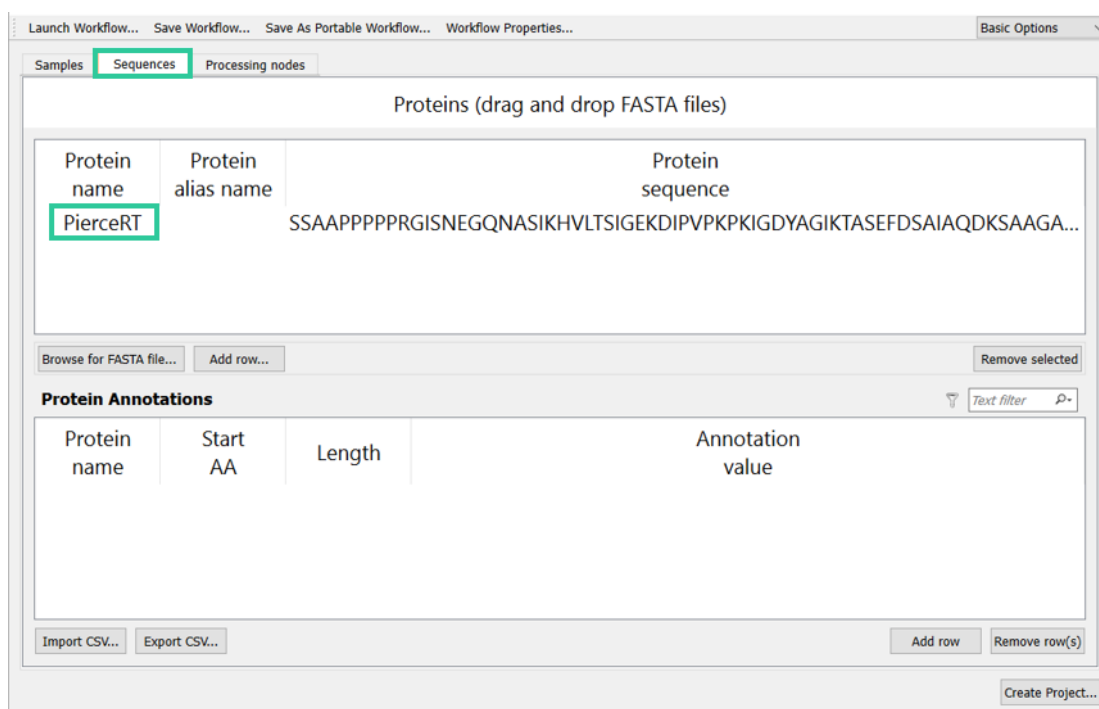


Figure 29: The Pierce RT standard is included in the Sequences tab of the System Suitability workflow.

The user can replace the Pierce RT standard with another protein if desired. The user can change this by clicking to highlight the line and click **Remove selected** (shown below). Another protein could be brought in using drag and drop (shown in Figure 26).

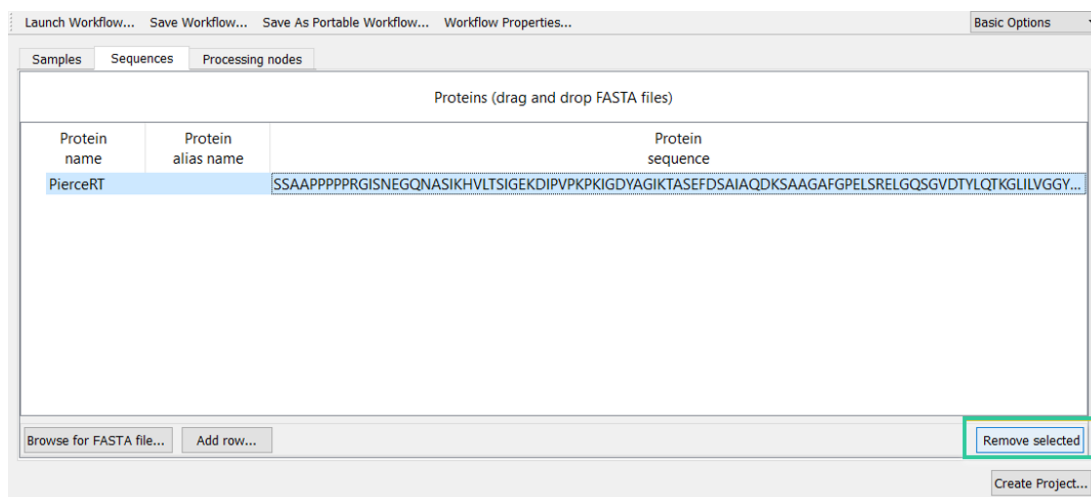


Figure 30: The user can change the default standard protein sequence.

Users can import **custom protein annotations** during Project Creation within the **Sequences** tab.

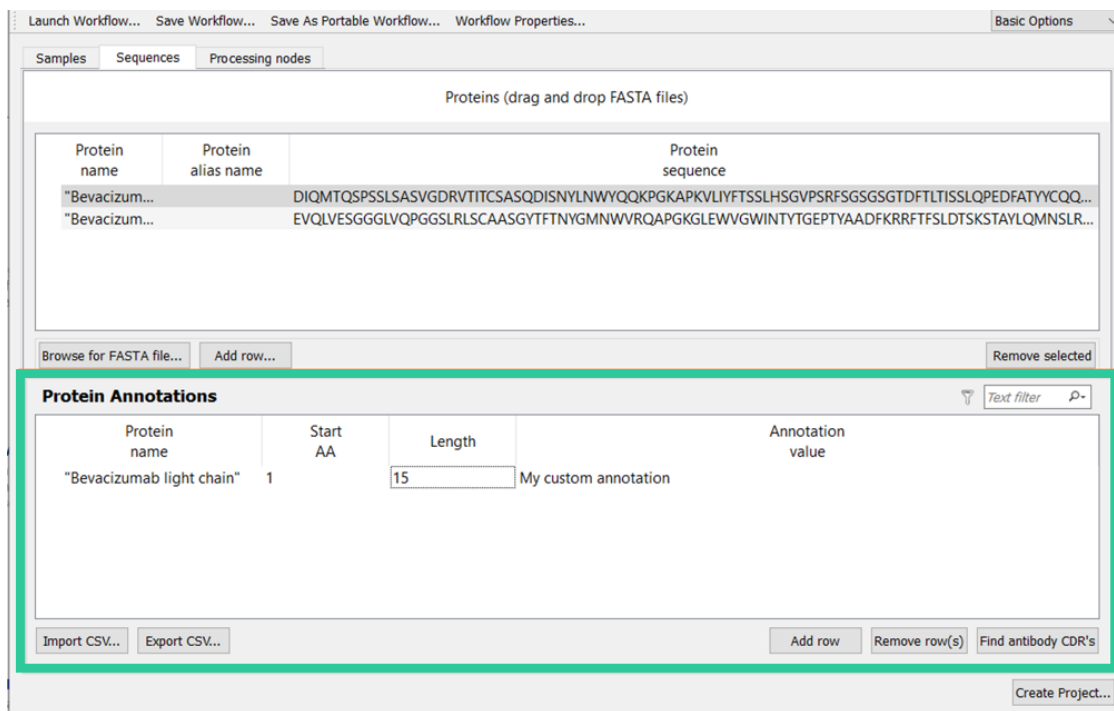


Figure 31: Protein Annotations

There is also an option to import or export protein annotations after Project Creation in the **Protein Annotations** table. This feature has been implemented in a backward compatible fashion so that users can import annotations from CSV file created in previous version of application.



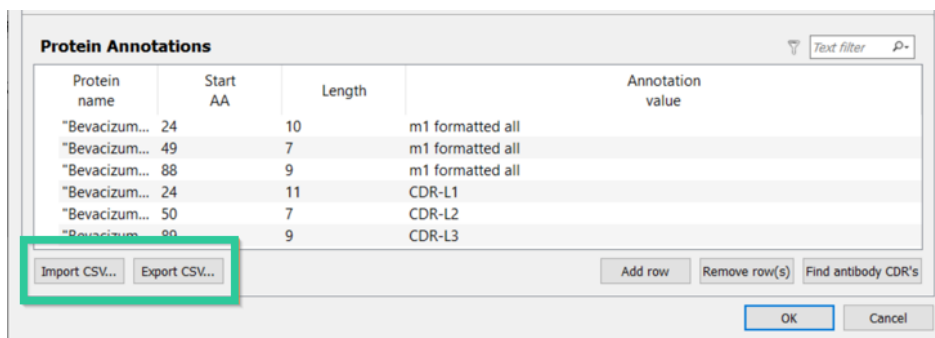


Figure 32: Import CSV for Protein Annotations

In the Intact, Reduced, and ADC workflows, this tab is named **Sequences and masses**. The user can add chains to the **Chains** table and based on these chains, add rows to the **Sequence combinations** table. The user enters a protein sequence either by dragging and dropping a FASTA file (shown in Figure 33 below), manually typing it in (shown in the Figure below), or selecting from a FASTA file (shown in Figure 35).

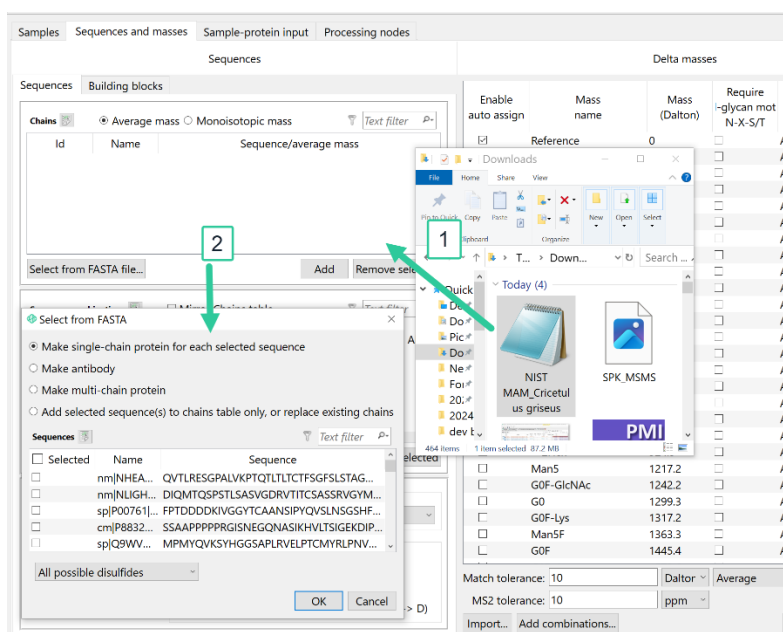
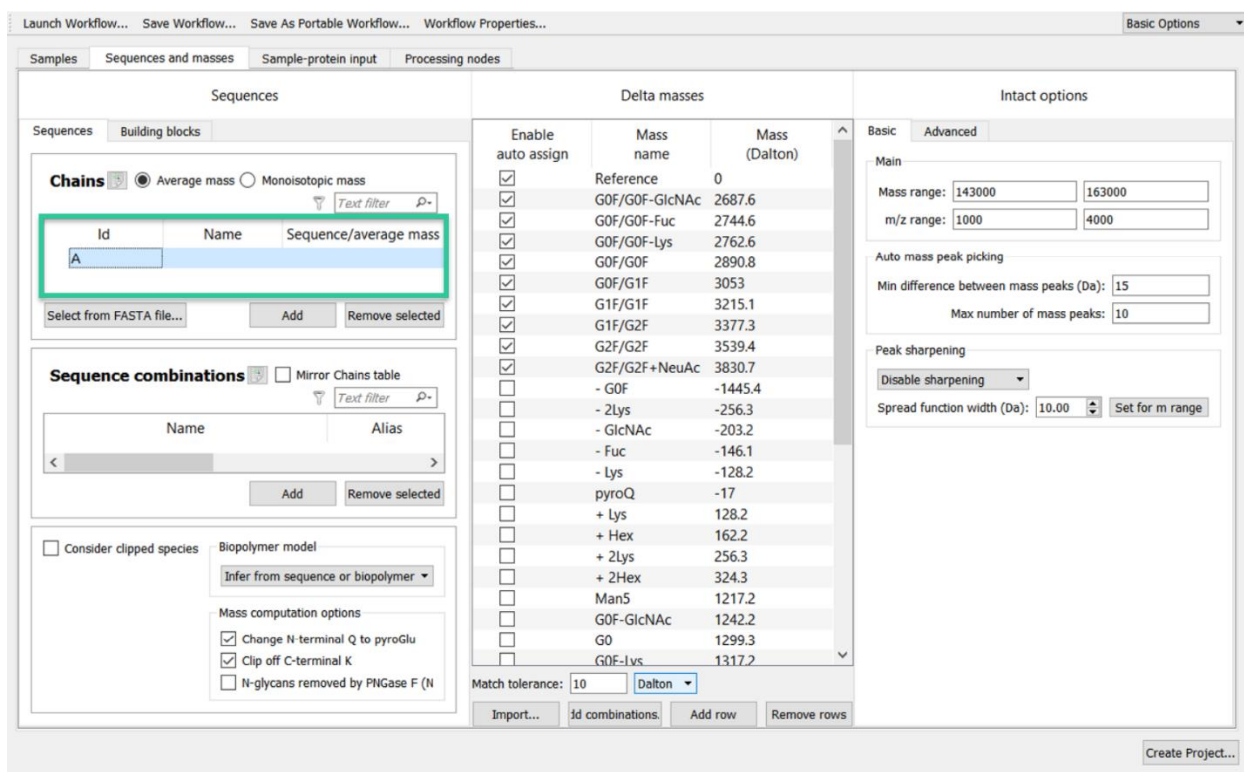


Figure 33: Drag and drop the FASTA file to be processed into the Sequences and masses tab.



Launch Workflow... Save Workflow... Save As Portable Workflow... Workflow Properties... Basic Options

Samples Sequences and masses Sample-protein input Processing nodes

Sequences Building blocks

Chains ☒ Average mass ☐ Monoisotopic mass

Id Name Sequence/average mass

A

Select from FASTA file... Add Remove selected

Sequence combinations ☐ Mirror Chains table

Name Alias

Add Remove selected

☐ Consider clipped species Biopolymer model

Infer from sequence or biopolymer

Mass computation options

☒ Change N-terminal Q to pyroGlu

☒ Clip off C-terminal K

☐ N-glycans removed by PNGase F (N)

Enable auto assign

Mass name Mass (Dalton)

Reference 0

G0F/G0F-GlcNAc 2687.6

G0F/G0F-Fuc 2744.6

G0F/G0F-Lys 2762.6

G0F/G0F 2890.8

G0F/G1F 3053

G1F/G1F 3215.1

G1F/G2F 3377.3

G2F/G2F 3539.4

G2F/G2F+NeuAc 3830.7

- G0F -1445.4

- 2Lys -256.3

- GlcNAc -203.2

- Fuc -146.1

- Lys -128.2

pyroQ -17

+ Lys 128.2

+ Hex 162.2

+ 2Lys 256.3

+ 2Hex 324.3

Man5 1217.2

G0F-GlcNAc 1242.2

G0 1299.3

G0F-Lys 1317.2

Match tolerance: 10 Dalton

Import... Id combinations Add row Remove rows

Intact options Basic Advanced

Main

Mass range: 143000 163000

m/z range: 1000 4000

Auto mass peak picking

Min difference between mass peaks (Da): 15

Max number of mass peaks: 10

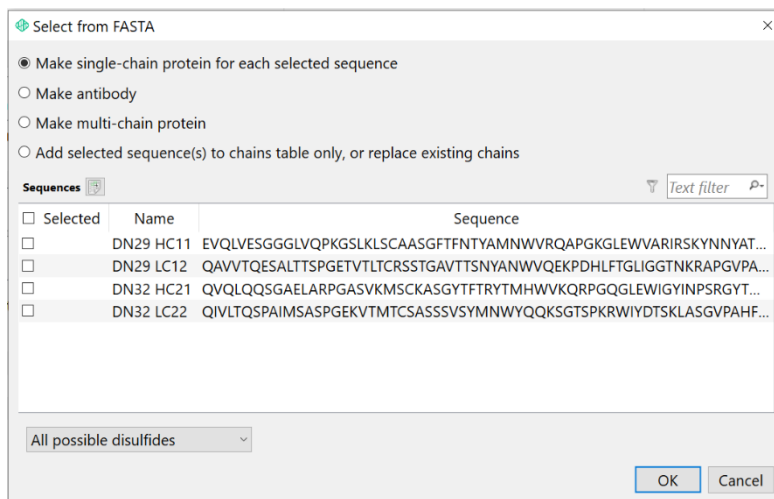
Peak sharpening

Disable sharpening

Spread function width (Da): 10.00 Set for m range

Create Project...

Figure 34: Click **Add** to add a row and manually enter the protein data (Id, Name, and Sequence/mass).



Select from FASTA

☒ Make single-chain protein for each selected sequence

☐ Make antibody

☐ Make multi-chain protein

☐ Add selected sequence(s) to chains table only, or replace existing chains

Sequences

Selected Name Sequence

☐ DN29 HC11 EVQLVESGGGLVQPKGSLKLSCAASGFTFTNYAMNWWVRQAPGKGLEWVARIRSKYNNYAT...

☐ DN29 LC12 QAVVTQESALTTSPGETVTLTCRSSTGAVTTSNYANWVQEKPDHLFTGLIGGTNKRAPGVPA...

☐ DN32 HC21 QVQLQSGAELARPGASVKMSCKASGYTFTRYTMHWVKRPGQGLEGWIGYINPSRGYT...

☐ DN32 LC22 QIVLTQSPAIMASPGKVTMTCSASSSVSYMNWYQQSGTSPKRWIYDTSKLASGVPAHF...

All possible disulfides

OK Cancel

Figure 35: Select from FASTA. Select the appropriate protein and disulfide option.

There are four options for FASTA sequences that result in different kinds of sequence combinations

- **Make single-chain protein for each selected sequence** generates a protein in the Sequence combinations table for each of the checked sequences:

**Chains** ☒ Average mass ☐ Monoisotopic mass Text filter

Id	Name	Sequence/average mass
A	Heavy.Chain_SS_mAb2	SEVQLVQSGGGVERPGGSLRLSCAASGFTFDDYAMSWVRQAPGKGLEWVSGINWQGGSTGYADSVKGRVTISRDNAKNSLYLQMNSLRAEDTAVYYCAKILGAGRGWYFDYWGKGT...
B	Light.Chain_SS_mAb1	SELTQDPAVSVLGGTQVTRITCSGDSLRSYYASWYQQKPGQAPVLIYGANNNRPSGIPDRFSGSSSGNTASLTITGAQAEADYYCNSADSSGNHVVFGGKTLTLVGGPKAAPSVTLFPP...

Select from FASTA file... Add Remove selected

---

**Sequence combinations** ☐ Mirror Chains table Text filter

Samples Id	Name	Alias	Composition	Disulfides	Average mass	Mono mass	Expected type
All samples	Heavy.Chain_SS_mAb2		A(1)	All possible disulfides	49115.99	49085.34	Desired
All samples	Light.Chain_SS_mAb1		B(1)	All possible disulfides	22309.43	22295.83	Desired

Add Remove selected

Figure 36: A single protein is generated per chain

- **Make antibody** generates an antibody from two copies of a checked heavy chain sequence and two copies of a checked light chain sequence:

**Chains** ☒ Average mass ☐ Monoisotopic mass Text filter

Id	Name	Sequence/average mass
A	Heavy.Chain_SS_mAb2	SEVQLVQSGGGVERPGGSLRLSCAASGFTFDDYAMSWVRQAPGKGLEWVSGINWQGGSTGYADSVKGRVTISRDNAKNSLYLQMNSLRAEDTAVYYCAKILGAGRGWYFDYWGKGT...
B	Light.Chain_SS_mAb1	SELTQDPAVSVLGGTQVTRITCSGDSLRSYYASWYQQKPGQAPVLIYGANNNRPSGIPDRFSGSSSGNTASLTITGAQAEADYYCNSADSSGNHVVFGGKTLTLVGGPKAAPSVTLFPP...

Select from FASTA file... Add Remove selected

---

**Sequence combinations** ☐ Mirror Chains table Text filter

Samples Id	Name	Alias	Composition	Disulfides	Average mass	Mono mass	Expected type
All samples	Antibody		A(2)B(2)	All possible disulfides	142846.82	142758.31	Desired

Add Remove selected

Figure 37: An antibody is generated from two heavy and light chains

- **Make multi-chain protein** generates a protein from the designated counts of the checked sequences:

Select from FASTA

☐ Make single-chain protein for each selected sequence  
☐ Make antibody  
☒ Make multi-chain protein  
☐ Add selected sequence(s) to chains table only, or replace existing chains

**Sequences** Text filter

Count	Name	Sequence
1	Be...	DIQMTQSPSSLSASVGRVITITCSASQDISNYLNWYQQKPGKAPKVLIIYFTSSLHSGVPSRFSGSGSGTDFLTISLQPEDFATYYCQ...
2	Be...	EVQLVESGGGLVQPGGSLRLSCAASGYFTFTNYGMNWRQAPGKGLEWVGVWINTYTGEPTYAADFKRRFTSLDTSKSTAYLQMNS...

All possible disulfides

OK Cancel

Figure 38: FASTA - Make multi-chain protein

**Chains** ☒ Average mass ☐ Monoisotopic mass Text filter

Id	Name	Sequence/average mass
A	Heavy.Chain_SS_mAb2	SEVQLVQSGGGVERPGGSLRLSCAASGFTDDYAMSWVRQAPGKLEWVSGINWQGGSTGYADSVKGRVTSRDNAKNSLYLQMNSLRAEDTAVYYCAKILGAGRGWYFDYWGKG...
B	Light.Chain_SS_mAb1	SELTQDPAVSVALGQTVRITCSGDSLRSYYASWYQQKPGQAPVLIYIGANNRPSGIPDRFSGSSSGNTASLTITGAQAEADADYYCNSADSSGNHVVFGGGLKTLVGLGQPKAAPSVTLFPP...

Select from FASTA file... Add Remove selected

---

**Sequence combinations** ☐ Mirror Chains table Text filter

Samples Id	Name	Alias	Composition	Disulfides	Average mass	Mono mass	Expected type
All samples	Multi-chain protein		A(1)B(2)	All possible disulfides	93732.84	93674.99	Desired

Figure 39: A protein is generated from the counts of checked sequences

- **Add selected sequence(s) to chains table only, or replace existing chains** performs two functions: first, when importing into an empty Chains table, the option loads the Chains table with the checked sequences without generating any sequence combinations:

**Chains** ☒ Average mass ☐ Monoisotopic mass Text filter

Id	Name	Sequence/average mass
A	Heavy.Chain_SS_mAb2	SEVQLVQSGGGVERPGGSLRLSCAASGFTDDYAMSWVRQAPGKLEWVSGINWQGGSTGYADSVKGRVTSRDNAKNSLYLQMNSLRAEDTAVYYCAKILGAGRGWYFDYWGKG...
B	Light.Chain_SS_mAb1	SELTQDPAVSVALGQTVRITCSGDSLRSYYASWYQQKPGQAPVLIYIGANNRPSGIPDRFSGSSSGNTASLTITGAQAEADADYYCNSADSSGNHVVFGGGLKTLVGLGQPKAAPSVTLFPP...

Select from FASTA file... Add Remove selected

---

**Sequence combinations** ☐ Mirror Chains table Text filter

Samples Id	Name	Alias	Composition	Disulfides	Average mass	Mono mass	Expected type
------------	------	-------	-------------	------------	--------------	-----------	---------------

Add Remove selected

Figure 40: No Sequence combinations are generated for this option

Second, this option can also be used to replace existing sequences added using any of the four entry options. To replace existing sequences, click **Select from FASTA file** again and select the original or a new FASTA file. Select the **Add selected sequence(s) to chains table only, or replace existing chains** radio option. Choose the rows containing the new sequences by checking the boxes in the **Selected** column. Use the dropdowns in the **Replace Chain Id** column to select the Ids of the sequences to replace (by default, capital letters):

**Select from FASTA**

☐ Make single-chain protein for each selected sequence  
☐ Make antibody  
☐ Make multi-chain protein  
☒ Add selected sequence(s) to chains table only, or replace existing chains

If adding sequences, leave Chain Id blank; if replacing sequences, choose the Chain Id that needs to be replaced.

**Sequences** Text filter

<input type="checkbox"/> Selected	Replace Chain Id	Name	Sequence
<input type="checkbox"/>		DN29 HC11	EVQLVESGGGLVQPKGSLKLSCAASGFTNTYAMNWWVRQAPGKLEWVAR...
<input type="checkbox"/>		DN29 LC12	QAVVTQESALTSPGETVTLTCRSSTGAVTTSNYANWVQEKPDHLFTGLIGG...
<input type="checkbox"/>	A	DN32 HC21	QVQLQSGAELARPGASVKMSCKASGYTFRTYTMHWVKQRPGQGLEWIL...
<input type="checkbox"/>	B	DN32 LC22	QIVLTQSPAIMASAPGEKVTMTCSASSSVSYMNWYQQKSGTSPKRWIYDTS...
	C		
	D		

OK Cancel

Figure 41: Replacing existing sequences from the original or a new FASTA file

## Processing nodes

Processing nodes is where the user specifies the processing parameters to be applied during project creation and report generation. Each workflow is populated with default values, designed for the user to review once and save for future projects. This empowers every member of a team, scientist within a lab, or colleagues across a project to complete the exact same analysis and generate identical reports regardless of Byos user-skill level, scientific expertise or experience, or location. Please note the next section will only have to be completed once for each analysis type and the workflow used for perpetuity.

## Customizing System Workflows

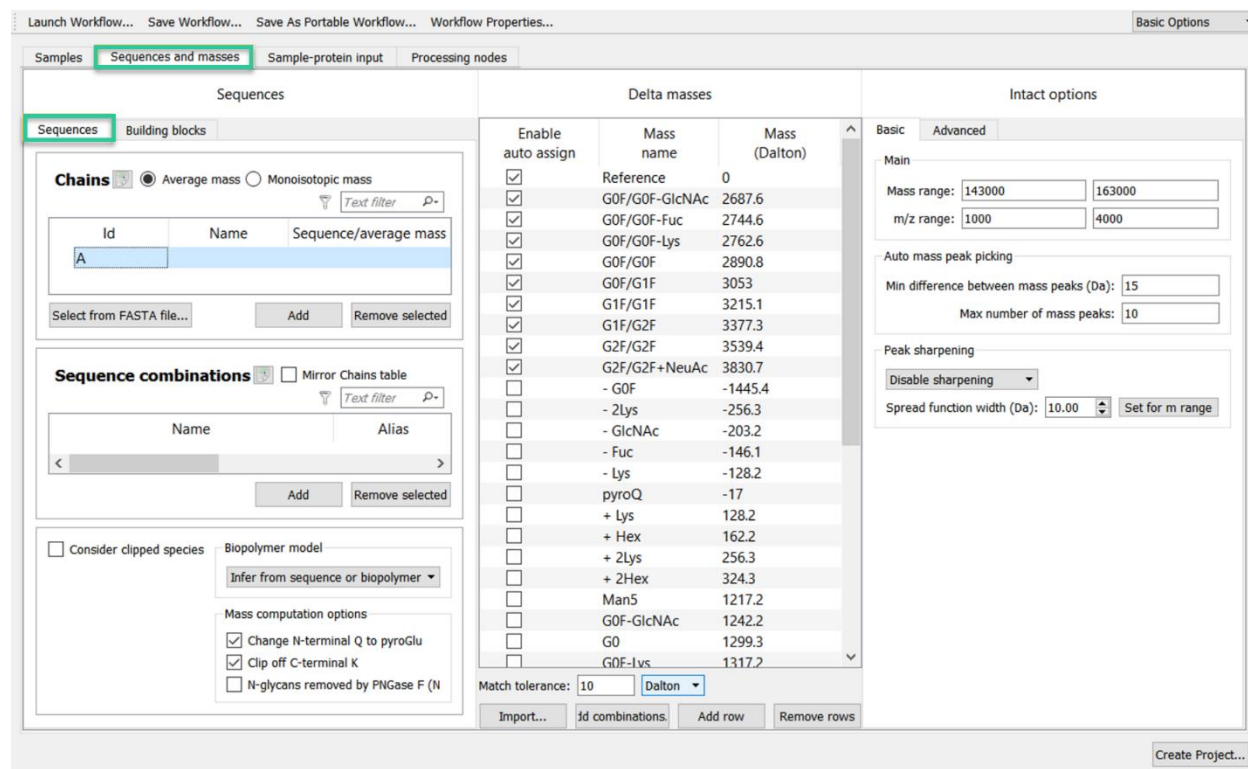
The user should review the parameters populated within the default set of System Workflows to generate customized My Workflows.

## Intact, Reduced, ADC, and icIEF-MS Workflows

The following section is relevant to analyses based upon the **Intact**, **Reduced**, **ADC**, and **IntaBio icIEF-MS** workflows. These workflows all have the Samples, Proteins and masses, Sample-protein input and Processing nodes tabs. The default parameters to review are included within the **Sequences and masses** and **Processing nodes** tabs.

## Sequences and masses Tab

### Sequences Sub-tab



Launch Workflow... Save Workflow... Save As Portable Workflow... Workflow Properties... Basic Options

Samples **Sequences and masses** Sample-protein input Processing nodes

Sequences

Sequences Building blocks

**Chains** ☒ Average mass ☐ Monoisotopic mass

Id	Name	Sequence/average mass
A		

Select from FASTA file...

**Sequence combinations** ☐ Mirror Chains table

Name	Alias
<	>

☐ Consider clipped species Biopolymer model

Mass computation options

☒ Change N-terminal Q to pyroGlu

☒ Clip off C-terminal K

☐ N-glycans removed by PNGase F (N)

**Delta masses**

Enable auto assign	Mass name	Mass (Dalton)
<input checked="" type="checkbox"/>	Reference	0
<input checked="" type="checkbox"/>	G0F/G0F-GlcNAc	2687.6
<input checked="" type="checkbox"/>	G0F/G0F-Fuc	2744.6
<input checked="" type="checkbox"/>	G0F/G0F-Lys	2762.6
<input checked="" type="checkbox"/>	G0F/G0F	2890.8
<input checked="" type="checkbox"/>	G0F/G1F	3053
<input checked="" type="checkbox"/>	G1F/G1F	3215.1
<input checked="" type="checkbox"/>	G1F/G2F	3377.3
<input checked="" type="checkbox"/>	G2F/G2F	3539.4
<input checked="" type="checkbox"/>	G2F/G2F+NeuAc	3830.7
<input type="checkbox"/>	- G0F	-1445.4
<input type="checkbox"/>	- 2Lys	-256.3
<input type="checkbox"/>	- GlcNAc	-203.2
<input type="checkbox"/>	- Fuc	-146.1
<input type="checkbox"/>	- Lys	-128.2
<input type="checkbox"/>	pyroQ	-17
<input type="checkbox"/>	+ Lys	128.2
<input type="checkbox"/>	+ Hex	162.2
<input type="checkbox"/>	+ 2Lys	256.3
<input type="checkbox"/>	+ 2Hex	324.3
<input type="checkbox"/>	Man5	1217.2
<input type="checkbox"/>	G0F-GlcNAc	1242.2
<input type="checkbox"/>	G0	1299.3
<input type="checkbox"/>	G0F-Lys	1317.2

Match tolerance: 10 Dalton

**Intact options**

Basic Advanced

Main

Mass range:

m/z range:

Auto mass peak picking

Min difference between mass peaks (Da):

Max number of mass peaks:

Peak sharpening

Spread function width (Da):

Figure 42: Sequences sub-tab.

The user can add chains to the Chains table, and based on these chains add rows to the Proteins table. The user enters a protein sequence either by dragging and dropping a FASTA file (shown in Figure 33), manually typing it in (shown in Figure 34), or selecting from a FASTA file (shown in Figure 35).

Users can designate peptide chain masses without sequences as average mass or monoisotopic mass. In the Chains section, a radio button selects between the two types of masses and the mass header updates accordingly:

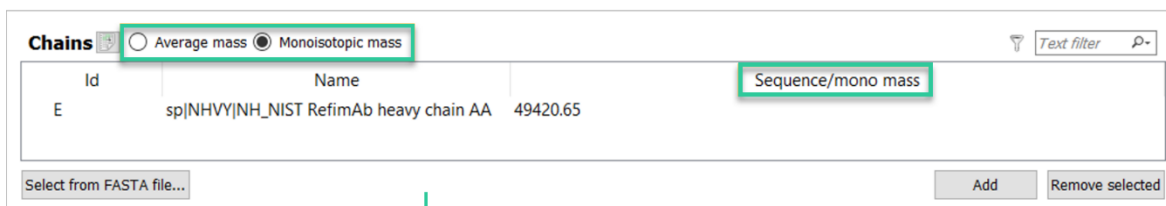
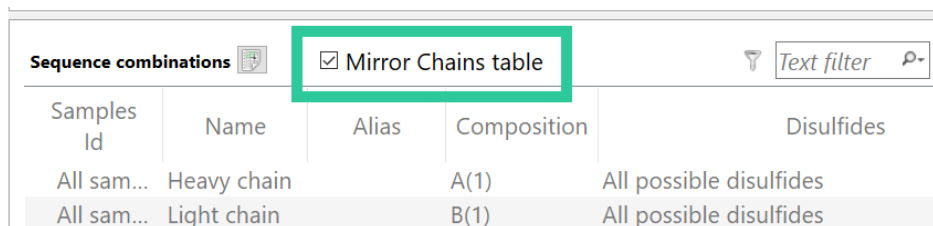


Figure 43: Average or monotypic mass assignment

The software automatically computes the average and monoisotopic mass for each selected FASTA entry, and these masses are then available for automatic peak assignment. The default computation assumes that Cys residues are disulfide bonded, but this default can be changed by the user. Protein sequences can be arbitrary if the user inputs the average or monoisotopic masses, overriding the computed masses. Average or monoisotopic masses give the software reference masses to assign mass peaks automatically based on mass deltas.

The **Mirror Chains table** check box populates the content from the Chains table into the Proteins table. This is shown below:



Samples Id	Name	Alias	Composition	Disulfides
All sam...	Heavy chain	A(1)	All possible disulfides	All possible disulfides
All sam...	Light chain	B(1)	All possible disulfides	All possible disulfides

Figure 44: Mirror Chains table option.

For each row in the “Chains” table, a corresponding row is added to the “Proteins/protein complexes” table (with single-chain composition and all possible disulfides). The user can also set the number of disulfides within the Proteins/protein complexes table. The “Disulfides” window is activated by double clicking within the box within the table, as shown below:

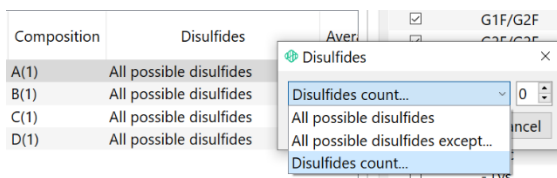


Figure 45: Disulfides.

To set a specific number, select “Disulfides count” and enter a value, as shown below:

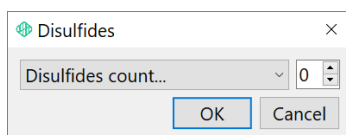



Figure 46: Disulfides count.

Sequence combinations are associated with all samples, by default. However, they can be associated with an individual sample. Click  after the **Sample Id** for the desired sequence combination to open the **Edit samples selection** dialog:

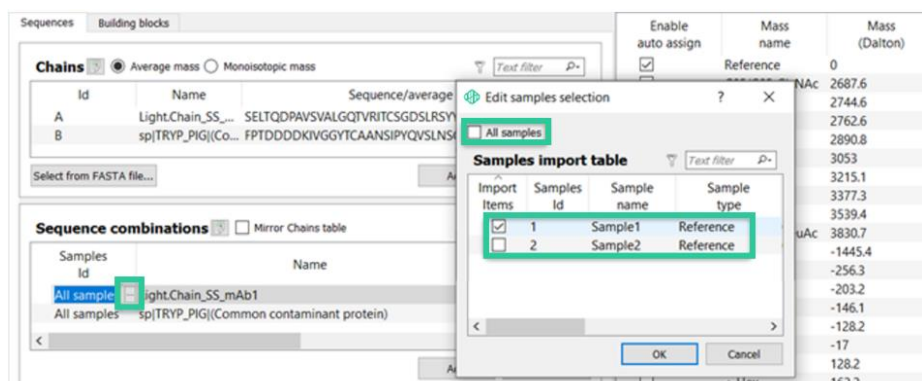


Figure 47: Edit sample selection for sequence combinations

Check either **All samples** or one of the sample rows and click **OK**. More than one checked sample for a sequence combination is not supported.

The user can also direct the software to consider clipping. As shown in figure below, select the box to activate – there are the options: 1. “Clip at specific sites” or 2. “Clip everywhere”. Click **Add** to enter specific residues depending on the radio button selected.

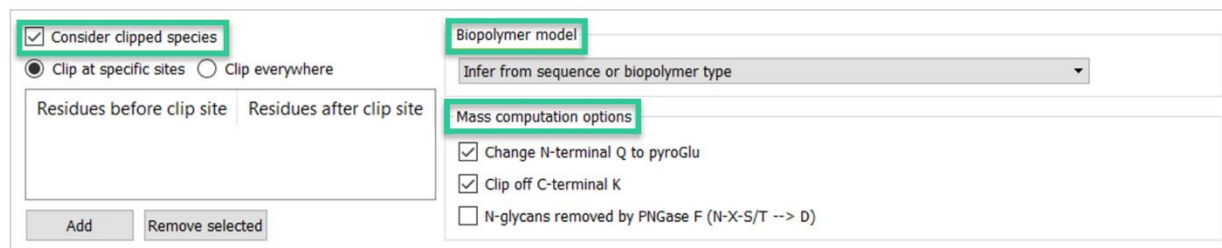


Figure 48: Consider clipped species, Biopolymer model and Mass computation options

Users can select the kind of biopolymer model used to convert average mass to monoisotopic mass and monoisotopic mass to average mass:

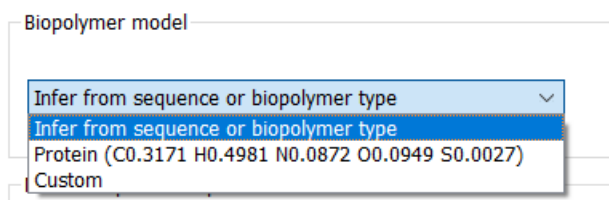


Figure 49: Biopolymer model choices

Biopolymer model options include:

- **Infer from sequence or biopolymer type** is the default selection, which uses the standard model that Intact Analysis has used to calculate mass.
- **Protein (C0.3171 H0.4981 N0.0872 O0.0949 S0.0027)** uses the displayed averagine formula to calculate mass.



- **Custom** opens a cell to the right of the selection to add a custom formula to calculate mass.

Note that the formula values refer to atom counts (molar ratios), not mass ratios.

There are also a few Mass computation options shown in figure below. The user should select which to consider when calculating the masses.

Within the **Sequences and masses** tab, there is also the **Building blocks** tab, as shown below:

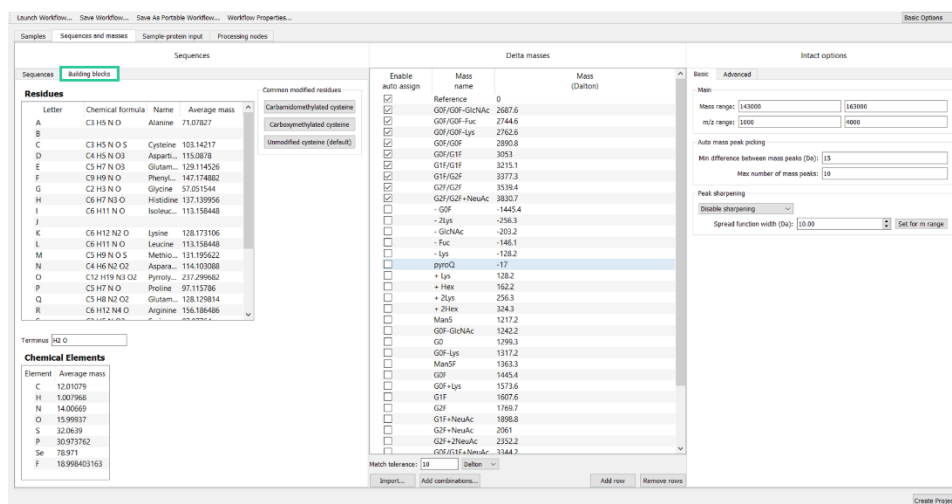


Figure 50: Building blocks sub-tab.

There are several parameters to consider in this tab. As shown in the three figures below, the cysteine modification can be clarified by selecting amongst the 3 options. The chemical formula will be adjusted accordingly.

Residues			Common modified residues
Letter	Chemical formula	Computed Mass	
A	C3 H5 N O	71.07827	Carbamidomethylated cysteine
B			Carboxymethylated cysteine
C	C3 H5 N O S	103.14217	Unmodified cysteine (default)

Figure 51: Unmodified cysteine (default).

Residues			Common modified residues
Letter	Chemical formula	Computed Mass	
A	C3 H5 N O	71.07827	Carbamidomethylated cysteine
B			Carboxymethylated cysteine
C	C5 H8 N2 O2 S	160.193714	Unmodified cysteine (default)

Figure 52: Carbamidomethylated cysteine.

Residues			Common modified residues
Letter	Chemical formula	Computed Mass	
A	C3 H5 N O	71.07827	Carbamidomethylated cysteine
B			Carboxymethylated cysteine
C	C5 H7 N O3 S	161.178426	Unmodified cysteine (default)

Figure 53: Carboxymethylated cysteine.

The user can also specify the average mass of each chemical element, as shown below:



### Chemical Elements

Element	Average mass
C	12.01079
H	1.007968
N	14.00669
O	15.99937
S	32.0639
P	30.973762
Se	78.971
F	18.998403163

Figure 54: Chemical Elements

### Delta masses

The Delta masses in the center of the Sequences and masses tab (shown in Figure 50) gives a table of likely mass differences between observed peaks and the reference mass (either input by the user or computed from an amino acid sequence, as completed in the Proteins section). Byos uses this information to assign peaks based on mass differences from Reference peaks. Check each group to include it for peak assignments. The mass names and Dalton values can be edited by selecting and typing new values. The Delta mass table is also useful in automated mass assignments when processing many data files.

The user can also add delta masses not available by default by clicking **Add row** and entering a mass name and mass manually. This is shown below in 5.

Delta masses

Enable auto assign	Mass name	Mass (Dalton)
<input type="checkbox"/>	+ 2Hex	324.3
<input type="checkbox"/>	Man5	1217.2
<input type="checkbox"/>	G0F-GlcNAc	1242.2
<input type="checkbox"/>	G0	1299.3
<input type="checkbox"/>	G2F+2NeuAc	2352.2
<input type="checkbox"/>	G0F/G1F+NeuAc	3344.2
<input type="checkbox"/>	G1F/G1F+NeuAc	3506.4
<input type="checkbox"/>	G1F/G2F+NeuAc	3668.5
<input type="checkbox"/>	G1F/G2F+2NeuAc	3959.8
<input type="checkbox"/>	G2F/G2F+2NeuAc	4121.9
<input type="checkbox"/>	G2F/G2F+3NeuAc	4413.2
<input checked="" type="checkbox"/>	Custom	4410.8

Match tolerance:  Dalton

Figure 55: **Add row** to add a delta mass.

Any row can be deleted by highlighting and clicking the **Remove rows** button. Select multiple rows by using Ctrl-click or Shift-click. This is shown below in 6.

Delta masses

Enable auto assign	Mass name	Mass (Dalton)
<input type="checkbox"/>	+ 2Hex	324.3
<input type="checkbox"/>	Man5	1217.2
<input type="checkbox"/>	G0F-GlcNAc	1242.2
<input type="checkbox"/>	G0	1299.3
<input type="checkbox"/>	G2F+2NeuAc	2352.2
<input type="checkbox"/>	G0F/G1F+NeuAc	3344.2
<input type="checkbox"/>	G1F/G1F+NeuAc	3506.4
<input type="checkbox"/>	G1F/G2F+NeuAc	3668.5
<input type="checkbox"/>	G1F/G2F+2NeuAc	3959.8
<input type="checkbox"/>	G2F/G2F+2NeuAc	4121.9
<input type="checkbox"/>	G2F/G2F+3NeuAc	4413.2
<input checked="" type="checkbox"/>	Custom	4410.8

Match tolerance:  Dalton

Figure 56: Remove rows.

The user can also import a custom list (.CSV) by clicking **Import** and selecting a saved delta mass table in CSV format. This is shown below:

Delta masses

Enable auto assign	Mass name	Mass (Dalton)
<input checked="" type="checkbox"/>	Reference	0
<input checked="" type="checkbox"/>	G0F/G0F-GlcNAc	2687.6
<input checked="" type="checkbox"/>	G0F/G0F-Fuc	2744.6
<input checked="" type="checkbox"/>	G0F/G0F-Lys	2762.6
<input checked="" type="checkbox"/>	G0F/G0F	2890.8
<input checked="" type="checkbox"/>	G0F/G1F	3053
<input checked="" type="checkbox"/>	G1F/G1F	3215.1
<input type="checkbox"/>	pyroQ	-17
<input type="checkbox"/>	+ Lys	128.2
<input type="checkbox"/>	+ Hex	162.2
<input type="checkbox"/>	+ 2Lys	256.3
<input type="checkbox"/>	+ 2Hex	324.3

Match tolerance:  Dalton

Choose csv file

« ExampleData » Intact

Organize ▾ New folder

Name	Date modified	Type
PMI-Suite		
Base		
Documents		
ExampleData		
Intact		
Supernovo		
Tools		

File name:  CSV Files (\*.csv)

Figure 57: Import a custom mass deltas list.

## Intact options

Intact options

Basic

Advanced

Main

Mass range: 143000 163000

m/z range: 1000 4000

Auto mass peak picking

Min difference between mass peaks (Da): 15

Max number of mass peaks: 10

Peak sharpening

Disable sharpening

Spread function width (Da): 10.00 Set for m range

Figure 58: Intact options.

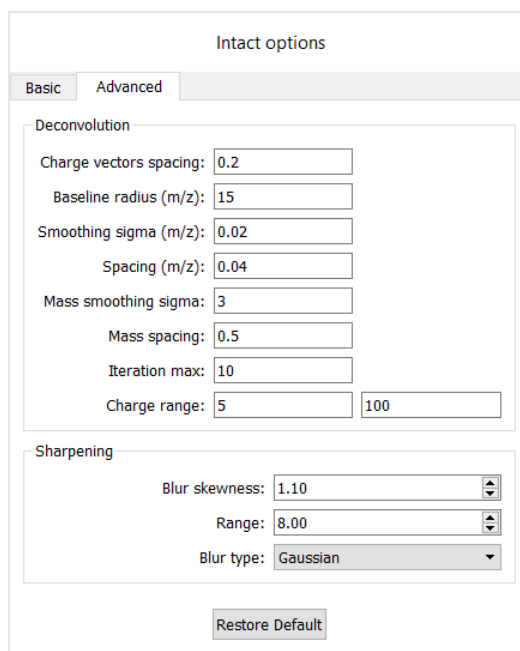
The Intact options section on the right side of the Deltamasses tab (shown in Figure 50) includes Basic and Advanced entries for Mass range, m/z range, mass peak picking, and peak sharpening parameters that are applied to all trace peaks.

The Basic tab contains the primary settings used in computing deconvolved masses. Mass range and m/z range set ranges for neutral masses and m/z, respectively. Mass range defines the range of neutral masses displayed in the Deconvolved mass spectrum. M/z range defines the segment of the MS1 spectrum used to compute neutral masses.

Within the Auto mass peak picking parameters, Min difference between mass peaks and Max number of mass peaks control peak picking. Min difference between mass peaks prevents the peak picker from picking multiple points on top of a ragged or isotope-resolved mass peak. Max number of mass peaks sets a limit on the number of picked peaks.

The Peak sharpening dropdown controls an optional “super-resolution step” that sharpens peaks beyond what is seen in the data. It deconvolves a “point spread function” to give super-resolved mass peaks and turn shoulders into separate peaks. The Spread function width sets the expected width of a peak when peak sharpening is enabled. A reasonable width is the square root of the peak mass in kilodaltons, for example, 12 for a spectrum from 140,000 to 150,000 Da. A too-narrow width will not sharpen much, and a too-wide width might split peaks. Peak sharpening on mass spectra that span a range of more than about 20,000 Da is not recommended, because no single width will be optimal for the entire range. Peak sharpening should also be avoided on isotope-resolved mass spectra.

The **Advanced** options mainly concern resolution. The user will only need to use this tab if the aim is to produce isotopically resolved neutral mass spectra.



Intact options

Basic Advanced

Deconvolution

Charge vectors spacing: 0.2

Baseline radius (m/z): 15

Smoothing sigma (m/z): 0.02

Spacing (m/z): 0.04

Mass smoothing sigma: 3

Mass spacing: 0.5

Iteration max: 10

Charge range: 5 100

Sharpening

Blur skewness: 1.10

Range: 8.00

Blur type: Gaussian

Restore Default

Figure 59: Advanced options.

**Charge vectors** give the charge assignment probabilities for each small interval of  $m/z$  points. Deconvolutions of most MS1 spectra will be almost exactly the same with any **Charge vectors spacing** from 0.2 to 1  $m/z$  units (Thomsons), but a narrow spacing of 0.1 may give better results on isotope-resolved MS1 spectra with interleaved signals, and a wide spacing of 2 may give better results for native MS with broad  $m/z$  peaks.

Intact Analysis removes an  $m/z$  baseline before deconvolution; this step and removal of charge one or stop-list peaks are the only steps in the algorithm that do not conserve ions. **Baseline radius** controls the stiffness of the baseline. A baseline radius of 8 gives a flexible baseline that will cut into  $m/z$  peaks broader than 8 Thomsons; this will often give better visual separation of neutral mass peaks, but may distort peak areas. A baseline radius of 30 will give a stiffer baseline that cannot cut into  $m/z$  peaks narrower than 30 Thomsons. An even larger value, 100 or more, may be needed for native MS. The default value of 15 is a compromise.

**Spacing (m/z)** controls the spacing of sample points in the  $m/z$  spectrum. The raw MS1 data is represented as a continuous piecewise-linear function that can support any spacing of sample points, but  $m/z$  spacing finer than the finest spacing in the original data will slow the computation without adding resolution. Reasonable values for Spacing (m/z) are in the range 0.005 – 0.05 for QTOF instruments, which have almost the same resolution at all  $m/z$ ; the low setting of 0.005 would be appropriate for Bruker maXis and the higher setting of 0.05 for older instruments with lower resolution. For Orbitrap, the resolution depends upon  $m/z$ ; the setting of 0.005 shown above is for isotope-resolved 25 kDa masses (antibody subunits). For native MS on Exactive EMR with  $m/z$ 's in the 5000 – 10,000 range, a spacing of 0.1 is fine enough.

**Smoothing sigma** is typically set to the same value as Spacing (m/z), but a larger value can be helpful for producing an appropriately smoothed neutral mass spectrum with less smoothing at lower mass and more smoothing at higher mass.

**Mass spacing** controls the spacing of points in the neutral mass spectrum. To preserve isotopic resolution, spacing should be set to 0.1 or even 0.05. If the MS1 spectrum does not have isotopic resolution, or isotopic resolution is not needed for analysis, mass spacing in the range 0.2 to 1 is best for target molecules below 200 kDa. Spacing of 10 Da or more is best for targets above 300 kDa. For mass spectra without isotopic resolution, **Mass smoothing sigma** in the range of 2 – 5 will smooth jittery peaks in the range 20 – 200 kDa; larger values will be needed for larger masses. For mass spectra with isotopic resolution 0.1 will work.

**Iteration max** set to 10 will work for most purposes. A larger value, for example 20 or 30, can be helpful for lower signal-to-noise spectra that take longer to converge.

**Charge range** is best set to a wide range, in which case the charge range will be implied by the mass and m/z ranges. The default range of 5 – 100 covers most applications, but 5 will need to be reduced for deconvolutions with mass range starting below 10 kDa, and 100 increased for targets that may have charges above 100.

**Sharpening** uses the Spread function width set in the Basic tab to deconvolve the data. **Blur skewness** controls the asymmetry of the point spread function; the default value of 1.1 means that the right tail has sigma (standard deviation) 10% bigger than the left tail. 1.2 gives even more tailing; 1.0 gives a symmetric point spread function. **Range** sets the length of the tails in standard deviations; a small value of 5 or 6 may work better in the case of Lorentzian point spread. **Blur type** has two choices: Gaussian (skinny tails) and Lorentzian (fat tails). Lorentzian should use a slightly smaller sigma than Gaussian.

## Sample-protein input Tab

The **Sample-protein input** tab allows for association of protein sequences with the samples that contain them:

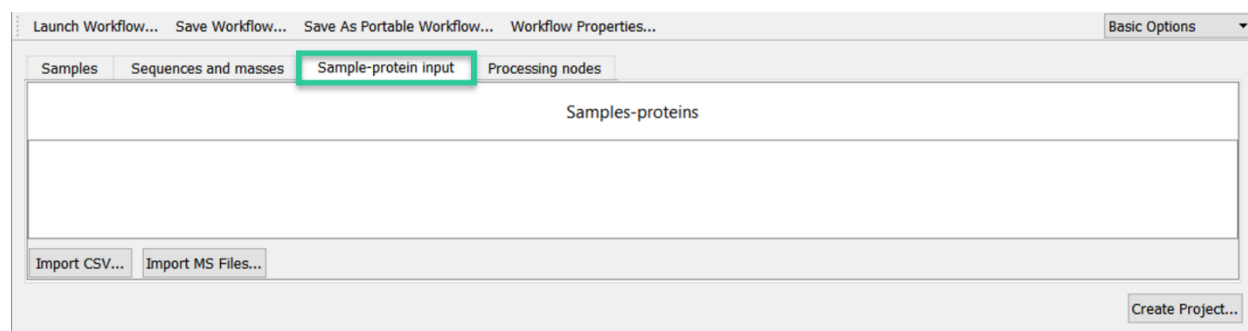


Figure 60: Sample-protein input tab

Sample-protein associations can be imported from \*.csv files and from MS files. To create csv files from other projects, use **File > Export > Generate MS path template CSV**. This capability is useful for making a single Intact Analysis project with many different samples. Imports to the tab support custom fields.

## Processing nodes

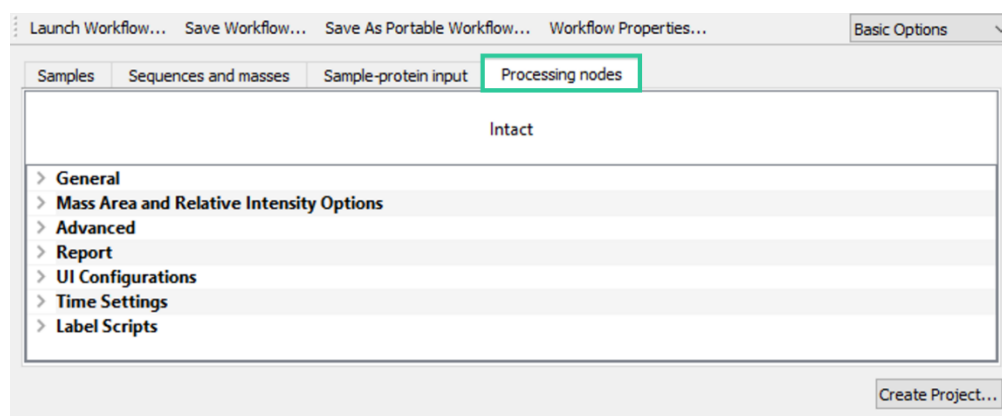


Figure 61: Intact Processing nodes tab.

## Intact

### • General

▼ General	
Samples	*
Enable Lock-Mass Calibration	— No
Lock Mass (m/z)	556.2771
Mass assignments	Auto charge deconvolution and mass assignments

Figure 62: General parameters.

- o **Samples** - The "\*" character applies all parameters to all samples dragged and dropped into the Samples tab.
- o **Enable Lock-Mass Calibration** can be set as yes or no.
- o **Lock Mass (m/z) sets** the calibrant m/z value. Several are available using the drop-down menu or the user can type in a numerical value. If empty, no calibration will be applied. The user can select:

1221.990637	▼
556.2771	
609.2812	
785.84206	
121.050873	
322.048121	
622.028960	
922.009798	
1221.990637	
1521.971475	

Figure 63: Lock mass drop-down values.

- o **Mass assignments**

Mass assignments	Auto charge deconvolution and mass assignments ▼
	No charge deconvolution
	Auto charge deconvolution
	Auto charge deconvolution and mass assignments

Figure 64: Mass assignments.

The Mass assignments parameter allows the user to turn on/off charge deconvolution and automatic mass assignment. The default selection is Auto charge deconvolution and mass assignments. For new projects, especially those with longer chromatography times, significant computation time may be saved by creating the project with the setting No charge deconvolution selected. The summed m/z spectra can then be viewed before deciding which elution peaks warrant deconvolution. Similarly, charge deconvolution without mass assignments will let the neutral mass spectra be viewed before mass peak assignment.

## • Mass Area and Relative Intensity Options

▼ Mass Area and Relative Intensity Options		
Compute Areas of Mass Peaks	✓ Yes	
Mass Area Width	500	
Report Intensities Relative to Local Base Peak	✓ Yes	
Window for Local Base Peak (%)	20	None
Minimum % of Local Base Peak	10	None
Generate zoomed-in segments	None	Using reference masses
Plot segment width	50.00	Using observed masses

Figure 65: Mass Area and Relative Intensity Options

- **Compute Areas of Mass Peaks**
- **Mass Area Width** compute peak area within a band around each mass, defined by the mass area width value
- **Report Intensities Relative to Local Bas Peak** reports intensities based upon the two below parameters
- **Window for Local Base Peak (%)** sets tolerance for local base peak (e.g. mass within +/- 20%),
- **Minimum % of Local Base Peak** filters out masses below a certain % of local base peak
- **Generate zoomed-in segments** The user has 3 options: None, using reference masses, Using observed masses (per highest local base peak).
- **Plot segment width** Segments can be set to be automatically generated around reference/observed masses for deconvoluted mass spectrum and MS1 plots during project creation.

## • Advanced

▼ Advanced	
Advanced configuration	
Ionization Mode	Positive
	Positive
	Negative

Figure 66: Advanced.

Several advanced commands can be applied to processing by adding in the **Advanced configuration** text box. Please refer to the **Advanced Commands** section of the relevant workflow manual. The user can select between Positive and Negative using the drop-down menu.

## • Report

▼ Report	
Report Configuration Path	nfig\Intact_Mass_Pivot_Report_withPlots (ReferenceDefault).rptc

Figure 67: Report Configuration Path

Each Byos default workflow includes a report template created by our Customer Success team that is optimized for the specific type of analysis. If the user prefers a customized report template, they can direct Byos to this file using the light blue “...” button. They will be prompted to select a file.

## • UI Configurations

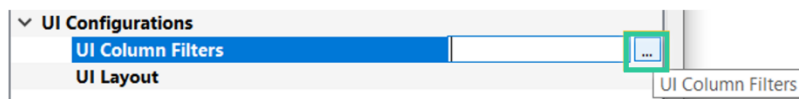


Figure 68: UI Configurations – UI Column Filters and UI Layout.

The user has the ability to import column filters as well as layout files. This is designed to standardize analyses across all users, labs, and sites. The user can direct Byos to the preferred file for each using the light blue “...” button. The user will be prompted to select a file.

### • Time Settings

▼ Time Settings	
Start time of interest	0.00
End time of interest	0.00
Alignment max time	1.00

Figure 69: Time Settings

- **Start time of interest** and **End time of interest** control the time limits within which the computation of the baseline and peaks are completed. The default values for both are set to 0.00, which means filtering will not be applied. The user can set them to apply the additional processing options.
- **Alignment max time** sets the maximum alignment value between plots (for example, UV and TIC). This value will limit the allowed alignment time between the two signals. The default value is set to 1.00.

### • Label Scripts

This feature allows users to customize peak labels for Trace plot and Deconvolved Mass Plot.

▼ Label Scripts	
Trace plot	[PeaksDynamicNumber] - <script>max(['ApexPositIntensity'], ['Sequence'])</script>
Deconvolved mass plot	<pre>&lt;script&gt; (function () {   var proteinName = templateFunctionsManager.detectValue(['ProteinAccessionName']).trim();</pre>

Figure 70: Label Scripts

Scripts related to trace plot are in C:\Program Files\ProteinMetrics\PMI-Suite\Base\labelscripts\traceplot folder, while the scripts related to deconvolved mass plot are in C:\Program Files\ProteinMetrics\PMI-Suite\Base\labelscripts\dmsplot folder. To load a script during project creation, select **Processing nodes**, expand **Label Scripts**, click on ... for Trace plot (or Deconvolved mass plot), then click **Load** to select a script, then click **OK** and **Create Project**. This will result in creation of the project with custom labels as specified in the script.

The user has an option to load scripts after project creation as well. To customize peak labels after project creation, select **Rendering options** icon, click **Edit Annotations**, then click **Load** to select a script, then click **Open** and click **OK** to display new custom labels.

### • Peak Construction options (older versions of Byos)

The Peak Construction option parameters formerly found in Byos are now set directly in the Samples tab with the buttons **Edit trace peak options** and **Edit global trace peak filters**.



## PTM, PTM (in-silico), PTM-DIA, HCP, S-S, SVA (C57) – Specific, SVA (C58) – Specific, HotSpot, System Suitability, Oxidative Footprinting, and MAM New Peak Detection

The following section details parameters in the **Processing nodes** tab for the combination of MS/MS IDs and Quant. These parameters are therefore relevant to the **HCP, HotSpot, PTM, PTM (in-silico), ptm-dia, SVA (C57) – Specific, SVA (C58) – Specific, S-S, System Suitability, Oxidative Footprinting, and MAM New Peak Detection** workflows.

NOTE: Parameters specific to certain workflows (including SVA and System Suitability) are called out.

### Processing nodes

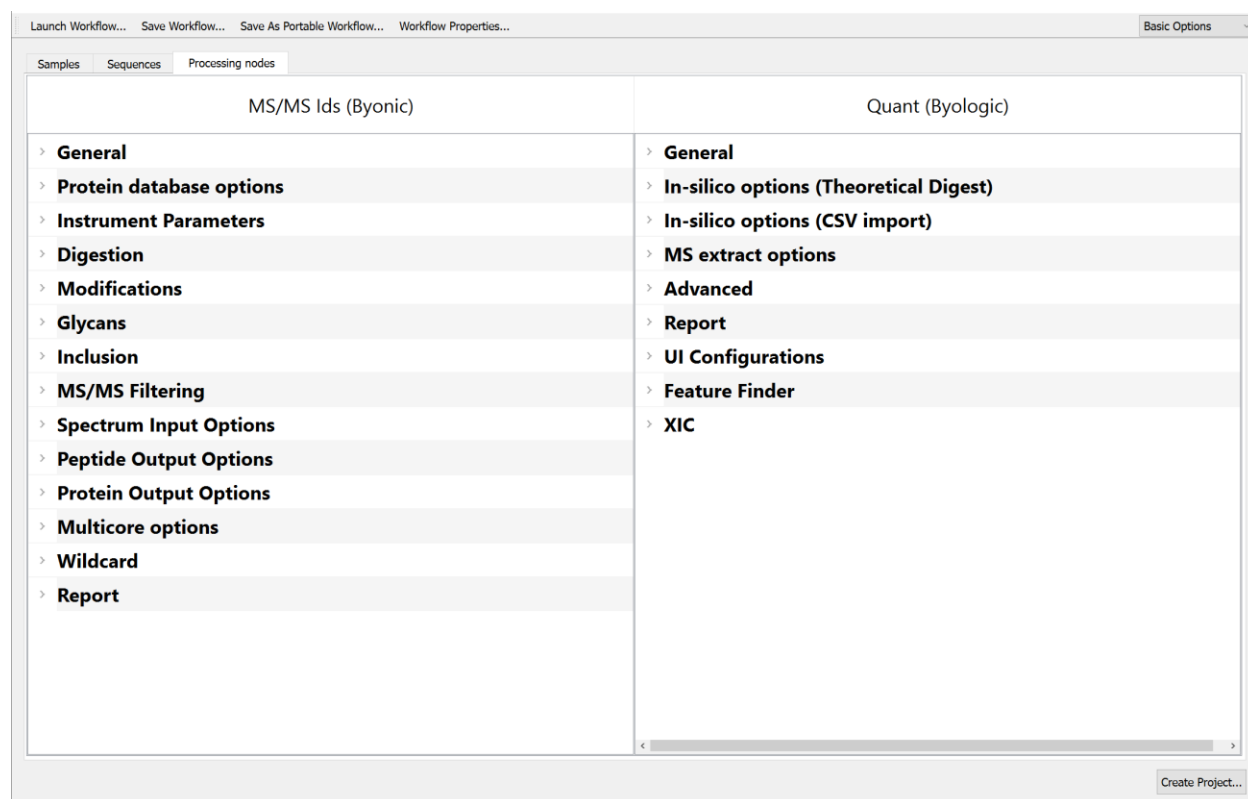


Figure 71: Peptide Analysis Processing nodes.

### MS/MS Ids

NOTE: These parameters are not part of the PTM (in-silico) workflow. Please refer directly to the below Quant section.

- **General**

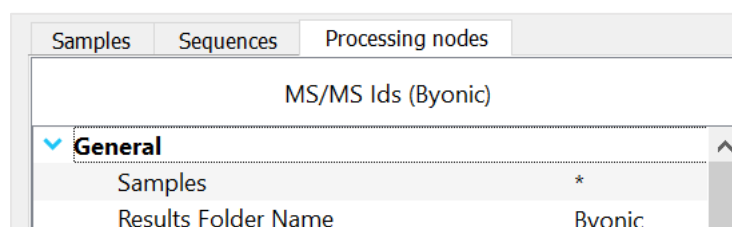


Figure 72: General.

- o **Samples** - The “\*” will apply all parameters to all samples dragged and dropped into the Samples tab.
  - o **Results Folder Name** creates a folder of that name to save results, set to “Byonic” by default.
- **Protein database options**

▼ Protein database options	
Add Common Contaminants	— No
Add Decoys	✓ Yes

Figure 73: Protein database options.

The protein database should contain both targets and decoys (recognized by protein names beginning >Reverse or >Decoy) for false discovery rate (FDR) estimation. Byonic will automatically add decoys if the **Add decoys** box is checked and contaminant proteins (e.g., trypsin, bovine serum albumin, and human keratins) if the **Add common contaminants** box is checked. Typical folders to store input files are: C:\data\_input\Mass\_Spectra and C:\data\_input\Protein\_Databases.

- **Instrument Parameters**

▼ Instrument Parameters	
Precursor Mass Tolerance	6.00 ppm
Fragmentation Type	QTOF / HCD
Fragment Mass Tolerance 1	20.00 ppm
Fragment Mass Tolerance 2	20.00 ppm
Recalibration (lock mass)	None

Figure 74: Instrument Parameters. Optimize for the experiment completed.

In the figure above, the user set 6.0 ppm **Precursor Mass Tolerance**, 20.00 ppm **Fragment Mass Tolerance**, and QTOF/HCD as the **Fragmentation Type**. Both Dalton and ppm mass tolerances for precursors and fragments are supported, along with several fragmentation types. The Dalton tolerance applies to measured mass for precursors but measured m/z for fragments. The way scoring is completed changes at fragment tolerances of 0.1 Da or 100 ppm or less: high-resolution MS/MS is assumed, meaning resolution sufficient to distinguish charge states of fragment ions. For this reason, fragment tolerances larger than 0.1 Da should be used with low-resolution (ion trap) MS/MS analysis.

Internal models for most fragmentation types are included – CID low energy (ion trap), QTOF / HCD (beam-type CID), and ETD / ECD (electron transfer and electron capture dissociation), as well as a number of combinations of types. These internal models determine which fragment peak types will be scored and annotated. For example, prominent c- and z-ions and small y-ions are expected for ETD. Prominent oxonium ions are expected from glycopeptides with QTOF / HCD fragmentation, but small or missing oxonium ions from CID low energy.

- o **Precursor Mass Tolerance**

▼ Instrument Parameters	
Precursor Mass Tolerance	6.00 ppm

Figure 75: Precursor Mass Tolerance.

The user can change this value by clicking within the text box and then clicking on the activated blue “...” square. The user can then modify the text value and mass accuracy, as required. .

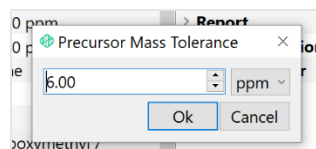


Figure 76: Modify the text value and mass accuracy.

- o **Fragmentation Type**

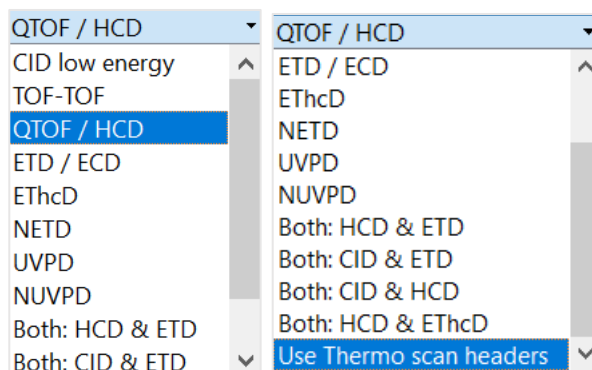


Figure 77: Fragmentation type.

The user can select from the available options using the drop-down menu. Additional options are visible using the scroll bar, including “Both:” for spectrum file(s) containing more than one fragmentation type and “Use Thermo scan headers” to read directly from Thermo raw data files.

- o **Fragment Mass Tolerance 1**- The user can set the value used to acquire data in either ppm or Da.
- o **Fragment Mass Tolerance 2** - If necessary, the user can set the value used to acquire data in either ppm or Da. This is only required for the “Both:” Fragmentation Types (selected as shown in Figure 74 above – the final entries). The user can set the value used to acquire data in either ppm or Da. No value will be applied if only a value for “Fragmentation Mass Tolerance 1” was entered.
- o **Recalibration (lock mass)**

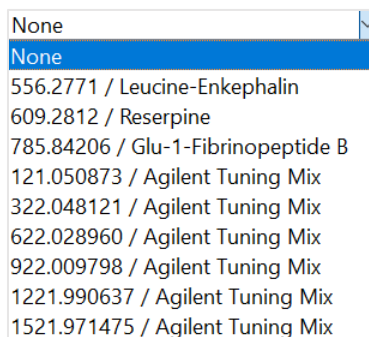


Figure 78: Lock mass calibration options.

The user can select from the available options using the drop-down menu.

- **Digestion**

▼ Digestion	
Cleavage Site(s)	RK
Cleavage Side	C-terminal
Digestion Specificity	Fully specific (fastest)
Missed Cleavages	2

Figure 79: Digestion options.

The Digestion settings allow the user to set the residues recognized by the digestion enzyme. In this example, the enzyme is trypsin, so the user entered RK for arginine and lysine for the Cleavage Site(s) and chose C-terminal for the Cleavage Side.

- **Cleavage Site(s)** - The user can change this value by entering text. If the user leaves the Cleavage Site(s) box empty, the only specific cleavage sites are protein termini.
- **Cleavage Side**

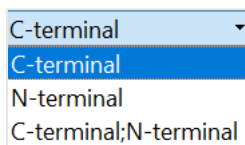


Figure 80: Cleavage Side options.

The user can change this by selecting another option from a drop-down menu. Click on the current selection to activate the drop-down menu to view the available options.

- **Digestion Specificity**

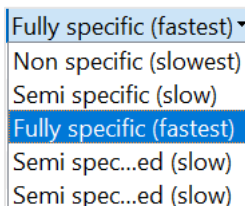


Figure 81: Digestion Specificity

In the figure above, the user chose a “Fully specific (fastest)” search, meaning that both the N- and C-terminal cleavages must be C-terminal to R or K. Nonspecific cleavage at either or both endpoints is supported. A nonspecific search with RK in the Cleavage Site(s) box searches all peptides but favors tryptic peptides; the user must leave the Cleavage Site(s) box empty for a true no-enzyme search. Digestion Specificity can be changed by selecting another option from a drop-down menu. The user can click on the current selection to activate the drop-down menu to view the available options.

- **Missed Cleavages** - The user selected 2 Missed Cleavages, as shown in Figure 79. This limits the maximum number of internal Rs and Ks not followed by P to 2; leaving Missed Cleavages at its default value of -1, which means any number of internal Rs and Ks. Missed Cleavages can be changed by entering text.

- **Modifications**

▼ <b>Modifications</b>	
Modifications	Carboxymethyl / +58.005479 @ C   fixed Oxidation / +15.994915 @
Total Common Max	1
Total Rare Max	1

Figure 82: Modifications options

Like most proteomics search engines, two types of modifications are supported: fixed and variable. A fixed modification is assumed to occur on all the residues of that type, but a variable modification is optional, so that each site for a variable modification is considered with and without the modification.

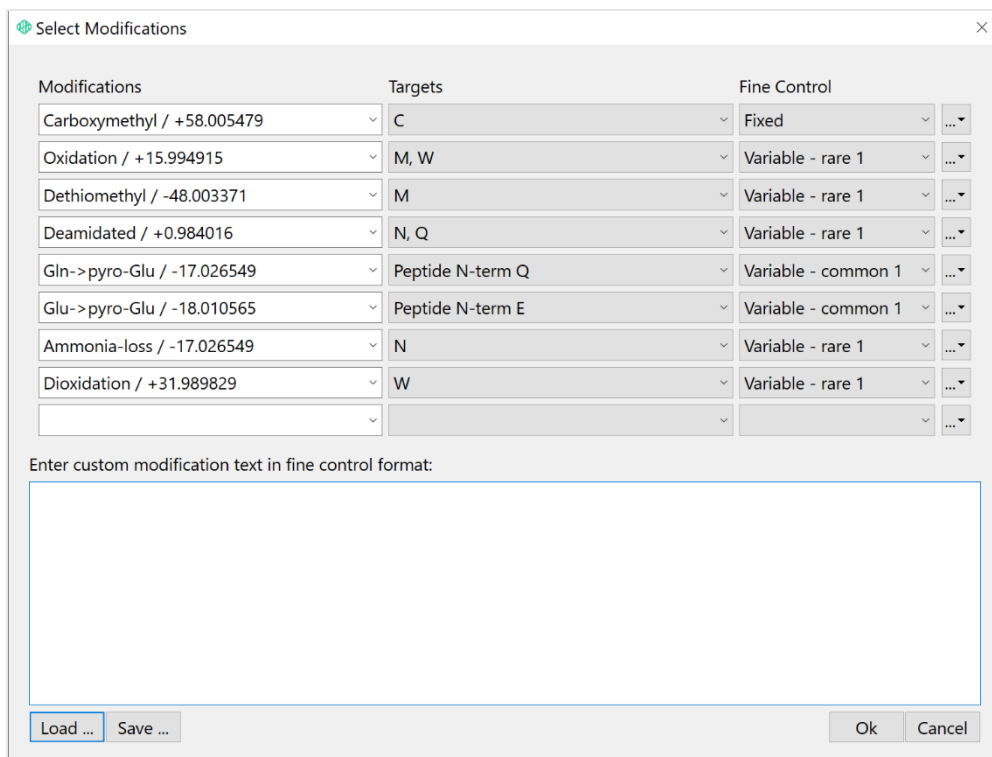
o **Modifications**



Figure 83: Click within the text box to activate the light blue “...”button

To view the list of modifications included with the default workflow, the user can click within the text box to activate the light blue “...”button. It will take a second or two for the window to open.

The user can then specify any number of modification rules via a pull-down menu containing all the modifications listed in [www.unimod.org](http://www.unimod.org). For convenience, frequently used modifications are listed twice, at the top and again in the complete list. The three pull-down menus in each row select modification type, target residues, and fine control. There is a fourth pull-down, which lets the user delete, invert (as in (De)Carbamidomethyl), or add “attributes” to modifications. Attributes allow the user to define protein-specific modifications.



The 'Select Modifications' window contains a table with three columns: Modifications, Targets, and Fine Control. It lists various modifications like Carboxymethyl, Oxidation, Dethiomethyl, etc., with their mass deltas and target residues. Each row has dropdown menus for selecting the modification, target, and fine control setting. At the bottom, there is a large text area for custom modifications and buttons for Load, Save, Ok, and Cancel.

Modifications	Targets	Fine Control
Carboxymethyl / +58.005479	C	Fixed
Oxidation / +15.994915	M, W	Variable - rare 1
Dethiomethyl / -48.003371	M	Variable - rare 1
Deamidated / +0.984016	N, Q	Variable - rare 1
Gln->pyro-Glu / -17.026549	Peptide N-term Q	Variable - common 1
Glu->pyro-Glu / -18.010565	Peptide N-term E	Variable - common 1
Ammonia-loss / -17.026549	N	Variable - rare 1
Dioxidation / +31.989829	W	Variable - rare 1

Enter custom modification text in fine control format:

Load ... Save ... Ok Cancel

Figure 84: Select Modifications window

- o **Total Common Max** and **Total Rare Max** - A unique feature not found in other search engines is offered: the user designates each variable modification as either “common” or “rare”, with the names suggesting their use. The user can define separate limits on the number of occurrences of each variable modification, so that “common 2” means at most two occurrences per peptides. Separate limits can also be set for the total number of common and rare modifications per peptide. A typical search allows a total of at most two common modifications and a total of at most one rare modification per peptide. To search for, say, three phosphoserines per peptide, the user can change Total common modification max to 3 or split phosphorylated serine between two rules: common2 and rare1. Depending upon the other modification rules, the latter approach may give a faster search. Please review the “Modification Fine Control” Application Note available at <https://www.proteinmetrics.com/resources/>.

NOTE: The single most important factor in search time is **Total Common Max** (shown in Figure 82). Roughly speaking, the search time grows as  $C \cdot T$  where  $C$  is the number of common modifications enabled and  $T$  is Total Common Max.

Conceptually, the search engine has one modification “slot” for each residue, along with slots for the peptide’s N- and C-termini. A variable modification such as +0.984016 @ N uses up the residue slot; a nonspecific terminal modification such as +57.021464 @ NTerm uses up the terminal slot; but residue-specific N-terminal modifications, such as -17.026549 @ NTerm Q, use up both the residue and the N-terminal slots.

The big open box (shown in the figure above) is a space for the user to type in custom modifications not listed in Unimod. The manual fine control format has the form:

Modification\_Name / Mass\_Delta @ Targets | Fine\_Control

Modification\_Name / is optional. The Targets field allows the 20 one-letter amino acid abbreviations, as well as four special locations: NTerm, CTerm, Protein NTerm, and

Protein CTerm, NTerm, CTerm, Protein NTerm, and Protein CTerm can also be used as modifiers of amino acid residues. Targets form a comma-separated list.

Here is an example of a real modification not (yet) in Unimod:

```
DehydroFormyl / +9.98435 @ NTerm S, NTerm T | rare1
```

A limited number of nonstandard amino acid residues can be supported by redefining one-letter amino acid abbreviations using fixed modifications. B, Z, U, O, J, and X are accepted within FASTA protein databases, with masses, respectively, of 114.042927 (same as N), 128.058578 (same as Q), 150.95363 (selenocysteine), 237.052645 (pyrrolysine), 100.0, and 110.05 (close to averagine). By placing, for example, a fixed modification of +13.04768 on J, the user can make J in a FASTA database have mass 113.04768, correct for hydroxyproline. However, the amino acid sequence is used to predict peak intensity, so this fixed modification on J will not give the same scores as a +15.9949 variable modification on P.

For comprehensive sequence variant searches, or other searches with large numbers of modifications, it is more convenient to paste in a list of modifications in the custom modification box than to add all the modifications via the drop-down menus. Sequence variant lists are available from Protein Metrics by contacting [support@proteinmetrics.com](mailto:support@proteinmetrics.com).

## • Glycans

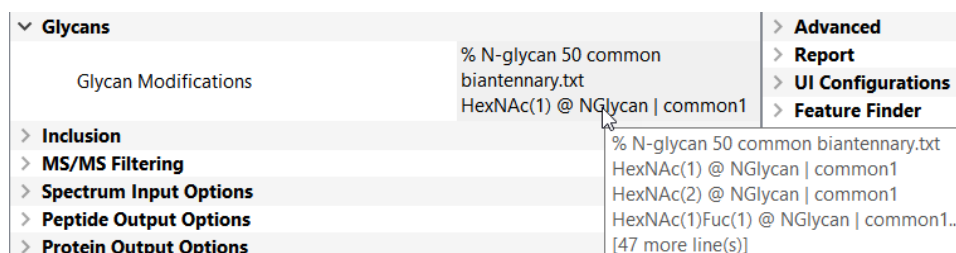


Figure 85: Glycans are loaded as a list of records

Three ways to define glycan modifications are offered: internal preset tables, external glycan databases, and user-defined glycans. Click the activated light blue “...” button to pop up a window labeled **Select Glycans**:

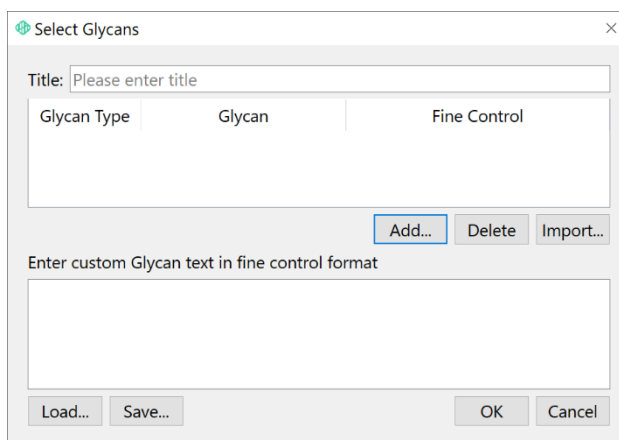


Figure 86: The Select Glycans

- o **Import** populates the dialog with a list of glycans from a glycan database text file:

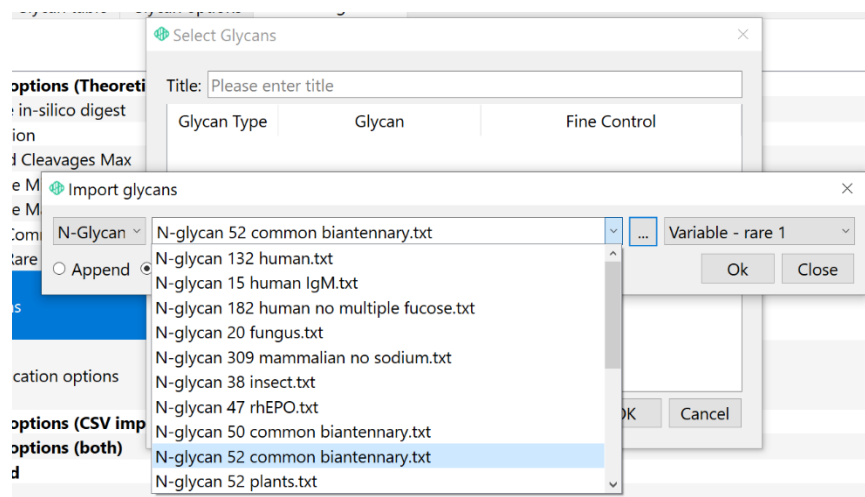


Figure 87: Importing glycan databases included with installation

Click the dropdown arrow to choose a glycan DB file. The dropdown displays a list of glycan database text files found in C:\Program Files\ProteinMetrics\PMI-Suite\Base\data\GlycanDatabases. These text files can be edited, and new glycan database text files can be added to the directory, where they become available in the dropdown (after closing and reopening Byos). This set of glycan databases is continually updated based on customer feedback. Please reach out to [support@proteinmetrics.com](mailto:support@proteinmetrics.com) to request additional content.

Alternatively, click the “...” button to open a custom glycan DB file from a different directory. The user can choose the Glycan type (N- or O-linked) and then set the Fine Control (rare1, common2, etc.).

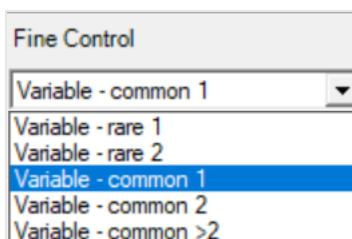


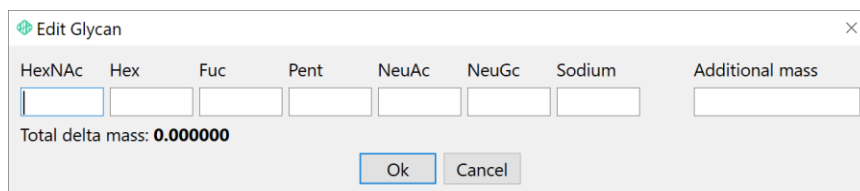
Figure 88: Glycan Fine Control options

The text files include one glycan composition per line; for example, the following gives five of the most common human O-glycans. Spaces between monosaccharides are optional, and unused monosaccharides can be left out or included with zero (0) occurrences.

```
HexNAc (1) Hex (0)
HexNAc (1) Hex (1) Fuc (0) NeuAc (0)
HexNAc (1) Hex (1) Fuc (0) NeuAc (1) NeuGc (0) Na (0)
HexNAc (1) Hex (1) Fuc (0) NeuAc (2) NeuGc (0) Na (0)
HexNAc (1) Hex (1) Fuc (1) NeuAc (0) NeuGc (0) Na (0)
```

- o **Add** opens the **Edit Glycan** dialog to create glycans from internal preset tables: .





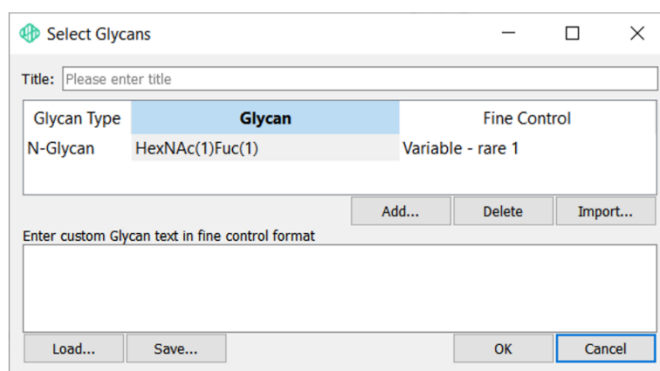
**Edit Glycan**

HexNAc	Hex	Fuc	Pent	NeuAc	NeuGc	Sodium	Additional mass
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Total delta mass: **0.000000**

Figure 89: Edit glycans.

Enter the count of each used monosaccharide. Six monosaccharide residues are allowed: HexNAc, Hexose, Fucose, Pentose (common in plants), NeuAc, and NeuGc (common in non-humans). There is also a box for Sodium because it is a common adduct on sialic acids. Unused monosaccharides can be left blank or included with zero (0) occurrences. Other glycan masses and modifications such as sulfation and acetylation can be defined with the Additional mass box; this mass is added to the mass of the monosaccharides. The total delta mass will automatically populate to six decimal places. Click **OK** to load the glycan to the Select Glycans dialog. Edit the Glycan Type and Fine Control, as needed.



**Select Glycans**

Title:

Glycan Type	Glycan	Fine Control
N-Glycan	HexNAc(1)Fuc(1)	Variable - rare 1

Enter custom Glycan text in fine control format

Figure 90: Glycans assembled from monosaccharides

- o A third way to enter glycans is to enter or paste glycan text in the **Enter custom glycan text in fine control format** box at the bottom. These glycans are entered using the same format as for individually added glycans: Monosaccharide(count) @ OGlycan or NGlycan | fine control option:

Enter custom glycan text in fine control format:

HexNAc(1)Hex(1)Fuc(1) @ OGlycan | rare1

Figure 91: Custom glycan entry format

Legacy Glycan DB file references can be converted into a list of custom glycans. When a glycan DB file and path are shown in the Glycan Processing node, click the “...” button to convert the reference to custom glycans:

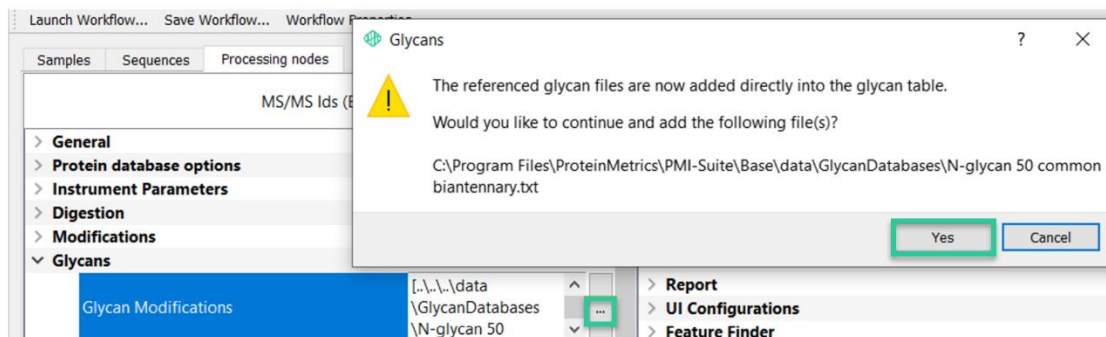
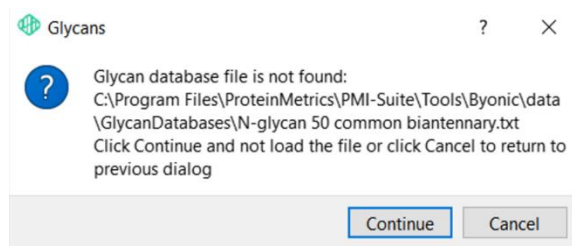


Figure 92: Converting glycan DB file reference to a glycan set

Click **Yes**, and the contents of the glycan DB file are converted into custom glycans. If the glycan DB file is no longer in the path specified, an error message gives the option to clear the obsolete file path reference:

Figure 93: Click **Continue** to clear an obsolete glycan DB file reference

Click **Continue** and the **Select Glycans** dialog opens, cleared of the glycan DB file reference. Glycans can now be added using one of the methods described above.

For some helpful examples and best practices for conducting N-linked and O-linked glycan searches, see our Application Notes at <https://www.proteinmetrics.com/resources/>.

- **Inclusion** is used to import a \*.csv file which defines m/z ratio ranges and/or elution time range segments

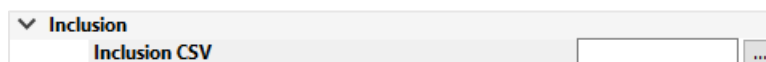


Figure 94: Inclusion option

- **MS/MS Filtering** creates MS/MS diagnostic peak filters.

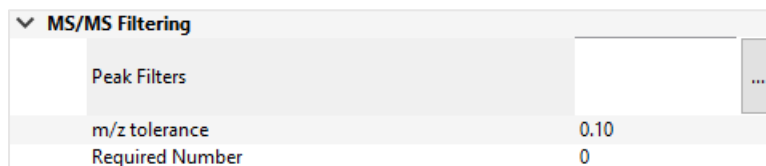


Figure 95: MS/MS diagnostic peak filtering

- o The **Peak Filters**  button in Byos opens the **Select Peaks** dialog:

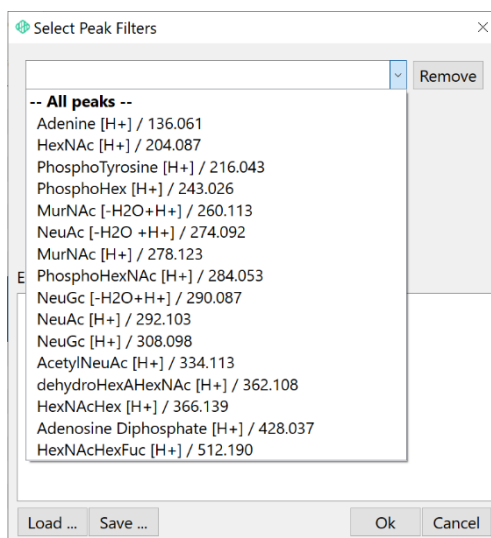


Figure 96: Select Peaks dialog for MS/MS Filtering, with dropdown

Click the drop-down arrow and select the diagnostic peak to filter. Alternatively, custom modifications can be entered manually in the box labeled **Enter custom diagnostic peaks**. The following is an example format of a diagnostic peak:

Kdn / 251.076

Set a maximum **m/z tolerance** value. **Required Number** sets a required count of diagnostic peaks to be applied from within the list. To set no minimum requirement, leave the value at zero.

- **S-S, Xlink** (NOTE: This set of parameters is unique to the S-S disulfide analysis workflow.)

S-S, Xlink	
Enable Disulfide	✓ Yes
Disulfide Text	Disulfide / -2.015650 @ C
Enable Trisulfide	— No
Trisulfide	Trisulfide / 29.956421 @ C
Enable Crosslink DSS	— No
Crosslink: DSS	DSS / 138.0680796 @ K   xlink
Enable Crosslink Custom	— No
Crosslink: Custom	
For FASTA Proteins	1,2

Figure 97: S-S and Xlink parameters.

This workflow allows the user to search for disulfide-bonded peptide pairs, trisulfide-bonded (also called persulfide-bonded) pairs, and more general cross-linking. This set of parameters also provides options to allow a user to search for expected and unexpected disulfide bonds. Enable S-S, Xlink analysis by setting **Enable Disulfide** parameter to “Yes”. Numerically designating which protein sequences from the FASTA database to consider in the **For FASTA Proteins** parameter, an in-silico digestion will be completed based on the digestion parameters selected. The Xlink analysis considers every peptide that contains a cysteine and look to pair with other cysteine-containing peptides. The separation of numbers in the “For FASTA proteins” field below indicates how the potential pairings should be considered. For example, “1” searches for

crosslinks in the first protein only. “4,5; 7” searches for all potential crosslinks on the 4th, 5th, and 7th protein, #4 and #5 may crosslink to each other, but not with #7.

- o The **Trisulfide** option allows a user to search for trisulfides within a single peptide and linking 2 peptides. Similarly, the Crosslink: DSS and Crosslink: Custom allow a user to search for crosslinks within a single peptide and linking 2 peptides.
- o The user can specify the modification fine control used in a custom crosslink search in the “Crosslink: Custom” text box.

#### • Spectrum Input Options

Spectrum Input Options help Byos cope with imperfect inputs. For example, on many MS instruments, precursor ion charges are uncertain for some or all spectra.

▼ Spectrum Input Options	
Apply Charges To	unassigned spectra
Charge States	
Precursor Isotope Off by X	No error check
Maximum Precursor Mass	12,000
Precursor and Charge Assignments	Compute from MS1
Maximum Number of Precursors per Scan	1
Smoothing Width (m/z)	0.01

Figure 98: Spectrum Input Options.

- o **Apply Charges To** - By default, the assigned charge will be used for all spectra with assigned charges and +1, +2, +3 will be used for all CID spectra and +2, +3, +4 for all ETD spectra without assigned charges. The **Apply Charges To** parameter allows the user to override this default setting by instead selecting unassigned spectra.
- o **Charge States** - All comma-separated charges detailed will be applied to each spectrum (based on the values entered into the Charge States box).
- o **Precursor Isotope Off By X**

Precursor Isotope Off by X	No error check ▼
Maximum Precursor Mass	No error check
Precursor and Charge Assignments	Too high (narrow)
Maximum Number of Precursors per Scan	Too high (wide)
Smoothing Width (m/z)	Too high or low (narrow)
Smoothing Width (m/z)	Too high or low (wide)

Figure 99: Precursor Isotope Off By X.

Similarly, on many instruments the nominal precursor mass may actually be the mass of a <sup>13</sup>C isotope peak rather than of the base (all <sup>12</sup>C monoisotopic) peak, so the true precursor mass will within 10 ppm of 2350.120 Da or within 10 ppm of 2351.123 Da. Precursor Isotope Off By X is a pulldown menu with several options.

**No error check** will use only the assigned precursor; **Too high (narrow)** will allow the assigned precursor to be up to 2 Da too high; **Too high (wide)** will allow the assigned precursor to be up to n Da too high for a precursor of mass at least 1000n Da; **Too high or low (narrow)** will allow the assigned precursor to be up to 2 Da too high or 2 Da too low; **Too high or low (wide)** will allow the assigned precursor to be up to n Da too high or 2 Da too low for a precursor of mass at least 1000n Da.

- o **Maximum Precursor Mass** sets the Maximum Precursor Mass to be considered.

- o **Precursor and Charge Assignments**

Precursor and Charge Assignments	Compute from MS1 ▾
Maximum Number of Precursors per Scan	Originally assigned
Smoothing Width (m/z)	Compute from MS1

Figure 100: Precursor and Charge Assignments.

The precursor and charge assignments will be calculated directly from the MS1 data or the originally assigned values will be applied. The user can set this using a drop-down menu.

- o **Maximum Number of Precursors per Scan** - Multiple precursors per scan can also be considered – it is recommended for the user to set this to 2 for complex samples and 5-10 if processing MS<sup>E</sup> or DIA data.
- o **Smoothing Width** - The user can enter a sigma value in Thomsons for Gaussian smoothing and centroiding of Waters or Sciex data. Half-width at peak half maximum (~0.01 m/z) works well and is the default value already entered for the user.

- **Peptide Output Options**

▼ Peptide Output Options	
Automatic Score Cut	✓ Yes
Manual Score Cut	0
Show All N-glycopeptides	— No

Figure 101: Peptide Output Options.

The Peptide Output Options parameters offer options for filtering the peptide-spectrum matches (PSMs) by score. By default, PSM filtering is deferred until after protein ranking, and then filters to control PSM FDR on the “true” proteins—those ranked above the top-ranking decoy protein. This method gains sensitivity while simultaneously reducing both protein and PSM FDRs. See:

Two-dimensional target decoy strategy for shotgun proteomics, Journal of Proteome Research 10 (12), 5296-5301, 2011.

- o **Automatic Score Cut** - To filter PSMs before protein ranking, the user can click **Yes** to activate Automatic score cut and type in a minimum score.
- o **Manual Score Cut** - The user can activate Automatic Score Cut and enter a minimum score. For example, a score threshold of 200 will remove weak matches and a threshold of 400 will remove all but the best matches. Filtering by score may be helpful in special cases, for example to eliminate from consideration all but the best wildcard PSMs.
- o **Show All N-Glycopeptides** - The user also has the option to Show All N-Glycopeptides. This will show N-Glycopeptide matches regardless of score or FDR. This is recommended for simple samples only. This can be especially useful for low energy CID data.

- **Protein Output Options**

▼ Protein Output Options	
Protein FDR	1% FDR (or 20 reverse count)
Create Focused Database	— No
Export mzIdentML	— No

Figure 102: Protein Output Options.

- o **Protein FDR**

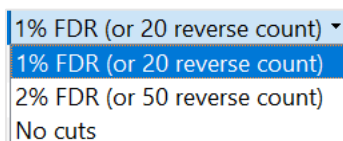


Figure 103: Protein FDR options

This gives the user control of the protein list cut-off. By default, the protein list is cut at 1% protein FDR or 20 decoy proteins, whichever comes last, but the user can ask for 2% protein FDR or a completely unfiltered (but still ranked) protein list, or No Cuts.

- o **Create a Focused Database** - If the user clicks this to select Yes, the software is directed to output a new FASTA file (labeled focused and appearing in the output objs directory) containing only the proteins found in the search, along with suitable decoys (>Reverse) for unbiased FDR estimation. The focused database can then be used for subsequent wide searches, including more modifications and/or a wildcard. Of course, the user can also create focused databases outside of the software by editing existing FASTA files.
- o **Export mzIdentML** - The user can also select to export this file as a mzIdentML file.

- **Multicore Options**

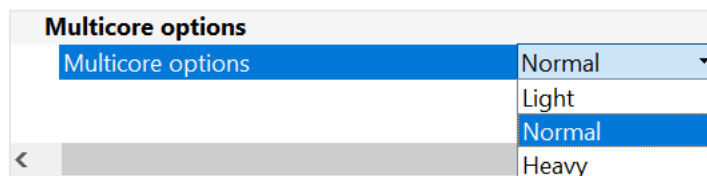


Figure 104: Multicore processing options.

The user can control the number of computer cores of the CPU used through the drop-menu. The Light searches uses one core, Normal search uses all available cores minus two, and Heavy search uses all available cores.

- **Wildcard**

▼ Wildcard	
Wildcard Search	Disabled
Minimum Mass	-40
Maximum Mass	100
Restrict to Residues	

Figure 105: Wildcard parameters.

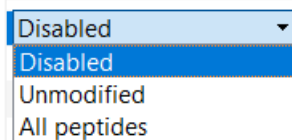


Figure 106: Wildcard Search drop-down menu options.

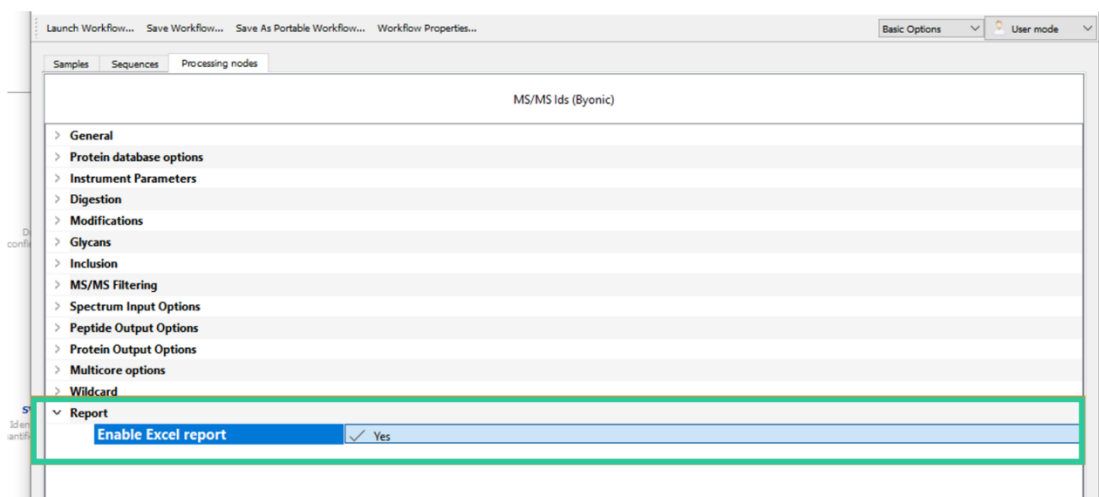
Wildcard lets the user turn on wildcard searches, set the range for the wildcard mass, and restrict the wildcard to certain residues if desired. The options Disabled, Unmodified, and All peptides are available through a drop-down menu.

The **Restrict to residues** box uses the common 20 single-letter residue abbreviations, and (lower case) n denotes peptide N-terminus and (lower case) c denotes peptide C-terminus (e.g. Kn searches lysine and N-terminus). Leaving the field blank searches all residues. A wildcard, even one with a mass range of only 50 or 60 Da, greatly increases the size of the search. It is best used with a focused database (see the Advanced tab section below) and used either alone or with only a few other modifications enabled. Most wildcard mass shifts will be recognizable by an expert; hence, a wildcard can be used to discover which known modifications should be enabled in a subsequent search. For more details about the wildcard search, see the application note “Byonic™: Wildcard Search™” at <https://www.proteinmetrics.com/resources/#application-notes>.

By specifying most modifications as rare, it is quite feasible to search for 10 – 20 modification types at once with Byonic. Even larger searches are possible with focused protein databases, for example with therapeutic proteins. Such a focused database easily allows efficient mutation searches with 200+ possible substitutions, or oxidative footprinting searches with 50+ types of oxidations. Glycans and wildcards can easily enlarge the search space by 2 to 3 orders of magnitude, so these options should be used with care, and in conjunction with only the most common variable modifications (such as oxidized methionine or pyro-Glu N-terminus). **NOTE:** The single most important factor in search time is Total common max. Roughly speaking, the search time grows as  $\frac{C}{T}$  where  $C$  is the number of common modifications enabled and  $T$  is Total common max.

The **Appendix** of the Byonic Manual provides examples of frequently found modifications and appropriate syntax for including those modifications in a Byonic search.

- **Report**



To enable Excel report, the user should go to **Processing nodes > Byonic column > Report** and check the **Enable Excel Report** option.

## Quant

NOTE: As mentioned in the above MS/MS Ids section, these parameters apply alone to the PTM (in-silico) Workflow.

- **General**



Figure 107: Samples to include in the project.

- o **Samples** - The "\*" will apply all parameters to all samples dragged and dropped into the Samples tab.
- **In-silico options (Theoretical Digest)**

<b>&gt; General</b>	
<b>▼ In-silico options (Theoretical Digest)</b>	
Enable in-silico digest	— No
Digestion	Trypsin @ R,K   C-term
Missed Cleavages Max	0
Peptide Minimum Mass	300.00 da
Peptide Maximum Mass	9,000.00 da
Total Common Max	1
Total Rare Max	1
Glycans	
Modification options	
<b>▼ In-silico options (CSV import)</b>	
In-silico peptides CSV	
<b>▼ In-silico options (both)</b>	
Skip if in-silico peptide is duplicate of MS2	✓ Yes

Figure 108: In-silico options processing parameters.

Parameters are separated into three sections to clarify which pertain to a **Theoretical Digest**, **CSV import** and **both**.

- o **Enable In-Silico digest** - The user has the option to click **Yes** to enable this parameter to generate an in-silico list of peptides. The below additional parameters must then be set.

NOTE: The option to add a list of in-silico peptides from a CSV file is offered through In-Silico Peptides CSV.

- o **Disulfide mode (S-S workflow only)** – restricts disulfide bonds to **Free peptides only**, **Disulfide complexes only** and **Both**:

Both ▼
Free peptides only
Disulfide c...plexes only
Both

Figure 109: S-S Disulfide mode options

- o Digestion



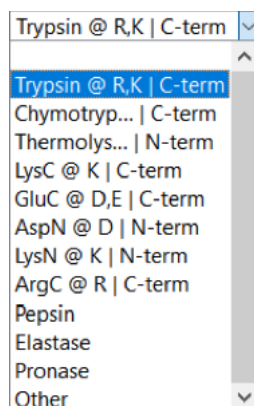


Figure 110: Digestion options.

The user can view several options through the drop-down menu. These are detailed as one letter residue codes adjacent to cleavage points. The user may enter a customized digestion by using the following syntax: Name @ Amino acid letters | C-term or N-term.

- o **Missed Cleavages Max**

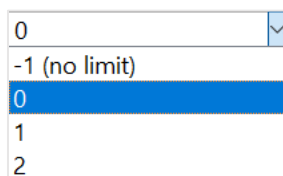
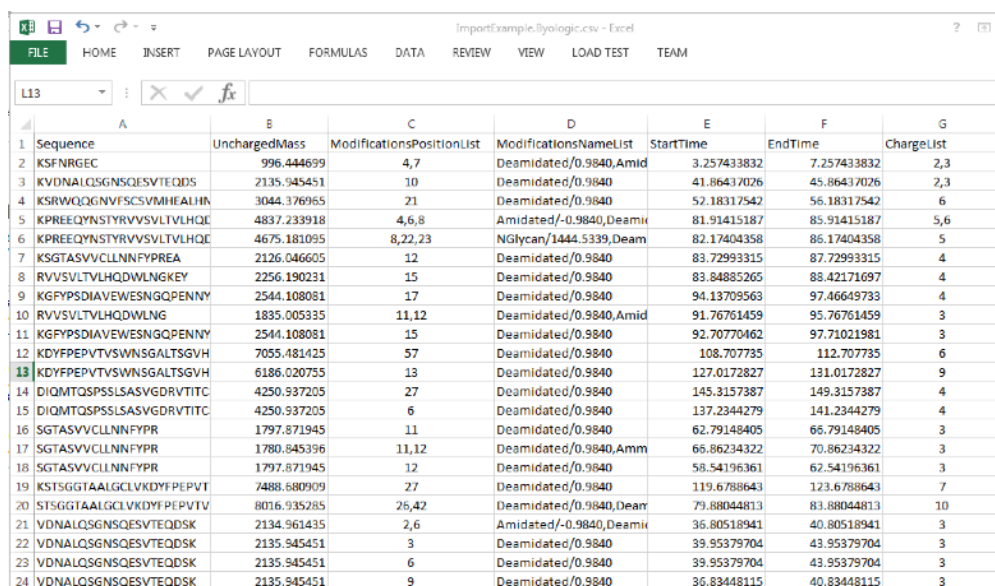


Figure 111: Missed Cleavages Max options.

The user can set the maximum number of missed cleavages per peptide by selecting a value using the drop-down menu. A value of -1 allows any number of cleavages.

- o **Peptide Minimum Mass:** The user can set the minimum value of the peptide mass range.
  - o **Peptide Maximum Mass:** The user can set the maximum value of the peptide mass range.
  - o **Total Common Max:** The user can set the maximum number of common modifications per peptide. The search size grows by an order of magnitude with each increase in Total Common Max.
  - o **Total Rare Max:** The user can set the maximum number of rare modifications per peptide. The search size grows by an order of magnitude with each increase in Total Rare Max.
  - o **Glycans:** The user can specify glycan databases or custom glycans to include in the in-silico generation. This is done the same as detailed in the MS/MS Ids section.
  - o **Modifications Options:** The user can specify modifications to include in the in-silico generation. This is done the same as detailed in the MS/MS Ids section
- **In-Silico Options (CSV Import)**
    - o **In-Silico Peptides CSV:** This option allows the user to add a list of in-silico peptides from an imported CSV file, thus there is no need to run an MS2 search if there are known modifications with known masses and retention times. The format for the CSV is shown below.



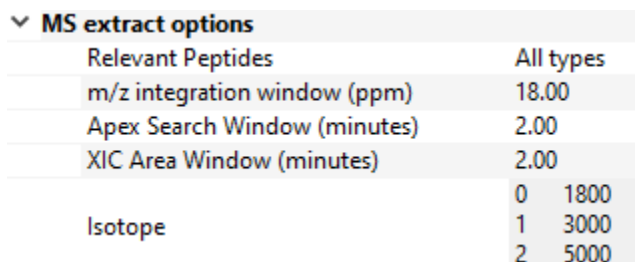
	A	B	C	D	E	F	G
	Sequence	UnchargedMass	ModificationsPositionList	ModificationsNameList	StartTime	EndTime	ChargeList
1	KSFNRGEC	996.444699	4,7	Deamidated/0.9840,Amid	3.257433832	7.257433832	2,3
2	KVDNALQSGNSQESVTEQDS	2135.945451	10	Deamidated/0.9840	41.86437026	45.86437026	2,3
3	KSRWQQGNVFSQSVMIHEALHN	3044.376965	21	Deamidated/0.9840	52.18317542	56.18317542	6
4	KPREEQYNSTYRVVSVLTVLHQE	4837.233918	4,6,8	Amidated/-0.9840,Deamid	81.91415187	85.91415187	5,6
5	KPREEQYNSTYRVVSVLTVLHQE	4675.181095	8,22,23	NGlycan/1444.5339,Deamid	82.17404358	86.17404358	5
6	KSGTASVVCLLNFFYPREA	2126.046605	12	Deamidated/0.9840	83.72993315	87.72993315	4
7	RVVSVLTVLHQDWLNGKEY	2256.190231	15	Deamidated/0.9840	83.84885265	88.42171697	4
8	KGFYPDSIAVEVESNGQPENNY	2544.108081	17	Deamidated/0.9840	94.13709563	97.46649733	4
9	RVVSVLTVLHQDWLNG	1835.005335	11,12	Deamidated/0.9840,Amid	91.76761459	95.76761459	3
10	KGFYPDSIAVEVESNGQPENNY	2544.108081	15	Deamidated/0.9840	92.70770462	97.71021981	3
11	KDYFPEPVTWSWNSGALTSVGH	7055.481425	57	Deamidated/0.9840	108.707735	112.707735	6
12	KDYFPEPVTWSWNSGALTSVGH	6186.020755	13	Deamidated/0.9840	127.0172827	131.0172827	9
13	KDYFPEPVTWSWNSGALTSVGH	6186.020755	13	Deamidated/0.9840	127.0172827	131.0172827	9
14	DIQMTQSPSSLSASVGDRVTITC	4250.937205	27	Deamidated/0.9840	145.3157387	149.3157387	4
15	DIQMTQSPSSLSASVGDRVTITC	4250.937205	6	Deamidated/0.9840	137.2344279	141.2344279	4
16	SGTASVVCCLNFFYP	1797.871945	11	Deamidated/0.9840	62.79148405	66.79148405	3
17	SGTASVVCCLNFFYP	1780.845396	11,12	Deamidated/0.9840,Amid	66.86234322	70.86234322	3
18	SGTASVVCCLNFFYP	1797.871945	12	Deamidated/0.9840	58.54196361	62.54196361	3
19	KSTSGGTAALGCLVKDYFPEPVT	7488.680909	27	Deamidated/0.9840	119.6788643	123.6788643	7
20	STSGGTAALGCLVKDYFPEPVT	8016.935285	26,42	Deamidated/0.9840,Deamid	79.88044813	83.88044813	10
21	VDNALQSGNSQESVTEQDSK	2134.961435	2,6	Amidated/-0.9840,Deamid	36.80518941	40.80518941	3
22	VDNALQSGNSQESVTEQDSK	2135.945451	3	Deamidated/0.9840	39.95379704	43.95379704	3
23	VDNALQSGNSQESVTEQDSK	2135.945451	6	Deamidated/0.9840	39.95379704	43.95379704	3
24	VDNALQSGNSQESVTEQDSK	2135.945451	9	Deamidated/0.9840	36.83448115	40.83448115	3

Figure 112: Format for an In-Silico Peptides CSV.

- In-Silico Options (Both)

- **Skip if in-silico peptide is duplicate of MS2** - The user can click to select **Yes**. This will skip any generated in-silico peptide if one already exists from the MS2 data. It is recommended to make this selection when generating in-silico peptides.

- MS extract options

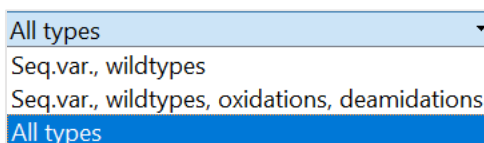


MS extract options	
Relevant Peptides	All types
m/z integration window (ppm)	18.00
Apex Search Window (minutes)	2.00
XIC Area Window (minutes)	2.00
Isotope	0 1800
	1 3000
	2 5000

Figure 113: MS extract options.

There are various aspects to the XIC of a peptide which can help in distinguishing a true from false identification, or whether the peptide hit is relevant based on intensity.

- **Relevant Peptides**



Relevant Peptides
All types
Seq.var., wildtypes
Seq.var., wildtypes, oxidations, deamidations
All types

Figure 114: Relevant Peptide options.

The user can select what types of peptides to consider relevant using the drop-down menu. Only peptides of interest will be considered for extracting additional information, such as XIC, isotope profile, and MS2 fragments. This pulldown menu option may affect the file size significantly. For SVA, the user may only need to choose Sequence variants and wildtypes. For peptide mapping, the user should choose All types.

- o **m/z integration window (ppm)** - The user can define the maximum m/z error tolerance to be applied to the search.
- o **Apex Search Window (minutes)** - The user can automatically define the search window for the peak apex.
- o **XIC Area Window (minutes)** - The user can automatically set the time window for integration. The defined time limits will be visible as two vertical lines for each XIC in the project. These lines can be dragged by the user's mouse to adjust the integration time if needed.
- o **Isotope**

XIC Area Window (minutes)	2.00
Isotope	0 1800
	1 3000
	2 5000

Figure 115: Isotope settings.

The user can set the mass ranges for the offset from the monoisotopic mass. Click within the box to activate the light blue “...” button and view the default values. These values can be modified by typing in values.

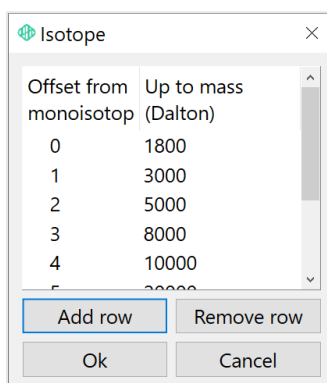


Figure 116: Isotope settings

### • Advanced

Advanced	
Enable Lock-Mass Calibration	No
Lock Mass (m/z)	1221.990637
Lock Mass tolerance (ppm)	100.00
Centroid Smoothing Width	0.02
Elution Prediction Score Min	300
Compute Fragment Coverage	Yes
Advanced configuration	

Figure 117: Advanced options

- o **Enable Lock-Mass Calibration** - The user can click to select between **No** and **Yes**.
- o **Lock Mass (m/z)**

1221.990637	▼
556.2771	
609.2812	
785.84206	
121.050873	
322.048121	
622.028960	
922.009798	
1221.990637	
1521.971475	

Figure 118: Lock mass m/z value options

The user can enter the calibrant m/z value. Several are also available using the drop-down menu or the user can type in a numerical value. If empty, no calibration will be applied.

- **Lock Mass tolerance (ppm)** - The user can enter the calibrant m/z mass tolerance value in ppm. The user can type in a numerical value or use the up and down arrow keys to increase or decrease the value. If empty, no calibration will be applied.
- **Centroid Smoothing Width** - The user can enter a sigma value in Thomsons for Gaussian smoothing and centroiding of Waters or Sciex data. Half-width at peak half maximum (~0.01 m/z) works well and is the default value already entered for the user.
- **Elution Prediction Score Min** - The user can set a value for this filter to find wildtype peptides with scores above this value to adjust the elution time.
- **Compute Fragment Coverage** - The user can click between **Yes** and **No** to compute fragmentation coverage at project creation.
- **Advanced Configuration** - The user can enter text commands to complete advanced processing. These are detailed in the Release Notes included with each quarterly release. Please reach out to [support@proteinmetrics.com](mailto:support@proteinmetrics.com) for additional details.

- **Report**

See [Report](#)

- **UI Configurations**

See [UI Configurations](#)

- **Feature Finder**

▼ Feature Finder	
Enable Feature Finder	— No
Minimum Isotope Corr.	0.90
Mass Range Min	500.00
Mass Range Max	8,000.00
Maximum Feature Count	300
Absolute minimum intensity	0.00
Minimum Isotope Count	3
Minimum Peak Width (min)	0.01
Minimum S/N Ratio	3
Minimum Scan Count	3
Exclude +1 Charge Only Features	— No
Exclude Features With MS2 Matches	— No
Mass matching tolerance (ppm)	20.00
Time matching tolerance (min)	0.20

Figure 119: Feature Finder

**Feature Finder** is an algorithm that allows the user to scan the MS1 domain for all existent peptides in a sample. In a typical data dependent acquisition assay, peptides are identified only if an MS2 scan is triggered from a precursor ion in an MS1 scan. Therefore, if a peptide ion does not trigger an MS2 scan, or its score excludes it from identification, it will not be detected. Feature Finder ensures that a peak will not go undetected by identifying all possible isotopic distributions in a sample that could potentially be a peptide.

Feature Finder also allows the user to match these features to the current identifications and to add a specified number of the unidentified features as unknowns to the project. Doing so will generate XICs for the unknowns and allow reporting on features such as retention time, mass, intensity, and fold change across samples.

- **Enable Feature Finder** - The user can click to select between **No** and **Yes** to enable Feature Finder.
- **Minimum Isotope Corr.** - The user can set the minimum cutoff to match the found isotopic distribution to the theoretical distribution using Pearson's correlation. If a feature has an isotopic distribution that is vastly different from its theoretical, this can be used to exclude it.
- **Mass Range Min** - The user can set the minimum mass for a feature to be considered.
- **Mass Range Max** - The user can set the maximum mass for a feature to be considered.
- **Maximum Features Count** - The user can set the maximum feature count to include the most intense features.
- **Absolute minimum intensity** - The user can set the minimum intensity for a feature to be considered.
- **Minimum Isotope Count** - The user can set the minimum number of isotope peaks required for a feature to be considered. This applies only if its charge is greater than 2.
- **Minimum Peak Width (min)** - The user can set the minimum peak width in minutes for a feature to be considered.
- **Minimum S/N Ratio** - The user can set the minimum S/N ratio for a feature to be considered.
- **Minimum Scan Count** - The user can set the minimum scans required across the peak for a feature to be considered.
- **Exclude +1 Charge Only Features** - The user can click to select between **Yes** and **No** to exclude singly charged features.
- **Exclude Features With MS2 Matches** - The user can click to select between **No** and **Yes** to exclude features with MS2 matches. Exclude will remove the feature if it coincides within the time and mass tolerances set – it will not be included in the top X features to avoid duplicates. It is recommended to check this box for projects containing MS2 data and Byonic searches.
- **Mass matching tolerance (ppm) and Time matching tolerance (min)** - These parameters are used to match and remove features that have already been identified by MS2 (when **Exclude Features With MS2 Matches** is set to Yes) so that they do not contribute to the maximum feature count. They help to match the retention time and mass of an unknown feature to a MS2 feature so that it may be excluded from the Peptides table in Peptide Analysis.

- **XIC**

**XIC** refers to the “Extracted Ion Chromatogram”, which represents the signal intensity of an ion or ions of interest over time extracted from the full mass spectrometry dataset. Fast XIC accelerates XIC processing.

XIC	
Enable Fast XIC	No
Granularity	100,000
Max Mz	3,010
Max Ion Count	3,000

Figure 120: Fast XIC parameters

If **Enable Fast XIC** is set to **Yes**, the below options are available:

Parameter	Description
Granularity	Integer value, relating to the precision. 10000 is 4 places past the decimal. For mass 1000.1236, a granularity of 1000 sets the value to 1000.124, while a granularity of 10000 keeps the value at 1000.1236. Default value is 100000.
Max Mz	Integer value. The maximum m/z value during data acquisition. To acquire data in the range 200-2000, set Max Mz = 2000, so that no ions are acquired past that value. Default value is 3010
Max Ion Count	Integer value. The maximum number of ions to be considered, sorted by intensity. Some files may contain ~2500 ions per scan. Max Ion Count = 1000 uses only the top 1000 ions by intensity. This is especially important for TOFs with ~25K ions per scan, where a higher value of 2500 is recommended. Default value is 3000.

## HDX Workflow

The parameters for the **HDX** workflow are the same as for Peptide Analysis workflows. Please see the section immediately above for details about the **Processing nodes** tab.

## Comparison Chromatography, Comparison Chromatography (in-silico), Reference Chromatography, Reference Chromatography (in-silico) Workflows

The following section details parameters for the combination of MS/MS IDs and Chromatographic Quant. They are therefore relevant to the **Reference Chromatography**, and **Characterization Chromatography** workflows.

## Processing nodes

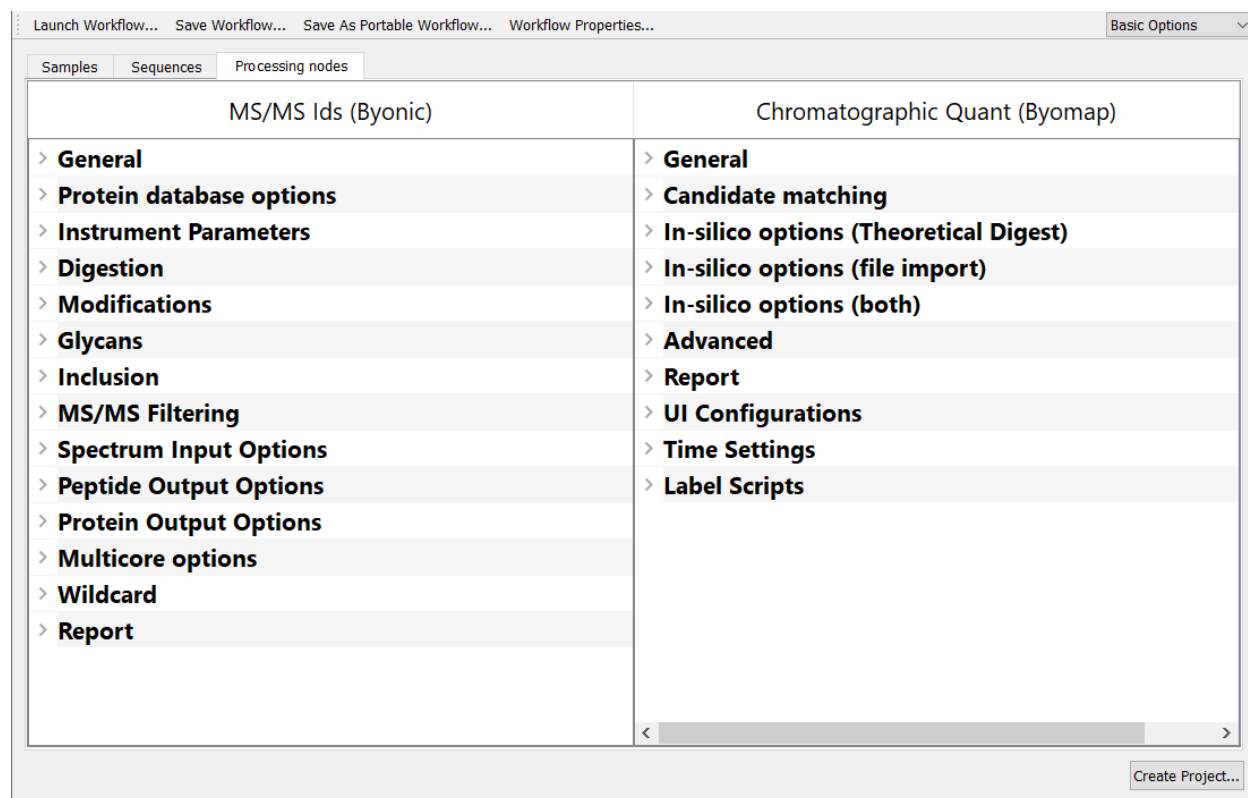


Figure 121: Processing nodes parameters – Reference Chromatography

### MS/MS Ids

NOTE: These parameters are not part of the Reference Chromatography (in-silico) or Characterization Chromatography (in-silico) workflows

- **General**  
See [General](#)
- **Protein database options**  
See [Protein database options](#)
- **Instrument Parameters**  
See [Instrument Parameters](#)
- **Digestion**  
See [Digestion](#)
- **Modifications**  
See [Modifications](#)
- **Glycans**  
See [Glycans](#)
- **Spectrum Input Options**

See [Spectrum Input Options](#)

- **Peptide Output Options**

See [Peptide Output Options](#)

- **Protein Output Options**

See [Protein Output Options](#)

- **Multicore Options**

See [Multicore Options](#)

- **Wildcard**

See [Wildcard](#)

## Chromatographic Quant

NOTE: As mentioned in the above MS/MS Ions section, these parameters apply alone to the Reference Chromatography (in-silico) and Characterization Chromatography (in-silico) workflows.

- **General**

See [General](#)

- **Candidate Matching**

Candidate matching	
Mass assignments	Auto candidate assignment
Minimum charge	1.00
Maximum charge	5.00
Tolerance (ppm)	10.00
Minimum peak intensity percent	10.00
Minimum m/z	250.00

Figure 122: Candidate matching

These parameters allow the user to set mass assignment, charge range and other candidate matching filters before project creation.

- **In-silico options (Theoretical Digest)**

See [In-silico Options \(Theoretical Digest\)](#)

- **In-Silico options (CSV import)**

See [In-Silico options \(CSV import\)](#)

- **In-Silico Options (Both)**

See [In-Silico Options \(Both\)](#)

- **Advanced**

See [Advanced](#)

- **Report**

See [Report](#)

- **UI Configurations**

See [UI Configurations](#)



- **Time Settings**

See [Time Settings](#)

- **Label Scripts**

See [Label Scripts](#)

- **Peak Construction options**

See [Peak Construction options](#)

(Pertains to older versions of Byos)

## Multi-Protein Quantitation, Multi-Protein Identification, Multi-Protein Quantitation

There are four **Multi-Protein Analysis** workflows available:

- Multi-Protein Quantitation
- Multi-Protein Quantitation-DIA
- Multi-Protein Preview
- Multi-Protein Identification

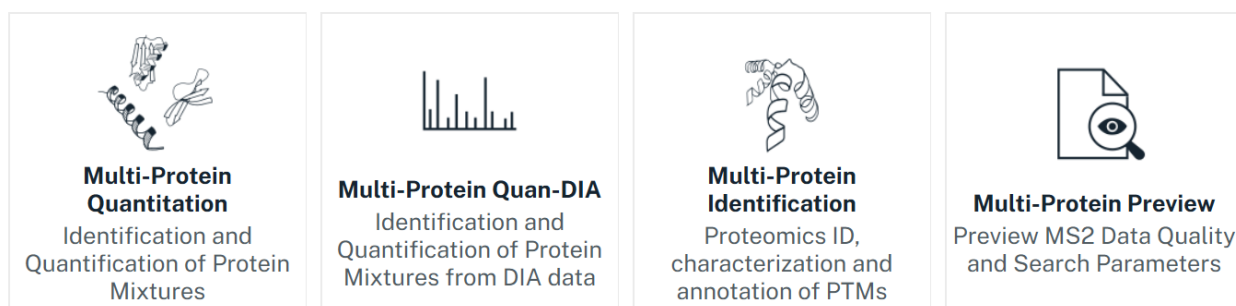


Figure 123: Multiprotein workflow icons

## Applications

- Users who deal with Host Cell proteins from Bioreactors in the early phases of development
- Users who have to deal with mixtures of proteins in a complex matrix
- Users who need to investigate protein mixtures and quantify proteins in that mixture
- Users who need to follow specific proteins in a complex mixture (tens of thousands of proteins) across each step of the purification process

## Multi-Protein Quantitation

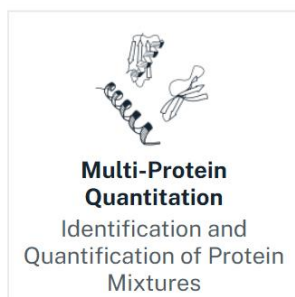
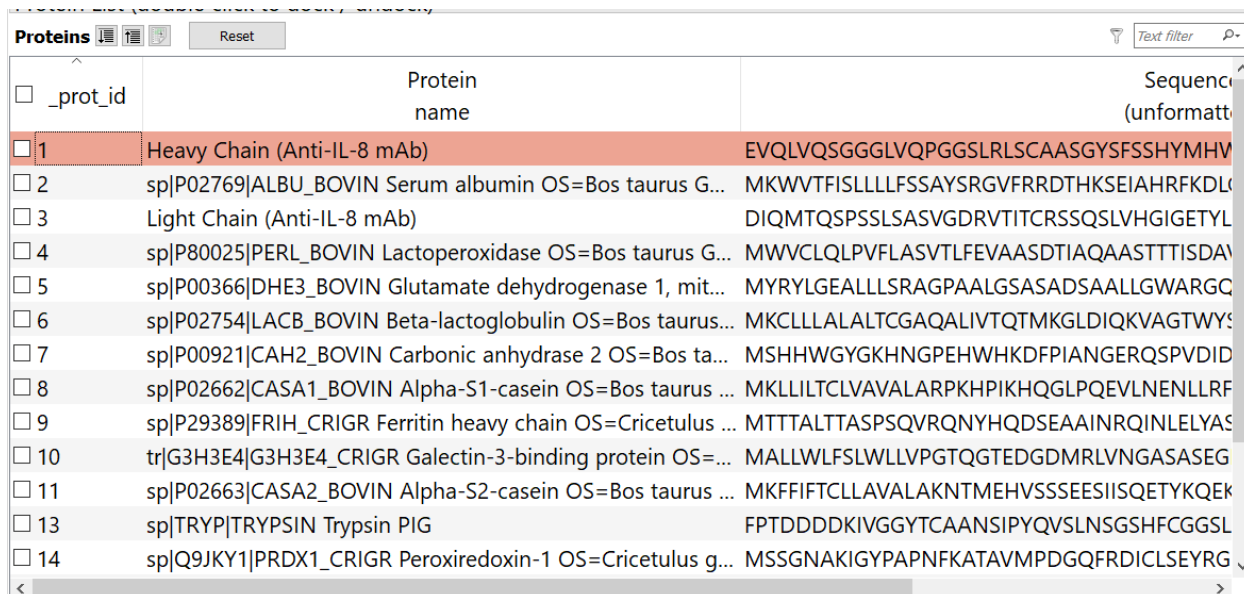


Figure 124: Multi-Protein Quantitation workflow icon

The **Multi-Protein Quantitation** workflow provides a mass spectrometry-based assessment of many proteins at the same time, with features such as:

- **Protein Quantitation:** Allows the user to provide Label-Free quantitation without the need to cope with complex sample preparation by using information from precursor extracted ion chromatograms (XICs)
- **High-throughput analysis:** Allows users to study datasets with a large number of proteins (tens of thousands) and quantify proteins from a subset of a database searched (thousands of proteins) at a time
- **Rapid Results:** Automatically curates peptide-level data from enzymatic digests to provide quantitative information rapidly, automatically, and report relative quantitation levels
- **Enhanced Reporting:** The interface allows users to rapidly review data generated, as well as interact with the data to produce templated, or customized reports...
- **New algorithms:** The Multi-Protein Quantitation workflow can automate multiple tasks, such as:
  - Check which proteins are quantified across an entire sample set
  - Match between runs to ensure that the same peptides are used for quantitation
  - Choose peptides that are consistent across a sample set
  - Automatically curate raw data to ensure that appropriate peptides are used
  - Calibrate raw data and suggest parameters to use, and use them automatically

An additional tool available to users in the Multi-Protein Quantitation workflow is **Proteins** view. This view enables the user to limit the number of displayed peptides in the Peptide table, which can be useful in improving performance when processing large projects. Clicking on a given protein in the Proteins view will cause the Peptides view to load only the Peptides relating to that protein, starting with the peptides with the top number of hits. Proteins not referenced in the Peptides table in the database are hidden in the Proteins view. Users can select the number of peptides they wish to see in the Peptides table.



<input type="checkbox"/> _prot_id	Protein name	Sequence (unformatted)
<input checked="" type="checkbox"/> 1	Heavy Chain (Anti-IL-8 mAb)	EVQLVQSGGGLVQPGGSLRLSCAASGYSFSSHYMHV
<input type="checkbox"/> 2	sp P02769 ALBU_BOVIN Serum albumin OS=Bos taurus G...	MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDL
<input type="checkbox"/> 3	Light Chain (Anti-IL-8 mAb)	DIQMTQSPSSLSASVGDRTITCRSSQSLVHGIGETYL
<input type="checkbox"/> 4	sp P80025 PERL_BOVIN Lactoperoxidase OS=Bos taurus G...	MWVCLQLPVFLASVTLFEVAASDTIAQAASTTTISDAV
<input type="checkbox"/> 5	sp P00366 DHE3_BOVIN Glutamate dehydrogenase 1, mit...	MYRYLGEALLSRAGPAALGSASADSAALLGWARGC
<input type="checkbox"/> 6	sp P02754 LACB_BOVIN Beta-lactoglobulin OS=Bos taurus...	MKCLLLALALTCAQALIVTQTMKGLDIQKVAGTWYS
<input type="checkbox"/> 7	sp P00921 CAH2_BOVIN Carbonic anhydrase 2 OS=Bos ta...	MSHHWGYGKHNGPEHWHKDFPIANGERQSPVDID
<input type="checkbox"/> 8	sp P02662 CASA1_BOVIN Alpha-S1-casein OS=Bos taurus ...	MKLLILTCLVAVALARPKHKHQGLPQEVNLNENLLRF
<input type="checkbox"/> 9	sp P29389 FRIH_CRIGR Ferritin heavy chain OS=Cricetulus ...	MTTALTASPSQVRQNYHQDSEAAINRQINLELYAS
<input type="checkbox"/> 10	tr G3H3E4 G3H3E4_CRIGR Galectin-3-binding protein OS=...	MALLWLFSLWLLVPGTQGTEDGDMRLVNGASASEG
<input type="checkbox"/> 11	sp P02663 CASA2_BOVIN Alpha-S2-casein OS=Bos taurus ...	MKFFIFTCLLAVALAKNTMEHVSSEESIISQETKQEK
<input type="checkbox"/> 13	sp TRYP TRYPSIN Trypsin PIG	FPTDDDDKIVGGYTCAANSIPYQVSLNSGSHFCGGSL
<input type="checkbox"/> 14	sp Q9JKY1 PRDX1_CRIGR Peroxiredoxin-1 OS=Cricetulus g...	MSSGNAKIGYPAPNFKATAVMPDQGFRDCLSEYRG

Figure 125: Protein View

The Peptides and Proteins view can be reset to default, which resets filtering by proteins and loads all peptides, by pressing the Reset button:

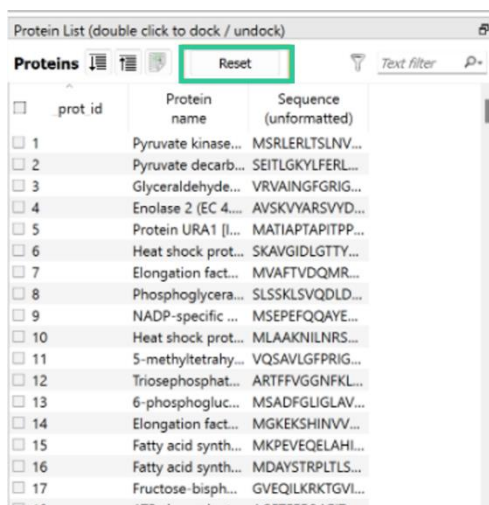


Figure 126: Reset Proteins view

Note: Be sure to 'Reset' the view if you desire to include all proteins in the generated report. It is recommended that users performing analysis with any of the Multi-Protein workflows use the High Performance Computing options specified in the [System Requirements](#).

## Processing nodes

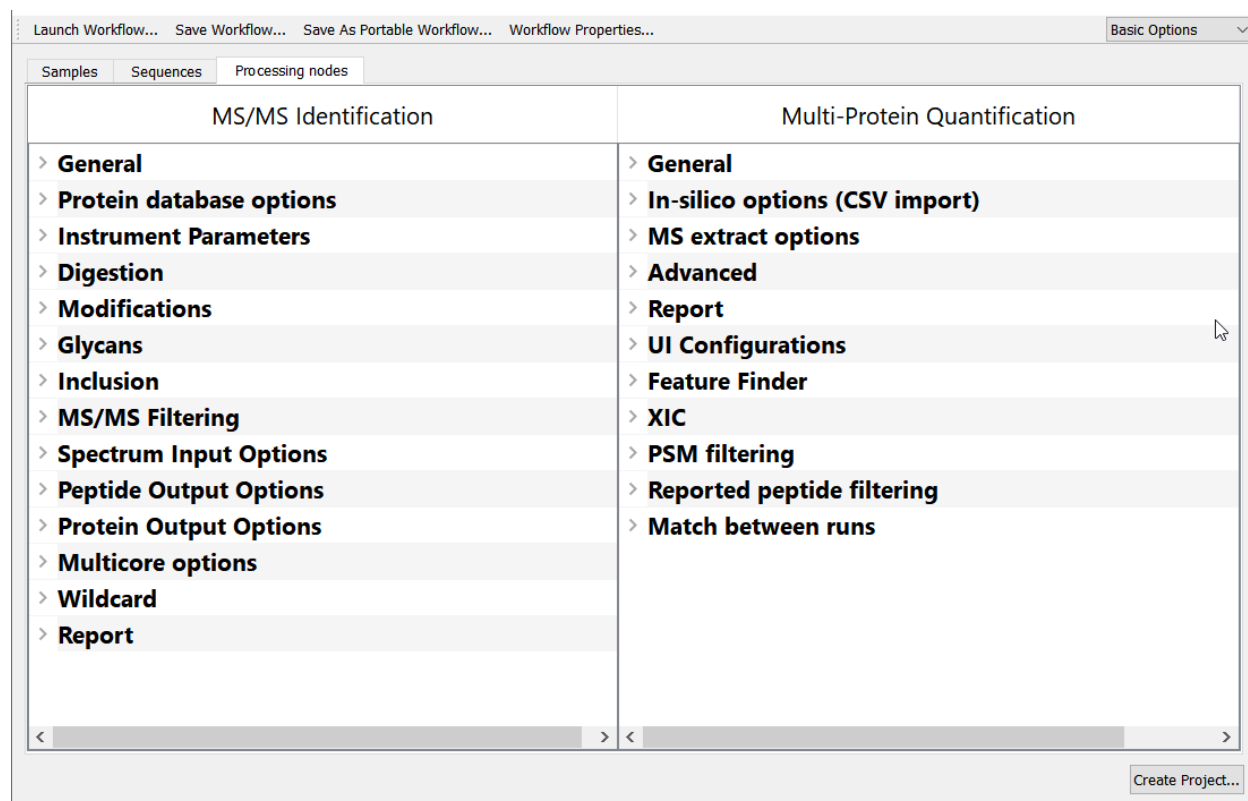


Figure 127: Processing nodes

## MS/MS Identification

- **General**

See [General](#)

- **Protein database options**

See [Protein database options](#)

- **Instrument Parameters**

See [Instrument Parameters](#)

- **Digestion**

See [Digestion](#)

- **Modifications**

For thousands of proteins in biopharmaceutical settings, a default set has been chosen. In some cases, the user may wish to optimize this for different types of organism or for plasma-proteins. For plasma proteins, glycosylation should be considered.

See [Modifications](#)

- **Glycans**

See [Glycans](#)

- **Inclusion** is used to import a \*.csv file which defines m/z ratio ranges and/or elution time range segments. **Inclusion list** allows the user to add/edit/delete the following values brought in via csv: m/z begin, m/z end, Elution time begin, Elution time end.

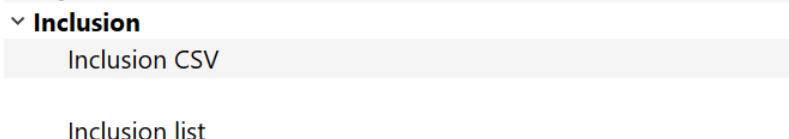


Figure 128: Inclusion

- **MS/MS Filtering**

See [MS/MS Filtering](#)

- **Spectrum Input Options**

See [Spectrum Input Options](#)

- **Peptide Output Options**

See [Peptide Output Options](#)

- **Protein Output Options**

See [Protein Output Options](#)

- **Multicore Options**

See [Multicore Options](#)

- **Wildcard**

See [Wildcard](#)

- **Report**

See [Report](#)

## Multi-Protein Quantification

- **General**

See [General](#)

- **In-Silico options (CSV import)**

See [In-Silico options \(CSV import\)](#)

- **MS extract options**

See [MS extract options](#)

- **Advanced**

See [Advanced](#)

Note: Autodecoys for the Multiprotein Quantitation workflow is = 100.

- **Report**

See [Report](#)

- **UI Configurations**

See [UI Configurations](#)

- **Feature Finder**

See [Feature Finder](#)

- **XIC**

See [XIC](#)

- **PSM filtering:** Pre-filtering of PSMs allows the user to quantify proteins based upon the peptide features that you designate. These filter settings cannot be adjusted once the project creation process has completed. The Multi-Protein Quantitation workflow is the only workflow that permits this type of filtering.

▼ <b>PSM filtering</b>	
Minimum peptide length	6
Maximum peptide length	32
Maximum missed cleavage count	2
Minimum peptide score	80.00
Minimum peptide replicate count	2
Keep homologous sequences	✓ Yes
Minimum peptide matching a protein count	2
Include modifications	None

Figure 129: PSM filtering

- **Minimum peptide length:** A minimum length of six residues helps ensure minimal likelihood of homologous sequences being selected for quantifying proteins
- **Maximum peptide length:** The maximum length was selected to ensure more confident peptide matches
- **Maximum missed cleavage count:** Ensures that peptides identified as a result of poor enzymatic efficiency are not used for quantifying proteins
- **Minimum peptide score:** This minimum score should be modified depending upon your specific application and mass spectrometer
- **Minimum peptide replicate count:** To ensure a more-likely peptide identification, the peptide must be found in at least two input raw files
- **Keep homologous sequences:** If your sequence FASTA database contains very low sequence homology in the representative proteins, it might be beneficial to exclude homologous sequences when quantifying proteins (in which the parameter should be changed to 'No')
- **Minimum peptide matching a protein count:** To preclude the occurrence of “one hit wonder” protein identifications, it is best to require a minimum of at least two peptides to be found for identifying a protein
- **Include modifications:** The most robust quantitation is generally achieved by excluding modified proteins for the purpose of quantification
- **Reported peptide filtering:** Allows the user to select the reported peptide filter set that best fits project needs for data review. These filters can be modified after opening the project result file.

▼ <b>Reported peptide filtering</b>	
Retain only primary MS2 peptide ID	✓ Yes
Use only top charge state for quantification	✓ Yes
Use the same top N peptides for quantification	✓ Yes
Top N count	3
Apply RSD limit for replicates	— No
Maximum acceptable %RSD for replicates	30.00

Figure 130: Reported peptide filtering

- **Retain only primary MS2 peptide ID:** Will 'save' only the highest scoring MS2 spectral match for confirmation purposes
- **Use only top charge state for quantification:** Ensures that only one charge state per peptide sequence will be used for quantifying the protein
- **Use the same top N peptides for quantification:** Ensures that the same peptide is used from each input raw file for the purpose of quantifying proteins, ensuring a more robust quantification
- **Top N count:** This parameter enables the user to specify the number of peptides they wish to use for quantifying proteins in their experiment
- **Apply RSD limit for replicates:** This parameter is useful if sample replicates produce very reproducible XICs to ensure that random peptide matches are not used for the purpose of quantifying proteins

- **Maximum acceptable %RSD for replicates:** This parameter is used by the application only if the “Apply RSD limit for replicates” is also used, and the value should be adjusted based upon the reproducibility of your particular HPLC/LC-MS/MS system
- **Match between runs:** Enabling Match between runs causes the software to search for MS1 peaks and create extracted ion chromatograms (XICs) in the sample replicates which did not already have a MS2 verified peptide identification.

▼ <b>Match between runs</b>	
Enable match between runs	✓ Yes
Mode	Between all samples
Enable sample-to-sample XIC alignment	✓ Yes

Figure 131: Match between runs

In each individual sample, some identifications could be missed due to:

- The ion being skipped in the experiment
- Only MS1 data available for some peptides
- Matrix effects swamping the detector

This feature tells the software to fill in gaps using the XIC of the precursor ion that matches the same identifications in the other samples and/or replicates of the same project.

The **Mode** option determines whether Match Between Runs is run between all samples (default) or within each condition.

As of Byos v5.9, a new **sample-to-sample XIC alignment** parameter is now enabled by default to improve XIC RT bounding for in-silico peaks added by the Match Between Runs parameter. Disabling this feature could reduce overall processing time by 10% or more, but the XIC RT bounding improvements will not be applied.

## Multi-Protein Identification

The new **Multi-Protein Identification** workflow has been optimized for analyzing mass spectrometry data which is expected to have thousands of proteins, and tens of thousands of peptide identifications. This analysis workflow is available from the default System Workflows:

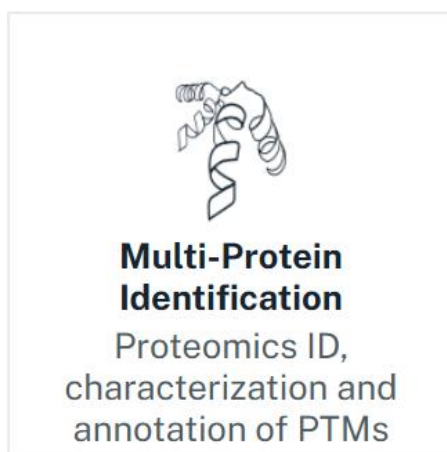


Figure 132: Multi-Protein Identification workflow icon

## Processing Nodes

The Multi-Protein Identification workflow consists of a unique processing node dedicated to the task.

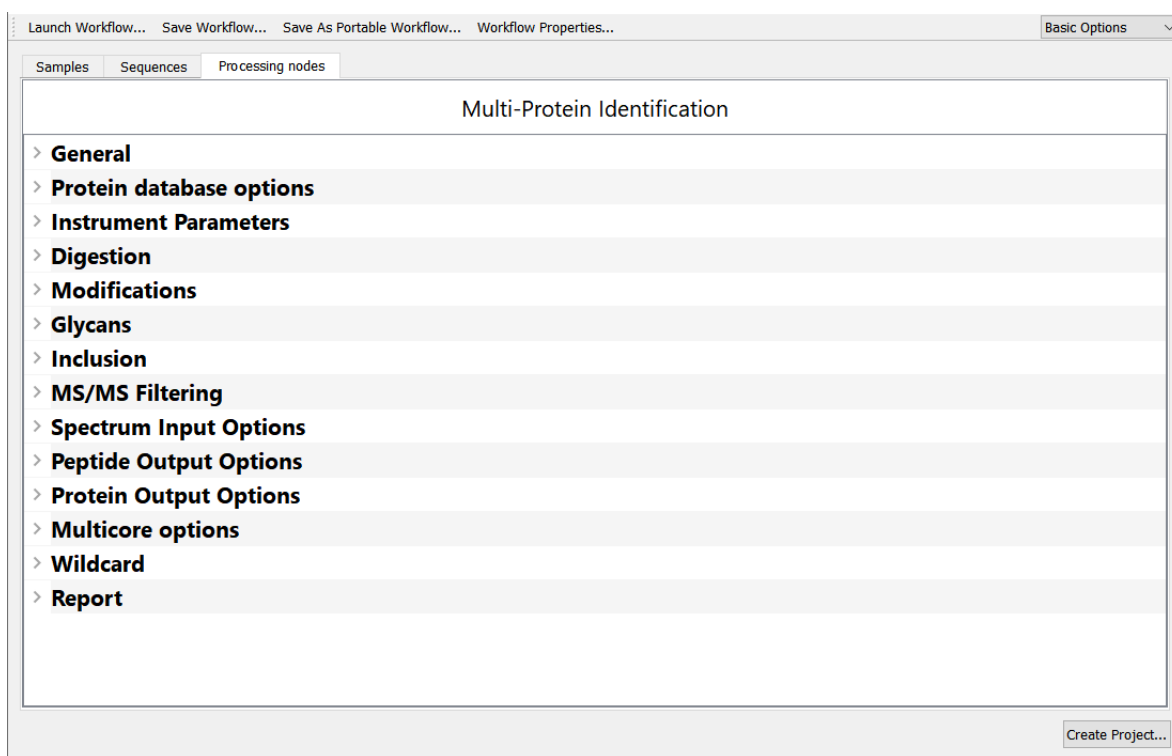


Figure 133: Multi-Protein Identification processing node

### Multi-Protein Identification

- **General**  
See [General](#)
- **Protein database options**  
See [Protein database options](#)
- **Instrument Parameters**

Instrument Parameters		
	Precursor Mass Tolerance	6.00 ppm
	Fragmentation Type	QTOF / HCD
	Fragment Mass Tolerance 1	20.00 ppm
	Fragment Mass Tolerance 2	20.00 ppm
	Recalibration (lock mass)	None

Figure 134: Instrument Parameters



- o **Precursor Mass Tolerance:** The user can change this value by clicking within the text box and then clicking on the activated blue “...” square. The user can then modify the text value and mass accuracy, as required. .

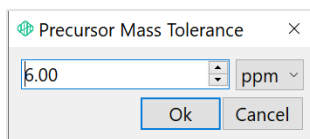


Figure 135: Precursor Mass Tolerance dialog

- o **Fragmentation Type**

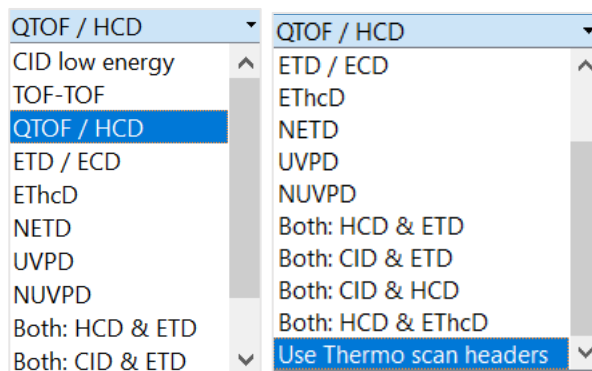


Figure 136: Fragmentation type.

The user can select from the available options using the drop-down menu. Additional options are visible using the scroll bar, including “Both:” for spectrum file(s) containing more than one fragmentation type and “Use Thermo scan headers” to read directly from Thermo raw data files.

- o **Fragment Mass Tolerance 1-** The user can set the value used to acquire data in either ppm or Da.
- o **Fragment Mass Tolerance 2 -** If necessary, the user can set the value used to acquire data in either ppm or Da. This is only required for the “Both:” Fragmentation Types (selected as shown in Figure 74 above – the final entries). The user can set the value used to acquire data in either ppm or Da. No value will be applied if only a value for “Fragmentation Mass Tolerance 1” was entered.
- o **Recalibration (lock mass)**

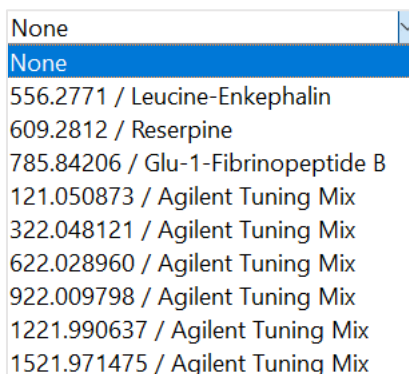


Figure 137: Lock mass calibration options.

The user can select from the available options using the drop-down menu.

- **Digestion**  
See [Digestion](#)
- **Modifications**  
See [Modifications](#)
- **Glycans**  
See [Glycans](#)
- **Inclusion**  
See [Inclusion](#)
- **MS/MS Filtering**  
See [MS/MS Filtering](#)
- **Spectrum Input Options**  
See [Spectrum Input Options](#)
- **Peptide Output Options**  
See [Peptide Output Options](#)
- **Protein Output Options**  
See [Protein Output Options](#)
- **Multicore Options**  
See [Multicore Options](#)
- **Wildcard**  
See [Wildcard](#)
- **Report**  
See [Report](#)

Note: The Multi-protein Identification report will not be generated by default as of Byos v5.9. Users should click on the **Generate Report** icon to create a report manually once the Project has been loaded.

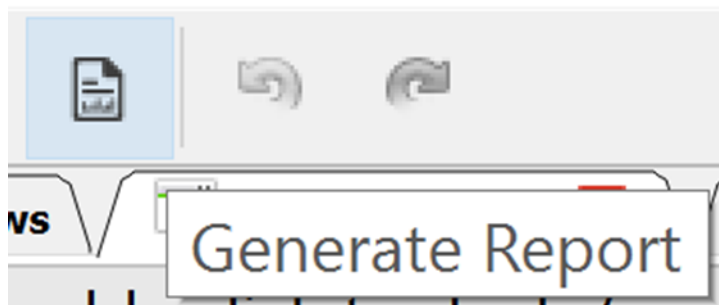


Figure 138: Generate report manually

## Multi-Protein Quantitation Data-Independent Acquisition (DIA)

The Multi-Protein Quantitation-DIA workflow, available with Multi-Protein Quantitation workflow license, enables the processing of data-independent acquisition (DIA) mass spec data via the MPQ pathway.

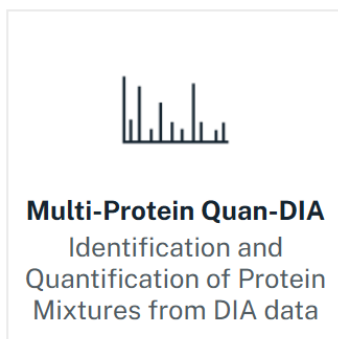


Figure 139: Multi-Protein Quantitation DIA icon

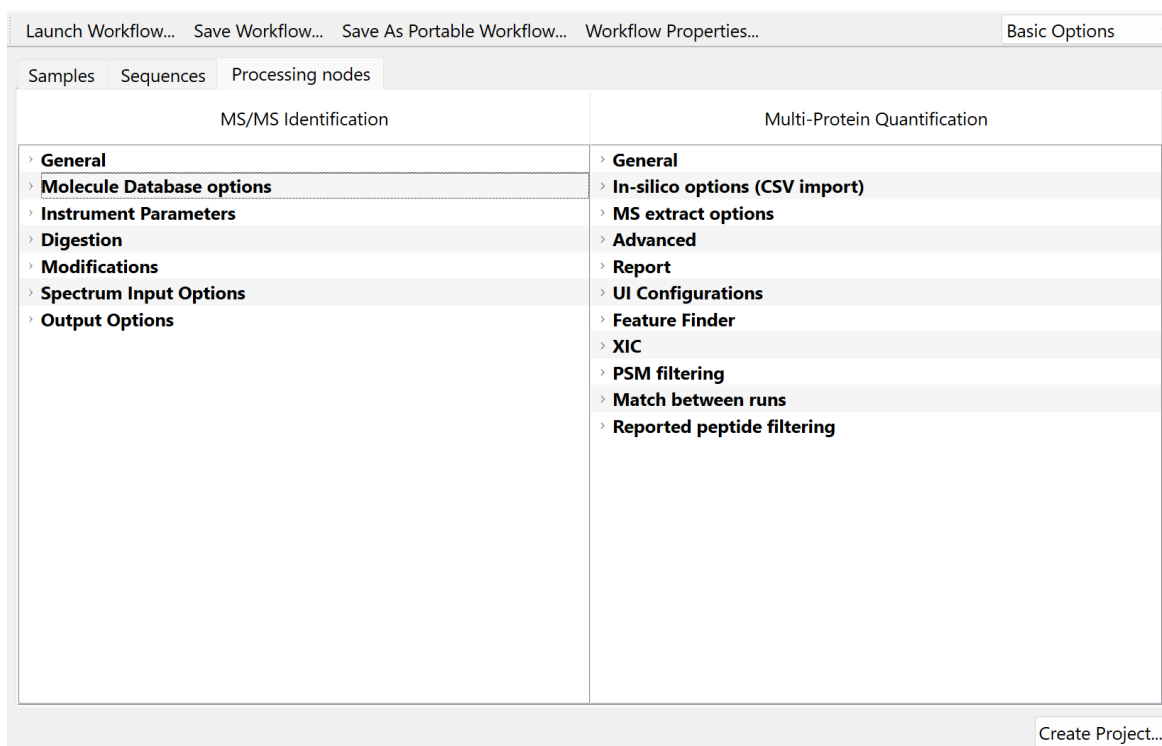


Figure 140: Processing nodes for the MPQ DIA workflow

### Processing nodes

#### MS/MS Identification

- **General**  
See [General](#)
- **Molecule database options**

**Generate Decoys:** Default value is 'Yes'; can be set to 'No' if you are searching a single protein sequence. Decoys are created not by reversing sequences but by mutating the target sequence.

- **Instrument Parameters**

See [Instrument Parameters](#). Note that tolerances are slightly wider than typically used with Byonic, with default values of 10 for **Precursor** and 15 for **Fragment mass tolerance**. If tolerances are too narrow (lower than expected scatter) the score will drop and matches will decrease.

- **Digestion**

See [Digestion](#)

- **Modifications**

For thousands of proteins in biopharmaceutical settings, a default set has been chosen. In some cases, the user may wish to optimize this for different types of organism or for plasma-proteins. For plasma proteins, glycosylation should be considered.

See [Modifications](#) for more details.

- **Spectrum Input Options**

- **Start/End time:** Needs to be checked and adjusted as necessary to ensure full/relevant time range is captured.
- **MS1/MS2 start and end time:** User will need to adjust m/z start and end range if acquiring with an extended mass range (EMR) type of mass spectrometer.
- **Resolution:** Should be based upon the analyzer/instrument used. Note that if the value is set too high, it can adversely impact processing time.
- **S/N:** Avoid S/N ratio lower than 1 for both MS1 and MS2 for MS2. MS1 S/N ratio will affect the determination of boundaries of XIC. If the threshold is set too high, the boundaries will be truncated. If the value is too low, expect degradation and lower performance. Setting S/N to 3 for MS2 will likely result in a loss of visibility of all peaks lower than the three lowest intensity peaks.
- **Smoothing:** Centroid data generates Gaussian blur which renders smoothing redundant. Profile data (e.g., TOF) will benefit from smoothing to alleviate the results of oversampling.

- **Output Options**

- **Minimum Score Threshold:** Setting this value too high could adversely impact results for log probability. Threshold limits how many decoy matches the user will get. Note that lowering the threshold might increase processing time. It is advised that the value not exceed 15 or to go lower than 5.

## Multi-Protein Quantification

- **General**

See [General](#)

- **In-Silico options (CSV import)**

See [In-Silico options \(CSV import\)](#)

- **MS extract options**

See [MS extract options](#)

- **Advanced**

See [Advanced](#)

Note: Autodecoys for the Multiprotein Quantitation workflow is = 100.

- **Report**

See [Report](#)

- **UI Configurations**

See [UI Configurations](#)

- **Feature Finder**

See [Feature Finder](#)

- **XIC**

See [XIC](#)

- **PSM filtering:** Pre-filtering of PSMs allows the user to quantify proteins based upon the peptide features that you designate. These filter settings cannot be adjusted once the project creation process has completed. The Multi-Protein Quantitation workflow is the only workflow that permits this type of filtering.

▼ <b>PSM filtering</b>	
Minimum peptide length	6
Maximum peptide length	45
Maximum missed cleavage count	2
Minimum peptide score	15.00
Minimum peptide replicate count	1
Keep homologous sequences	✓ Yes
Minimum peptide matching a protein count	2
Include modifications	None

Figure 141: PSM filtering

- **Minimum peptide length:** A minimum length of six residues helps ensure minimal likelihood of homologous sequences being selected for quantifying proteins
- **Maximum peptide length:** The maximum length was selected to ensure more confident peptide matches
- **Maximum missed cleavage count:** Ensures that peptides identified as a result of poor enzymatic efficiency are not used for quantifying proteins
- **Minimum peptide score:** This minimum score should be modified depending upon your specific application and mass spectrometer
- **Minimum peptide replicate count:** To ensure a more-likely peptide identification, the peptide must be found in at least two input raw files
- **Keep homologous sequences:** If your sequence FASTA database contains very low sequence homology in the representative proteins, it might be beneficial to exclude homologous sequences when quantifying proteins (in which the parameter should be changed to 'No')
- **Minimum peptide matching a protein count:** To preclude the occurrence of "one hit wonder" protein identifications, it is best to require a minimum of at least two peptides to be found for identifying a protein

- **Include modifications:** The most robust quantitation is generally achieved by excluding modified proteins for the purpose of quantification
- **Reported peptide filtering:** Allows the user to select the reported peptide filter set that best fits project needs for data review. These filters can be modified after opening the project result file.

#### ▼ Reported peptide filtering

Retain only primary MS2 peptide ID	✓ Yes
Use only top charge state for quantification	✓ Yes
Use the same top N peptides for quantification	✓ Yes
Top N count	3
Apply RSD limit for replicates	— No
Maximum acceptable %RSD for replicates	30.00

Figure 142: Reported peptide filtering

- **Retain only primary MS2 peptide ID:** Will 'save' only the highest scoring MS2 spectral match for confirmation purposes
  - **Use only top charge state for quantification:** Ensures that only one charge state per peptide sequence will be used for quantifying the protein
  - **Use the same top N peptides for quantification:** Ensures that the same peptide is used from each input raw file for the purpose of quantifying proteins, ensuring a more robust quantification
  - **Top N count:** This parameter enables the user to specify the number of peptides they wish to use for quantifying proteins in their experiment
  - **Apply RSD limit for replicates:** This parameter is useful if sample replicates produce very reproducible XICs to ensure that random peptide matches are not used for the purpose of quantifying proteins
  - **Maximum acceptable %RSD for replicates:** This parameter is used by the application only if the "Apply RSD limit for replicates" is also used, and the value should be adjusted based upon the reproducibility of your particular HPLC/LC-MS/MS system
- **Match between runs:** Enabling Match between runs causes the software to search for MS1 peaks and create extracted ion chromatograms (XICs) in the sample replicates which did not already have a MS2 verified peptide identification.

▼ Match between runs	
Enable match between runs	✓ Yes
Mode	Between all samples

Figure 143: Match between runs

In each individual sample, some identifications could be missed due to:

- The ion being skipped in the experiment
- Only MS1 data available for some peptides
- Matrix effects swamping the detector

This feature tells the software to fill in gaps using the XIC of the precursor ion that matches the same identifications in the other samples and/or replicates of the same project.

The **Mode** option determines whether Match Between Runs is run between all samples (default) or within each condition.

## Multi-Protein Preview

The **Multi-Protein Preview** workflow provides a rapid overview of the data to indicate analysis parameters that are likely to be optimal.

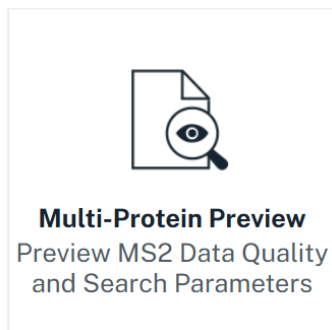


Figure 144: Multi-Protein Preview icon

This workflow provides the user with a way to preview complex protein data to:

- Preview the 'top hits' (proteins) in their data
- Measure the accuracy and precision of their mass spec calibration
- Assess the efficiency of their protein digestion
- Understand the modifications and artifacts in their data

This workflow includes suggested Byonic workflow parameters (\*.byparms file), Spectrum.identifications.csv (target list), and Recalibration details (precursor and fragment coefficients) to assist the user in their analysis.

## Processing nodes

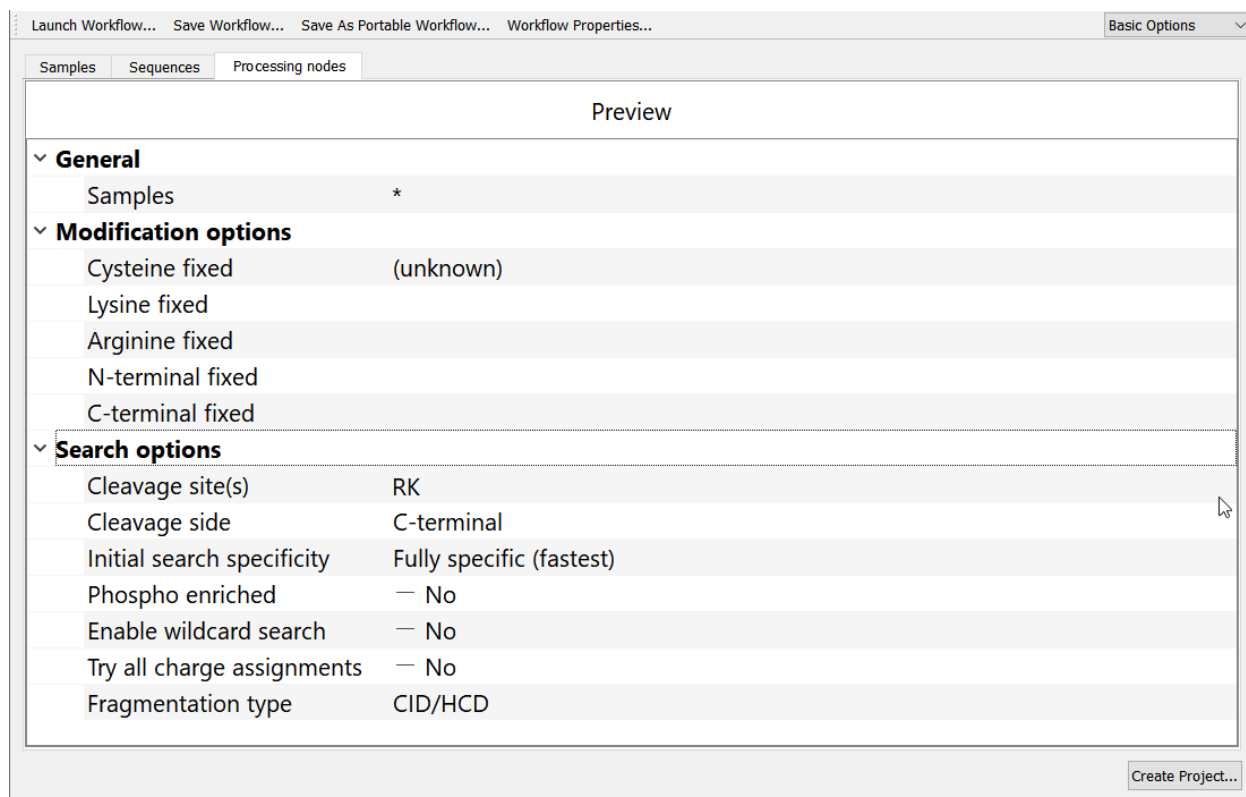


Figure 145: Processing nodes

### Preview

- **General**

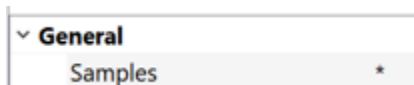


Figure 146: Samples to include in the project

- **Samples** - The “\*” character applies all parameters to all samples dragged and dropped into the Samples tab.

- **Modification options**

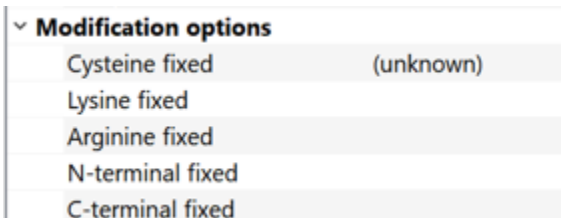


Figure 147: Modification options

Under Modification options, the user can specify which modifications to include in their search. For standard cysteine treatments (+0, +46, +57, +58, and +71), input is not always necessary. If the user selects “(unknown)”, Preview usually determines the correct cysteine treatment from the data.



- **Search options**

▼ **Search options**

Cleavage site(s)	RK
Cleavage side	C-terminal
Initial search specificity	Fully specific (fastest)
Phospho enriched	— No
Enable wildcard search	— No
Try all charge assignments	— No
Fragmentation type	CID/HCD

Figure 148: Search options

- **Cleavage site(s) and side** set the point of cleavage and whether to cleave on the **C-terminal** or **N-terminal** side. Enter one-letter abbreviations for residues on either side of the cleavage point. (in this example, trypsin, on the C-terminal side of arginine or lysine). For a broader search, use nonspecific cleavage at one or both termini. Nonspecific digestion can vary from negligible to ubiquitous depending upon endogenous peptidases, and missed cleavages vary widely depending upon sample processing conditions.
- **Initial search specificity** sets the level of specificity, and thus the search speed. The options are:

Initial search specificity: Fully specific (fastest) ▼

Phospho enriched: Non specific (slowest)

Enable wildcard search: Semi specific (slow)

Try all charge assignments : Fully specific (fastest)

Semi specific N-ragged (slow)

Semi specific C-ragged (slow)

○ Figure 149: Search specificity options

- Fully specific searches are recommended for all digested samples. Nonspecific initial searches may perform better for undigested (peptidomic) samples. In Fully specific searches (the default), both peptide termini must agree with the input digestion cleavages. In N-ragged searches, only the C-terminus must agree, and in C-ragged searches, only the N-terminus must agree. In Semi specific searches, one the two termini can disagree. In Non specific searches, both termini can disagree.
- Check **Phospho enriched** to optimize Preview when the sample is composed predominantly of phosphopeptides.
- Check **Enable wildcard search** to enable searches of the spectra with a wild-card modification. When it is checked, the program performs a blind modification search that tries each integer mass shift from -50 to +150 on any one residue. The mass of the modification will be reported to the accuracy of the precursor mass.
- When **Try all charge assignments** is checked, the charge assignments in the spectrum file are ignored. The program runs every spectrum using  $z = +1, +2, +3$  for each CID spectrum, and  $z = +2, +3, +4$  for each ETD spectrum. You can run Preview twice, once with this box checked and once without, to test the reliability of the charge assignments.
- The **Fragmentation type** options are:

Fragmentation type: CID / HCD

CID / HCD  
ETD / ECD

Figure 150: Fragmentation type options

- CID / HCD represent b and y ions, and ETD / ECD represent c and z ions.

## Preview Results

After a successful run, Preview displays the results as an html page in the default web browser. The initial page shows the summary results, also displayed when **Summary** is clicked in the title bar:

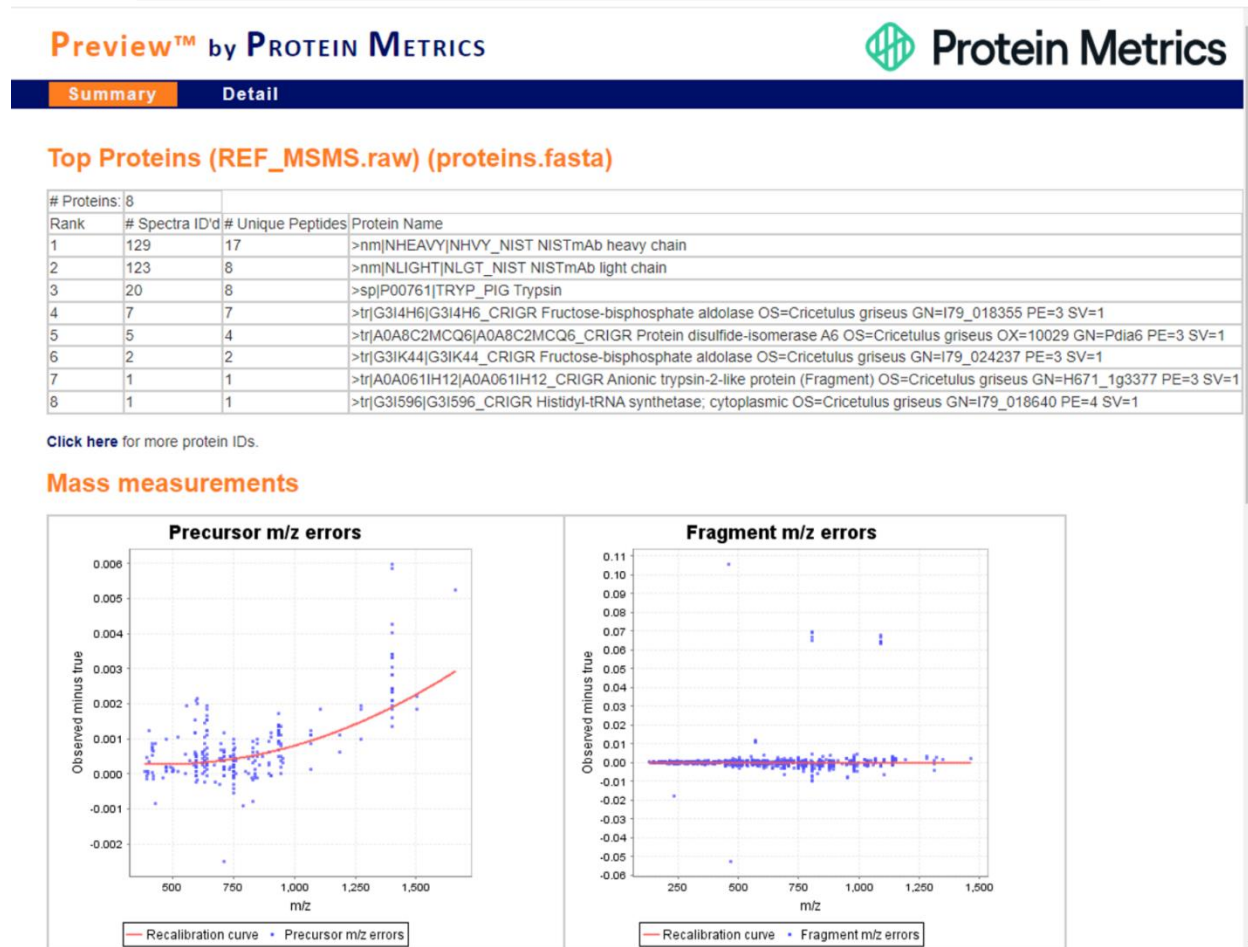



Figure 151: Example Preview report

Separate results will open in the default browser for each input data file previewed. The HTML report contains two pages, the **Summary** page and the **Detail** page.

## Index of Z:\Multi-Protein Preview\UNK\_MSMS.raw\_Preview

 [parent directory]








	Name	Size	Date Modified
	objs/		8/8/24, 2:00:59 PM
	index.html	7.5 kB	8/8/24, 2:00:59 PM
	result_detail.html	12.2 kB	8/8/24, 2:00:59 PM
	result_summary.html	7.5 kB	8/8/24, 2:00:59 PM
	Spectrum.identifications.csv	16.3 kB	8/8/24, 2:00:59 PM
	UNK_MSMS.raw_Preview_240808_140043.prv	504 B	8/8/24, 2:00:59 PM
	UNK_MSMS.raw_Preview_240808_140043.prv.byparms	3.2 kB	8/8/24, 2:00:59 PM

Figure 152: HTML index of Preview results

There is also a link to the result folder, permitting access to the **Byparms** file, which can be imported into protein identification nodes in Byos workflows

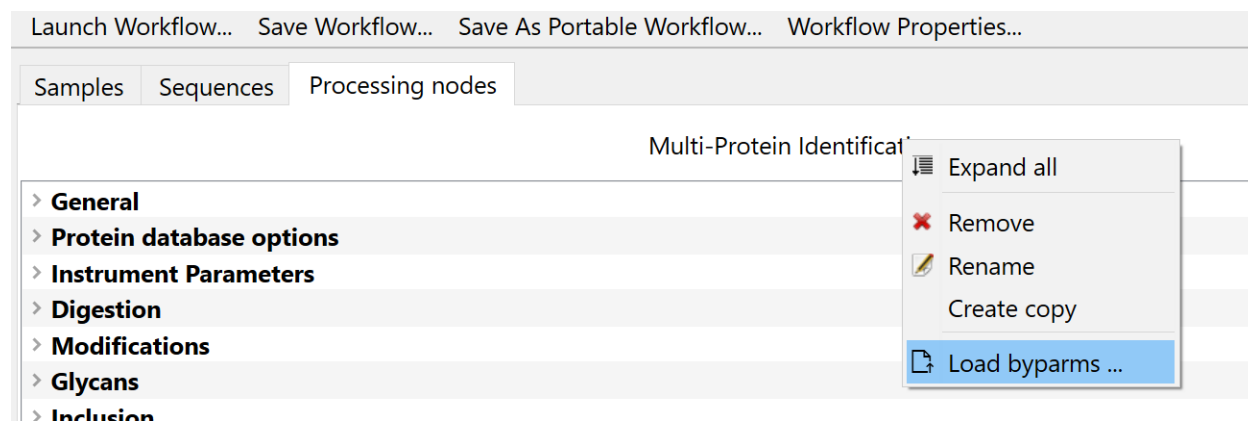


Figure 153: Load Byparms

## Licensing Requirements

Multi-Protein Quantitation requires a dedicated license. The licensing is in addition to the Peptide, Chromatogram, and Intact licenses available for Byos or Byosphere and can be added as an 'add-on' for users who already have other licenses.

Contact Protein Metrics at [sales@proteinmetrics.com](mailto:sales@proteinmetrics.com) for additional information.

## Supported Data

- Mass spec vendor formats are supported from the leading providers of mass spectrometry instruments: .d; .lcd; .raw; .wiff; .wiff2
- Data that contains peptide-level species from enzymatic digests (principally single-enzyme digests are expected, but multi-enzyme digests are supported)
- Studies of relevance include Host Cell Protein (HCP) assays, Disease comparisons with a few thousand proteins, searching for and identifying impurities in protein mixtures, and unknown proteins in the presence of known Biopharmaceutical products
- Data import from other search results (ByrsIt files)

## Released Glycans (N-Linked, O-Linked, and IgG) Workflows

The following section details parameters in the Glycan table, Glycan options, and Processing nodes tabs for these three workflows.

## Glycan Table

Each default workflow is populated with a table of Glycans specific to the workflow – N-linked, O-linked, or IgG. The “Glycan alias name” and “Glycan composition” are detailed for each entry. Please note the start and end times are left blank by default.

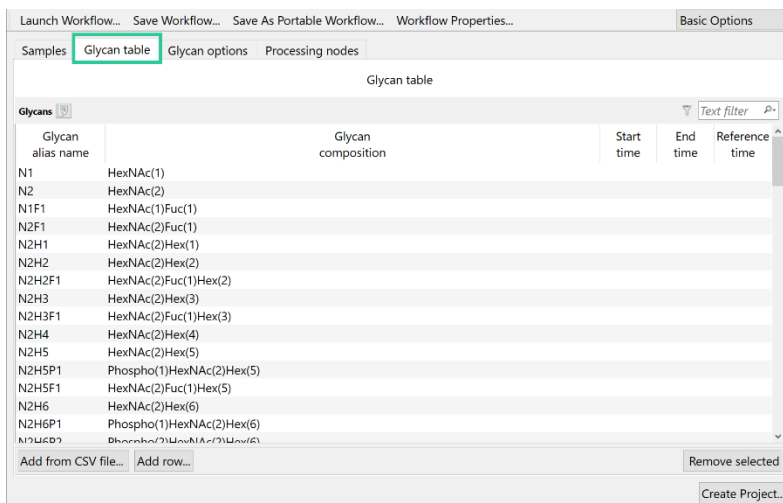


Figure 154: Glycan table tab in the Released Glycan workflows

The user has many options while working with these databases – the current content is editable. The user may also opt to add to the current content or even delete the current content and replace it with a customized database. This is done by highlighting the content and clicking on Remove all selection, as shown below:

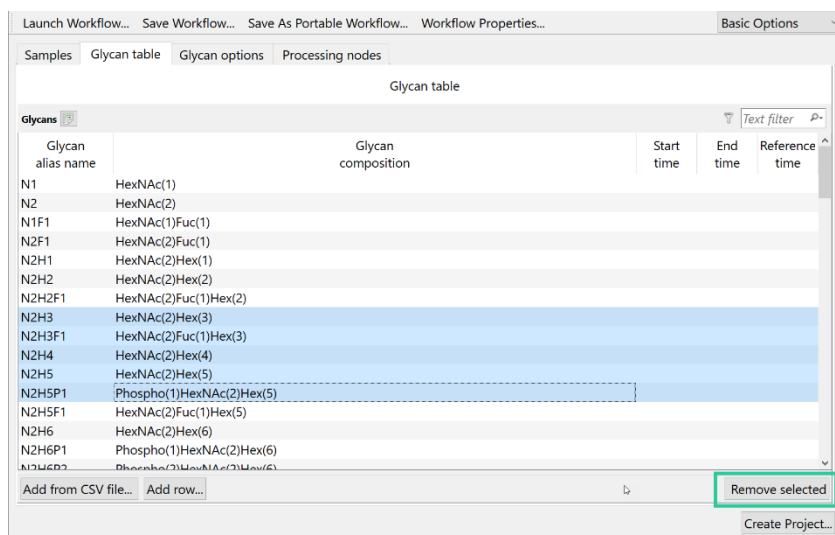


Figure 155: Remove all selection

The user would then click Add from CSV file and direct Byos to the desired database, as shown below:

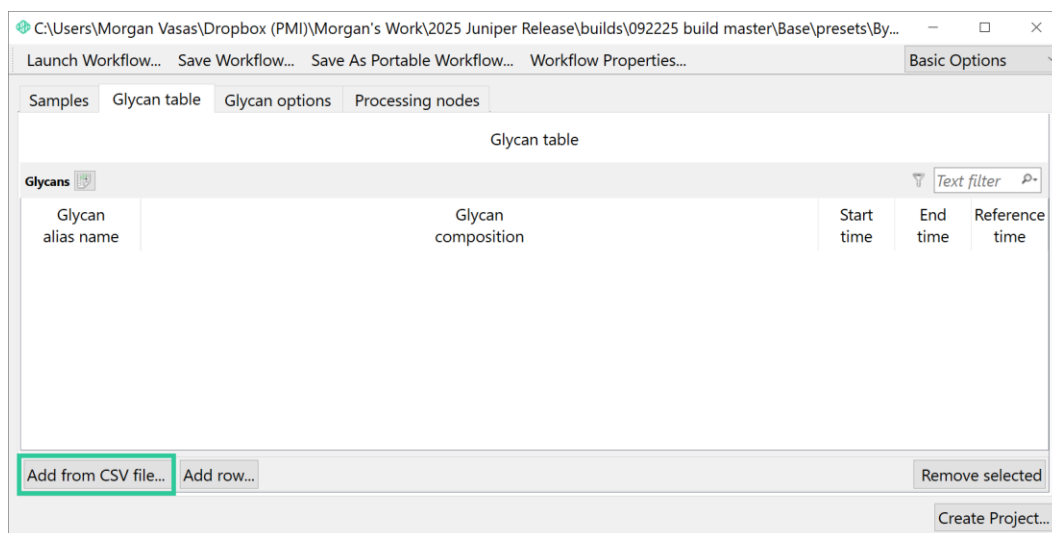


Figure 156: Glycan table, Add from CSV file to import a custom Glycan database

## Glycan Options

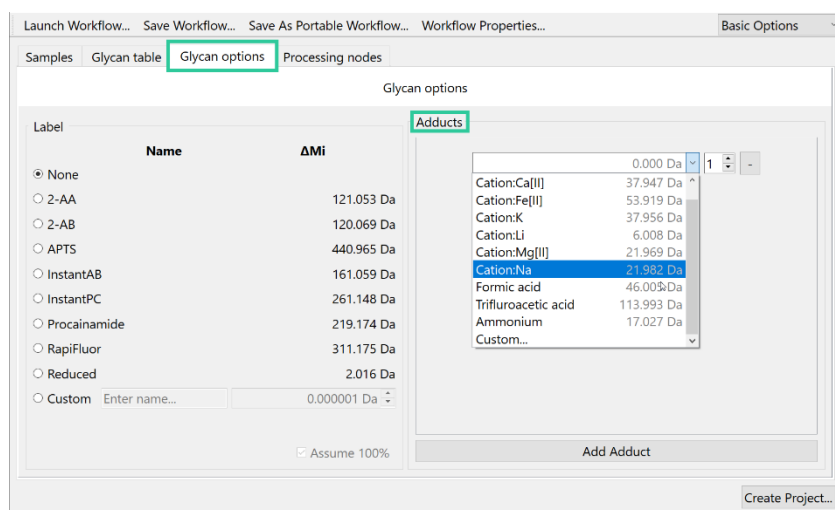


Figure 157: Glycan options tab – identify the label used and set Adducts

The user has the option to define the label used during the experiment. The delta mass of each tag is defined. The user can also define 100% labeling.

The user can also include adducts to consider in the mass calculation in this same tab. Click **Add Adduct** will make the displayed drop-down menu available. The user can click the button again to add another adduct.

## Processing nodes

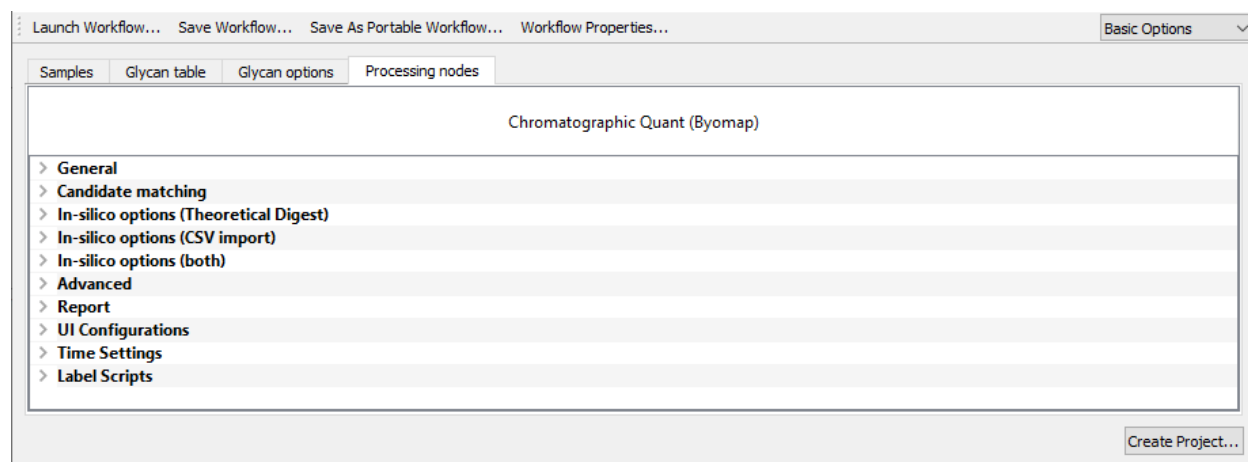


Figure 158: Processing nodes

## Chromatographic Quant

- **General**

See [General](#)

- **Candidate matching**

See [Candidate matching](#)

- **In-silico options (Theoretical Digest)**

See [In-Silico options \(Theoretical Digest\)](#)

- **In-silico options (CSV Import)**

**In-Silico Peptides CSV** - This option allows the user to add a list of in-silico peptides from an imported CSV file, thus there is no need to run an MS2 search if there are known modifications with known masses and retention times. The format for the CSV is shown below.

	A	B	C	D	E	F	G
1	Sequence	UnchargedMass	ModificationsPositionList	ModificationsNameList	StartTime	EndTime	ChargeList
2	KSFNRGEC	996.444699	4,7	Deamidated/0.9840,Amid	3.257433832	7.257433832	2,3
3	KVDNALQSGNSQESVTEQDS	2135.945451	10	Deamidated/0.9840	41.86437026	45.86437026	2,3
4	KSRWQQGNVFCSSVMHEALHN	3044.376965	21	Deamidated/0.9840	52.18317542	56.18317542	6
5	KPREEQYNSTYRVVSVLTVHLQ	4837.233918	4,6,8	Amidated/-0.9840,Deamid	81.91415187	85.91415187	5,6
6	KPREEQYNSTYRVVSVLTVHLQ	4675.181095	8,22,23	NGlycan/1444.5339,Deamid	82.17404358	86.17404358	5
7	KSGTASVVCLLNNFYPREA	2126.046605	12	Deamidated/0.9840	83.72993315	87.72993315	4
8	RVVSVLTVHLQDNLNGKEY	2256.190231	15	Deamidated/0.9840	83.84885265	88.42171697	4
9	KGPYPSPDIAVEWESNGQPENNY	2544.108081	17	Deamidated/0.9840	94.13709563	97.46649733	4
10	RVVSVLTVHLQDNLNG	1835.005335	11,12	Deamidated/0.9840,Amid	91.76761459	95.76761459	3
11	KGPYPSPDIAVEWESNGQPENNY	2544.108081	15	Deamidated/0.9840	92.70770462	97.71021981	3
12	KDYFPEPVTWNSGALTSQGVH	7055.481425	57	Deamidated/0.9840	108.707735	112.707735	6
13	KDYFPEPVTWNSGALTSQGVH	6186.020755	13	Deamidated/0.9840	127.0172827	131.0172827	9
14	DIQMTQSPSSLSASVGDRVTTC	4250.937205	27	Deamidated/0.9840	145.3157387	149.3157387	4
15	DIQMTQSPSSLSASVGDRVTTC	4250.937205	6	Deamidated/0.9840	137.2344279	141.2344279	4
16	SGTASVVCLLNNFYP	1797.871945	11	Deamidated/0.9840	62.79148405	66.79148405	3
17	SGTASVVCLLNNFYP	1780.845396	11,12	Deamidated/0.9840,Amid	66.86234322	70.86234322	3
18	SGTASVVCLLNNFYP	1797.871945	12	Deamidated/0.9840	58.54196361	62.54196361	3
19	KSTGGTAALGCLVKDYFPEPVT	7488.680909	27	Deamidated/0.9840	119.6788643	123.6788643	7
20	STSGGTAAALGCLVKDYFPEPVT	8016.935285	26,42	Deamidated/0.9840,Deamid	79.88044813	83.88044813	10
21	VDNALQSGNSQESVTEQDSK	2134.961435	2,6	Amidated/-0.9840,Deamid	36.80518941	40.80518941	3
22	VDNALQSGNSQESVTEQDSK	2135.945451	3	Deamidated/0.9840	39.95379704	43.95379704	3
23	VDNALQSGNSQESVTEQDSK	2135.945451	6	Deamidated/0.9840	39.95379704	43.95379704	3
24	VDNALQSGNSQESVTEQDSK	2135.945451	9	Deamidated/0.9840	36.83448115	40.83448115	3

Figure 159: Format for an In-Silico Peptides CSV

- **In-Silico Options (Both)**

See [In-Silico Options \(Both\)](#)

- **Advanced**

See [Advanced](#) except for the following option:

- **Ionization mode**

- **Report**

See [Report](#)

- **UI Configurations**

See [UI Configurations](#)

- **Time Settings**

See [Time Settings](#)

- **Label Scripts**

See [Label Scripts](#)

- **Peak Construction options**

See [Peak Construction options](#)

(Pertains to older versions of Byos)

## De novo Sequencing Workflow

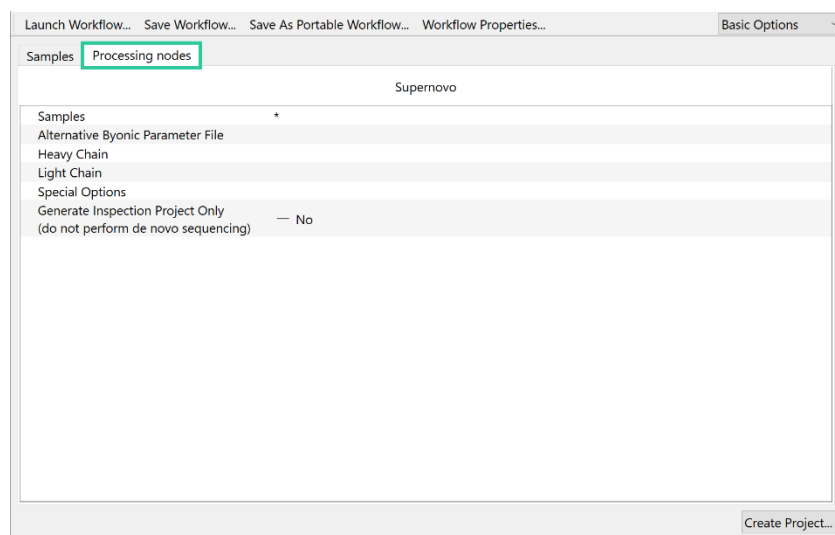


Figure 177: Processing nodes

## Processing nodes

### Supernovo

- **Samples** - The “\*” will apply all parameters to all samples dragged and dropped into the Samples tab.
- **Alternative Byonic Parameter File** [OPTIONAL]

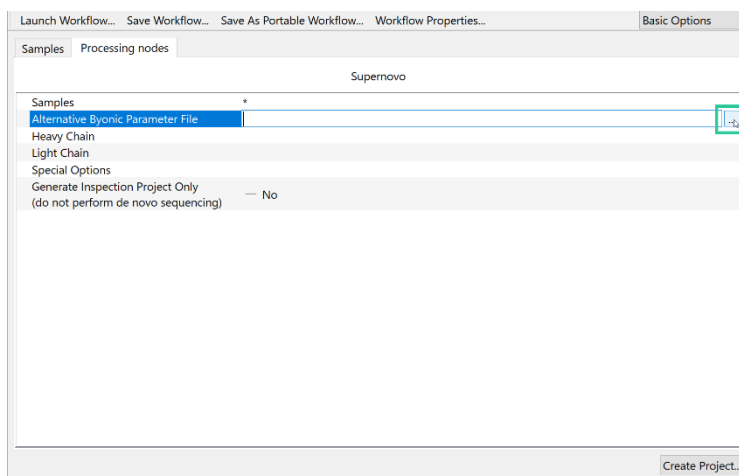


Figure 160: Alternative Byonic Parameter File parameter

Specify an alternative Byonic parameter file (\* .byparms) by clicking within the box to activate the light blue “...”button, as detailed in the Figure above. If left blank, Byos uses a default Byonic parameter file, which specifies 10 ppm precursor mass tolerance, 20 ppm fragment mass tolerance, and carbamidomethyl (delta mass +57) for the cysteine alkylation.

- **Heavy Chain and Light Chain [OPTIONAL]**

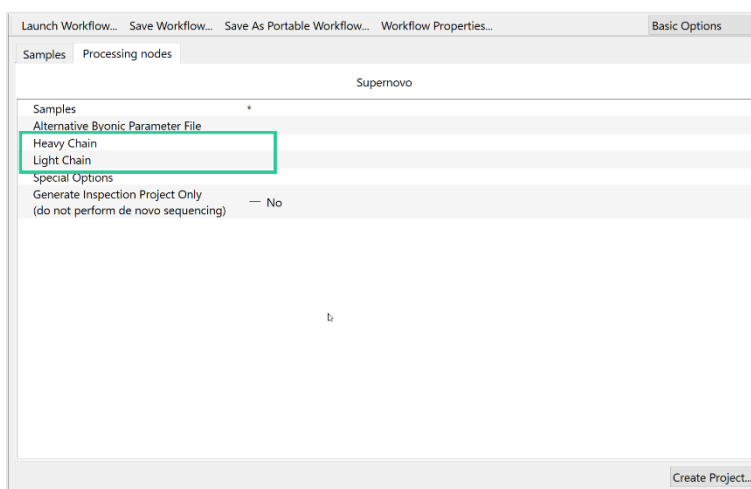


Figure 161: The sequences for the Heavy Chain and Light Chain can be pasted in as text

If a starting sequence template is to be used, paste the Heavy Chain and Light Chain sequences. If left blank, Byos will automatically determine an antibody scaffold sequence as the starting point for de novo analysis. This works best for antibodies from human, mouse, rat, or related species.

- o **Special Options [OPTIONAL]** - Leave this blank generally. Please reach out to [support@proteinmetrics.com](mailto:support@proteinmetrics.com) for additional details.
- o **General Inspection Project Only** - the user can toggle between Yes/No. Check this box to only generate metrics / visualizations for a particular specified sequence. Byos will not perform de novo sequencing for this inspection project. Please note this type of project runs faster.

Click **Create Project** to start analysis and report generation. Please note Byos will save projects based upon this type of processing as a .blgc file.



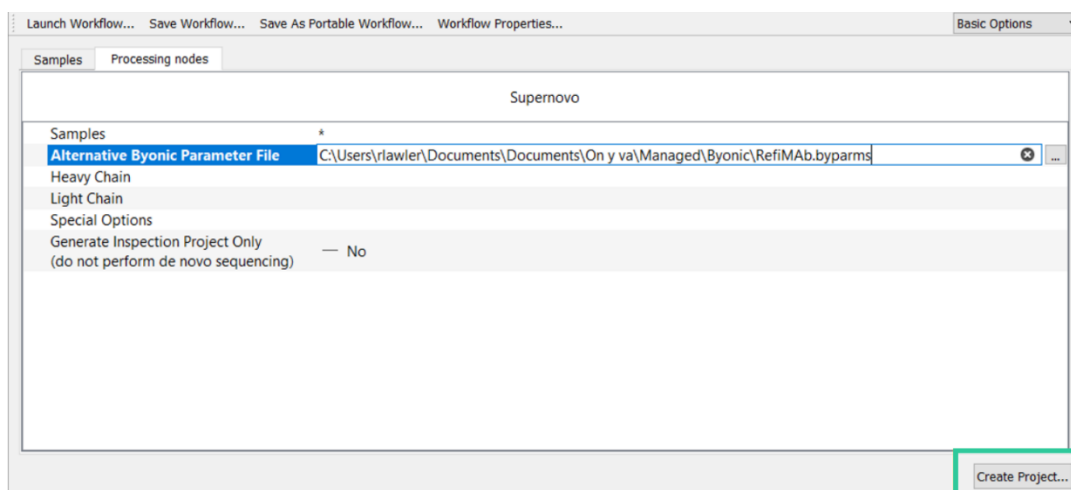


Figure 180: Click Create Project

Processing will commence. Once complete, the project and associated report will appear as additional tabs after the “Byos Workflows” tab, as shown in figure below.

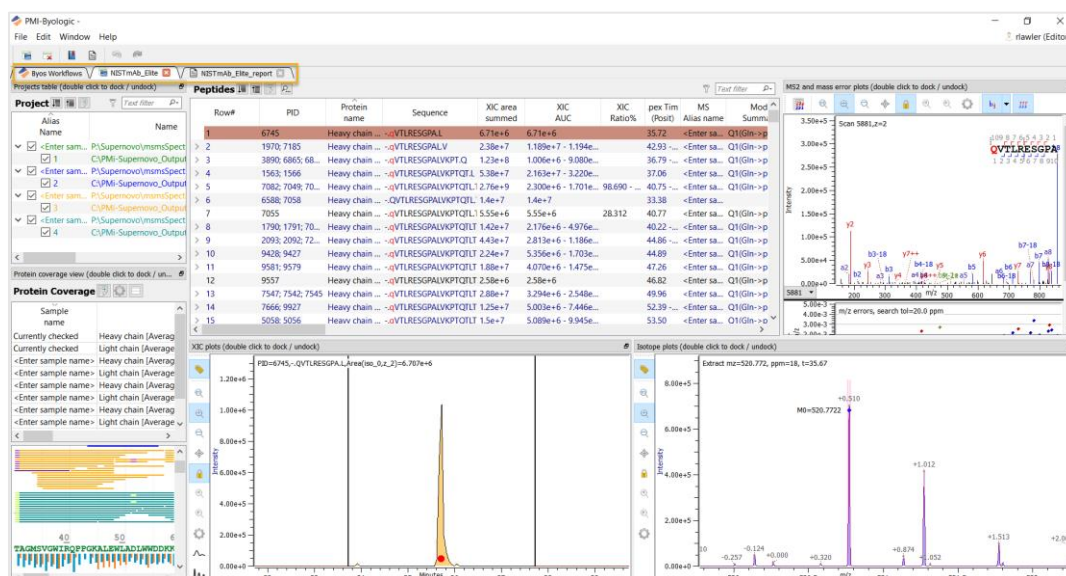


Figure 182: Byos opening tab, completed project, and associated report in separate tabs

Please note the analysis can take hours. Early in the analysis, a constant region depth of coverage is calculated. The number of distinct peptide sequences covering each amino acid residue in the constant regions of the heavy chain and the light chain is counted, and the average is calculated to give the average depth of coverage. The depth of coverage provides an estimate of how likely the full analysis will find the complete sequence accurately and provides feedback to the user on the sequence depth of their data set.

- o Greater than 20 – good chance of success.
- o Between 10 and 20 – moderate chance of success.
- o Less than 10 – poor chance of success; recommend collecting more data and/or revising data collection procedure.

## Oligonucleotide Analysis

A workflow to launch Oligonucleotide analysis is available. Users can input MS1 data of Oligonucleotides to identify different molecules within the sample. Byos software has a novel algorithm that deconvolves the Oligonucleotide data to generate the mass spectrum. The Oligonucleotide analysis project files are stored as \*.olms files. An additional workflow has been added in Byos 5.4 that supports Digested Oligonucleotide analysis. Please see the **Oligo User Guide** for more information.

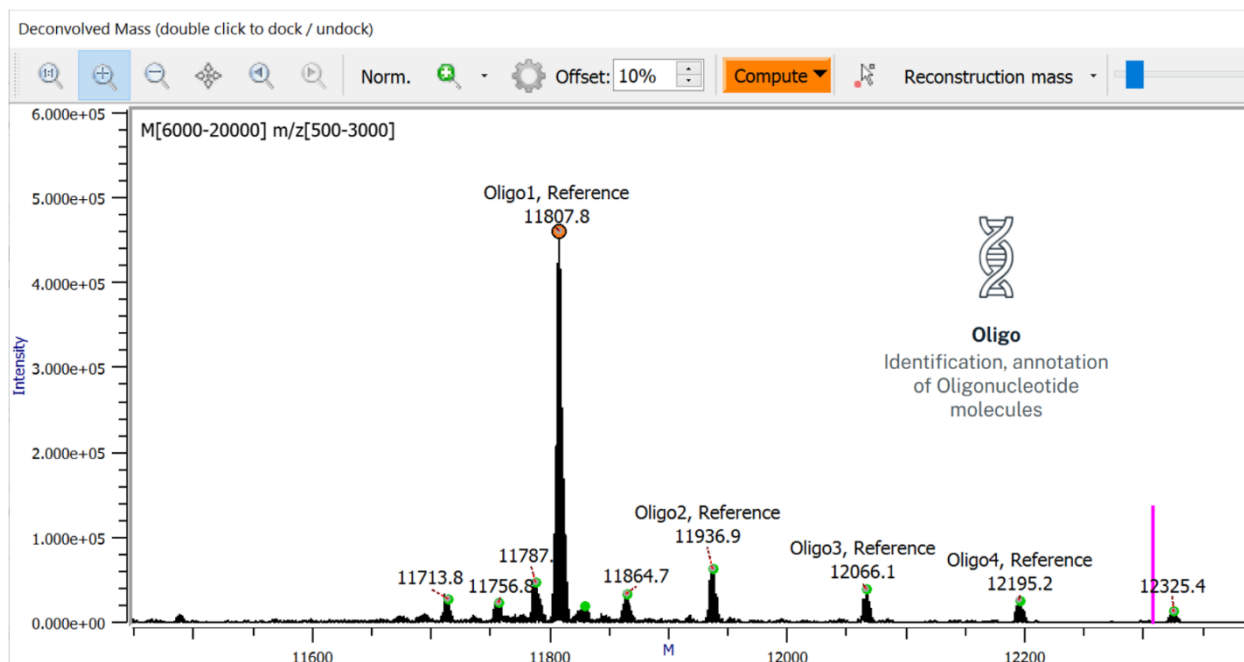


Figure 183: Oligonucleotide Deconvolved Mass data analysis

## MOBILion High Resolution Ion-Mobility (HRIM) Analysis

Three new workflows are now available to support the MOBILion high-resolution ion mobility (HRIM) data. The MOBILion raw data has \*.mbi file extension, while the project files generated are generic \*.ntms (for HRIM-Intact), and \*.bmap (for HRIM-Peptide and HRIM-Glycan workflows):

- o HRIM Intact workflow to support intact analysis of MOBILion HRIM data
- o HRIM Peptide workflow to support peptide analysis of MOBILion HRIM data
- o HRIM Glycan workflow to characterize released glycan analysis of MOBILion HRIM data

Please see the MOBILion HRIM Analysis User Guide for more information, available through [support@proteinmetrics.com](mailto:support@proteinmetrics.com).

## Charge Variant Analysis with Reconstruction



Figure 162: Charge Variant Analysis with Reconstruction

The Charge Variant Analysis workflow allows the user to correlate cIEF traces with peptide mapping data. The workflow searches for PTMs that commonly modify the pI of proteins and then constructs a theoretical cIEF trace and overlays it onto the experimentally collected trace. This allows the user to visually compare the traces and identify species that are not currently accounted for in the peptide mapping search space.

**The** modifications taken into account include: succinamide, pyroGlu, pyroGln, deamidation, proline amidation, C-term Lys loss, C-term Lys, glycation, and salic acid incorporated glycans. All these modifications are already preset within the Charge Variant Analysis with Reconstruction workflow.

The workflow supports any cIEF file in X,Y format such as .arw files, along with mass spectrometry raw files. For the first step in the analysis both file names need to be renamed. Add peptide to the end of the name for the MS raw file and change the file extension to .pl for the cIEF file as shown below:



 iCIEF_pl_trace.pl	2/6/2024 11:41 AM	PI File	35 KB
 NIST_A1_peptide.raw	1/31/2024 6:47 AM	RAW File	231,555 KB

Figure 163: Renamed files

Both files can then be dragged and dropped into the Samples table:

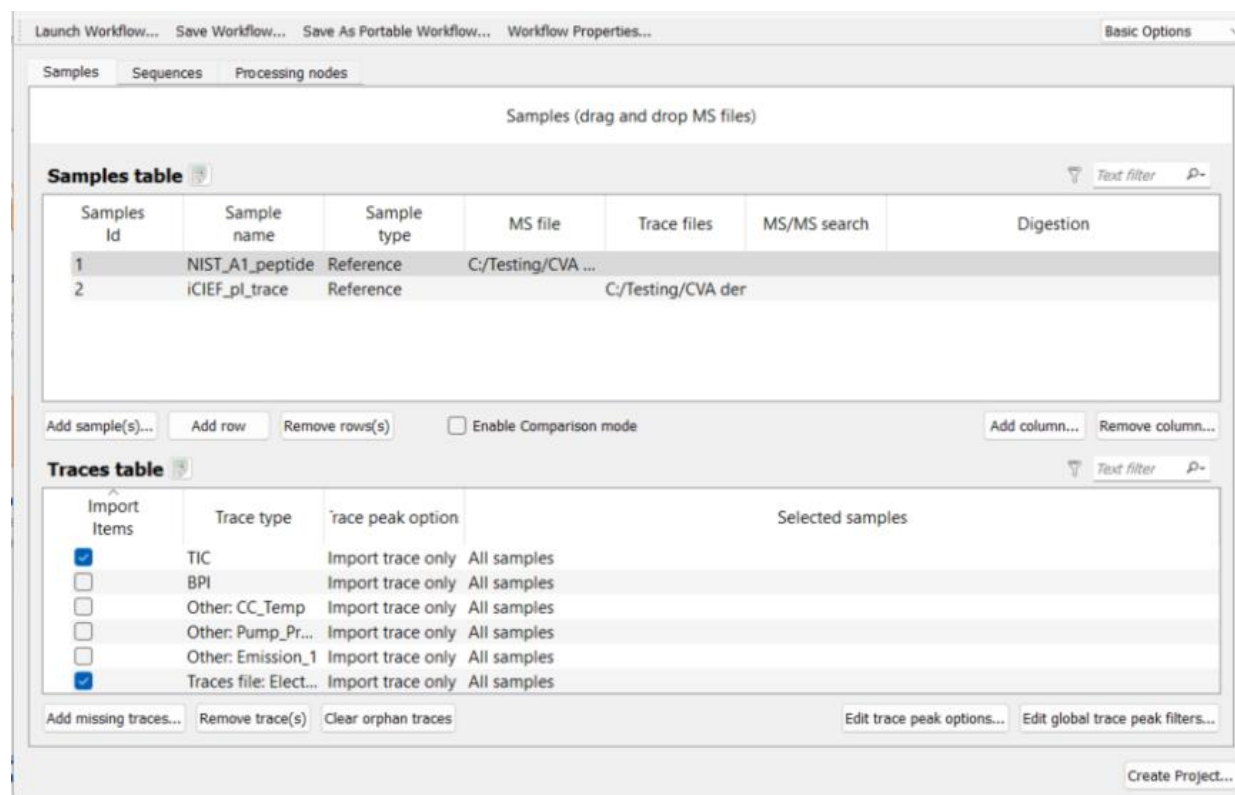
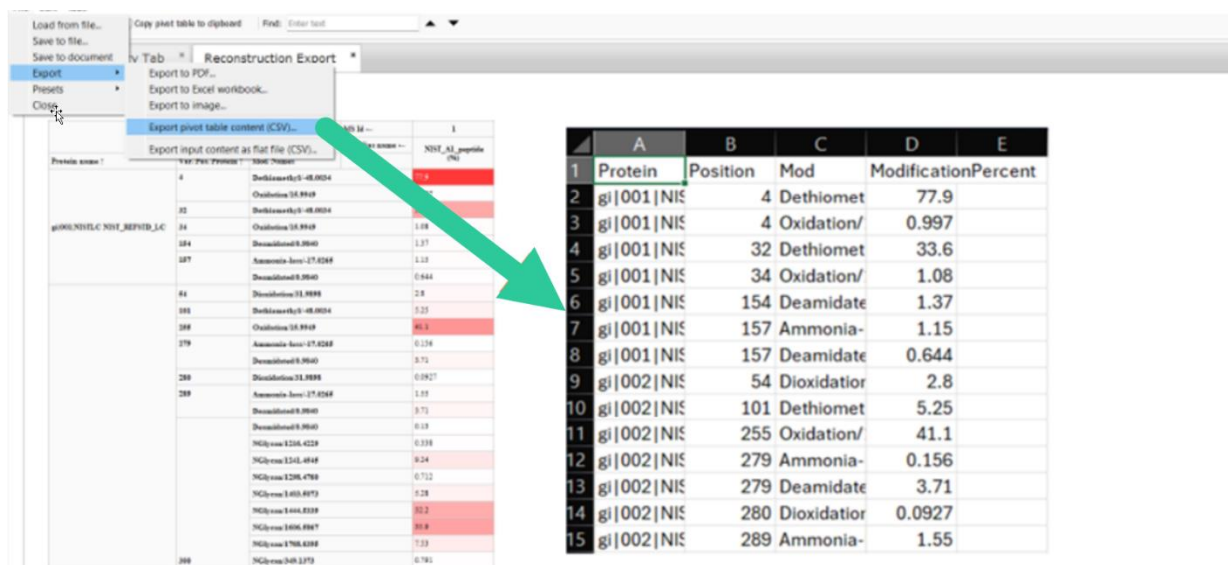


Figure 164: Samples added

Next, add protein sequences and configure the Byonic node with appropriate settings including instrument parameters, digestion, modifications including alkylating agent, etc. and create the project.

The workflow will output both a peptide (.blgc) and chromatogram (.bmap) project. The peptide mapping project can be manually validated. Once completed, open the report and view the "reconstruction export" tab. This tab is formatted for importing into the reconstruction feature. To export, select File->Export->Export pivot table content CSV.

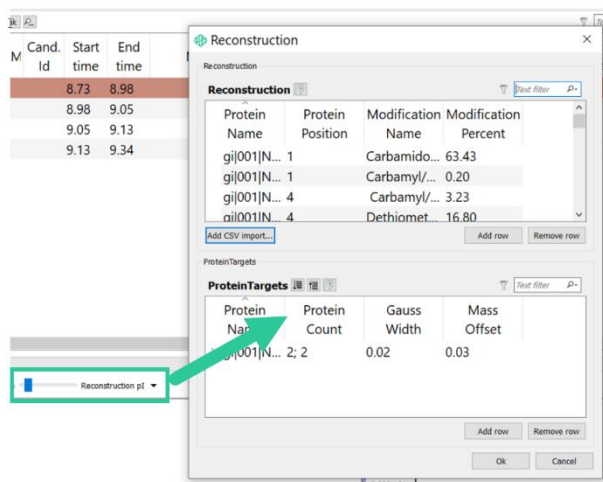
The resulting CSV will have the following output format:



Protein	Position	Mod	ModificationPercent
gi 001 NIS	4	Dethiomet	77.9
gi 001 NIS	4	Oxidation/	0.997
gi 001 NIS	32	Dethiomet	33.6
gi 001 NIS	34	Oxidation/	1.08
gi 001 NIS	154	Deamidate	1.37
gi 001 NIS	157	Ammonia-	1.15
gi 001 NIS	157	Deamidate	0.644
gi 002 NIS	54	Dioxidator	2.8
gi 002 NIS	101	Dethiomet	5.25
gi 002 NIS	255	Oxidation/	41.1
gi 002 NIS	279	Ammonia-	0.156
gi 002 NIS	279	Deamidate	3.71
gi 002 NIS	280	Dioxidator	0.0927
gi 002 NIS	289	Ammonia-	1.55

Figure 165: Reconstruction CSV contents

It is important to make sure the modification percent at any given position does not add up to greater than 100 percent or an error will occur when creating the reconstruction. Otherwise, the output format can be taken directly into the reconstruction feature:



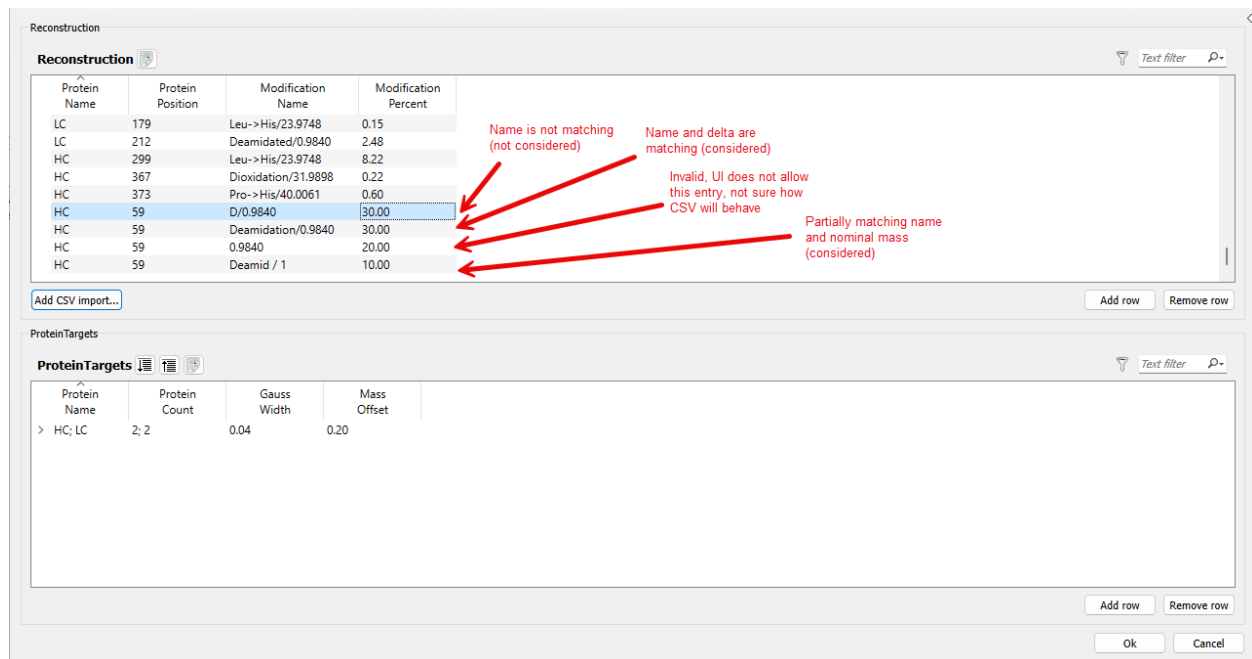





Figure 166: Reconstruction dialog

Once the CSV has been added the protein targets will need to be configured. The protein target will need to match the construct used in the cIEF trace, in this case a native antibody was used, so 2 counts of the heavy and light chains are added in the Protein Targets table as shown below:






Reconstruction

**Reconstruction**   Text filter 

Protein Name	Protein Position	Modification Name	Modification Percent
gi 001 NISTL...	4	Dethiomethyl/-...	77.90
gi 001 NISTL...	4	Oxidation/15.99...	1.00
gi 001 NISTL...	32	Dethiomethyl/-...	33.60
gi 001 NISTL...	34	Oxidation/15.99...	1.08
gi 001 NISTL...	154	Deamidated/0....	1.37
gi 001 NISTL...	157	Ammonia-loss/...	1.15

[Add CSV import...](#) [Add row](#) [Remove row](#)

ProteinTargets

**ProteinTargets**     Text filter 



Protein Name	Protein Count	Gauss Width	Mass Offset
gi 002 NIST...	2; 2	5.00	
gi 002 N...	2	5.00	
gi 001 N...	2	5.00	

[Add row](#) [Remove row](#)

[Ok](#) [Cancel](#)



Next, the Gauss Width must be adjusted to better fit cIEF data. 0.01 is often a good value to use:

Reconstruction

**Reconstruction**  Text filter 

Protein Name	Protein Position	Modification Name	Modification Percent
gi 001 NISTL...	4	Dethiomethyl/-...	77.90
gi 001 NISTL...	4	Oxidation/15.99...	1.00
gi 001 NISTL...	32	Dethiomethyl/-...	33.60
gi 001 NISTL...	34	Oxidation/15.99...	1.08
gi 001 NISTL...	154	Deamidated/0....	1.37
gi 001 NISTL...	157	Ammonia-loss/...	1.15

ProteinTargets




**ProteinTargets**  Text filter 

Protein Name	Protein Count	Gauss Width	Mass Offset
gi 002 NIST...	2; 2	0.01	
gi 002 N...	2	0.01	
gi 001 N...	2	0.01	

Lastly, Mass Offset can be used to align the theoretical cIEF trace over experimentally collected trace. In this case 0.03 was used:








Reconstruction

**Reconstruction**   Text filter 

Protein Name	Protein Position	Modification Name	Modification Percent
gi 001 NISTL...	4	Dethiomethyl/-...	77.90
gi 001 NISTL...	4	Oxidation/15.99...	1.00
gi 001 NISTL...	32	Dethiomethyl/-...	33.60
gi 001 NISTL...	34	Oxidation/15.99...	1.08
gi 001 NISTL...	154	Deamidated/0....	1.37
gi 001 NISTL...	157	Ammonia-loss/...	1.15


Add CSV import... Add row Remove row

ProteinTargets

**ProteinTargets**     Text filter 

Protein Name	Protein Count	Gauss Width	Mass Offset
> gi 002 NISTL...	2; 2	0.01	0.03

Add row Remove row Ok Cancel



Once applied the reconstruction will be complete:

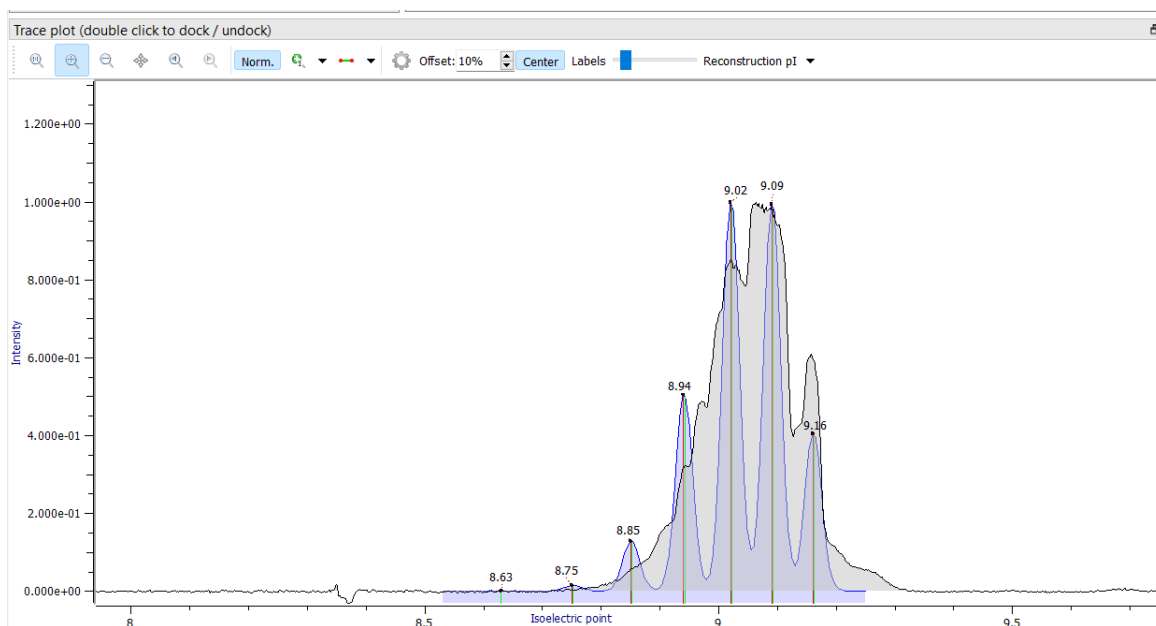


Figure 167: Result of Charge Reconstruction

## Byos Preview Workflow

As of Byos v5.9, a workflow is now available within Byos encompassing the tools provided by the Preview standalone application.

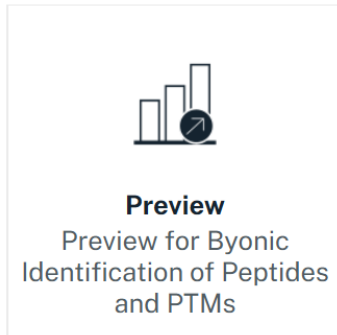


Figure 168: Byos Preview workflow icon

Preview is a standalone desktop application that performs a first-pass, prospecting style search to identify key parameters that can subsequently be used for Byonic peptide search. Mass errors, digestion specificity, and common modifications are among the parameters that Preview will suggest. More information about the standalone desktop application can be found in **General-08-Preview-Manual**.

## My Workflows

### Saving a Customized Workflow or Portable Workflow

Once the user has reviewed the default parameters and made all desired changes as detailed in the “Customizing System Workflows to Create My Workflows” section, he or she is ready to save this workflow as one of the “My Workflows”.

## Save as Portable Workflow

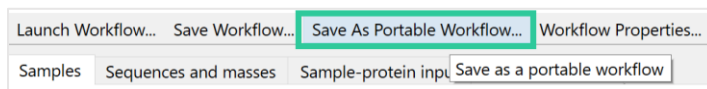


Figure 184: Save as Portable Workflow

The user can click **Save As Portable Workflow** to save the customized workflow file.

The user will be prompted to define the folder name. The user must save the portable workflow into a preferred folder location.

NOTE: The portable workflow folder contains the layout, report template, and workflow comprising the Portable Workflow, as shown below. The customized content is contained within a folder.

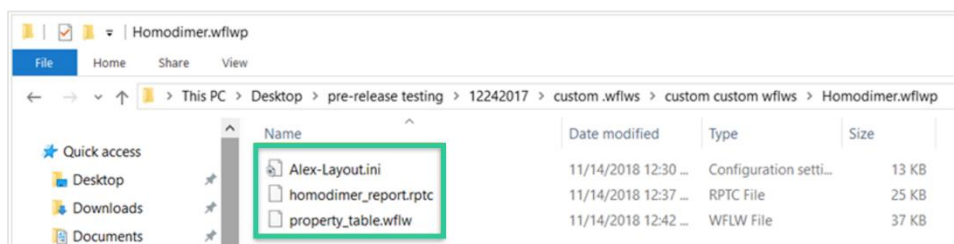


Figure 186: Content of portable workflow folder

## Populate “My Workflows”

The user is now ready to populate the “My Workflows” section. The user will do this by directing Byos to the folder where their customized portable workflow file(s) are saved. This is done by clicking on the “...” button and selecting workflow folder to define the folder directory.

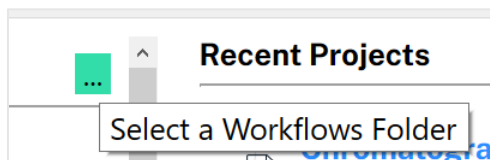


Figure 187: Select the My Workflows folder location

## Create Project to Generate an Analysis and Report

Once the user has saved customized portable workflows into their My Workflows section, a Byos workflow can be launched to generate customized projects and reports. Please note these same steps could be followed using one of the default System Workflows as well. My Workflows simply provides the user with the option to customize.

The below example is for an Intact project; please note the steps to follow are exactly the same for every workflow.

1. Click **Create Project** to start processing and to generate the report.

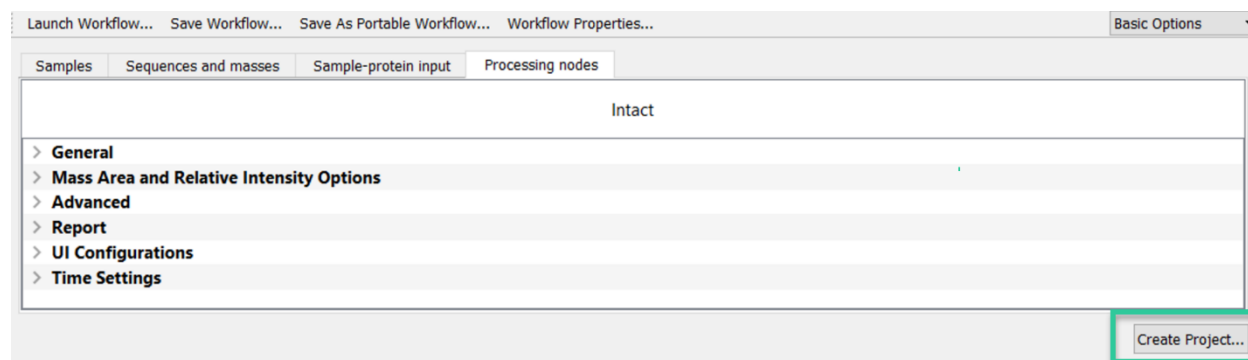


Figure 188: Click Create Project

2. The user will be prompted to name and save the project file. Please note Byos will save projects based upon the type of processing as a .ntms file, .blgc file, or .bmap file.

Processing will commence. Once complete, the project and associated report will appear as additional tabs after the “Byos Workflows” tab.

## Byos UNIFI API Integration

Sample results stored in the Waters UNIFI Application Programming Interface (API) can be loaded directly into Byos. The integration requires installation of the Waters UNIFI API and configuration of both UNIFI and Byos. The **Add UNIFI sample(s)** button will then appear in the **Samples** tab in Byos:

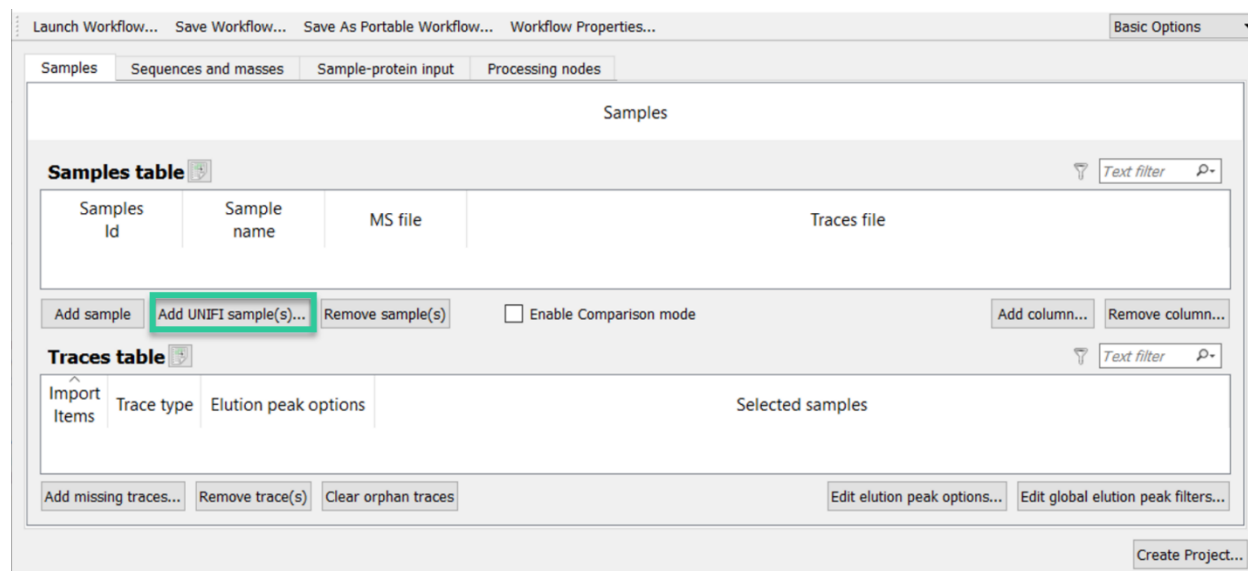


Figure 191: Byos Samples tab configured for UNIFI sample addition

This opens the **Select UNIFI sample** dialog where MS data stored in UNIFI can be browsed and selected for download to Byos:

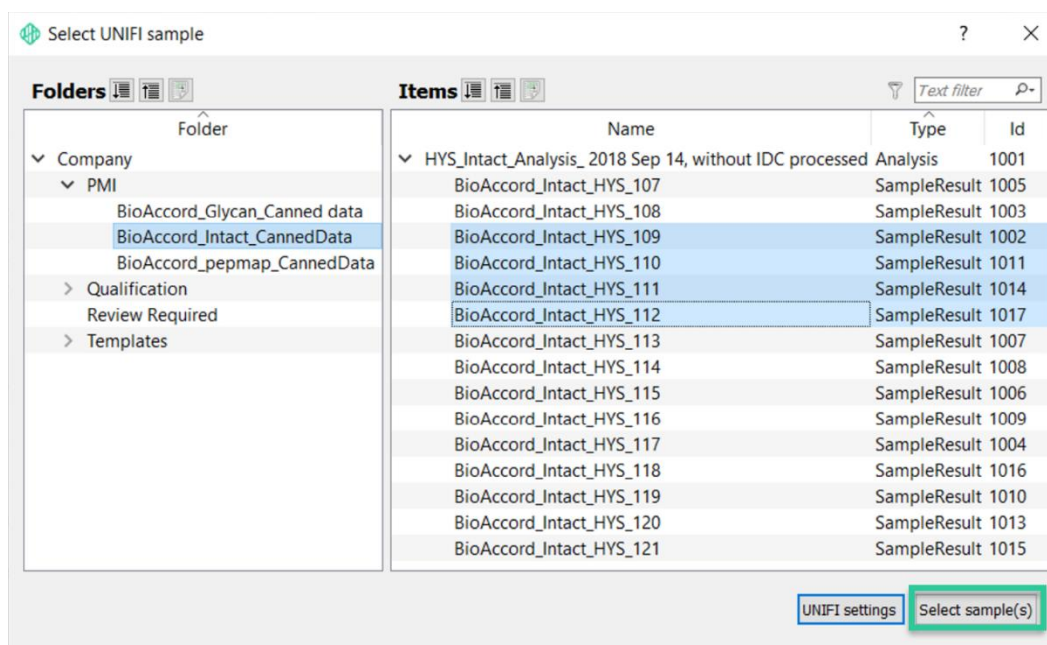


Figure 192: Select UNIFI sample dialog opened from within Byos

To learn how to access UNIFI from Byos, please refer to **PMI UNIFI Integration User Guide.docx**. Please start with the **PMI Byos UNIFI Integration Prerequisites Guide.docx**.

## Customized Workflows

Contact [support@proteinmetrics.com](mailto:support@proteinmetrics.com) for information on a variety of existing and new workflow templates.

Up to date release and product information are listed on our website. Please visit:  
<https://www.proteinmetrics.com>.

## Appendix

### Byos: Advanced Commands

Byos includes ways to customize the functionality to fit specific needs. Protein Metrics uses Advanced Commands to test new ideas, beta-test new features, and enable specialized options, without adding complexity to the graphical user interface. This section describes several text-format Advanced Commands that will enable finer control over processing. Advanced Commands may be entered during project creation in the **Advanced configuration** panel of the **Advanced** tab or after project creation by choosing **Edit -> Advanced configuration**.

1. **Factors Dialog** when enabled, processes sample names in the Samples tab to split them by a delimiter ("\_" by default), opening a preview dialog. The preview dialog contains the resulting table and features that enable users to alter delimiter, column names and columns order.

```
[BYOS]
EnableFactorsDialog=true
```