



# Byonic™

## User Manual

October 2025

5.11

Protein Metrics LLC, Boston, Massachusetts, USA

# Contents

Overview .....	3
System Specifications .....	4
Java Requirement .....	4
Byonic Workflow in Byos .....	7
Byonic Search Screen .....	8
Input and Output sections .....	9
Digestion and Instrument Parameters tab .....	10
Sample digestion .....	10
Instrument parameters .....	11
Modifications tab .....	12
Select Modifications .....	13
Glycans tab .....	19
S-S, Xlink tab .....	24
Inclusion tab .....	26
MS/MS Filtering tab .....	26
Advanced tab .....	27
Spectrum input options .....	28
Peptide output options .....	29
Protein output options .....	29
Progress tab .....	30
Additional Settings .....	30
Load, Save, and Reset Parameters .....	30
Multicore options .....	31
Run .....	31
Byonic Menu Bar .....	31
File Menu .....	31
Edit Menu .....	31
Help Menu .....	32
Byonic Results .....	33
Byos and Byonic Viewer Results .....	33
Byonic Viewer Menu Bar .....	35
File Menu .....	35
Edit Menu .....	35
Window Menu .....	35
Help Menu .....	36
Table Menus .....	36
Table Right-Click Menus .....	37

Proteins Table .....	38
Protein Coverage Table and Menu .....	38
Peptides Table and Menu.....	40
Spectrum View and Menu .....	40
Spectrum Menus .....	42
Spectrum Right-Click Menus .....	44
Excel Report .....	46
Byonic Field Descriptions .....	48
Modifications .....	50
False Discovery Rate .....	52
Appendix .....	52
Common Modifications .....	52
Byonic: Advanced Commands .....	54

## Overview

Byonic™ is a software package for identifying peptides and proteins by tandem mass spectrometry. Byonic plays the same role as Mascot, SEQUEST, and X!Tandem, but offers greater accuracy, sensitivity, and flexibility. Byonic provides several major features not found in the other search engines:

**Modification Fine Control™** enables the user to search for 10s or even 100s of modification types at a time without a combinatorial explosion. For example, a user might allow up to three phosphoserines, S[+80], per peptide, but allow at most one beta elimination, S[-18], and at most one deamidated asparagine, N[+1]. To further reduce the search, the user might allow at most one of *either* S[-18] *or* N[+1], that is, disallowing peptides containing one of each. Modification Fine Control empowers the user to tailor the search to the sample, and thus avoid overly narrow searches that miss interesting peptides and overly broad searches that run for hours or days and produce “noisy” results with many false positives.

**Wildcard Search™** enables the user to search for unanticipated, or even unknown, modifications. A wildcard can modify any residue by any mass delta within a user-settable range. Wildcard masses occur at roughly 1.0 Dalton spacing just like molecular masses. There is a limit of one wildcard per peptide.

**Glycopeptide search** enables the user to search for glycosylated peptides, without prior knowledge of either glycosylation sites or glycan masses. Byonic now offers three ways to specify glycopeptide searches: internal tables, external tables, and modification fine control. Byonic's internal tables contain the most likely N- and O-linked glycan compositions. The other two options allow the user to customize the list of glycans and/or allow more than one glycan per peptide.

**Top-down, middle-down, and bottom-up proteomics** – Byonic is uniquely capable of top-down, middle-down, and bottom-up searches, but it relies upon isotope-resolved precursors in order to determine precursor ion charges. Without isotope-resolved precursors, the user will need to assign charge and set a large precursor mass tolerance.

**Disulfide bonds, trisulfide bonds, and general crosslinking** – Byonic has disulfide bond, trisulfide bond, and general crosslink search capability. It is designed to search for all disulfide pairs (both expected and shuffled disulfide bonds) and the user can constrain which protein chains to consider when doing the search.

## System Specifications

- **Recommended PC:**
  - Windows 10/11 64-bit (Windows 10 must be version 1809 or higher)
  - 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

Two Xeon CPUs (recommend 8 to 16 physical cores)

## Java Requirement

Java is required to run Byonic. Either Oracle JRE or OpenJDK needs to be installed. Byonic uses the system environmental variable PMI\_JAVA\_HOME to determine the version of Java to use. If undefined, then the default Oracle Java registered in the system is used.

To load Java:

- Oracle JRE:

Download and install the 64-bit version from <https://www.java.com/en/>

- OpenJDK:
  1. If it does not already exist, create the subfolder “java” under the “C:\Program Files” folder. This requires Administrator privileges on the machine.

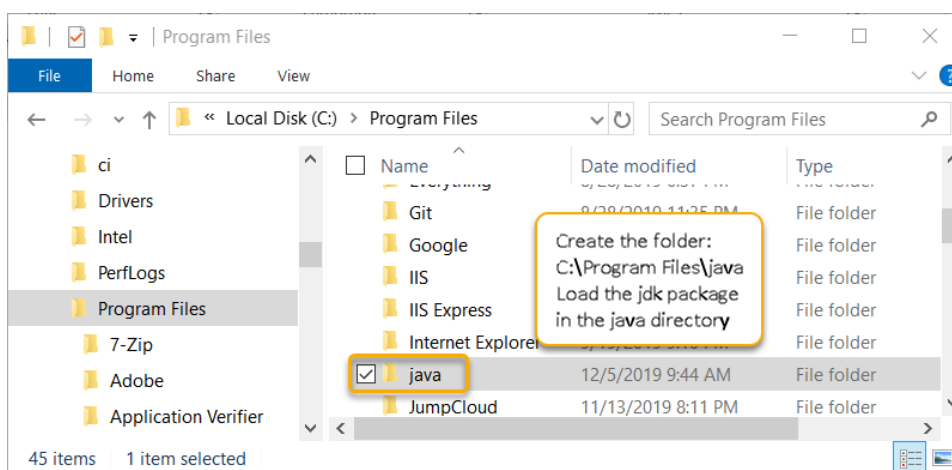


Figure 1: Create a “java” folder

2. Go to <https://jdk.java.net/archive/> and download the latest Windows/x64 version. **Note:** OpenJDK version 13.0.1 is the current general-availability release when this manual was written. The older versions 11.0.2 and 12.0.2 are also supported.
3. Unzip and copy the content of the downloaded package to “C:\Program Files\java”.

4. Create the system environmental variable 'PMI\_JAVA\_HOME' and set the value to the OpenJDK installation directory: e.g: PMI\_JAVA\_HOME=C:\Program Files\java\jdk-13.0.1. In older versions of Windows, open **Control Panel -> System Properties -> Advanced -> Environmental Variables**. In Windows 10, open **Control Panel -> System and Security -> System** and click **Advanced system settings**.

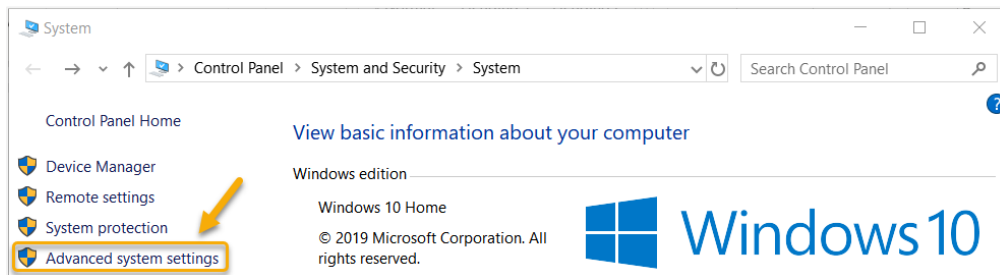


Figure 2: Windows 10 System settings

5. In the Advanced tab, click **Environment Variables**.

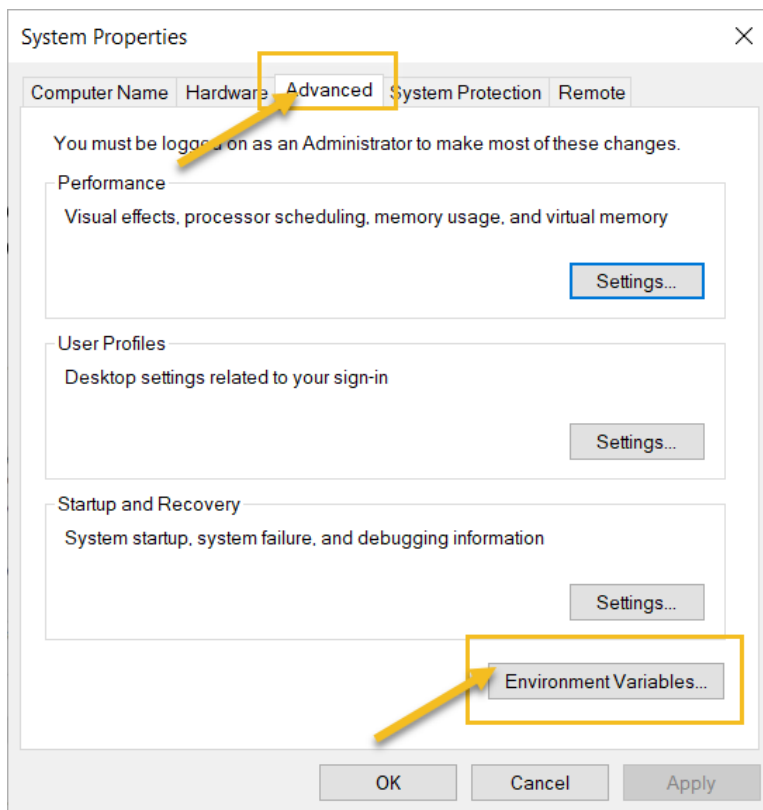


Figure 3: Advanced System Properties from the Control Panel

6. In the Environmental Variables dialog, click **New**.

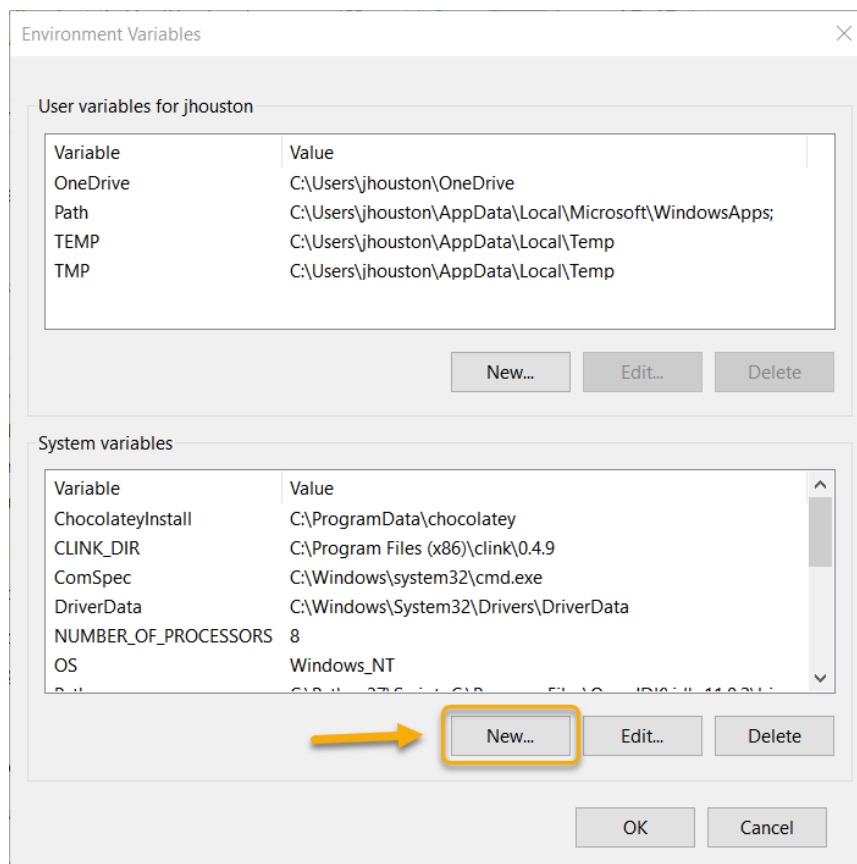


Figure 4: Environmental Variables dialog

7. In the Edit System Variable dialog, enter Variable name "PMI\_JAVA\_HOME". Click **Browse Directory** and navigate to the jdk folder (C:\Program Files\java\jdk-XX.X.X) and click **Open**.

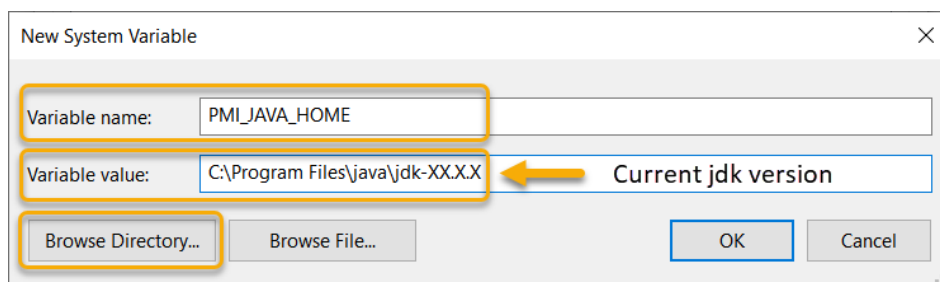


Figure 5: New System Variable creation

Click **OK** to create the variable.

Note: If the PMI\_JAVA\_HOME variable name already exists with an older version of jdk that was replaced, select the variable name, click **Edit**, click **Browse Directory**, navigate to the jdk folder, click **Open**, and then click **OK** to save the edited variable value.

## Byonic Workflow in Byos

As of the release of Byos v5.0 a **Byonic workflow** is now available in **Byos**.

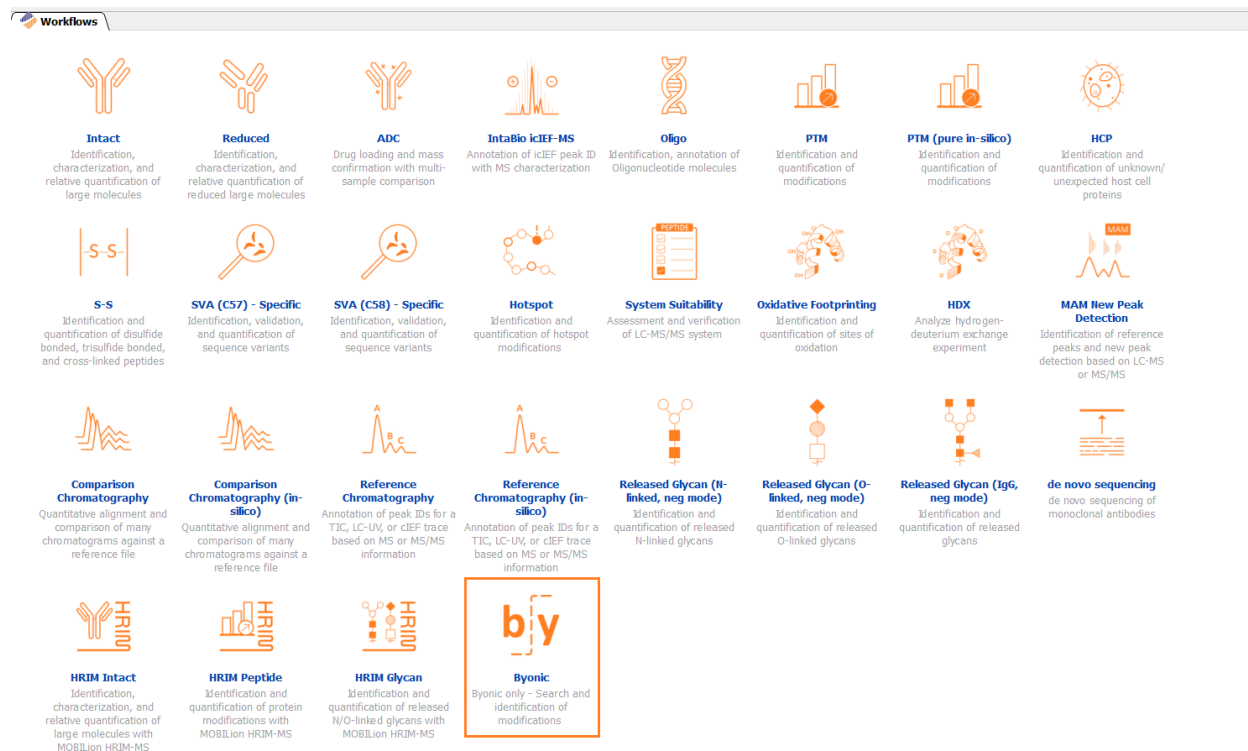


Figure 6: Byonic workflow icon in Byos

The Byonic workflow includes the following settings found under the Processing Nodes tab. These menu items emulate the Byonic tabs interface:

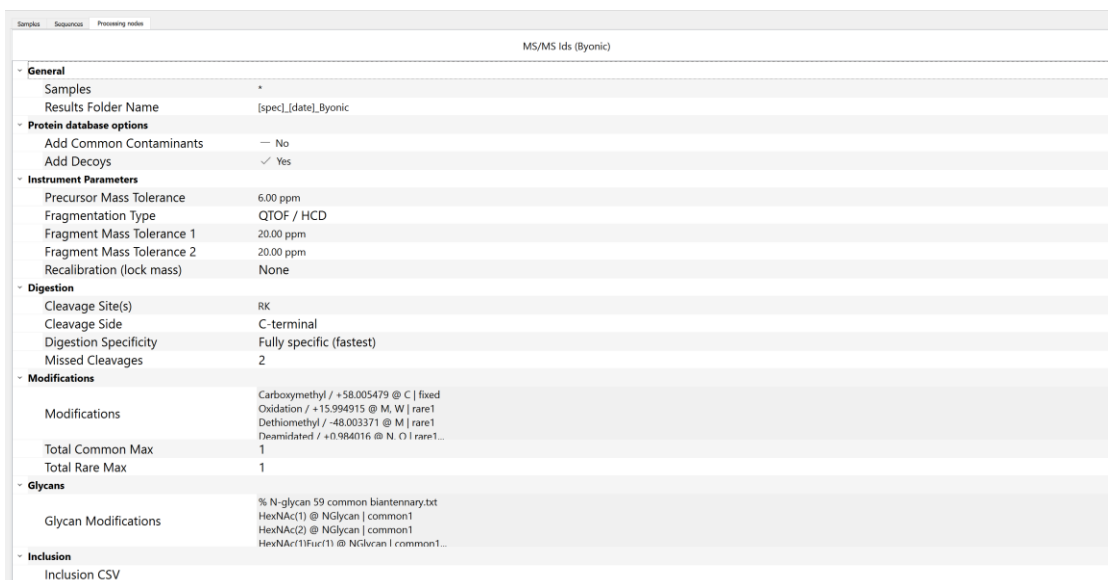
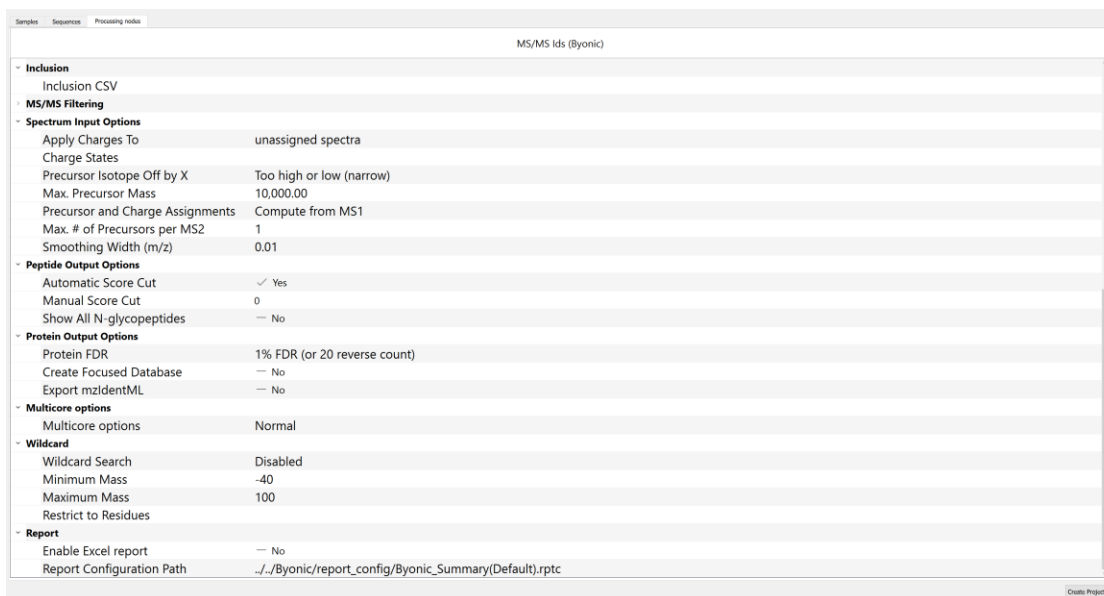


Figure 7: Byonic settings in Project Creation






MS/MS Ions (Byonic)

- Inclusion**
  - Inclusion CSV
- MS/MS Filtering**
- Spectrum Input Options**
  - Apply Charges To: unassigned spectra
  - Charge States
  - Precursor Isotope Off by X: Too high or low (narrow)
  - Max. Precursor Mass: 10,000.00
  - Precursor and Charge Assignments: Compute from MS1
  - Max. # of Precursors per MS2: 1
  - Smoothing Width (m/z): 0.01
- Peptide Output Options**
  - Automatic Score Cut: ☒ Yes
  - Manual Score Cut: 0
  - Show All N-glycopeptides: ☐ No
- Protein Output Options**
  - Protein FDR: 1% FDR (or 20 reverse count)
  - Create Focused Database: ☐ No
  - Export mzIdentML: ☐ No
- Multicore options**
  - Multicore options: Normal
- Wildcard**
  - Wildcard Search: Disabled
  - Minimum Mass: -40
  - Maximum Mass: 100
  - Restrict to Residues
- Report**
  - Enable Excel report: ☐ No
  - Report Configuration Path: ./Byonic/report\_config/Byonic\_Summary(Default).rptc

Create Project...

Figure 8: Byonic settings in Project Creation continued

As of Byos v5.5, a new report template is now available for Byonic-only workflows, titled Byonic\_Summary(Default).rptc. The user must click on the **Report**  icon to create the report.

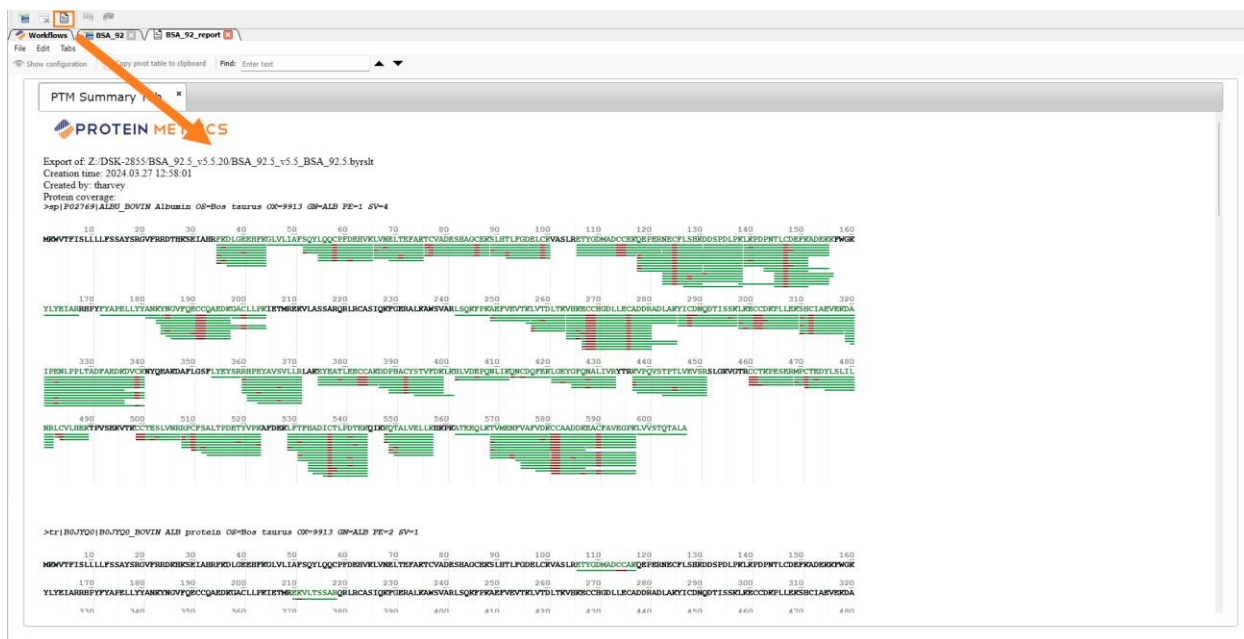


Figure 9: Byonic report

## Byonic Search Screen

An example of the Byonic search screen is shown in the figure below:

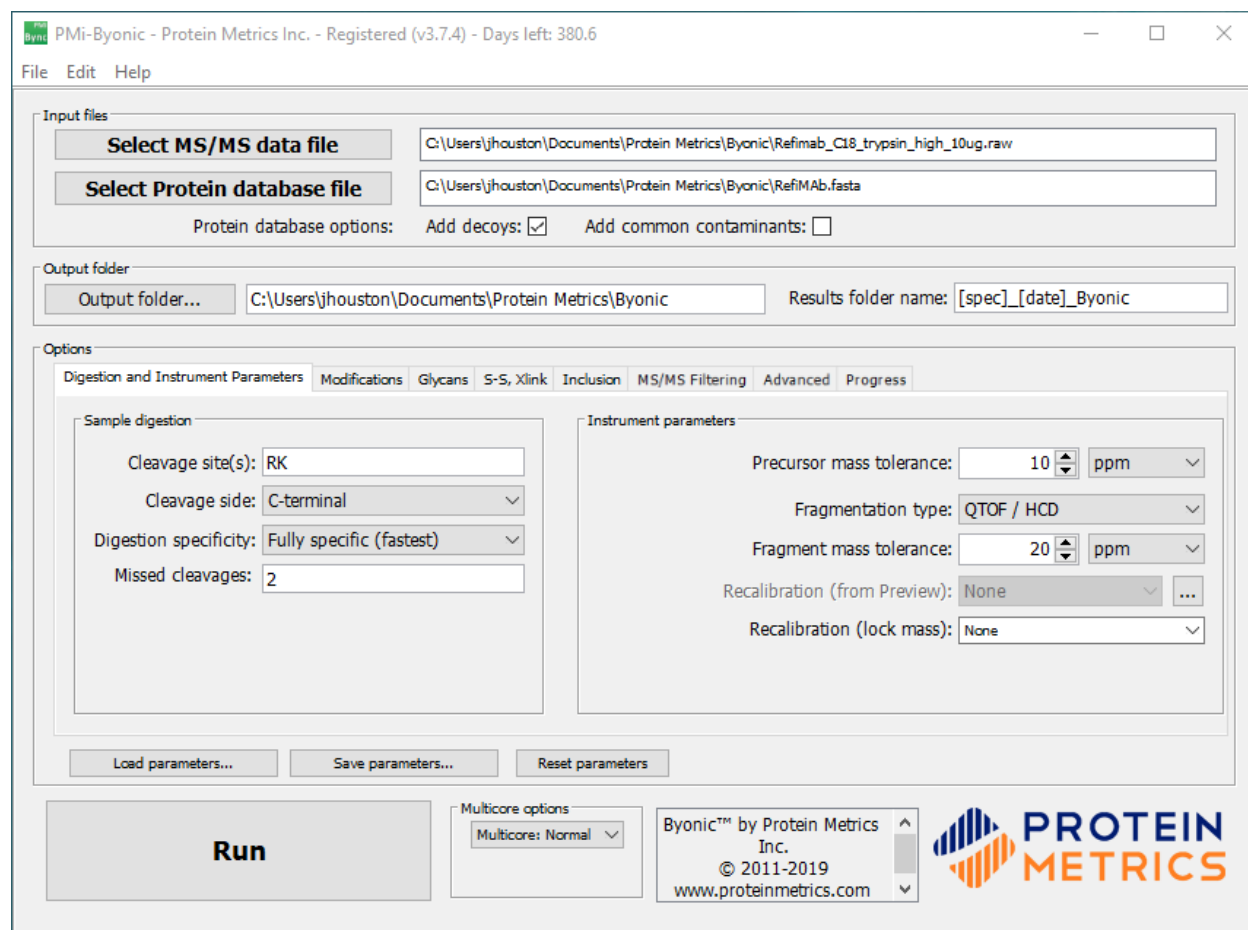


Figure 10: Byonic search screen

The search screen includes input file management, output file management, several tabs for setting search options or parameters, and finally, search management actions.

## Input and Output sections

The **Input files** section at the top locates the spectrum data file and protein database file, along with database search options:

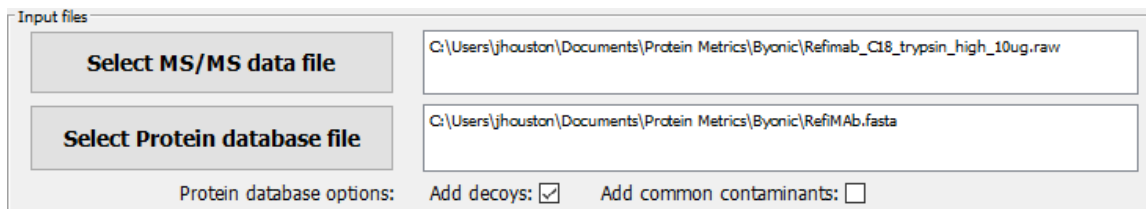


Figure 11: Input files section

To load a mass spectrum data file, click **Select MS/MS data file**, navigate to the desired file, and click **Open**. The spectrum data file can be one of a number of standard formats: \*.MGF, \*.mzML, \*.mzXML, Thermo \*.raw, Waters \*.raw, Sciex \*.wiff, Bruker \*.d, Agilent \*.d, and Shimadzu \*.lcd, etc.

To load a protein database, click **Select Protein database file**, navigate to the desired FASTA file, and click **Open**. 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.

In the **Output folder** section, appearing next, the user selects the directory under which a results folder will be created and chooses the naming convention for that folder and the results files:

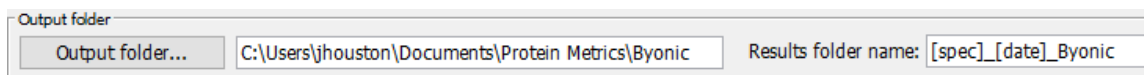


Figure 12: Output folder section

To select the directory where the Byonic search output folder will be created, click **Output folder**, navigate to the target folder, and click **Open**. (Due to a Windows DOS limitation, the total folder path name must be less than 256 characters; please note long .raw file names may exceed the 256-character limit) The standard output folder is C:\data\_results. In the **Results folder name** cell, the user can choose the format for naming the search output folder and the search result files saved to that folder. Free text will appear in folder and file names as written. The items in square brackets represent supported objects that can become part of these names. The supported objects are:

- [spec] the name of the mass spectrum data file,
- [fasta] the name of the FASTA file,
- [date] the current date (year, month, day)
- [time] the current time (hour, minute)

In the example in Figure 12, if the mass spectrum data file was named “project123.d” and the search was run on 02/15/2020, the output folder name would be “project123.d\_20200215\_Byonic”.

## Digestion and Instrument Parameters tab

The first tab in the **Options** section is the **Digestion and Instrument Parameters** tab:

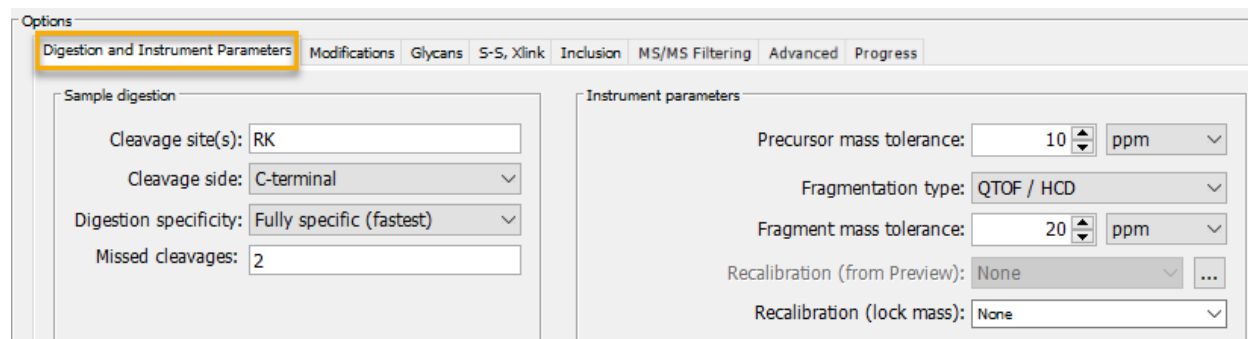


Figure 13: Digestion and Instrument Parameters tab

### Sample digestion

The **Sample digestion** section allows the user to set the residues recognized by the digestion enzyme as shown. For example, in Figure 13, the enzyme is trypsin, so the user entered RK for arginine and lysine and chose C-terminal for the cleavage side.

- **Cleavage site(s)** identifies specific cleavage sites. When the Cleavage site(s) box is blank, specific cleavage sites are limited to protein termini.

- **Cleavage side** designates the C-, N- or both termini for the cleavage side. For example, RK; DE for Cleavage site(s) along with C-terminal; N-terminal for Cleavage side would be appropriate for digestion by a combination of trypsin and Asp-N.

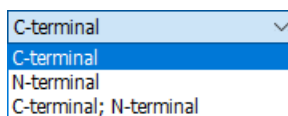


Figure 14: Cleavage side options

- **Digestion specificity** balances search specificity and performance. For example, the “Fully specific (fastest)” search selected in the figure below, requires that both the N- and C-terminal cleavages be C-terminal to R or K. Nonspecific cleavage at either or both endpoints is supported.

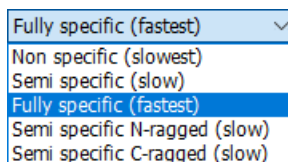


Figure 15: Digestion specificity

A nonspecific search will search all peptides but favors specified digestions. In Figure 13, Cleavage Site(s) is set to RK which searches all peptides but favors tryptic peptides. For a true no-enzyme search, leave the Cleavage Site(s) box blank.

- **Missed cleavages** limits the maximum number of missed cleavages. In Figure 13, a setting of two missed cleavages limits internal R's and K's not followed by P to 2. The default value for Missed cleavages is -1, which allows any number of internal R's and K's.

## Instrument parameters

The **Instrument parameters** section sets values controlling precursor mass, fragmentation and recalibration. Both Dalton and ppm mass tolerances for precursors and fragments are supported, along with several fragmentation types

- **Precursor mass tolerance** sets the maximum mass difference between observed and theoretical (calculated) peptide mass.
- **Fragmentation type** sets the model(s) used to scan and score fragment peaks.

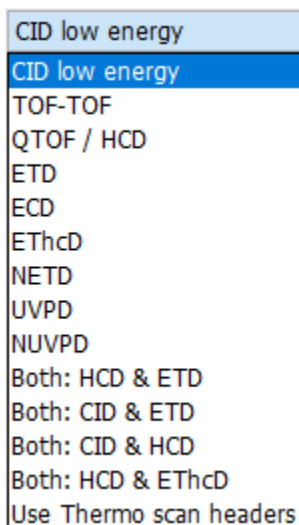


Figure 16: Fragmentation type

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 these 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.

If a spectrum file contains more than one fragmentation type, for example, alternating HCD and ETD on a Thermo Orbitrap, Byonic will match the labels on the scans to the chosen fragmentation types, so that the MS/MS scans will be scored with the appropriate models. If a scan does not match the selected model, it will be ignored. For example, if the .raw file contains both HCD and ETD scans, the user must select “Both: HCD & ETD” or “Use Thermo scan headers” in order to score both types of scans with the correct model. “Use Thermo scan headers” is a new option that supports all standard combinations of ETD, CID, and HCD, along with choices of mass analyzers. For example, Byonic can now handle “tribrid” data with Orbitrap HCD and EThcD and ion trap CID.

If a spectrum file contains only a single fragmentation type, or as in the case of \*.mgf files, does not specify fragmentation type, then the user can choose any single fragmentation model for scoring. For example, HCD can be scored as CID low energy, or EThcD scored as pure ETD. Switching fragmentation models is occasionally advantageous, for example, when EThcD scans use extremely low collision energy.

- **Fragment mass tolerance** is the maximum difference in m/z for matching fragments. Byonic supports both relative (ppm) and absolute (Dalton) tolerances. Relative tolerances are generally recommended for time-of-flight and FTMS mass analyzers, and absolute tolerances for ion trap mass analysis. Fragment tolerances of 0.1 Da or 100 ppm or less apply to high-resolution MS/MS analysis, providing sufficient resolution to distinguish charge states of fragment ions. Fragment tolerances larger than 0.1 Da should be used with low-resolution (ion trap) MS/MS analysis.

## Modifications tab

The **Modifications tab** defines specific modification parameter values and Wildcard searches.

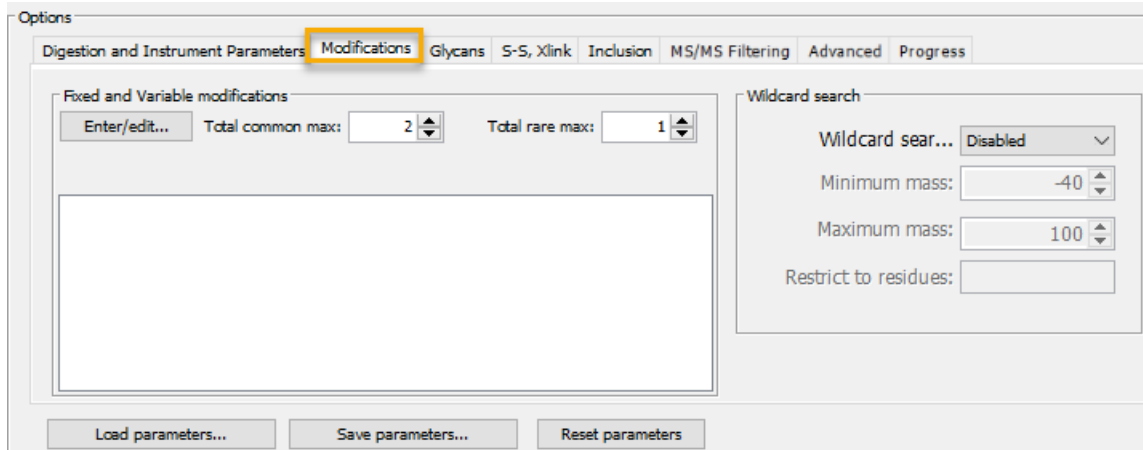


Figure 17: Modifications tab

Like most proteomics search engines, Byonic supports two types of modifications: 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.

Byonic also offers a unique feature not found in other search engines: the user designates each variable modification as either “**common**” or “**rare**”, with the names suggesting their use. Byonic has separate limits on the number of occurrences of each variable modification, so that “common2” means at most two occurrences per peptides. Byonic also has separate limits on 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.

## Select Modifications

To add fixed and variable modifications, click **Enter/Edit** to open the **Select Modifications** dialog:

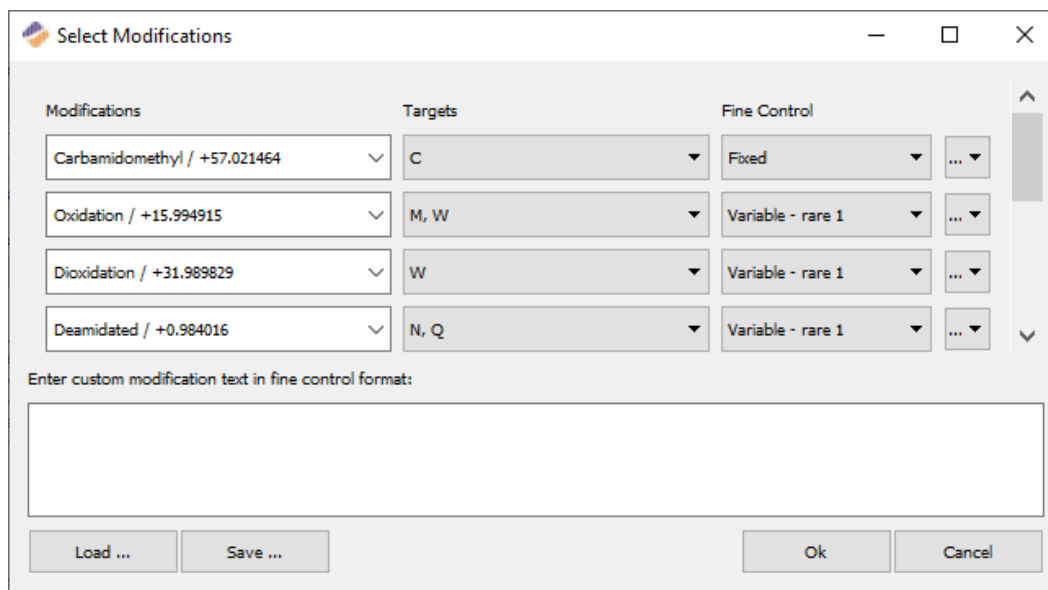


Figure 18: Select Modifications dialog

The modification set can then be edited and extended. Modifications include name and delta mass pairings, modification targets and fine control settings that set modifications as fixed or variable, as well as rare or common. Additional actions can also be performed on the modifications.

- **Modifications** displays a list of common modifications and their corresponding delta masses. The list includes all candidates found in [www.unimod.org](http://www.unimod.org).

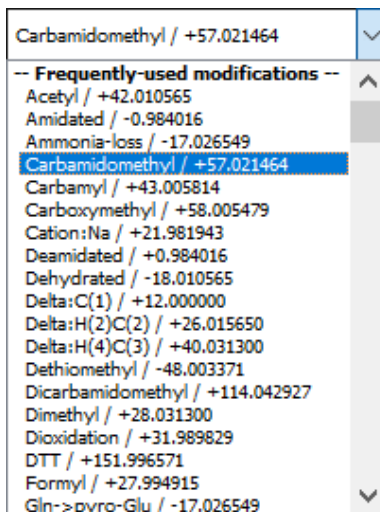


Figure 19: Modification Name / Mass Delta

For convenience, frequently used modifications are listed twice, once at the top and again in the complete list.

- **Targets** displays a list of possible target locations associated with the selected modification name. 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.

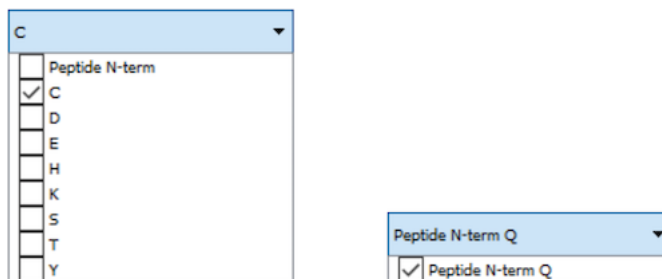


Figure 20: Modification Targets

- **Fine control** marks the modification as “Fixed” or “Variable”, whether the variable modification is considered “Rare” or “Common”, and the maximum count of that modification in the peptide.

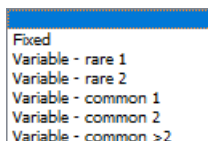



Figure 21: Fine Control

- The  button in the last column exposes additional actions that can be performed on modifications:

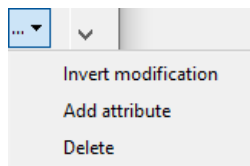


Figure 22: Additional Modification actions

- Invert modification** defines a variable modification that is the removal of a fixed modification (for example, under-alkylation. The modification is labeled with the prefix “(De)”, as shown in the figure below:

Modifications	Targets	Fine Control	
(De) Carbamidomethyl / -57.021464	C	Fixed	...
Oxidation / +15.994915	M, W	Variable - rare 1	<input checked="" type="checkbox"/> Invert modification
Dioxidation / +31.989829	W	Variable - rare 1	Add attribute
			Delete

Figure 23: Invert modification

- Add attribute** allow the user to define protein-specific modifications. For example, adding the attribute “ProteinLabel{collagen}” to Oxidation on P allows hydroxyproline only on proteins with “collagen” in their protein names:

Modifications	Targets	Fine Control	
Carboxymethyl / +58.005479	C	Fixed	...
Oxidation / +15.994915	P	Variable - rare 1	<input type="checkbox"/> Invert modification
Dioxidation / +31.989829	W	Variable - rare 1	Add attribute
			Delete

Figure 24: Add modification attribute

- Delete** removes that modification record.

To add a new modification, scroll to the bottom of the modification records and add entries for Modifications, Targets and Fine Control:



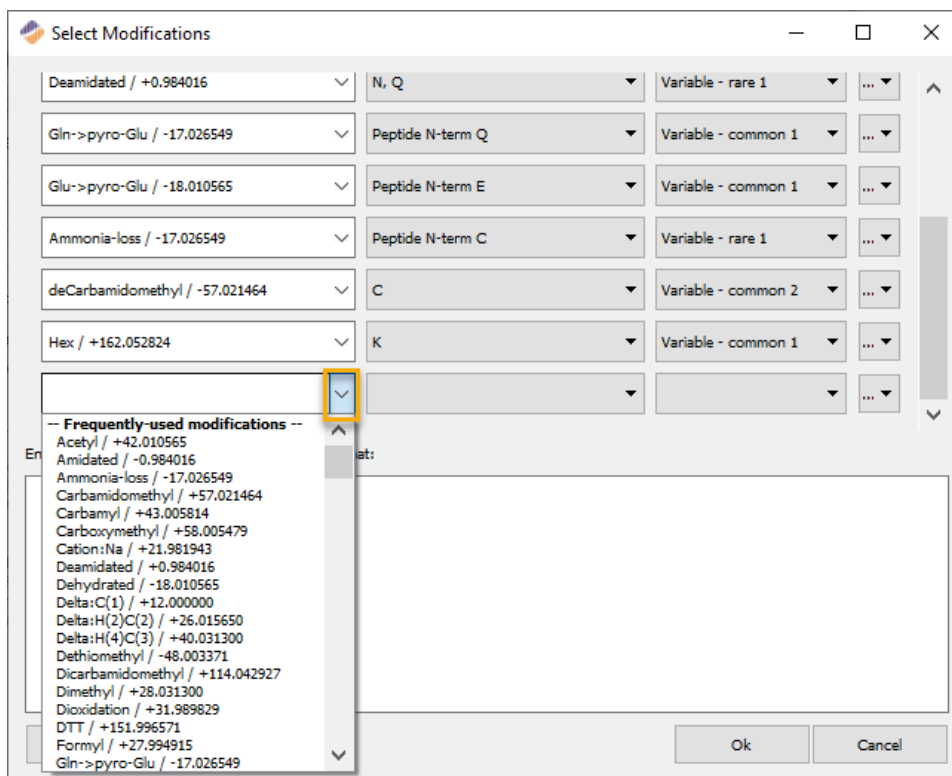


Figure 25: Adding a new modification

An entry for Modification name is optional. The resulting modification record has the format:

```
Modification_Name / Mass_Delta @ Targets | Fine_Control
```

Custom modifications can be manually entered in this format in the box labeled **Enter custom modification text in fine control format**. The following is an example of a real modification not (yet) in Unimod:

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

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).

Click **Save** to export the selected and custom modifications as a text file. This text file can then be loaded into another Byonic session by clicking **Load**. When the list of modifications is complete, click **Ok**. The modification strings are now listed in the Modifications tab.

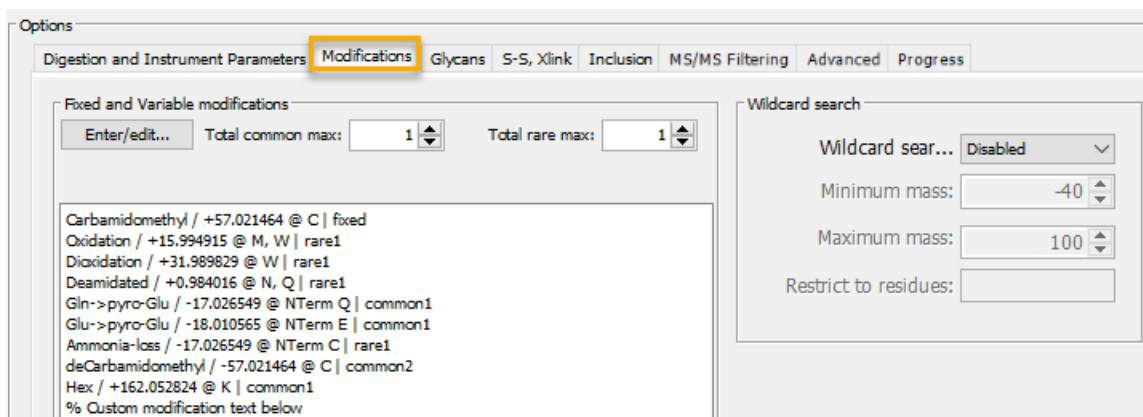


Figure 26: Example set of Fixed and Variable modifications.

To set maximum allowed modifications, set values for **Total common max** and **Total rare max**. To set no maximum, leave the value at zero.

In Figure 26, the user specified Carbamidomethyl / 57.021464 @ C | fixed, meaning carbamidomethylated cysteine (camC). The user also specified Oxidation / +15.994915 @ M | common2, directing the program to consider each methionine residue with and without this modification, up to a limit of 2 such modifications per peptide. In addition, the user specified Ammonia-loss / -17.026549 @ N-term C | rare1, indicating that the program also considers this modification for any N-terminal cysteines as a rare variable modification. Variable modifications are added on top of fixed modifications, so the total mass added to these N-terminal C's will be 57.021+14.016 = 71.037, which represents cysteine propionamide. One way to represent incomplete carbamidomethylation is with these two rules: Carbamidomethyl / +57.021464 @ C | fixed and (De)Carbamidomethyl / -57.021464 @ C | common2.

The rule Carbamidomethyl / +57.021464 @ NTerm | rare1 specifies a common artifact (over-alkylation) on the peptide N-terminus.

The next two rules, +0.984016 @ N | common2 and +0.984016 @ Q | common1, represent deamidation; here the user is allowing up to two deamidated asparagines (the more common deamidation) but only one deamidated glutamine per peptide.

The rule Gln->pyro-Glu / -17.026549 @ NTerm Q | rare1 specifies a modification that occurs only on peptides with N-terminal glutamine.

Conceptually, Byonic 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.

Byonic can support a limited number of **non-standard amino acid residues** by redefining one-letter amino acid abbreviations using fixed modifications. Starting with version 2.9.77, Byonic accepts B, Z, U, O, J, and X in FASTA protein databases. The non-standard amino acid masses are:

Abbreviation	Mass	Comments
B	114.042927	Same as N
Z	128.058578	Same as Q
U	150.95363	Selenocysteine
O	237.147730	Pyrrolysine
J	100.0	
X	110.05	Close to averagine

For example, by placing a fixed modification of +13.04768 on J, the user can make J in a FASTA database have mass 113.04768, correct for hydroxyproline. Byonic does, however, use amino acid sequence 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 **backwards compatibility**, Byonic still accepts the syntax from earlier versions:

```
[Residues][Mass_Delta], Fine_Control
```

Examples of the earlier syntax are [ST][+79.966], common2 for phosphorylation and N-terminal S[+9.984], rare1 for DehydroFormyl.

The **Wildcard search** section of the Modifications tab lets the user turn on wildcard searches, set the range for the wildcard mass, and restrict the wildcard to certain residues if desired.

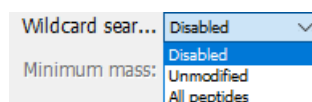


Figure 27: Wildcard search

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. When an invalid entry for Restrict to residues will give an error message designating valid entries. The tooltip also specifies the entry format:

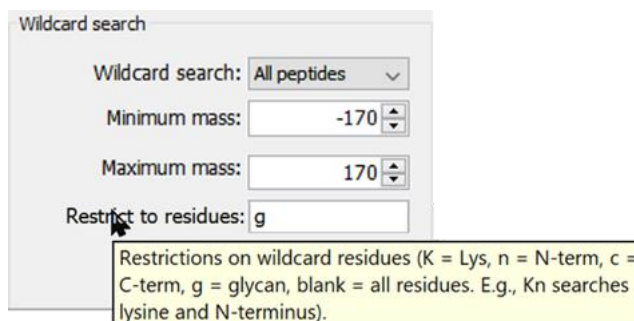


Figure 28: Restrict to residue entry format

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  $C^T$  where  $C$  is the number of common modifications enabled and  $T$  is Total common max.

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

The wildcard search has been expanded to amino acid positions with an already present glycan mass modification. N-linked, O-linked, and other glycan motifs are supported.

The user must use the below parameters:

- **Wildcard search:** All peptides
- **Minimum** and **Maximum mass:** Set values
- **Restrict to residues:** set to "g" to search for wildcard modifications on top of glycan modifications.

## Glycans tab

The **Glycans tab** allows the user to associate a glycan database with the Byonic search:

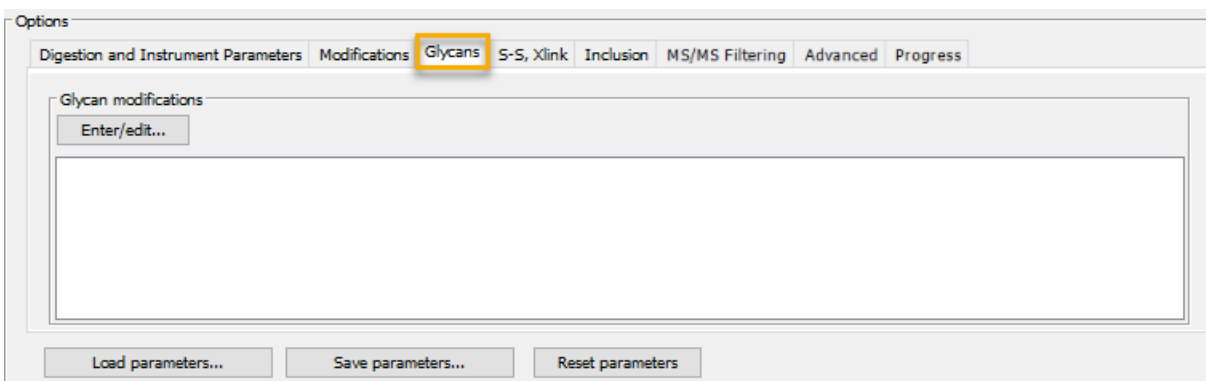


Figure 29: Glycan tab

Byonic offers three ways to define glycan modifications: additions from internal preset tables, imports from external glycan databases, and user-defined glycans. To add glycans, click **Enter/Edit** to open the **Select Glycans** dialog:

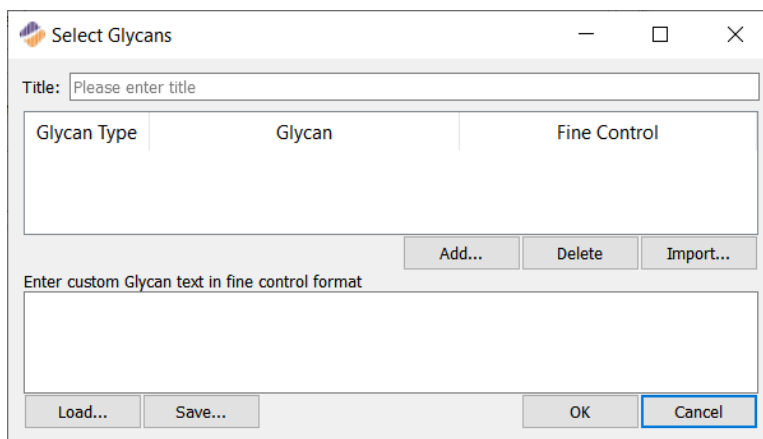
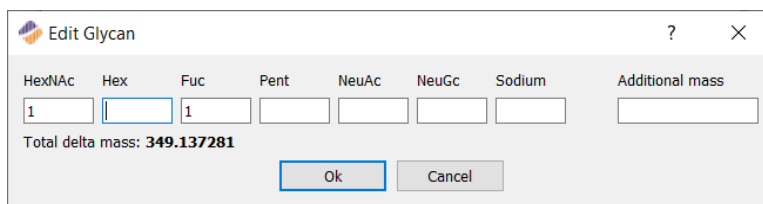


Figure 30: Select Glycans

Select Glycans allows the user to set glycan types, import glycans from a glycan database text file, and define variables as rare or common. The dialog also allows the free entry of glycans in the proper format.

To enter glycans from the internal preset tables, click **Add**. The **Edit Glycan** dialog opens:



The 'Edit Glycan' dialog box contains input fields for the following components:

HexNAc	Hex	Fuc	Pent	NeuAc	NeuGc	Sodium	Additional mass
1	1	1					

Total delta mass: 349.137281

Buttons: Ok, Cancel

Figure 31: Adding a glycan in the Edit Glycan dialog

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 this is a common adduct on sialic acids. Unused monosaccharides can be left blank or included with zero (0) occurrences. In the resulting text, spaces between monosaccharides units are optional. For example, five common human O-glycans are represented as:

HexNAc (1)

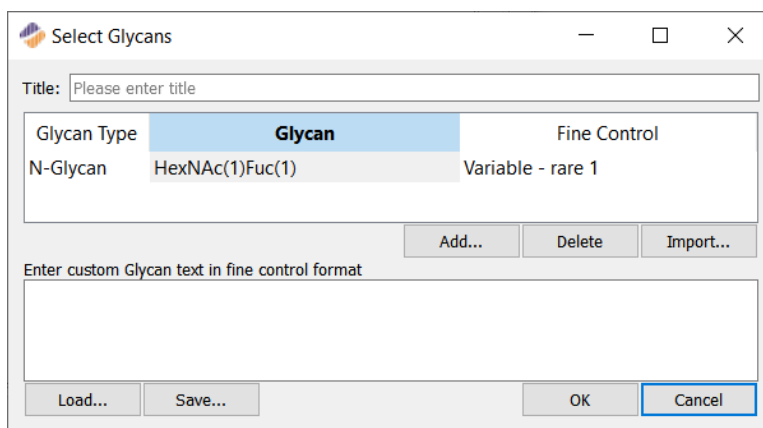
HexNAc (1) Hex (1)

HexNAc (1) Hex (1) NeuAc (1)

HexNAc (1) Hex (1) NeuAc (2)

HexNAc (1) Hex (1) Fuc (1)

**Total delta mass** displays below for the added monosaccharides to six decimal places. Glycan modifications such as acetylation, phosphorylation, or sulfation or less common monosaccharides such as deaminated neuraminic acid can be added by typing in the correct keywords, “Acetyl”, “Phospho”, “Sulfo”, and “Kdn” respectively. Byonic scoring and annotation will consider peaks for likely placements of these modifications, for example, Acetyl on NeuAc. We are constantly adding to our list of keywords, so it is a good idea to check with [support@proteinmetrics.com](mailto:support@proteinmetrics.com) for unusual glycan components. To support as-yet-undefined glycan components, Byonic also allows arbitrary mass deltas to be input to the Additional mass box. Keywords are preferred over delta masses; a delta mass of 42.010565 does not give the same scoring and spectrum peak annotation as a keyword of “Acetyl”. Click **Ok** to create the glycan in the Select Glycans dialog:



The 'Select Glycans' dialog box shows the following structure:

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

Buttons: Add..., Delete, Import...

Enter custom Glycan text in fine control format

Buttons: Load..., Save..., OK, Cancel

Figure 32: Glycan components added to the Select Glycans dialog

The new glycan can be further modified. To change the attachment type, double-click in the cell under **Glycan Type** and click the dropdown arrow:

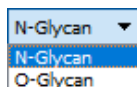


Figure 33: Glycan Type

Glycan type sets it as either **N-Glycan** or **O-Glycan**. N-glycans are allowed on  $Nx\{S/T\}$  motif where x is anything except P. O-glycans are allowed on any S/T.

Fine Control designates the glycan as common or rare and sets a limit on the count of that glycan per peptide. To change the fine control, double-click in the cell under **Fine Control** and click the dropdown arrow:

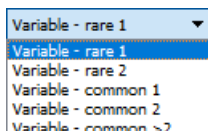


Figure 34: Glycan Fine Control options

To edit the glycan components, double-click in the cell under **Glycan** to open the **Edit Glycan** dialog again. Click **Delete** to remove the selected glycan record.

The second way to generate glycans is to load them from a glycan database text file. To load a set of glycans from a glycan DB file, click **Import** (Note: **Load** is no longer used since this only creates a glycan file path reference, and this mechanism is no longer used.):

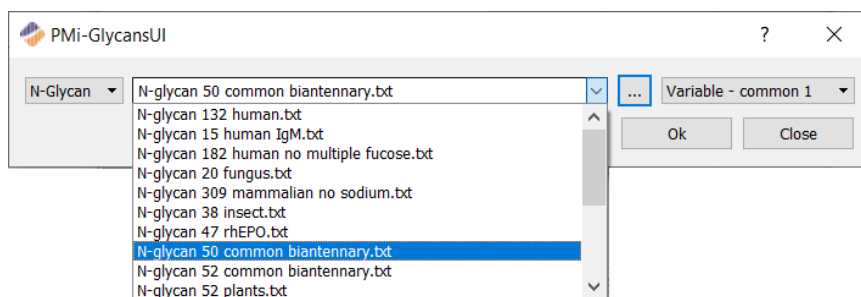


Figure 35: Glycan database options

The database dropdown references a list of glycan database text files found in `Program Files\ProteinMetrics\PMI-Suite\Tools\Byonic\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 Byonic). 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.

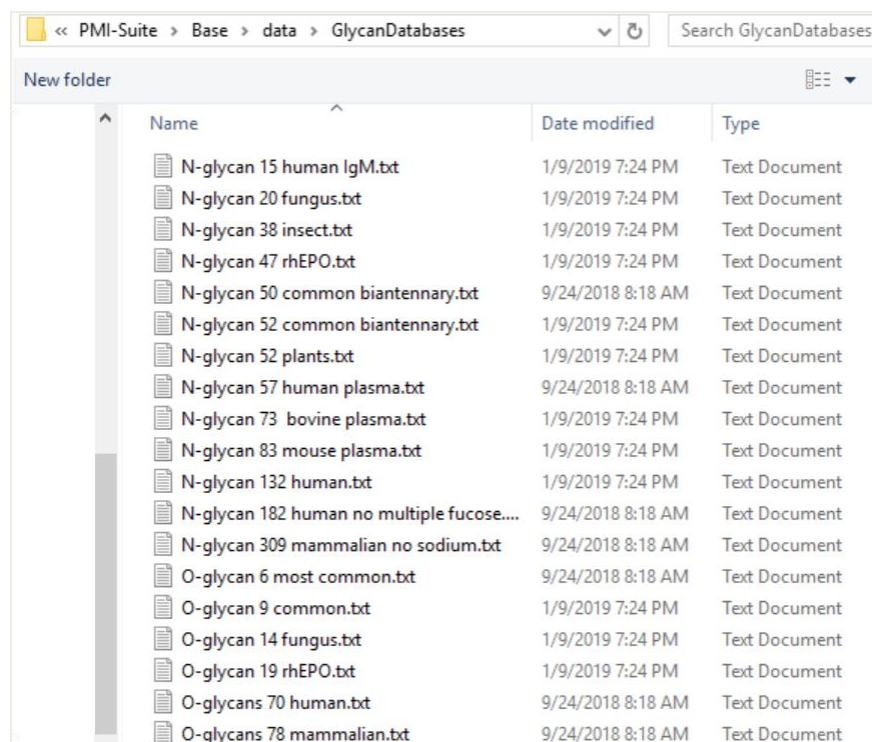



Figure 36: Glycan databases included with installation

Alternatively, a glycan database file can be selected from another directory. To open a new glycan DB, click the  button, navigate to the desired directory, select database text file, and click **Open**.

The first dropdown sets the Glycan Type for all the DB text file glycans as either N-Glycan or O-Glycan. The last dropdown selects the same Fine Control setting for all the DB text file glycans. The options are described above. Click **Ok** to load the database of glycans into the Select Glycans dialog:

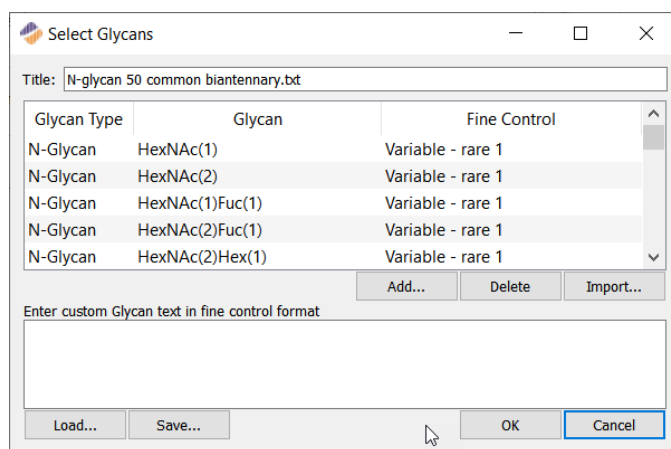


Figure 37: Glycans loaded from a glycan DB text file

Double-click in the **Glycan Type**, **Glycan** or **Fine Control** cells to modify any of these value for specific glycans.

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

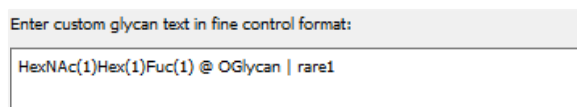


Figure 38: Manually entered or pasted custom glycan

As with modifications, glycans can be saved and reloaded in other Byonic searches. Click **Ok** to add the various glycans to the Byonic query:

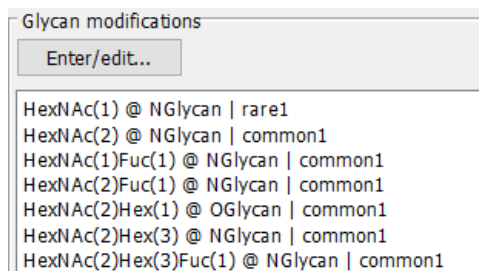


Figure 39: Example of entered glycans

Legacy Glycan DB file references can be converted into a list of custom glycans:

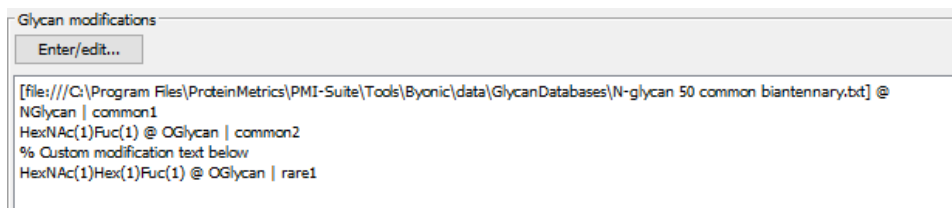


Figure 40: Legacy Glycan parameter showing glycan DB file path

When a glycan DB file and path are shown in the Glycan modifications window, click **Enter/edit** to open a dialog to convert the reference to custom glycans:

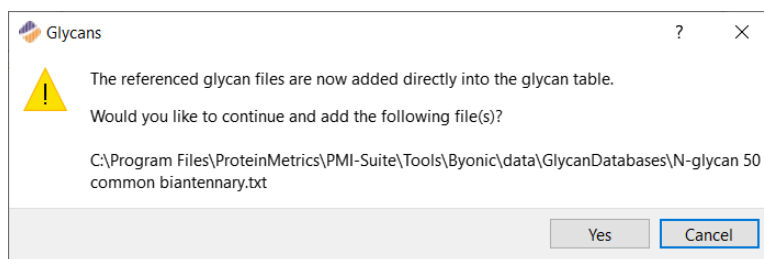


Figure 41: Converting glycan DB file reference to glycan set

Click **Yes**, and the contents of the glycan DB file are converted into custom glycans:



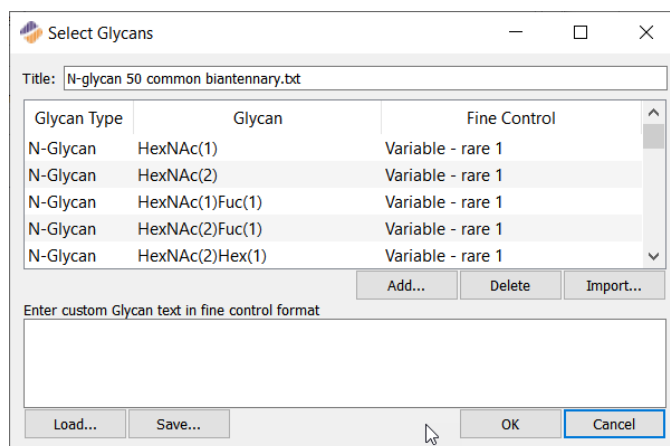


Figure 42: Glycan set generated from a legacy glycan DB file reference

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:

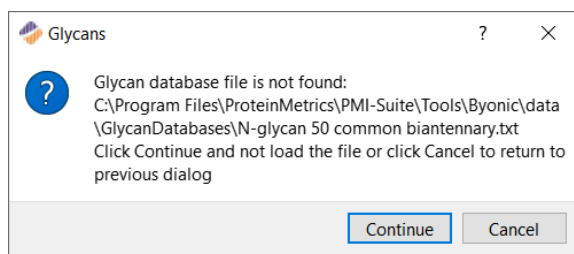


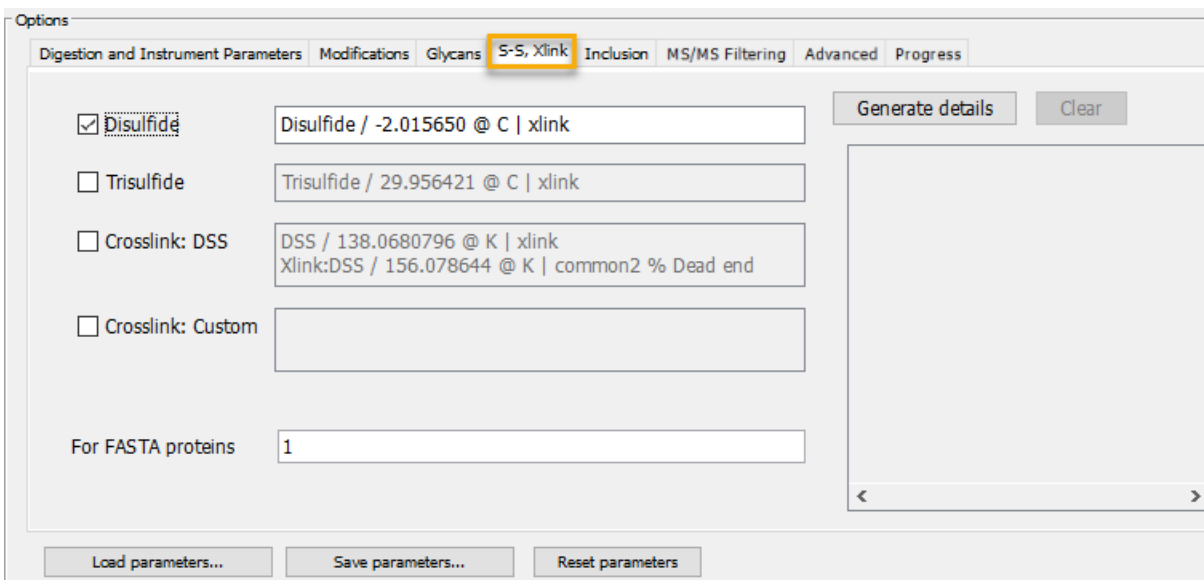
Figure 43: Click **Continue** to clear an obsolete glycan DB file reference

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.

Both “OGlycan” and “NGlycan” are modification sites in the sense that modifications other than glycosylation may make use of these keywords. For example, “Heavy deamidation / 2.98826 @ NGlycan | common2” would allow deamidation in O18 water only at N within Nx{S/T}. Conversely, one can place glycans on glycosylation sites besides NGlycan and OGlycan. For example, “HexNAc(1)Hex(1) @ Y | common2” allows (O-linked) glycosylation on Y. 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/#application-notes>.

## S-S, Xlink tab

The **S-S, Xlink** tab allows the user to search for disulfide-bonded peptide pairs, trisulfide-bonded (also called persulfide-bonded) pairs, and more general cross-linking. This tab provides options to allow a user to search for expected and unexpected disulfide bonds.



Options

Digestion and Instrument Parameters Modifications Glycans **S-S, Xlink** Inclusion MS/MS Filtering Advanced Progress

☒ **Disulfide** Disulfide / -2.015650 @ C | xlink

☐ **Trisulfide** Trisulfide / 29.956421 @ C | xlink

☐ **Crosslink: DSS** DSS / 138.0680796 @ K | xlink  
Xlink:DSS / 156.078644 @ K | common2 % Dead end

☐ **Crosslink: Custom**

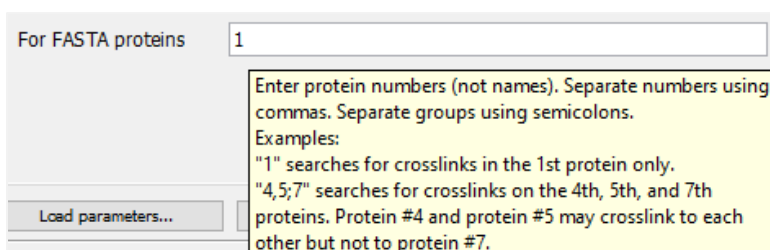
For FASTA proteins 1

Generate details Clear

Load parameters... Save parameters... Reset parameters

Figure 44: S-S, Xlink tab.

Check the **Disulfide** box on the top left to attempt to pair each peptide that contains a cysteine with each other cysteine-containing peptide in an in-silico digestion. To indicate which protein sequences from the FASTA database to consider for pairings, enter the protein numbers in the **For FASTA proteins** box:



For FASTA proteins 1

Enter protein numbers (not names). Separate numbers using commas. Separate groups using semicolons.  
Examples:  
"1" searches for crosslinks in the 1st protein only.  
"4,5;7" searches for crosslinks on the 4th, 5th, and 7th proteins. Protein #4 and protein #5 may crosslink to each other but not to protein #7.

Load parameters...

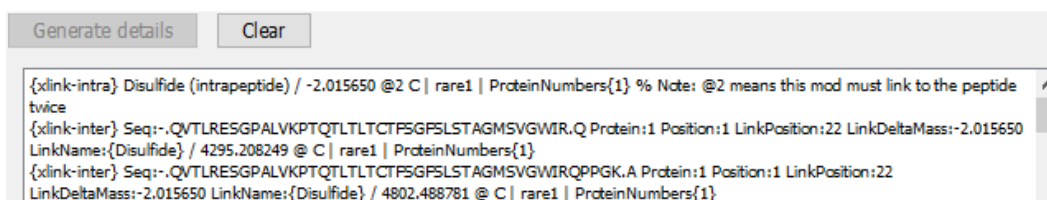
Figure 45: For FASTA proteins entry for designating numerically which protein sequences from the FASTA database to be considered

Commas between protein numbers allow crosslinks between those proteins. Semicolons between protein numbers indicate internal linking for the two proteins but no crosslinking between them.

The **Trisulfide** checkbox includes searches for trisulfides within a single peptide or linking between two peptides.

The **Crosslink: DSS** and **Crosslink: Custom** checkboxes include searches for crosslinks within a single peptide or linking between two peptides.

The **Generate details** button displays all the linked peptides with their mass deltas to be included in the Byonic search, based on the previous settings. To fine-tune the search, this text may be edited.



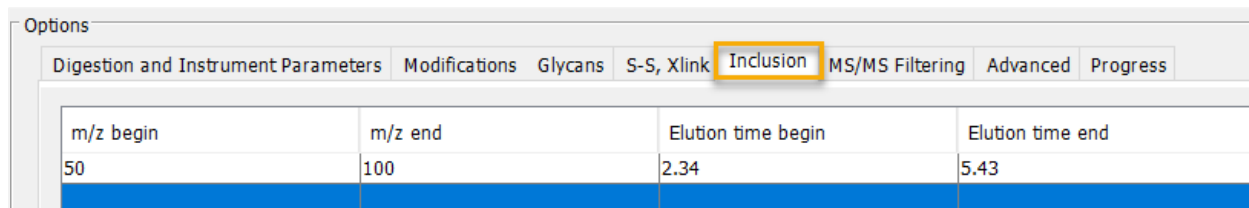
Generate details Clear

```
{xlink-intra} Disulfide (intrapeptide) / -2.015650 @2 C | rare1 | ProteinNumbers{1} % Note: @2 means this mod must link to the peptide twice
{xlink-inter} Seq:-QVTLRSGPALVKPTQTTLTCTFSGFSLTAGMSVGWIR.Q Protein:1 Position:1 LinkPosition:22 LinkDeltaMass:-2.015650
LinkName:{Disulfide} / 4295.208249 @ C | rare1 | ProteinNumbers{1}
{xlink-inter} Seq:-QVTLRSGPALVKPTQTTLTCTFSGFSLTAGMSVGWIRQPPGK.A Protein:1 Position:1 LinkPosition:22
LinkDeltaMass:-2.015650 LinkName:{Disulfide} / 4802.488781 @ C | rare1 | ProteinNumbers{1}
```

Figure 46: Generate details for S-S, Xlink

## Inclusion tab

The **Inclusion** tab allows the user to define a m/z ratio range and/or elution time range segments to include in the search. Enter the begin and end values for one or both range types. A new row is automatically added for additional range definitions.



m/z begin	m/z end	Elution time begin	Elution time end
50	100	2.34	5.43

Figure 47: Inclusion tab

## MS/MS Filtering tab

The **MS/MS Filtering** tab introduces MS/MS diagnostic peak filtering:

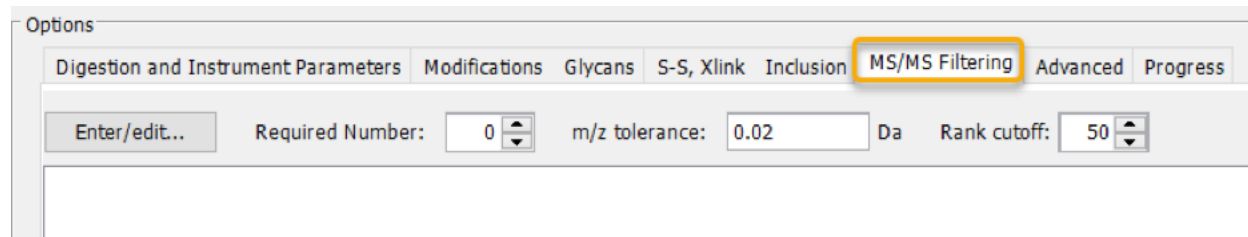
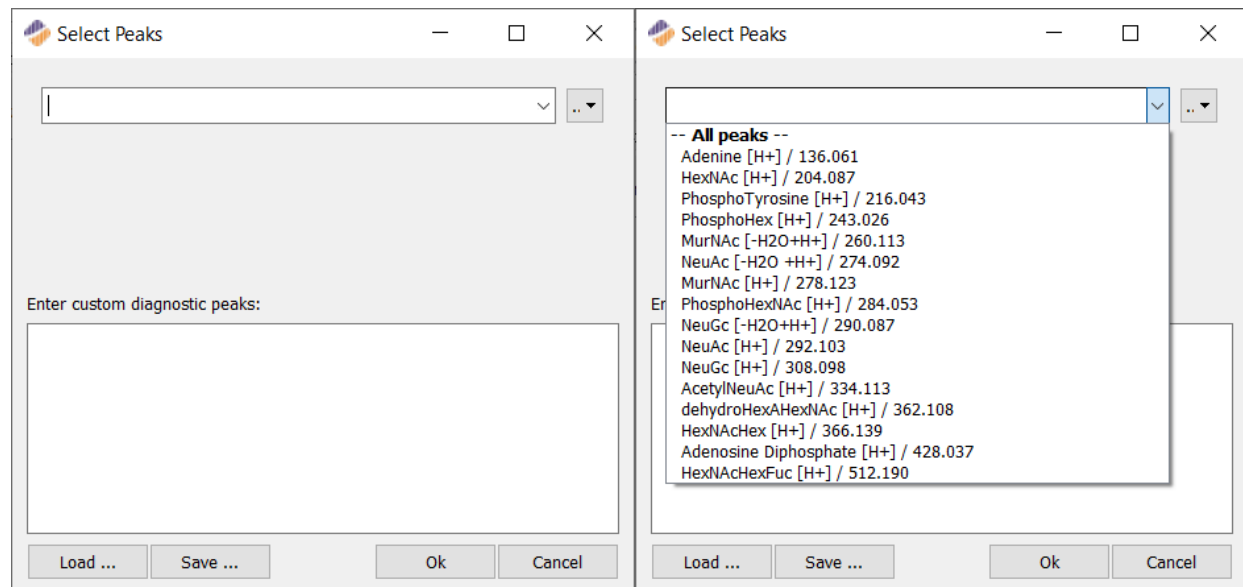


Figure 48: MS/MS Filtering tab

To create an MS/MS diagnostic peak filter, click the **Enter/edit** button. The **Select Peaks** dialog opens:



Enter custom diagnostic peaks:


Load ... Save ... Ok Cancel

Load ... Save ... Ok Cancel

-- All peaks --

- Adenine [H+] / 136.061
- HexNAc [H+] / 204.087
- PhosphoTyrosine [H+] / 216.043
- PhosphoHex [H+] / 243.026
- MurNAc [-H2O+H+] / 260.113
- NeuAc [-H2O+H+] / 274.092
- MurNAc [H+] / 278.123
- PhosphoHexNAc [H+] / 284.053
- NeuGc [-H2O+H+] / 290.087
- NeuAc [H+] / 292.103
- NeuGc [H+] / 308.098
- AcetylNeuAc [H+] / 334.113
- dehydroHexAHexNAc [H+] / 362.108
- HexNAcHex [H+] / 366.139
- Adenosine Diphosphate [H+] / 428.037
- HexNAcHexFuc [H+] / 512.190

Figure 49: Select Peaks dialog with diagnostic peak drop-down

Click the down arrow to expose the available diagnostic peak. When a peak is selected a new row appears for diagnostic peaks. The  button in the last column exposes additional actions that can be performed on diagnostic peaks:

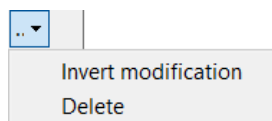


Figure 50: MS/MS peak inversion option

Custom modifications can be manually entered in this format in the box labeled **Enter custom diagnostic peaks**. The following is an example format of a diagnostic peak:

Kdn / 251.076

The selected and custom diagnostic peaks can be exported as a text file by clicking **Save**. This text file can then be loaded into another Byonic session by clicking **Load**. When the list of diagnostic peaks is complete, click **Ok**. The diagnostic peaks strings are now listed in the MS/MS Filtering tab:

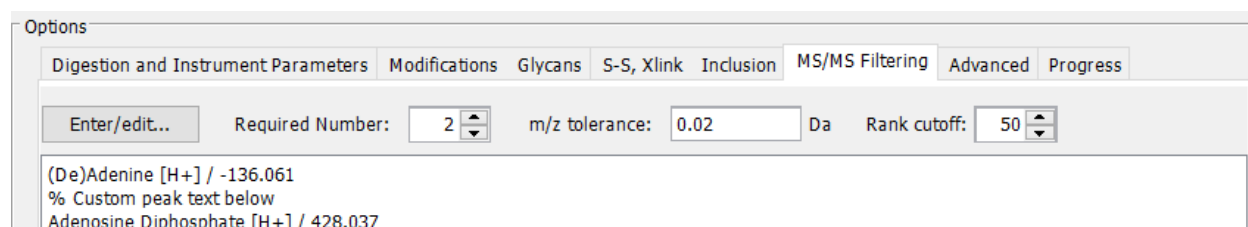


Figure 51: MS/MS Filtering tab with added diagnostic peaks

To set a required count of diagnostic peaks from within the list, enter a value in **Required Number**. To set no minimum requirement, leave the value at zero. A maximum **m/z tolerance** can also be set. **Rank cutoff** sets the maximum count of peaks, filtered by intensity (this was previously set using Advanced commands).

## Advanced tab

The **Advanced** tab contains a variety of spectrum, peptide and protein options. This tab helps Byonic respond to imperfect inputs. For example, on many MS instruments precursor ion charges are uncertain for some or all spectra

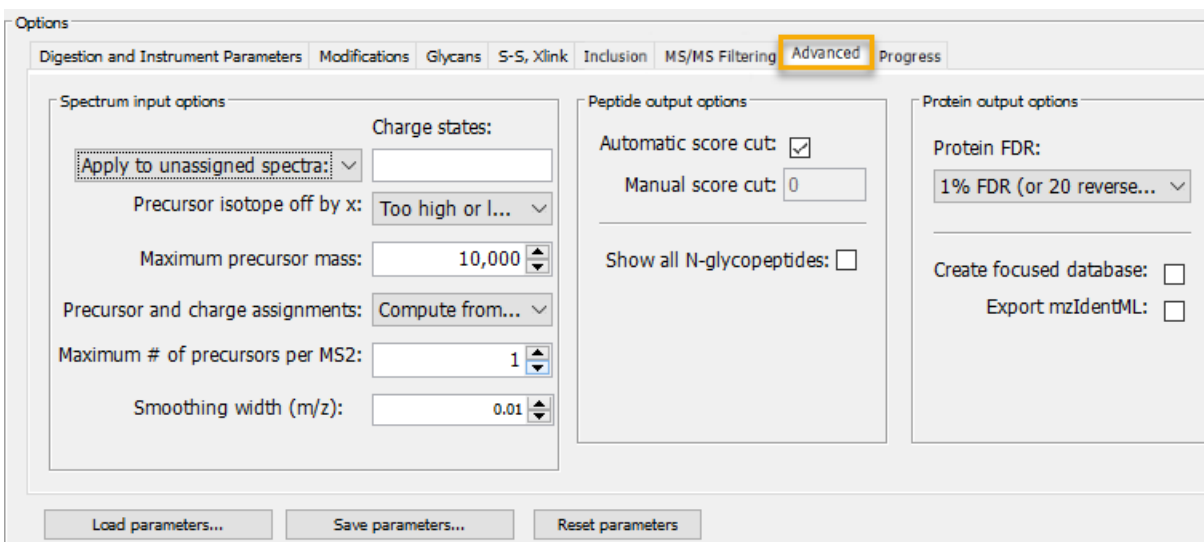


Figure 52: Advanced tab.

## Spectrum input options

The **Spectrum input options** section includes precursor and other spectrum settings that are to be applied to the Byonic query.

- **Apply to all spectra/unassigned spectra** allows the user to assign custom charge states to either all spectra or only to unassigned spectra.

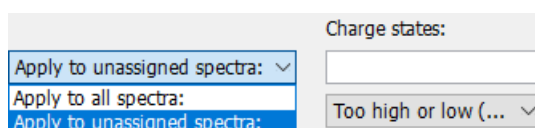


Figure 53: Apply charges to spectra

- **Charge states** sets custom charge states to the applied spectra. By default, the charges +1, +2, +3 are assigned for all CID spectra and the charges +2, +3, +4 are assigned for all ETD spectra. To assign different charges to the spectra, list the charge values, separated by commas.
- **Precursor isotope off by x** allows the user to set an error check for precursor isotope offsets. 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.

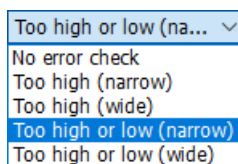


Figure 54: Precursor Isotope off by x

- The tooltip for this setting explains each setting:

No error check: Assume correct monoisotope.  
 Too high (narrow): (0-1000Da) -> no error check, (1000-2500Da) -> +1 error check, (2500+Da) -> +2 error check.  
 Too high (wide): Error check =  $\text{+floor}((\text{mass in Da}) / 1000)$ .  
 Too high or low (narrow): Error check =  $\text{+/-floor}((\text{mass in Da}) / 4000)$ .  
 Too high or too low (wide): Error check =  $\text{+/-floor}((\text{mass in Da}) / 2000)$ .

Figure 55: Tooltip for Precursor Isotope off by x

- **No error check** will use only the assigned precursor; Too high (narrow) allows no errors for precursor up to 1000 Da. For precursors from 1000 to 2500 Da, it allows the assigned precursor to be 1 Da too large in addition to the precursor mass tolerance set on the Digestion and Instrument Parameters tab. For precursors above 2500 Da, the assigned precursor may be 2 Da too large.
- **Maximum precursor mass** allows the user to set a maximum allowed precursor mass
- **Precursor and charge assignments** allows the user to choose between using the pre-assigned precursor and charge assignment values or calculating them from the MS1 data.

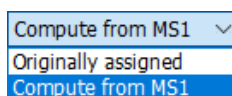


Figure 56: Precursor and Charge Assignments

- **Maximum # of precursors per MS2** sets this limit. It is recommended to set this limit to 2 for complex samples and 5-10 when processing MSe or DIA data.
- **Smoothing width (m/z)** applies a sigma value in Thomsons for Gaussian smoothing and centroiding of Waters or Sciex data. The default value (0.01 m/z) for half-width at peak half maximum works well in most cases.

## Peptide output options

The **Peptide output options** section includes options for filtering the peptide-spectrum matches (PSMs) by score.

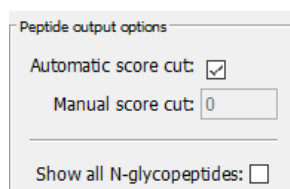


Figure 57: Peptide output options

- **Automatic score cut** uses Byonic to compute this value. By default, PSM filtering is deferred until after protein ranking. It is then filtered 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. (For more information, see: “Two-dimensional target decoy strategy for shotgun proteomics”, *Journal of Proteome Research* 10 (12), 5296-5301, 2011.)
- **Manual score cut** overrides Byonic’s algorithm. When Automatic score cut is unchecked, PSMs are filtered before protein ranking, using the minimum score entered into the Manual score cut box. 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.
- **Show all N-glycopeptides** shows the 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

The **Protein output options** section includes protein settings applied after the Byonic query is run.

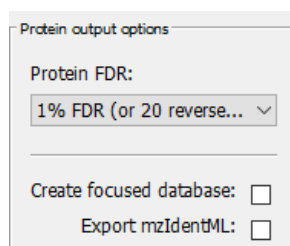


Figure 58: Protein output options

- **Protein FDR** sets the protein list cut-off. By default, the protein list is cut off at 1% protein FDR or 20 decoy proteins, whichever comes last. The user can set the cutoff at 2% protein FDR or 50 decoy proteins, whichever comes last. Finally, the user can select **No Cuts** to produce a completely unfiltered (but still ranked) protein list.

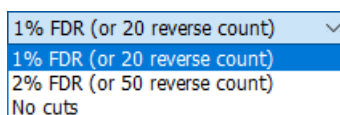


Figure 59: Protein FDR options

- **Create focused database** outputs a new FASTA file 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 additional modifications and/or a wildcard. Of course, the user can also create focused databases directly by editing existing FASTA files. The FASTA file is named as focused and saved to the designated output folder.
- **Export mzIdentML** creates mzIdentML and .mgf files that can then be input into third-party tools such as Skyline and Scaffold.

## Progress tab

The **Progress** tab sets output options. When the program is launched, this tab will show a progress bar of the current search.

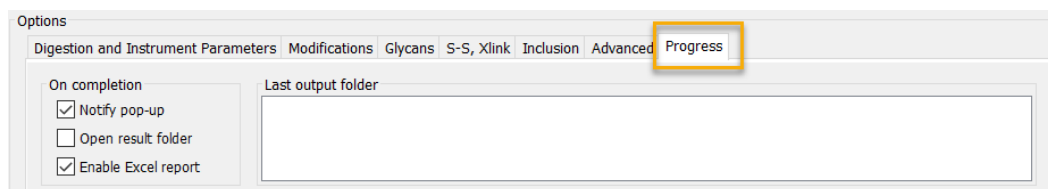


Figure 60: Progress tab

- **Notify pop-up** opens an .html page when the search is done.
- **Open result folder** automatically opens the result folder when the search is done.
- **Enable Excel report** adds an Excel report of the search results that can be shared with collaborators.

## Additional Settings

### Load, Save, and Reset Parameters

Because Byonic has many options and capabilities, writing modification rules and setting parameters can be a nontrivial task. The **Save parameters** button allows the user to save all of the inputs described above. Files are saved with the extension \*.byparms. **Load parameters** loads a previously saved set of inputs, which can then be edited as desired. **Reset parameters** clears edited inputs and restores the defaults.

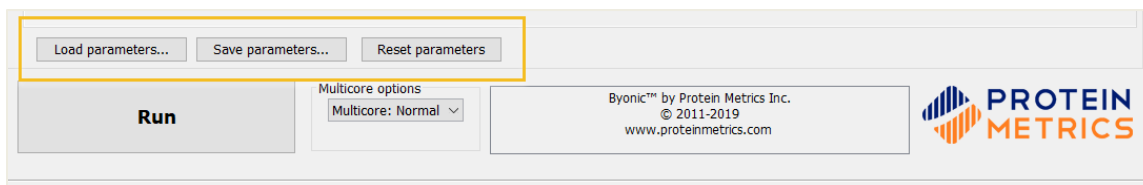


Figure 61: Load, Save, and Reset parameters

## Multicore options

The user can control the number of computer cores of the CPU used in the Byonic search. The Light search uses one core, Normal search uses all available cores minus two, and Heavy search uses all available cores. The Light search uses the least CPU resources, while the Heavy search runs the fastest.

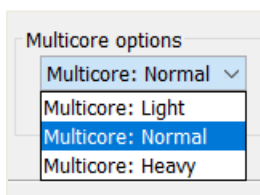


Figure 62: Multicore options

## Run

Click the **Run** button to start the analysis.

## Byonic Menu Bar

The main menu bar includes three items: **File**, **Edit**, and **Help**.

### File Menu

The **File** menu contains three items: **Load params**, **Save params**, and **Exit**.

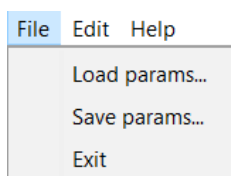


Figure 63: File menu

**Save params** saves a copy of all the parameters entered in the various tabs of Byonic query to a file with extension \*.byparms. **Load params** loads query parameters saved from a previous session. The loaded parameters will replace existing parameters. These functions are the same as the **Load parameters** and **Save parameters** buttons below the tabs. **Exit** closes the application.

### Edit Menu

The **Edit** menu formerly contained options that have since been moved to the search tabs. See the **Glycan Tab** section for more information about these Edit options.



## Help Menu

The **Help** menu provides information about the software:

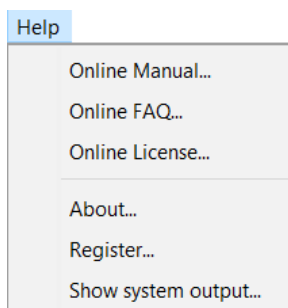


Figure 64: Help menu

The first three menu items open web pages found on [www.proteinmetrics.com](http://www.proteinmetrics.com). **Online Manual** opens a web version of this document, **Online FAQ** opens up summary questions and answers, and **Online License** opens the End User License Agreement for Protein Metrics software.

The **About** menu displays the software version number and system information, which is useful when reporting issues. **Register** is used to activate the software upon first use, along with other license actions. See the **PMI End-User License Manual** for more information. **Show system output** displays logs for the current Byonic search; these are helpful for troubleshooting problems together with Protein Metrics, Inc. staff via [support@proteinmetrics.com](mailto:support@proteinmetrics.com).

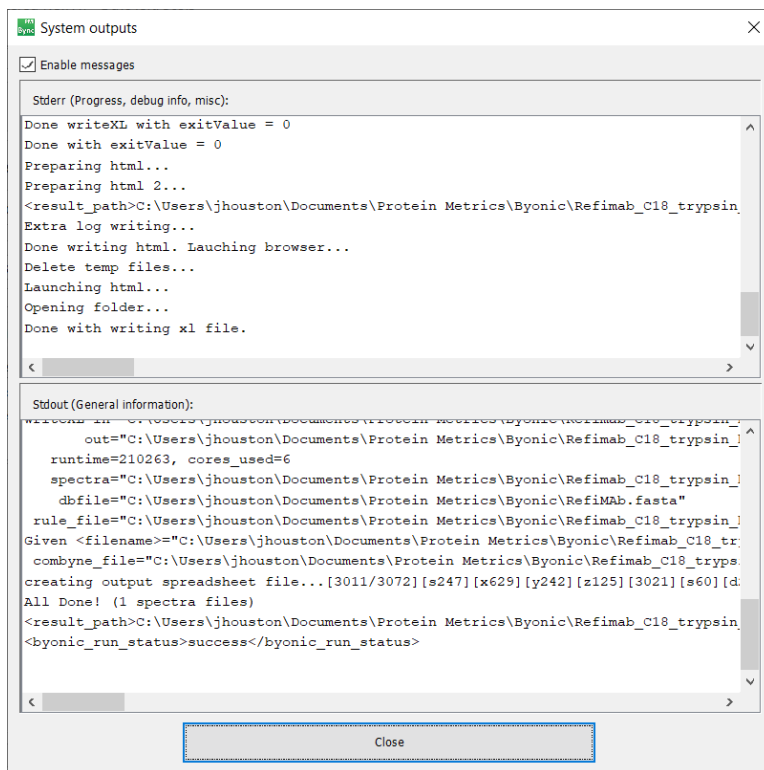


Figure 65: System outputs search logs

## Byonic Results

Byonic generates a series of files to the directory defined under **Output folder** in the Search screen. The folder and files are named using the conventions set in the **Results folder name** entry in the Search screen (by default, the MS file name, the date and the term “Byonic”). The Byonic results file (extension \*.byrslt) can be viewed and explored interactively using Byonic™ Viewer, a separate program from the Byonic search engine. Byonic also writes the output data to an Excel spreadsheet (.xlsx) for viewing, sorting, importing into other programs, and sharing with collaborators. In both output formats, Byonic organizes its findings into two lists, one for proteins and one for peptide-spectrum matches (PSMs).

Byonic also generates a reformatted spectrum file (extension \*.byspec2) that can be loaded into other Protein Metrics applications, such as Peptide Analysis (Byologic®), Intact Analysis (Intact Mass™) or Chromatogram Analysis (Byomap™). Byonic also creates an output of all the parameters used (extension \*.byparms) which can be used to reproduce the search later. Finally, the search run creates a subfolder named “objs”, which will contain log files and miscellaneous other output files.

## Byos and Byonic Viewer Results

Byonic search result file can be opened in either the Byos® or Byonic™ Viewer applications. Starting with version 3.7.5, a \*.byrslt file opens in Byos by default. Therefore, newer versions of Byonic require a co-installation of Byos or Byos Viewer to function correctly. An example \*.byrslt file opened in Byos is shown in the figure below:

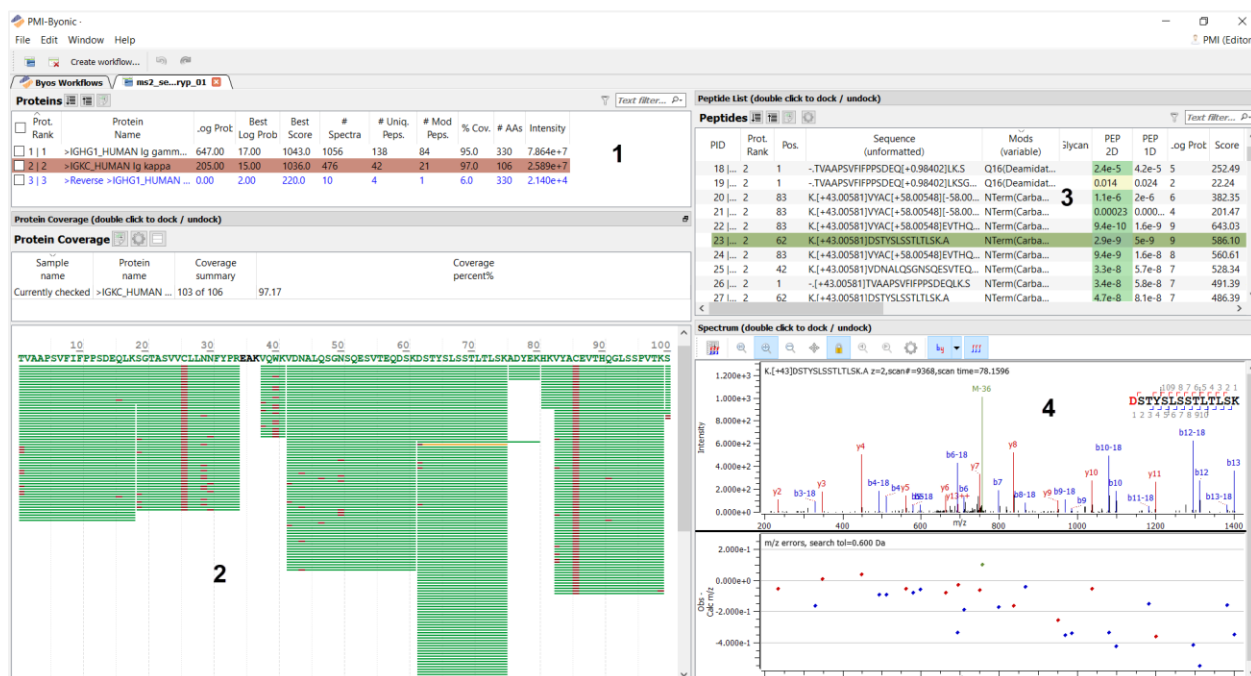


Figure 66: The four Byonic views as displayed in Byos: 1) Proteins table; 2) Protein Coverage map for the selected protein, 3) Peptide table (PSMs) for the selected protein(s), and 4) Annotated Spectrum for the selected PSM.

An example \*.byrslt file opened in Byonic Viewer is shown in the figure below:

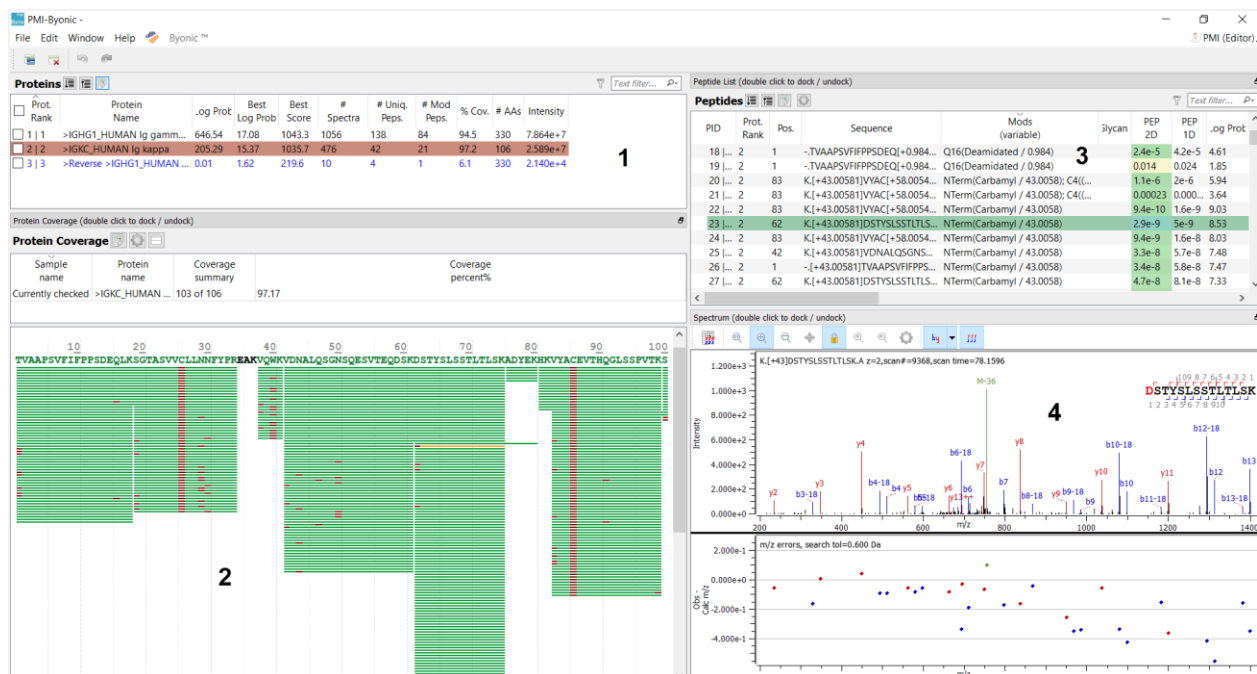



Figure 67: The four Byonic views as displayed in Byonic Viewer: 1) Proteins table; 2) Protein Coverage map for the selected protein, 3) Peptide table (PSMs) for the selected protein(s), and 4) Annotated Spectrum for the selected PSM.

The four views are interconnected: when a protein is selected in the Proteins table (view 1), the Protein Coverage map (view 2) and the Peptides table PSMs (view 3) are displayed for that protein. When a PSM is selected in the Peptides table, the annotated Spectrum (view 4) is displayed for that peptide. To view all the proteins at once, check the box at the top left of the Proteins table. All of the PSMs will display in the Peptides table. The four views can be undocked and rearranged for a customized screen layout; this is especially useful for double-headed displays. To undock (detach) a pane from the others, click the  icon at the top right of the pane:

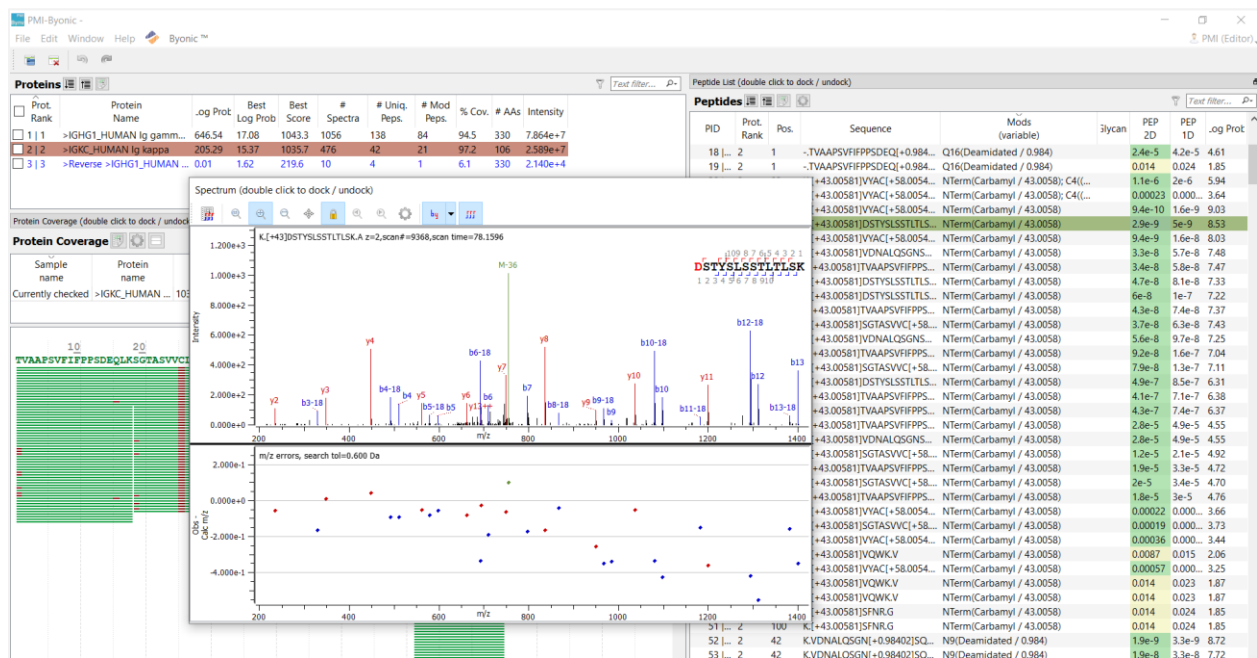


Figure 68: Byonic Viewer with the Spectrum view (MS2 plot and fragment errors) undocked

To return the panes to their original locations, choose **Windows > Restore Default Layout**.

## Byonic Viewer Menu Bar

The main menu bar includes three items: **File**, **Edit**, **Window**, and **Help**.

### File Menu

The **File** menu manages Byonic project files:

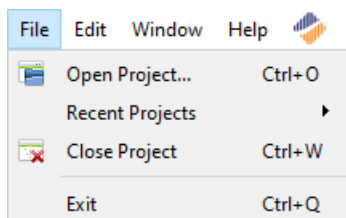






Figure 69: File menu

The File menus includes items to open previously created projects from saved files, close a project (yet leave the application open), and exit Byonic. Mouse over the **Recent Projects** menu to open a side window with a list of recently visited projects that can be reopened. The  button and  button below the File and Edit menus can also be used to open and close projects, respectively.

### Edit Menu

The **Edit** menu contains the **Undo** and **Redo** features. The  button and  button below the Window and Help menus also perform undo and redo, respectively.

### Window Menu

The **Window** menu manages layouts:

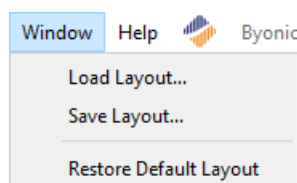




Figure 70: Window Menu

**Load Layout** opens a saved layout for how the different views are arranged and customized. These are stored to files with extension \*.ini. **Save Layout** saves the current layout to an .ini file for use in other sessions. **Restore Default Layout** applies the layout used when the application is first opened.

The current layout is customized by hovering over the vertical or horizontal edges between table and plot views so that the cursor changes to arrows:  or . Left click and drag the edges up or down or right or left to change the sizes of the views. Columns in the table views can be made larger or smaller in same manner. To reorder columns in a table, left click the name of the column and drag it to the new location. To do a simple sort, left click the header once to sort ascending and twice to sort descending. Mouse over a column header to see a description of that field.

To hide or display columns, right-click any column header. The **Column Header Editor** opens:

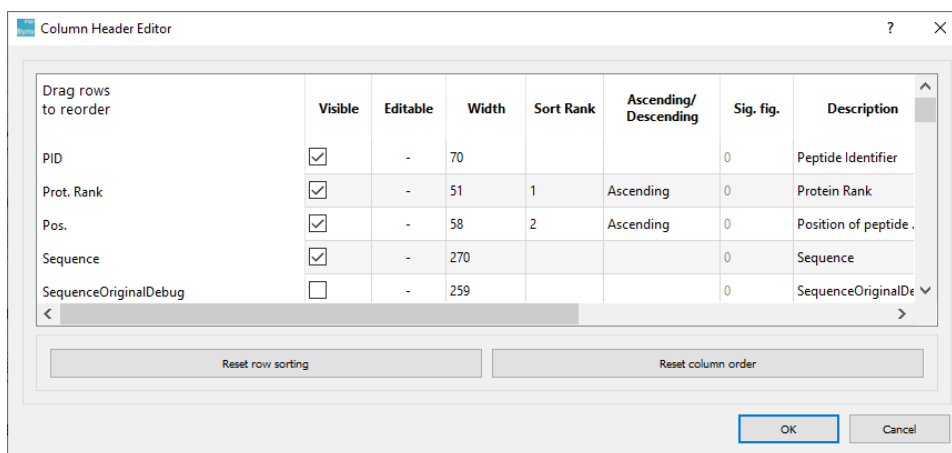


Figure 71: Column Header Editor

Check the column fields to display and uncheck the fields to hide. The dialog contains brief descriptions of each field. Alternatively, hover over the column names to display the descriptions as tooltips. Note that the Byonic Viewer layout persists on exit.

## Help Menu

The **Help** menu provides information about the software:

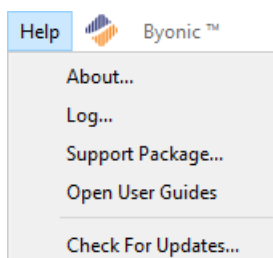






Figure 72: Help Menu

The **About** menu shows the software version number, which is needed when reporting issues. The **Log** menu opens a log containing recent activity; this is helpful for troubleshooting problems together with Protein Metrics staff via [support@proteinmetrics.com](mailto:support@proteinmetrics.com). **Support Package** collects all relevant information helpful for troubleshooting problems together with Protein Metrics staff via [support@proteinmetrics.com](mailto:support@proteinmetrics.com). **Open User Guides** opens the Document directory where this and other useful guides can be found, including PMI Reporting Manual.pdf. **Check For Updates** checks whether software updates are available on [www.proteinmetrics.com](http://www.proteinmetrics.com). Note that the **Register** function is available in the Help menu from the Byonic search application.

## Table Menus

The menu bars at the top of the three table views, Project, Peaks, and Candidates, share icons that manage hierarchical lists and filter on data.

- The  icon expands rows to show “sub-rows”.
- The  icon collapses rows to hide sub-rows. The use of sub-row depends on the table. For example, in the Peptides table, a row is a peptide record and a sub-row is the data resulting from an individual sample that contained the peptide.

- The  icon exports the table to a .csv file for opening with Excel.
- The  icon opens a dialog to create custom filters specific to columns in that data table:

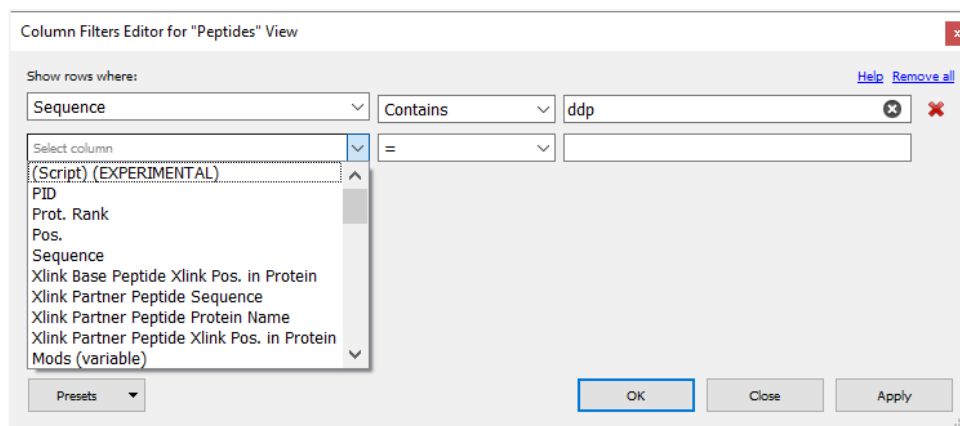




Figure 73: Column Filters Editor

Select a column in the first cell, select an operator in the second cell, and enter text in the third cell. A second filter row becomes available for further entry. Click the red X after a filter row to delete it. Custom filters allow masses to be filtered by mass range, annotation, intensity, and so forth. Custom filters can be stored with the project document or exported and imported using the **Presets** dropdown.

- The  cell filters the entered text across the content of all columns in the protein and peptide tables. The records are filtered automatically as text is entered. Click the  icon to search the string as a whole word or as case sensitive.
- To sort the contents of a column by ascending value, click the column header. Click the column header again to sort by descending value.

## Table Right-Click Menus

The Project and Peptides tables have context menu items revealed by a right-click on the rows within the tables. (Recall that a right-click on the header of these tables opens the Column Header Editor dialog.)

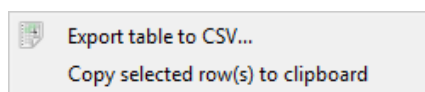


Figure 74: Table right-click menus

- **Export table to CSV** saves the table content and headers to a .csv file using parameters chosen in the **Export data** dialog.
- **Copy selected row(s) to clipboard** copies selected table rows, with their headers to be available for pasting into another application. (This menu is not available in the Protein Coverage table.)
- Table context menus contain a set of query options for searching the clicked text value using any of the available search operators. This is an easy way to filter a table based on a specific field value:

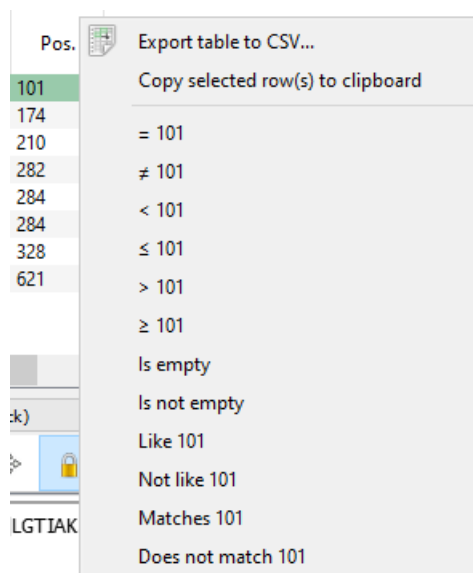
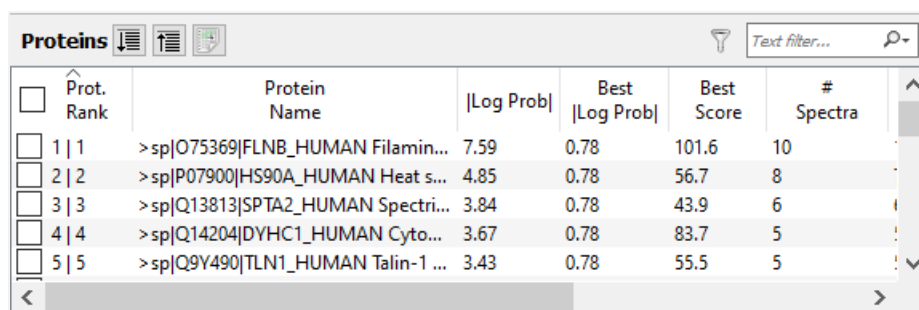


Figure 75: Search table values through a right-click menu

The context menu shows all the possible queries for that data in that column, in this case, for value 101 in the Pos. column of the Peptides table.

## Proteins Table

The **Proteins** table contains information about the identified proteins, including name, rank and scores:



<input type="checkbox"/>	Prot. Rank	Protein Name	Log Prob	Best  Log Prob	Best Score	# Spectra
<input type="checkbox"/>	1   1	>sp O75369 FLNB_HUMAN Filamin...	7.59	0.78	101.6	10
<input type="checkbox"/>	2   2	>sp P07900 HS90A_HUMAN Heat s...	4.85	0.78	56.7	8
<input type="checkbox"/>	3   3	>sp Q13813 SPTA2_HUMAN Spectri...	3.84	0.78	43.9	6
<input type="checkbox"/>	4   4	>sp Q14204 DYHC1_HUMAN Cyto...	3.67	0.78	83.7	5
<input type="checkbox"/>	5   5	>sp Q9Y490 TLN1_HUMAN Talin-1 ...	3.43	0.78	55.5	5

Figure 76: Proteins table

To display or remove the associated views of a protein in the Protein Coverage and Peptides tables, check or uncheck the box at left of the protein. To view all of the proteins together, check the box at the top left.

## Protein Coverage Table and Menu

The **Protein Coverage** table displays peptide sequences and corresponding protein coverage. The accompanying display maps the protein sequences of detected peptides to colored lines below each sequence:





Protein Coverage   			
Sample name	Protein name	Coverage summary	Coverage percent%
Currently checked	>sp P07900 HS90A_...	69 of 732	9.43
>sp P07900 HS90A_HUMAN Heat shock protein HSP 90-alpha ( ^			
10 20 30 40 50 MPEETQTQDQPMEEVEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELIS			
60 70 80 90 100 NSSDALDKIRYESLTDPKLDGKELHINLIPNKQDRTLTIVDTGIGMTK			
110 120 130 140 150 ADLINNLTIAKSGTKAFMEALQAGADISMIGQFGVGFYSAYLVAEKVTV			

Figure 77: Protein Coverage table

The Protein Coverage table also includes the following specialized menu icon button:

- The  icon opens the Protein coverage rendering options dialog:

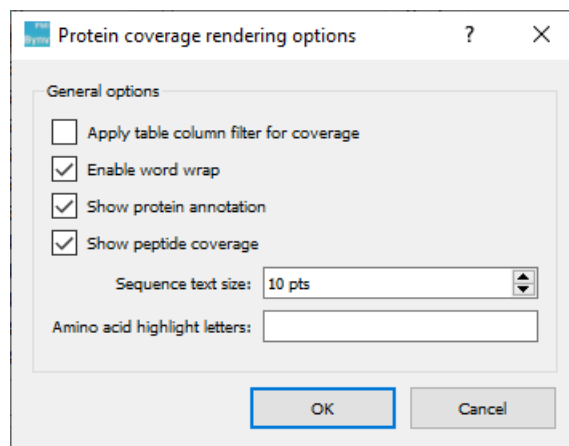


Figure 78: Protein coverage rendering options

- The  icon turns on and off the display of the tabular protein coverage data.

The Protein Coverage view has context menu items revealed by a right-click on the graphical portion of the display:

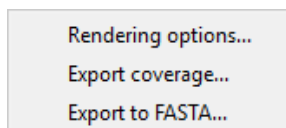


Figure 79: Protein Coverage right-click menu

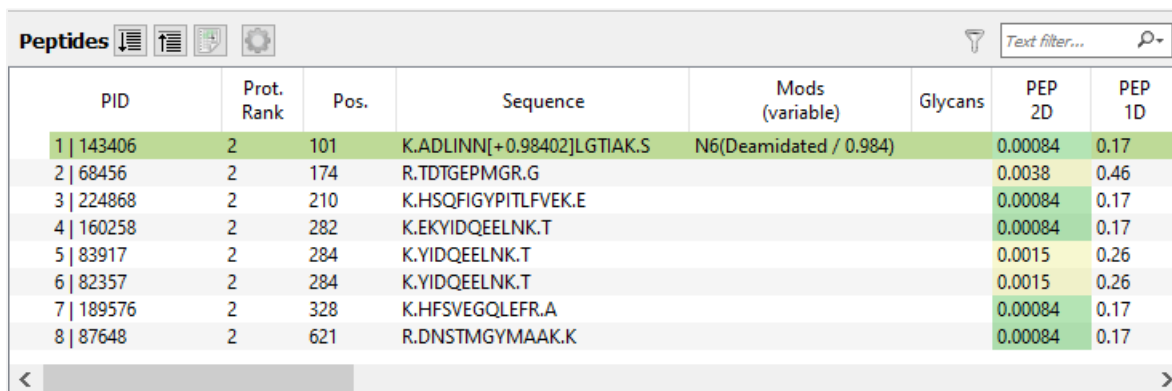
- Rendering options** also opens the Protein coverage rendering options dialog show above.
- Export coverage** creates a \*.png file of the graphical protein coverage display.



- **Export to FASTA** creates a \*.fasta file of the proteins found in the Protein Coverage view. The \*.fasta file is then available for import into new projects.

## Peptides Table and Menu


The **Peptides** table contains detailed information about all identified peptides for the protein(s) checked or selected in the Proteins table:



PID	Prot. Rank	Pos.	Sequence	Mods (variable)	Glycans	PEP 2D	PEP 1D
1   143406	2	101	K.ADLINN[+0.98402]LGTIAK.S	N6(Deamidated / 0.984)		0.00084	0.17
2   68456	2	174	R.TDTGEPMGR.G			0.0038	0.46
3   224868	2	210	K.HSQFIGYPITLFVEK.E			0.00084	0.17
4   160258	2	282	K.EKYDQEELNK.T			0.00084	0.17
5   83917	2	284	K.YIDQEELNK.T			0.0015	0.26
6   82357	2	284	K.YIDQEELNK.T			0.0015	0.26
7   189576	2	328	K.HFSVEGQLEFR.A			0.00084	0.17
8   87648	2	621	R.DNSTMGYMAAK.K			0.00084	0.17

Figure 80: Peptides (PSM) table

The **Peptide** table includes a large number of possible column fields. The most important ones are the peptide sequence, Byonic score and log probability, and scan number (often also found in the Comment field). In the case of an identification that includes a glycan from the predefined glycan tables, the Glycan column displays the glycan's monosaccharide composition as a sequence of 6 numbers: the counts of HexNAc, Hex, Fuc, NeuAc, NeuGc, and sodium. The order of monosaccharides is displayed in the tooltip for the Glycan column header.

In addition to the usual table menu buttons, the Peptides table contains the  icon button which opens the Peptide view options dialog. This allows the table to be grouped by unique peptides:

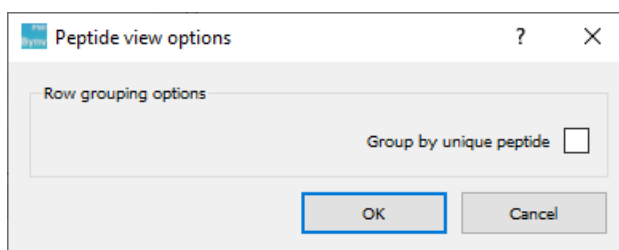


Figure 81: Peptide view options

## Spectrum View and Menu

The annotated **Spectrum** view allows the user to inspect identified peaks and associated fragment errors to manually validate peak identifications and modification placements.

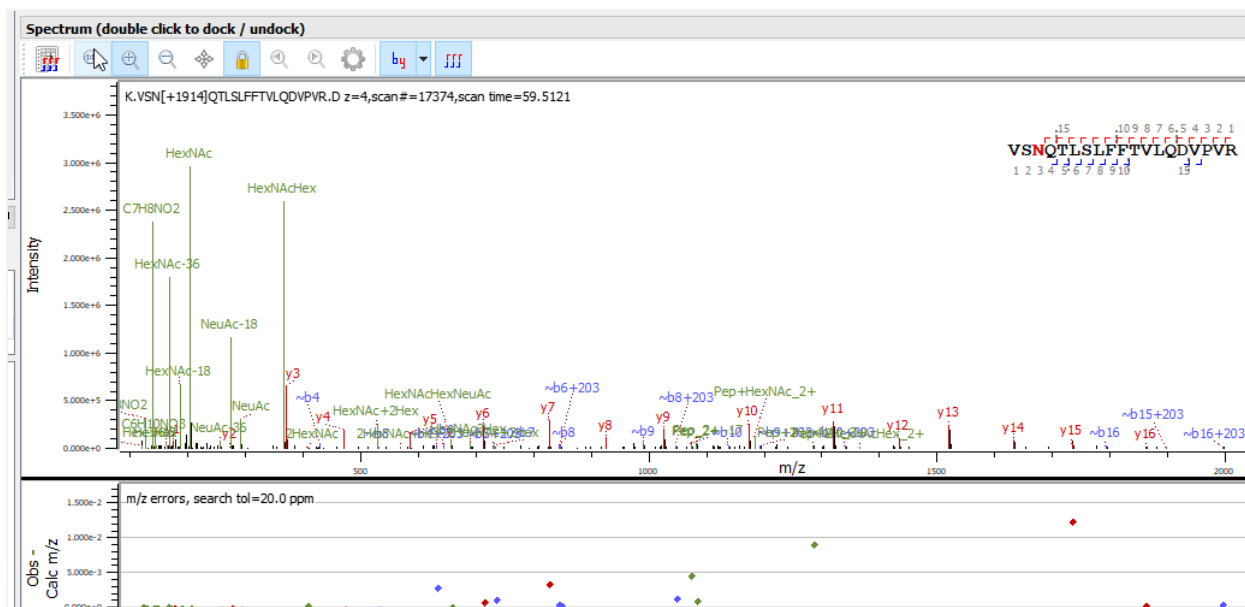


Figure 82: Annotated Spectrum view

Byonic annotates key elements of the spectrum. It annotates the most commonly observed ion series (a, b, y, c, z, b++, y++, etc.) for the type of fragmentation. It annotates common neutral losses from the precursor (e.g., M-98 for loss of phosphoric acid). It annotates immonium ions for certain amino acids and oxonium ions for glycans. Finally, Byonic annotates key diagnostic peaks such as imm\_pY at 216 Da for phosphotyrosine, 243 for phopho-Hex, and 284 for phospho-HexNAc. For Byonic, M alone refers to a precursor in its original charge state, so M+e in an ETD spectrum represents a charge-reduced precursor, and M\_2+ - HexNAcHex in a spectrum of a plus-three precursor represents the loss of HexNAc, Hexose, and a proton. In a glycopeptide spectrum, Pep refers to the neutral (possibly modified) peptide without glycans. For example, Pep+HexNAc\_2+ represents the peptide with HexNAc and two protons, that is, the Y1++ ion.

Most of Byonic's annotations of product ions are standard, but a few of the conventions may require additional explanation. If Byonic does not indicate charge, the charge is plus one. If Byonic does not indicate isotopes, the peak is the monoisotopic peak. For high-mass fragments, Byonic will append annotations to identify higher-isotope peaks, for example, \_iso1, \_iso2. Common neutral losses are shown by removed nominal masses, for example, -17 = ammonia, -18 = water, -36 = two water molecules, -54 = three water molecules (a common loss from glycosylated lysine), -80 = loss of SO<sub>3</sub> from sulfation, and -98 = loss of H<sub>3</sub>PO<sub>4</sub> from phosphorylation. An apostrophe before a z-ion, for example, 'z6 shows loss of (alkylated) side-chain on an N-terminal cysteine. A tilde shows loss of labile modifications, for example, ~y7 means y7 with loss of all labile modifications. Labile modifications are defined to be O- and N-linked glycosylation, sulfation, histidine phosphorylation, and gamma-carboxylation. (In order to accommodate all the various types and modifications of O-linked glycans, Byonic treats any sufficiently heavy modification on S/T as labile.) For GalNAc- or GlcNAc-initiated glycosylation, Byonic will annotate peaks with a single base HexNAc, for example, ~y7+203. Byonic annotates oxonium ions from glycans, either with names like "NeuAc-18" or atomic formulas when names are not definitive. The Byonic Viewer also annotates peaks at 284 and 446 for sulfated HexNAc and sulfated HexNAc-Hex. These are small peaks and they do not always show clearly.

Byonic supports scoring and annotation for user-specified custom peaks. The theoretical peaks for the peptide fragments will contain these custom peaks. If the spectrum matches the custom peaks, they will be included in the Peptide Score.


The custom peaks are constrained to belong to a particular modification rule in the **Modifications** tab in Byonic. An example of how to specify the custom peaks:

```
[Woo2/+346.14575 @S,T | common2 | CustomPeaks{IsoTaG:347.1531, IsoTaG-36: 311.1320, IsoTaG-90: 257.11017}
```



- **Woo2/+346.145** - name/composition of the modification
- **@ S, T** - target amino acid for modification
- **common2** - type of modification (common, rare etc)
- **CustomPeaks{peak\_label: peak\_mz}** - comma separated list of peak\_label: peak\_mz strings where you specify the strings of peak\_label (ex IsoTaG-36) and peak\_mz (ex 311.132)

## Spectrum Menus

The menu bar at the top includes icons for operations on plots. These icons appear in almost all Protein Metrics, Inc. software products.

- The  icon switches between the default plot view and the mass table view:



Spectrum (double click to dock / undock)








  ☒ Calculated ☐ Observed ☐ Delta



PID=150619: VNIGQGSHPPQK

#	a calc.	b calc.	b-18 calc.	b++ -HexNAc calc.	b++ -HexN calc.	b-HexN calc.	Seq.	y calc.	y++ calc.	#
1	72.0808	100.0757	82.0651				V			11
2	186.1237	214.1186	196.1081				N	1065.5436	533.2754	10
3	299.2078	327.2027	309.1921				I	951.5006	476.2540	9
4	<b>356.2292</b>	384.2241	366.2136				G	838.4166	419.7119	8

Figure 83: Mass table view of the Spectrum view

When the **Mass table** view is shown, the observed values are shown in **RED** by default. The user can optionally expand to show observed and Delta masses in separate columns. The  button that allows an analyst to copy the table to the clipboard. The  button returns to the spectrum view.

- The  icon resets the plot to default zoom level. Shortcut = double left-click.
- The  icon enables zooming in. The cursor changes to this icon. Drag across the range of the plot to display to zoom to that x-range. By default, the plot's y-range scales according to the maximum y-value within the x-range, but the software also supports freeform zooming as enabled under the  icon.
- The  icon enables zooming out. Click anywhere in the plot to zoom out.
- The  icon enables moving (panning) across the plot. The cursor changes to this icon. Click the plot and drag up or down, right or left to view a part of the plot that is off-screen.
- The  icon locks the x-axis for stacked plots in a plot view to use the same ranges. Unselecting this allows stacked plots to apply independent x-axis ranges.
- The  icon performs an undo of the last zoom step. Shortcut = Shift-left arrow key.

- The  icon performs a redo of the last zoom step. Shortcut = Shift-right arrow key.
- The  icon manages how plots are displayed (render options), as well as zoom modes. The Render and zoom options include:

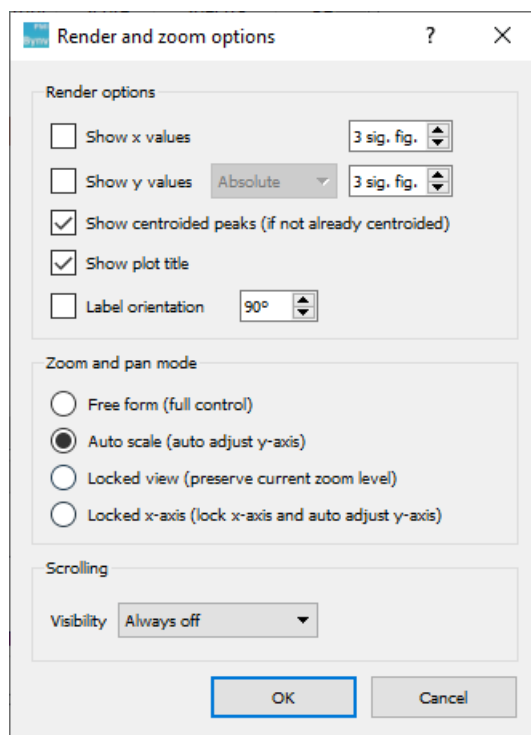


Figure 84: Render and zoom options


- **Show x values** and **Show y values** display the x- or y-coordinates beside plot peaks to the chosen number of significant figures. When x-coordinates are displayed, hovering the cursor over an x-coordinate label, will display x-coordinate (that is, m/z differences). This feature is invaluable for identifying unlabeled peak. For example, a mass delta of 18.011 indicates water loss (or gain).
- **Show centroided peaks** darkens the peak positions that are determined to be centroided.
- **Show plot title** displays the title of the plot, as generated from field values.
- **Label orientation**, when checked, orients labels to some angle other than horizontal.

The **Zoom and pan mode** options include:

- **Free form** mode to manually select the desired y range as well as x range
- **Auto scale** mode to select only the x range (the y range is then adjusted to the value of the highest peak)
- **Locked view** mode to keep the current x range (for either *m/z* and/or *m*) when moving between elution peaks. In Isotope plots
- **Locked x-axis** mode turns off autoscaling for the x-axis (but not the y-axis) and applies the current x-axis scale across all Peptide table selections

Under **Scrolling**, the **Visibility** setting controls display of the scroll bars:

- **Always on** displays scroll bars even when the full plot is shown

- **Always off** turns off display of the scroll bars even when a partial plot is shown
- **Show as needed** displays scroll bars only when a partial plot is shown
- The  icon switches between showing and hiding the b and y ion labels. The icon is accompanied by a drop-down arrow that reveals three sub-menus:
  - **Open annotation options** opens a dialog that allows the user to set options for determining the assignment of the MS2 annotations:

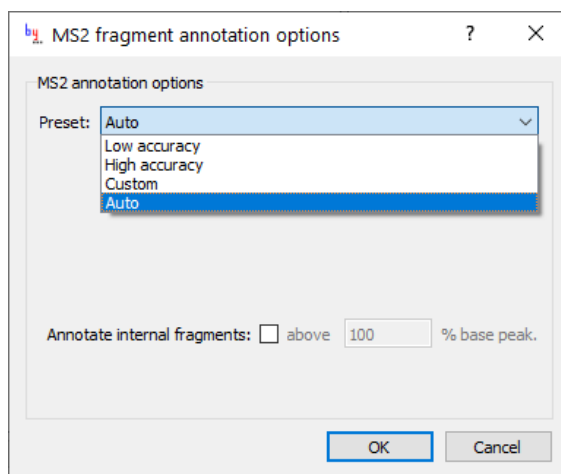



Figure 85: MS2 fragment annotation options

The MS2 mass-accuracy tolerance options can use presets for Low accuracy, High accuracy, or Custom, in which the user sets these values. The dialog defaults to the Auto preset, in which the software determines the MS2 fragment annotation settings.

To annotate internal fragments, check **Annotate internal fragments** and set the threshold using the field next to the checkbox. The internal fragments would be annotated as  $i(n-m)$ , where  $n$  and  $m$  are the positions of the first and last amino acids of the internal fragment. For example, for the peptide sequence **PEPTIDEK**, the internal fragment **TIDE** would be annotated as  $i(4-7)$ .

- **Remove fragment coverage cache** clears the existing cache of detected fragments used for the protein fragment coverage rendering.
- **Export fragment to CSV**, exports the fragment sequence to a \*.csv file.
- The  icon turns on or off the fragment sequences displayed in the top right of the plot. The cleavage diagram for the amino acid sequence is a standard summarization of the evidence for the identification

## Spectrum Right-Click Menus

The Spectrum plot has a variety of context menus for plot styling and exporting revealed by a right-click on the plot:

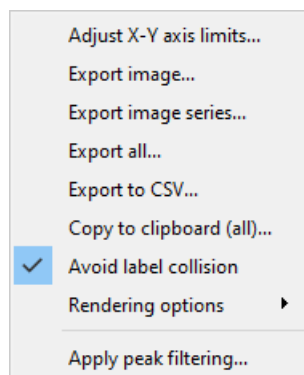




Figure 86: Spectrum context menus

Figures exported can be useful for reports, publications, regulatory filings, or internal communications. There are many options for rendering and exporting so that the user has much freedom to prepare a plot or figure style as needed.

- **Adjust X-Y axis limits** opens a dialog for the user to manually edit the plots x and y maxima and minima. This is a less convenient but more precise (and reproducible) alternative to the  and  icons.
- **Export image** allows the user to save the plot as a \*.pdf or \*.png file. It opens the **Plot Exporting Settings** dialog, which controls image size, file name and folder, and x and y minima and maxima:

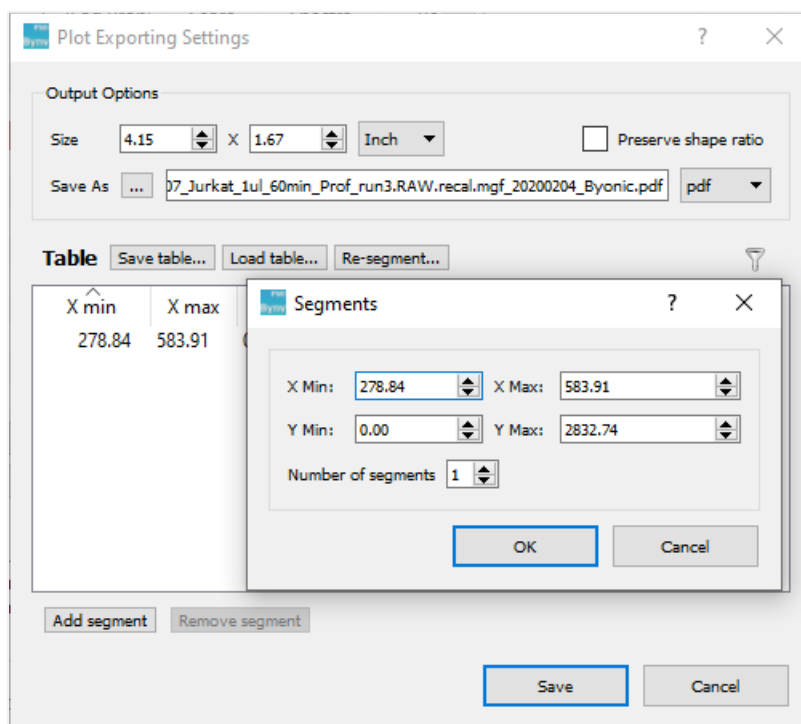


Figure 87: Plot export settings

Sometimes a user wishes to display a wide range across the x-axis and yet retain significant detail. This can be done by breaking up the plot into a series of panels. The **Add segment** button creates segments with user-defined x and y values. A series of these segments eliminates unneeded portions of the plot and increases the effective detail in the image. This is also a useful function for

automated reporting. The **Re-segment** button allows edits to the reported x and y minima and maxima for existing segments

- **Export image series** also opens the Plot Exporting Settings dialog, except that the segments are prepopulated with six equal sized segments. This simplifies the edits of the segments. The **Add segment** and **Remove segment** buttons control the number of segments that divide up the plot.
- **Export all** also opens the Plot Exporting Settings dialog, except that the wildtype peptide plots are exported along with the peptide plots. Segment editing functions are disabled for this option.
- **Export to CSV** exports the plot trace x-y points to a \*.csv file.
- **Copy to clipboard (all)** enables pasting the plot image into another application.
- **Avoid label collision** staggers the label text to reduce overlap.
- **Rendering options** controls how a plot is displayed in the application. It includes several sub-menus:

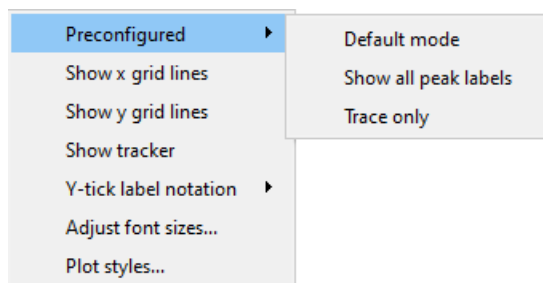


Figure 88: Rendering options sub-menu

- **Preconfigured** sub-menus control what is displayed in the plots. **Default mode** displays the trace, peak labels and plot title. **Show all peak labels** does exactly that. **Trace only** turns off the peak labels and plot title.
- **Show x grid lines** turns on and off the x grid lines.
- **Show y grid lines** turns on and off the y grid lines.
- **Show tracker** displays a vertical dotted line that follows the cursor when moved in the plot. This allows the user to line up annotated peaks with their associated m/z errors.
- **Y-tick label notation** toggles between y-axis notation options of absolute amounts or as percentages of the highest peak.
- **Adjust font sizes** controls the font size for each type of text on the plot.
- **Plot styles** allows changes to the graphic styling of the plot, including trace width, axis width, and total m/z dot colors to be used.
- **Apply peak filtering** sets the signal-to-noise ratio (set to zero by default).

When the cursor is positioned exactly over a spectrum peak, the reported m/z and intensity of that peak is shown inside curly brackets { }. When the cursor is not directly over a peak, these numbers give the m/z and intensity (x, y coordinates) of the cursor position.

## Excel Report

A Byonic search generates an Excel report summarizing much of tabular data found in the \*.byrs1t file:

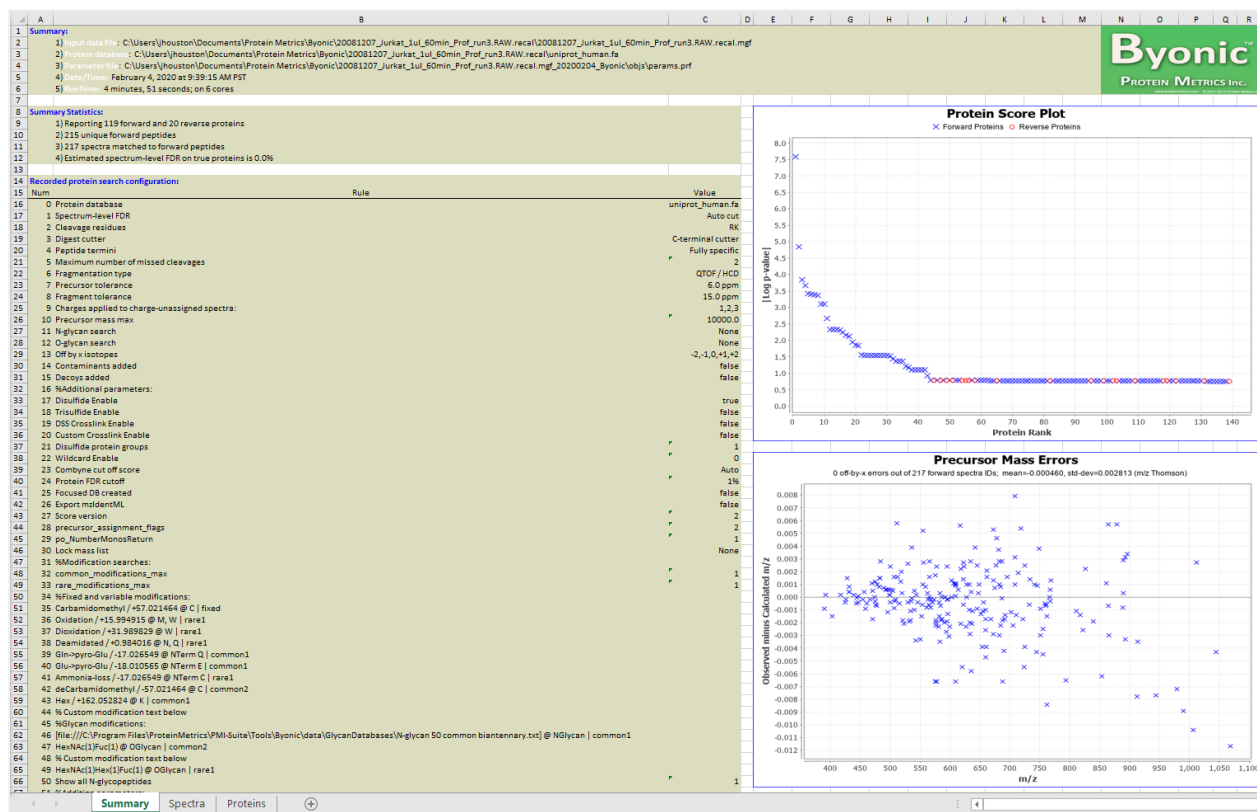


Figure 89: Byonic results (Summary sheet) as viewed in Excel

The \*.xlsx file is saved to the same download directory as the \*.byrs1t file. Note that the Byonic logo graphic may need to be enlarged from a single cell to be viewed. The shaded areas of the Summary sheet display the initial search parameters. The top Protein Score plot charts protein rank vs. log p-value. The bottom Precursor Mass Errors plot charts m/z vs. observed minus calculated m/z.

The **Spectra** sheet displays the Peptides (PSM) table data:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	Protein Rank	Peptide < ProteinMetrics Confidential >	Glycans NHFAgNa	Modification Type(s)	Observed m/z	z	Observed (M+H)	Calc. mass (M+H)	Off- by-x error	Mass error (ppm)	Starting position	Cleavage	Score	Delta	Delta Mod	Log Prob	# of unique peptides	Protein Name
1	Query #2	1 K.IFFAGDTIPK.S			554.806	2	1108.605	1108.604	0	0.9	430	Specific	32.4	32.4	32.4	3.07	10	>sp O75369 FLNB_HUMAN Filamin-B OS=Hom
2	01801.2	1 K.AAGSGELGVTMK.G			560.787	2	1120.566	1120.567	-0.7	481	Specific	101.6	101.6	101.6	3.08	10	>sp O75369 FLNB_HUMAN Filamin-B OS=Hom	
3	07645.2	1 R.VNIGQSGHPQK.V			582.808	2	1164.609	1164.612	-3.0	734	Specific	33.1	33.1	33.1	3.07	10	>sp O75369 FLNB_HUMAN Filamin-B OS=Hom	
4	01108.2	1 K.VFQGVVER.S			430.735	2	860.463	860.463	0	0.5	747	Specific	45.0	45.0	45.0	2.83	10	>sp O75369 FLNB_HUMAN Filamin-B OS=Hom
5	03481.2	1 K.PGTYYVIYV.F			534.298	2	1067.588	1067.588	-0.3	1681	Specific	63.4	63.4	63.4	3.07	10	>sp O75369 FLNB_HUMAN Filamin-B OS=Hom	
6	04871.2	1 R.DAGYGGISLAVEGPKS.V			760.883	2	1520.759	1520.759	-0.4	2029	Specific	51.9	51.0	51.0	3.08	10	>sp O75369 FLNB_HUMAN Filamin-B OS=Hom	
7	08416.2	1 K.VSYFPTVGYIVYVTK.F			878.983	2	1756.958	1756.952	3.3	2059	Specific	14.7	14.7	14.7	3.08	10	>sp O75369 FLNB_HUMAN Filamin-B OS=Hom	
8	04210.2	1 K.FADEHVPVGSPTVTK.I			765.883	2	1530.759	1530.759	0	0.0	2075	Specific	16.3	16.3	16.3	3.07	10	>sp O75369 FLNB_HUMAN Filamin-B OS=Hom
9	02678.2	1 K.VNQPFASFAIR.L			551.803	2	1102.599	1102.600	-1.3	2302	Specific	40.2	21.5	21.5	3.07	10	>sp O75369 FLNB_HUMAN Filamin-B OS=Hom	
10	02986.2	1 K.YGGPNHIVGSPFK.A			686.854	2	1372.701	1372.701	-0.1	2456	Specific	37.2	37.2	37.2	3.07	10	>sp O75369 FLNB_HUMAN Filamin-B OS=Hom	
11	06443.2	2 K.ADLINL+0.984 LGTIAK.S		N[+1]	622.350	2	1243.692	1243.689	0	2.2	101	Specific	2.5	2.5	0.0	3.07	7	>sp P07900 HS90A_HUMAN Heat shock protei
12	06512.2	2 K.TDTGERMGR.G			482.214	2	963.422	963.420	1.6	174	Specific	17.9	17.9	17.9	2.43	7	>sp P07900 HS90A_HUMAN Heat shock protei	
13	00010.2	2 K.HSQFIVGIVTLFVEKE			889.978	2	1778.948	1778.948	0	0.2	210	Specific	56.8	56.8	56.8	3.08	7	>sp P07900 HS90A_HUMAN Heat shock protei
14	00997.2	2 K.EKVIDEELNK.T			704.850	2	1408.693	1408.696	-1.5	282	Specific	16.6	16.6	16.6	3.07	7	>sp P07900 HS90A_HUMAN Heat shock protei	
15	00771.2	2 K.VIDDEELNK.T			576.282	2	1151.556	1151.558	-1.5	284	Specific	13.6	9.4	9.4	2.83	7	>sp P07900 HS90A_HUMAN Heat shock protei	
16	00933.2	2 K.VIDDEELNK.T			576.282	2	1151.556	1151.558	-1.9	284	Specific	22.5	0.4	0.4	2.83	7	>sp P07900 HS90A_HUMAN Heat shock protei	
17	04512.2	2 K.HFSVEGQLFRA			674.836	2	1348.665	1348.664	0	0.5	328	Specific	12.8	12.8	12.8	3.07	7	>sp P07900 HS90A_HUMAN Heat shock protei
18	01320.2	2 K.DNMTKCVKMAV.V			604.364	2	1208.601	1208.601	1.6	631	Specific	12.1	12.1	12.1	3.07	7	>sp P07900 HS90A_HUMAN Heat shock protei	
19	01320.2	2 K.DNMTKCVKMAV.V			604.364	2	1208.601	1208.601	1.6	631	Specific	12.1	12.1	12.1	3.07	7	>sp P07900 HS90A_HUMAN Heat shock protei	
Summary Spectra Proteins																		

Figure 90: Excel Spectra sheet displaying Peptides table data

The **Proteins** sheet displays the Proteins table data:



	A	B	C	D	E	F	G	H	I	J	K	L
	Protein Rank	Description	[Log Prob]	Best [Log Prob]	Best score	Total Intensity	# of spectra	# of unique peptides	# of mod peptides	Coverage %	# AA's in protein	Protein DB number
1	>sp O75369 FLNB_HUMAN Filamin-B OS=Homo sapiens GN=FLNB PE=1 SV=2		7.59	0.78	101.60	858541.0	10	10	0	4.57	2602	6470
2	>sp P07900 HS90A_HUMAN Heat shock protein HSP 90-alpha OS=Homo sap		4.85	0.78	56.70	1373049.5	8	7	1	9.43	732	7978
3	>sp Q13813 SPTA2_HUMAN Spectrin alpha chain, brain OS=Homo sapiens		3.84	0.78	43.90	492532.1	6	6	0	2.22	2472	16438
4	>sp Q14204 DYHC1_HUMAN Cytoplasmic dynein 1 heavy chain 1 OS=Homo s		3.67	0.78	83.70	482097.6	5	5	0	1.08	4646	5263
5	>sp Q9Y490 TLN1_HUMAN Talin-1 OS=Homo sapiens GN=TLN1 PE=1 SV=3		3.43	0.78	55.50	307395.7	5	5	2	2.91	2541	17491
6	>sp P14625 ENPL_HUMAN Endoplasmic reticulum protein OS=Homo sapiens GN=HSP90B1 PE=		3.41	0.78	69.20	639699.4	5	5	0	6.10	803	5569
7	>sp Q01082 SPTB2_HUMAN Spectrin beta chain, brain 1 OS=Homo sapiens		3.39	0.78	81.80	360233.0	5	5	0	3.00	2364	16441
8	>sp P26038 MOES_HUMAN Moesin OS=Homo sapiens GN=MSN PE=1 SV=3		3.37	0.78	33.80	1399723.2	6	6	1	8.32	577	10723
9	>sp P26358 DNMT1_HUMAN DNA (cytosine-5)-methyltransferase 1 OS=Hom		3.11	0.78	90.90	301334.2	4	4	0	2.29	1616	5029
10	>sp P46940 IQGAP1_HUMAN Ras GTPase-activating-like protein IQGAP1 OS=		3.11	0.78	91.90	370119.4	4	4	0	2.41	1657	8560
11	>sp O75643 U520_HUMAN U5 small nuclear ribonucleoprotein 200 kDa hel		2.66	0.78	82.20	327786.5	4	4	0	2.48	2136	18382
12	>sp P14618 KPYM_HUMAN Pyruvate kinase isozymes M1/M2 OS=Homo sapi		2.34	0.78	81.80	223783.8	3	3	0	6.59	531	9293
13	>sp Q92598 HS105_HUMAN Heat shock protein 105 kDa OS=Homo sapiens		2.33	0.78	91.50	328619.5	3	3	0	4.31	858	7962

Figure 91: Excel Proteins sheet displaying Proteins table data

## Byonic Field Descriptions

The Byonic Viewer, as well as the Excel report, contain many fields in the Proteins table and sheet and the Peptides table and Spectra sheet. The following tables contain descriptions of these Proteins and Peptides fields.

Proteins Field	Description
Prot. Rank	The rank of the protein by p-value. The p-value is the likelihood of the protein arising by random chance. For example, for a protein p-value of 0.001, a search against a database containing 10,000 independent proteins, should find only about ten p-values better than 0.001 arising at random. Byonic's p-values are only as accurate as the probabilistic model, so the user should also check the ranking of the proteins relative to the decoy (>Reverse) proteins. Byonic actually outputs the absolute value of the logarithm base 10 of the p-value. Logarithms are useful to prevent numerical underflow, and the absolute value offers compatibility with follow-on tools that expect a larger-is-better score. Byonic's protein table includes a number of columns that reflect the quality of the PSMs for each protein:
Protein Name	The name of the protein
[Log Prob]	Absolute value of the log base 10 of the protein p-value
Best [Log Prob]	Largest [Log Prob] of an individual PSM assigned to the protein
Best Score	Largest Byonic score of a PSM assigned to the protein. For "one-peptide-per-protein" samples, for example top-down proteomics or N-glycopeptide samples, sorting by Best Score may give a better protein list (lower-ranking decoys) than protein [Log Prob]
# Spectra	Total number of PSMs, including duplicate PSMs
# Uniq. Peps.	Total number of PSMs, discounting duplicates (the same modification in a different place is treated as a distinct PSM)
# Mod Preps	Total number of modified peptides
% Cov.	Percent of the protein sequence covered by PSMs. Byonic groups together if their spectra match exactly. An ambiguous PSM, one matching a peptide that is found in two or more proteins, is always assigned to the higher-ranking of those proteins. For example, if PSM P1 has separate evidence but P2 does not, P2 will not be shown. If P1 has more separate evidence than P2, then P1 will be ranked according to all its evidence, but P2 will be ranked according to only its separate evidence. For this reason, as well as others, none of Byonic's outputs (# of spectra, intensity, etc.) is an accurate measure of protein abundance.
# AAs	Number of amino acids

Intensity	Total intensity of MS/MS peaks
ProteinId	Internal Byonic ID used to identify the protein records

Peptides/Spectra Field	Description
PID	Byonic ID used to identify the peptide records
Prot. Rank	Protein Rank (see Proteins Field description, above)
Pos.	Position within the protein sequence of the N-terminal residue of the peptide
Sequence	The peptide sequence
Xlink Base Peptide Xlink Pos. in Protein	The amino acid position of the link in the base protein
Xlink Partner Peptide Sequence	Crosslinked pairs include "base" (under Sequence column) and "partner" (this column)
Xlink Partner Peptide Protein Name	The crosslink bridge's "partner" protein
Xlink Partner Peptide Xlink Pos. in Protein	The amino acid position of the link in the partner protein
Mods (variable)	Modification list (common and rare)
Glycans	Glycan composition. If in numeric form (e.g. 0 1 2 0 0 0): HexNAc Hex Fuc NeuAc NeuGc Sodium
PEP 2D	Protein-aware posterior error probability
PEP 1D	Protein-oblivious posterior error probability
Log Prob	The absolute value of the log10 of the posterior error probability (PEP). The PEP takes into account the Byonic score, delta, precursor mass error, digestion specificity, and so forth (10 features in all). For PSMs with non-negligible error probabilities, say error probabilities > 0.0001, and hence  Log Prob  < 4.0, PEPs are in good agreement with "local FDR" measured by the decoy sequences.  Log Prob  incorporates more evidence than Byonic scores, so that sorting by  Log Prob  will almost always give a better ROC curve than sorting by score.
Score	Byonic score, the "raw" indicator of PSM correctness. Byonic scores reflect the absolute quality of the peptide-spectrum match, not the relative quality compared to other candidate peptides. Byonic scores range from 0 to about 1000, with 300 a good score, 400 a very good score, and PSMs with scores over 500 almost sure to be correct.
Xlink Score	The score of the lower-scoring peptide in a crosslinked pair of peptides
Delta Score	The drop in Byonic score from the top-scoring peptide to the next <i>distinct</i> peptide. In this computation, the same peptide with different modifications is not considered distinct.
Delta Mod. Score	The drop in Byonic score from the top-scoring peptide to the next peptide that is different in any way, including placement of modifications. DeltaMod gives an indication of whether modifications are confidently localized; DeltaMod over 10.0 means that there is a high likelihood that all modification placements are correct.
z	Charge
Obs. m/z	Observed m/z
Calc. m/z	Calculated m/z

ppm err.	ppm mass error (observed - calculated), or more specifically, $10^6 \times (M_{\text{Observed}} - M_{\text{Computed}}) / (M_{\text{Computed}})$ . The ppm mass error is computed after correcting for off-by-x errors
Off -By-X	$[M_{\text{Observed}} - M_{\text{Computed}}]$ , where $M_{\text{Observed}}$ is the observed M+H (singly charged) precursor mass and $M_{\text{Computed}}$ is the computed M+H precursor mass, and $[ ]$ means closest integer
Obs. MH	Observed M+H (singly protonated mass)
Calc. MH	Calculated M+H (singly protonated mass)
Cleavage	Digestion specificity, where Specific means fully specific, Nragged refers to nonspecific at the N-terminus, Cragged refers to nonspecific at the C-terminus, and Non refers to nonspecific at both termini.
Glycans Pos.	Glycan position within the peptide
Z-Score	Z-Score
Protein Name	Protein Name
Prot. Id	Protein record identifier
Scan Time	Scan time, usually reported in seconds (depending on the input data file format). This is usually also the retention time (R.T.) in LC-MS experiments.
Scan #	Scan number(s) for the MS/MS spectrum
QID	Query Identifier (QueriesId)
Mods (fixed)	Modification list (fixed)
Comment	Comment is derived from TITLE= [line] of the source data file. This sometimes contains the scan number from the raw file.
Fragment Type	Fragmentation type (MS2)
Obs.M	Observed neutral mass
Calc.M	Calculated neutral mass
FDR 2D	(# Decoy PSMs) / (# Target PSMs) in full PSM list ranked by PEP 2D
FDR 1D	(# Decoy PSMs) / (# Target PSMs) in full PSM list ranked by PEP 1D
FDR uniq. 2D	(# Decoy Peptides) / (# Target Peptides) in full peptides list ranked by PEP 2D
FDR uniq. 1D	(# Decoy Peptides) / (# Target Peptides) in full peptides list ranked by PEP 2D
q-value 2D	Estimated FDR for the full list of PSMs thresholded at this PSM by PEP 2D
q-value 1D	Estimated FDR for the full list of PSMs thresholded at this PSM by PEP 1D

## Modifications

The following contains additional details about Byonic's modification rules.

**Bundled Modifications.** A modification rule can specify more than one target residue at a time. For example, the one rule

```
Phospho / +79.966331 @ S, T | common2
```

is identical to the two rules

```
Phospho / +79.966331 @ S | common2
```

```
Phospho / +79.966331 @ T | common2
```

**Protein Terminal Modifications.** The following rule specifies acetylation on the protein, but not peptide, N-terminus:

```
Acetyl / +42.010565 @ Protein NTerm | rare1
```

**Protein-Specific Modifications.** The following rule specifies hydroxyproline only on proteins that include the string “collagen” in their FASTA file names.

```
Oxidation / +15.994915 @ P | common3 | ProteinLabel[icase]{collagen}
```

The last field is an example of a modification attribute; to add an attribute pull down the menu highlighted in an orange box in Figure 18: Select Modifications dialog

The keyword [icase] specifies case-insensitive match, so that collagen will also match “Collagen”, “procollagen”, “collegenase”, etc. The keyword [case] specifies case-sensitive match. To match a single protein, use a unique identifier such as the accession number. Protein-specific modifications offer even finer control of the size and focus of a Byonic search and can be extremely useful when search space would otherwise be unnecessarily large, as mentioned below.

**Glycopeptide searches.** The fully automatic, pre-set, glycopeptide searches (the checkboxes under Glycan preset tables) allow only one glycan per peptide. The limitation of one glycan per peptide is not a severe restriction for N-linked glycosylation, because few peptides contain two N-glycosylation motifs, that is, two occurrences of NX{S/T}. However, this limitation is a serious restriction for O-linked glycosylation, especially mucin glycosylation, and for O-GlcNAc-ylation on S/T, because sites for these modifications often cluster together. The following rule gives a reasonable O-GlcNAc search:

```
HexNAc / +203.079373 @ S, T | common3, or alternatively,
HexNAc @ OGlycan | common3.
HexNAc @ S, T | common3
```

The third expression lets the software compute the mass applicable to the specific cases of glycan compositions only.

For a faster but narrower search, use:

```
HexNAc @ S, T | common2
```

For an intermediate search, designate the modification as both common2 and rare1.

An example of a slightly expanded search rule is:

```
HexNAc / +203.079373 @ NGlycan, S, T | common3
```

NGlycan as a modification target means the asparagine in the N-glycosylation motif NX{S/T}, where X is any residue except proline. The expanded rule searches for truncated N-glycans; with the earlier rule a peptide with a truncated N-glycan is likely to be mistaken for an O-GlcNAc peptide.

The NGlycan keyword enables the user to search for more than one N-glycan per peptide. The keyword also enables the user to customize the N-glycosylation search to specific glycan masses, rather than rely on Byonic’s predefined tables. For example, a researcher may have already acquired detailed knowledge of the glycan masses through detached glycan analysis. For human blood serum samples, one of the most common N-glycans is usually HexNAc(4)Hex(5)NeuAc(2) at 2204.77 Da.

Mucin-type O-glycosylation sites tend to cluster, so that a single tryptic peptide may have 5 or more glycosylation sites. Searches allowing 4 or more common modifications are computationally intensive, so these searches are best run with small protein databases and limited lists of glycan masses. With Total common modification max set to 5, a search involving 10 proteins (typically the protein of interest, along with trypsin and contaminants) and 10 likely O-glycans (for example, HexNAc(1), HexNAc(1)Hex(1), HexNAc(1)Hex(1)NeuAc(1), HexNAc(1)Hex(1)NeuAc(2), HexNAc(2)Hex(1), etc.) may already be an undesirably large search. Specifying fully specific digestion, no more than one missed cleavage, and

making the glycosylation protein-specific using the ProteinLabel attribute are ways to help control the combinatorial explosion by limiting the large glycosylation part of the search to just the relevant protein(s).

Byonic allows the user to combine handcrafted modification rules with one or both of the predefined glycopeptide searches. One strategy is to use the predefined searches for exploration, and then iteratively move the modification masses found in confident identifications into handcrafted rules in order to focus the range of modification masses while simultaneously increasing the number of allowed instances of each modification.

**Wildcard mass defect.** If the precursor mass tolerance is low (less than 100 ppm) Byonic obtains the exact mass of the wildcard from the difference between the observed precursor mass and the calculated mass of the candidate peptide. If the precursor mass tolerance is high (within 100 ppm), Byonic uses a mass defect (fractional part of the mass) characteristic of an organic molecule. In either case, Byonic will show the wildcard mass it used in the output. With precursor masses good to 5 ppm, a modest-size wildcard, say in the range -40 Da to 40 Da, may pinpoint the elemental composition.

## False Discovery Rate

Byonic uses a two-dimensional false discovery rate (2D-FDR) estimation method. For more details on this method, see the journal article: "Two-dimensional target decoy strategy for shotgun proteomics," by M. Bern and Y. Kil in J. Proteome Res., 2011, vol 10(12), 5296-5301. PMID 22010998

Internally, Byonic retains all PSMs and all proteins. For presentation to the user, however, Byonic by default cuts the protein list after the 20<sup>th</sup> decoy protein or at the point in the list at which the protein FDR first reaches 1%, whichever cut gives more proteins. On most data sets, almost all the true proteins will be on this truncated list. Researchers with special knowledge of their samples, along with the time to examine protein identifications manually, may choose to make a lower protein cut using the options on the Advanced tab of the Byonic input window. (For making a focused database for subsequent searches, a 2% FDR is generally recommended to make the resulting FASTA file liberal in inclusion of protein sequences.)

After making the protein cut, Byonic makes a PSM cut in order to keep the spectrum-level FDR to a reasonable level, that is, in order to discard false matches to the reported proteins. By default, Byonic makes this cut by discarding the *n* PSMs with the lowest Byonic scores, where *n* is the expected number of random PSMs matching the reported proteins. Byonic then estimates the spectrum-level FDR of the remaining PSMs to the reported proteins; this FDR will typically be in the range 0 – 5%. Alternatively, using the Advanced tab, the user can ask Byonic to make a manual score cut in order to obtain a desired spectrum-level FDR. Warning: A significant number of true PSMs may be lost by imposing an arbitrary low FDR limit rather than accepting Byonic's automatic cutoff.

And a final word of caution: The target-decoy strategy gives an accurate estimate of the rate of completely false identifications at both the protein and PSM levels, but it does not give any estimate of "partially correct" identifications, for example, wrong homologues or splice variants in the case of proteins, or misplaced or incorrect modifications in the case of peptides. Peptide identifications that may be only partially correct can usually be recognized by low DeltaMod (often zero), and manual inspection of the annotated spectrum may in some cases resolve the ambiguity. Even with all of Byonic's advances over the previous state of the art, human judgment remains the ultimate arbiter of subtle identification clues.

## Appendix

### Common Modifications

Below are examples of modifications that are often encountered and the appropriate syntax for including those modifications in a Byonic search.

**Table 1. Cysteine Treatments + Artifacts of Treatment**

The modifications in **blue** apply to samples with the cysteine treatment **in black** above them. The syntax for the rules from the pull-down menu are also accepted in the custom modification entry box as well as the “old syntax” pre-dating the new pull-down menu interface.

Rule from Pull-Down Menu	Explanation
Carbamidomethyl / +57.021464 @ C   fixed	Iodoacetamide treatment
(De)Carbamidomethyl / -57.021464 @ C   common1	Under-alkylation
Ammonia-loss / -17.026549 @ NTerm C   rare1	Pyro-glu from camC
Methyl / +14.01565 @ C   rare1	Propionamide (from gel)
DTT / +151.996571 @ C   rare1	Total mass delta $\approx$ 209 Da
Carbamidomethyl / +57.021464 @ NTerm, H, K   common2	Over-alkylation
Carboxymethyl / +58.005479 @ C   fixed	Iodoacetic acid treatment
DTT / +151.996571 @ C   rare1	Total mass delta $\approx$ 210 Da
Methylthio / +45.987721 @ C   fixed	MMTS treatment
Propionamide / +71.037114 @ C   common2	Common in gel samples

**Table 2. Other Chemical Treatments**

Mixtures of isotopically labeled and unlabeled peptides (e.g., SILAC) are best run as two searches with fixed modifications, rather than as variable modifications, because peptides should be either completely labeled or completely unlabeled.

Rule from Pull-Down Menu	Explanation
Propionyl / +56.026215 @ NTerm, K   fixed	Propionylation
TMT / +224.152478 @ NTerm, K   fixed	TMT0 labeling
TMT2plex / +225.155833 @ NTerm, K   fixed	TMT2 labeling
iTRAQ4plex / +144.102063 @ NTerm, K   fixed	iTRAQ labeling
Label:13C(6)15N(2) / +8.014199 @ K   fixed	Heavy Lysine for SILAC
Label:13C(8)15N(2) / +10.020909 @ R   fixed	Heavy Arginine for SILAC

**Table 3. Common In Vitro Modifications**

Rule from Pull-Down Menu
Deamidated / +0.984016 @ N, Q   common2
Carbamyl / +43.005814 @ NTerm   common1
Gln->pyro-Glu / -17.026549 @ NTerm Q   common1
Glu->pyro-Glu / -18.010565 @ NTerm E   common1
Delta:H(2)C(2) / +26.01565 @ NTerm   common1
Oxidation / +15.994915 @ M   common2
Dioxidation / +31.989829 @ M, W   rare1
Trioxidation / +47.984744 @ C   rare1
Dethiomethyl / -48.003371 @ M   common2

Rule from Pull-Down Menu
Cation:Na / +21.981943 @ D, E   common1
Methyl / +14.01565 @ E   common1
Dehydrated / -18.010565 @ S, T   rare1

**Table 4. Common In Vivo (Biological) Modifications**

Rule from Pull-Down Menu
Oxidation / +15.994915 @ P   common3   ProteinLabel[icase]{collagen}
Phospho / +79.966331 @ S, T   common3
Phospho / +79.966331 @ Y   common2
Methyl / +14.01565 @ K, R   common2
Dimethyl / +28.0313 @ K   common2
Acetyl / +42.010565 @ K   common2
Trimethyl / +42.04695 @ K   common2
Acetyl / +42.010565 @ Protein NTerm   common1
Amidated / -0.984016 @ Protein CTerm   rare1
Sulfo / +79.956815 @ C, S, T, Y   common2
HexNAc / +203.079373 @ S, T   common2
HexNAc(1)Hex(1) @ OGlycan   common2
HexNAc(4)Hex(5)NeuAc(2) @ NGlycan   rare1
HexNAc(4)Hex(5)NeuAc(2)Sodium(1) @ NGlycan   rare1
Hex / +162.052824 @ K   common2

## Byonic: Advanced Commands

Byonic 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. Advanced commands may be entered in the **Modifications** tab by clicking **Enter/Edit** and entering the commands in the **Enter custom text box**.

1. **MS Convert** runs mzML files or uses the legacy file reader in general.

```
ms_convert = pwiz
```

2. **Max Score Keep** is useful if the FASTA file has many sequences that are highly homologous. Use this string in the custom modification area to keep common peptides from being removed from the identified sequence coverage (e.g. mAb has two different HCs).

```
cMaxScoreKeepDesired=1000
```

3. **Cleavage flags** turns off the default, whereby Byonic treats the first and last residues in a FASTA entry as optional. The default is convenient for proteins without initial M or mAbs missing heavy chain C-terminal K, for example. To override the default, set this to 0 (disabled).

```
cleavage_flags=0
```



- During the analyses of proteins that share repetitive sequences (homologs), Byonic will generate a peptide spectrum match (PSM) for every protein from which an identified peptide can originate. Once proteins are scored by summing the log probability for all associated PSMs, only the PSM associated with the highest scoring protein in a group of homologs is kept. There are now two ways to change this behavior in Byonic.

First, you can tell Byonic to exclude homologous peptides when calculating the protein score, meaning that only unique peptides will contribute to the protein score. This provides an unbiased mechanism for assigning homologs to the appropriate protein. To turn the feature on, add the following line to the modifications edit box:

```
dont_score_homologous_peptides = 1
```

Second, it is now possible to prioritize the protein assignment of a homologous peptide to a specific protein. The prioritized targets are defined in the Byonic modifications window. To turn the feature on, add the following line to the modifications edit box:

```
target_protein=example_protein
```

where **example\_protein** can be any substring that matches the fasta header of the target protein(s).

- Advanced command for improved scoring of target proteins**

A bonus score for peptides from target proteins to break ties from equivalent peptide spectral matches (PSMs) in contaminant proteins is available for Byonic searches. This can be done using the advanced commands **target\_protein\_score\_boost** and **target\_protein**.

```
target_protein_score_boost=16.0  
target_protein=NIST
```

In the example above, any PSM from a protein with **NIST** in its protein name will get the score boost mentioned in the custom modifications box.

- Advanced command customization of the N-glycan motif to consider non-canonical glycosites**

Customization of the N-glycan motif to consider non-canonical N-glycosites is now available. This is done by entering `allowed_third_letters_in_nglycan_motif=[x]` into the Byonic modifications field, where [x] is a string of amino acid letters that are allowed in the +2 position. For example, using `allowed_third_letters_in_nglycan_motif=STC`, Byonic will search for that N-glycan mods only in cases where there is the motif NX{S/T/C}.