

PEAKS Studio User Manual (v5.3)

PEAKS Team

PEAKS Studio User Manual (v5.3)

PEAKS Team

Publication date 2011

Table of Contents

I. Basic Operations	1
1. Welcome to PEAKS®	4
1. Main Functions	4
2. Major New Features in 5.3	4
3. Guidelines for Using this Manual	5
4. Scope	6
5. Service and Support	6
2. Installation and Registration	7
1. Package Contents	7
2. System Requirements	7
3. Installation on a Windows Computer	7
4. Registration	8
4.1. Registration with Internet Connection	9
4.2. Registration without Internet Connection	12
4.3. Re-registering PEAKS	14
4.4. Common Errors during Registration	14
5. Adjusting PEAKS' Memory Usage	15
6. What's Next	15
3. A 15-Minute Walkthrough	16
1. Overview	16
2. The Main User Interface	16
3. Examining the Analysis Results	17
4. Conducting the Data Analysis	21
5. What's Next	22
4. Loading Data to a PEAKS Project	23
1. Overview	23
2. Supported Data Formats	23
3. Vendor Specific Requirements	24
3.1. Thermo Data	24
3.2. Waters/Micromass (MassLynx) Data	24
3.3. Agilent Data	25
3.4. Applied Biosystems/Sciex Data	25
3.4.1. QSTAR or QTRAP	25
3.4.2. Instrument Preferences for WIFF	25
3.4.3. ABI 4700/4800	25
3.5. Bruker Data	26
3.5.1. Instrument Preferences for Bruker Data	26
3.6. Shimadzu Data	27
3.7. Varian	27
3.8. PEAKS 5.x Projects	28
4. Creating a New Project	28
5. Adding Data to an Existing Project	30
6. Changing the Default Project Location	30
5. Data Visualization	32
1. Overview	32
2. MS View	32
3. MS/MS View	33
4. Heat Map	34
4.1. Blur / Unblur Heat Map	35
4.2. Highlight Feature / Hide Feature	35
4.3. Mark Feature / Unmark Feature	36
4.4. Show MS2 / Hide MS2	37
4.5. Show PID	37
4.6. Show 3D View	38
4.7. Noise Level	39

6. Adding a Sequence Database	40
1. Configuring Sequence Databases	40
2. Databases to be Used in PEAKS' inChorus Function	42
II. Basic Data Analyses	43
7. Data Refinement	45
1. Overview	45
2. Data Refinement Parameters	45
8. Peptide <i>De Novo</i> Sequencing	47
1. Overview	47
2. <i>De Novo</i> Sequencing Parameters	48
2.1. Mass Error Tolerance	48
2.2. Enzyme Specificity	49
2.3. Fixed and Variable PTMs	49
2.4. Other Parameters	50
2.5. Saving the Parameters for Future Use	50
3. Understanding PEAKS <i>De Novo</i> Sequencing Result	50
3.1. Summary View	51
3.2. <i>De Novo</i> Peptide View	52
3.2.1. Peptide Table	52
3.2.2. Spectrum Annotation	53
3.2.3. Ion Table	55
3.2.4. Error Map	56
3.2.5. Spectrum Alignment	56
3.2.6. Parent Scan	56
4. Filtering <i>De Novo</i> Sequencing Result	56
5. Export <i>De Novo</i> Results	57
6. Run Auto <i>De Novo</i> Sequencing on a Single Spectrum	57
7. Manual <i>De Novo</i> Sequencing	57
7.1. Manual <i>De Novo</i> Graphical User Interface	57
7.2. Manual <i>De Novo</i> Operations	59
9. PEAKS DB	65
1. Overview	65
2. Set PEAKS DB Parameters	66
3. Understanding PEAKS DB Result	67
3.1. The Peptide and Protein Scores	67
3.2. Summary View	67
3.3. Peptide View	70
3.3.1. Peptide Table	71
3.3.2. Peptide Spectrum Match	71
3.3.3. Protein	73
3.4. Protein View	73
3.4.1. Protein Table	74
3.4.2. Peptides	74
3.4.3. Coverage	75
3.4.4. Tool Box	75
3.5. <i>De Novo</i> Only View	75
4. Filter PEAKS DB Result	75
5. Export PEAKS DB Results for Publication	57
6. Comparison of PEAKS DB Results	77
6.1. Comparison Result	77
6.2. Peptide Comparison	78
6.3. Protein Comparison	79
6.4. Statistical Charts	80
6.5. Exporting Comparison Results	80
10. PEAKS InChorus	82
1. PEAKS inChorus Overview	82
2. Understanding PEAKS inChorus Result	83
3. Filter PEAKS inChorus Result	85

III. PEAKS Q - Protein Quantification	86
11. PEAKS Q - MS Level	88
1. Overview	88
2. Setting Parameters	88
3. Understanding the Result	90
3.1. Summary View	90
3.2. Protein View	90
3.3. Peptide View	91
12. PEAKS Q - MS/MS Level	92
1. Overview	92
2. Setting Parameters	92
3. Understanding the Result	93
3.1. Summary View	94
3.2. Protein View	94
3.3. Peptide View	94
13. PEAKS Q - Label Free	96
1. Overview	96
2. Setting Parameters	96
3. Understanding LFQ Result	98
3.1. Summary View	98
3.2. Protein View	99
3.2.1. Extracted Ion Chromatogram	100
3.2.2. Heat Map	100
3.2.3. 3D View	101
3.2.4. MS2 Annotation	101
3.2.5. Isotope	102
4. Filter LFQ Result	102
5. Export Quantification Result to Other Formats	102
6. Replicate Analysis in LFQ	103
6.1. Assign Replicate Number to a Sample	103
6.2. Run Replicate Analysis	105
6.3. Understand Replicate Analysis Results	106
6.4. Export Replicate Analysis Result	108
IV. Advanced Data Analyses	109
14. PTM Finder	111
1. Summary	111
2. Setting Up PTM Finder Parameters	111
3. Understand PTM Finder Result	112
15. Homology Search with SPIDER	113
1. Set SPIDER Parameters	113
1.1. Run SPIDER on PEAKS DB Result	113
1.2. Run SPIDER Independently	115
2. Understand SPIDER Results	115
2.1. SPIDER Peptide View	116
2.2. SPIDER Protein View	116
16. Workflow	118
1. Identification Workflow	118
2. Quantification Workflow	119
3. inChorus Workflow	120
17. Exporting Data/Reports and Printing	121
1. Export Data	121
2. Export <i>De Novo</i> Result	121
2.1. Export Summary and Peptides	122
2.2. Export Images	122
3. Export PEAKS DB Result	123
3.1. Export Summary, Proteins and Peptides	123
3.2. Export Images	124
4. Export Quantification Results	124

4.1. Export Labeled Quantification Results	125
4.2. Export Label-Free Quantification Results	125
4.2.1. Export Result in Excel or HTML	125
4.2.2. Export Summary Page	126
5. Export SPIDER Result	126
6. Export inChorus Result	127
18. Configuration and Preferences	129
1. PEAKS Environment Preferences	129
1.1. General Preferences	129
1.1.1. Display Options	130
1.1.2. RMI Connections	130
1.1.3. Derby Database	130
1.1.4. Performance	131
1.2. Instrument Preferences	131
1.2.1. ABI (.wiff)	132
1.2.2. Bruker (.yep/baf, fid)	132
1.2.3. Shimadzu Axima (.run)	132
1.2.4. Varian (.xms)	132
1.3. Search Engine Preferences	132
1.3.1. Mascot Settings	132
1.3.2. X!Tandem Settings	133
1.3.3. OMSSA Settings	133
1.3.4. Sequest Settings	133
1.4. Spectrum Annotation Preferences	134
2. PEAKS Configuration	135
2.1. Enzyme Configuration	135
2.2. PTM Configuration	136
2.3. Database Configuration	137
2.4. Instrument Configuration	137
2.5. Parameter Configuration	139

Part I. Basic Operations

Table of Contents

1. Welcome to PEAKS®	4
1. Main Functions	4
2. Major New Features in 5.3	4
3. Guidelines for Using this Manual	5
4. Scope	6
5. Service and Support	6
2. Installation and Registration	7
1. Package Contents	7
2. System Requirements	7
3. Installation on a Windows Computer	7
4. Registration	8
4.1. Registration with Internet Connection	9
4.2. Registration without Internet Connection	12
4.3. Re-registering PEAKS	14
4.4. Common Errors during Registration	14
5. Adjusting PEAKS' Memory Usage	15
6. What's Next	15
3. A 15-Minute Walkthrough	16
1. Overview	16
2. The Main User Interface	16
3. Examining the Analysis Results	17
4. Conducting the Data Analysis	21
5. What's Next	22
4. Loading Data to a PEAKS Project	23
1. Overview	23
2. Supported Data Formats	23
3. Vendor Specific Requirements	24
3.1. Thermo Data	24
3.2. Waters/Micromass (MassLynx) Data	24
3.3. Agilent Data	25
3.4. Applied Biosystems/Sciex Data	25
3.4.1. QSTAR or QTRAP	25
3.4.2. Instrument Preferences for WIFF	25
3.4.3. ABI 4700/4800	25
3.5. Bruker Data	26
3.5.1. Instrument Preferences for Bruker Data	26
3.6. Shimadzu Data	27
3.7. Varian	27
3.8. PEAKS 5.x Projects	28
4. Creating a New Project	28
5. Adding Data to an Existing Project	30
6. Changing the Default Project Location	30
5. Data Visualization	32
1. Overview	32
2. MS View	32
3. MS/MS View	33
4. Heat Map	34
4.1. Blur / Unblur Heat Map	35
4.2. Highlight Feature / Hide Feature	35
4.3. Mark Feature / Unmark Feature	36
4.4. Show MS2 / Hide MS2	37
4.5. Show PID	37
4.6. Show 3D View	38
4.7. Noise Level	39
6. Adding a Sequence Database	40

1. Configuring Sequence Databases	40
2. Databases to be Used in PEAKS' inChorus Function	42

Chapter 1. Welcome to PEAKS®



1. Main Functions

PEAKS is an integrated software platform for several common proteomics analyses using tandem mass spectrometry data from all major mass spectrometry vendors. It integrates the following functions in a user-friendly graphical user interface:

- Peptide *de novo* sequencing

De novo sequencing is PEAKS' best-known function. It derives the peptide's sequence from the MS/MS spectrum, without using a protein database. *De novo* sequencing is instrumental to the identification of novel or mutated peptides that are not included in the database, and to the study of the organisms without a protein database.

- Peptide/protein identification with database search

PEAKS can also identify peptides and proteins by searching a sequence database with the MS/MS spectra. Variable PTMs (post-translational modifications) are allowed for the search. PEAKS uses a unique approach to achieve superior accuracy and sensitivity.

- PTM analysis

After the proteins are identified, a second-round search using the included PTM Finder tool can find more peptides with additional PTMs.

- Homology analysis

The included SPIDER tool can identify peptides with an inexact database. Therefore, when the target organism does not have a protein database, the database of a closely homologous species can be used. SPIDER can also help identify single amino acid mutations and protein database errors.

- Protein quantification

PEAKS Q, an optional node, quantifies proteins using a large variety of methods (ICAT, iTRAQ, LFQ, SILAC, and user defined labels).

2. Major New Features in 5.3

Throughout this manual, new and existing features will be discussed in considerable detail. Below is a short list of major new features implemented in PEAKS 5.3:

- Unprecedented accuracy and sensitivity for database search

The revamped database search engine (PEAKS DB) results in substantially improved accuracy and sensitivity for peptide identification, resulting in the identification of more peptides with a reduced false discovery rate (FDR). In particular, the preview version of the PEAKS DB engine produced excellent results in the ABRF/iPRG 2011 study for ETD data analysis.

- Comprehensive result visualization

Numerous improvements were made to support the visual examination of the results from many different angles. Being so confident about PEAKS's performance, BSI is pleased to allow users to have a new powerful mechanism to examine their results so vigorously, and yet so conveniently.

A major addition to most of the analysis results is a summary view that displays the result statistics. Through the charts, users can easily answer important questions such as whether the target-decoy FDR calculation is reliable, and whether the instrument is well-calibrated. Moreover, the peptide and protein tables are rewritten to provide convenient sorting and searching functions, allowing easy location and examination of certain particularly interesting peptides or proteins.

- Built-in result validation

PEAKS DB provides a seamlessly built-in result validation with an enhanced target-decoy method. The score threshold can be conveniently selected from the FDR curve, avoiding the guess work familiar in competitive products.

- Enhanced result reporting

The new summary view provides a central place for specifying the score thresholds to filter the results. Results can be filtered quickly and efficiently with the summary view showing the changes at a glance. The filtered result can be easily exported to several csv and html files for publication or result sharing with non-PEAKS users. Alternatively, the results can be saved as a PEAKS project that a collaborator can view with a free PEAKS Viewer.

- *De novo* sequence tag generation

By specifying a confidence threshold with a user-friendly sliding bar, users can promptly convert the acclaimed PEAKS *de novo* sequencing results to high-confidence *de novo* tags. Users can even export these tags for the integration into their own in-house analysis workflow.

- Improved inChorus search

The inchorus function combines the search results of several search engines. It now supports Mascot 2.3. An intuitive engine icon displays the different engines that identified the same peptide. The new filtration rule allows users to filter the results flexibly, including by individual engines' scores. A venn diagram and the side-by-side FDR curves display each engine's contribution to the final result.

- Heatmap view of quantification results

The heatmap provides a bird's-eye view of the protein quantity changes across different groups of samples. The proteins are automatically clustered according to their quantity change patterns, facilitating the quick identification of possible interesting proteins and patterns.

- Heatmap view for LC-MS data

The data heatmap provides a bird's-eye view of the peaks in LC-MS data, including the peptide features highlighted.

- New statistical peptide score

A rigorous statistical score, $-10\lg P$, is introduced to measure the significance of a peptide-spectrum match in database search. The score is -10 times the common logarithm of the P-value.

3. Guidelines for Using this Manual

This manual is intended to assist in the use of PEAKS 5.3. It outlines functionalities, provides instruction on how to customize PEAKS to a specific application, provides a task based reference, and offers troubleshooting recommendations. Each main function of PEAKS is discussed in one designated chapter.

Note

It is highly recommended that users begin by going through the Walkthrough provided in Chapter 3, *A 15-Minute Walkthrough* to quickly become familiar with the most commonly used features of PEAKS.

4. Scope

It is assumed that PEAKS users are familiar and comfortable with using computers and their respective operating systems. Given this, it is beyond the scope of this manual to instruct the user on the use of Windows, dialogue boxes, menus, file storage, etc. Please refer to the operating system's manual or one of the numerous computer help books for such information. Similarly, PEAKS users are expected to be familiar with mass spectrometry, standard operating practices, data acquisition and analysis.

5. Service and Support

In addition to reading this manual, it is recommended that users take the time to view the accompanying training videos that explain the main features of PEAKS visually and in detail here.

<http://www.bioinfor.com/products/peaks/support/tutorials.php>

Please send technical questions to <support@bioinfor.com>.

We strongly encourage users to provide BSI with any suggestions or comments, as BSI is consistently improving and updating PEAKS to meet the future needs of the scientific community here.

<http://www.bioinfor.com/corporate/contactform.php>

Chapter 2. Installation and Registration

This section of the manual will guide users through the installation and registration of PEAKS 5.3.

1. Package Contents

The PEAKS 5.3 package contains:

- This manual
- PEAKS 5.3 Software
- Quick reference guide for PEAKS 5.3
- Quick reference sheet for mass spectrometry

2. System Requirements

PEAKS 5.3 runs and has been tested on Windows XP, Vista, and 7. PEAKS runs with the following hardware requirements:

Minimum: A dual core processor, 2GB RAM and 100GB free hard drive space.

Recommended: A quad core processor, 4GB RAM, 500GB free hard drive space and 64-bit OS.

Note

When configured correctly, one CPU license of PEAKS 5.3 can take the most advantage of the computing power provided by a quad core processor. Refer to Section 5, “Adjusting PEAKS' Memory Usage” for adjusting the memory usage of PEAKS.

3. Installation on a Windows Computer

Important

Please uninstall any older version of PEAKS currently installed on the system before proceeding.

Important

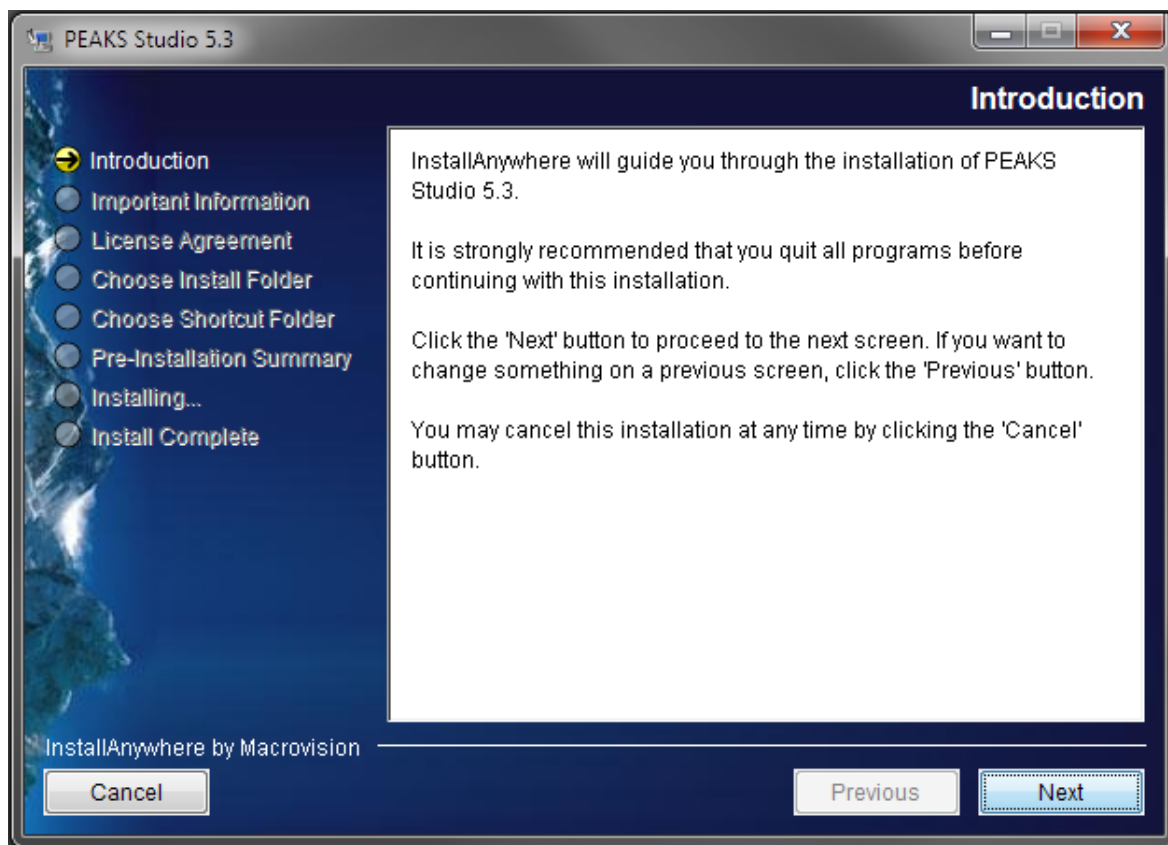
Avoid installing PEAKS in any directory that contains a white space, for example the `Program Files` directory as some features may not function correctly in such situations. Please make sure that the user account has full access permissions (read/write/execute) on the selected directory.

Important

To open an instrument's raw data using PEAKS, it is necessary to install PEAKS on the same computer where the instrument vendors' own software is installed. Refer to Section 3, “Vendor Specific Requirements” for the vendor-specific requirements for raw data loading.

1. Close all programs that are currently running.
2. Insert the PEAKS 5.3 disc into the CD-ROM drive. Or double click on the downloaded PEAKS installation file and move ahead to step 4.

3. The installation window should automatically appear after the CD-ROM is inserted. If it does not, find the CD-ROM drive and open it to access the disc. Double-click on `PEAKS_Studio_Installation.exe`.
4. A menu screen will appear. Select the top item “PEAKS Installer”. The installation utility will launch the installer. When the PEAKS 5.3 installation dialogue appears, click the “Next” button.



5. Follow the on-screen instructions to finish the installation.
6. Adjust the amount of memory utilized by PEAKS as outlined in Section 5, “Adjusting PEAKS' Memory Usage”.

4. Registration

All users are required to register in order to use PEAKS. A license wizard will appear to guide the registration process the first time PEAKS is run. Normally, the registration process is straightforward and involves the following steps. And the detailed instructions of registration with an Internet connection can be found in Section 4.1, “Registration with Internet Connection”.

1. Provide the registration key to the license wizard (or request for the 30 day evaluation without a key).
2. Receive an email containing the license file.
3. Copy the content of the license file into the license wizard.

However, when the computer does not have an Internet connection (or is behind a firewall that blocks the registration), the registration process requires the assistance of another computer with an Internet connection (or outside the firewall), and consists of the following main steps. The detailed instructions can also be found in Section 4.2, “Registration without Internet Connection”.

1. Provide the registration key to the license wizard (or request the 30 day evaluation without a key) on the computer you want to run PEAKS.

2. Save a generated *request file* to a removable storage device (e.g. a USB memory key).
3. From another computer with an Internet connection, upload the request file to BSI's license server.
4. Receive an email containing the license file.
5. Transfer the license file to the computer running PEAKS, and import the license file into the license wizard.

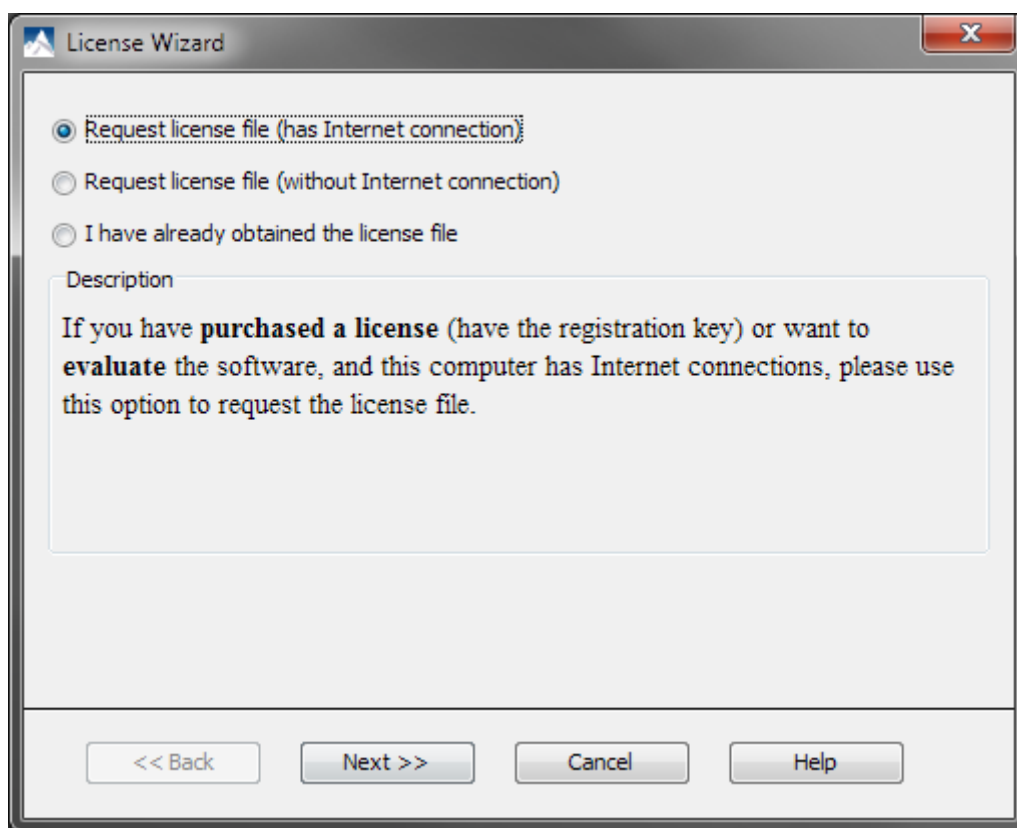
Important

Keep your registration key safe. After a computer hardware upgrade, it might be required to re-register to get a new license file.

4.1. Registration with Internet Connection

Follow the following instructions when the computer has the Internet connection.

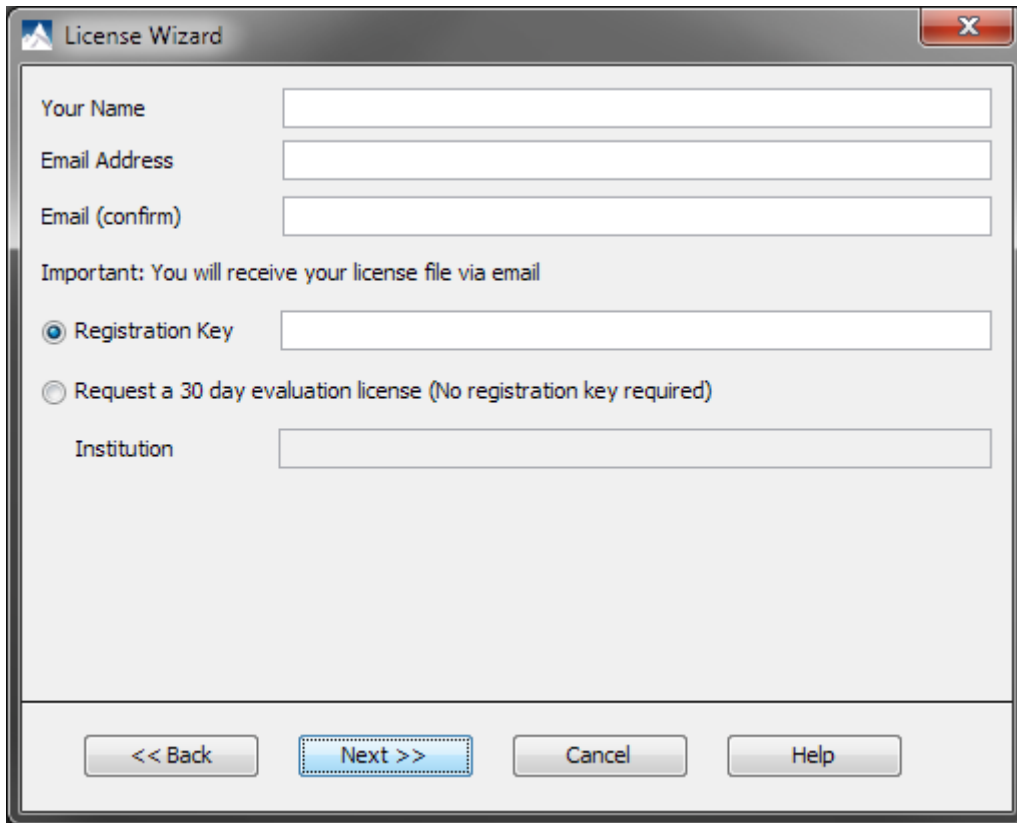
1. When an unregistered PEAKS is run, the following license wizard appears. Select "Request license file (has Internet connection)" and click "Next".



2. The following window will appear. Input the required information. To evaluate the software without a registration key, select the 30 day evaluation option. Click "Next".

Important

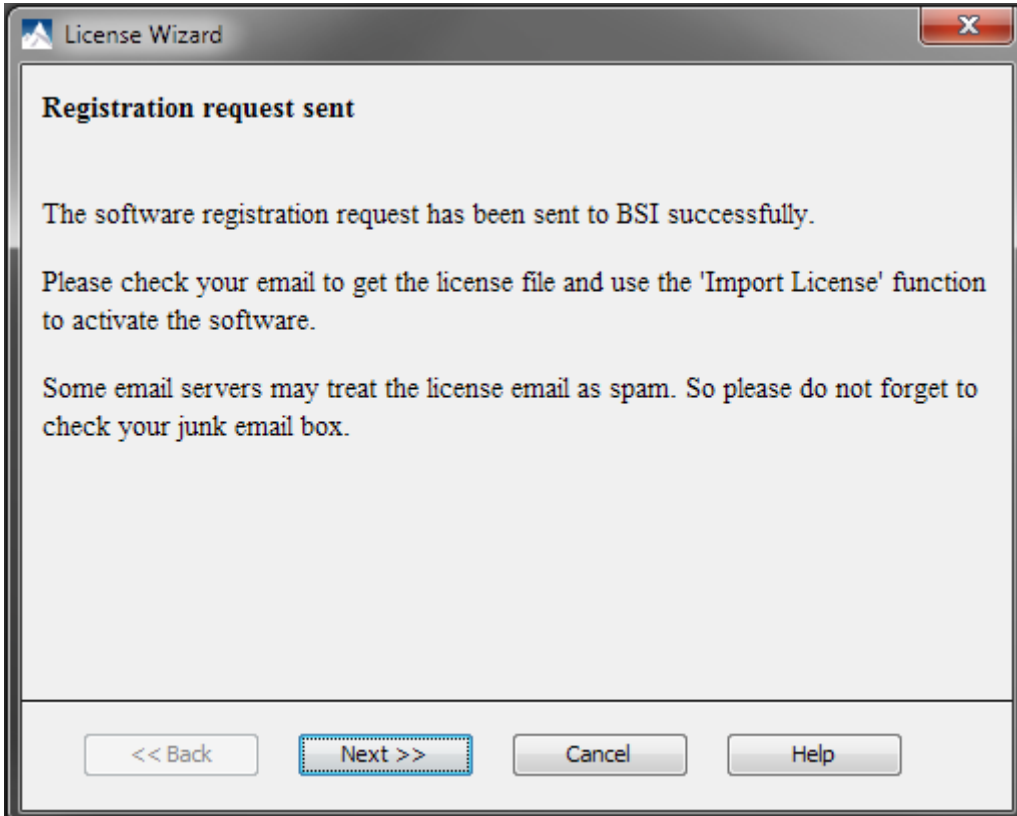
Input ONLY English characters as other character encoding will corrupt the license file.



The 'License Wizard' window contains the following fields and options:

- Text input fields for 'Your Name', 'Email Address', and 'Email (confirm)'.
- A message: 'Important: You will receive your license file via email'.
- Two radio button options:
 - ☒ Registration Key (with an adjacent text input field)
 - ☐ Request a 30 day evaluation license (No registration key required)
- A text input field for 'Institution'.
- Navigation buttons at the bottom: '<< Back', 'Next >>' (highlighted with a dashed border), 'Cancel', and 'Help'.

3. If the license file is requested successfully, the following window will appear. Otherwise, an error message will occur. Make sure that there is no typo in the registration key field in step 2. The registration key is case-sensitive.



The 'License Wizard' window displays the following success message:

Registration request sent

The software registration request has been sent to BSI successfully.

Please check your email to get the license file and use the 'Import License' function to activate the software.

Some email servers may treat the license email as spam. So please do not forget to check your junk email box.

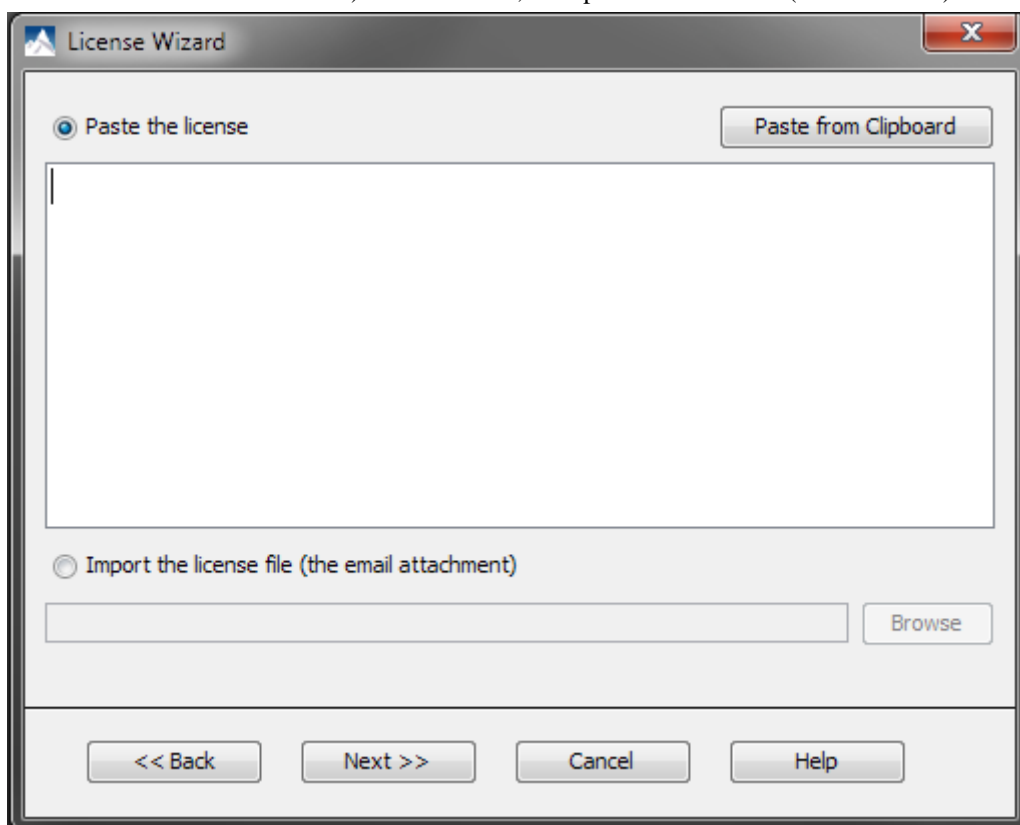
Navigation buttons at the bottom: '<< Back', 'Next >>' (highlighted with a dashed border), 'Cancel', and 'Help'.

4. An automated BSI service will generate the license file (`license.lcs`) and send an email (from `<support@bioinfor.com>`) to the email address provided from the License Wizard. Either save the attached license file or copy the content between '====>' and '<====' in the email to Windows clipboard.

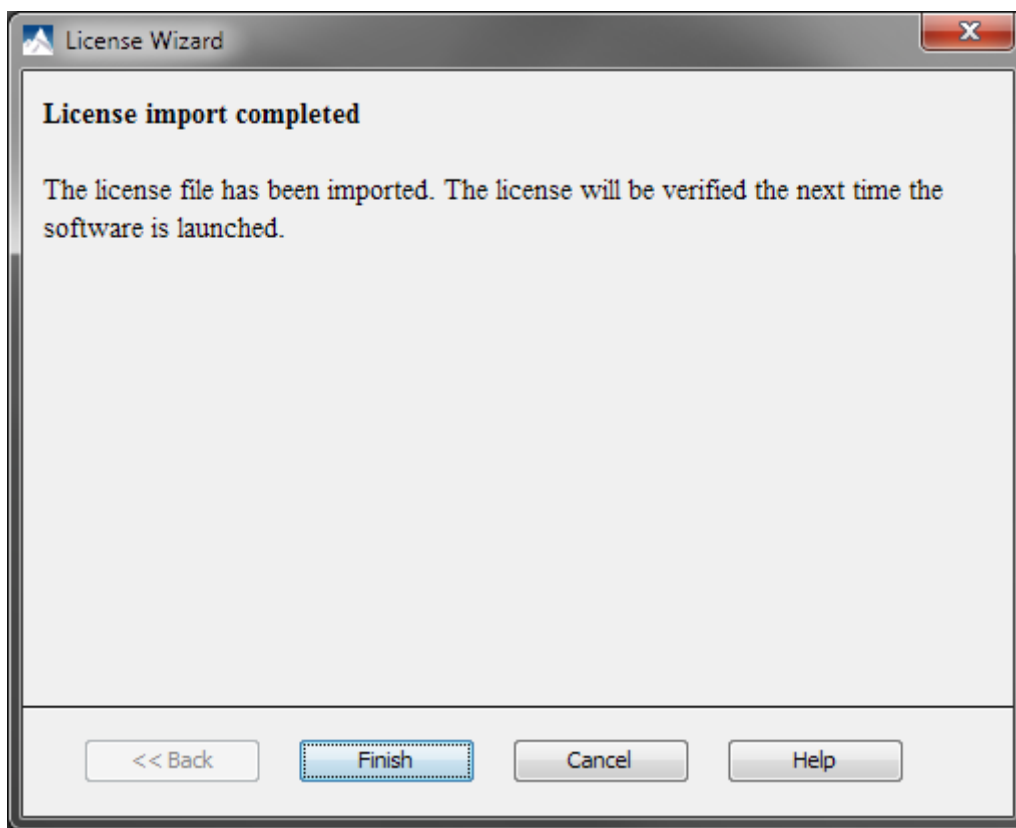
Important

If an email is not received in several minutes, check the junk email box as some email servers treat the license email as spam.

5. Click "Next" in the window of step 3, the following window will appear. Either copy/paste the license content (between '====>' and '<====' in the email) to the text area, or import the license file (`license.lcs`). Click "Next".



6. The following window will open. Click "Finish".

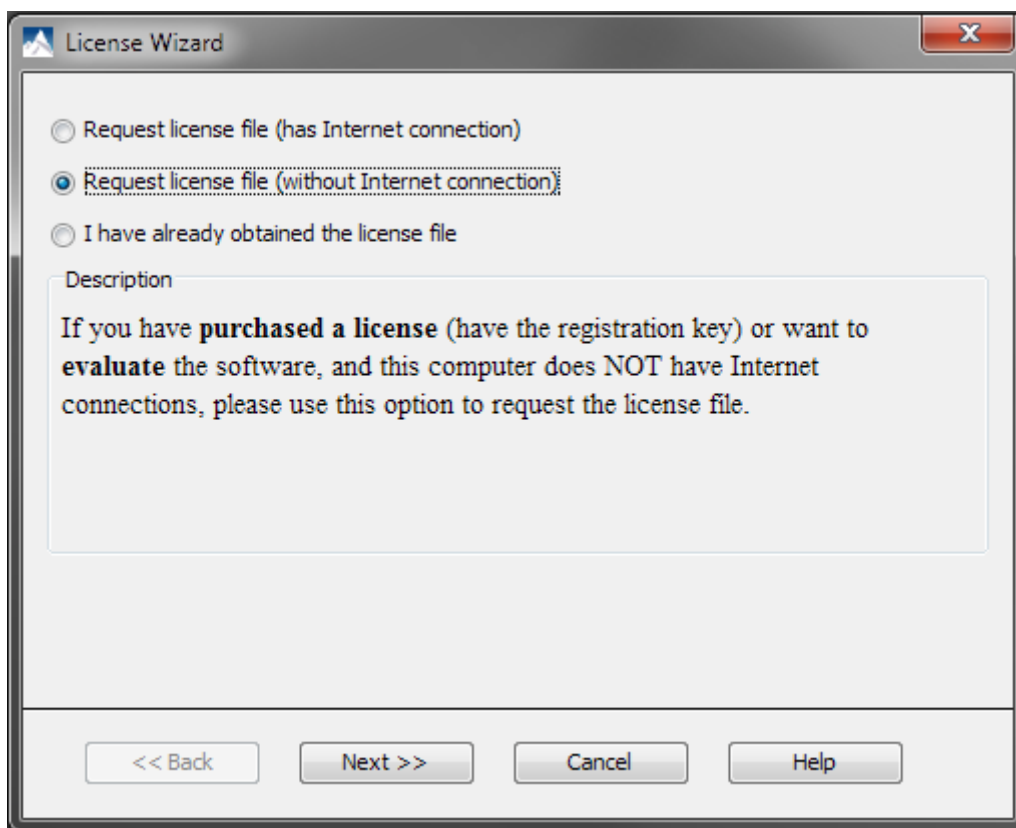


7. Restart PEAKS to begin using the software.

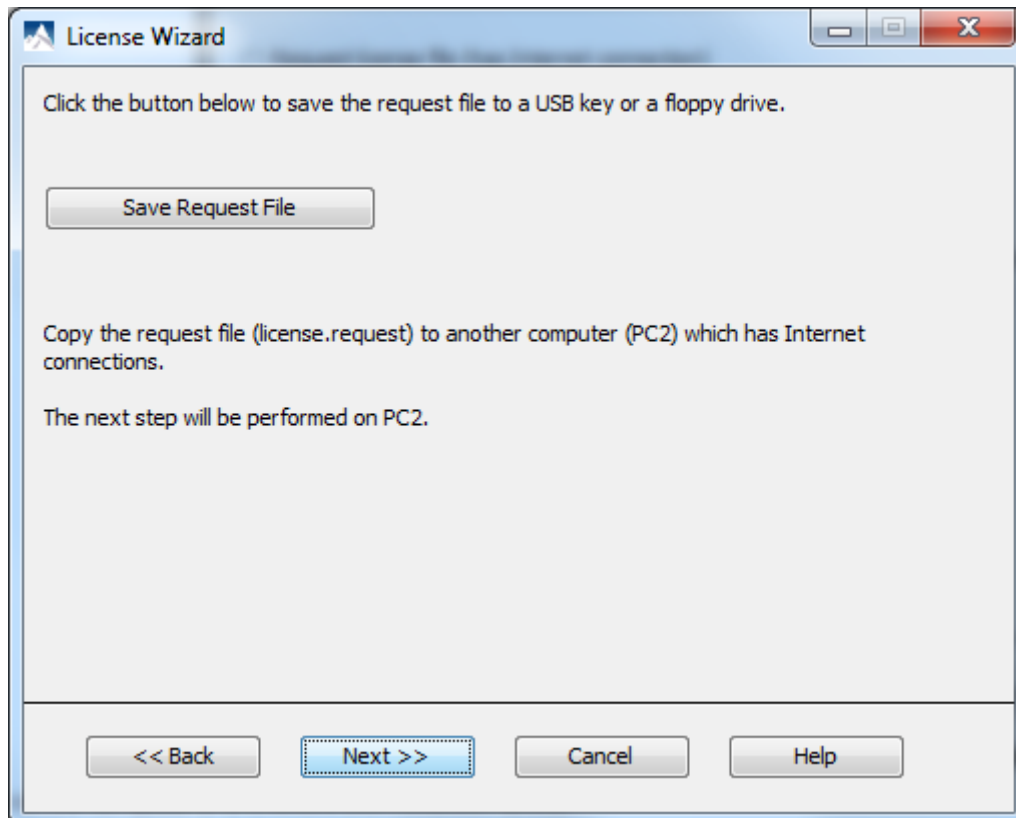
4.2. Registration without Internet Connection

If the computer does not have internet connection, or it is behind a proxy/firewall, or failed using the first option, please follow these instructions.

1. Use the registration interface to "Request license file (without Internet connection)" and follow the instructions.



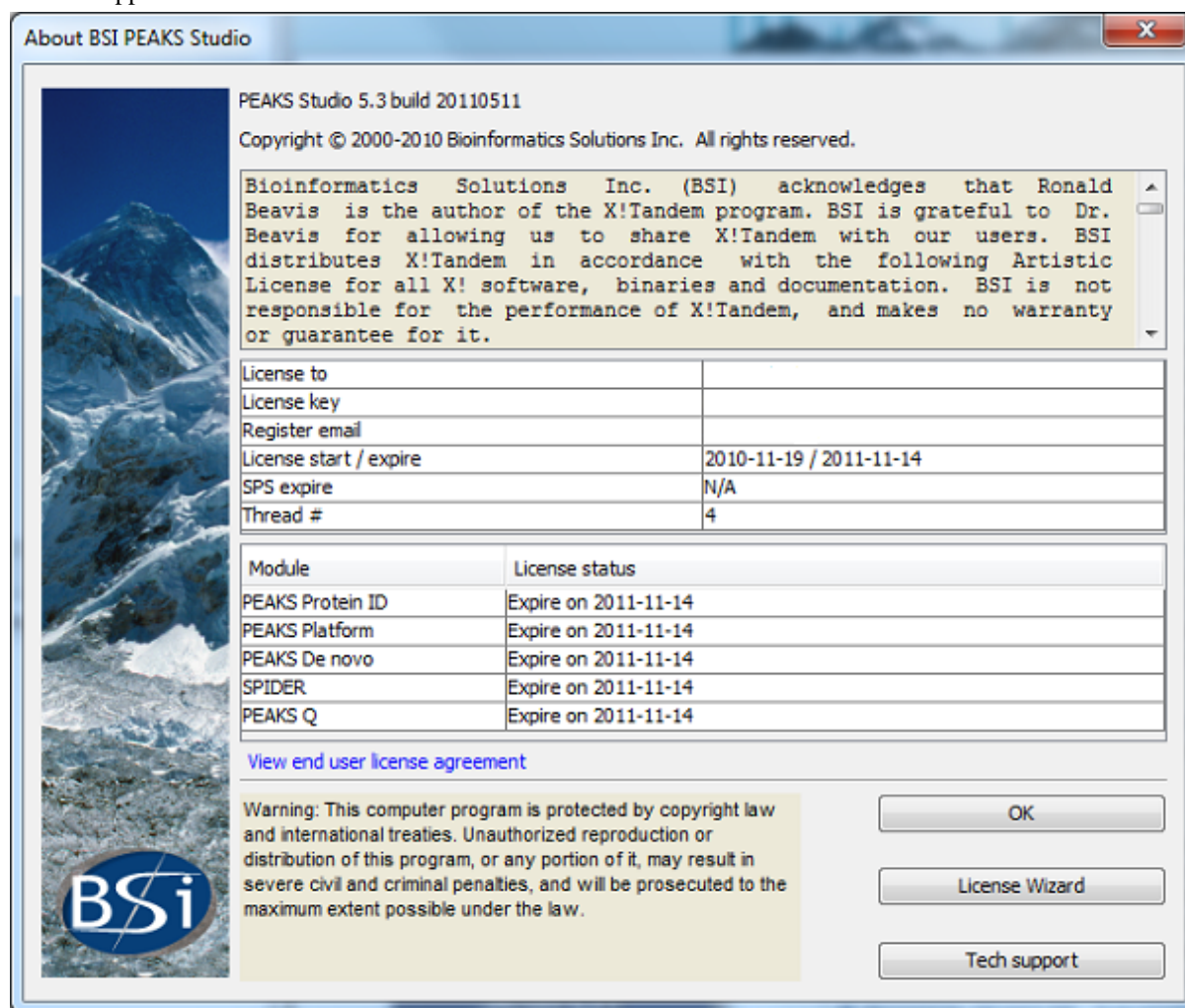
2. This allows users to save a request file on the computer (PC1) where PEAKS was installed.



3. Transfer the license.request file from this computer (PC1) onto a computer with an Internet connection (PC2) using a USB key or a removable storage device. On PC2, go to <http://www.bioinfor.com/lcs20/index.jsp> and follow the onscreen instructions. The license file will be sent to the registered email address.
4. After the license email is received on PC2, save the attachment, `license.lcs`, as is and transfer the file to the computer with PEAKS (PC1). If the license email was not found in the inbox, please check the junk mail folder.
5. Complete the license wizard on PC1 to register.

4.3. Re-registering PEAKS

Re-registering PEAKS may be necessary when an additional software module was purchased or SPS was renewed. BSI will modify the license information accordingly on the server side. A new license file is required to make the changes effective. Select “About PEAKS” from the Help menu. The “About BSI PEAKS Studio” dialogue box will appear:



Click the “License Wizard” button to continue. Then follow the instructions in Section 4.1, “Registration with Internet Connection” or Section 4.2, “Registration without Internet Connection” for re-registering PEAKS.

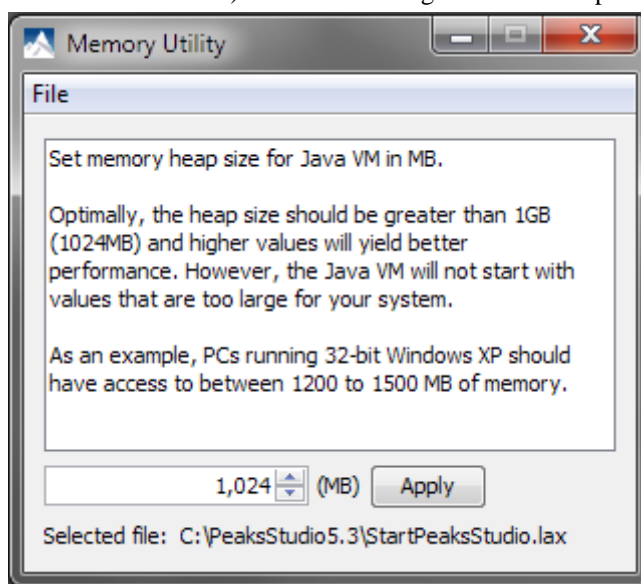
4.4. Common Errors during Registration

1. The registration key is case-sensitive and it is recommended to copy (ctrl-C) / paste (ctrl-V) the registration key whenever possible.
2. The user information can only contain English characters (letters, digits and symbols). Characters from a non-English encoding may cause a registration to fail.

3. If the computer is behind a firewall or has other internet connection problems, the registration may fail. Please follow the on screen instructions or refer to Section 4.2, “Registration without Internet Connection”.

5. Adjusting PEAKS' Memory Usage

By default PEAKS will utilize up to 1GB (1024MB) memory on the computer. If the system is equipped with more memory, it is recommended that this limit should be adjusted as a larger memory will allow PEAKS to take more advantage of the computing power provided by the modern multi-core processor. After installation, the PEAKS directory (e.g., C:\PeaksStudio5.3) contains a file called `Memory Utility.exe`. Click to open this file (it also can be accessed from the Windows Start Menu) and the following window will open.



While this default value (1,024 MB) works fine on most of today's computers, it is recommended to try to increase the value to determine the highest optimal value for the computer. Keep in mind that trial and error may be needed as Java will not start if the value set is too large for the system. Please do not hesitate to contact BSI to ask any questions about the memory usage.

6. What's Next

You are almost done. Depending on the data formats and the type of analysis needed, there may still be two additional configuration steps before data analysis can be conducted.

- To read the instrument's raw data formats, it might be required to install specific software libraries to support the instrument. See Section 3, “Vendor Specific Requirements” for the requirement.
- To conduct database search for protein identification, a protein (or EST) sequence database must be configured. See Chapter 6, *Adding a Sequence Database*.

Or if you are eager to try PEAKS now, leave these two configuration steps aside for a while, and try out the 15-minute walkthrough to get familiar with PEAKS' GUI and basic operations. See Chapter 3, *A 15-Minute Walkthrough*.

Chapter 3. A 15-Minute Walkthrough

1. Overview

In this chapter we quickly walk through some of PEAKS' basic functions by playing with a sample project included in the PEAKS installation. Three concepts are covered:

1. The main user interface.
2. How to examine the analysis results.
3. How to conduct data analysis.

By trying out the few basic tasks in this chapter, users can become reasonably familiar with the PEAKS user interface and some of the important concepts in PEAKS. This will prepare them for the more advanced analyses introduced in the following chapters and save tremendous time later on during the use of PEAKS.

2. The Main User Interface

When you first open PEAKS, the main graphical user interface (GUI) is divided into the following sections, as illustrated in the screenshot.

- A: The user menu and toolbar
- B: The project tree

The data analyses from one experiment are grouped into one *project*. The data are grouped in *samples* and *fractions* of samples. Data analysis is done by selecting a data node and clicking a button in the toolbar. Analytical results are added as *result nodes* on the tree after analysis is done.

- C: The data/result views

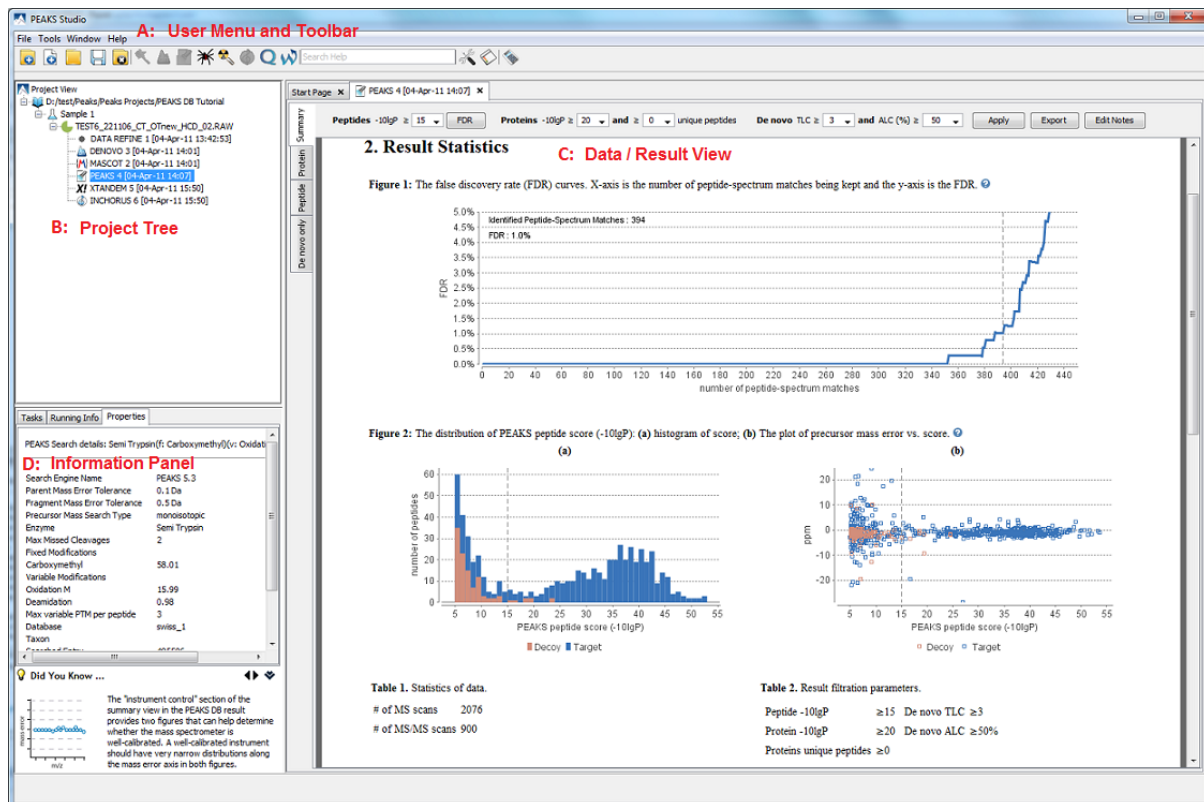
Double clicking a data node or a result node will bring up a panel in this section to display the corresponding data or results. Depending on the types of the data or result nodes; the views may be different and are introduced in more detail in their corresponding chapter of this manual. Multiple views can be opened at the same time as tabs.

Note

It is recommended to close the unused projects and data/result views to preserve memory.

- D: The information panel

This panel provides information corresponding to the data in the progress tree, including parameters and running progress.



3. Examining the Analysis Results

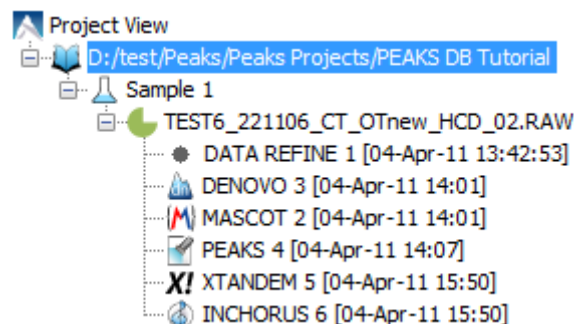
Next let us open a sample project included in the PEAKS installation.

1. Click the open button at the toolbar.



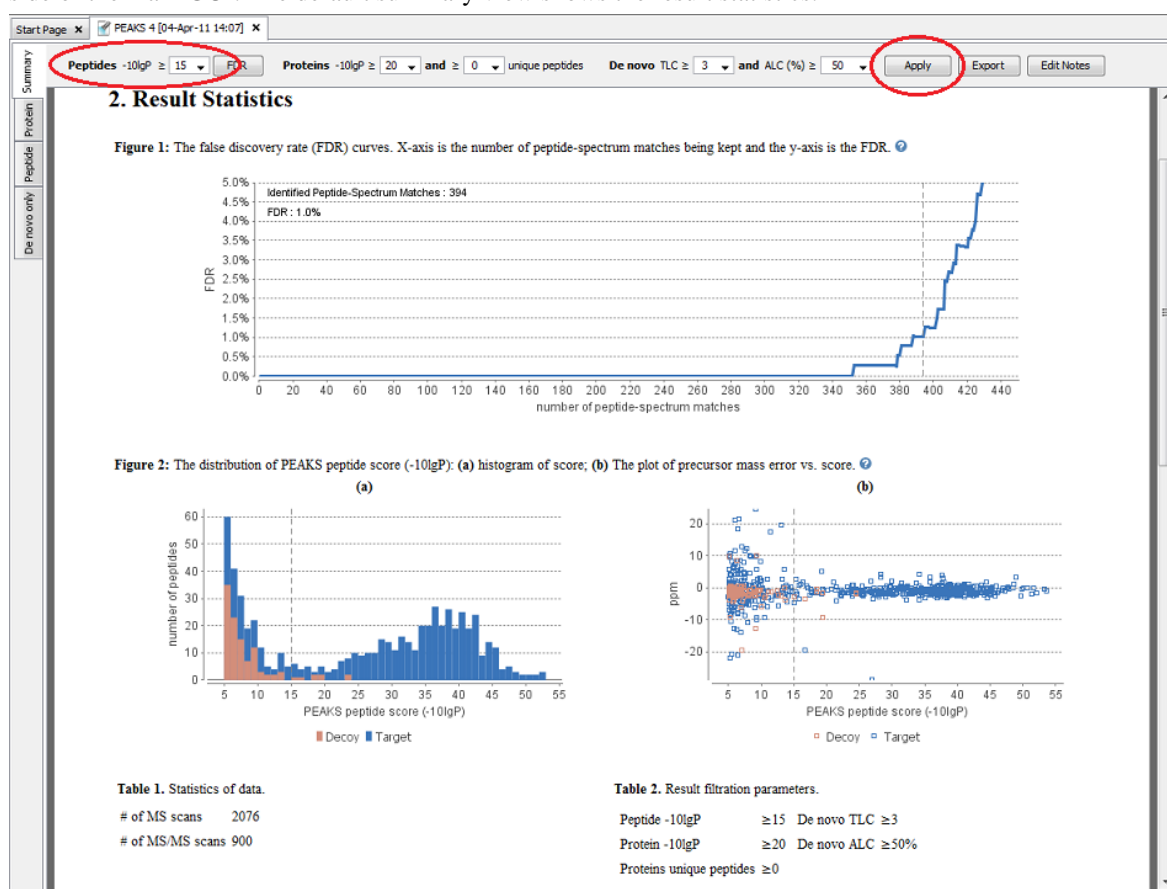
2. Select `PEAKS_HOME/derbyServer/PEAKS DB Tutorial/` in the file chooser. Here you need to replace `PEAKS_HOME` with the directory where you installed PEAKS. If you have not changed the default directory during installation it should be the `C:/PeaksStudio5.3/` directory.
3. Click OK in the file chooser.

The project is now opened in the project tree as shown below.

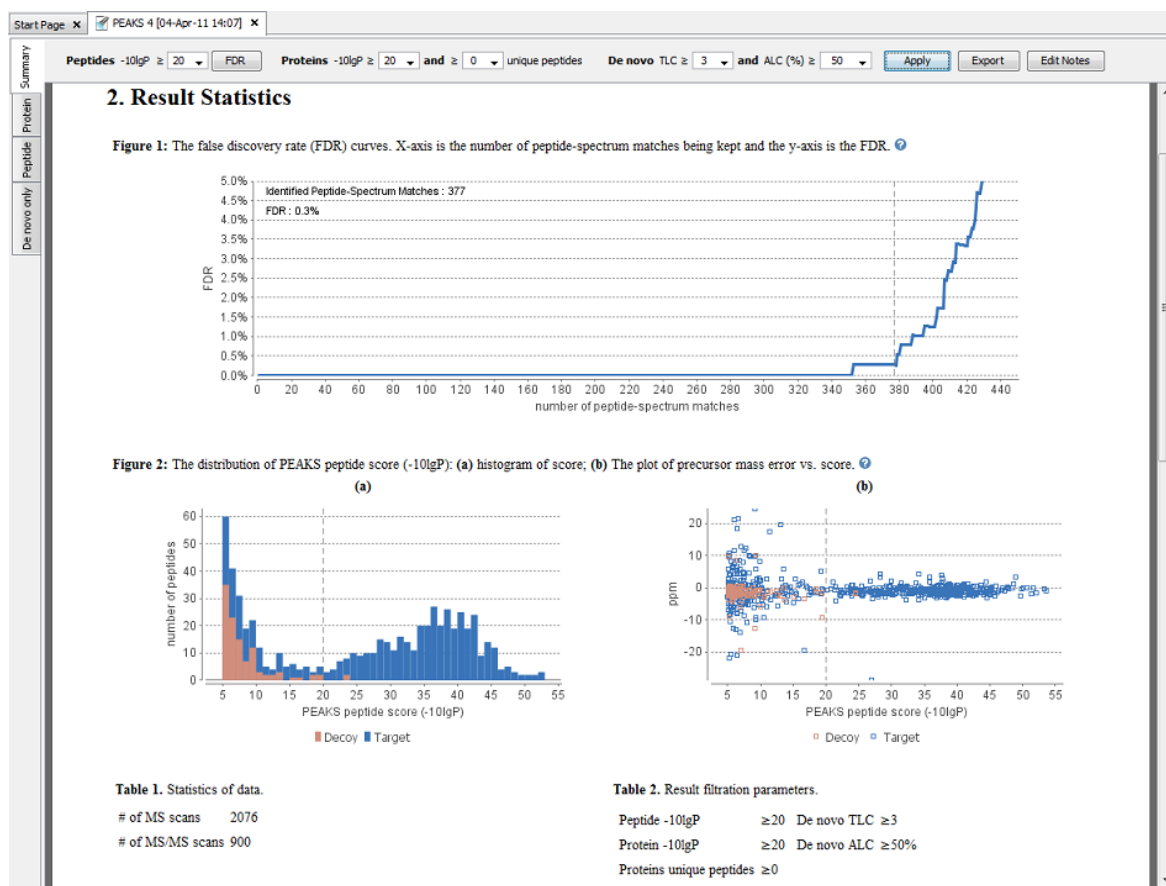


There are already several analysis results in the sample projects. Open a database search result node, PEAKS 4, as an example.

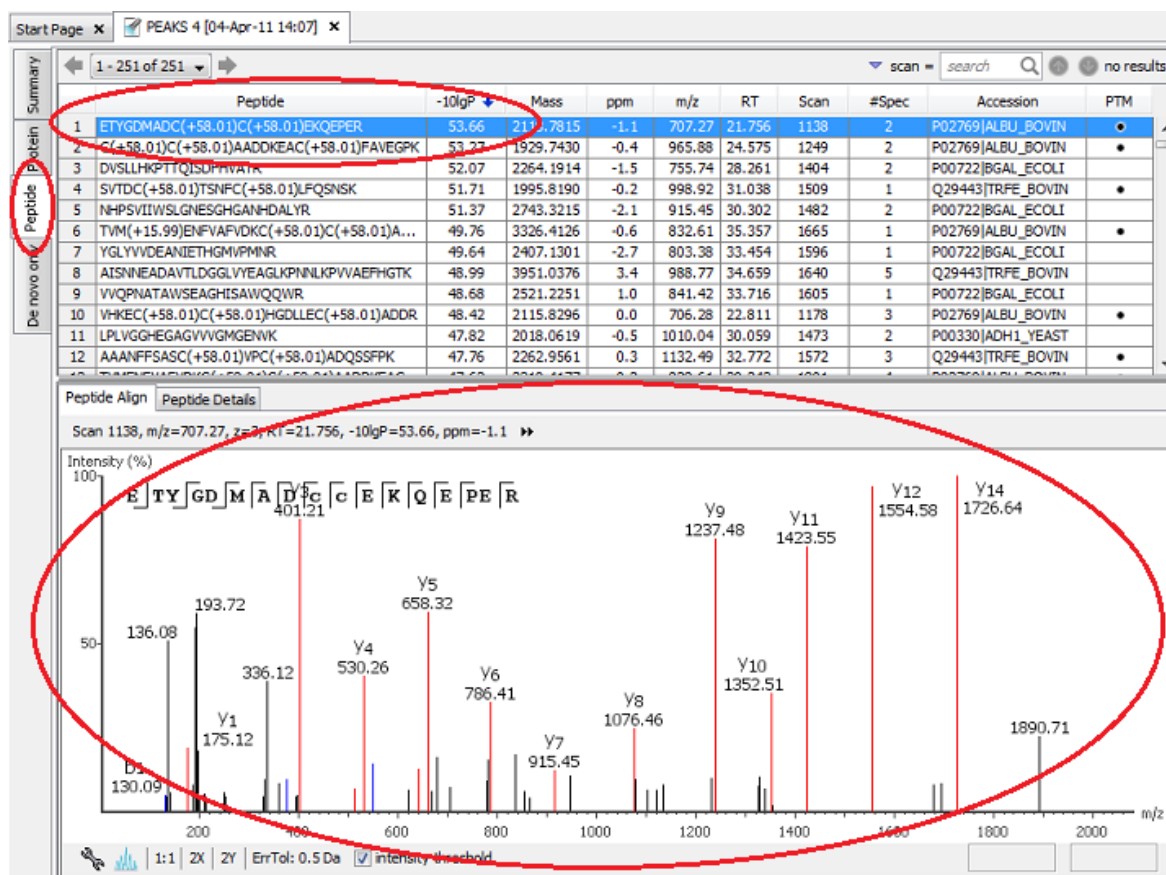
1. Double click the PEAKS 4 result node on the project tree. The following result view is brought up at the right side of the main GUI. The default summary view shows the result statistics.



2. Change the score threshold beside the $-10\lg P$ from 20 to 30. Click Apply button. This will apply a more stringent filter to the peptide identification results. Notice the changes in the result statistics.



- Click the peptide tab to view the identified peptides. Click a peptide to check the spectrum annotation below the table.

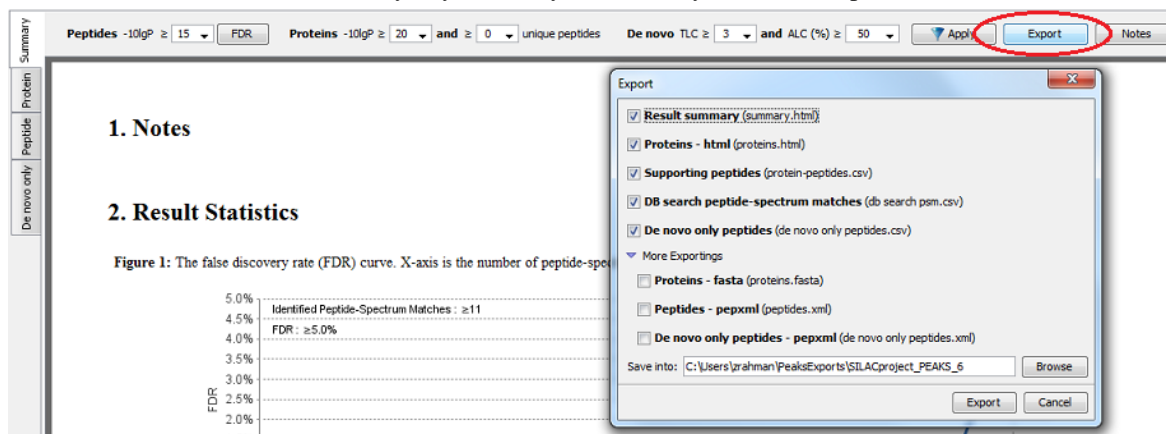


4. Try the following ways to zoom and navigate the annotated spectrum using the mouse.

- Click and drag horizontally to zoom to a region.
- Scroll the mouse wheel button (if your mouse has a wheel) above the x-axis to change peak height, and below the x-axis to zoom along m/z.
- After zooming into a particular spectrum, hold the scroll bar below the spectrum and slide to view the rest of the spectra.
- Double click the spectrum to zoom back to the whole spectrum.

If your lab regularly requires sharing of results with a non-PEAKS user, you may want to try the exporting function as follows. Otherwise, go to Section 4, “Conducting the Data Analysis” for conducting data analysis.

1. Click the summary view tab again. Click the export button to export the results into html and csv (text file viewable in Excel). Select a directory in your file system where you want to export the results.

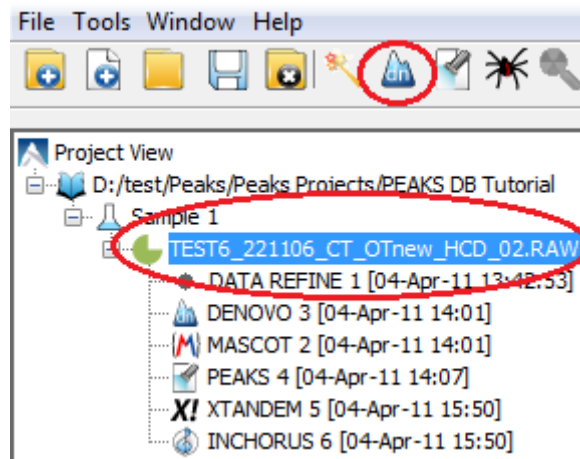


2. Click the index.html file in the target directory to view the exported results with a web browser. You can zip the whole directory and email it to your colleague; or put the whole directory under your own website.

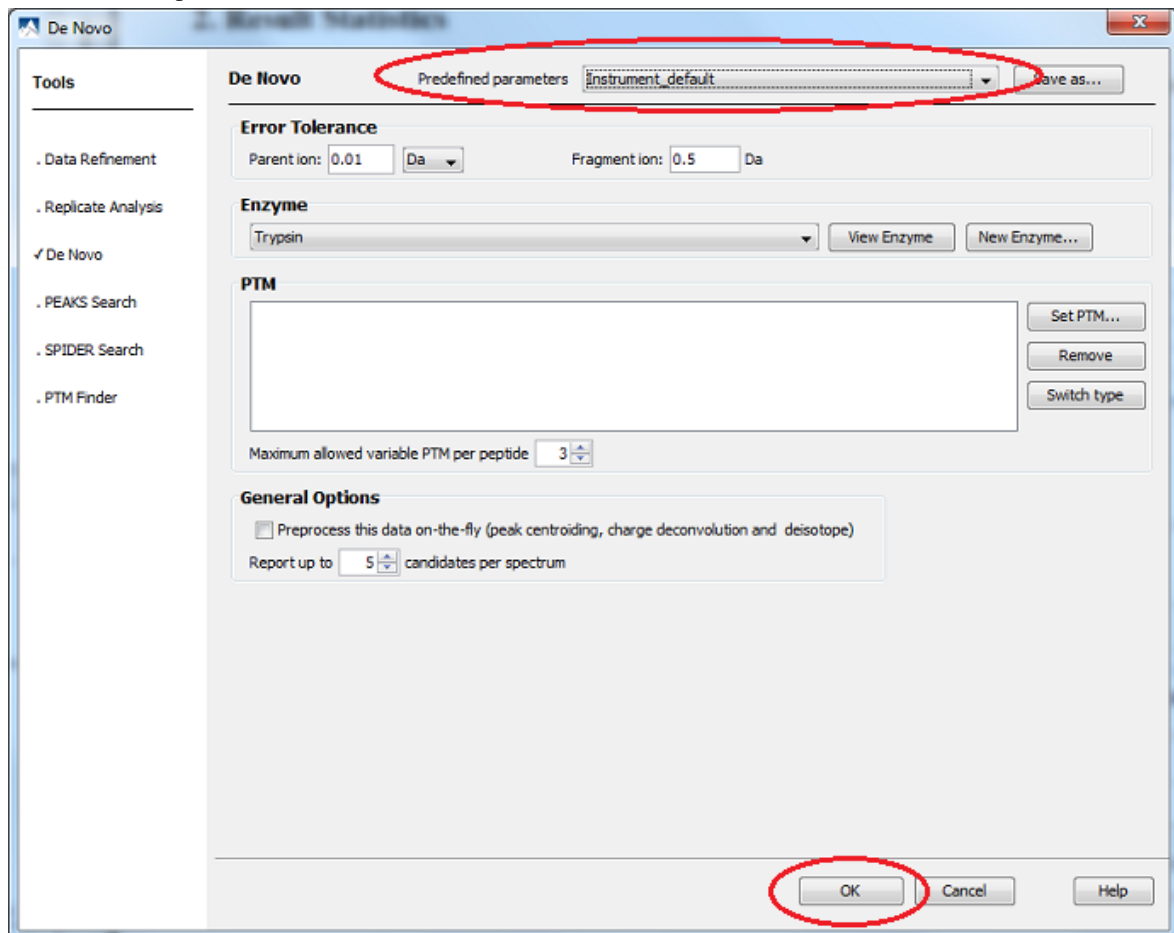
4. Conducting the Data Analysis

In this section we illustrate how to conduct data analysis using *de novo* sequencing as an example. Note that PEAKS has many more functions than just *de novo* sequencing. These will be introduced in later chapters of this manual.

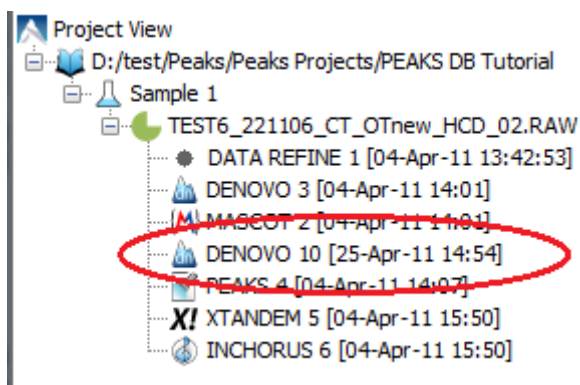
1. If the sample project is not opened, open it as explained in Section 3, “Examining the Analysis Results”.
2. Select the fraction 1 data node. Click the *de novo* sequencing button at the toolbar.



3. Use the default parameters. Click OK.



4. Wait for a few minutes. A new *de novo* sequencing result node shows up at the project tree.



5. Double click the node to examine the *de novo* sequencing results.

5. What's Next


Now you've learnt some of the very basic operations of PEAKS with the sample project. To start using PEAKS for analyzing your own data, and explore the full functions of PEAKS, the following additional work is highly recommended.

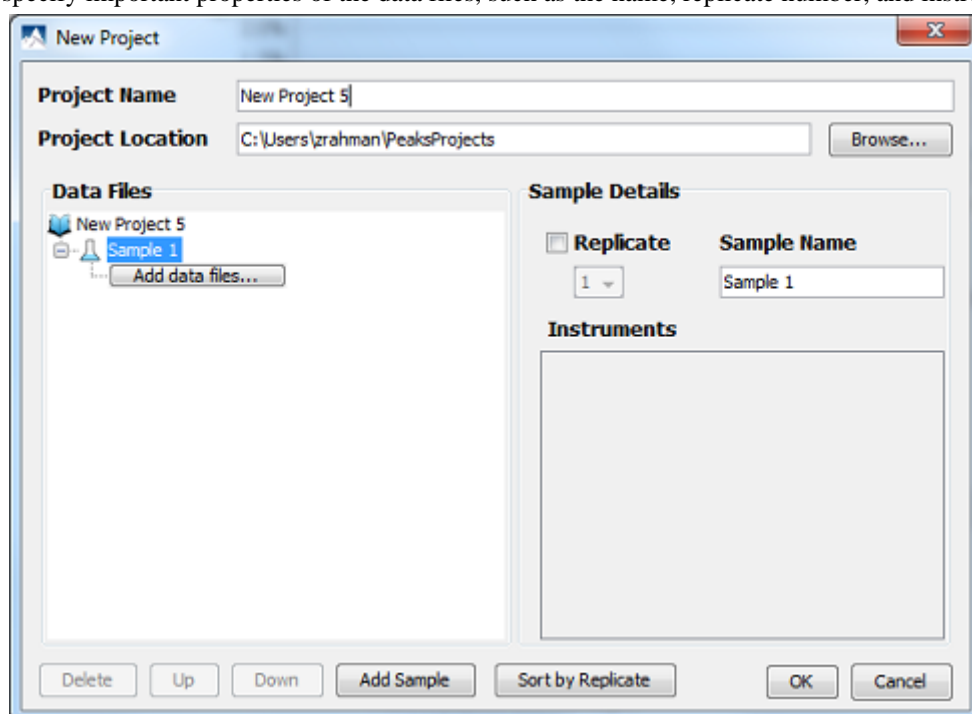
- Read Section 4, “Creating a New Project” for creating a PEAKS project with your own data. This chapter also discusses some additional configuration that may be needed for PEAKS to read the instruments' raw data.
- If you need to identify proteins from a sequence database, read Chapter 6, *Adding a Sequence Database* for configuring a protein or EST sequence database.
- Attend a free training webinar to go through PEAKS' basic functions. The webinar information can be found at BSI's website (<http://www.bioinfor.com/products/peaks/support/webinar.php>).
- Read the corresponding chapters for the analytical tasks you want to conduct. It is recommended to become familiar with the *de novo* sequencing and PEAKS DB analyses before moving on to the more advanced.

Chapter 4. Loading Data to a PEAKS Project

1. Overview


Mass spectrometry data needs to be loaded into a PEAKS project before any analysis can be done. After creation, a PEAKS project is shown as a project node at the project tree of the PEAKS user interface. On the computer's file system, a project is saved as a directory that contains multiple files that contain the compressed spectral data and the analysis results. It is possible to transfer the whole project directory to another user to open with PEAKS Studio or the free PEAKS Viewer.

To create a new project simply click the new project button  on the toolbar. The following “New Project” dialog will appear, where new samples and data files (a.k.a. fractions of samples) can be added. Users also get the chance to specify important properties of the data files, such as the name, replicate number, and instrument type.



PEAKS supports different instrument vendors' raw data formats. A list of supported formats can be found in Section 2, “Supported Data Formats”. Some vendors' formats may require the vendors' specific software to be installed on the same computer that PEAKS is running on. Before you trying creating a project with your own data, ensure that the vendor-specific requirements discussed in Section 3, “Vendor Specific Requirements” are met.

Upon clicking the OK button in the "New Project" dialog, PEAKS will make an effort to import the vendors' raw MS data into the PEAKS project. Once the data is loaded it becomes a part of that project so you can manipulate/delete the original data files without affecting the analysis in PEAKS.

To close a project that you are working on, select the “Close Project” command from the file menu or use the close project icon  from the tool bar. It is recommended to close the unused projects to preserve computer memory.

In the rest of this chapter we discuss the details of data loading and project creation.

2. Supported Data Formats

Following is the list of supported data formats in PEAKS. PEAKS supports these formats at three different levels:

- Native Support - PEAKS can read the following files directly without additional work:
 - mzXML
 - mzData
 - mzML
 - .DTA file or a directory of .DTA files
 - .MGF
 - .PKL
 - PEAKS 5.x projects. See Section 3.8, “PEAKS 5.x Projects”.
- Library Level Support - The instrument vendor's software library is required to be installed at the same computer. PEAKS will call the software library to read the data directly.
 - .RAW file (Thermo Fisher Scientific). See Section 3.1, “Thermo Data”.
 - .D directory (Agilent instruments). See Section 3.3, “Agilent Data”.
- Convertor Level Support - Third party convertors are required. Users may need to install the required convertors and let PEAKS know their locations. This only needs to be set up once. PEAKS will call the convertor to convert the data to another supported format before loading. The actual conversion process is invisible to the user.
 - .RAW directory (Waters QTOF instruments). See Section 3.2, “Waters/Micromass (MassLynx) Data”.
 - .WIFF file (AB/Sciex QSTAR and QTRAP instruments). See Section 3.4.1, “QSTAR or QTRAP”.
 - .T2D file (AB 4700/4800 series). See Section 3.4.3, “ABI 4700/4800”.
 - .LIFT or .D directory (Bruker instruments). See Section 3.5, “Bruker Data”.
 - .RUN folders from Shimadzu instruments. See Section 3.6, “Shimadzu Data”.
 - .XMS files from Varian instruments. See Section 3.7, “Varian”.

3. Vendor Specific Requirements

Most vendors provide tools for MS analysis software to read their raw data format. PEAKS works best with raw data because it is unprocessed. This allows it to use the data pre-processing tools built in to the software designed to maximize identification results. Listed below are the requirements to load raw data from each supported vendor.

3.1. Thermo Data

RAW data from Thermo Fisher Scientific mass spectrometers can be loaded, provided that the XCalibur software or the Thermo MSFileReader package is installed on the same computer as PEAKS 5.3. MSFileReader is publicly available and can be found at the following link:

<http://sjsupport.thermofinnigan.com/public/detail.asp?id=586>

Converting with MSFileReader will only work if there are only English characters in the file path.

3.2. Waters/Micromass (MassLynx) Data

RAW data from Waters instruments can be imported provided that MassLynx 4.1 software is installed on the same computer as PEAKS. MassLynx 4.0 users, can download a different version of wolf.exe. Command line can be used to convert raw files to mzXML with wolf.exe. You can also replace the file: PeaksStudio5.3/wolf.exe with the program compatible with MassLynx 4.0. For links to different versions of Wolf visit the link below:

<http://www.bioinfor.com/products/peaks/support/watersmicromass.php>

3.3. Agilent Data

Agilent QTOF data can be loaded provided that MassHunter software is installed on the same computer.


Agilent Ion Trap data can be loaded provided that CompassXport is loaded on the same computer as PEAKS. The instrument produces a .d folder. The spectral data within this folder is in the .yep file. Select this file to load the raw data.

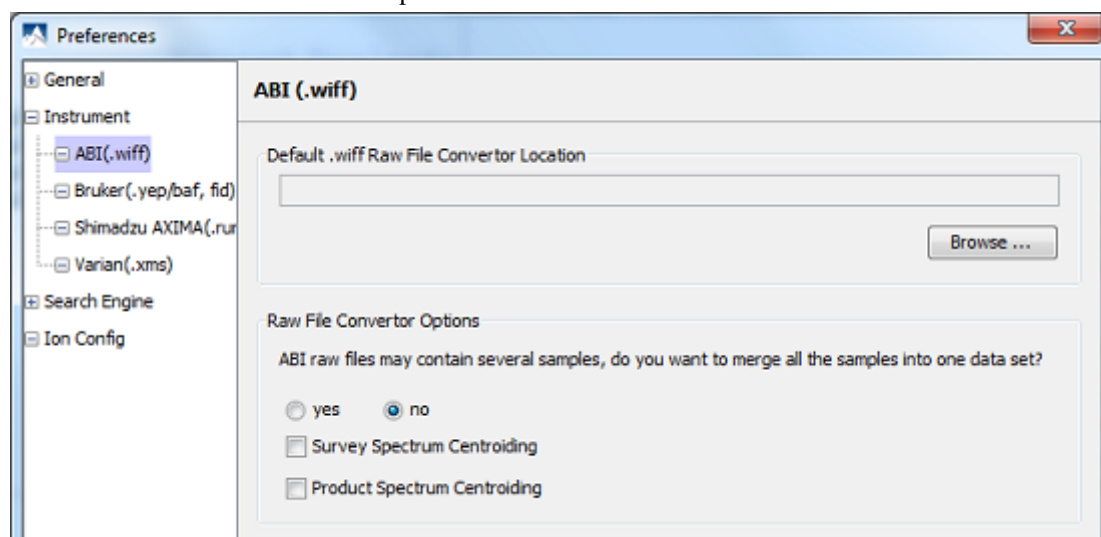
3.4. Applied Biosystems/Sciex Data

3.4.1. QSTAR or QTRAP

WIFF data from Applied Biosystems/Sciex QSTAR or QTRAP mass spectrometers can be loaded. Analyst QS for QSTAR or Analyst 1.4 for QTRAP and the MSX plug-in must be installed on the same computer as PEAKS. The MSX tool is produced and sold by Infochromics Ltd., and is available (at cost) from Bioinformatics Solutions Inc. Please contact a BSI sales representative to obtain an evaluation or full license.

3.4.2. Instrument Preferences for WIFF

To set WIFF related preferences in PEAKS, click the “Preferences” toolbar icon  or select “Preferences” from the “Window” menu to open the “Preferences” window. Clicking on “Instrument” and then “ABI (.wiff)” on the menu on the left hand side will show the preferences for ABI instruments.



Default .wiff Raw File Converter Location. Click “Browse” to tell PEAKS the location of the default .wiff raw file converter. *MSX*, a commercial .wiff converter developed by Infochromics, is available from BSI.

Raw File Converter Options. ABI raw files may contain several samples. By default, these samples are not merged into one data set. Select “yes” if you would like PEAKS to merge all the samples into one data set. Select “Survey Spectrum Centroiding” if centroiding has been performed before loading the data into PEAKS. Select “Product Spectrum Centroiding” if centroiding has been performed on the product spectrum before loading into PEAKS. This is important to insure PEAKS performs optimally.

3.4.3. ABI 4700/4800

T2D files can be extracted and imported into PEAKS with a free tool created by BSI. If you require this separate free tool, contact your sales representative. Instructions on how to use this tool follow.

System Requirements. This extractor can be installed on the same machine as ABI 4700 Explorer and the Oracle database or another machine that has direct network access to the 4700 SERVER. There cannot be a firewall or proxy between the computers. Windows 2000 or Windows XP is recommended for use with this tool.

Configuration. Start the ABI 4700 converter tool. Choose “Settings” from the “File” menu. Configuration requires the following:

- **4700 SERVER Name or IP Address:**input “localhost” if the Extractor is running on the same computer as ABI 4700 Explorer (this is the default value), otherwise enter the IP address of the 4700 SERVER.
- **The socket used by the 4700 SERVER:**this is the port that the Oracle database listens to (the default is 1521).
- **Username to access the Oracle database:**most likely we do not need to change this (the default is “tsquared”).
- **Password to access the Oracle database:**mostly likely you do not need to change this either.

Data Extraction Procedure. The data extraction requires


1. Load Spot Set List from the database: (Do this via menu File | Load Spot Set List). The extractor will export the peak list of a spot set into a PKL file.
2. Open a Spot Set: (menu File | Open Spot Set) Spot Set Chooser will help the user to choose a spot set. After selecting a spot set, click 'OK' to open it. The job run information of a spot set will be shown.
3. Select a job to run: There is a button to select before each run. Only the MS/MS job run can be selected for export, as the precursor information is needed. Select a job run and click 'Convert' to do the extraction.
4. Choose a filename to save: After clicking the 'Convert' button, the user needs to input a file name and the peak lists of the selected job run will be exported.

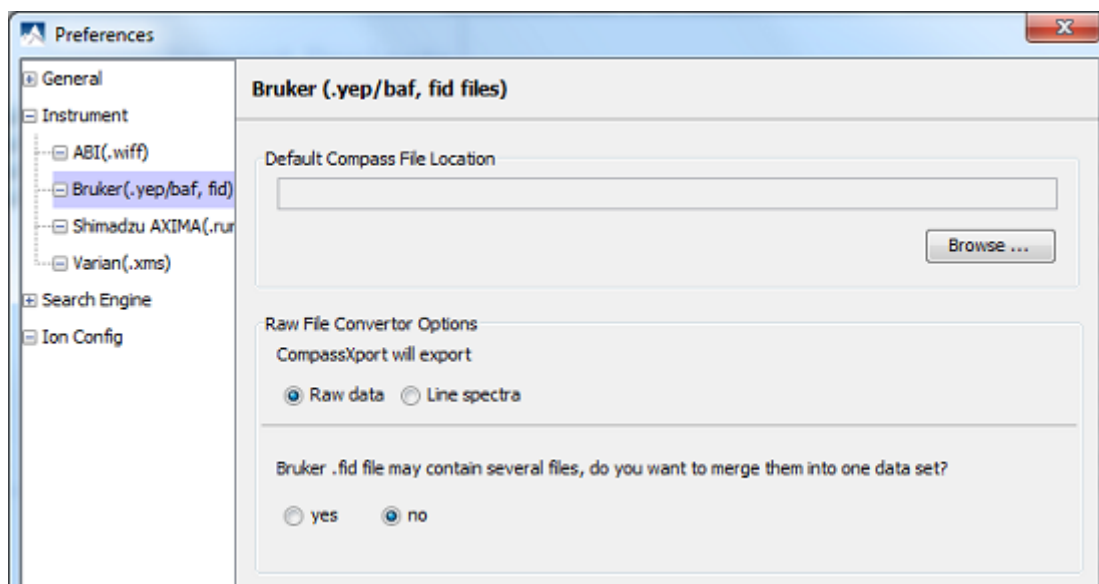
3.5. Bruker Data

.D and .LIFT directories from Bruker mass spectrometers can be imported provided that CompassXport software is installed on the same computer as PEAKS. The spectral data will be contained in the .yep, .baf, or fid file. If loading fid files, you can select the top level folder to load them all at once (make sure the merge option is set in the Bruker instrument preferences refer to Section 3.5.1, “Instrument Preferences for Bruker Data”)

CompassXport 3.0 is readily available on the Bruker Daltonics web site. PEAKS currently cannot import raw data from CompassXport 3. You can either import line spectra instead (refer to Section 1.2.2, “Bruker (.yep/baf, fid)”) or you can install CompassXport 1.3. This version is fully compatible with Peaks. You may need to contact your Bruker representative to obtain CompassXport 1.3.

3.5.1. Instrument Preferences for Bruker Data

To set Bruker data related preferences in PEAKS, click the “Preferences” toolbar icon  or select “Preferences” from the “Window” menu to open the “Preferences” window. Clicking on “Instrument” and then “Bruker (.yep/baf, fid)” on the menu on the left hand side will show Bruker instrument preferences on the right hand side.




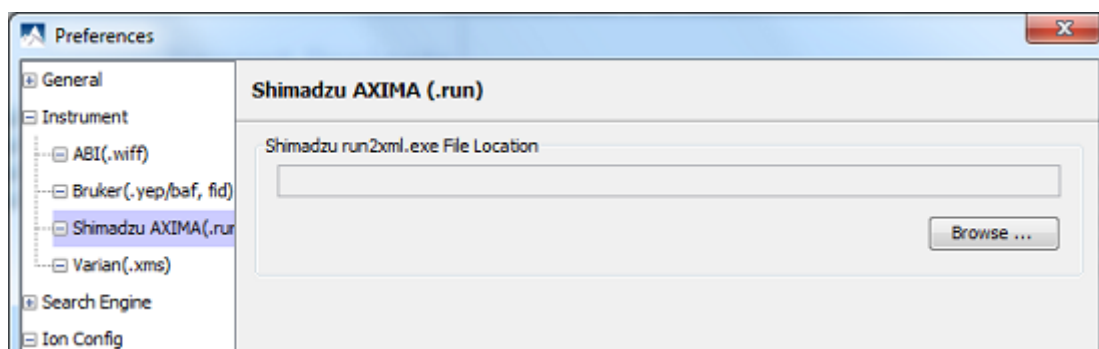
Default Compass File Location. Click “Browse” to tell PEAKS the location of the *CompassXport* file converter.

Raw File Converter Options. CompassXport by default will export raw data. If you attempt to load raw data and no spectra are displayed choose to export line spectra. Bruker .fid files may contain several samples. By default, these samples are not merged into one data set. Select “yes” if you would like PEAKS to merge all the samples into one data set.

3.6. Shimadzu Data

RUN files from Shimadzu mass spectrometers can be loaded provided that the Shimadzu software is installed on the same computer as PEAKS.

Instrument Preferences for Shimadzu Data. To set Shimadzu data related preferences in PEAKS, click the “Preferences” toolbar icon  or select “Preferences” from the “Window” menu to open the “Preferences” window. Clicking on “Instrument” and then “Shimadzu AXIMA (.run)” in the menu on the left hand side will show the Shimadzu instrument preferences on the right hand side. Click “Browse” to tell PEAKS the location of the Shimadzu run2xml.exe file.




3.7. Varian

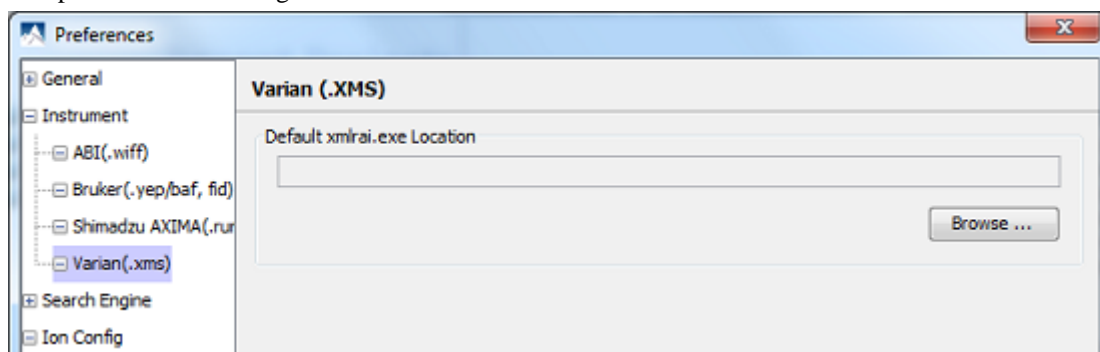
A conversion tool is embedded into Varian’s data acquisition software which allows the conversion of Varian raw data into .pkl files which can be immediately read by PEAKS.

The .trans type data (raw) is converted in Varian programs by clicking "File", "Save As" and selecting the .pkl file format or by clicking "File", right clicking "Export" and selecting ".pkl". If you are viewing a chromatogram with

the Varian software, all the spectral data in the viewed chromatogram is converted to the .pkl format. If you are viewing a single spectrum and choose to convert the data, only the viewed spectra will be converted.

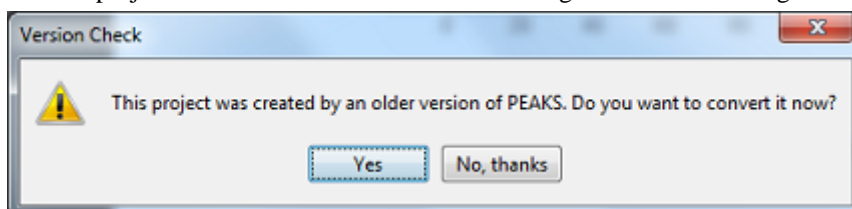
Importing raw data that has not been preprocessed will produce better results when using the preprocessing options native to PEAKS.

Instrument Preferences for Varian Data. To set Varian data related preferences in PEAKS, click the “Preferences” toolbar icon  or select “Preferences” from the “Window” menu to open the “Preferences” window. Clicking on “Instrument” and then “Varian (.xms)” in the menu on the left hand side will display the Varian instrument preferences on the right hand side. Click “Browse” to tell PEAKS the location of the xmlrai.exe file.

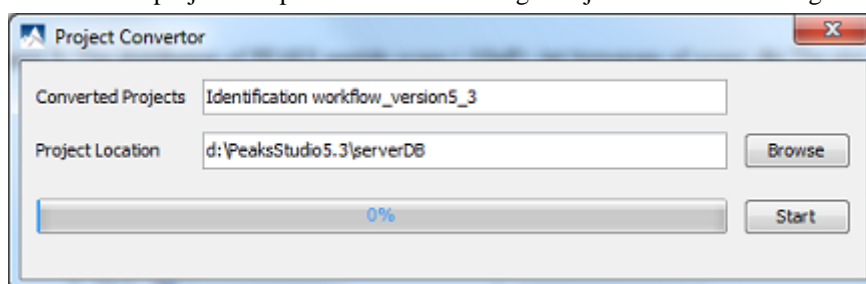


3.8. PEAKS 5.x Projects

Projects created in any PEAKS 5 series software can be opened in PEAKS 5.3. To convert the project to a PEAKS 5.3 project, open the project in the same way you would open any existing PEAKS project. The project will be recognized as a PEAKS project from an older version. The following confirmation dialog box will popup.




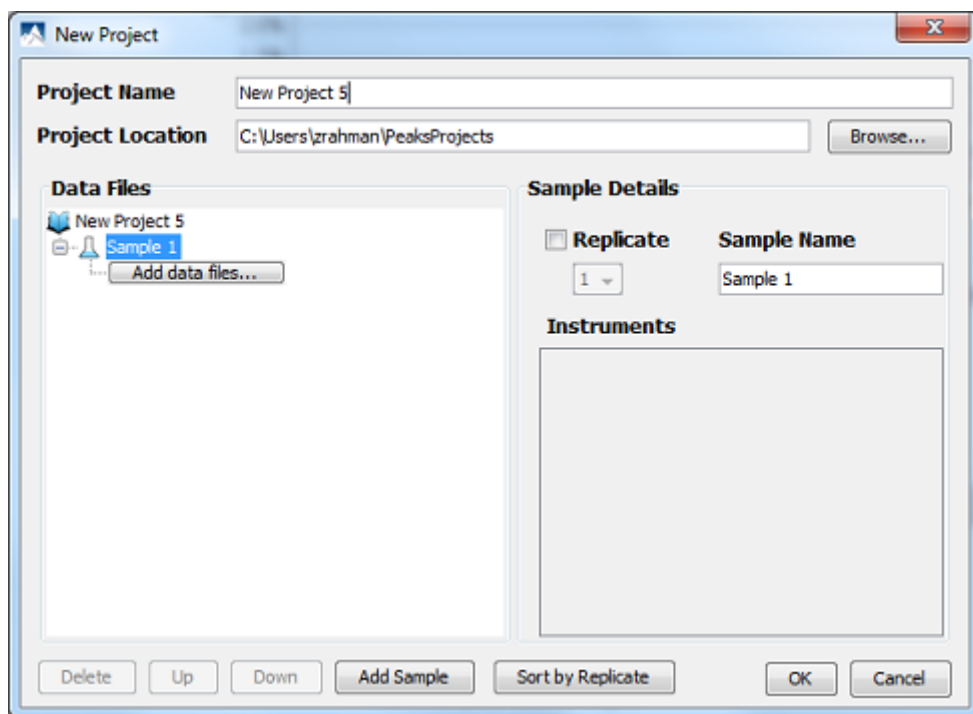
Choose “Yes” to convert the project and proceed. The following “Project Converter” dialog will appear.



Choose the converted project name and location. Click “Start” to begin the conversion process. A new version of the project will be created at the new location. The old project is not altered.

4. Creating a New Project

1. To create a new project, select “New Project” from the file menu or using the new project icon  on the toolbar. The “New Project” dialog appear.





2. Use the “Project Name” field to name your job. Click “Browse” to set the location to save the project in the “Project Location” text box.



Note

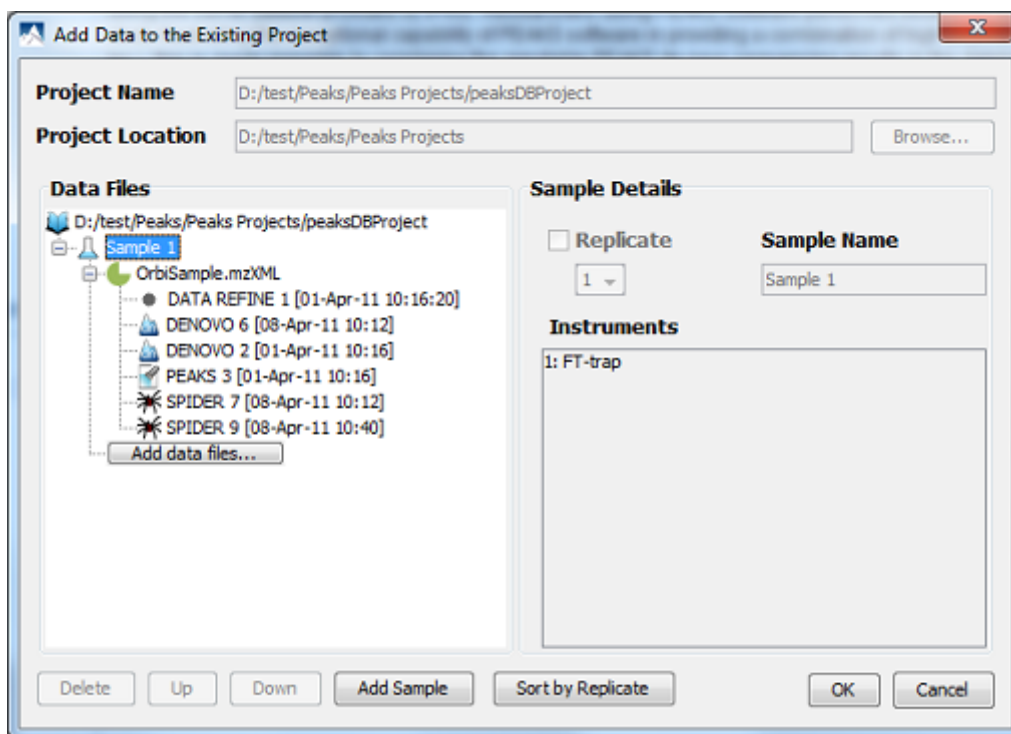
Refer to Section 6, “Changing the Default Project Location” for changing the default save location for projects.

3. Use the “Add data files...” button to browse to the location of the files you wish to load. Select the files you wish to load and click “Open”. Once the data file appears, select the “Instrument Vendor” and “Instrument Type” that was used to generate the experimental data from the drop-down lists. Selecting the “All Instruments” option from the “Instrument Vendor” drop-down list will display some general instrument types in the “Instrument Type” drop-down list. If no fragmentation mode is specified in the instrument type name (in brackets), the default setting is CID. If you would like to apply the same instrument configuration to all of the files in the sample, or to all of the samples in the project click on the “whole sample” button or “whole project” button, respectively.
4. To add another sample, click on the “Add Sample” button. To add a data file to Sample 2 click on the “Add data files...” button. Select the instrument vendor and type from the drop-down menus unless you had previously applied the instrument configuration to the whole project in step 3. These separate samples can be used to get batch results for multiple files in the samples. They can also be used to batch export dta, mgf, or pkl files containing all the data in the sample. Separating into samples is also necessary for label free quantification (refer to Chapter 13, *Label Free Quantification (LFQ)*).
5. To declare a sample as a replicate, click on the sample node and select the replicate check box and set a replicate number using the replicate drop down menu. You can set up to 3 samples to be replicates of the same experiment. Setting replicates allows you to use the “Replicate Analysis” tools (refer to Section 6, “Replicate Analysis in LFQ”).
6. To delete a sample or data file, select the appropriate node (sample or data file) and click the “Delete” button. You can also change the order of the samples within a project or data files within a sample using the “Up” and “Down” buttons.
7. Click the “OK” button once all data files and samples are added to the project.

8. The project will appear in the “Project View” panel. The outlined  symbol indicates that the file is still loading. The solid  symbol indicates that the file has finished loading .

5. Adding Data to an Existing Project

1. To open a saved project, select “Open Project” or “Open Recent Project” command from the file menu or  from the toolbar.
2. To add data to an existing project, choose the project from the “Project View” panel and select “Add Data” command from the file menu or use the add data icon  from the toolbar. The original project window will open.

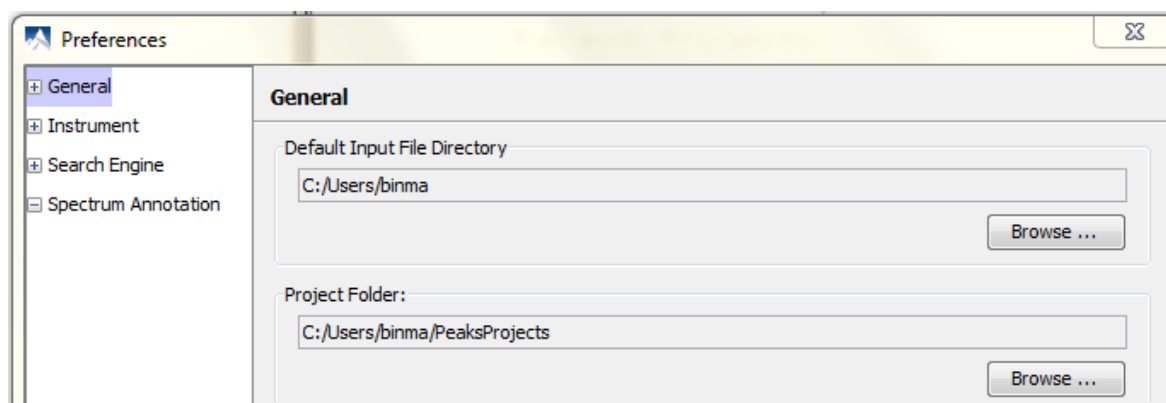


3. You can add more files to an existing sample using “Add data files...” button or create additional samples using “Add Sample” button.
4. You will need to select the instrument vendor type. For more information on adding files/samples or setting up the instrument configuration refer to Section 4, “Creating a New Project”.

6. Changing the Default Project Location

If many projects are to be created, it is convenient to change the default project location to the directory where all the projects are stored.

1. Click  from the toolbar. The following "Preference" dialog pops up.



2. Select "General" in the in the "Preference" dialog and click the "Browse" button below "Project Folder" to specify the default location.

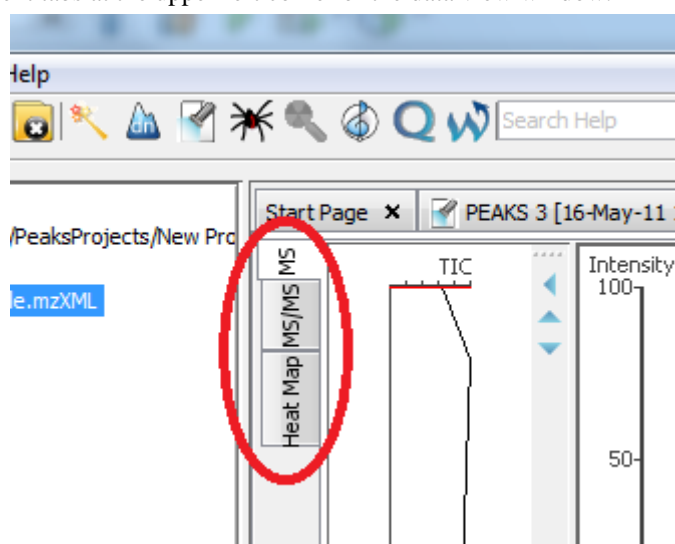
Chapter 5. Data Visualization

1. Overview

After the project is created, the spectral data can be visually examined. For a typical LC-MS/MS fraction, three views are provided:

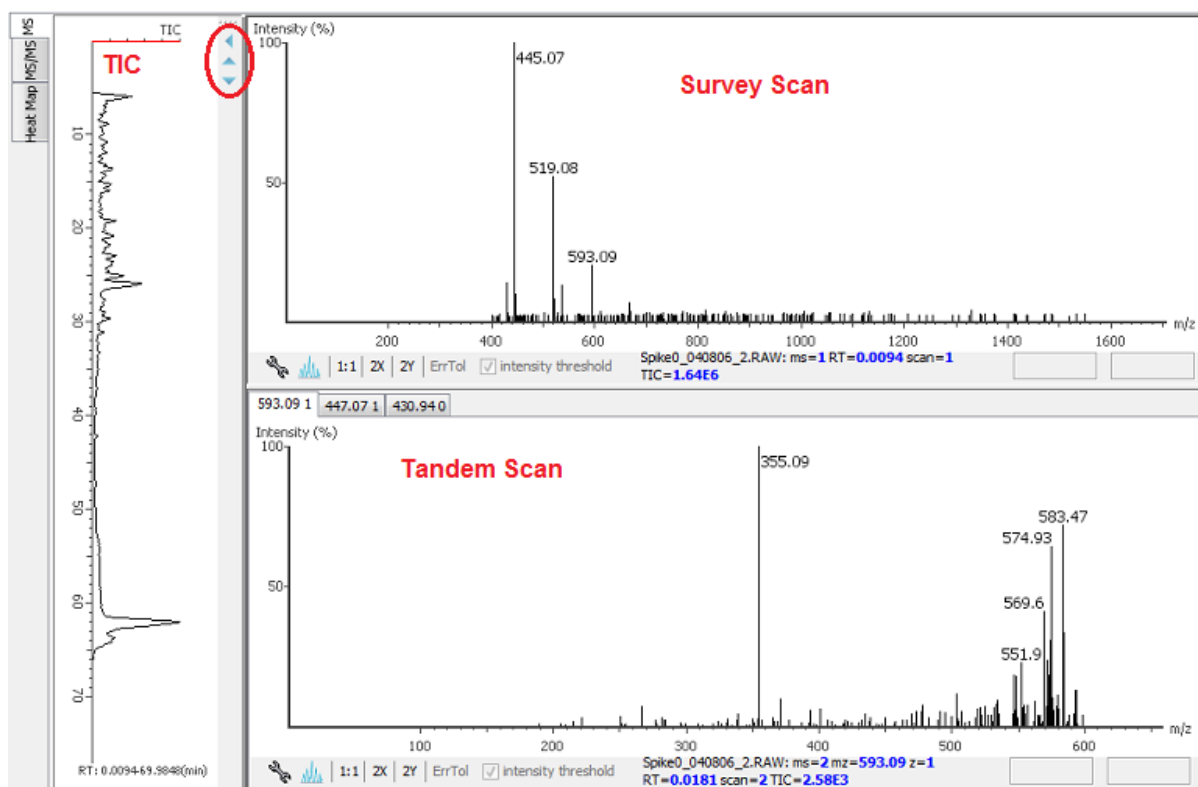
- **MS** - this view shows the TIC (total ion chromatogram) plot and all the MS scans. For each MS scan, the corresponding MS/MS scans are also displayed.
- **MS/MS** - this view lists all the MS/MS scans. For each MS/MS scan, the corresponding MS scan is also displayed.
- **Heatmap** - this provides a bird's eye view of the whole LC-MS dataset.

After opening a data file by double clicking the data node on the project tree, the choice of different views can be made by choosing different tabs at the upper-left corner of the data view window.



2. MS View

The MS View contains TIC and all scans. The total ion chromatogram (TIC) is displayed in the left of MS view. The navigation buttons are circled in the figure. To collapse the TIC chart, click the left navigation button. To navigate the survey scans, use the up and down navigation buttons. Clicking on a specific position in TIC will display the corresponding survey scan. The survey scans can also be navigated by using up and down arrow of the keyboard. The tandem scans associated with the current survey scan are shown in the bottom right panel.

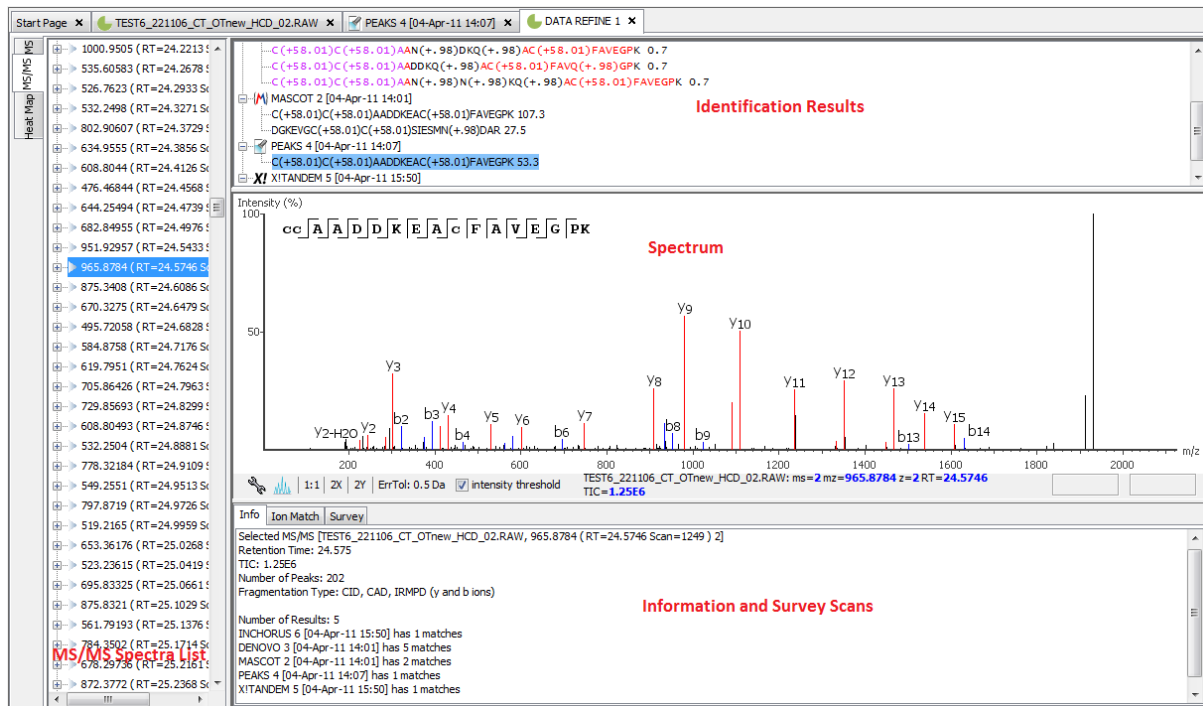


The survey scans and tandem scans provide a few convenient way to zoom and navigate in the spectrum:

- Zoom to an m/z region - click the desired start m/z and drag horizontally to the desired end m/z, release the mouse button.
- Zoom in/out smoothly - place the mouse pointer at a particular m/z value (right below the x-axis line), scroll the mouse wheel button.
- Increase the peak intensity - place the mouse pointer in the spectrum, scroll the mouse wheel button.
- See the whole spectrum - double click in the spectrum, or click the 1:1 button.

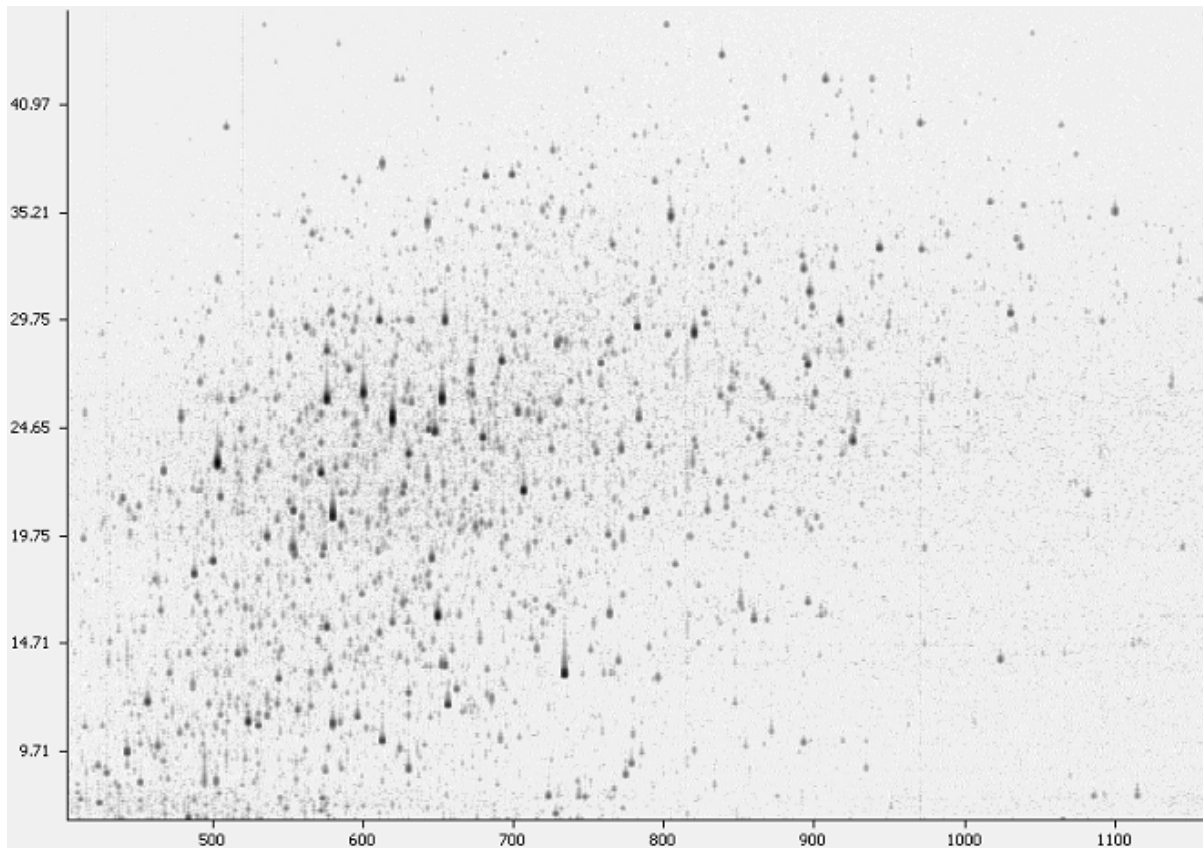
3. MS/MS View

The MS/MS View shows the list of tandem scans on the left. For each MS/MS scan, the list of identification results, the spectrum and its survey scans are shown on the right. Zoom options are the same as described in Section 2.

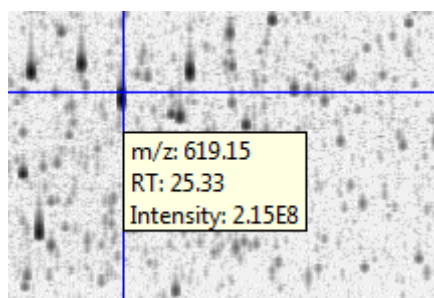


4. Heat Map

Heat Map view shows the distribution of LC-MS signals/features.



Placing the cursor on the heat map will show the m/z value, retention time and intensity of that point in a pop-up window.

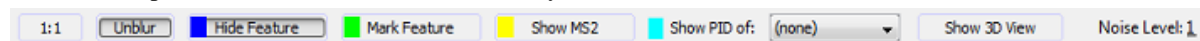


The Heat Map view provides a few convenient ways to zoom and navigate the LC-MS features in the data:

- Zoom to a specific Heat Map area - click the desired start (m/z value, RT) position and drag the cursor to the desired end (m/z value, RT) position, release the mouse button.
- Zoom in/out smoothly - place the mouse pointer at a particular (m/z value, RT) position, scroll the mouse wheel button.
- See the whole Heat Map - click the “1:1” button.

4.1. Blur / Unblur Heat Map

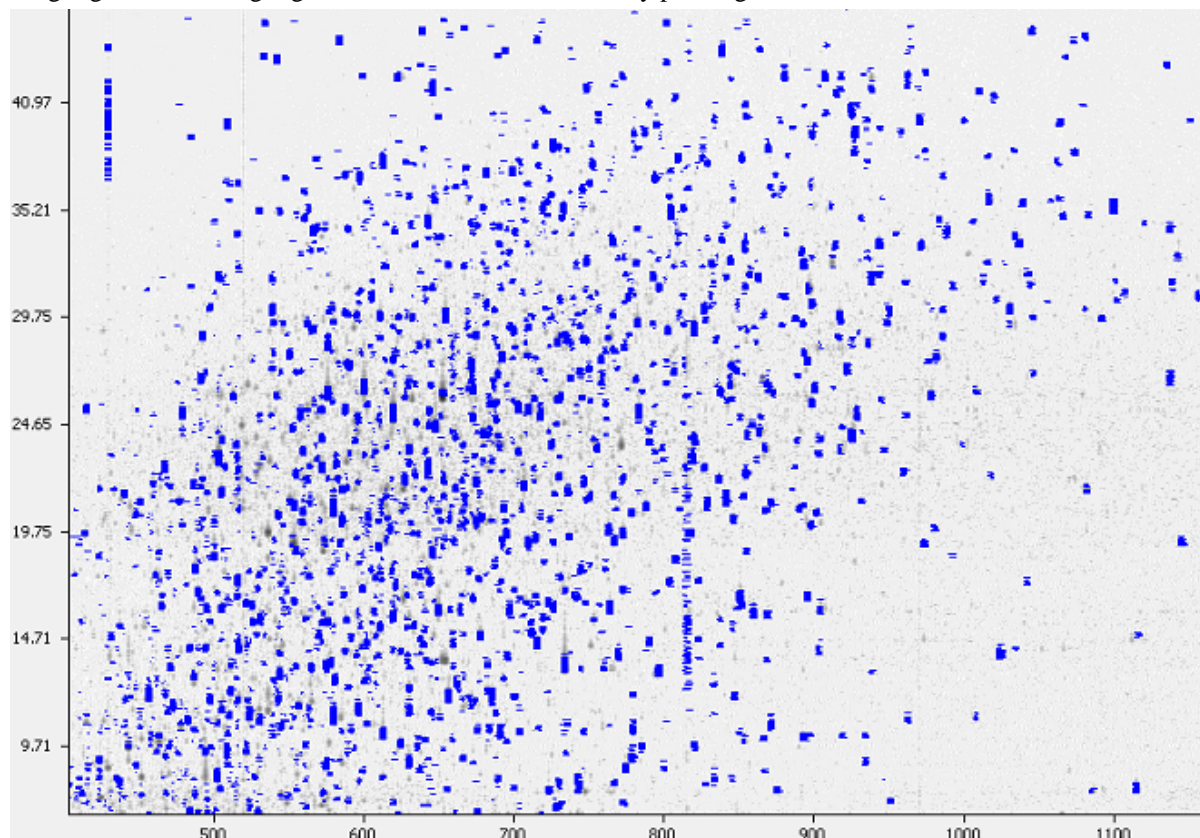
The Heat Map view offers various controls to study the LC-MS data features.



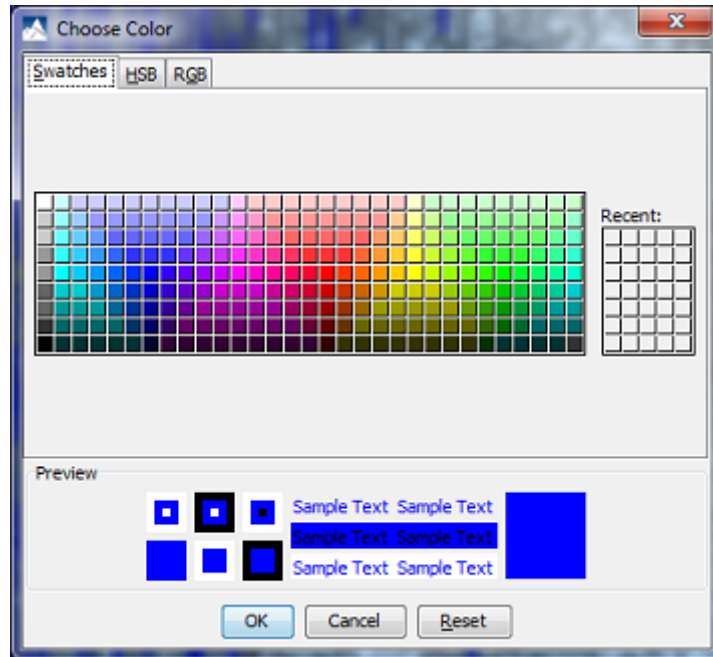
For smoother view of the Heat Map choose “Blur” and for sharp contrasted view choose “Unblur”.

4.2. Highlight Feature / Hide Feature

“Highlight Feature” highlights the identified MS features by painting them with a chosen colour.

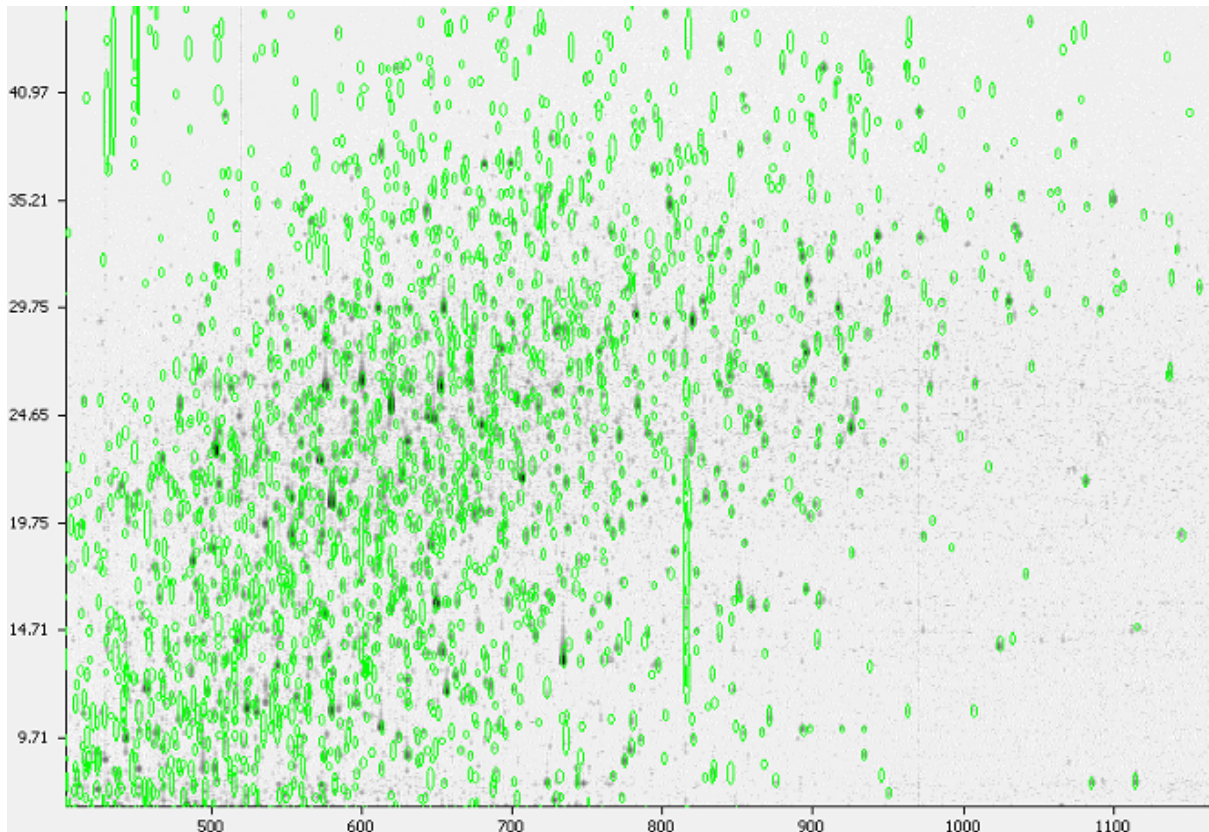


To change the default highlight colour, click on the colour icon of the highlight feature button to display the color palette. Select the preferred color from the color palette.



4.3. Mark Feature / Unmark Feature

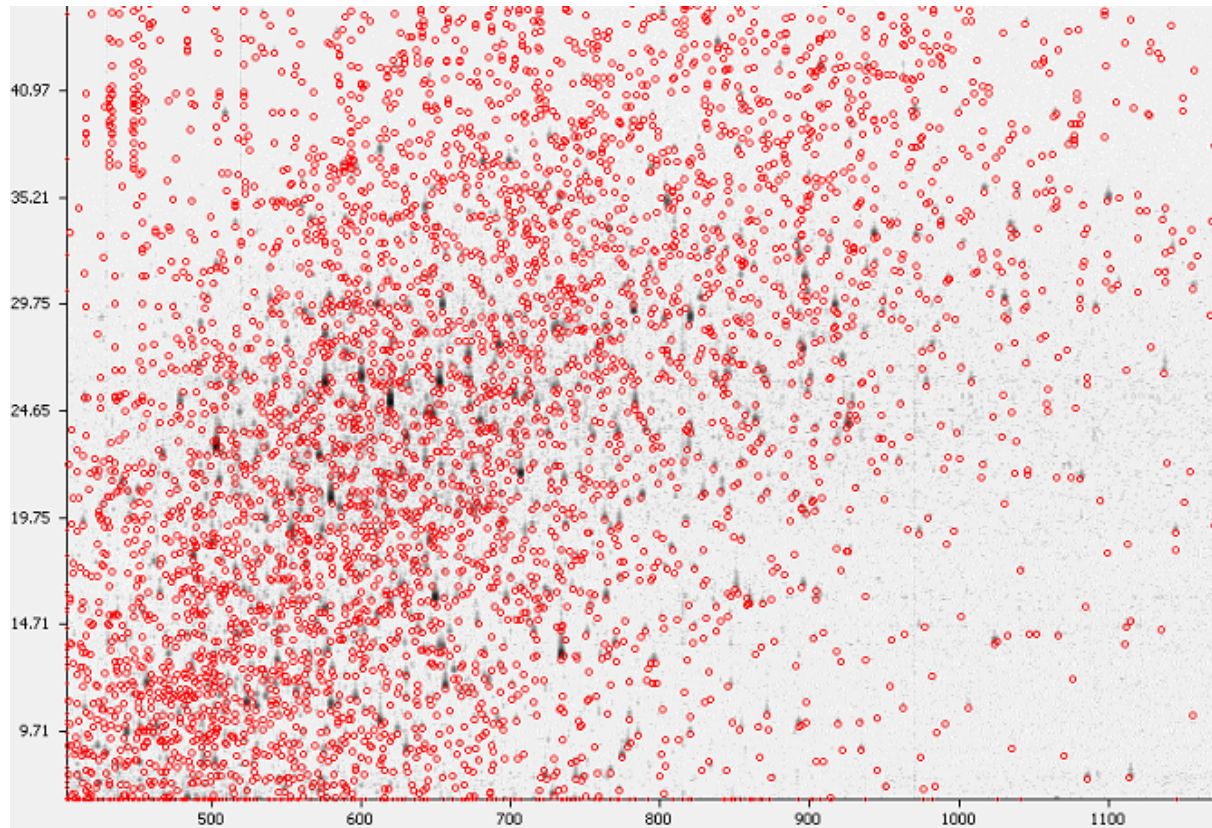
“Mark Feature” marks the identified features by circling around them using a selected colour.



To change the default mark colour, click on the colour icon of the “Mark Feature / unMark Feature” button to display the colour palette. Select the preferred colour from the colour palette.

4.4. Show MS2 / Hide MS2

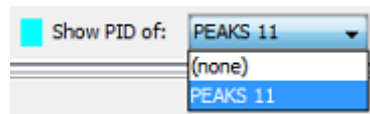
“Show MS2” displays the features associated with tandem Mass scan by marking them with the selected colour.



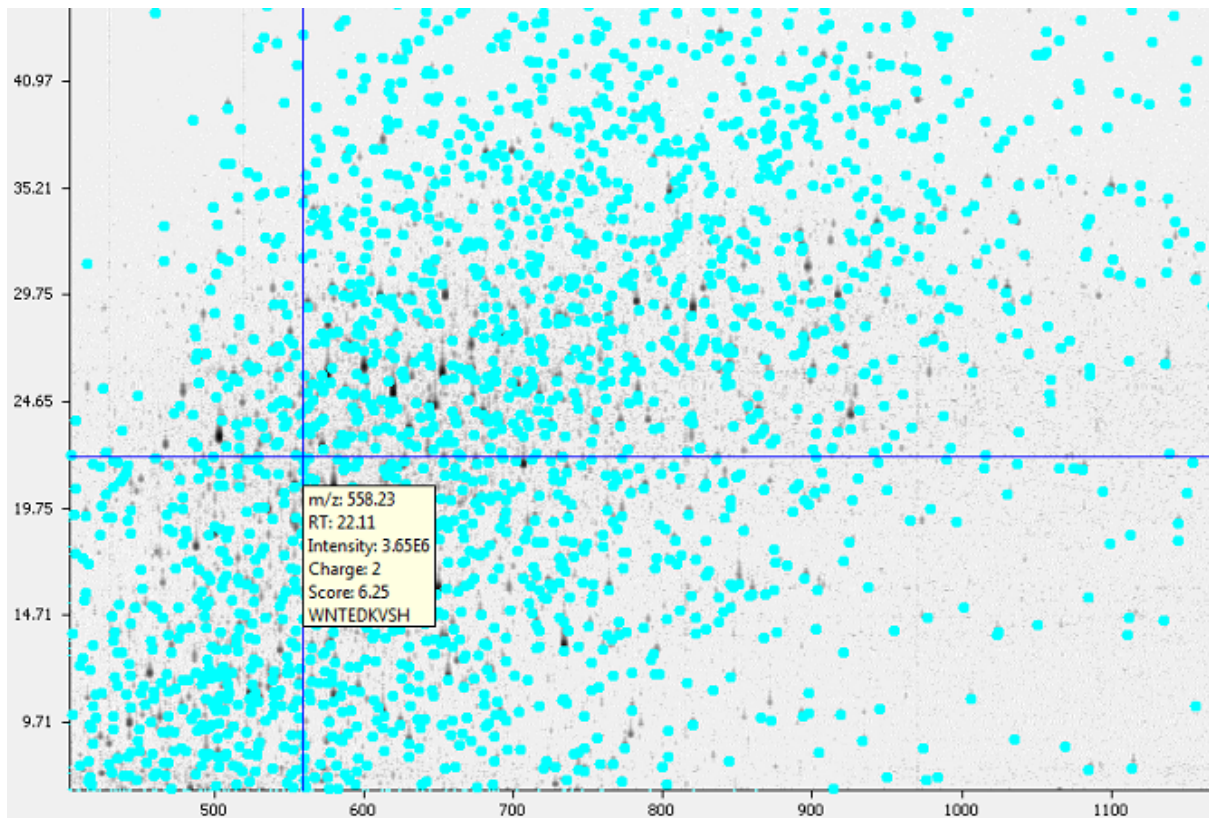
To change the default colour, click on the colour icon of the “Show MS2 / Hide MS2” button to display the colour palette. Select the preferred colour from the colour palette.

4.5. Show PID

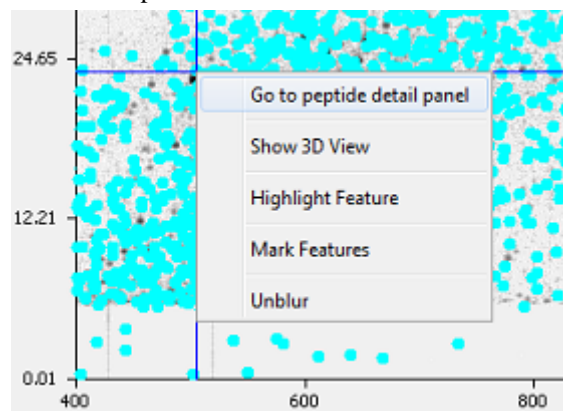
“Show PID” displays the peptide identifications from a PEAKS DB search. Select the PEAKS DB search result from the dropdown list.



The peptide identifications are marked with the selected colour. Placing the cursor on a marked peptide displays more information on the identified peptide in a pop-up window.



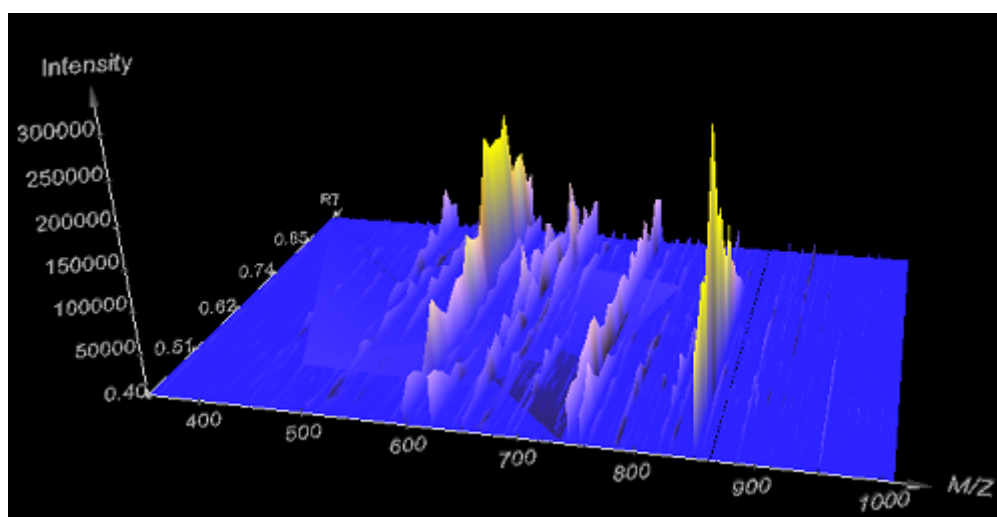
To change the default colour, click on the colour icon of the “Show PID” button to display the colour palette. Select the preferred colour from the color palette.



To view the peptide details of a peptide, place the cursor on a marked peptide, right click to display a pop-up menu and select the command “Go to peptide detail panel”. This will show the peptide details in the “MS/MS View” panel. (see Section 3, “MS/MS View”).

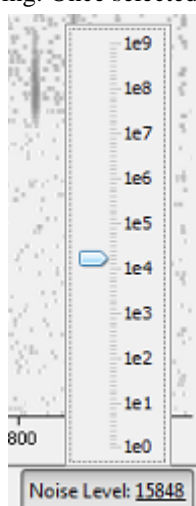
4.6. Show 3D View

The selected Heat Map area will be displayed in a 3D view.



4.7. Noise Level

Select the appropriate threshold for noise filtering. Once selected the Heat Map view will reflect the changes.

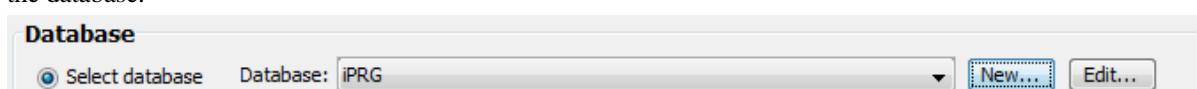



Chapter 6. Adding a Sequence Database

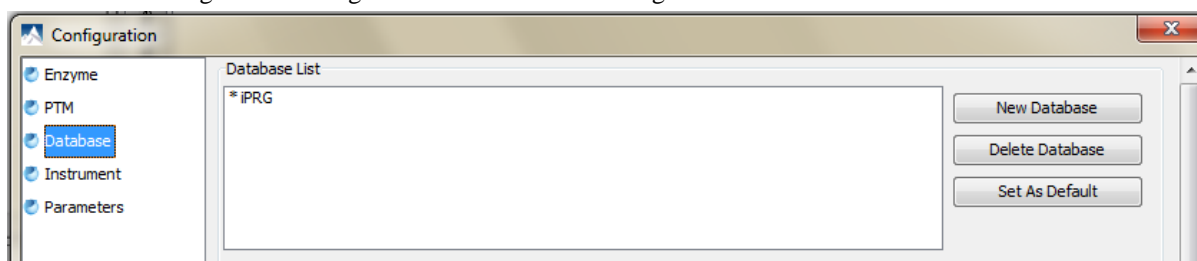
1. Configuring Sequence Databases

In addition to *de novo* sequencing of peptides, PEAKS 5.3 has the ability to search through a database to identify proteins. In order to use this function, PEAKS must have access to a protein or EST database in FASTA format (the standard format for popular public sequence databases). PEAKS can be configured to use existing databases on the system or download from servers. Additionally, taxonomy may be specified with certain databases.

Adding or editing a sequence database can be done by - clicking the "New" and "Edit" buttons when specifying the parameters for a database search analysis. A database configuration dialog will pop-up for you to configure the database.



Alternatively, clicking the  icon in the main toolbar and selecting "Database" from the left-hand side will open the database configuration dialog. The database can be configured in the area below the "Database List".



In both cases, the database configuration parameters appear as following:

FASTA format database: NCBI nr Validate Database

Basic Options

Database name: ☐ Validated

Path: Browse or Download Database

☐ EST database

Advanced Options - Fasta Title Format

Rule to parse accession/id from FASTA title:

Rule to parse description from FASTA title:

Accession/id URL:
 <Accession/ID>

Delimiter:

Taxonomy Options

taxonid Browse Download

taxdmp Browse Download

Follow these steps to fill up or update the required fields to configure a database:

1. Select the database from the “FASTA Format Database” drop-down list, or select “Other” if the desired format is not present and a custom one is to be defined.
2. If the database FASTA file is already on the local system, skip to step 6. In the “Basic Options” panel, enter a name for the database and select “Download Database”. A window will appear confirming the database chosen to be downloaded from the appropriate FTP or website.
3. Click “Yes” to invoke the default FTP client software and download the database automatically. Click “No” to copy the URL to the system clipboard. If “No” was selected, click “Ok” on the dialog detailing the copy to the clipboard. Next, open a browser and paste the URL into the address bar. When the file download window opens, click “Save”.
4. Once the database has been downloaded, check to see if it is compressed. If so, extract the file using a program such as *WinZip*, or *WinRar*. The desired result is a FASTA format text file (a .fas or a .fasta file).
5. Move the database file into a directory that PEAKS can access.
6. Click “Browse” to inform PEAKS about the location of the database file.
7. If the selected database is an EST database, check the box labeled “EST database”. If not, ensure that it is blank.
8. Based on the selected format from the “FASTA Format Database” list in Step 2, the accession number information and parsing rules for the database headers are automatically entered in the textboxes in the “Advanced Options- Fasta Title Format” panel below.

If “Other” was selected in Step 2, enter the parsing parameters into the corresponding textboxes. Alternatively, if the database format is similar to one of the public databases such as *NCBI-nr*, the parsing rules can be filled up by selecting the similar database from the drop-down list and edited to set the desired parsing rules.
9. If the configuration dialog was invoked from the toolbar, click the “Add/Update” button and then “OK”. If the configuration was invoked when specifying DB search parameters, simply click “OK”.

Note

Apart from starting with a “greater than” symbol, the precise syntax of the FASTA title line varies from database to database. For this reason, PEAKS uses Java Regular Expressions to define how the accession string and the description text should be parsed from the FASTA title line.

Note

Some databases use one entry to represent multiple protein entries. The FASTA headers are concatenated with a delimiter. Since some of these databases use unprintable control codes as delimiters, PEAKS will use the equivalent ASCII decimal code to represent them. For example the NCBI NR database uses CTRL-A as a delimiter so the user should input "1" as the equivalent decimal delimiter.

To be able to do PEAKS DB using a specific taxonomy, corresponding files must be downloaded and then referenced by PEAKS in the “Taxonomy Options” panel.

1. To download the “taxonid” file, click the “Download” button. A window will appear confirming the FTP or website which has been identified as the location of the desired database. To invoke the default FTP client software and download the file automatically, click “Yes”. Click “No” to copy the URL to the system clipboard. If “No” was selected, click “OK” on the dialog detailing the copy to the clipboard. Next, open a browser and paste the URL into the address bar. When the file download window opens, click “Save”. Be sure to save the file to a location that is accessible by PEAKS. Please note that it is not necessary to decompress the taxonomy files.
2. Now that the taxonomy files have been downloaded, PEAKS must be given access to them by clicking the “Browse” button and selecting the file.

2. Databases to be Used in PEAKS' inChorus Function

The database configured here will also be used in PEAKS' inChorus function to call the X!Tandem and OMSSA search engines. However, Mascot search depends on Mascot's databases only. When using these third-party software tools, note the following with care:

- **X!Tandem** - At the time of writing, X!Tandem has difficulty in searching through large databases and may crash. It is therefore suggested that X!Tandem only be used with small databases. If used with a large database, a taxon should be specified. The *NCBInr* and *SwissProt* databases are ideal for this purpose.
- **OMSSA** - At the time of writing, OMSSA can not be used with databases that are not in NCBI, or SwissProt format, and have those results available to inChorus. Also, a bug in OMSSA prevents easy use of databases with OMSSA when they are stored in a folder that contains a space in its path. This creates problems when PEAKS creates temporary databases on your behalf. To avoid this, best practices suggest that all our databases are put in a folder `C:\peaksdatabases`. Note that the folder `C:\My Documents\databases` wouldn't work as it contains a space between 'My' and 'Documents'. Using spaces in the database file name causes the same problem. Once databases have been downloaded and extracted, save the database file as `ncbinr.fas`, or `ncbi_nr.fas` rather than `ncbi nr.fas`.
- **Mascot** - The database used by Mascot will have to be identical to the database configured in PEAKS in order for inChorus to parse Mascot results correctly.

Part II. Basic Data Analyses

Table of Contents

7. Data Refinement	45
1. Overview	45
2. Data Refinement Parameters	45
8. Peptide <i>De Novo</i> Sequencing	47
1. Overview	47
2. <i>De Novo</i> Sequencing Parameters	48
2.1. Mass Error Tolerance	48
2.2. Enzyme Specificity	49
2.3. Fixed and Variable PTMs	49
2.4. Other Parameters	50
2.5. Saving the Parameters for Future Use	50
3. Understanding PEAKS <i>De Novo</i> Sequencing Result	50
3.1. Summary View	51
3.2. <i>De Novo</i> Peptide View	52
3.2.1. Peptide Table	52
3.2.2. Spectrum Annotation	53
3.2.3. Ion Table	55
3.2.4. Error Map	56
3.2.5. Spectrum Alignment	56
3.2.6. Parent Scan	56
4. Filtering <i>De Novo</i> Sequencing Result	56
5. Export <i>De Novo</i> Results	57
6. Run Auto <i>De Novo</i> Sequencing on a Single Spectrum	57
7. Manual <i>De Novo</i> Sequencing	57
7.1. Manual <i>De Novo</i> Graphical User Interface	57
7.2. Manual <i>De Novo</i> Operations	59
9. PEAKS DB	65
1. Overview	65
2. Set PEAKS DB Parameters	66
3. Understanding PEAKS DB Result	67
3.1. The Peptide and Protein Scores	67
3.2. Summary View	67
3.3. Peptide View	70
3.3.1. Peptide Table	71
3.3.2. Peptide Spectrum Match	71
3.3.3. Protein	73
3.4. Protein View	73
3.4.1. Protein Table	74
3.4.2. Peptides	74
3.4.3. Coverage	75
3.4.4. Tool Box	75
3.5. <i>De Novo</i> Only View	75
4. Filter PEAKS DB Result	75
5. Export PEAKS DB Results for Publication	57
6. Comparison of PEAKS DB Results	77
6.1. Comparison Result	77
6.2. Peptide Comparison	78
6.3. Protein Comparison	79
6.4. Statistical Charts	80
6.5. Exporting Comparison Results	80
10. PEAKS InChorus	82
1. PEAKS inChorus Overview	82
2. Understanding PEAKS inChorus Result	83
3. Filter PEAKS inChorus Result	85


Chapter 7. Data Refinement

1. Overview

The raw LC-MS/MS data often contains noise, redundancy, as well as errors due to sample preparation and instrument approximation. PEAKS Data Refinement tool can be used to improve the overall quality of the data. Upon users' choice, all or some of the following functions can be applied to the data in a project. A correct data refinement, especially the precursor m/z correction, can often result in significant improvement in the final analysis result.

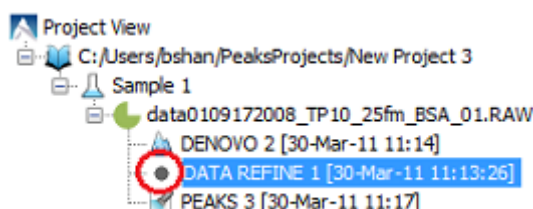
- Merging scans - the redundant MS/MS scans from the same precursor m/z and similar retention time will be merged together.
- Precursor m/z correction - the precursor m/z value given by some instruments is often not of the monoisotopic ion. This creates problem in the downstream analysis. By examining the isotope shapes in the corresponding MS scans, this function can accurately correct the precursor m/z to be monoisotopic.
- Precursor charge correction - occasionally the data provides wrong or no charge information for the precursor ions. This function makes an attempt to correct the charge information.
- Low quality spectrum removal - this function attempts to remove the "apparently junk" spectra. This will save some analysis time. Use this function with caution as it may also remove a small portion of identifiable spectra.
- Centroiding and charge deconvolution - centroiding the peaks and deconvolute the multiple charge ions to singly charged in the MS/MS scans.

If the data is not refined within PEAKS, most analysis functions (such as *de novo* sequencing or PEAKS DB) will ask you to input the refinement parameters before the analysis is done. However, you can run the data refinement function separately by selecting fraction, sample or project on the project tree and all the fraction(s) under the selected node will be refined. The use of this function is outlined in the following.


1. Select a project, sample or a fraction node. Click the "Data Refinement" button  on the tool bar.

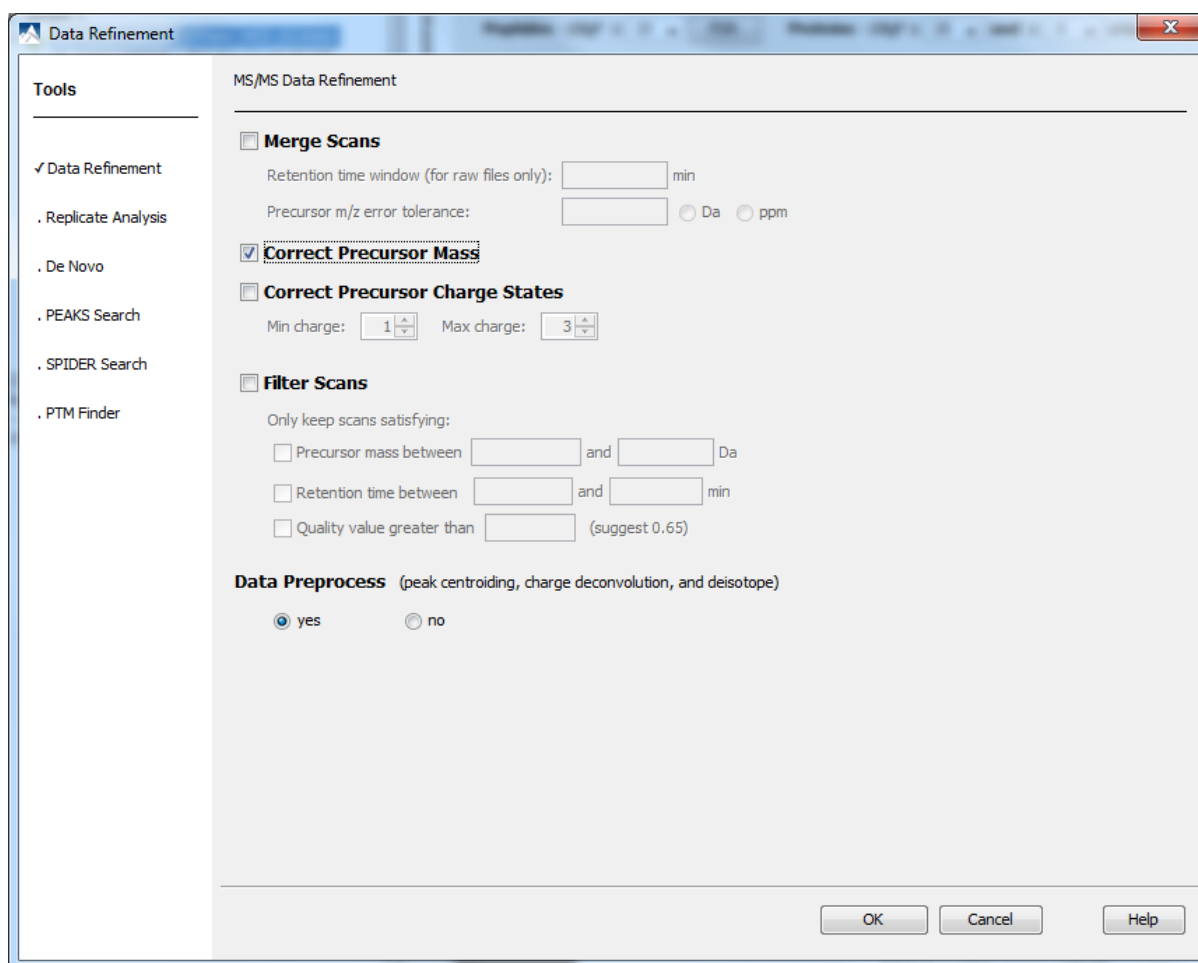


2. Specify the Data Refinement parameters in the popup dialog and click "OK". Most of the parameters are self-explanatory and the default parameters provide a good starting point for the analysis.
3. Wait for the analysis is done. A new Data Refinement node will appear at the project tree. Later analysis on this fraction will be based on the refined data.



2. Data Refinement Parameters

After selecting a data node in the project tree, click the data refinement toolbar icon . The "Data Refinement Parameters" dialogue window will appear:



- **Retention Time Window:** It defines the maximum difference of retention time between two spectra to be merged.
- **Precursor m/z Error Tolerance:** The tolerance of the difference in m/z between two spectra to be merged.
- **Min Charge:** It defines the minimum charge of a precursor ion.
- **Max Charge:** It defines the maximum charge of a precursor ion.
- **Precursor Mass Range:** It defines the precursor mass region to select scans for further analysis.
- **Retention Time Range:** It defines the retention time region to select scans for further analysis.
- **Quality Threshold:** It defines the spectrum quality threshold to select scans for further analysis. The recommended value is 0.65.
- **Data Preprocess:** If the data is already pre-processed, select “no, already done”. Otherwise, select “yes” to pre-process the data for all further analysis. Or, select “no” to do on-the-fly pre-processing for further analysis.

Once all parameters are set, press the “OK” button to initiate data refinement process.


Chapter 8. Peptide *De Novo* Sequencing

1. Overview

De novo sequencing is not only a preferred method for identifying peptide sequences yet to be included in databases, but also a proven method to measure alongside database findings. PEAKS is the most adopted tool for *de novo* sequencing in mass spectrometry labs. The PEAKS automated *de novo* sequencing can process over 10 spectra per second on a moderate desktop PC. Moreover, users can use the manual *de novo* sequencing tool to assist the manual interpretation of an individual spectrum. Most importantly, the automated *de novo* sequencing results are used in several other analyses in PEAKS, including PEAKS DB for database search and SPIDER for homology search. For example, in PEAKS DB, the "de novo only" table lists the highly-confident *de novo* sequences for the spectra that do not match any database peptides. This provides a convenient way to identify novel or mutated peptides in your sample.

PEAKS assigns a *local confidence* score for each amino acid in the *de novo* sequence. This local confidence ranges from 0% to 99%, indicating how confident the algorithm is about the particular amino acid. The whole peptide is evaluated by two measures: the ALC (Average of Local Confidence) and TLC (Total of Local Confidence) scores. Roughly speaking, ALC reflects the average correct ratio for the amino acids in the sequence, and TLC reflects the expected total number of correct amino acids in the sequence.

The use of automated *de novo* sequencing is outlined in the following. Details of each step can be found in later sections of this chapter.

1. Select a project, a sample, or a fraction on the project tree. Click the automatic *de novo* toolbar icon  or select "De novo" from the "Tools" menu.



Note

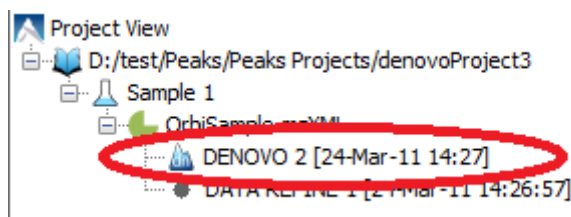
Refer to Chapter 4, *Loading Data to a PEAKS Project* for how to create a project.

2. Specify the PEAKS *de novo* parameters in the *de novo* parameters dialog and click "OK". If your data is not refined yet, you also need to specify the data refinement parameters first and click next. Most of the parameters are self-explanatory and the default parameters provide a good starting point for the analysis.

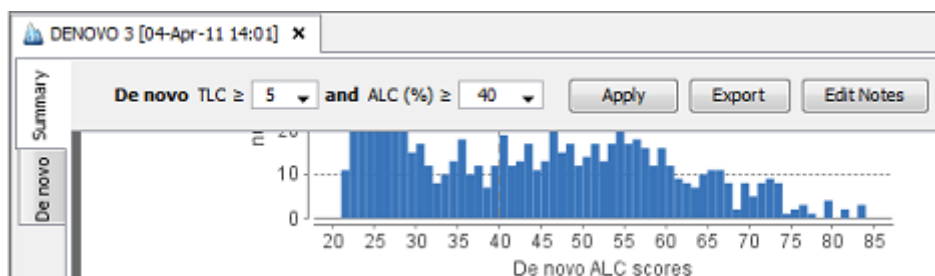
Note

Refer to Chapter 7, *Data Refinement* for PEAKS' data refinement function.

3. Wait for the analysis to finish. A new *de novo* result node will appear in the project tree. Double click the node to open the result file.




4. The result contains two different views: Summary and *De novo*. The summary view allows you to specify rules to filter the results, and provides statistics of the results. The *de novo* view shows the *de novo* sequencing results in greater details.



5. You can optionally export the results to other formats by using the summary view.

2. *De Novo* Sequencing Parameters

In the “Project View Frame”, select the data file(s) or project containing the spectra that you wish to sequence by auto *de novo*. Note that users can run *de novo* sequencing on a fraction or sample level by selecting the fraction node or sample node respectively.

Click the automatic *de novo* toolbar icon  or select “*De novo*” from the “Tools” menu. The auto *de novo* parameters dialogue window will appear.

The meaning of each parameter is discussed in the following sections.

2.1. Mass Error Tolerance

Parent mass error tolerance: The parent/precursor ion mass errors that PEAKS will allow for during the analysis.

Fragment mass error tolerance: The error tolerance for the peaks in the MS/MS spectrum.

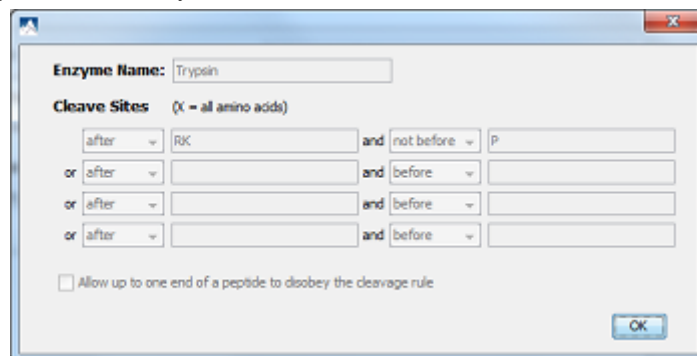
2.2. Enzyme Specificity

This informs PEAKS as to what type of enzyme was used to digest the sample. Choose from a drop-down list of enzymes.

Note

There are several "semi" versions of common enzymes in the dropdown list. For example, "semi trypsin" has the same digestion rule as trypsin, except that "semi trypsin" allows one termini to violate the digestion rule. These semi versions are recommended due to some degree of non-specificity of the digestion enzymes.

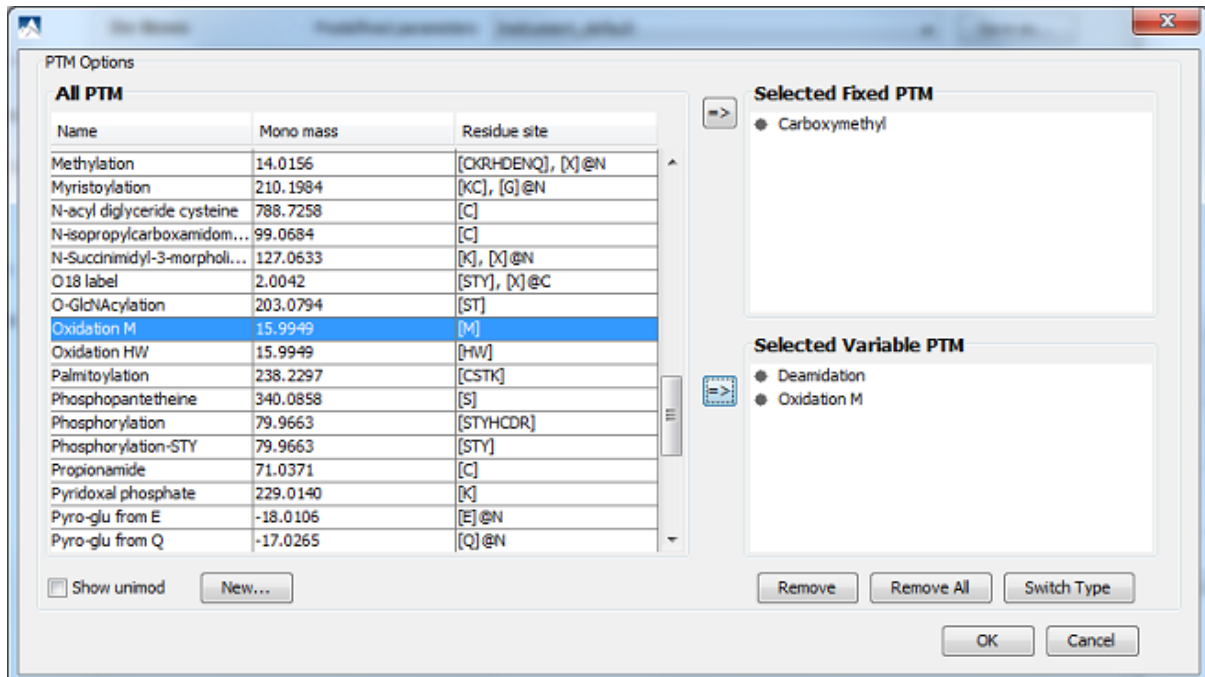
If your enzyme (or combination of enzymes) is not in the list, click the "New Enzymes..." button to define the enzyme used in the experiment in "Enzyme Editor" window.



You can provide the name of the new enzyme and define the custom cleavage rules/sites and select if you would like to allow up to one end of a peptide to disobey the cleavage rule.

2.3. Fixed and Variable PTMs

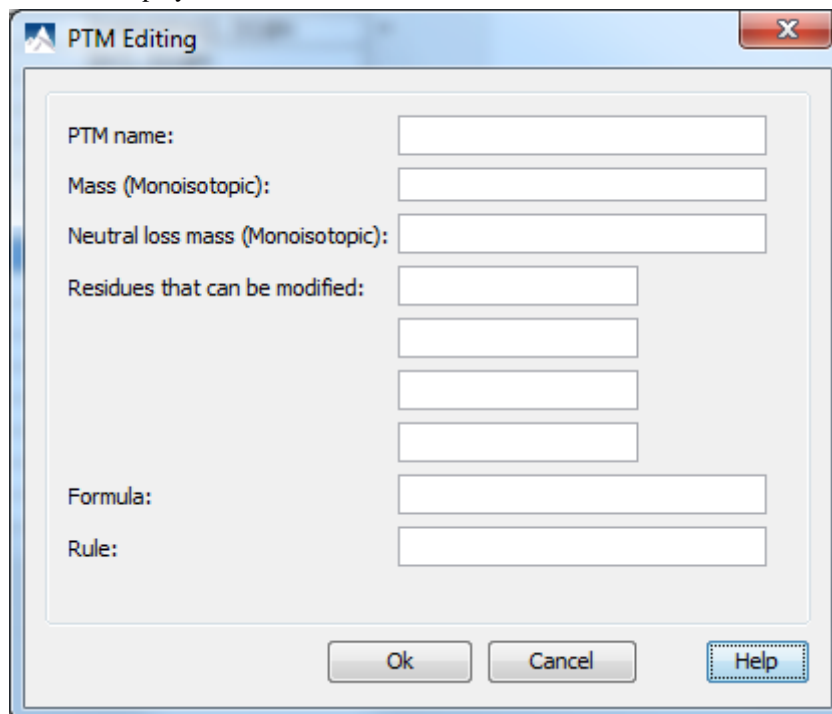
To select the PTMs for the *de novo* sequencing, click the "Set PTM..." button to open the "PTM Setup" window.



The "PTM Options" list displays all available PTMs. To view additional modifications, select the "Show unimod" checkbox. To select a PTM as *Fixed* or *Variable*, click the PTM from the "All PTM" list and click the arrow beside the "Selected Fixed PTM" box or the "Selected Variable PTM" box respectively. To remove a wrongly selected

PTM, click the PTM from the “Selected Fixed PTM” or “Selected Variable PTM” lists and press the “Remove” button. The “Switch Type” button can switch a selected PTM between fixed and variable.

If a desired PTM does not appear on the list or is different than what is listed, select the “New” button and the “PTM Editing” window will open where you can fill in the information pertaining to your particular PTM. The newly edited PTM will be displayed in the “All PTM” list.

The image shows a software window titled "PTM Editing". It contains several input fields: "PTM name:", "Mass (Monoisotopic):", "Neutral loss mass (Monoisotopic):", "Residues that can be modified:" (with four stacked text boxes), "Formula:", and "Rule:". At the bottom, there are three buttons: "Ok", "Cancel", and "Help".

Maximum Number of Variable PTMs per Peptide. To reduce uncertainty, limit PEAKS' *de novo* sequencing 'vocabulary' by restricting the number of variable PTMs found on a peptide. Specify a number by typing it in the textbox beside “Maximum allowed variable PTM per peptide”.

2.4. Other Parameters

Report up to (#peptides): Set how many peptide sequences PEAKS will report from its *de novo* sequencing analysis per spectrum.

Preprocess this data “on the fly”: PEAKS has its own built-in preprocessor for removing noise, centroiding and deconvolution. Check this box to turn preprocessing on, only if you have not done the Data Refinement Step.

Note

If you have already pre-processed your data in the data refinement step, you do not need to do this again.

2.5. Saving the Parameters for Future Use

After setting up the desired parameters, you can save them for future use. Click the “Save as...” button at the top of the window, and define a name for these preferences for future use/reference when prompted. Any parameters that are saved will be available in the drop-down list at the top of the window. To examine the contents, select a saved parameters file and the parameter values will be automatically updated and displayed.

3. Understanding PEAKS *De Novo* Sequencing Result

Once *de novo* sequencing is completed, a new *de novo* result node will appear at the project tree. Double click the node to open the result file. The following results will be available to view:

3.1. Summary View

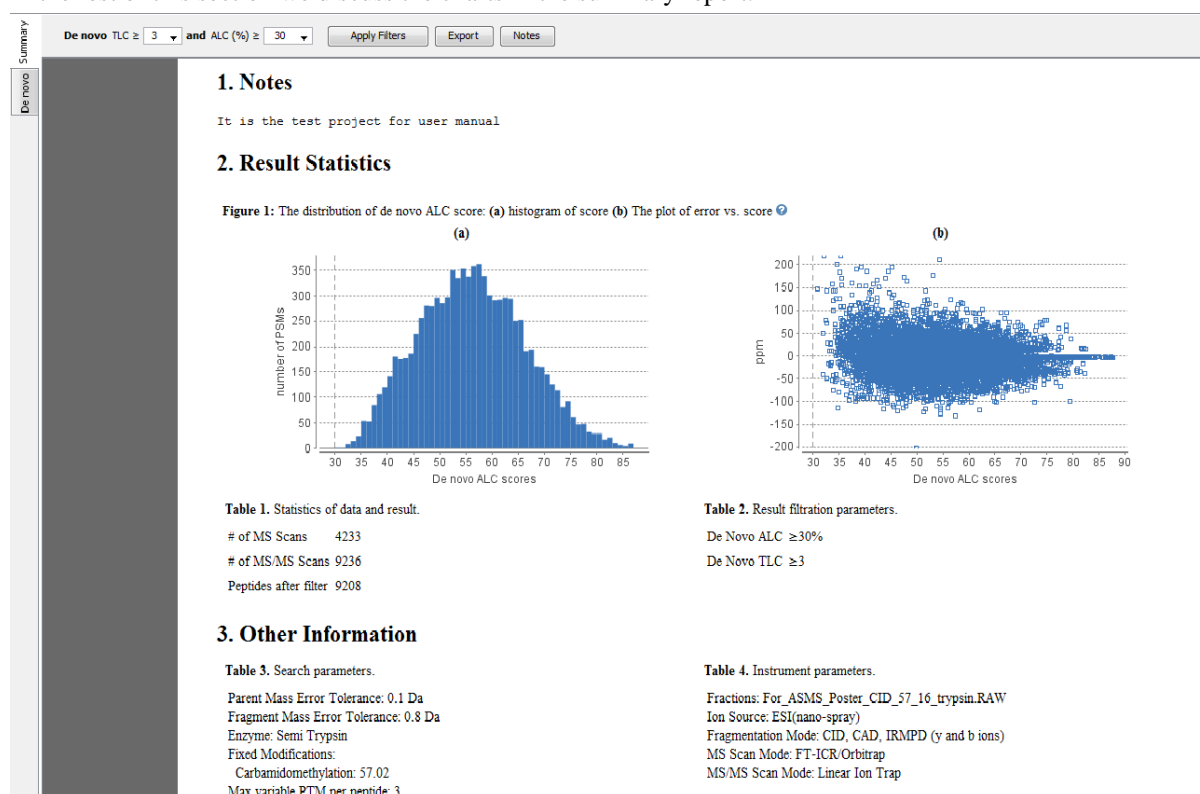
The summary view provides three main functions:

1. Result filtration - this is achieved by specifying the filtration rules in the area at the top of the summary view. The filtration function is discussed in a separate section, Section 4, "Filtering *De Novo* Sequencing Result".
2. Result exporting - this is achieved by clicking the "Export" button at the top of the summary view. The exporting function is discussed in a separate section, Section 2, "Export *De Novo* Result".
3. Summary report - several statistical charts assist the user to get an overall picture of the results and assess the result quality. This function is the focus of this section.

The charts in the report are divided into three sections:

1. Notes - A user can enter a special text note regarding the experiment. Click the "Notes" button at the upper right corner of the summary view to edit the note.
2. Result Statistics - The figures and tables summarize the data and results.
3. Other Information - The search parameters and MS instrument information are given here.

In the rest of this section we discuss the charts in the summary report.

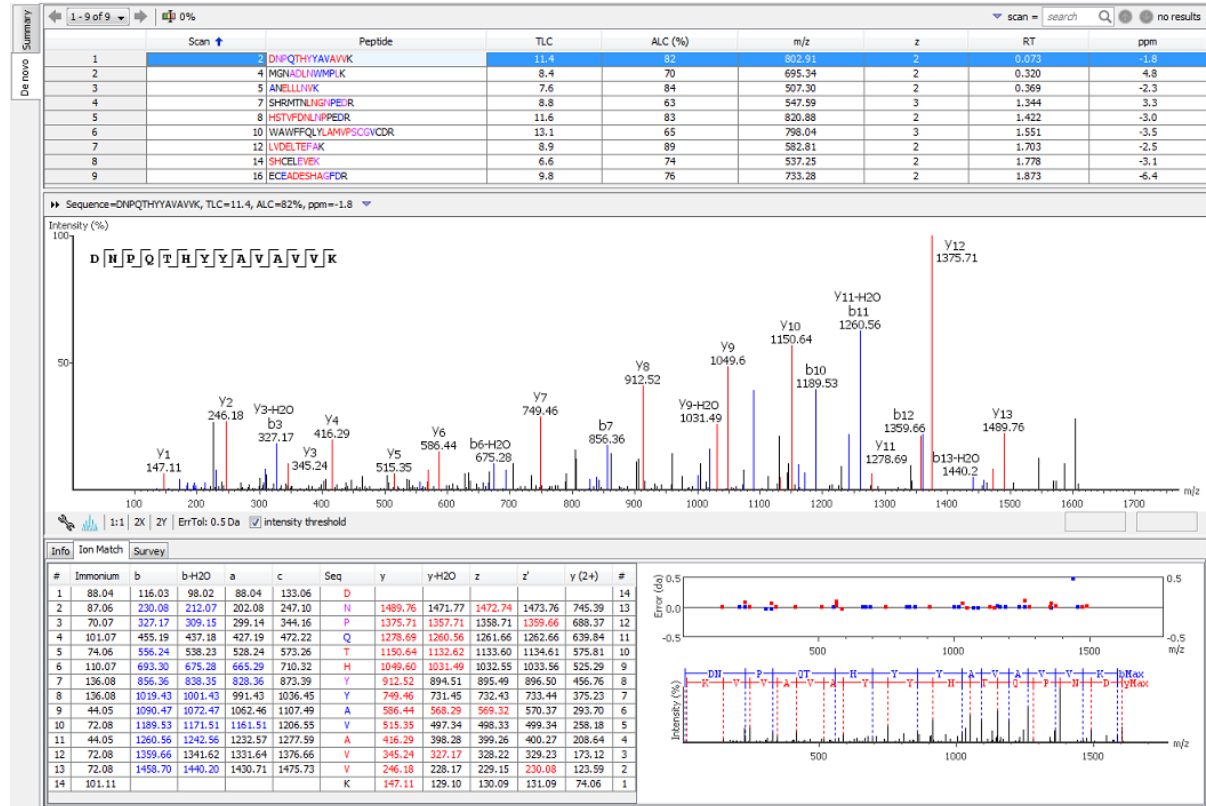


Histogram of Score (ALC). The histogram of ALC scores is a graphical representation showing a visual impression of the distribution of ALC scores of the identified peptides. The peptides are binned in 5% interval of scores.

Mass Error Distribution. Mass errors (in ppm) of the identified peptides are plotted against their ALC scores. The mass error is calculated as a ratio of observed mass error (difference between observed mass and theoretical mass) and the theoretical mass and is expressed in ppm.

3.2. *De Novo* Peptide View

The *de novo* peptide view displays the *de novo* sequencing results in more detail as shown in the next figure. The table at the top displays all the *de novo* sequences, and the bottom half of the view provides additional information about the peptide-spectrum match. The use of this view is explained in the following sections.



3.2.1. Peptide Table


PEAKS displays the peptide sequence candidates at the top of the screen. You can sort the results by clicking on any of the titles of the columns. For example, to sort the peptide sequence candidates by the scan number click on the title bar of the “Scan” column.

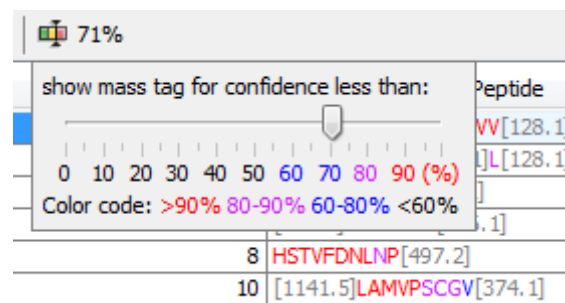
The following list describes the contents of the columns in the “Peptide Candidates Frame”. The first column is a unique index for the peptides in the list.

- **Scan:** Scan number.
- **Peptide:** The amino acid sequence of the peptide as determined by *de novo* sequencing. If there is any PTM on an amino acid, the amino acid is followed by a pair of parentheses enclosing the delta mass of the PTM.
- **TLC:** Total local confidence. It is calculated by adding the local confidence for each amino acid in the peptide sequence.
- **ALC:** Average local confidence (TLC divided by the peptide length).
- **m/z:** The measured mass/charge value, in Daltons, for the spectrum.
- **z:** The calculated charge value for the peptide.
- **RT:** Retention time (elution time) for the spectrum as recorded in the data.
- **ppm:** The precursor mass error, calculated as $10^6 \times (\text{observed mass} - \text{theoretical mass}) / \text{theoretical mass}$.

Confidence Scores. Next to the proposed sequence candidates, the auto *de novo* “Total Local Confidence” (TLC) and “Average Local Confidence” (ALC) confidence scores are shown. The local confidence scores for each amino acid (that is, confidence that the correct residue in each position has been identified) are represented by color coding. Red represents a very high confidence (greater than 90%), purple represents a high confidence (80 to 90%) blue represents a medium confidence (60 to 80%) and black represents a low confidence (less than 60%). For more detailed positional confidence, place the cursor over the sequence of interest and a “Position Confidence Table” will appear, showing the confidence that each amino acid/pair of amino acids are correctly identified.

2	DNPQTHYYAVVK	11.4
4	MGNADLNWMLK	9.4
5	ANELLNVK	6
7	SHRMTNLNGNPEDR	8
8	HSTVFDNLNPPEDR	11.6

Mass Tags. The low confidence residues can be displayed as mass tags by setting up the threshold score using the button  in the title bar of the “Peptide Candidates Frame”. If the score is set at 0.0, all of the amino acids in the peptide sequences will be displayed. Increasing the threshold will display a mass in square brackets if the residues do not satisfy the threshold.



Modifications. Consider the following sequence: SHM(+15.99)TNLNGNPEDR. The (+15.99) in brackets refers to a position where a modification may have occurred. If you forgot the PTMs you specified before running *de novo*, check Table 3 in the summary view.

Search for a Peptide. The peptide candidates can be searched by entering the value in the search tool located in the top right corner of the title bar of the “Peptide Candidates Frame”. The available options for a search are scan number, m/z value and retention time (RT value). The resultant peptide candidates can be iterated by clicking the circled up and down arrow buttons in the search tool.

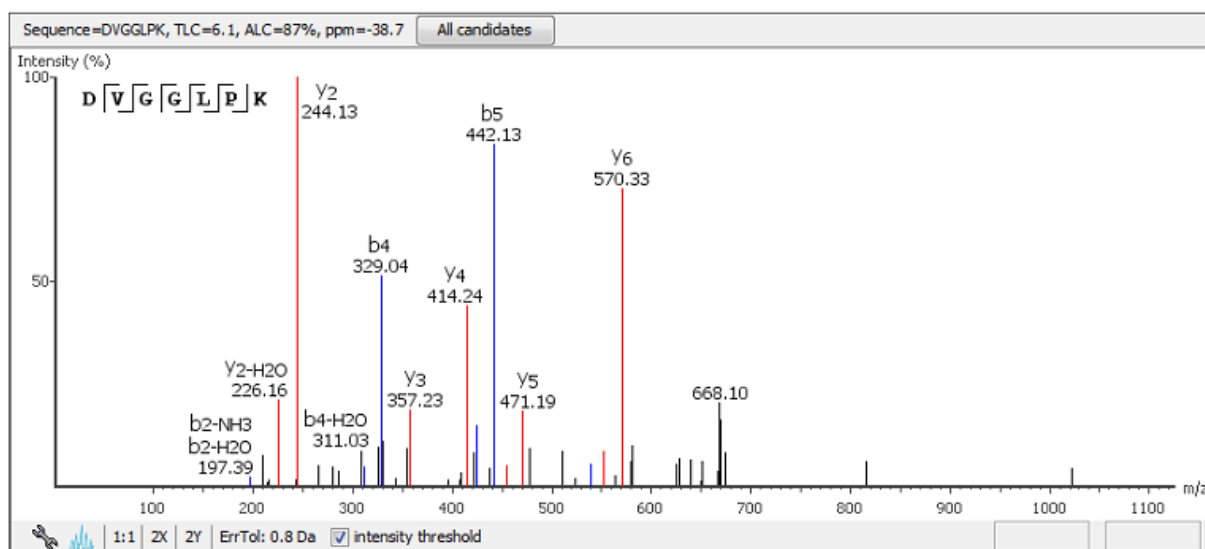
m/z ≈ 733		1 / 1
scan =		ppm
m/z ≈	3	3.1
RT ≈	9	-7.2
	0	8.6

Note

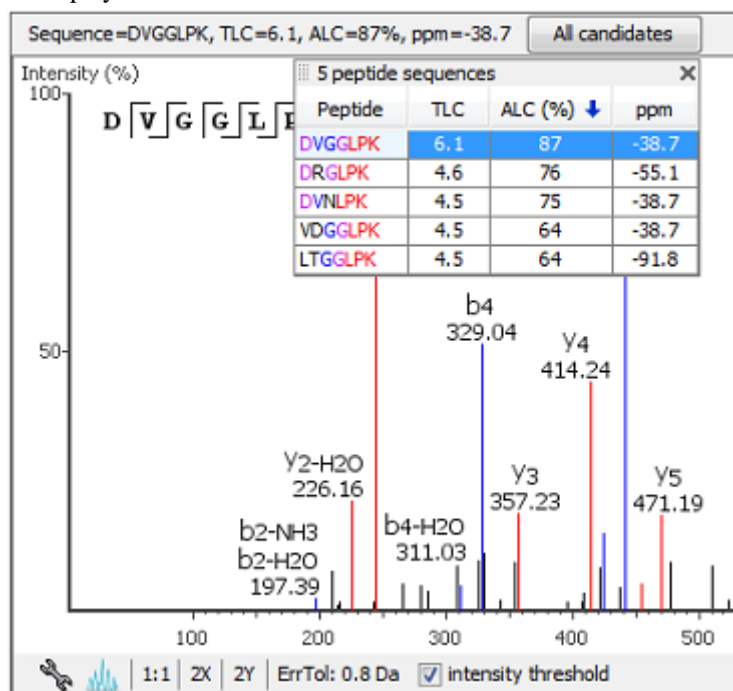
To search with an approximate mass value, type only the necessary number of digits after the decimal points. For example, 130.3 will match any value from 130.25 to 130.35 (exclusive). And 130 will match from 129.5 to 130.5 (exclusive).

3.2.2. Spectrum Annotation

The spectrum annotation displays a graphical representation of the peptide-spectrum spectrum.



The title bar shows the peptide sequence of the spectrum that is being displayed. Press the “All candidates” button in the title bar to open a pop-up window to display all alternative peptides. Click on a peptide sequence in the pop-up window to select and display the annotation.





Moving the mouse over the spectrum will display a “tooltip” to show the annotation, the m/z ratio and the relative height/intensity (as a percentage of 100) of that particular peak. Both the m/z ratio and the height of the peak can also be found at the right hand side of the bottom bar of the spectrum annotation panel.


The annotation provides a few convenient ways to zoom and navigate in the spectrum:

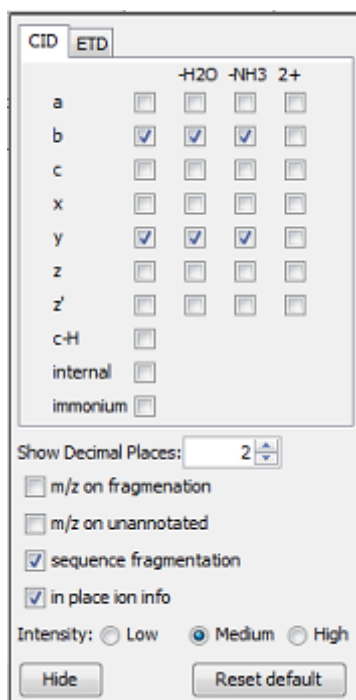
- Zoom to a m/z region - click the desired start m/z and drag horizontally to the desired end m/z, release the mouse button.
- Zoom in/out smoothly - place the mouse pointer at a particular m/z value (right below the x-axis line), scroll the mouse wheel button.
- Increase the peak intensity - place the mouse pointer in the spectrum, scroll the mouse wheel button.
- See the whole spectrum - double click in the spectrum, or click the 1:1 button.

The “ErrTol” is used to adjust the error tolerance to view the display of matched ions.

You can use the profile  and peak  buttons to switch the spectrum view from profile mode to peak mode and vice versa.

The “intensity threshold” check box provides an option to annotate the lower intensity peaks.

You can change the spectrum annotation preferences by clicking the  button to open the “Spectrum Annotation Preferences” window. Refer to Section 1.4, “Spectrum Annotation Preferences” for more details on how to set spectrum annotation preferences.



3.2.3. Ion Table

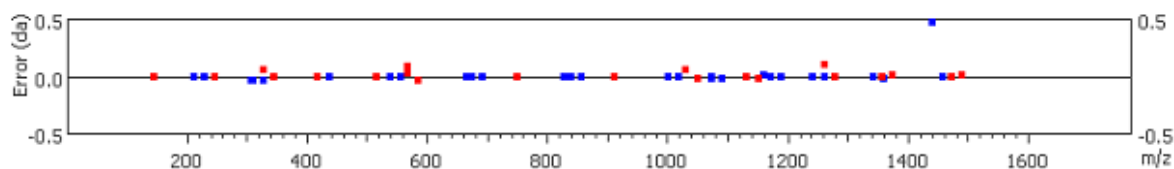
The “Ion Match” tab at the bottom panel of the *de novo* peptide view contains the “Ion Table” that shows the proposed ions with their corresponding masses. If an ion is found in the corresponding spectrum, it must first pass two criteria before being displayed in a specific color (blue for N-terminal ions and red for C-terminal ions). It must be found within the mass error tolerance chosen by the user and must have an intensity of greater than 2% of intensity of the ion with the highest intensity. The ion types to display in the table are controlled by the same configuration as the spectrum annotation (Section 3.2.2, “Spectrum Annotation”).

#	b	b-H ₂ O	b-NH ₃	Seq	y	y-H ₂ O	y-NH ₃	#
1	116.03	98.02	99.01	D				14
2	230.08	212.07	213.05	N	1489.76	1471.77	1472.74	13
3	327.17	309.15	310.14	P	1375.71	1357.71	1358.71	12
4	455.19	437.18	438.16	Q	1278.69	1260.56	1261.66	11
5	556.24	538.23	539.20	T	1150.64	1132.62	1133.60	10
6	693.30	675.28	676.27	H	1049.60	1031.49	1032.55	9
7	856.36	838.35	839.33	Y	912.52	894.51	895.49	8
8	1019.43	1001.43	1002.40	Y	749.46	731.45	732.43	7
9	1090.47	1072.47	1073.43	A	586.44	568.29	569.32	6
10	1189.53	1171.51	1172.50	V	515.35	497.34	498.33	5
11	1260.56	1242.56	1243.54	A	416.29	398.28	399.26	4
12	1359.66	1341.62	1342.62	V	345.24	327.17	328.22	3
13	1458.70	1440.20	1441.67	V	246.18	228.17	229.15	2
14				K	147.11	129.10	130.09	1

Clicking the header of a column in the Ion Table highlights the corresponding points on the error map and peaks in the spectrum annotation.

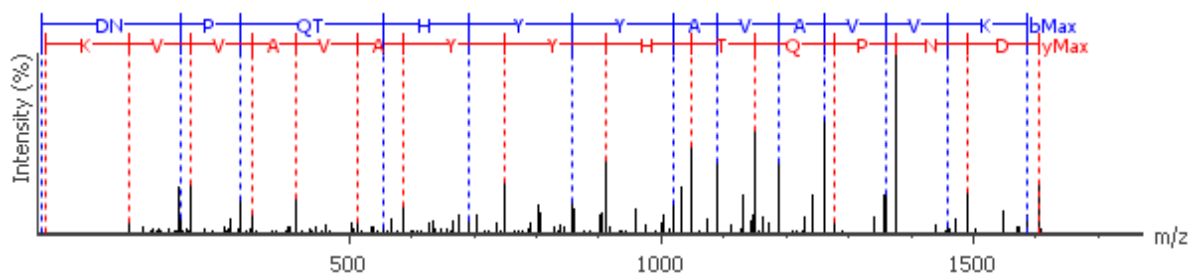
3.2.4. Error Map

The “Error Map” shows the mass errors of the annotated ions and it is displayed on the right hand side of the “Ion Table”. The m/z ratio is displayed on the x-axis and the error is listed on the y-axis in Daltons. The most confident results lie on the centerline.



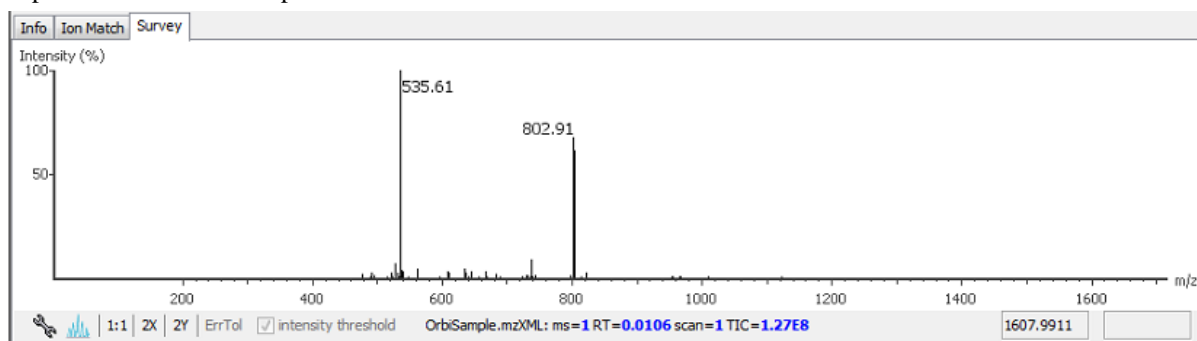
3.2.5. Spectrum Alignment

The “Spectrum Alignment” is displayed under the “Error Map” and it shows the entire spectrum and is used as a tool to help us navigate the “Spectrum Annotation”. The blue bar along the horizontal m/z axis of the alignment indicates the range of the spectrum in the “Spectrum Annotation”. This alignment displays how the proposed ions align with the spectrum. By default, the “Spectrum Alignment” displays b-ions and y-ions. The b-ions are shown right to left in blue, while the y-ions are shown left to right in red.



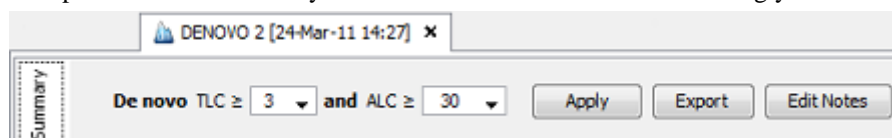
3.2.6. Parent Scan

The “Survey” tab displays the precursor ion spectrum. The buttons that appear in this section are identical to those explained above in the “Spectrum Annotation” section.



4. Filtering *De Novo* Sequencing Result

PEAKS *de novo* sequencing result can be filtered based on TLC (Total local confidence) and ALC (Average local confidence) score filters. You can set the appropriate values for the filters by changing the filtration parameter values from the drop-down lists in the title bar of the “Summary” view panel and clicking on the “Apply” button. The result will be updated in the “Summary” view and the “*De novo*” view accordingly.



Note

Whenever you changed a score threshold, the "Apply" button changes color to remind you to apply the filter by clicking it.

5. Export *De Novo* Results

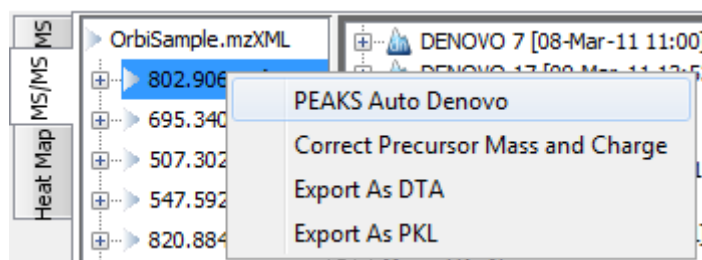
The "Export" button at the top of the summary view allows exporting of the filtered results into a list of html (text file readable in a web browser) and csv (text file readable in Excel) files. This provides the opportunity to supplement the results in a publication, or put up the results on your website. To export the filtered results:

1. Click the "Export" button at the top of the summary view. A file chooser appears.
2. Choose the location and directory name where you want to put the exported files. Click OK.

This will create a collection of files in the target directory. Refer to Section 2, "Export *De Novo* Result" for details.

6. Run Auto *De Novo* Sequencing on a Single Spectrum

To perform auto *de novo* sequencing on a single spectrum, select the spectrum in the MS/MS view of the sample and click the right button to display a pop-up menu. Select the "PEAKS Auto Denovo" command from the pop-up menu.



7. Manual *De Novo* Sequencing

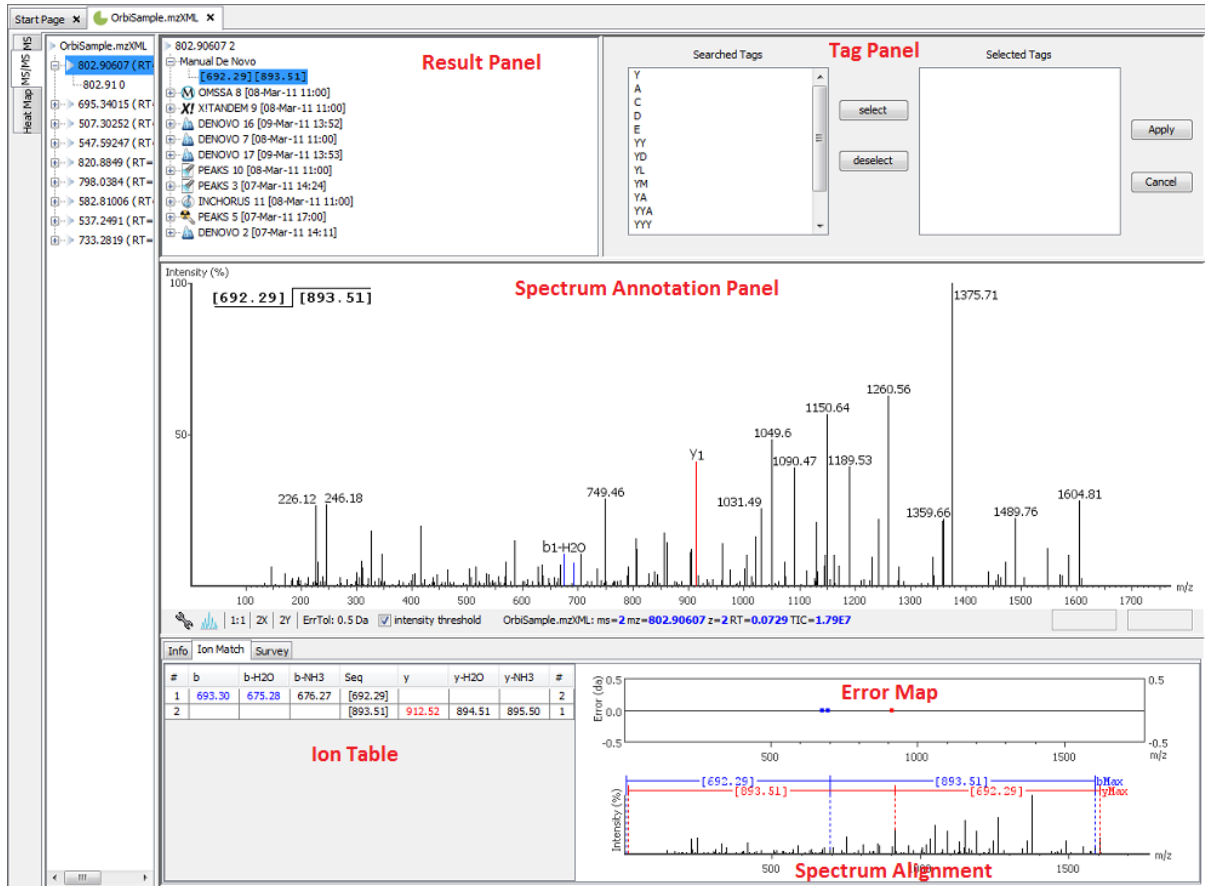
PEAKS 5.3 provides a set of tools to help you manually sequence a peptide using graphic cues from the spectrum.

Note

ETD manual *de novo* is not supported.

7.1. Manual *De Novo* Graphical User Interface

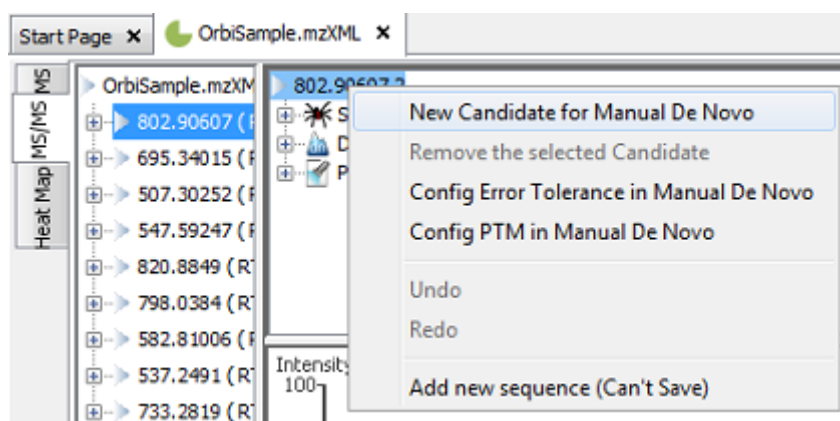
The figure below shows the main panels related to manual *de novo*. The five main panels are indicated in the figure below:



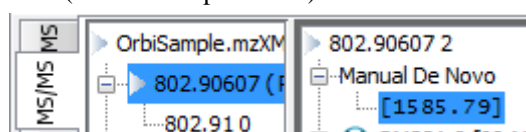
The panels are briefly described below:

- **Result Panel:** The “Result Panel” shows all the sequencing results. The results of manual *de novo* are listed in the sub-tree with root “Manual *De Novo*”.
- **Spectrum Annotation Panel:** The “Spectrum Annotation” shows a graphical representation of the spectrum, the peaks in the spectrum, the user-selected peaks and assigned ions. Users can pick a peak on the panel and assign ions or tags to it in manual *de novo*.
- **Ion Table Panel:** The “Ion Table” shows the proposed ions with their corresponding masses. The default Ion Table will display immonium, b, b-H2O, b-NH3, y, y-H2O, and y-NH3 ions.
- **Spectrum Alignment and Error Map Panel:** The “Spectrum Alignment” shows how the proposed ions assigned in manual *de novo* align with the spectrum. By default, the “Spectrum Alignment” displays b-ion and y-ion. The b-ions are shown right to left in blue, while the y-ions are shown left to right in red. The “Error Map” displays the confidence assigned to each ion.
- **Tag Panel:** The “Tag Panel” will appear when you search tags or ions in the spectrum. You can select the tags in the list using the “Select” button. Clicking “Apply” will add the selected tags to the sequence candidate.

To create a new peptide candidate for sequencing, select the m/z value in the “Result Panel” and right click to bring up a pop-up menu.



Select “New Candidate for Manual *De Novo*” from the pop-up menu. A new candidate will be created under the 'Manual *De Novo*' heading. The new candidate will not have been sequenced, so it will be represented by the mass of the peptide less the mass of water (see an example below).

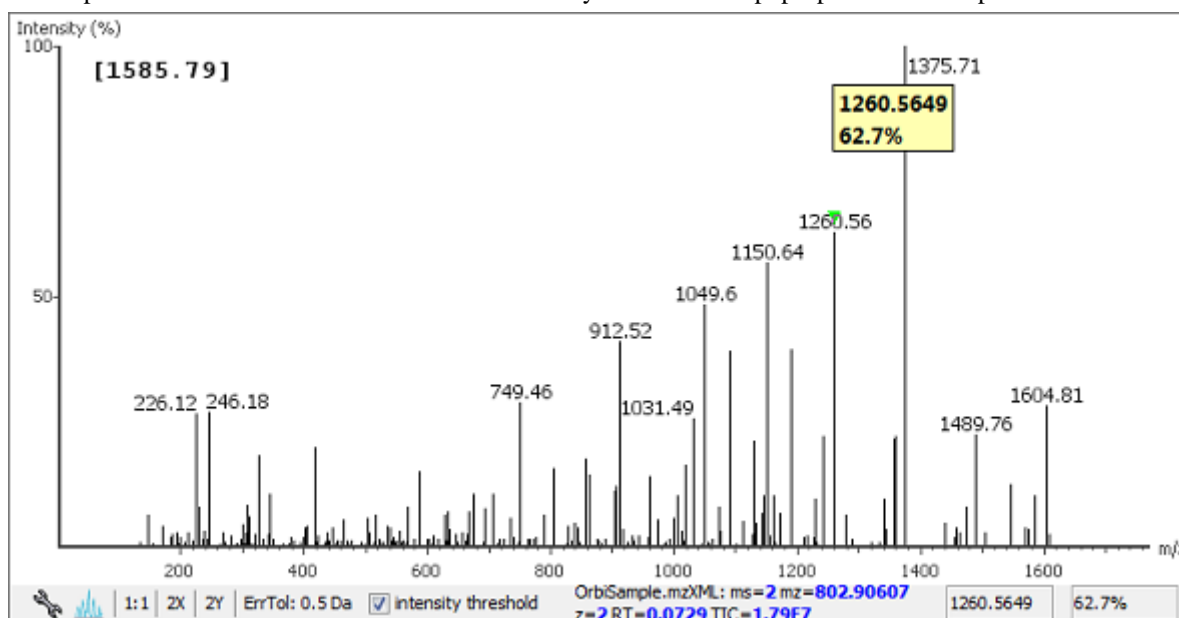


Note

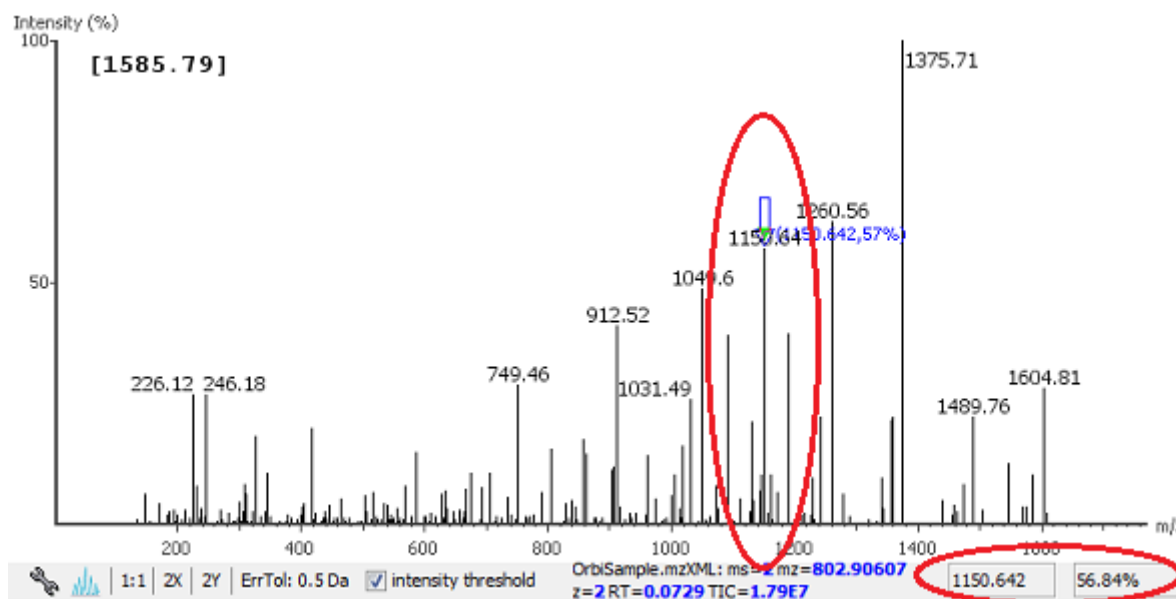
The pop-up menu will not be accessible if you have highlighted any of the results in the “Result Panel”.

7.2. Manual *De Novo* Operations

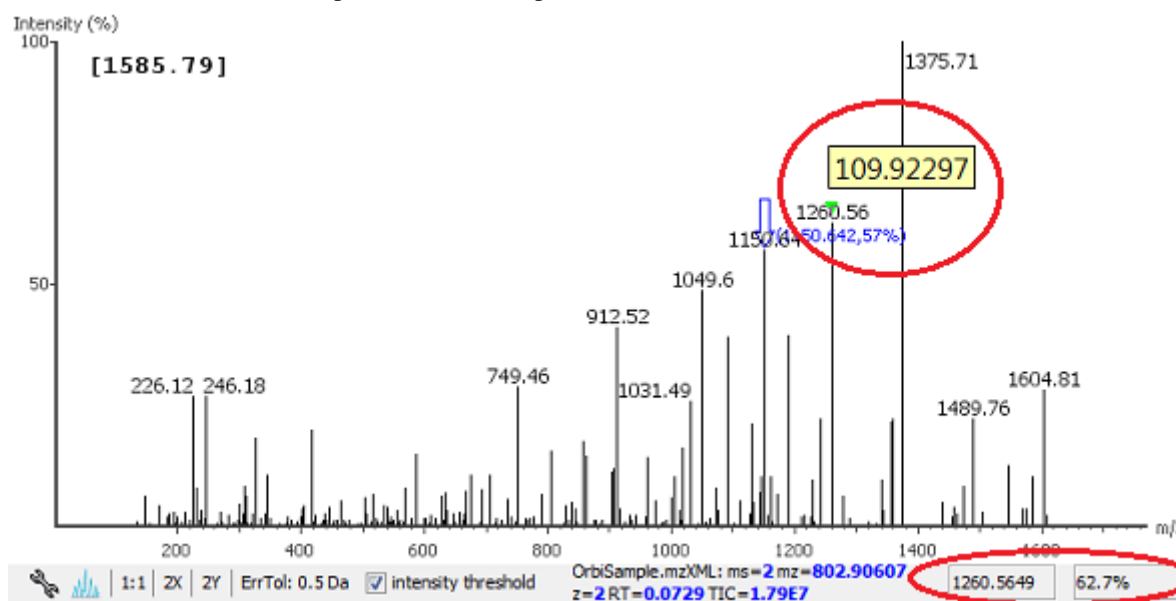
When the mouse cursor is placed in the “Spectrum Annotation” panel, a green (by default) arrow follows the movement of the mouse. This is the *Position Bar* and it is used as a cursor for all manual *de novo* operations. The cursor's position on the *m/z* scale and its relative intensity are shown in a pop-up window on top of the *Position Bar*.



Selecting a peak. To select a peak, simply click on it. A blue (by default) arrow, called the *Freeze Bar*, indicates the selected peak. Alternatively an ion peak can be selected by clicking on its corresponding cell in the Ion Table.



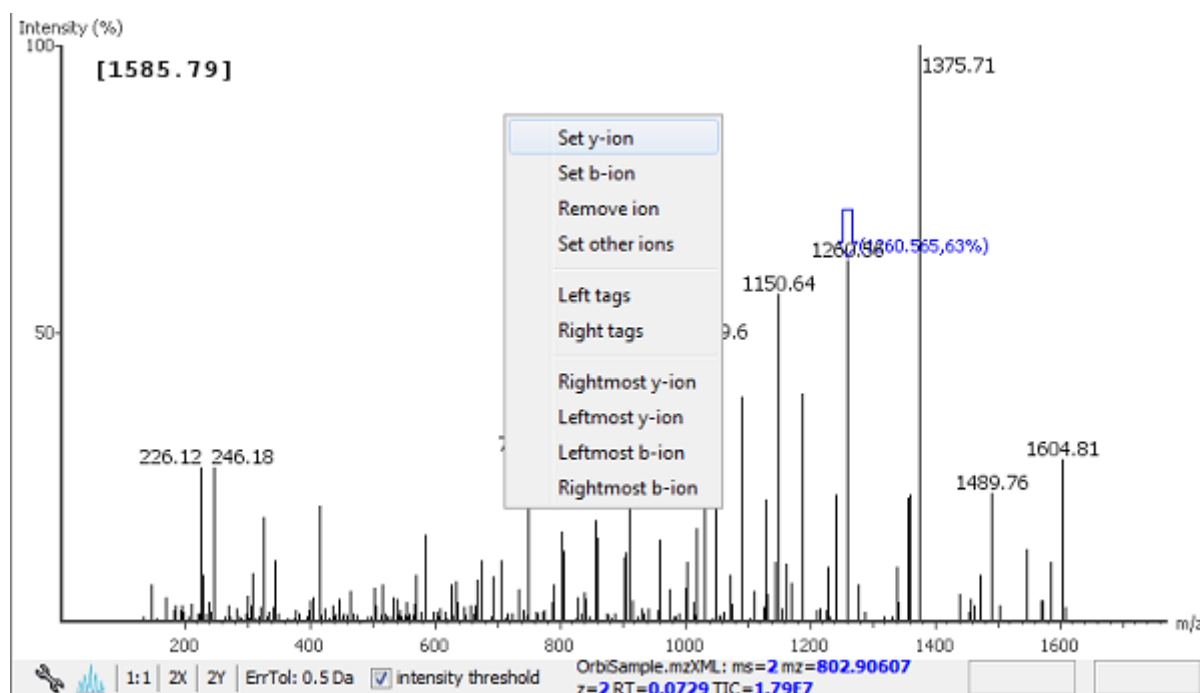
Measure the m/z difference between two peaks. Select a peak (blue arrow) with the *Freeze Bar* and move the mouse to the left or right. Hold the *Position Bar* (green arrow) above another peak. A pop-up window displays the difference between the two peaks (in the example below the difference is 109.92297).



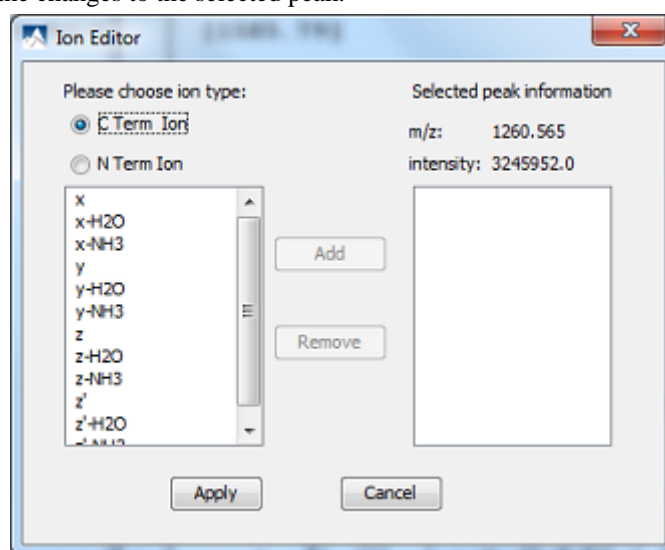
Deselect a peak. Double click anywhere in the “Spectrum Annotation” panel to deselect a peak.

Zoom in on part of the spectrum. In the “Spectrum Annotation” panel, click and drag the mouse horizontally. The selected area will be enhanced and shown in the “Spectrum Annotation” panel. Click on the “1:1” button to return to the default view.

Setting/removing ions to/from a peak. Select a peak, and then right click the mouse anywhere in the “Spectrum Annotation” panel. Select “Set y-ion” from the pop-up menu to designate the peak as a y-ion, or “Set b-ion” from the pop-up menu to designate the peak as a b-ion. Click on “Remove ion” to remove the ion that you have previously set.

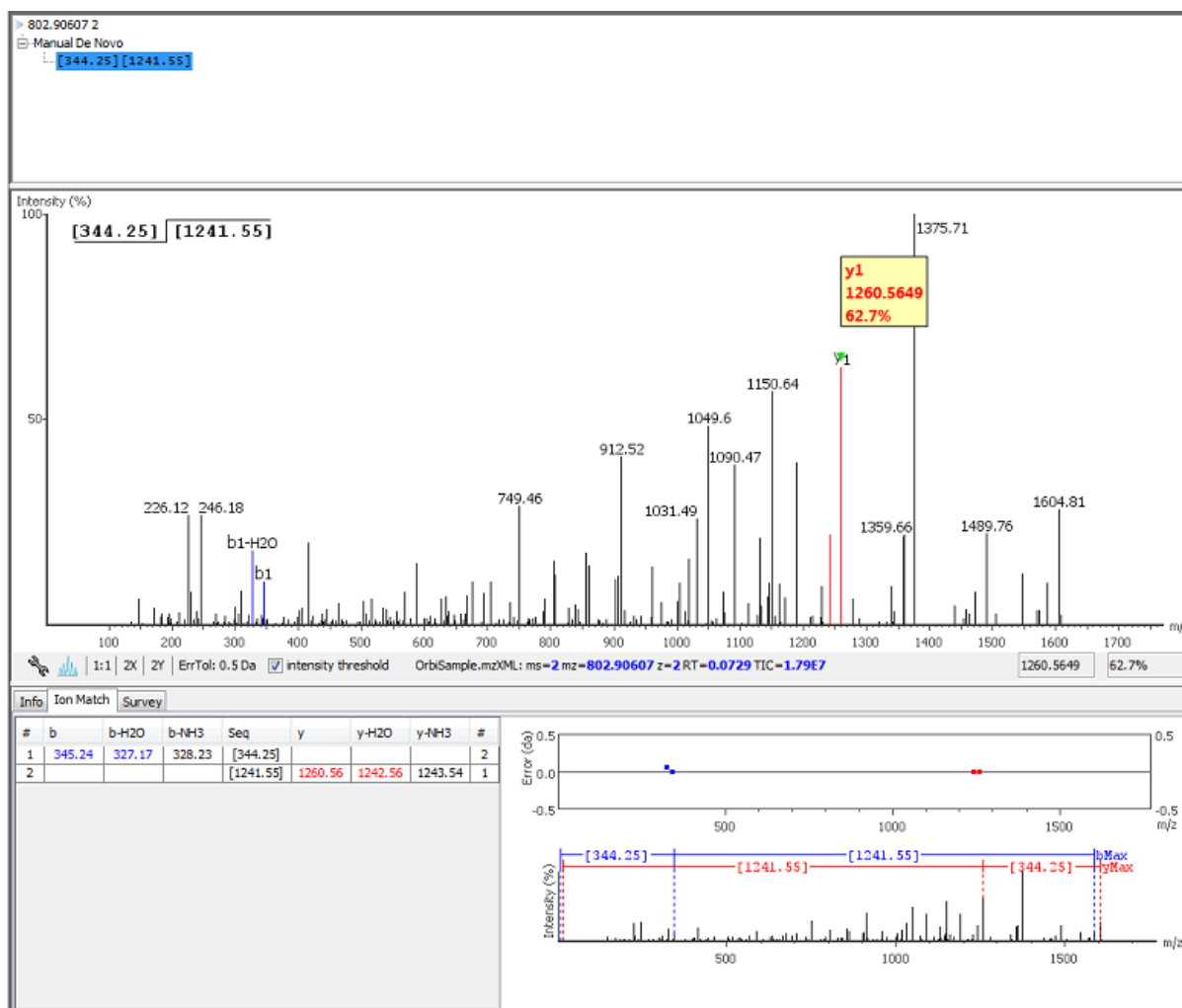


Select “Set other ions” from the pop-up menu to view the “Ion Editor” dialog box. The “Ion Editor” dialogue allows you to add or remove ion designations to/from a peak. Select either “C Term Ion” or “N Term Ion” to see the C and N terminal ions respectively. Then select an ion from the ion list and press the “Add” button to add it to the selected ion list. Remove an ion from the selected ion list by selecting it and pressing the “Remove” button. Click “Apply” to apply the changes to the selected peak.



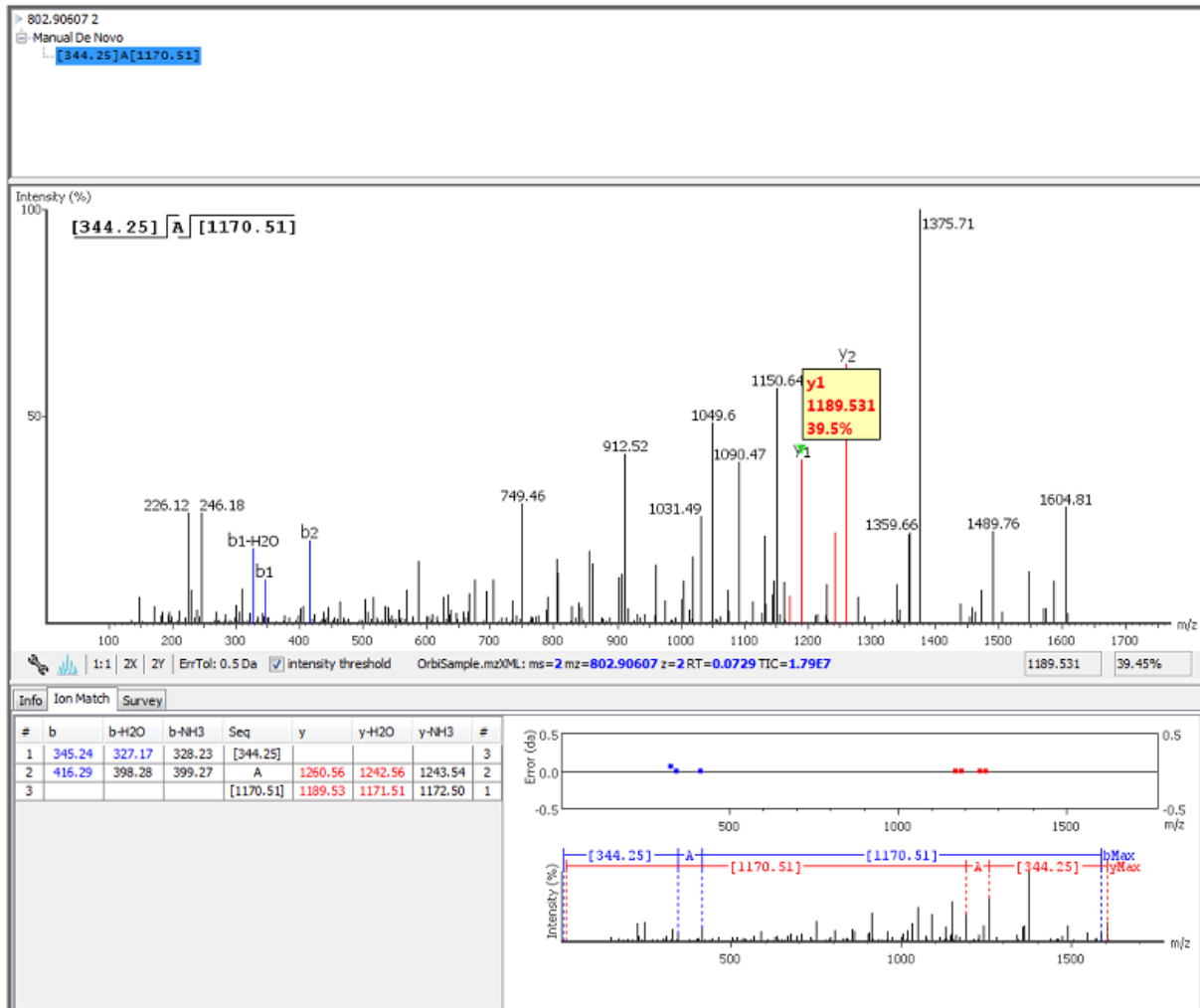
After setting an ion, the “Spectrum Annotation” panel, the “Spectrum Alignment and Error Map” panel and the “Ion Table” panel will reflect the changes. The *peptide sequence candidate* name (as displayed in the “Result” panel and on the top of the “Spectrum Annotation” panel) will also change to reflect the mass remaining to be sequenced on either side of the ion.

In the example below, the selected peak at 1260.5649 m/z was designated as a y-ion.



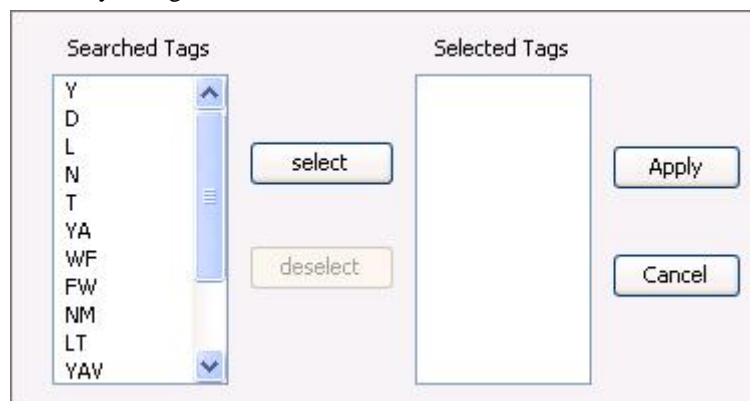
Note that the manual *de novo* candidate information is updated in the “Result” panel, “Ion Table” panel and “Spectrum Alignment and Error Map” panel. The selected ions are also annotated and colour coded in the “Spectrum Annotation” panel.

After setting two ions, PEAKS will estimate the residue found between them (if a residue corresponds closely to the mass difference). The *peptide sequence candidate* name will change to show the residue and the mass remaining to be sequenced on either side of the residue. All other panels will also reflect the changes.



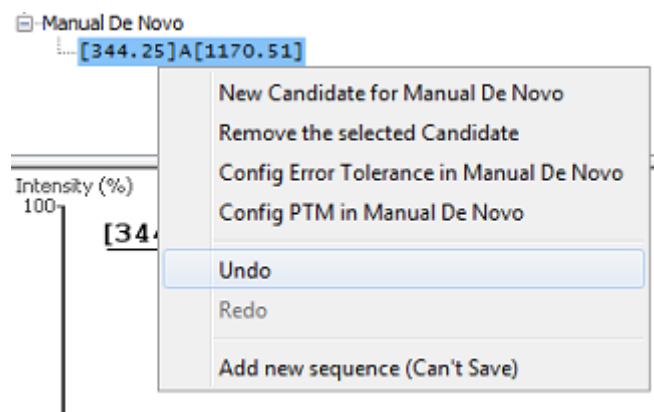
Searching the left or right side of the spectrum for the first/last y or b ion (Search a sequence tag). Select a peak, and then right click the mouse anywhere in the Spectrum View Frame to trigger the popup menu. From the menu, select either “Left tags” or “Right tags”. PEAKS will select the appropriate terminal tags and show them in the “Tag” panel (see below).

You can test the suitability of a tag by highlighting it in the “Searched Tags” list; the corresponding information for the tag will be shown in the “Spectrum Annotation” panel, the “Ion Table” panel and the “Spectrum Alignment and Error Map” panel. You can insert one or more tags by highlighting the desired tags, clicking “Select” to move them into the “Selected Tags” list and then clicking the “Apply” button. Press the “Cancel” button at any time to exit the search and discard any changes.



Undoing an edit. If you have made an error in your sequencing it is possible to undo the change. With the peptide candidate still selected in the “Result” panel, right click the mouse and select the “Undo” command from

the popup menu to return to the previous peptide sequence. You can use this button multiple times to return to earlier stages in your edit.



Redoing an edit. With the peptide candidate still selected in the “Result” panel, right click the mouse and select the “Redo” command from the popup menu if you have undone one too many changes. You can click this button multiple times to return to later stages in your edit.

Error Tolerance. To set the mass error tolerance in manual *de novo* sequencing, with the peptide candidate selected in the “Result” panel, right click the mouse and select the “Config Error Tolerance in Manual *De Novo*” command from the popup menu to open a dialog where the error tolerance can be set.

PTM configuration. To mention the types of post-translational modifications (PTM) to include in manual *de novo* sequencing, with the peptide candidate selected in the “Result” panel, right click the mouse and select the “Config PTM in Manual *De Novo*” command from the popup menu to open the “PTM Setup” window. To know more about the PTM configuration using the “PTM Setup” window, refer to Section 2.3, “Fixed and Variable PTMs”.

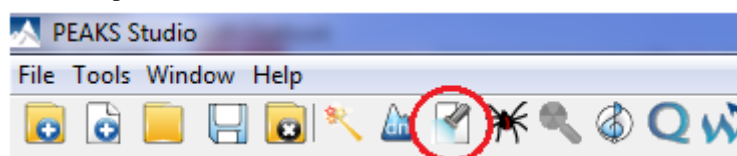
Chapter 9. Protein Database Search with PEAKS DB

1. Overview

The PEAKS DB function identifies peptides/proteins from a protein sequence database. If your target proteins are in a known protein sequence database, this is the recommended method for analyzing your MS data. PEAKS DB belongs to the "database search" approach in MS analysis. However, it takes advantage of the PEAKS *de novo* sequencing results to achieve a higher sensitivity and accuracy than other software purely based on database search. PEAKS DB includes a built-in result validation to ensure the quality of the reported results. Additionally, PEAKS DB automatically generates a list of highly confident *de novo* sequences whose spectra do not match database peptides. These are possible novel peptides in the sample.

The use of this function is outlined in the following section. Details of each step can be found in later sections of this chapter.

1. Select a project node or a sample node. Click the PEAKS DB button on the tool bar.



Note

Refer to Chapter 4, *Loading Data to a PEAKS Project* for how to create a project.

2. Specify the PEAKS DB parameters in the popup dialog and click OK. Most of the parameters are self-explanatory and the default parameters provide a good starting point for the analysis.

Note

If your data is not refined yet, you also need to specify the data refinement parameter first and click next. Refer to Chapter 7, *Data Refinement* for data refinement.

3. Wait for the analysis to finish. A new result node will appear in the project tree. Double click the node to open the result file.

A screenshot of the PEAKS Studio interface. On the left is the 'Project View' pane showing a tree structure: 'C:/Users/binma/Dropbox/tmp/paper_0415_cid' > 'Sample 1' > '20090115_SMPA14_2.RAW' > 'DATA REFINE 1 [15-Apr-11 15:23:]' > 'DENOVO 2 [15-Apr-11 16:09]' > 'PEAKS 4 [15-Apr-11 16:21]' > 'PEAKS 6 [15-Apr-11 16:44]'. The 'PEAKS 6' node is selected. On the right is the 'PEAKS 6 [15-Apr-11 16:44]' results window. It has a 'Summary' tab selected. The window shows a table with columns 'Accession', '-10lgP', and 'Coverage'. The table lists several protein entries with their accession numbers, -10lgP values, and coverage percentages. A 'Show' dropdown is set to 'top' and 'proteins in each group' is selected. The table data is as follows:


Accession	-10lgP	Coverage
gi 15599581	330.38	57%
gi 15599462	300.09	48%
gi 15598358	291.85	41%
gi 15599955	287.56	51%
gi 15596984	280.50	36%
gi 15599466	274.51	24%
gi 15600749	263.10	40%
gi 15596793	260.29	43%
gi 15599461	258.95	55%

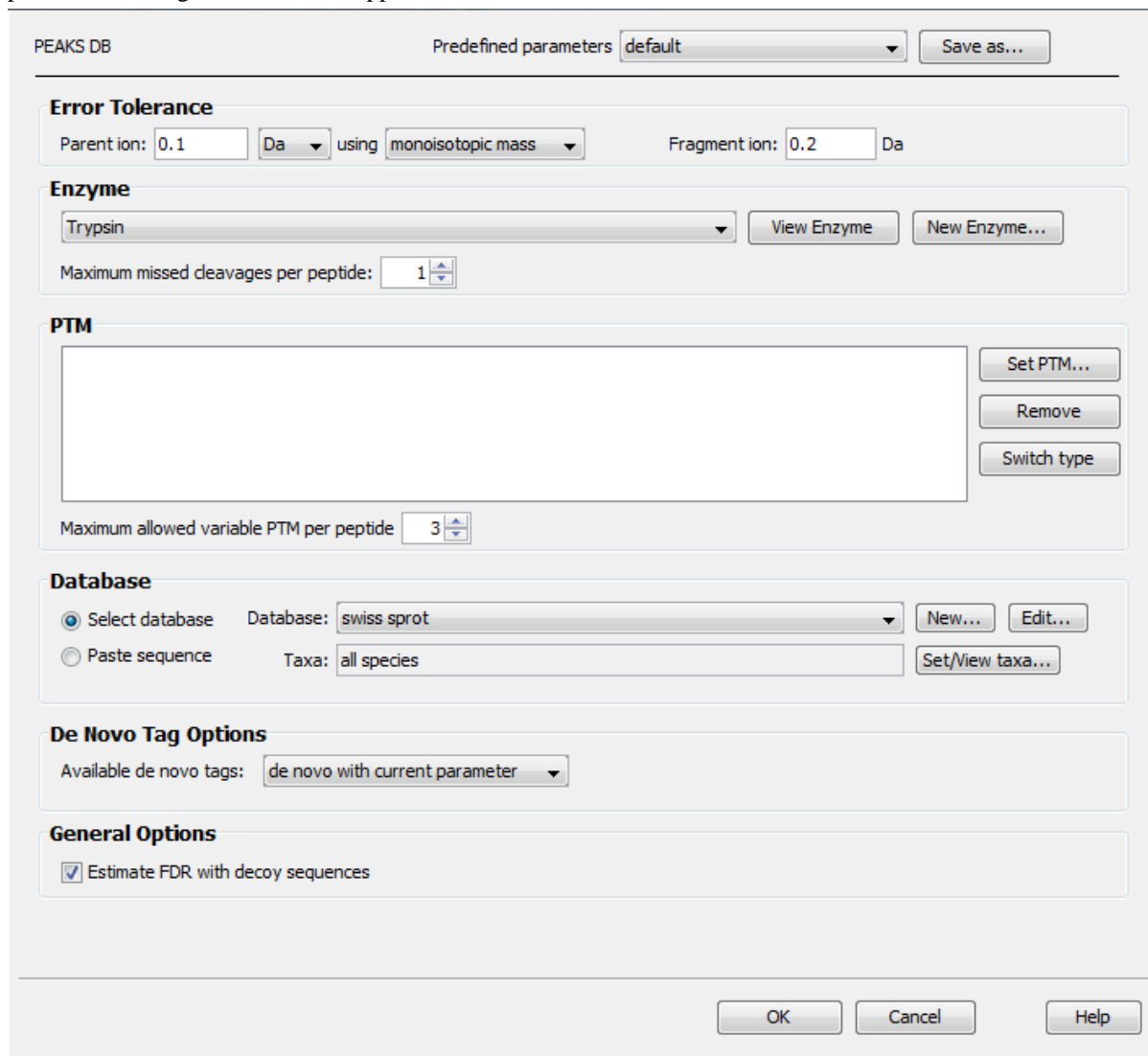
4. The result contains four different views: Summary, Peptides, Proteins, and *De novo* only. These provide different angles for examining the analysis. In particular, the summary view allows you to specify rules to filter

the results, and provides information for you to judge the quality of the experiment. The *de novo* only view shows the high confident *de novo* sequence tags that do not have any high-confidence database matches. These may be novel peptides in your sample.

5. You can optionally export the results to other formats by using the summary view.

2. Set PEAKS DB Parameters

After selecting a data node in the project tree, click the PEAKS DB toolbar icon . The protein identification parameters dialogue window will appear:



PEAKS DB

Predefined parameters: default Save as...

Error Tolerance

Parent ion: 0.1 Da using monoisotopic mass Fragment ion: 0.2 Da

Enzyme

Trypsin View Enzyme New Enzyme...

Maximum missed cleavages per peptide: 1

PTM

Set PTM...
Remove
Switch type

Maximum allowed variable PTM per peptide: 3

Database

☒ Select database Database: swiss sprot New... Edit...

☐ Paste sequence Taxa: all species Set/View taxa...

De Novo Tag Options

Available de novo tags: de novo with current parameter

General Options

☒ Estimate FDR with decoy sequences

OK Cancel Help

Note

If your data is not refined in PEAKS yet, you will be prompted to specify the data refinement parameters. Refer to Chapter 7, *Data Refinement* for data refinement parameters.

Error Tolerance. The mass error tolerance of the parent (precursor) and fragment ions. The parent ion error tolerance can be specified in either Da or ppm, and using monoisotopic or average mass.

Enzyme. Choose the enzyme used to digest the proteins, and the number of missed cleavages allowed in each peptide. Refer to Section 2.2, “Enzyme Specificity” for details.

PTM. Click the “Set PTM...” button to choose a list of PTMs to be used during the search. Refer to Section 2.3, “Fixed and Variable PTMs” for details.

Database. Select the protein sequence database for the search. You can select from the list of databases which have been configured, and set the taxonomy if applicable. To configure a new sequence database, refer to Chapter 6, *Adding a Sequence Database*.

Or if you have only a few protein sequences, you can choose to paste the protein sequences from a Windows clipboard.

De novo tag options. PEAKS DB requires the *de novo* sequencing results to improve its search speed, sensitivity and accuracy. You can choose to perform a fresh new *de novo* sequencing, or select from the existing *de novo* sequencing results, if there are any.

Estimate FDR. Selecting this option enables PEAKS DB to validate the search results with an enhanced target-decoy method. A few important statistical charts in the summary view of the PEAKS DB result will depend on this. Uncheck this only if you want to do your own result validation.

Perform on-the-fly preprocessing. Check this box if you have not pre-processed your data in the data refinement step, and uncheck it if you have already done so.

3. Understanding PEAKS DB Result

After PEAKS DB is done, two result nodes will be generated. One is a database searching result, and the other is from the subroutine - *de novo* sequencing (when not using an existing *de novo* tag). The result of PEAKS DB consists of:

- **Summary** - outline of PEAKS DB outputs with statistics. Also a place for result filtration.
- **Peptide** - list of peptide identification.
- **Protein** - list of protein/protein groups inferred by identified peptides
- **De novo only** - list of good *de novo* sequences without a good assignment from database search.

3.1. The Peptide and Protein Scores

Peptide score (-10lgP). The scoring schema of peptide identification involves matched peaks and their intensities, precursor mass error, enzyme specificity, *de novo* sequence, and peptide size, etc. A statistical evaluation, -10lgP, is given for each peptide-spectrum match. Here lg() is the common logarithm with base 10, and P is the probability that a false identification of the current search has the same or better significance.

Protein score (-10lgP). The protein -10lgP score in PEAKS is the weighted sum of -10lgP score of all supporting peptides. After removing redundancies, PEAKS DB sort those peptides from the same protein according to their -10lgP scores. In the weighted sum, the k-th ranked peptide gets a weight 1/k.

3.2. Summary View

The summary view provides three main functions:

1. Result filtration - this is achieved by specifying the filtration rules in the area at the top of the summary view. The filtration function is discussed in a separate section, Section 4, “Filter PEAKS DB Result”.
2. Result exporting - this is achieved by clicking the "Export" button at the top of the summary view. The exporting function is discussed in a separate section, Section 5, “Export *De Novo* Results”.

3. Summary report - several statistical charts assist the user to get an overall picture of the results, assess the result quality, and examine the reliability of the mass spectrometer. This function is the focus of this section.

The charts in the report are divided into four sections:

1. Notes

A user can enter a special text note regarding the experiment. Click the "Notes" button at the upper right corner of the summary view to edit the note.

2. Result Statistics

The first three figures provide important information for validating the database search result. Given the large volume of MS data, we cannot over-emphasize the importance of statistical result validation. Without it, the analysis result is simply not trustworthy.

Then four tables summarize the data and results, such as the number of confidently identified peptides and how many contain a particular PTM.

3. Instrument Control

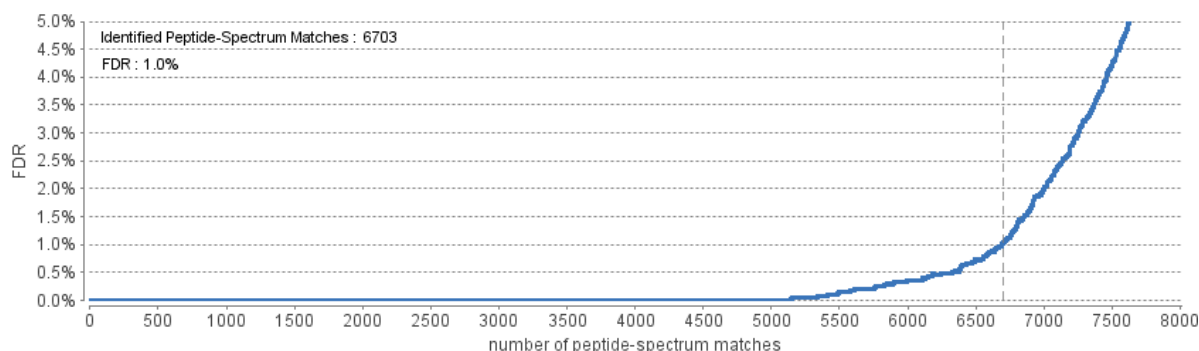
Two figures plot the precursor ion mass error distribution, revealing how well the instrument is calibrated.

4. Other Information

The search parameters and MS instrument information are given here.

In the rest of this section we discuss the most important charts in the summary report.

False Discovery Rate (FDR) Curve. Figure 1 in the summary page is the FDR curve for the identified peptide-spectrum matches (PSM). PEAKS keeps at most one peptide for each spectrum (peptides with only I/L isoform difference are counted as one). Thus, the number of PSMs is the same as the number of spectra with assigned peptides. The PSMs are sorted according to their $-10\lg P$ scores. The curve shows the FDR with respect to the number of PSMs to be kept in the final result. If a score threshold has been provided in the result filtering, a vertical dashed line indicates the score threshold.



Normally a $<1\%$ FDR is recommended for score filtering. If you notice a rapid growth of FDR around the the 1% FDR threshold, you may decide to sacrifice several PSMs to significantly reduce the FDR.

The FDR curve is estimated with the decoy fusion method (an enhanced target-decoy method that is more conservative in keeping results) performed together with the PEAKS DB search. The "Estimate FDR" checkbox must be checked in the search parameters to enable this function.

Note

The decoy hits are removed from the counting of the number of PSMs in the FDR curve. Similarly, unless otherwise specified, all the counts in the summary view have excluded the decoy hits. By default, the false hits are also excluded from the peptide and protein views, as well as the exported results.

PSM Score Distribution. Figure 2 (a) and (b) help you assess the quality of the results and the effectiveness of the enhanced target-decoy method (decoy fusion). It is strongly recommended to turn on the "Estimate FDR"

checkbox in the search parameters, so that both the target and decoy PSMs are shown in the same figure with different colors.

Figure 2(a) shows the number of PSMs at each score interval. If the target-decoy method worked as promised, then you should observe a similar number of the target (blue) and decoy matches (brown) in the low score region. If the search result is of high confidence, then you should observe very few decoy matches (brown) in the high score region. The vertical dashed line indicates the user-specified score threshold.

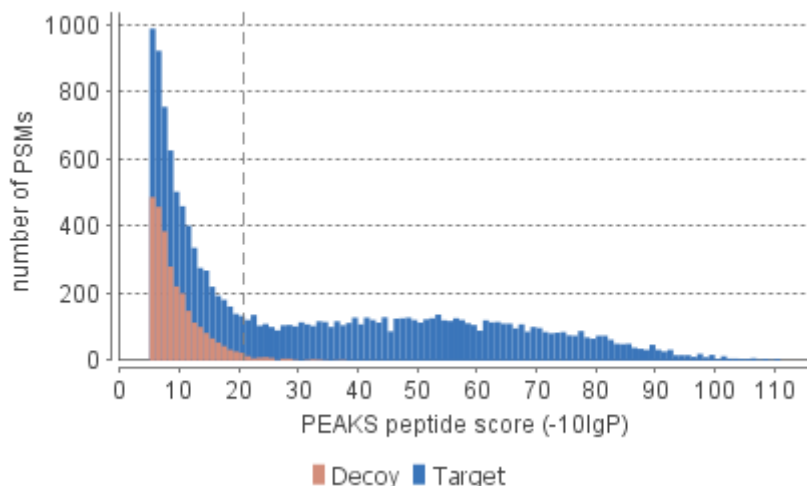
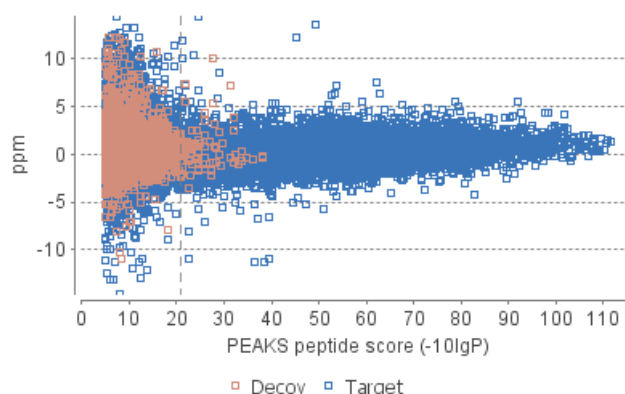


Figure 2(b) plots the precursor mass error versus score for all the PSMs. This figure is the most useful for the high mass resolution instruments. Generally you should see that the high-scoring points are centered around the mass error 0. And only below a certain score threshold the data points start to scatter to have bigger mass error. The vertical dashed line indicates the user-specified score threshold.



Statistics of Data and Results. Tables 1-4 shows the statistical numbers of the data and results.

Table 1. Statistics of data.

# of MS scans	7737
# of MS/MS scans	20319

Table 3. Statistics of filtered result.

Peptide-Spectrum Matches	7036
Peptide Sequences	1820
Proteins Groups	632
Proteins	638
Proteins unique peptides	220 (>2); 117 (=2); 287 (=1);
FDR (Peptide-Spectrum Matches)	1.0%
FDR (Peptide Sequences)	2.9%
denovo only	2022

Table 2. Result filtration parameters.

Peptide -10lgP	≥20.9	De novo TLC	≥3
Protein -10lgP	≥20	De novo ALC	≥50%
Proteins unique peptides	≥0		

Table 4. PTM Profile.

Q(+0.98,-17.03)	3
Q(-17.03)	71
Q(+0.98)	51
M(+15.99)	82
N(+0.98)	94
C(+57.02)	660

Most entries in these tables are self-explanatory. A few worth-mentioning are:

- **Peptide Sequences** (Table 3)

This is the number of distinct peptides in the filtered result. Peptides with the same primary sequence but different PTMs are counted separately. But several peptides differentiated with only I/L isoform are counted as one. Since the same peptides may be identified by multiple spectra (due to data redundancy and different charge states), this number is usually smaller than the number of Peptide-Spectrum Matches.

- **Protein Groups** (Table 3)

PEAKS DB groups the proteins identified by the same set of peptides (or a subset) into the same group, as there is not enough information to determine which of them contribute to the identified peptides in the sample. This number in the table shows the number of protein groups in the filtered result.

- **Proteins (#Unique Peptides)** (Table 3)

These show the number of identified proteins with the specific number of unique peptides. A unique peptide is a peptide that passes the user's peptide filtration score threshold and appears in only one protein group.

- **PTM Profile** (Table 4)

Beside each PTM is the number of the identified peptide sequences (not PSMs) containing the PTM.

Experiment Control. Figures 3(a) and 3(b) plots the precursor m/z error of the identified PSMs. These plots can help determine whether the MS instrument functioned properly. Figure 3(a) is the histogram of the mass errors. If the instrument worked properly, then the histogram should be concentrated around 0 ppm. Figure 3(b) plots each PSM using its m/z (x-axis) and mass error (y-axis). For a well-calibrated instrument, the data points should be distributed within a narrow horizontal band centered at the 0 ppm horizontal line. Table 5 shows the number of peptides by number of miss cleavages which indicates the efficiency of the enzyme.

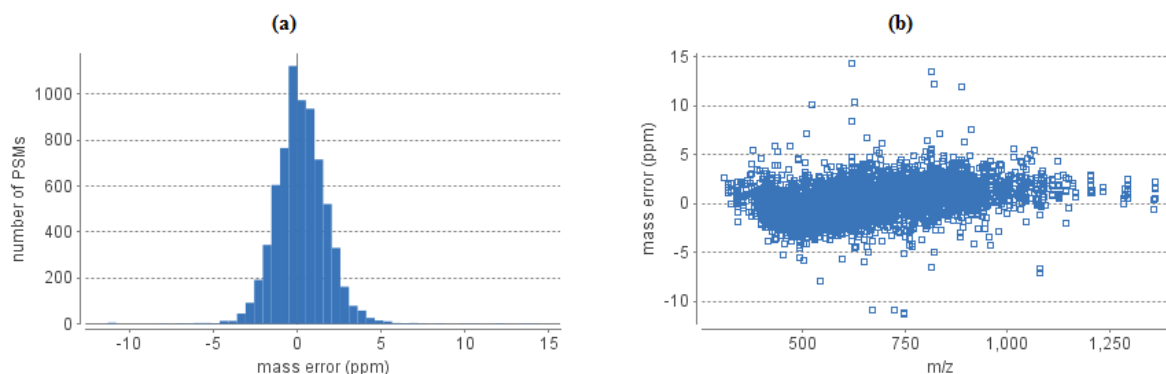


Table 5. number of peptides by number of missed cleavages

Missed Cleavages	1	2	3
Peptides	173	26	0

3.3. Peptide View

The “Peptide View” shows identified peptides. The interface contains a peptide table that supports sorting and the search for a peptide. Selecting any peptide in the table will display the peptide-spectrum matching details at the bottom half of the peptide view.

3.3.1. Peptide Table

All peptides above the user-specified peptide score threshold are listed in the table. If there are more than 1000 peptides, the list is broken into multiple pages.

Summary	1 - 1000 of 1934											scan = <input type="text" value="search"/> no results		
		Peptide	-10lgP	Mass	ppm	m/z	RT	Scan	#Spec	Accession	PTM			
Protein	1	C(+57.02)YGSTNPVNVVYATFK	110.59	1818.8611	1.3	910.44	142.051	19458	20	gi 15599442				
	2	VYLVTGWTTGGYGVGK	105.99	1726.8678	1.2	864.44	111.836	14330	3	gi 15600385				
Peptide	3	VAAAGPLLAELDALK	105.89	1578.8981	0.3	790.46	160.238	21983	11	gi 15595749				
	4	LNAVSAAGNDGLQAAPAAEPK	104.63	1949.9806	1.4	976.00	48.779	3286	4	gi 15596353				
De novo only	5	EAPLHVSNAIFNTETSK	104.51	1955.9952	1.6	979.01	98.308	11904	6	gi 15599448				
	6	QTFPDVQVIGGNIATAEAAK	103.48	2029.0480	1.6	1015.53	130.443	17557	5	gi 15598965				
	7	ANDAAGDGTATVLAQAIVNEGLK	103.36	2400.2131	1.9	1201.12	164.791	22523	6	gi 15599581				
	8	GTDLVAVTGVIHPLGAYSSK	101.78	2113.1055	0.0	1057.56	130.815	17618	3	gi 15600698				
	9	AITVNGQTVNIPSYQVK	101.61	1830.9839	0.5	916.50	111.680	14300	7	gi 15599435				
	10	HVTDEAALAGLTDSAK	100.98	1597.7947	-0.2	799.90	64.681	5840	2	gi 15595265				
	11	AGFLSATVKGATSLIDK	100.27	1916.0618	0.9	959.04	146.860	20205	6	gi 15596984				
	12	ATVTDKQVSYEEAVKPAFAPOAA	98.95	2402.1965	1.7	1202.11	58.018	4678	4	gi 15596997				

The table provides the following controls:

- **Sorting by the column:** Table can be sorted by clicking the headers.
- **Going to a different page:** Use the combo box or the left/right arrows located at the left upper corner of the table.
- **Searching for a specific peptide:** First select the search criterion by clicking the triangle beside the search box, and then type in the value in the search box. Search criteria include scan id, partial sequence, m/z, retention time (RT), and PTM delta mass. Once a search is done, click the circled up and down arrows to navigate in the matched peptides.

For each peptide sequence in the table, several columns are given:

- **Peptide:** The amino acid sequence of the peptide. If there is any PTM on an amino acid, the amino acid is followed by a pair of parentheses enclosing the delta mass of the PTM.

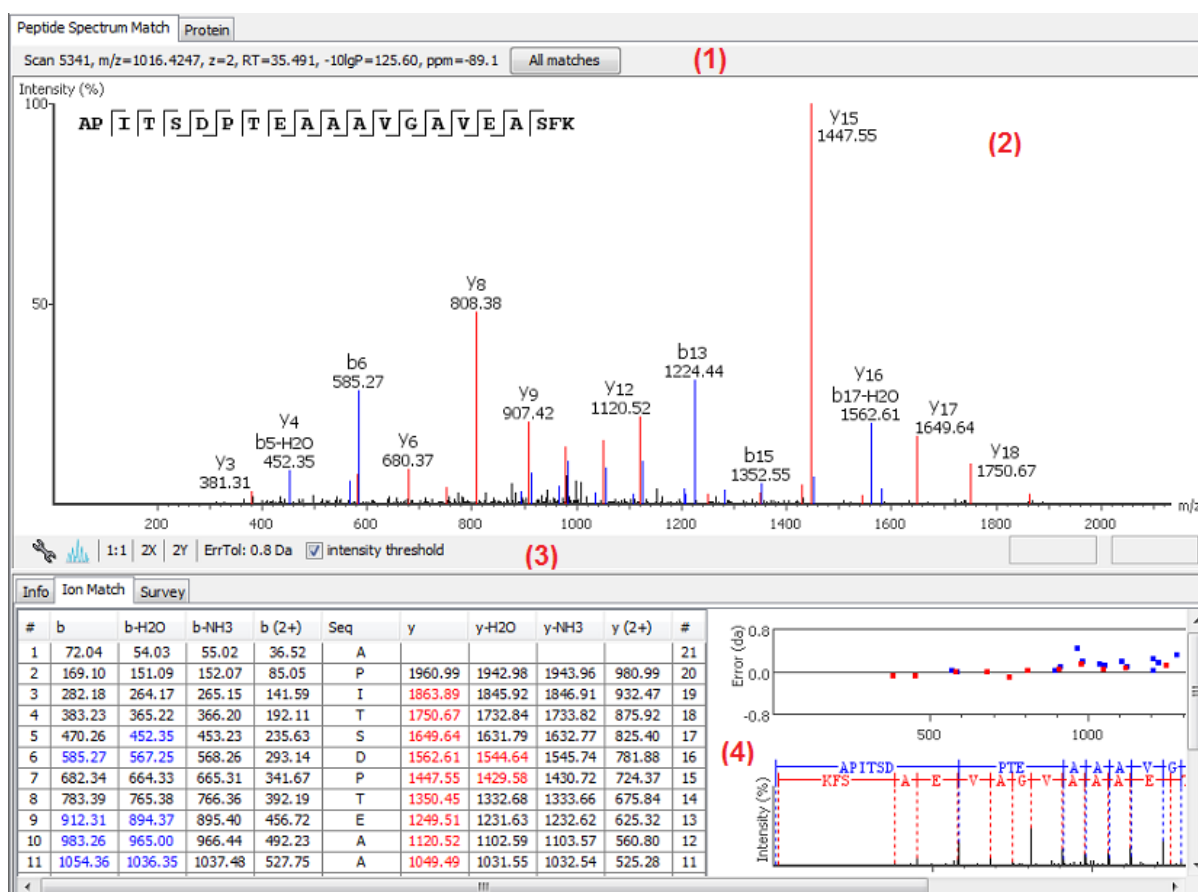
Note

If multiple PSMs have the same sequence, then only the top scoring one is displayed. The #Spec column shows how many spectra are assigned to the same peptide. The other PSMs can be examined by selecting the peptide. See Section 3.3.2, “Peptide Spectrum Match” for details.

- **-10lgP:** The peptide matching score.
- **Mass:** The theoretical mass of the peptide (including the H₂O but not the extra proton for positive charge).
- **ppm:** The precursor mass error, calculated as $10^6 \times (\text{observed mass} - \text{theoretical mass}) / \text{theoretical mass}$.
- **m/z:** The precursor mass to charge ratio.
- **RT:** Retention time.
- **Scan:** Scan number.
- **#Spec:** Number of spectra assigned to the peptide.
- **Accession:** The accession number of the highest-scoring protein containing this peptide. The other proteins containing this peptide can be examined in the Peptide Detail panel. See Section 3.3.3, “Protein”.
- **PTM:** A dot indicates the peptide contains a PTM.

3.3.2. Peptide Spectrum Match

For each peptide, the “Peptide Spectrum Match” shows the peptide-spectrum matching details.



The default display is divided in four areas:


1. The spectrum information.

When multiple spectra match the same peptide, the top-scoring spectrum is chosen by default. The spectrum information, including the peptide-spectrum matching score and mass error, are displayed in this area. The other spectra can be examined by clicking the "All matches" button.

2. The spectrum annotation. The annotation provides a few convenient ways to zoom and navigate in the spectrum.

- Zoom to a m/z region - click the desired start m/z and drag horizontally to the desired end m/z, release the mouse button.
- Zoom in/out smoothly - place the mouse pointer at a particular m/z value (right below the x-axis line), scroll the mouse wheel button.
- Increase the peak intensity - place the mouse pointer in the spectrum, scroll the mouse wheel button.
- See the whole spectrum - double click in the spectrum or click the "1:1" button.

3. The controls for the spectrum annotation.

- Click the  button to decide the fragment ion types to be annotated in the spectrum.
- Click the "ErroTol" to set the mass error tolerance to annotate fragment ions.
- Deselect the "intensity threshold" checkbox to turn on the low intensity peak annotation.

4. The ion match table, error plot, and peptide-spectrum alignment.

- Clicking the header of an ion type column in the ion table will let the spectrum annotation and the error plot only display only that particular ion type.
- The error plot shows the mass error and m/z of each annotated peak. A good peptide spectrum match should have these dots centered at error=0 line.

3.3.3. Protein

Besides the Peptide Spectrum Match, the Protein tab shows the proteins that contain the selected peptide. If the peptide is found in multiple proteins, the top-scoring protein is displayed by default. The other proteins can be examined by selecting from the drop down list at the top of the Protein tab. For each selected protein, this view shows

- the alignment between the *de novo* sequence and the database search sequence for the same spectrum, and
- the primary sequence of the protein.

Peptide Spectrum Match Protein

KPYM_MOUSE

De Novo: <S(+79.97)N>TSDPTEAAVGAVEASF

Database: <API>TSDPTEAAVGAVEASF

DB peptides shown in blue

PKPHSEAGTA FIOTOQLHAA MADTFLEHMC RLDIDSAPIT ARNTGIICTI GPASRSVEML

1 KEMIKSGMNV ARLNFSHGTH EYHAETIKNV REATESFASD PILYRPVAVA LDTKGPEIRT

GLIKSGTAE VELKKGATLK ITLDNAYMEK CDENILWLDY KNICKVVEVG SKIYVDDGLI

61 SLOVKEKGAD FLVTEVENG SLGSKKGVNL PGAAVDLPV SEKDIDLKF GVEQDVMVF

ASFIRKAADV HEVRKVLGEK GKNIKIISKI ENHEGVRRFD EILEASDGIM VARGDLGIEI

121 PAEKVFLAQK MMIGRCNRAG KPVICSTQML EIMIKKPRPT RAEGSDVANA VLDGADCIML

SGETAKGDYP LEAVRMOHLI AREAIAIYH LQFEEELRRL APITSDPTEA AAVGAVEASF

181 KCCSGAIIVL TSGRSAHQV ARYRPRAPII AVTRNPOTAR OAHLYRGIFP VLCKDAVLNA

WAEDVDLRVN LAMDVGKARG FFKKGDVVIV LTGWRPGSGF TNTMRVVPVP

Color Code: >=75% >=50% >0%

3.4. Protein View

The Protein View contains a table of identified proteins. For each selected protein in the table, three tabs are provided at the bottom half of the view:

- Peptides: the peptides identified from this protein.
- Coverage: a graphical display of the protein coverage.
- Tool Box: useful tools such as a BLAST search on NCBI's website.

Start Page x

PEAKS 6 [15-Apr-11 16:44] x

Summary

Protein

Peptide

De novo only

Show top proteins in each group

accession contains search no results

Accession	-10lgP	Coverage	#Peptides	#Unique	Mass	Description	Mark
gi 15599581	330.38	<div><div></div></div> 57%	29	29	57085.961	gi 15599581 ref NP_253075.1 ...	<input checked="" type="checkbox"/>
gi 15599462	300.09	<div><div></div></div> 48%	18	18	77784.531	gi 15599462 ref NP_252956.1 ...	<input checked="" type="checkbox"/>
gi 15598358	291.85	<div><div></div></div> 41%	20	20	61869.980	gi 15598358 ref NP_251852.1 ...	<input checked="" type="checkbox"/>
gi 15599955	287.56	<div><div></div></div> 51%	20	20	68403.273	gi 15599955 ref NP_253449.1 ...	<input checked="" type="checkbox"/>
gi 15596984	280.50	<div><div></div></div> 36%	19	19	93628.242	gi 15596984 ref NP_250478.1 ...	<input checked="" type="checkbox"/>
gi 15599466	274.51	<div><div></div></div> 24%	25	25	150837....	gi 15599466 ref NP_252960.1 ...	<input checked="" type="checkbox"/>
gi 15600749	263.10	<div><div></div></div> 40%	17	17	55393.250	gi 15600749 ref NP_254243.1 ...	<input checked="" type="checkbox"/>
gi 15596793	260.29	<div><div></div></div> 43%	18	18	71668.055	gi 15596793 ref NP_250287.1 ...	<input checked="" type="checkbox"/>
gi 15599461	258.95	<div><div></div></div> 55%	20	20	43369.684	gi 15599461 ref NP_252955.1 ...	<input checked="" type="checkbox"/>
gi 15596353	258.77	<div><div></div></div> 20%	9	9	107106....	gi 15596353 ref NP_249847.1 ...	<input checked="" type="checkbox"/>
gi 15596997	253.95	<div><div></div></div> 41%	15	15	48582.117	gi 15596997 ref NP_250491.1 ...	<input checked="" type="checkbox"/>
gi 15599465	250.19	<div><div></div></div> 20%	20	20	154382	gi 15599465 ref NP_252959.1 ...	<input checked="" type="checkbox"/>

Peptides

Coverage

Tool Box

	Peptide	Unique	-10... ↓	Mass	ppm	m/z	RT	Scan	#Spec	Start	End	PTM
1	K.ANDAGDGTTTATVLAQAIIVNEGLK.A	•	103.36	2400.2...	1.9	1201...	164....	22523	6	81	105	
2	K.ATVAIVAQLKELAKPC(+57.02)ADTK.A	•	95.77	2126.1...	2.1	1064...	125....	16757	7	123	142	•
3	R.QIVANAGDEPSVVDK.V	•	91.90	1639.8...	0.2	820...	54....	4099	10	453	468	
4	R.Q(-17.03)IVANAGDEPSVVDKVK.Q	•	91.07	1849.9...	1.3	926...	97....	11750	7	453	470	•
5	K.RVVINKENTTIIDGAGVQADIEAR.V	•	83.50	2581.3...	1.6	861...	96....	11558	3	322	345	
6	R.GIDKATVAIVAQLKELAKPC(+57.02)ADTK.A	•	82.71	2539.4...	1.4	847...	156....	21561	14	119	142	•
7	K.ENTTIIDGAGVQADIEAR.V	•	80.63	1871.9...	0.7	936...	93....	11101	5	328	345	
8	R.QIVANAGDEPSVVDKVK.Q	•	79.29	1867.0...	1.0	934...	66....	6177	15	453	470	
9	R.NVLDKSFSGAPITTK.D	•	78.03	1588.8...	-1.2	795...	78....	8282	19	37	51	
10	K.MLVGVNVLADAVK.A	•	77.57	1327.7...	0.5	664...	137....	18668	2	16	28	
11	K.Q(-17.03)IEETTSYDREK.L	•	75.87	1595.6...	2.8	798...	38....	2076	3	352	364	•
12	R.NVLDKSFSGAPITTKDGVSAK.E	•	73.04	2245.2...	0.6	749...	101....	12555	9	37	58	
13	K.GDNEEQNVGIALLR.R	•	71.86	1526.7...	1.4	764...	97....	11697	6	431	444	

3.4.1. Protein Table

Each row of the table is a group of proteins that share the same set (or a subset) of identified peptides. A dark blue node at the beginning of the row indicates that the group has multiple proteins. To expand the group, click the "+" button at the left. The table's columns are:

- **Accession:** The accession number of the protein entry in the database.
- **-10lgP:** Protein confidence score.
- **Coverage:** The number of amino acids spanned by the assigned peptides divided by the protein length $\times 100$. The blue blocks indicate assigned peptides at particular positions in the protein. Darker blocks indicate high-confidence (passing the user's filtration score threshold) peptides.
- **#Peptides:** The number of high-confidence peptides assigned to the protein.
- **#Unique:** The number of high-confidence peptides that are unique to the group of proteins (not found in other protein groups).
- **Mass:** The calculated mass of this protein.
- **Description:** The part of the protein's header information as parsed from the database.

Note

For the counting of #Peptides and #Unique, two peptides with the same starting and ending positions in the protein are counted as one, regardless of their PTM forms. This seemingly counter-intuitive counting rule is to follow the MCP (Molecular & Cellular Proteomics) guideline.

3.4.2. Peptides

The "Peptides" tab displays the supporting peptides assigned to the protein. The table is almost the same as the peptide table in the Peptide View, except that three additional columns are added:

- Start: the start position of the peptide in the protein.
- End: the end position (inclusive) of the peptide in the protein.
- Unique: whether this peptide is unique to the current protein group.

Additionally, the peptides from the protein and below the user-specific score threshold are also displayed in the table, but in grey color. Although their correctness is questionable, they are worth-examining once an interesting protein is confidently identified by the other high-confidence peptides.

3.4.3. Coverage

The coverage tab contains the protein sequence with the supporting peptide sequences represented by the underlining blue bars. Placing the cursor over a blue bar shows some more details about the corresponding peptide, including the sequence with PTM information, score, scan number and number of spectra reporting this peptide.

The screenshot shows the 'Coverage' tab in the PEAKS DB interface. At the top, there are three tabs: 'Peptides', 'Coverage', and 'Tool Box'. Below the tabs, the protein sequence for 'P02769|ALBU_BOVIN Serum albumin OS=Bos taurus GN=ALB PE=' is displayed. The sequence is: 1 SV=4 MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA FSQYLQQCPF. Below the sequence, several peptide sequences are highlighted with blue bars, indicating their coverage. A tooltip is visible over one of the bars, displaying the peptide sequence 'FKDLGEEHFK', PTM '29.80', scan number '1140', and the number of spectra '1'.

3.4.4. Tool Box

Two tools are provided to

- search the current protein sequence by NCBI BLAST
- search the protein accession/name by NCBI Entrez

3.5. De Novo Only View

The de novo only view displays high-confidence *de novo* sequences whose corresponding spectra only have low confidence database matches. High-confidence *de novo* sequence means the TLC and ALC score of the sequence passes the corresponding user-specified score threshold. Low confidence database match means the peptide -10lgP score is below the user-specified score threshold. The table is identical to the peptide table in a de novo sequencing result node. Refer to Section 3.2, "De Novo Peptide View" for how to use it.

4. Filter PEAKS DB Result

Through the summary view, users can effectively filter the database search results to ensure the result quality by specifying score thresholds for peptides, proteins, and *de novo* sequences.

Note

Whenever you changed a score threshold, the "Apply" button changes color to remind you to apply the filter by clicking it.

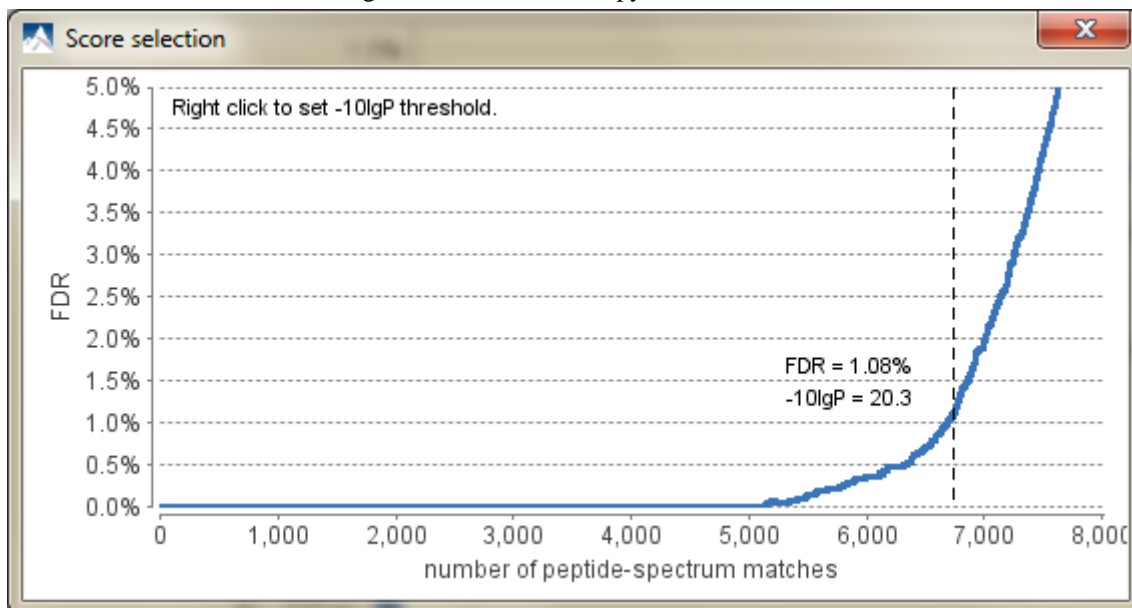
Start Page x PEAKS 5 [19-May-11 10:10] x

Summary

Peptides -10lgP \geq 15 FDR Proteins -10lgP \geq 20 and \geq 0 unique peptides

De novo Only TLC \geq 3 and ALC (%) \geq 50 and peptide -10lgP \leq 8 Apply Filters Export Notes

Peptides. The threshold here will affect both peptide and protein views and therefore has to be chosen with caution (For peptide view, only peptides with -10lgP score above the threshold will be kept in the table. For protein view, the number of supporting and unique peptides is based on the filtered peptide results). If the "FDR estimation" option was turned on in the PEAKS DB search parameters, the score threshold for peptides can be easily chosen by clicking the FDR button. An FDR curve will popup. Move the mouse cursor along the curve. When the desired FDR is reached, right click and select "Copy score threshold".



If the "FDR estimation" was turned off, then an empirical threshold is needed. Usually a score of 15 is a good choice. At -10lgP=15, the equivalent P value is 0.01.

Note

P-value and FDR are two very different concepts. In PEAKS DB search, P-value is defined as the probability that a false identification in the current search achieves the same or better matching score. A 1% P-value does not automatically correspond to a 1% FDR. For more details, please see http://peaks.bioinfor.com/doc/peaks53/peaks_db_paper.php

Proteins. Empirical thresholds for protein -10lgP score and the number of unique peptides are needed here. A protein score of 20 or higher is recommended. The unique peptides are the high-confidence peptides that are unique to the group of proteins (not found in other protein groups). To achieve confident results, at least one unique peptide is needed for a protein group. The thresholds here do not affect the peptide and *de novo* only views.

De novo Only. The minimum TLC and ALC *de novo* sequencing scores and the maximum peptide -10lgP score for a peptide to possibly appear in the *de novo* only view. *De novo* sequences with TLC and ALC scores above the threshold and whose corresponding spectra only have database matches with -10lgP score below the threshold will be shown in the *De novo* Only view. The thresholds here do not affect the peptide and protein views. Again, empirical thresholds are needed. A peptide -10lgP score of 8 is recommended. Recall that roughly TLC is the estimated number of correct amino acids and ALC is the estimated percentage of correct amino acids in the *de novo* sequence. Check Section 1, "Overview" for more explanation about TLC and ALC.

5. Export PEAKS DB Results for Publication

The "Export" button at the top of the summary view allows exporting of the filtered results into a list of html (text file readable in a web browser) and csv (text file readable in Excel) files. This provides the opportunity to supplement the results in a publication, or put up the results on your website.

Note

A better way to share results is to share the whole PEAKS project directory. It can be opened in our free PEAKS Viewer (<http://www.bioinfor.com/peaks/viewer/index.php>) that has the same GUI as PEAKS Studio.

Note

Labs with in-house software can easily make use of the csv files in their own analysis workflow.

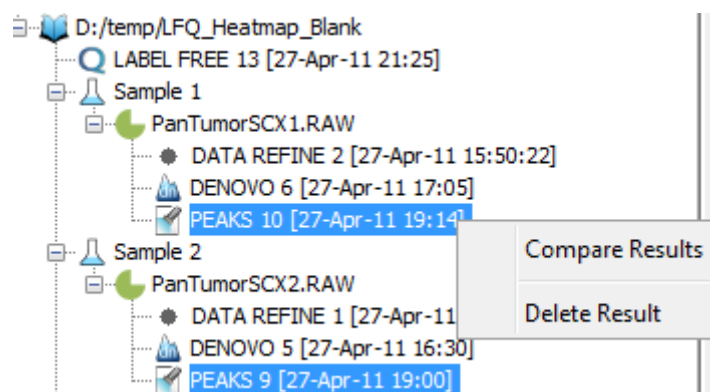
To export the filtered results:

1. Click the "Export" button at the top of the summary view. A file chooser appears.
2. Choose the location and directory name where you want to put the exported files. Click OK.

This will create a collection of files in the target directory. Refer to Section 3, "Export PEAKS DB Result" for details.

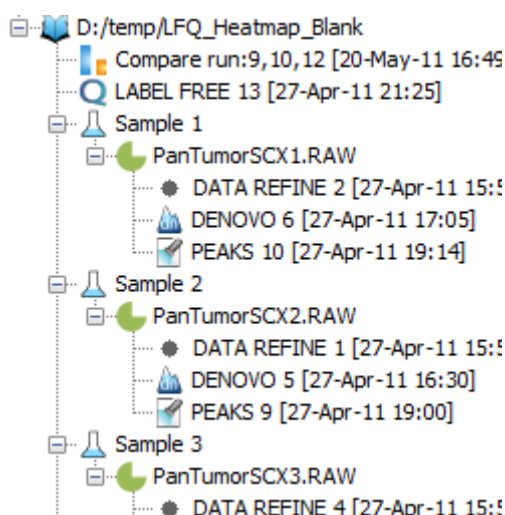
6. Comparison of PEAKS DB Results

In PEAKS 5.3, we support comparisons on at most three PEAKS DB results (including filtered results) in one project. To do such a comparison, select those PEAKS DB nodes and right click. Click on "Compare Results" and the comparison will be done automatically.



6.1. Comparison Result

After comparison is finished, a comparison node will be added to the project as shown in the following picture.



The result panel will be opened automatically after completing the comparison. Since the comparison run is done on the fly it won't be saved, it is suggested to export the results before closing the result panel. The details of exporting will be given in the next subsection.

The result consists of three parts: peptide comparison, protein comparison and statistical charts. Below is an outline of each.

6.2. Peptide Comparison

All the peptides identified in up to three PEAKS DB searches are displayed in the table. We show m/z, retention time, peptide score, charge and whether there are multiple hits for each peptide. The 'cover map' is a quick graphical illustration of the presence of the given peptide in one or both PEAKS DB results. A solid icon indicates a successful detection of the peptide.

You can also select to show only the common peptides of those PEAKS DB results, or the unique peptides of each PEAKS DB result by changing the display settings at the bottom of the panel.

We also provide filters on the peptide comparison results. After inputting the PEAKS score threshold on each PEAKS DB result and clicking the "Apply Threshold" button, those peptides below the threshold will be filtered out.

The following screenshot is a typical peptide comparison result:

Peptide Sequence	Cover Map	PEAKS 9 [27-Apr-11 19:00]					PEAKS 10 [27-Apr-11 19:14]					PEAKS 12 [27-Apr-11 19:54]				
		M/Z	RT	-10logP	Z	MULTI-HIT	M/Z	RT	-10logP	Z	MULTI-HIT	M/Z	RT	-10logP	Z	MULTI-HIT
VFLNVIR	■ ■ ■						495.29135	39.7717	44.76	2	true	495.291	38.4524	36.85	2	true
SPFEVVDK	■ ■ ■											542.26984	28.4963	47.67	2	true
FSMVVQGVK	■ ■ ■											611.82666	35.0541	48.84	2	true
ALDTMNFVVK	■ ■ ■											633.822	43.2472	41.94	2	false
EIVDSVLPVLDIK	■ ■ ■	865.5011	82.2145	60.6	2	false						865.50085	82.1928	61.86	2	true
AEELGLPILVLR	■ ■ ■	690.4142	69.7953	47.29	2	false						690.4143	68.3736	49.89	2	false
GAGTGGGLAIVEGPSEAK	■ ■ ■	785.9052	26.8409	72.26	2	true	785.90485	26.8903	85.12	2	true	785.90533	25.4232	71.49	2	false
ISVYNEATGGK	■ ■ ■	651.3205	19.51	35.03	2	false	651.3209	19.5598	44.54	2	false	651.3207	18.5345	64.55	2	true
LQTSSVLVSGLR	■ ■ ■	630.368	33.1726	36.7	2	false	630.3677	33.0669	45.33	2	false	630.3655	31.5186	39.08	2	false
VELAPLSPVQPVK	■ ■ ■	760.9248	47.423	48.51	2	false	760.92554	47.5334	48.38	2	true	760.9249	46.4147	55.76	2	true
ETDLLDSSLVSIFGNR	■ ■ ■						953.9895	73.1017	66.44	2	false	953.98987	73.3267	71.69	2	true
GYISPIPIITSK	■ ■ ■											695.3538	37.3655	53.56	2	false
ELADNRSNNEGK	■ ■ ■	716.8508	36.088	16.28	2	true	716.8515	35.6881	13.25	2	true	716.8505	35.3759	15.28	2	true
LTILSEER	■ ■ ■											473.7818	16.3971	32.03	2	false
ILLEGESR	■ ■ ■						466.73563	8.4792	32.5	2	true	466.7362	8.1263	46.34	2	true
PSEAEVPIPIK	■ ■ ■						650.86554	49.3616	19.02	2	false					
IMAKSTLK	■ ■ ■											502.81247	41.7395	12.43	2	false
DLTNVQLQK	■ ■ ■											586.80884	17.3177	21.29	2	false
LAADIDASATNESSWNGR	■ ■ ■											967.94336	53.1444	5.04	2	false
IIAFVGSFVEDIEK	■ ■ ■	759.3939	34.552	23.71	2	true	759.39325	34.4955	30.52	2	true					
SLGEISALTSK	■ ■ ■	553.3064	30.595	55.37	2	true	553.30646	30.5446	63.31	2	true	553.3069	28.9616	55.37	2	true
PLPLFDR	■ ■ ■						454.2539	47.8436	24.63	2	true					
VESLQEEIAFLK	■ ■ ■	703.379	53.16	54.9	2	true	703.38	52.7723	60.77	2	true	703.38043	51.3037	56.8	2	true
VATLNSEESDPTTK	■ ■ ■	890.4166	16.7659	64	2	false										
VSVELTSLFK	■ ■ ■											618.84375	42.7779	36.53	2	false
MGVEIETISPGGR	■ ■ ■											487.5689	12.8798	12.37	3	false
TFQQLTPDESK	■ ■ ■	640.8053	19.1792	39.87	2	true	640.8045	19.2152	43.74	2	true	640.80743	18.2499	33.24	2	false
VVSEDFLQDVSAATK	■ ■ ■	812.90485	42.137	69.46	2	true						812.9061	41.0045	59.46	2	false
ILLADOMGLK	■ ■ ■											573.31366	40.4774	5.31	2	false
ISDFEPIK	■ ■ ■											499.7619	46.8365	38.74	2	true
TALDAGIASLTTAEVIVTEIPKEEK	■ ■ ■											956.86176	83.1193	45.79	3	false
AVLVLEPQTK(+15.99)DSVR	■ ■ ■	809.41	34.4498	57.27	2	true	809.4083	34.5133	69.37	2	false					
TLQQLAWVEHQVIR	■ ■ ■	638.6531	64.9527	24.41	3	true	638.6536	64.3657	19.18	3	true					
AEQPDVAIDLK	■ ■ ■	605.8175	28.2688	63.5	2	true	605.8171	28.0168	66.79	2	false	605.8166	26.8709	60.6	2	false
WPSKPEAKSIPVSK	■ ■ ■	570.3503	59.1537	12.75	3	true										
IVFQSVSR	■ ■ ■	453.2843	35.105	14.74	2	false										
TFQQLSNIK	■ ■ ■											523.77545	9.944	38.51	2	false
ADLNILGTIAK	■ ■ ■	621.8556	39.11	58.2	2	true	621.8543	39.0524	58.05	2	true	621.8541	37.9902	56.72	2	true

6.3. Protein Comparison

The top proteins identified in the PEAKS DB results are displayed in the table. The display setting, score filter and cover map function the same as in the peptide comparison frame. The following information is also displayed for each protein:

Score: PEAKS protein score

#Spec: the number of spectrum on which this protein has been detected.

#Pep: the number of supporting peptides of the protein

#Uniq: the number of unique peptides of the protein

%Spec: the ratio of detected peptides to the theoretical numbers

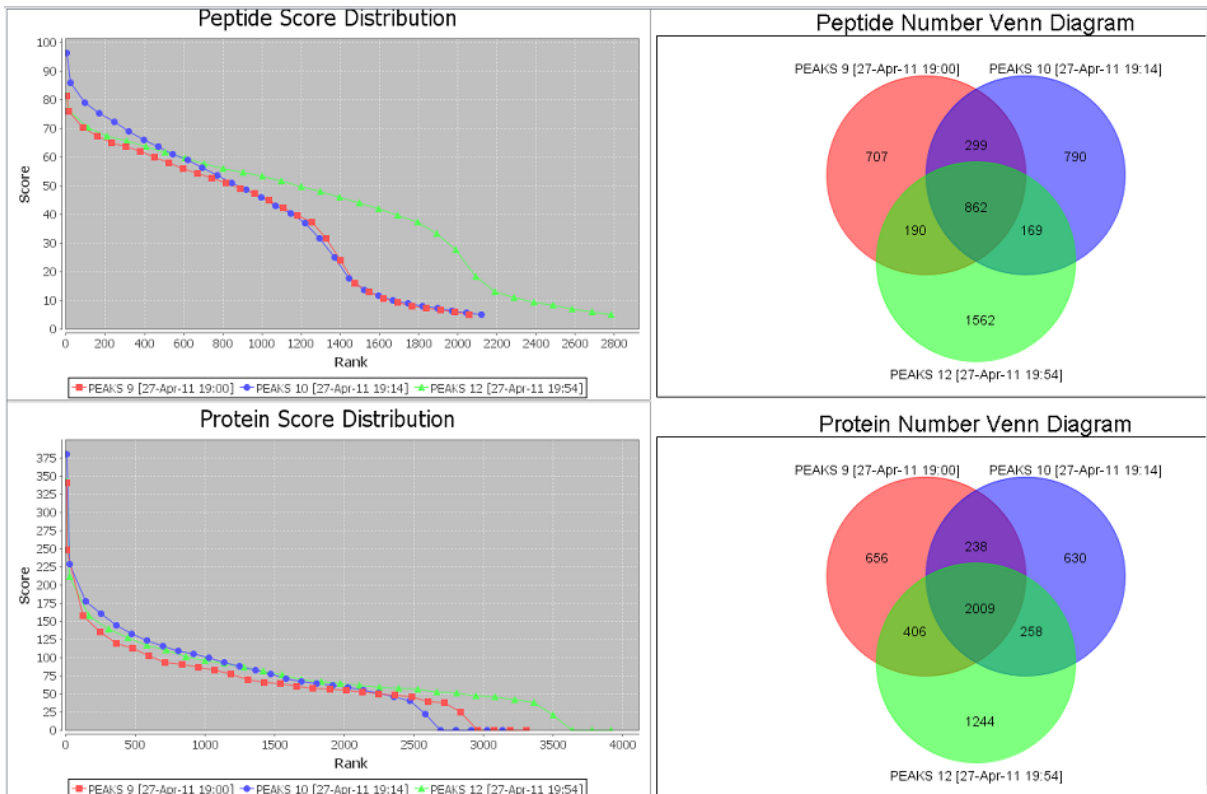
%Cov: the peptide coverage of the protein

The following screenshot is a typical results tab for protein comparisons:

Protein_ID	Cover Map	PEAKS 9 [27-Apr-11 19:00]					PEAKS 10 [27-Apr-11 19:14]					PEAKS 12 [27-Apr-11 19:54]				
		Score	#Spec	#Pep	#Uniq	%Spec	Score	#Spec	#Pep	#Uniq	%Spec	Score	#Spec	#Pep	#Uniq	%Spec
P42656[RA24_SCHPO]												64.13	8	2	0	6.9
P0C736[LRFN1_RAT]							0	1	1	1	1.41	1.83				
Q684U9[PRDX1_MYOLU]		125.49	16	6	5	26.09	131.11	16	6	5	26.09	139.03	25	7	0	30.43
Q9NY65[TBAB_HUMAN]		90.82	6	2	0	5	7.13	133.01	6	3	0	7.5	11.8	5	3	0
A2A3N6[IPSL_HUMAN]		63.27	1	1	1	1.2	1.97									
Q758U6[VID21_NEUCR]							0	1	1	1	0.62	0.69				
Q85QH5[ADT2_BOVIN]		53.64	3	1	1	2.38	4.36	81.82	4	2	1	4.76	7.38	51.5	3	1
P10096[G3P_BOVIN]		59.15	2	1	1	2.78	4.2	118.76	6	2	2	5.56	10.51			
Q52QH4[NCCL1_CHECK]												51.43	2	1	1	2.22
P30084[IECHM_HUMAN]												19.87	2	2	2	6.06
P35998[IRIS7_HUMAN]		47.77	2	2	2	6.06	13.45	37.47	1	1	1	3.03	6.21			
A1J811[IRNFC_YERE8]												48.99	2	2	2	3.33
P24724[HP90_THEPA]		56.02	2	1	0	0.92	1.94	57.79	1	1	1	1.18	1.58			
Q8FD23[CH60_XANCP]								0	2	2	2	3.23	5.68	83.52	3	2
A44W89[IRAK_RHOS5]																
A76DF3[IR54_SOLS1]								0	1	1	1	0.97	2.28	65.32	5	1
Q8CGP0[H2B3B_MOUSE]		89.94	10	2	1	7.14	19.05									
P09103[PDIA1_MOUSE]		138.91	12	5	5	7.94	12.38	153.41	20	5	5	7.94	11	155.03	22	7
Q28DR4[H4_XENTR]		50.05	5	1	1	4	11.65	104.49	10	3	3	12	31.07	77.91	4	2
Q77819[ROCK1_RABIT]														40.73	1	1
Q25538[1433_NEOCA]														64.13	8	2
Q8NDF8[PAPOS_HUMAN]		16.93	2	1	1	1.52	3.5									
Q6VAF4[ITB89_GOSHI]								55.2	1	1	0	2.7	3.15			
P22492[HIT_HUMAN]		65.87	4	1	1	1.92	5.31	73.17	3	1	1	1.92	5.31	65.17	4	1
P12303[ITTHY_SHEEP]														0	1	1
Q92597[INDG1_HUMAN]		63.45	1	1	1	3.7	3.55	67.78	2	2	2	7.41	7.61	64.4	1	1
Q00799[BP2_PLAIB]		0	1	1	1	0.23	0.59									
Q6494[PSME1_PIG]		124.48	8	4	4	10	20.48	119.77	9	3	3	7.5	15.26	137.5	16	5
P13491[LDHA_RABIT]								39.91	2	1	1	2.63	2.71	43.13	2	2
P63112[HBA_PAPCY]		55.93	2	1	1	6.25	8.45	62.37	2	1	1	6.25	8.45	85.95	7	2
Q3U962[COSA2_MOUSE]														69.02	3	2
P00739[PTIR_HUMAN]		59.47	3	1	1	2.86	3.45	72.06	5	2	2	5.71	7.18	83.49	6	2
P08814[PTMS_BOVIN]		51.66	5	1	1	5.88	10.78	66.86	7	1	1	5.88	10.78	51.35	3	1
Q5R887[MIL3_PONAB]														57.09	3	1
Q4QU6[ISPF30_RAT]		36.46	2	1	1	2.86	4.2									
Q43103[ISD2_USTMA]		0	1	1	1	0.26	0.32									
A8YUJ3[ATPA_LACH4]		39.42	1	1	1	1.92	2.39									
Q7U82[IRPOBC_HELPH]								0	1	1	1	0.26	0.31			

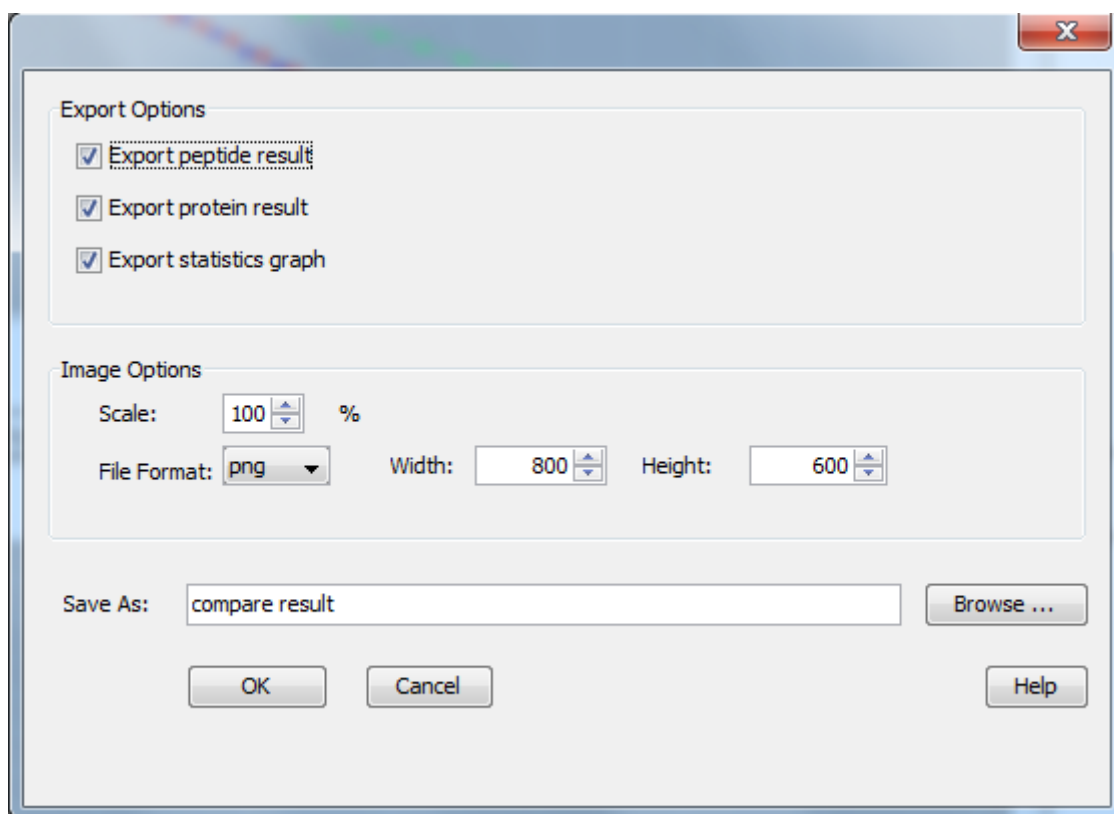
6.4. Statistical Charts

PEAKS 5.3 provides a number of statistical charts which are easily exported for use in publications. The peptide score distribution, protein score distribution, peptide number venn diagram and protein number venn diagram help users to validate their results.



6.5. Exporting Comparison Results

To export the comparison results of PEAKS DB searches, please right click on the comparison run node and choose to export to Excel file. Here you can choose image quality and filter the content you want to export.




Chapter 10. Combining Multiple Database Search Engines with PEAKS inChorus

1. PEAKS inChorus Overview

It is well-recognized that properly combining the results from different database search engines can enhance the accuracy and sensitivity of peptide identifications. PEAKS inChorus is such a tool to invoke or import the results of the database search engines SEQUEST (v27 rev12), Mascot (v2.3), X!Tandem (v2010.12.01.1) and OMSSA (v2.1.8). PEAKS inChorus uses a probabilistic model to combine multiple engines' results.

We assume that a reader is familiar with PEAKS DB (Chapter 9, *Protein Database Search with PEAKS DB*) before reading this chapter. The use of this function is outlined in the following overview. Details of each step can be found in later sections of this chapter.

1. Select a project node or a sample node. Click the PEAKS inChorus button  on the tool bar.

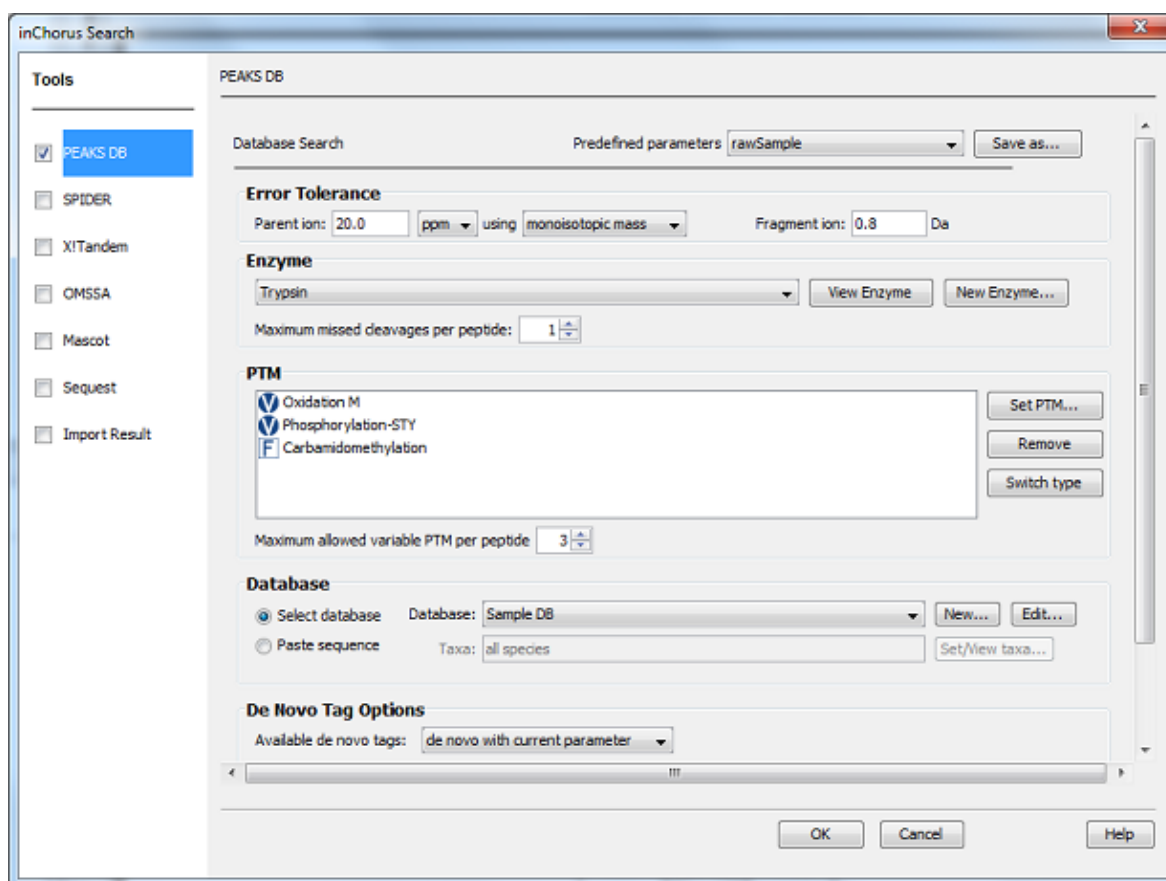


2. Check each engine that you would like to use at the left column of the parameter setting panel. Specify the search parameters for each engine in the right side of each panel. Each engine's parameter setting interface in PEAKS is kept very similar to their native interface. Please refer to third party software's user manuals for how to use them. For PEAKS DB, refer to Section 2, "Set PEAKS DB Parameters".

If you already have a search engine's result in a separate file or opened in current PEAKS project, check the "Import Result" at the bottom of the engine list.

Important

The other search engine's result should be based on the same refined data node in order to do inChorus.



Note

If your data is not refined by PEAKS yet, you will be asked to specify the data refinement parameters first and click next.

3. Wait for the analysis to complete. A new result node will appear in the project tree. Double click the node to open the result file.
4. The result presentation is similar to the PEAKS DB result, with additional information to show which peptide is identified by which engine(s).

2. Understanding PEAKS inChorus Result


The inChorus result is displayed in a very similar format of the PEAKS DB result (Section 3, “Understanding PEAKS DB Result”). This section only highlights the differences.

Peptide Score. The first noticeable difference is that the inChorus peptide score is not the $-10\lg P$ score used in PEAKS DB. Instead, a percentage confidence score is used to reflect the probability that this peptide-spectrum match is correct. The percentage score is calculated in accordance with the empirical calculation used in PeptideProphetTM (Keller et al. Anal. Chem. 2002, 74:5383-92).

1. The method of PeptideProphetTM is applied to each engine's result to estimate the probability of correctness for each peptide identification.
2. If a peptide-spectrum match is identified by multiple engines, the scores for all those engines are added up with a weighted sum.
3. The weighted sum scores of all peptides are converted to a probability by using the PeptideProphetTM method again.

Protein Score. The protein score is also a percentage score. It's calculated as follows. The scores of the peptides from a protein are added up by a weighted sum. Then the PeptideProphetTM method is applied to the weighted sum scores of all proteins to convert to a probability score.

Engines Icon. For each peptide in the peptide view, an "engines" icon is displayed to show the engine(s) that identified the peptide. Each engine is represented by a letter code and the block background color.

Start Page x DATA REFINE S x INCHORUS 27 [27-May-11 13:35] x														
1 - 1000 of 1784														
Peptide	Scan	m/z	RT	Mass	ppm	#Spec	Score (%)	Engines	P...	Se...	Ma...	XIT...	O...	PTM
1 SHC(+57.02)IAEVEN(+98)DEMPADLP... 2 HLNVSSTSGSIDMVDSPQLADEVSASLAK 3 FFESFGDLSTPDVIMGNPK 4 AQPQVQVAEGSEPDGFWEALGGK 5 EQLQDMGLVDLFSPEK 6 KVPQVSTPTLVEVSR 7 SC(+57.02)HLPWASGLETLDSLGGVLEASGY... 8 DLLFKDSAHGFLKVPVR 9 LVDSKINFDDYMK 10 YKSDSDVPMVLGNK 11 LADFALLC(+57.02)LDGK	F5:965 F7:1... F3:835 F1:868 F1:976 F9:513 F9:1... F12:685 F5:521 F5:552 F2:810	992.5... 1081.8... 1030.0... 1136.5... 924.9... 820.5... 1188.2... 970.5... 737.8... 808.4... 668.5...	27.... 29.... 23.... 23.... 26.... 13.... 44.... 19.... 13.... 14.... 23....	2974.3... 3242.5... 2057.9... 2271.0... 1847.8... 1638.9... 3560.7... 1939.0... 1473.6... 1614.8... 1334.6...	12... -3... 9... -1... -1... 10... 28... 2... 2... 3... 28...	10 11 10 8 5 18 40 3 5 9 7	99.95 99.96 99.97 99.97 99.97 99.97 99.97 100.00 99.98 99.98 99.97		31.35 53.63 61.27 68.33 63.14 54.25 52.23 55.53 48.61 53.98 52.30	4.94 3.84 - - - - - - - - -	46.40 82.44 102.35 87.31 112.54 70.58 81.33 55.79 65.03 81.27 81.31	1.5E-3 1.2E-4 1.2E-5 1.9E-7 4.5E-5 6.2E-5 2E-5 1.12E-9 5.2E-3 1.6E-3 1.7E-3	5.68E-10 2.48E-10 2.81E-9 5.16E-9 1.43E-6 2.79E-5 1.12E-9 1.22E-4 2.54E-6 1.19E-8 2.69E-10	- - - 63.64 60.27 50.35 35.65 52.93 49.28 51.08 44.22

- The **blue background** means the engine identified the peptide with high-confidence (above the engine's own filtration score threshold. See Section 3, "Filter PEAKS inChorus Result").
- The **white background** means the engine identified the peptide with low-confidence (below the engine's own filtration score threshold).
- A **dash symbol** means the engine did not identify the peptide.

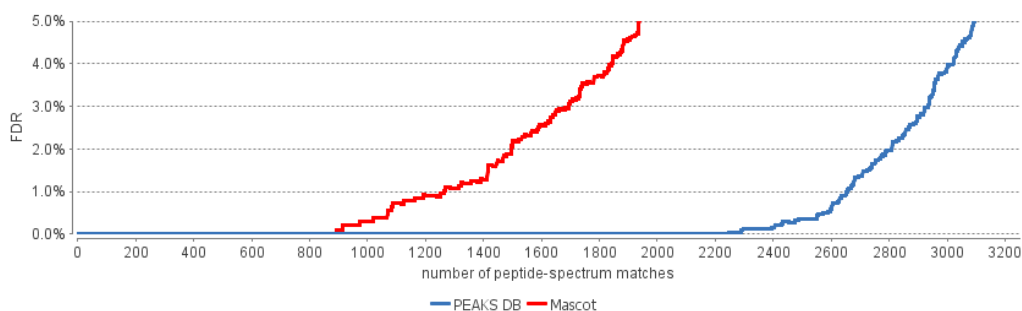
Different engines are coded by different letters, as follows:

- P: PEAKS DB
- M: Mascot
- S: Sequest
- X: X!Tandem
- O: OMSSA
- R: SPIDER

Individual Engine's Score. In the peptide view, each engine's own score is displayed. A dash symbol means the peptide is not found by the engine.

FDR curves. If MascotTM is part of the inChorus result, and decoy validation was chosen in the Mascot search parameters, then Mascot's FDR curve is displayed together with PEAKS DB in the summary view.

Figure 1: The false discovery rate (FDR) curve. X-axis is the number of PSMs (peptide-spectrum matches) being kept and the y-axis is the FDR.



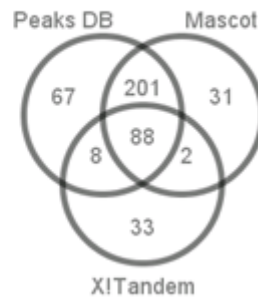
Venn Diagram. A Venn diagram shows the number of peptides identified by each possible combination of engines.

Note

These numbers do not distinguish high or low confidence for each engine's score. For example, if a peptide got a high inChorus score but low scores in every individual engine, the peptide is still counted

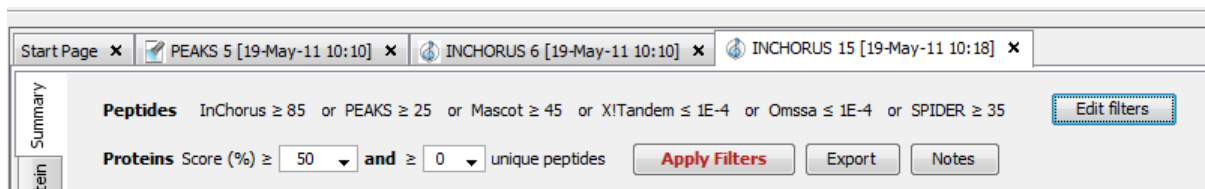
in the intersection of all engines in the Venn diagram. For this reason, this Venn diagram is NOT the best way to compare different engines' performance.

Figure 2: Venn diagram for confident PSMs (up to three selected search engines). ?

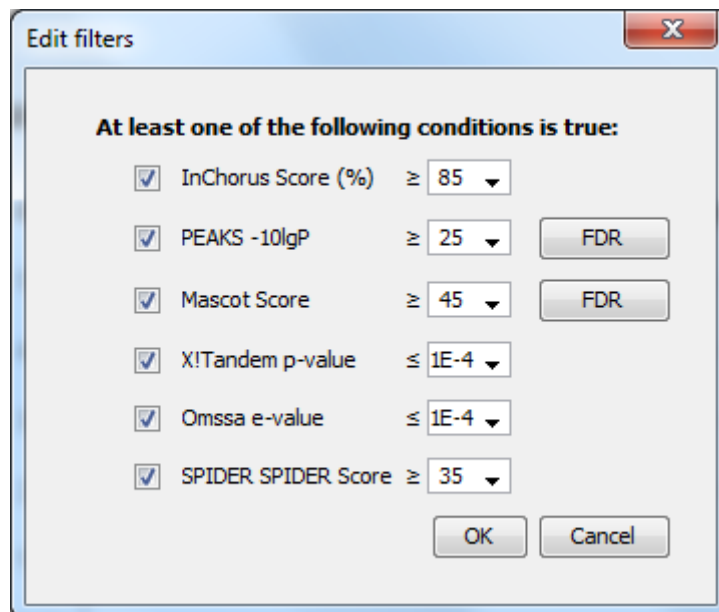


3. Filter PEAKS inChorus Result

The filtration differs from the PEAKS DB filtration mostly by the peptide filters. The peptides can be filtered by the inChorus score and each individual engine's score.



Click the "Edit filters" button in the summary view to specify the peptide filtration rules. A peptide is kept as long as one of the specified rules is satisfied.



Part III. PEAKS Q - Protein Quantification

Table of Contents

11. PEAKS Q - MS Level	88
1. Overview	88
2. Setting Parameters	88
3. Understanding the Result	90
3.1. Summary View	90
3.2. Protein View	90
3.3. Peptide View	91
12. PEAKS Q - MS/MS Level	92
1. Overview	92
2. Setting Parameters	92
3. Understanding the Result	93
3.1. Summary View	94
3.2. Protein View	94
3.3. Peptide View	94
13. PEAKS Q - Label Free	96
1. Overview	96
2. Setting Parameters	96
3. Understanding LFQ Result	98
3.1. Summary View	98
3.2. Protein View	99
3.2.1. Extracted Ion Chromatogram	100
3.2.2. Heat Map	100
3.2.3. 3D View	101
3.2.4. MS2 Annotation	101
3.2.5. Isotope	102
4. Filter LFQ Result	102
5. Export Quantification Result to Other Formats	102
6. Replicate Analysis in LFQ	103
6.1. Assign Replicate Number to a Sample	103
6.2. Run Replicate Analysis	105
6.3. Understand Replicate Analysis Results	106
6.4. Export Replicate Analysis Result	108

Chapter 11. Quantification with Labels at MS Level (e.g. SILAC and ICAT)

1. Overview


Quantification with isotope labels at the MS-level is one of the three quantification modes that are supported by the optional PEAKS Q module of PEAKS Studio. In this mode, the isotope labels with different mass values are introduced to two or more samples. The samples are then analyzed together in an LC-MS/MS experiment. The same peptide from different samples is recognized by a set of precursor ion peaks with similar retention time and mass differences within the retention time window and error tolerance set by the user. The ratio is calculated from the intensities of those peaks. PEAKS Q supports user-defined labels and commercial quantification labels.

The quantification analysis is based on a PEAKS DB database search result (See Chapter 9, *Protein Database Search with PEAKS DB*). Ensure that you specified the isotope labels as PTMs when you performed the database search. After database search is done, follow these steps:

1. Select a PEAKS DB result node in the project tree. Click the PEAKS Quantification tool bar icon .


Note

Refer to Section 2, “Quantification Workflow” for how to conduct PEAKS DB and quantification in a single workflow.

2. Select the quantification protocol 'Label at MS level' and specify the PEAKS quantification parameters in the dialog and click “OK”.
3. Wait for the analysis to finish. A new quantification result node  will appear in the project tree. Double click the node to open the result file.

The quantification result display is similar to the PEAKS DB result, A ratio is added to each quantifiable peptide and protein.

2. Setting Parameters

The following parameter dialogue pops up when clicking the quantification tool bar icon . Select “Label at MS level eg. SILAC” from the left hand side.

Quantification

Tools

- ☐ Label at MS/MS level
eg. iTRAQ/TMT
- ☒ Label at MS level
eg. SILAC
- ☐ Label Free

SILAC silac-test Save as...

Basic Options

Mass Error Tolerance: Da Upper Bound of Precursor Charge:

Retention Time Range: min. Peptide Score Threshold (-10logP):

Label Options

Sample Name	Added Mass	Residues	Labelling Efficiency (%)
No label		0.0R	100.0
No label		0.0K	100.0
Light		4.03K	100.0
Light		6.02R	100.0
Heavy		8.01K	100.0
Heavy		10.01R	100.0

Add Label Delete Label

OK Cancel Help


The following parameters are available in the “Basic Options” section of the quantification window:

- **Mass Error Tolerance:** This parameter is used to locate the precursor ion peak group of an identified peptide in the survey scans. Since in a SILAC or ICAT experiment, we are usually dealing with non-centroided survey scans, mass error tolerance should be set a little wider than the parent ion error tolerance used in the PEAKS DB database search.
- **Upper Bound Precursor Charge:** The precursor ion of an identified peptide may have sibling ions of different charge states. Only those sibling ions with a charge less than the upper bound precursor charge will be considered for quantification of the identified peptide.
- **Retention Time Range:** The retention time range is the maximum elution time range that is considered for the quantification of an identified peptide.
- **Peptide Score Threshold (-10lgP):** Only peptides with a score above this threshold are used to quantify the identified proteins.

The labels used in the experiment are defined in the “Label Options” section. To add a label click on the “Add Label” button. To delete a label from the list, select the label and press the “Delete Label” button. Each label is defined by sample name, added mass, target residue, and labeling efficiency. If one sample has multiple labels with different mass shifts, add multiple labels with the same sample name. These labels will contribute to the same number in the ratio.

Clicking the “Save As” button at the top right allows the user to save parameters for ease of use when regularly performing quantification with the same parameters.

3. Understanding the Result

Once completed, the protein quantification result will be displayed in the quantification node  in the project tree. Double click on this node to open the result that contains three views: Summary view, Protein view and Peptide view. The "Summary" view tab will appear by default.

3.1. Summary View

The MS level quantification results are summarized in one page in the "Summary" view. In the heatmap, proteins are clustered into a tree structure. Move the mouse to the tree to select a cluster and left click to show the variation trend chart for that cluster.

1. Heatmap View



Cell colour represents the log₂ ratio to the Control Sample (Control Sample is marked with (*))

2. Notes

3. Result Statistics

Table 1. Statistics of data and unfiltered result.

of MS Scans 5591
of MS/MS Scans 12443

Table 2. Result filtration parameters.

Protein fold change ≥ 5

4. Other Information

Table 4. Search parameters.

Quantification Type: ICAT/SILAC
Quantification Mass Tolerance: 0.1Da
Quantification RT Range: 1.0min.
Upper Bound Charge: 4
Peptide Score Threshold:15.0
No Label: K (100.0%) R (100.0%)
Light: K(+4.03) (100.0%) R(+6.02) (100.0%)
Heavy: K(+8.01) (100.0%) R(+10.01) (100.0%)

Table 5. Instrument parameters.

Fractions: VPS4B1J.RAW
Ion Source: ESI (nano-spray)
Fragmentation Mode: CID, CAD, IRMPD (y and b ions)
MS Scan Mode: FT-ICR/Orbitrap
MS/MS Scan Mode: Linear Ion Trap

3.2. Protein View

The "Protein" view shows a list of proteins that are identified in the database search, together with their identified peptides in the window below. The quantification ratios for quantifiable proteins are displayed in the ratio columns with sample names incorporated into the header. (e.g. "Ratio Heavy:No Label") The ratio is calculated from the unique peptides of the protein. Proteins with no unique peptides will not be assigned a ratio. The SD is the standard deviation of the peptide ratios in the protein. The peptides of the selected protein, together with their ratios, are displayed at the bottom half of the protein view.

Summary		proteins in each group													accession contains		search		no results	
Protein	Peptide	Accession	Score (%)	-10lgP	Cov...	#P...	#Unique	Mass	Description	Ratio Light:No Label	Ratio Heavy:No Label	SD Light:No Label	SD Heavy:No Label	Mark						
Proteins	Peptides	P13637 AT1A3_HUMAN	99.2	155.42	20%	20	5	111748.461	Sodium/potassium-tra...	40.00	8.84	0.00	3.25	✓						
		tr B7Z3U6 B7Z3U6_HU...	99.2	148.70	15%	15	4	109550.250	ATPase, Na+/K+ tran...	-	0.39	0.00	0.12	✓						
		tr B3KR11 B3KR11_HU...	99.2	142.30	18%	17	6	103561.984	cDNA FLJ34298 fs, d...	-	0.18	1.22	0.18	✓						
		P23396 RS3_HUMAN	99.2	131.59	68%	19	18	26688.326	40S ribosomal protein ...	0.51	0.64	0.15	0.17	✓						
		P16403 H12_HUMAN	99.1	129.22	31%	14	2	21364.732	Histone H1.2 OS=Ho...	0.48	0.63	0.04	0.00	✓						
		P10412 H14_HUMAN	99.1	128.27	37%	14	1	21865.252	Histone H1.4 OS=Ho...	0.36	0.49	0.00	0.00	✓						
		P35908 K22E_HUMAN	99.2	123.08	23%	14	8	65865.328	Keratin, type II cytos...	0.39	-	0.15	0.00	✓						
		Q00610 CLH1_HUMAN	99.2	118.03	7%	11	10	191613.281	Clathrin heavy chain ...	0.27	0.39	0.13	0.03	✓						
		tr Q96IR1 Q96IR1_HU...	99.2	115.32	43%	13	13	27259.803	RPS4X protein (Fragm...	0.70	0.72	0.09	0.11	✓						
		P35527 K1C9_HUMAN	99.1	114.16	18%	12	8	62129.383	Keratin, type I cytosk...	6.35	-	0.00	0.00	✓						
		P61247 RS3A_HUMAN	99.2	112.46	48%	16	14	29944.910	40S ribosomal protein ...	0.76	0.75	0.10	0.17	✓						
		Q07955 RSF1_HUMAN	99.2	110.80	45%	14	9	27744.570	Serine/arginine-rich ...	0.73	0.78	0.19	0.15	✓						
		tr B4DG62 B4DG62_HU...	99.1	109.92	11%	10	7	102311.719	cDNA FLJ56506, highl...	0.55	0.60	0.10	0.00	✓						
		P13645 K1C10_HUMAN	99.2	109.63	19%	11	8	59510.629	Keratin, type I cytosk...	0.49	0.70	0.00	0.03	✓						
		O95782 AP2A1_HUMAN	99.2	108.29	13%	13	11	107554.867	AP-2 complex subunit ...	0.79	0.26	0.79	0.21	✓						
		tr B4DZRO B4DZRO_HU...	99.0	107.24	8%	6	4	97561.695	cDNA FLJ55529, highl...	-	0.64	0.00	0.00	✓						
		P12036 NFH_HUMAN	99.2	105.37	10%	11	8	112479.805	Neurofilament heavy ...	0.35	0.34	0.00	0.00	✓						
		P63244 GILP_HUMAN	99.1	104.82	28%	10	9	35076.738	Guanine nucleotide-bi...	0.59	0.68	0.34	0.12	✓						
		P62917 RLB_HUMAN	99.1	102.29	8%	8	7	28024.654	60S ribosomal protein ...	0.71	0.70	0.15	0.16	✓						
		P15880 RS2_HUMAN	99.1	102.20	28%	9	7	31324.432	40S ribosomal protein ...	0.70	0.78	0.08	0.08	✓						
		tr B4DG62 B4DG62_HU...	99.1	102.10	28%	10	7	30855.613	cDNA FLJ56544, highl...	0.55	0.60	0.10	0.00	✓						
Peptides		Coverage	Tool Box																	
Protein	Peptide	Unique	-10lgP	Mass	ppm	m/z	RT	Scan	#Spec	Light:No Label	Heavy:No Label	Start	End	PTM						
Proteins	Peptides	1	R.ELAEDGYSGVEIVR.V	•	44.47	1422.6626	27.5	712.36	66.418	9048	1	3.14	4.96	28	40					
		2	R.ELAEDGYSGVEIVR(+10.01).V	•	44.16	1432.6709	31.7	717.37	77.792	10787	2	0.33	0.56	28	40	•				
		3	K.GC(+57.02)EVVYSGK.L	•	43.70	933.4589	20.0	467.75	20.230	1877	3	0.56	0.73	133	141	•				
		4	R.GLC(+57.02)AIQAESLR.Y	•	42.89	1287.6605	28.2	644.86	77.926	10808	7	0.90	0.77	95	106	•				
		5	K.DELPTTPISEKQ.G	•	40.84	1469.7614	26.4	735.91	66.351	9037	3	0.41	0.46	215	227	•				
		6	K.K(+4.03)PLPDHVSIVEPK(+4.03).D	•	38.40	1465.8744	36.0	489.65	95.310	13428	9	0.32	0.41	202	214	•				
		7	R.GLC(+57.02)AIQAESLR(+10.01).Y	•	37.64	1297.6688	27.5	649.86	77.932	10809	5	0.90	0.77	95	106	•				
		8	K.GKPEPPAMPQVPVTA	•	37.07	1572.7970	38.1	787.44	58.736	7803	1	0.30	0.36	228	243	•				
		9	K.DELPTTPISEKQ(+8.01).G	•	36.66	1477.7755	34.3	739.92	66.381	9042	2	0.42	0.46	215	227	•				
		10	R.GLC(+57.02)AIQAESLR(+6.02).Y	•	35.79	1293.6805	27.9	647.87	79.935	11097	2	0.72	0.66	95	106	•				
		11	K.KPLPDHVSIVEPK.D	•	34.86	1457.8242	37.2	486.97	94.700	13333	12	0.42	0.39	202	214	•				
		12	K.GC(+57.02)EVVYSGK(+4.03).L	•	34.62	937.4841	22.5	469.76	20.032	1850	2	0.57	0.69	133	141	•				
		13	K.K(+8.01)PLPDHVSIVEPK(+8.01).D	•	33.73	1473.8525	35.1	492.31	94.882	13361	9	0.37	0.39	202	214	•				
		14	K.GC(+57.02)EVVYSGK(+8.01).L	•	33.06	941.4731	27.9	471.76	20.400	1899	3	0.57	0.83	133	141	•				
		15	R.TEIIILAYR(+10.01).T	•	31.87	1038.6312	27.6	520.34	84.300	11736	1	0.56	0.62	46	54	•				
		16	K.AELNEFLTR.E	•	29.77	1091.5610	44.1	546.81	69.164	9449	21	0.78	0.60	19	27	•				
		17	R.K(+8.01)FVADGIFK(+8.01).A	•	28.16	1039.6036	34.3	520.83	102.182	14426	10	0.58	0.92	10	18	•				
		18	R.KFVADGIFK.A	•	27.50	1023.5753	33.7	512.81	102.187	14427	7	0.57	0.92	10	18	•				
		19	K.AELNEFLTR(+6.02).E	•	27.03	1097.5812	29.0	549.81	69.662	9524	1	0.71	0.68	19	27	•				
		20	R.ELTAIVQK.R	•	26.85	886.5123	41.6	444.28	32.973	3665	3	0.56	0.63	68	75	•				
		21	R.ELAEDGYSGVEIVR(+6.02).V	•	25.77	1428.6827	31.8	715.37	77.804	10789	1	0.33	0.56	28	40	•				
		22	R.FIM(+15.99)ESGAK(+8.01).G	•	24.73	905.4408	27.2	453.74	18.563	1666	9	-	-	125	132	•				

3.3. Peptide View

The peptide view displays all the identifiable peptides and their quantification ratios. The interface is similar to the peptide table in a PEAKS DB result. The quantification ratios for quantifiable peptides are displayed in the ratio columns with sample names incorporated into the header (e.g. "Heavy: No Label").

Summary		scan = <input type="text" value="search"/> <input type="button" value="no results"/>												
Protein	Peptide	Peptide	-10lgP	Mass	ppm	m/z	RT	Scan	#Spec	Accession	Light:No Label	Heavy:No Label	PTM	
Protein	Peptide	1	GVGIISGEINTEYDIAAR	62.43	1828.9166	28.9	915.49	100.803	14229	31	P13637 AT1A3_HUMAN	-	-	
		2	NITFFSTNC(+57.02)VEGTAR	54.30	1715.7937	29.6	858.93	89.020	12439	1	P13637 AT1A3_HUMAN	-	-	●
		3	FSSSSSGYGGSSSR	54.15	1234.5214	25.5	618.28	19.989	1843	3	P35527 K1C9_HUMAN	-	-	
		4	SPQFTNENPLETR	53.34	1518.6951	28.0	760.38	66.497	9061	3	tr B7Z3U6 B7Z3U6_HUMAN	0.16	0.32	
		5	GENLVSM(+15.99)TEVGPPPK	52.50	1569.7709	31.6	785.92	54.029	7043	8	tr Q66K91 Q66K91_HUMAN	2.53	0.39	●
		6	FAAMATGATPIAGR	51.76	1202.6407	25.5	602.34	45.865	5724	1	tr C939K3 C939K3_HUMAN	3.80	0.66	
		7	GAAM(+15.99)YTAIVAYR	51.47	1124.5648	37.5	563.31	39.978	4752	2	tr B4DG62 B4DG62_HUMAN	0.30	0.19	●
		8	GFSSGSAIVSGGSR	51.18	1253.6000	42.5	627.83	36.825	4252	4	P35908 K22E_HUMAN	-	-	
		9	GENLVSMTEVGPPPK	51.06	1553.7759	31.8	777.92	68.331	9337	8	tr Q66K91 Q66K91_HUMAN	0.40	0.44	
		10	GQFSTDELVAEIVR(+6.02)	50.48	1712.8676	37.9	857.47	108.498	15364	3	Q00610 CLH1_HUMAN	-	-	●
		11	VLATTFDPLGGR	50.45	1408.7350	34.0	705.40	94.400	13286	5	tr B4DZRO B4DZRO_HUMAN	0.12	0.06	
		12	EDIYSGGGGGGSSR	50.35	1210.5214	28.9	606.29	21.733	2085	2	Q13151 ROAO_HUMAN	0.55	0.58	

Peptide Spectrum Match

Protein

Chapter 12. Quantification with Labels at MS/MS Level (e.g. iTRAQ and TMT)

1. Overview


Quantification with isotope labels at MS/MS level is one of the three quantification modes that are supported by the optional PEAKS Q module of PEAKS Studio. In this mode, isotope labels with the same mass are introduced to several samples. The samples are then analyzed together in an LC-MS/MS experiment. The same peptides from different samples will have the same precursor m/z and is fragmented together. However, in the MS/MS, labels from the different samples will produce different reporter ions; which can then be used to calculate the quantification ratio between samples. User-defined labels are also supported in PEAKS Q as well as commercial labels such as iTRAQ and TMT.

The quantification analysis is based on a PEAKS DB database search result (See Chapter 9, *Protein Database Search with PEAKS DB*). Ensure that you specified the isotope labels as PTMs when you performed the database search. After database search is done, follow these steps:

1. Select a PEAKS DB result node in the project tree. Click the PEAKS Quantification tool-bar icon .

Note

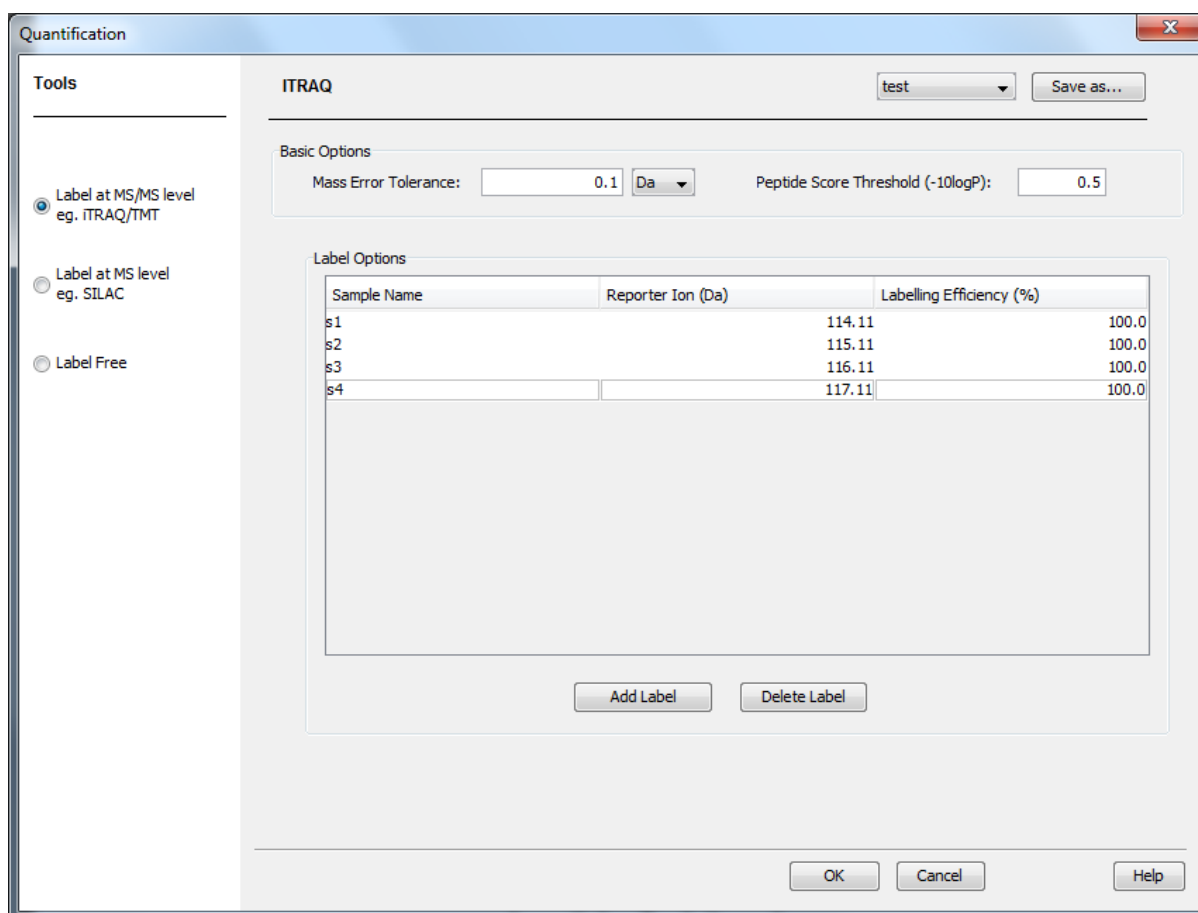
Refer to Section 2, “Quantification Workflow” for how to conduct PEAKS DB and quantification in a single workflow.

2. Select the quantification protocol as ‘Label at MS/MS level’ and specify the PEAKS quantification parameters in the dialog box on the right and click “OK”.
3. Wait for the analysis to finish. A new quantification result node  will appear at the project tree. Double click the node to open the result file.

The quantification results with labels are displayed in a format that is similar to the PEAKS DB result, except that a ratio is added to each quantifiable peptide and protein.

2. Setting Parameters

This is for quantification based on the relative intensities of fragment peaks at fixed m/z values within an MS/MS spectrum. Select “Label at MS/MS level eg. iTRAQ” from the left hand side under “Tools” heading in the quantification window. Enter the relevant MS/MS level labeling quantification parameters on the right hand side of the window.




The following parameters are available in the “Basic Options” section of the quantification window:

- **Mass Error Tolerance:** This parameter is used to locate the reporter ion peaks in the MS/MS spectrum. If the MS/MS spectrum is centroided, use the fragment ion tolerance set in the PEAKS DB database search. Otherwise the mass error tolerance should be set a little wider than the fragment ion error tolerance in the PEAKS DB database search.
- **Peptide Score Threshold (-10lgP):** Only those peptides with a score above this threshold are used to quantify the identified proteins.

The labels used in the experiment are configured in the “Label Options” section. To add label, click on the “Add Label” button. To delete a label from the list, select the label and press the “Delete Label” button. Each label is defined by the following parameters: sample name, reporter ion mass, and the labeling efficiency. The labeling efficiency defines rate at which of the chemical reaction add labels to the peptides. It is used to adjust the ratios calculated from the MS data.

Clicking the “Save As” button at the top right allows the user to save parameters for ease of use when regularly performing quantification with the same parameters.

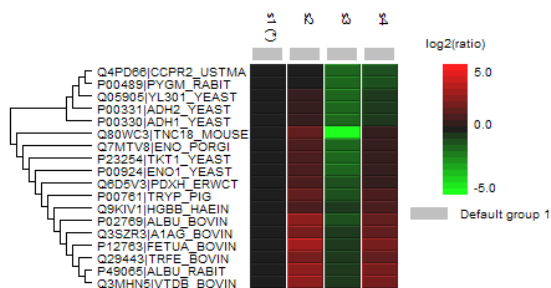
3. Understanding the Result

Once completed, the protein quantification result will be displayed in a quantification node  in the “Project View” panel. Double click on this node to open the result that contains three views: Summary view, Protein view and Peptide view. The “Summary” view tab will appear by default.

3.1. Summary View

The MS/MS labeled quantification results are summarized in one page in the "Summary" view. In the heatmap, proteins are clustered into a tree structure. Move the mouse onto the tree in order to select a cluster and left click to show the variation trend chart for that cluster.

1. Heatmap View



Cell colour represents the log₂ ratio to the Control Sample (Control Sample is marked with *)

2. Notes

3. Result Statistics

Table 1. Statistics of data and unfiltered result.

of MS Scans 11982
of MS/MS Scans 22175

Table 2. Result filtration parameters.

Protein fold change ≥ 2

4. Other Information

Table 4. Search parameters.

Quantification Type: ITRAQ
Quantification Mass Tolerance: 0.2Da
Quantification RT Range: 0.0min.
Upper Bound Charge: 0
Peptide Score Threshold: 20.0
Label 0: 114.11 (100.0%)
Label 1: 115.11 (100.0%)
Label 2: 116.11 (100.0%)
Label 3: 117.11 (100.0%)

Table 5. Instrument parameters.

Fractions: inj1.RAW, inj2.RAW, inj3.RAW, inj4.RAW, inj5.RAW
Ion Source: ESI(nano-spray)
Fragmentation Mode: CID, CAD, IRMPD (y and b ions)
MS Scan Mode: FT-ICR/Orbitrap
MS/MS Scan Mode: Linear Ion Trap

3.2. Protein View

The "Protein" view shows a list of proteins that are identified in the database search, together with their identified peptides in the window below. The quantification ratios of those quantifiable proteins are displayed in the ratio columns with reporter ion mass as the header eg. "Ratio 117:114". This is the ratio calculated from the unique peptides from the protein. The SD is the standard deviation of the peptide ratios in the protein. The peptides of the select protein, together with their ratios, are displayed at the bottom half of the protein view.

Start Page: X QUANTITATION 14 [22-Mar-11 13:03] X											
Protein	Accession	ID	Mass	Coverage Display	PEAKS(Score %)	p-value	Coverage(%)	Query matched	Description	Marked	Ratio 117.115:114.112
DB Search	P00722 BGAL_ECOLI	1	116351.600		98.8	1.99E2	4.3	0	Beta-galactosidase - Escherchia col.		4.31
	#DECOY#P0C0L4 CO4A_HUMAN	2	192770.160		60.4	2.5E1	0.46	0	Complement C4-A precursor (Acidic com...		
	#DECOY#Q8NDH6 ICA1L_HUMAN	3	54407.082		55.8	2.43E1	1.24	0	Islet cell autoantigen 1-like protein (Amy...		
	Q9NYC9 DIH9_HUMAN	4	511935.780		48.4	-	0.2	0	Ciliary dynein heavy chain 9 (Axonemal ...		N/A
	#DECOY#P23786 CPT2_HUMAN	5	73777.055		47.5	-	2.43	0	Carnitine O-palmitoyltransferase 2, mito...		
Peptides Coverage Tool Box											
ID	Sequence	PEAKS(Confidence %)	-10Log(p-value)	Confidence	M/Z	Z	Mr(Calc)	Delta(Mass)	Error(ppm)	File	RT Scan Quality Frag. Mode Start End #Hits 117.115:114.112
Spectrum 1 A(+144.10) PLDNDIGVSEATR		100.0	119.31	0.0	801.3545	2	1600.8179	0.1234	77.0997	ITRAQSample.mzML, 0.047	2 0.79 CID, CAD, ... 787 800 4.59
Spectrum 5 D(+144.10) WENPGYVQNR		100.0	104.26	0.0	786.8787	2	1571.7814	0.0386	24.5423	ITRAQSample.mzML, 0.42	8 0.79 CID, CAD, ... 15 26 3.94
Spectrum 7 M(+144.10) SGIFR		99.7	30.55	0.0	427.7355	2	853.4602	0.0037	4.2909	ITRAQSample.mzML, 0.59	11 0.777 CID, CAD, ... 205 210
Spectrum 8 V(+144.10) DEDQFFPAVVK(+144.10)		100.0	60.06	0.0	815.4257	2	1628.8654	0.0284	17.4618	ITRAQSample.mzML, 0.763	13 0.791 CID, CAD, ... 506 517

3.3. Peptide View

The peptide view displays all the identifiable peptides and their quantification ratios. The interface is similar to the peptide table in a PEAKS DB result. The quantification ratios of those quantifiable peptides are displayed in

PEAKS Q - MS/MS Level

the ratio columns with reporter ion mass as the header eg. "Ratio 117:114". Right click on a peptide and use "show original spectrum" popup menu to jump to the data view to check the reporter ions in the original spectrum.

Summary		1 - 1000 of 1396										scan = <input type="text" value="search"/> no results			
Peptide	Protein	Peptide	-10lgP	Mass	ppm	m/z	RT	Scan	#Spec	Accession	115.11:114.11	116.11:114.11	117.11:114.11	PTM	
1		T(+144.10)AGIQIVADDLTVTNPK(+144.10)	97.06	2043.1455	2.9	1022.58	48.812	F1:2894	70	P00924 ENO1_YEAST	2.21	0.51	1.48		
2		D(+144.10)AIPENLPLTADFAEDK(+144.10)DVC[...]	90.81	2889.4795	4.0	1445.75	50.602	F2:3172	191	P02769 ALBU_BOVIN	6.48	0.53	4.00		
3		L(+144.10)ITAIGDGVNHDPVVGDR	87.96	2033.1027	4.9	1017.56	55.036	F2:3664	144	P00489 PYGM_RABIT	1.22	0.32	0.74		
4		I(+144.10)EEELGDNAVFAGENFHHGDK(+144.10)L	86.57	2728.3335	2.6	910.45	49.599	F2:3061	54	P00924 ENO1_YEAST	1.95	0.28	1.10		
5		S(+144.10)GETEDFTIADLVGLR	86.35	1965.0176	3.5	983.52	69.426	F1:4939	49	P00924 ENO1_YEAST	1.81	0.32	1.10		
6		Y(+144.10)PIVSIEDFFAEDDWEAWSH	86.19	2549.1509	1.9	1275.59	61.163	F2:4297	5	P00924 ENO1_YEAST	1.61	0.31	1.02		
7		F(+144.10)DEFFSAGC(+57.02)APGSFR	82.40	1787.8058	5.0	894.91	49.045	F2:2996	3	Q29443 TRFE_BOVIN	3.96	0.68	2.57		
8		Y(+144.10)NGVFQEC(+57.02)C(+57.02)QAEDK[...]	82.37	2034.9019	-0.2	1018.46	43.745	F5:2174	150	P02769 ALBU_BOVIN	5.92	0.59	3.60		
9		W(+144.10)VDTQVWLAMPYDTPVPGYR	81.44	2450.2427	2.4	1226.13	56.767	F4:3460	10	P00489 PYGM_RABIT	0.75	0.22	0.37		
10		A(+144.10)VDOFLSLDGTANK(+144.10)	81.24	1865.9978	1.7	934.01	56.264	F2:3791	93	P00924 ENO1_YEAST	1.33	0.38	0.76		
11		A(+144.10)VDOFLSLDGTANK(+144.10)	81.24	1865.9978	1.7	934.01	56.264	F2:3791	93	P00925 ENO2_YEAST	1.33	0.38	0.76		
12		G(+144.10)AAGGLGLAVQYAK(+144.10)	80.90	1649.9344	6.7	825.98	46.340	F4:2354	2	P00330 ADH1_YEAST	1.40	0.24	0.92		

Note



If there is no special view for reporter ions, then you can select a peptide and zoom to the reporter ion region of the MS/MS to examine the reporter ions.

Chapter 13. Label Free Quantification (LFQ)

1. Overview

Label free quantification is one of the three quantification modes supported by the optional PEAKS Q module. This quantification type is based on the relative intensities of extracted ion chromatograms (XICs) for precursor ions of identified peptides in multiple data sets. No chemical label is required. Different samples are measured separately in the same instrument. The same peptides from different samples are correlated by their m/z and elution time. Label free quantification relies on the assumption that the changes in analyte signals reflect their concentrations in one sample relative to another. This technology employs overall spectral intensity normalization by interpreting signals of molecules that do not change concentration from sample to sample. PEAKS Q uses the overall protein concentration in each sample for the normalization, making spiking unnecessary.

Label free quantification is based on the PEAKS DB search results (See Chapter 9, *Protein Database Search with PEAKS DB*). The use of this function is outlined in the following overview.


1. Select a  PEAKS DB, fraction, sample or project node in the “Project View” frame. Click the PEAKS Quantification toolbar icon  or select “Quantification” from the “Tools” menu.

Important

In order to use the label free quantification analysis of PEAKS Q, the survey scans in the data have to be in profile (un-centroided) mode.

Note

Refer to Section 2, “Quantification Workflow” for how to conduct PEAKS DB and quantification in a single workflow.

2. Select the quantification protocol as label free, specify the PEAKS quantification parameters in the right panel of the window, and click “OK”.
3. Wait for the analysis to finish. A new quantification result node  will appear at the project tree. Double click the node to open the result file.

2. Setting Parameters

Select “Label Free” from the left hand side under the “Tools” heading in the quantification window to view the label free quantification parameters on the right hand side.

Quantification

Tools

☐ Label at MS/MS level
eg. iTRAQ

☐ Label at MS level
eg. SILAC

☒ **Label Free**

Label Free default Save as...

Basic Options

Mass Error Tolerance: Da
 Peptide Score Threshold (-10logP):

Retention Time Range: min.
 Protein Score Threshold (-10logP):

Upper Bound of Precursor Charge:
☒ do normalization

Parameter Table

Project Name	Sample Name	Fraction Number	File Name	Protein ID	Add to quantification
AndrewWatsonLFQ				PEAKS 32 [26-Apr...	
	Sample 1			PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>
		1	d11a.RAW	PEAKS 32 [26-Apr...	
	Sample 2			PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>
		1	d11b.RAW	PEAKS 32 [26-Apr...	
	Sample 3			PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>
		1	d11c.RAW	PEAKS 32 [26-Apr...	
	Sample 4			PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>
		1	d11d.RAW	PEAKS 32 [26-Apr...	
	Sample 5			PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>
		1	d11e.RAW	PEAKS 32 [26-Apr...	

The following parameters are available in the “Basic Options” section of the quantification window:

- **Mass Error Tolerance:** Used to locate the precursor ion peak group of an identified peptide in the survey scans. We deal with non-centroided survey scans in a LFQ experiment. So, set the mass error tolerance a little wider than the parent ion error tolerance in the PEAKS DB database search.
- **Retention Time Range:** The retention time range is the maximum elution time range that is considered for the quantification of an identified peptide. This also defines the search range for peptide feature pairing across samples.
- **Upper Bound Precursor Charge:** The precursor ion of an identified peptide may have sibling ions of different charge states. Only those sibling ions with charge less than the upper bound precursor charge will be considered for quantification of the identified peptide.
- **Peptide Score Threshold:** Only identified peptides with a score above this threshold will be used in quantification.
- **Protein Score Threshold:** Only identified proteins with a score above this threshold will be used in quantification.
- **do normalization:** If selected, normalization of protein ratios based on total ion intensity will be done automatically.

The “Parameter Table” includes the following information:

- **Project Name:** name of the project selected for quantification

- **Sample Name:** names of samples in the project

Note


You need to have at least 2 samples with at least 1 file/fraction in each sample.

- **Fraction Number:** the number of the fractions in the sample
- **File Name:** name of the data file
- **Protein ID:** PEAKS DB result that will be used in quantification. Select the PEAKS DB result to be used from the drop down list containing all available results.
- **Add to quantification:** Check/uncheck to add the sample to the quantification. There must be at least two samples in label free quantification and the number of fractions within each sample must be the same.

Clicking the “Save As” button at the top right allows the user to save parameters for ease of use when regularly performing quantification with the same parameters. All the parameters in quantification will be saved except the “Parameter Table” which will change from one project to another.

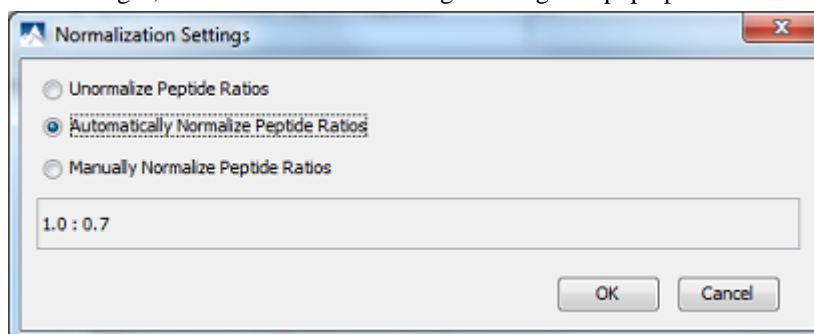
The “Assign replicates” button helps to assign the samples a replicate number. This enables PEAKS to perform replicate analysis. Refer to Section 6, “Replicate Analysis in LFQ” for details on how to assign replicates and perform replicate analysis.

3. Understanding LFQ Result

Once completed, the label free quantification result will be displayed in the quantification node . Double click on this node and the “Summary” view tab will appear by default. Right click on the result node to find more operations supported for the a label free quantification result.

PEAKS supports export of the label free quantification results to Excel or HTML file by right clicking the result node and choose the corresponding function. Please refer to Section 4.2, “Export Label-Free Quantification Results” for details.

PEAKS also supports changing the normalization factor of the protein ratio. Right click on the result node and select “Normalization Settings”, the “Normalization Settings” dialog will pop up:

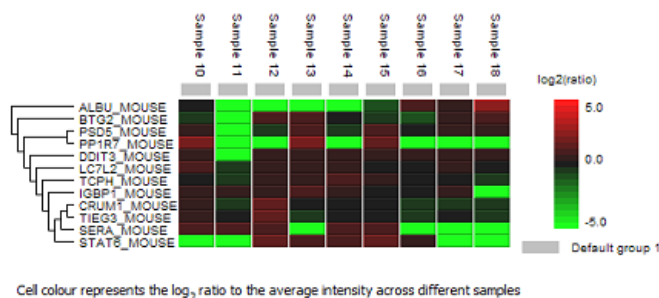


If you select “Unnormalize Peptide Ratios”, the protein ratio will be calculated from peptide ratios without normalization. PEAKS will normalize the result when “Automatically Normalize Peptide Ratios” is selected and the normalization factor will be displayed. You can also set the normalization factor manually by clicking the “Manually Normalize Peptide Ratios” and by inputting the ratios into the text field. The format of ratios should be numbers separated by colons and the number of ratios should be the same as the number of samples in the quantification result.

3.1. Summary View

The label free quantification results are summarized in a one page summary as shown in the next figure.

1. Heatmap View



2. Notes

3. Result Statistics

Table 1. Statistics of data.

# of MS Scans	24874
# of MS/MS Scans	65632

Table 2. Result filtration parameters.

Protein fold change ≥ 2

Table 3. Sample Selection.

Default group 1:

Sample 10
Sample 11
Sample 12
Sample 13
Sample 14
Sample 15
Sample 16
Sample 17
Sample 18

4. Other Information

Table 4. Search parameters.

Quant type: Label free quantification
Mass error tolerance: 0.2 Da
Retention time Range: 3.0 min
Upper bound charge: 3
Minimum peptide score: 20.0

Table 5. Instrument parameters.

Fractions: Spike5_040806_2.RAW,Spike0_040806_2.RAW,Spike1_040806_2b.RAW,Spike2_040806_2.RAW,Spike3_040806_2.RAW,Spike4_040806_2.RAW,Spike6_040806_2.RAW,Spike7_040806_2.RAW,Spike8_040806_2.RAW,Spike5_040806_1.RAW,Spike0_040806_1.RAW,Spike1_040806_1.RAW,Spike2_040806_1.RAW,Spike3_040806_1.RAW,Spike4_040806_1

The summary includes an expression profile with candidate proteins assorted in a heat map, result statistics tables, a list of instrument parameters, and a list of search parameters. To add a summary note, click on the “Edit Notes” button to open a “Notes Entry” editor where you can edit the notes to be displayed on summary page.

The summary page can be exported to other formats by clicking the “Export” button. For more details, refer to Section 4.2.2, “Export Summary Page”.

Heat map. The hierarchical clustering of proteins is represented as a heat map depicting relative protein abundance (normalized SC values logged to base 2) of the protein list with filters. The hierarchical clustering is measured with a euclidean distance similarity measurement of the log₂ ratios of the samples relative to a canonical sample.

3.2. Protein View

Click the “Protein View” tab. The quantified proteins, supporting peptides of each protein, and peptide features in the survey spectra from each sample will be displayed in the result panel. The quantified proteins will appear in the top panel with homologous proteins clustered together. The ratio of Sample 1: Sample 2 appears in the “Ratio” column and the standard deviation of Sample 1: Sample 2 appears in the “SD” column.

Q LABEL FREE 15 [22-Mar-11 09:46] x										
Accession	ID	Mass	Score	Coverage	Matched	Description	Ratio	SD	Marked	
Quantification Result										
P27824 CALX_HUMAN	21	67,568.453	0.8	3.37	5	Calnexin precursor (M...	1.00 : 1.06	0.00 : 0.01		
P08238 HS90B_HUMAN	4	83,264.234	0.7	4.28	4	Heat shock protein HS...	1.00 : 1.00	0.00 : 0.05		
O60832 DKC1_HUMAN	213	57,674.223	0.3	3.89	3	H/ACA ribonucleoprote...	1.00 : 1.10	0.01 : 0.00		

Peptides						
Sequence	Peaks	QScore	M/Z	Z	Retention Time	Ratio
Support Peptides						
Unique Peptides						
QKS(+79.97)DAEEDGGTVS(+79.97)QEEEDR	0.2	1134.9127	2		[32.62, 33.36]	1.00 : 0.62
QKS(+79.97)DAEEDGGTVSQEEEDR	0.2	1094.9296	2		[33.34, 34.39]	1.00 : 0.70
QKS(+79.97)DAEEDGGTVSQEEEDR	0.2	730.289	3		[33.31, 34.36]	1.00 : 0.75
SDAEEDGGTVS(+79.97)QEEEDR	0.2	966.85284	2		[33.34, 34.07]	1.00 : 0.70
Outlier						
QKS(+79.97)DAEEDGGTVS(+79.97)QEEEDR	0.1	756.9444	3		[32.62, 33.47]	1.00 : 0.00
Multi Cluster Supports						

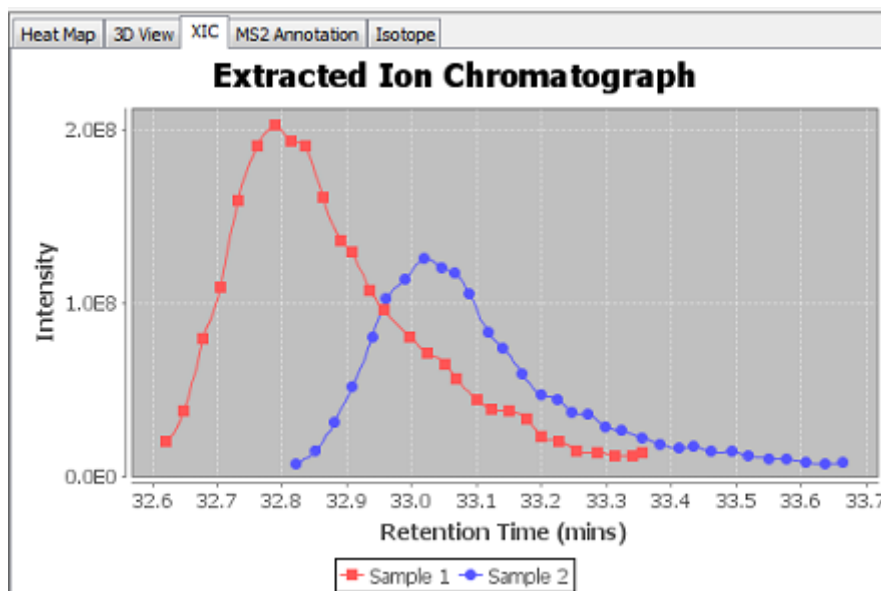
The supporting peptide is shown under the “Peptides” tab. The retention time is shown for the specific peptide as well as the peptide ratio from Sample 1: Sample 2.

Click on the “+” beside the “Outlier” folder to see the peptides that were not included in the ratio.

To see which peptides were used to identify the protein during the PEAKS DB search, select the “Coverage” tab. The entire sequence of the protein is shown and the matching peptides are highlighted in blue. In this example the total matched part accounts for 3.37% of the protein. This information can be found in the “Coverage” column above.

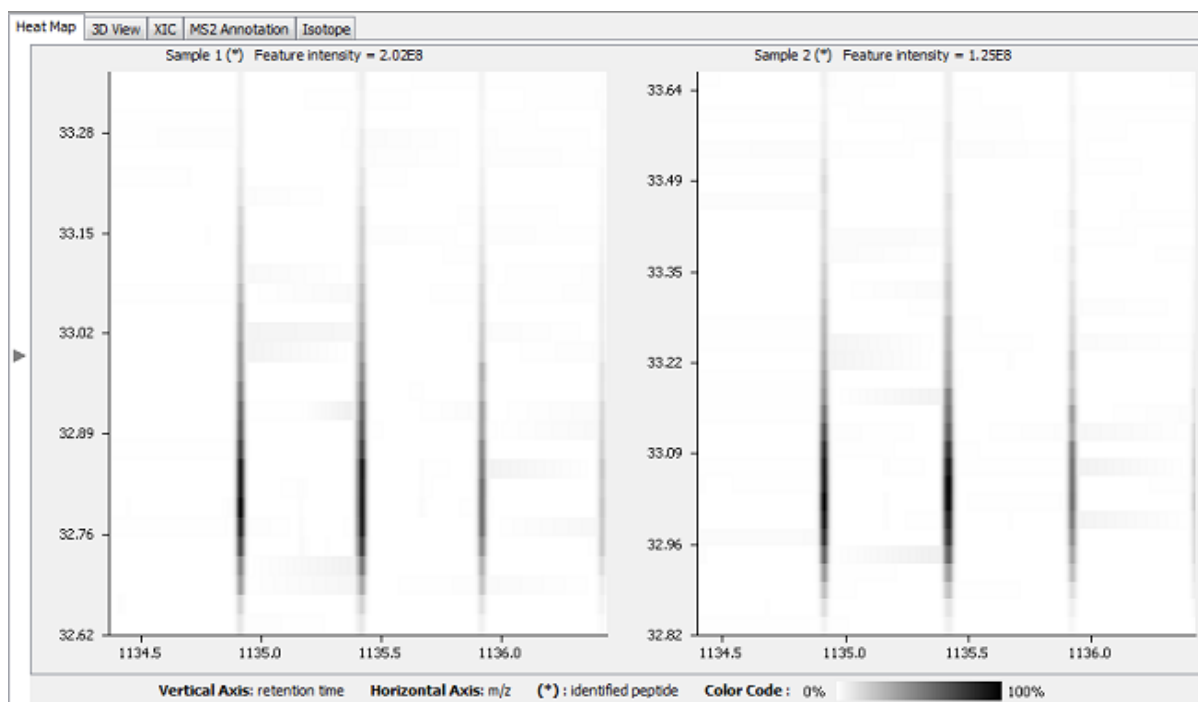
3.2.1. Extracted Ion Chromatogram

The reconstructed “Extracted Ion Chromatogram” chart will appear by default in the bottom panel: These display the shape of the peptide features over the retention time range where they were identified.



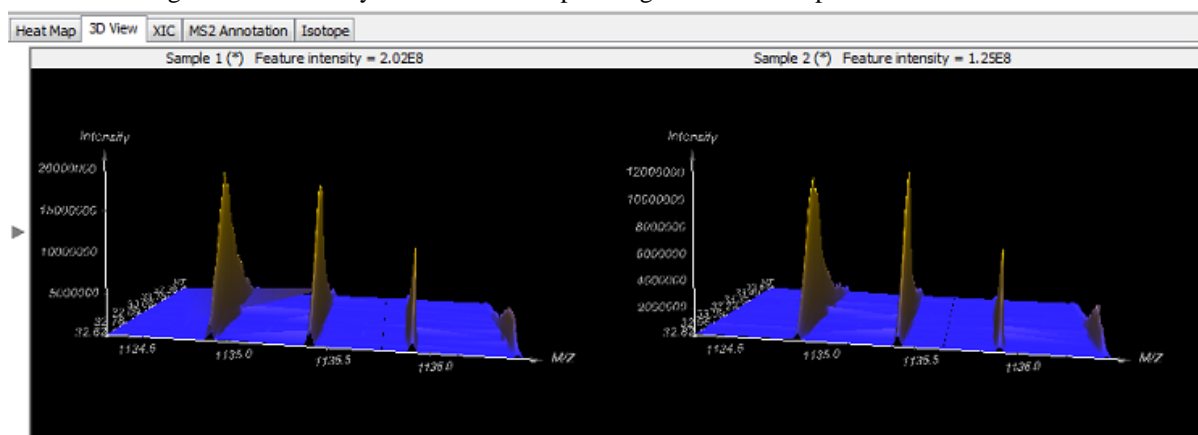
3.2.2. Heat Map

Select “Heat Map” tab to view the 2D heat map. When viewing the heat map in colour, red represents high intensity and yellow represents low intensity. The grayscale heat map displays high intensity in black while white represents low intensity. If the peptide is identified in PEAKS DB, there will be a star after the sample name.



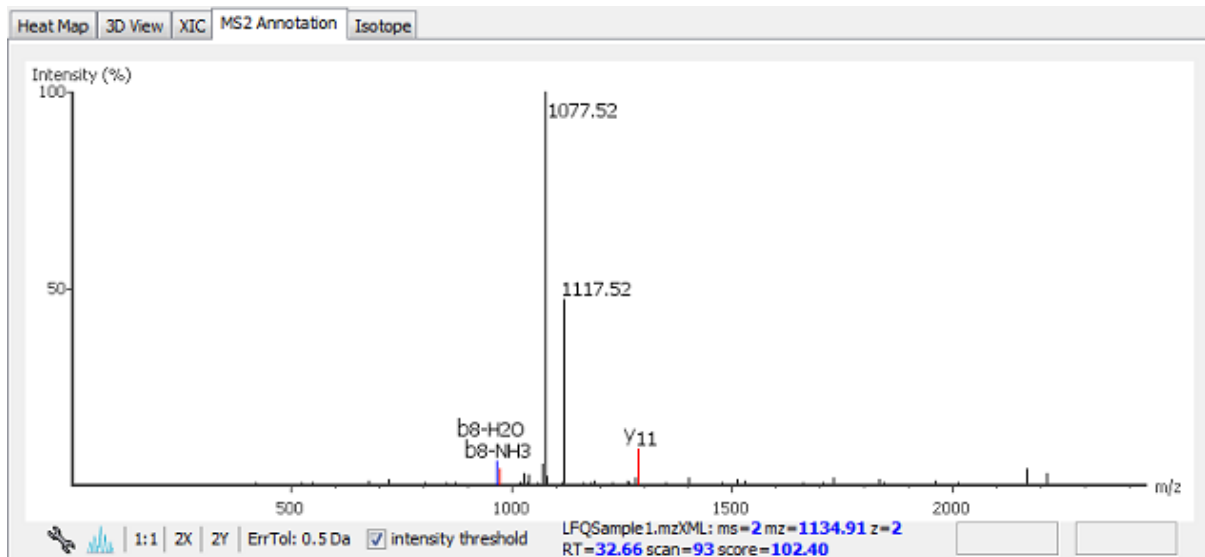
3.2.3. 3D View

Click on the “3D View” tab to display a 3D View of the peptide features for sample 1 and sample 2. Intensity is displayed on the y-axis, m/z on the x-axis and retention time on the z-axis. Click and move the cursor to rotate around the image. Notice that as you move one sample image the other sample moves to the same location.



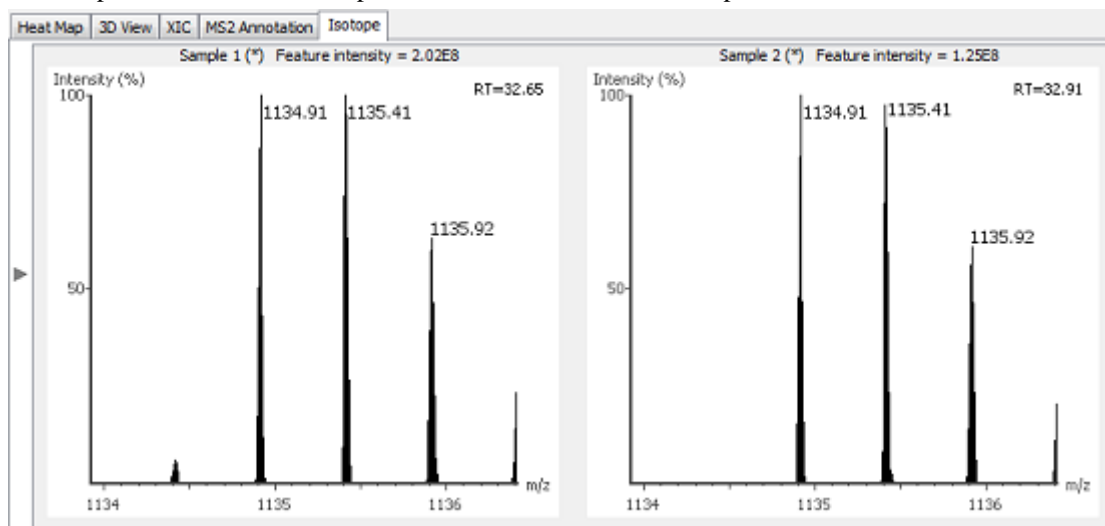
3.2.4. MS2 Annotation

Select the “MS2 Annotation” tab to view a graphical representation of the spectrum annotation. This is similar to the *de novo* results and PEAKS database search results spectrum annotation. Please refer to Section 3.2.2, “Spectrum Annotation” for more details.



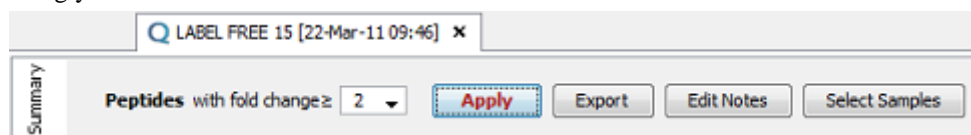
3.2.5. Isotope

Select “Isotope” tab to view the isotope distribution detected in the samples.



4. Filter LFQ Result

PEAKS quantification results can be filtered to show all peptides with a certain fold change. You can set the appropriate value for the filter by changing the filtration parameter value from the drop-down list in the title bar of the “Summary” view panel. Click the “Apply” button to refresh the results. The results will be updated in all views accordingly.



5. Export Quantification Result to Other Formats

PEAKS quantification results can be exported to other supported formats. To export the result, press the “Export” button in the title bar of the “Summary” view panel. Refer to Section 4.2, “Export Label-Free Quantification Results” for details.

6. Replicate Analysis in LFQ

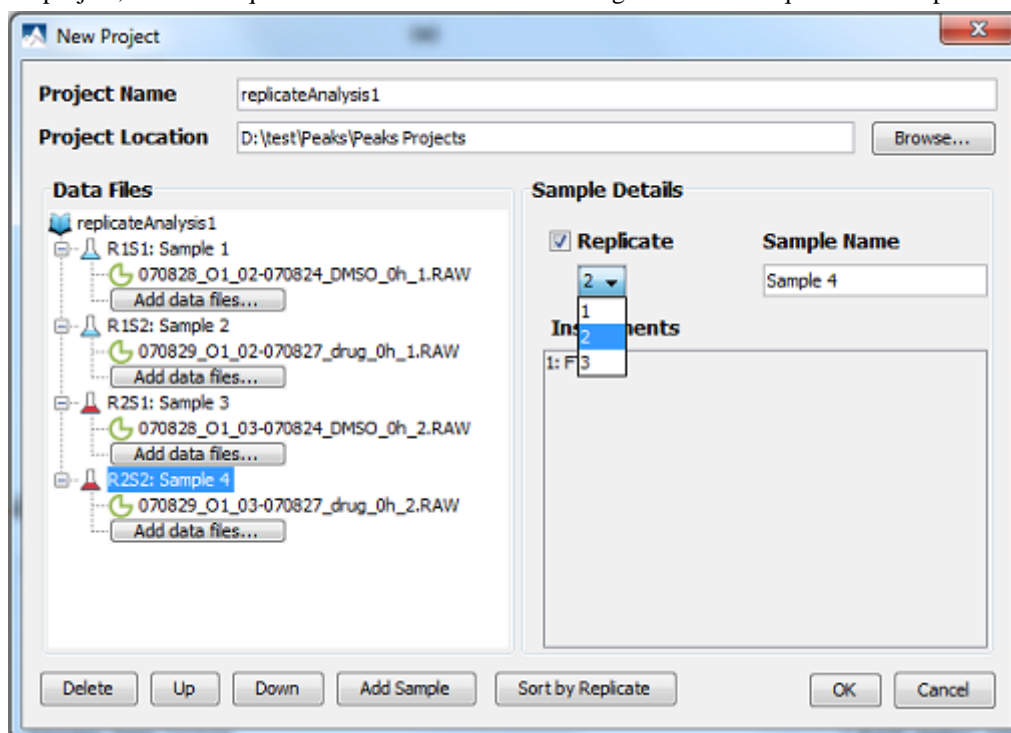
In liquid chromatography-mass spectrometry (LC-MS) based proteomics, multiple samples from different groups are often analyzed in parallel. Tools that access the quality of proteomics data based on sound statistical principles are needed in this field. In PEAKS 5.3, comparison functions are provided in three levels:

- Assess the reproducibility of MS data from technical replicates
- Perform compare analysis of peptides and proteins
- Assess the reproducibility of protein quantification from biological/technical replicates

This section is organized to first introduce how to assign replicate numbers to samples in the project. The replicate analysis of MS data comparisons and label free quantification are done together and so each function will be introduced together in one section.

6.1. Assign Replicate Number to a Sample

A sample can be assigned with a replicate number in two ways: in the “New Project” window when adding a sample to a project, and in the quantification window when setting the label free quantification parameters.



To assign the replicate number in the “New Project” window, select the sample from the project view on the left hand side, select the “Replicate” check box and click the drop-down list below the check box to select a number. Once assigned, the name of the sample will be changed to indicate its replicate number and the sample number in the replicate. The sample node icon colour also will be changed to display the replicates.

Quantification

Tools

- ☐ Label at MS/MS level
eg. ITRAQ
- ☐ Label at MS level
eg. SILAC
- ☒ **Label Free**

Label Free default Save as...

Basic Options

Mass Error Tolerance: Da Peptide Score Threshold (-10logP):

Retention Time Range: min. Protein Score Threshold (-10logP):

Upper Bound of Precursor Charge: ☒ do normalization

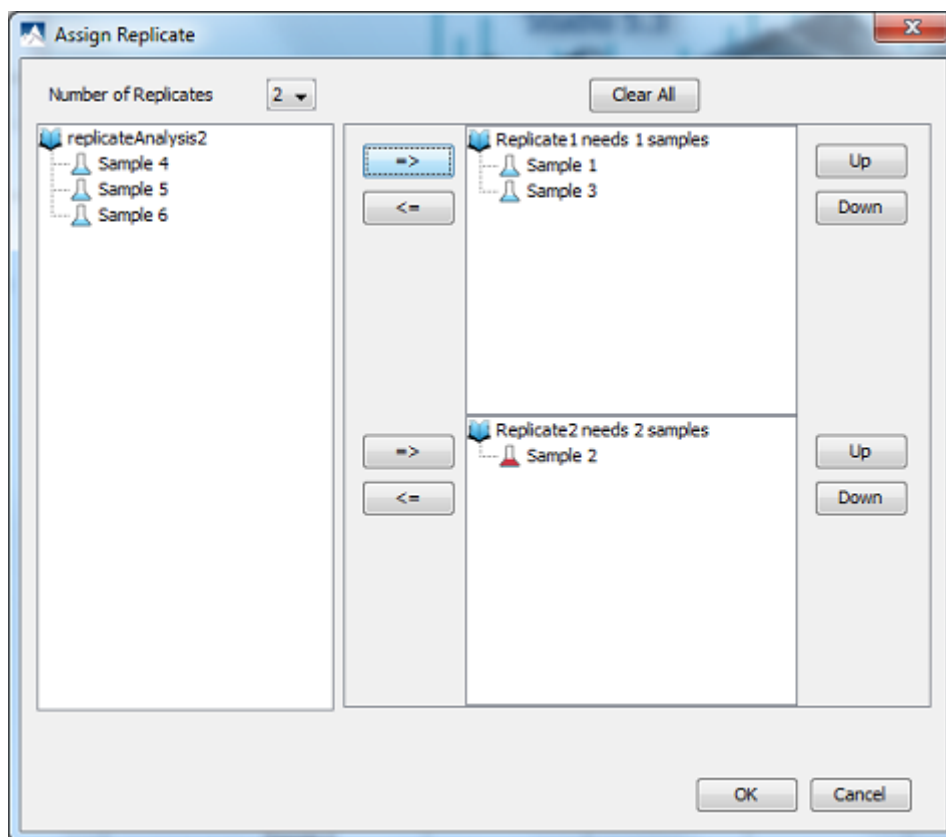
Parameter Table

Project Name	Sample Name	Fraction Number	File Name	Protein ID	Add to quantification
AndrewWatsonLFO				PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>
	Sample 1	1	d11a.RAW	PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>
	Sample 2	1	d11b.RAW	PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>
	Sample 3	1	d11c.RAW	PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>
	Sample 4	1	d11d.RAW	PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>
	Sample 5	1	d11e.RAW	PEAKS 32 [26-Apr...	<input checked="" type="checkbox"/>

up down Assign replicates

OK Cancel Help

To assign the replicate number in the quantification window, click the “Assign replicates” button below the “Parameter Table” on the right hand side. This will open the “Assign Replicate” dialogue where the replicates can be defined.



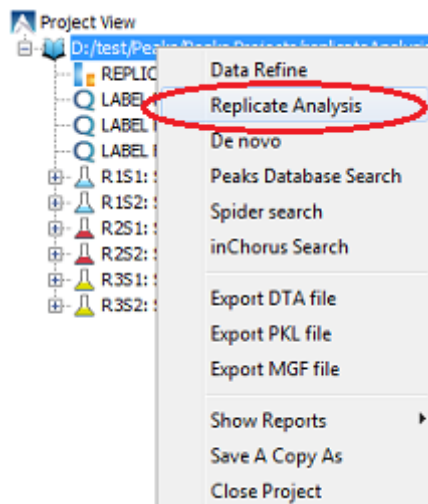
Select the number of replicates from the “Number of Replicates” drop-down list on top of the window. All available samples are listed in the unassigned samples list on the left hand side. The list of samples in each replicates are displayed on the right hand side. To assign a sample to a replicate, select a sample from the unassigned sample list and click on the “=>” button beside the list of samples in the corresponding replicate.

To remove a sample from a replicate, select the sample and click the “<=” button beside that replicate. To remove all assignments click on “Clear All” button.

Relative order of a sample in a replicate can be controlled by “Up” and “Down” buttons beside the corresponding replicate.

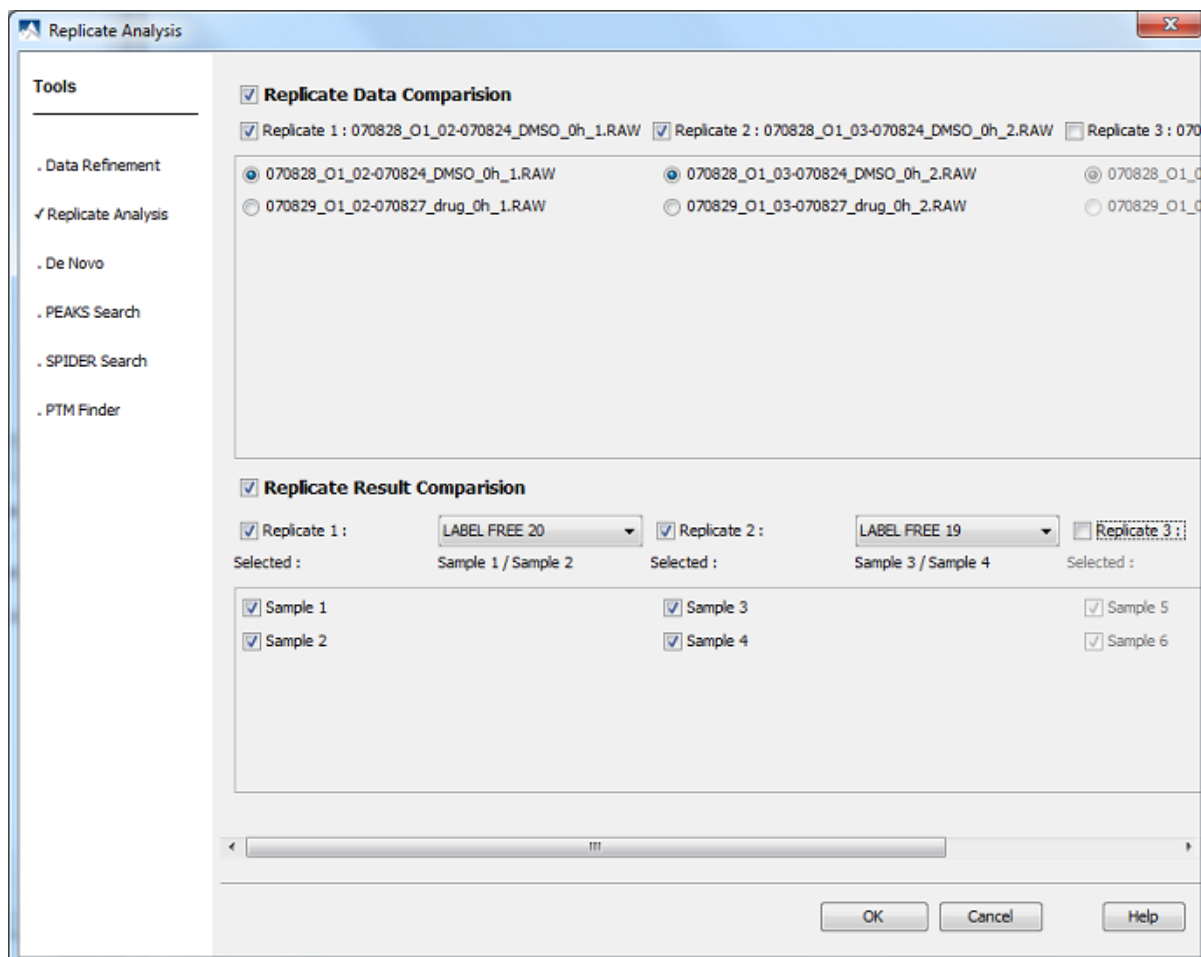
6.2. Run Replicate Analysis

Select the project from the “Project View” and right click on the project node. Select “Replicate Analysis” from the pop-up menu. This will pop up the “Replicate Analysis” window.



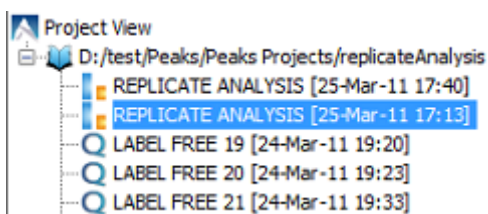
Replicate Data Comparison: Select the replicate and samples on which you want to perform data comparison analysis. Only two replicates can be selected for data comparison analysis.

Replicate Result Comparison: Select the label-free quantification results and samples on which you want to perform result comparison analysis. You need to select one label-free quantification result for each replicate and two samples you want to compare. Once a sample is selected, all the samples with the same index in other replicates will be selected.



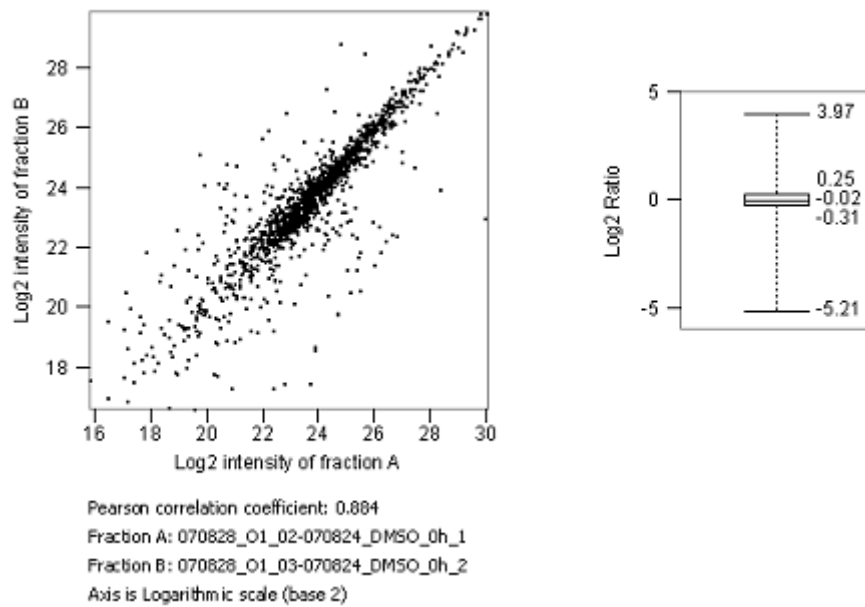
6.3. Understand Replicate Analysis Results

Once the replicate analysis is completed, a new replicate analysis result node will appear in project view. Double click on the result node to view the result.



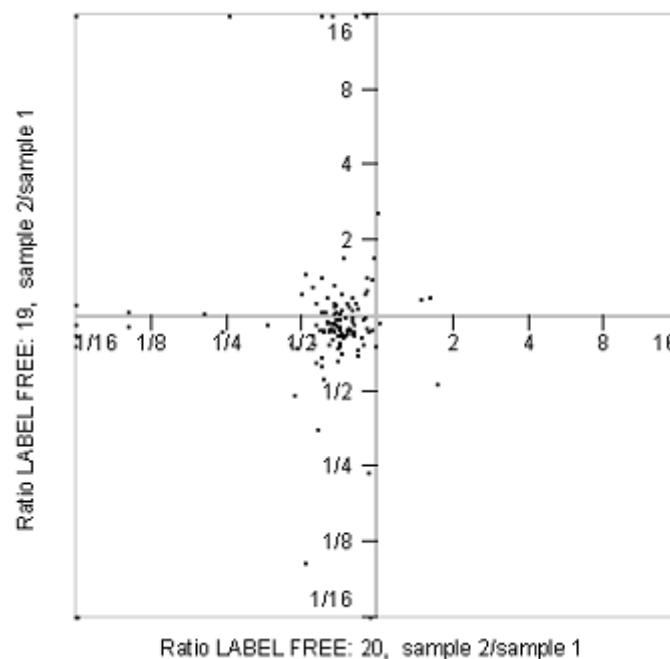
The results consist of a few charts to compare the data and results of the two samples. If you selected both the data and result comparisons when setting up replicate analysis, the following charts will appear:

Feature Comparison. The feature comparison scatter plot represents each feature vector which consists of two features detected in the two data files you want to analyze and aligned in the label free quantification. The x-axis is the log intensity of the feature detected in the first data file and the y-axis is the log intensity of the feature detected in the second data file. The Pearson Correlation Coefficient is calculated and listed under the chart. The standard box plot is shown on the right side of the scatter plot.



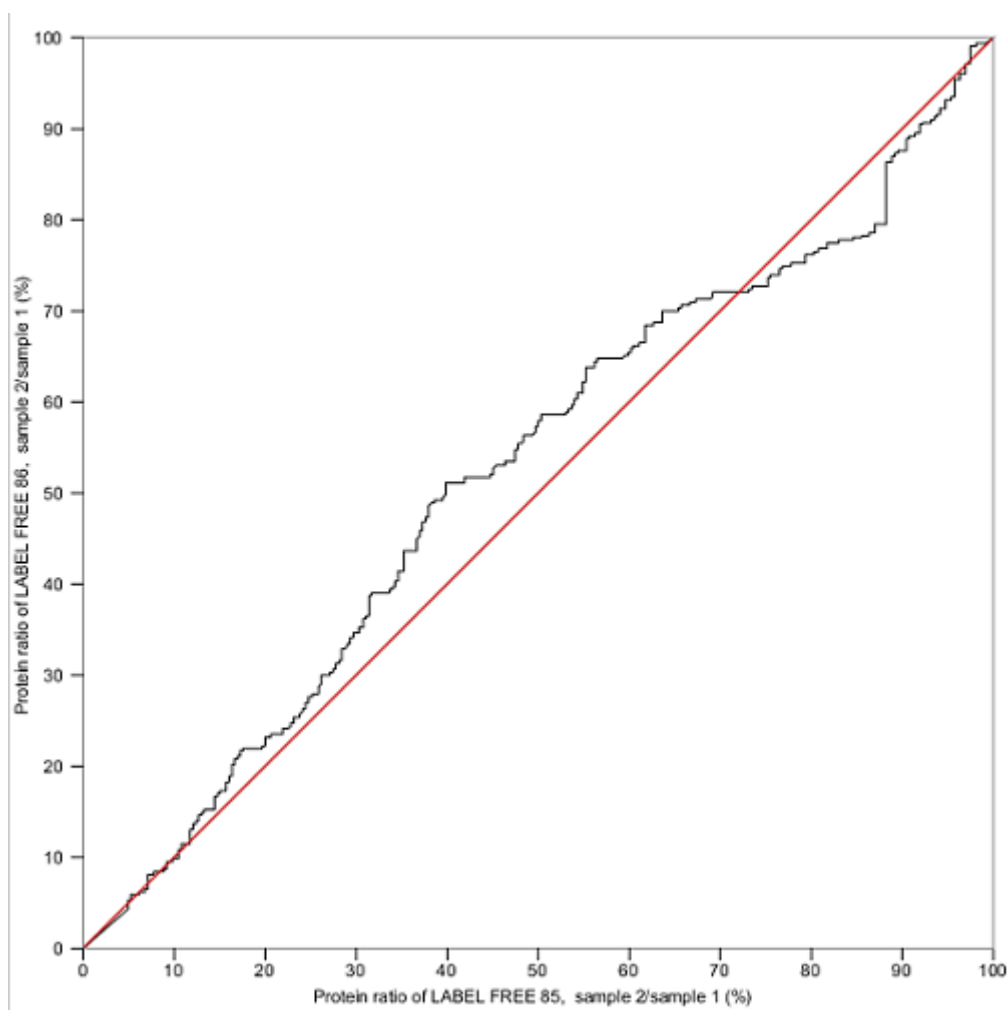
Feature Venn Diagram. The feature Venn diagram is a standard Venn diagram showing the number of common peptide features and unique peptide features of the two data files.

Peptide Scatter Plot. The peptide scatter plot compares the peptides quantified in two label free quantification results. The x-axis is the ratio of the peptide of label free quantification result 1 and the y-axis is the ratio of the same peptide relative intensity ratios in corresponding samples of label free quantification result B.



Peptide Venn Diagram. The peptide Venn diagram is a standard Venn diagram comparing the number of quantified common peptides and unique peptides of label free quantification results.

Protein Q-Q Plot. The protein Q-Q plot is a standard quantile plot comparing the protein ratios from selected samples of label free quantification results. The ratios of the proteins in the first sample is plotted against the ratios of the proteins in the second ratio, both in ascending order of size, and scaled from 0 to 100. In the ideal case, both replicates should result in the same ratios for the proteins, and thus the expected result is represented by the diagonal line in red.



6.4. Export Replicate Analysis Result

You can export the replicate analysis plots and diagrams to an image file. To export to an image file, position the cursor on any of the plots or diagram in the result panel and click the right mouse button to view the pop-up menu and select the “Export Image” command from the menu. Refer to Section 2.2, “Export Images” for details.

Part IV. Advanced Data Analyses

Table of Contents

14. PTM Finder	111
1. Summary	111
2. Setting Up PTM Finder Parameters	111
3. Understand PTM Finder Result	112
15. Homology Search with SPIDER	113
1. Set SPIDER Parameters	113
1.1. Run SPIDER on PEAKS DB Result	113
1.2. Run SPIDER Independently	115
2. Understand SPIDER Results	115
2.1. SPIDER Peptide View	116
2.2. SPIDER Protein View	116
16. Workflow	118
1. Identification Workflow	118
2. Quantification Workflow	119
3. inChorus Workflow	120
17. Exporting Data/Reports and Printing	121
1. Export Data	121
2. Export <i>De Novo</i> Result	121
2.1. Export Summary and Peptides	122
2.2. Export Images	122
3. Export PEAKS DB Result	123
3.1. Export Summary, Proteins and Peptides	123
3.2. Export Images	124
4. Export Quantification Results	124
4.1. Export Labeled Quantification Results	125
4.2. Export Label-Free Quantification Results	125
4.2.1. Export Result in Excel or HTML	125
4.2.2. Export Summary Page	126
5. Export SPIDER Result	126
6. Export inChorus Result	127
18. Configuration and Preferences	129
1. PEAKS Environment Preferences	129
1.1. General Preferences	129
1.1.1. Display Options	130
1.1.2. RMI Connections	130
1.1.3. Derby Database	130
1.1.4. Performance	131
1.2. Instrument Preferences	131
1.2.1. ABI (.wiff)	132
1.2.2. Bruker (.yep/baf, fid)	132
1.2.3. Shimadzu Axima (.run)	132
1.2.4. Varian (.xms)	132
1.3. Search Engine Preferences	132
1.3.1. Mascot Settings	132
1.3.2. X!Tandem Settings	133
1.3.3. OMSSA Settings	133
1.3.4. Sequest Settings	133
1.4. Spectrum Annotation Preferences	134
2. PEAKS Configuration	135
2.1. Enzyme Configuration	135
2.2. PTM Configuration	136
2.3. Database Configuration	137
2.4. Instrument Configuration	137
2.5. Parameter Configuration	139

Chapter 14. Identifying More PTMs with PTM Finder

1. Summary

The PEAKS DB database search function can already identify modified peptides with user-selected PTMs. However, specifying very many variable PTMs in PEAKS DB is not the best practice since it will significantly increase the running time. PTM Finder is provided to fill the gap. PTM Finder uses the same algorithm as the PEAKS DB search engine. To reduce the running time when many PTMs are specified, PTM Finder limits the search in two ways:

1. It only searches using the spectra that have highly-confident *de novo* sequence tags. The other spectra typically do not have the required quality to confidently identify a modified peptide.
2. It only searches in the proteins that were identified by a previous PEAKS DB search. Thus, the purpose of PTM Finder is to identify more modified peptides and increase the protein coverage. It will also not search scans that have already been well identified.

The use of this function is outlined in the following steps. Details can be found in later sections of this chapter.

1. Select a PEAKS DB results node. Click on the PTM Finder  button on the toolbar.

Note

You cannot execute PTM Finder on a raw file or *de novo* results.

2. Specify parameters for PTM Finder in the popup dialog and click OK.
3. After the analysis complete, a new result node will appear in the project tree.
4. Results are viewed in the same user interface as PEAKS DB.

2. Setting Up PTM Finder Parameters

The parameters of used in a PTM Finder search are very similar to a PEAKS DB search. The differences are:

- The database selection pane disappears since the search will be on the identified proteins of the selected PEAKS DB search result.
- The filter option asks for the minimum *de novo* tag score (ALC) and maximum database search score ($-10\lg P$) in order to determine which spectra are searched. If the ALC of the *de novo* result is low, then the spectrum is unlikely to provide a significant hit. If the database search score is high, then the spectrum was already assigned confidently in the PEAKS DB search step.

PTM Finder

Predefined parametersInstrument_default

Save as...

Mass Options

Parent Mass Error Tolerance:0.1Da

Precursor Mass Search Type:
☒ Monoisotopic ☐ Average

Fragment Mass Error Tolerance:0.2Da

Enzyme

Trypsin

View Enzyme

New Enzyme...

PTM

Set PTM...

Remove

Switch type

Maximum allowed variable PTM per peptide3

General Options

Max Missed Cleavages:1

Filter Options

Filter the spectra which satisfy the following conditions for use in the PTM search:

De novo ALC (%) score greater than:50% recommend 50%

Peptide score (-10logP) less than:15 recommend 15

3. Understand PTM Finder Result

The results from a PTM finder search are presented in a similar format to those seen in a PEAKS DB search. Please refer to Section 3, “Understanding PEAKS DB Result” for more information on interpreting the search results.

Chapter 15. Homology Search with SPIDER

After having obtained *de novo* sequences for peptides that are novel or from a species without a corresponding protein database, it is possible to look for a homologous peptide in a database from a related species (Han et al., 2005). This process is called homology search and can help you to learn more about the proteins in your sample. Homologous proteins can provide clues as to possible functions of your novel peptides. SPIDER reconstruction can use both your *de novo* sequence tags and a homologous peptide to reconstruct the probable real sequence. To search with SPIDER you must first have good quality *de novo* sequences (either on their own or in conjunction with a PEAKS DB result).

With SPIDER, there are two options for searching:


1. Search in an existing PEAKS DB result.

The highly-confident *de novo* sequence tags of the unassigned spectra are searched against the proteins identified in the PEAKS DB result. This way, we can potentially increase coverage, get additional evidence for a particular peptide match, or identify new mutations in an existing protein using reconstruction.

2. Search in a protein sequence database.

The highly-confident *de novo* sequence tags are searched against a protein sequence database. This allows you to find protein matches from a homologous database even if the studied organism does not have a protein database.


The use of this function is outlined in the following steps. Details can be found in later sections of this chapter.

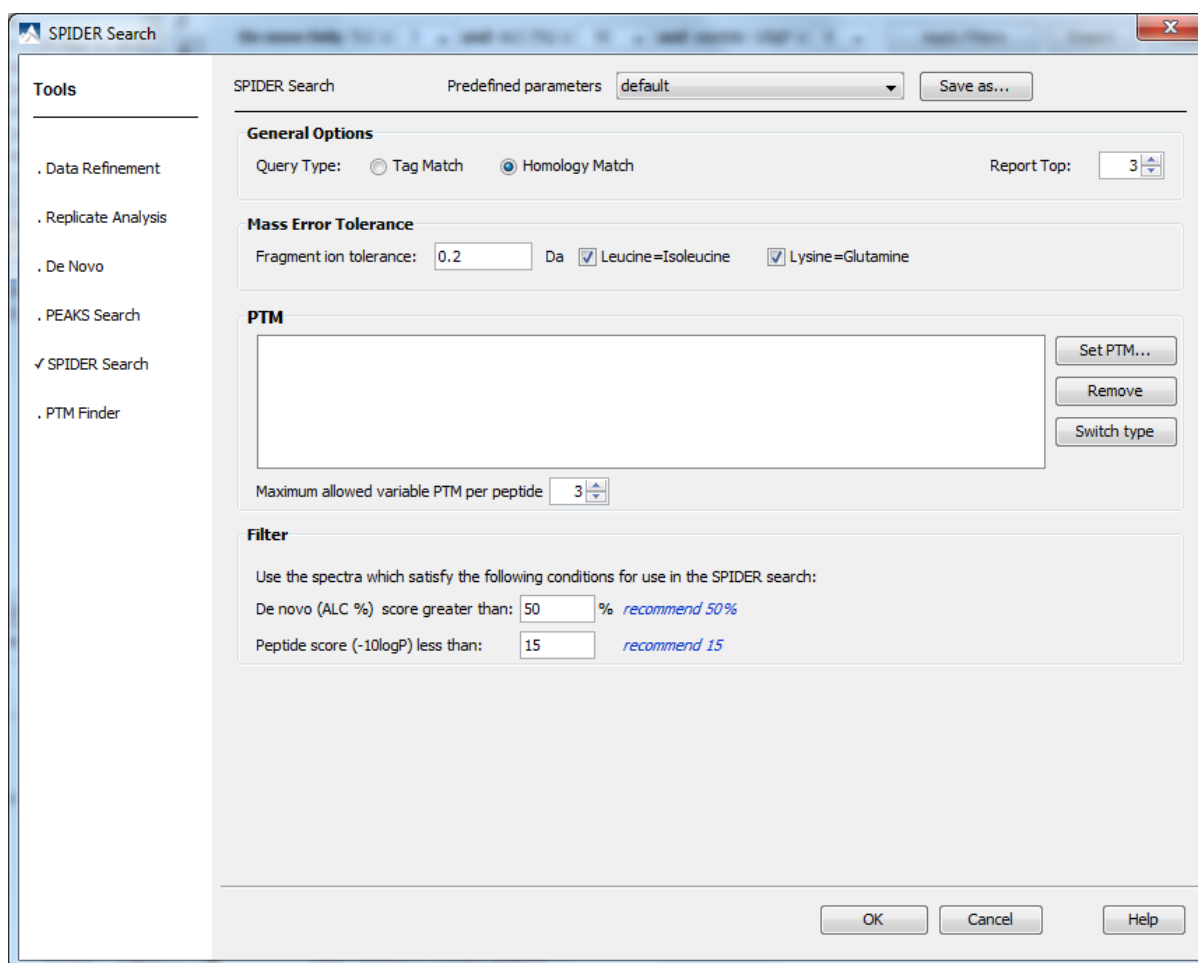
1. Select a PEAKS DB result node(search in an existing PEAKS DB result) or select a data node(search in a protein sequence database). Click on the SPIDER button  in the toolbar.
2. Specify parameters for SPIDER in the popup window and click OK.
3. After the analysis complete, a new result node will appear in the project tree.
4. Results are viewed in almost the same user interface as PEAKS DB.

The main change is, for the search in an existing PEAKS DB result, we will display three views: protein view, SPIDER peptide view and PEAKS DB peptide view. In the protein view, coverage column distinguishes between SPIDER hits, PEAKS hits, and overlapping hits by colour (Red, blue, and purple respectively).

1. Set SPIDER Parameters

1.1. Run SPIDER on PEAKS DB Result

The configuration panel of a restricted SPIDER Search will be invoked by selecting a PEAKS DB result and clicking the SPIDER icon on the toolbar  or choosing SPIDER Search from the Tools menu. Running SPIDER on a PEAKS DB result searches only proteins that have been identified by PEAKS DB. This allows you to efficiently expand coverage of your proteins and explain additional spectra that may not have matched to the database using database search. The configuration panel in this case appears as follows:



Note that when a SPIDER search is launched in this manner a protein database or a *de novo* run does not need to be configured. The SPIDER search will search based on the *de novo* run used for a particular PEAKS DB run and will search against the proteins identified in that result. Since SPIDER can be computationally intensive, filters are provided to limit the SPIDER search to high quality results based on user-specified thresholds.

The rest of this section will describe the different options that you have when setting up the parameters for your SPIDER search. After setting up the desired parameters, we can save them for future use. Click the "Save Parameters" button, and define name for these preferences for future use/reference. Any parameters that you save will be available in the drop-down list at the top of the window. To retrieve parameters, select a saved parameters file and the corresponding settings will be shown.

Query Type. There are two query types for SPIDER.


- **Tag Match** - This is not a true mutation search; instead, it will search the database with the *de novo* tags peptides while taking into account possible *de novo* sequencing errors. This search mode allows for the use of fixed modifications.
- **Homology Match** - This is a more rigorous (and a more computationally intensive) search mode, taking into account all types of mutations and the positional confidence scores. This search mode also creates reconstructed peptides that use information from both the *de novo* sequences and the database sequence in order to characterize the real sequence. This search mode allows for the use of both fixed and variable modifications.

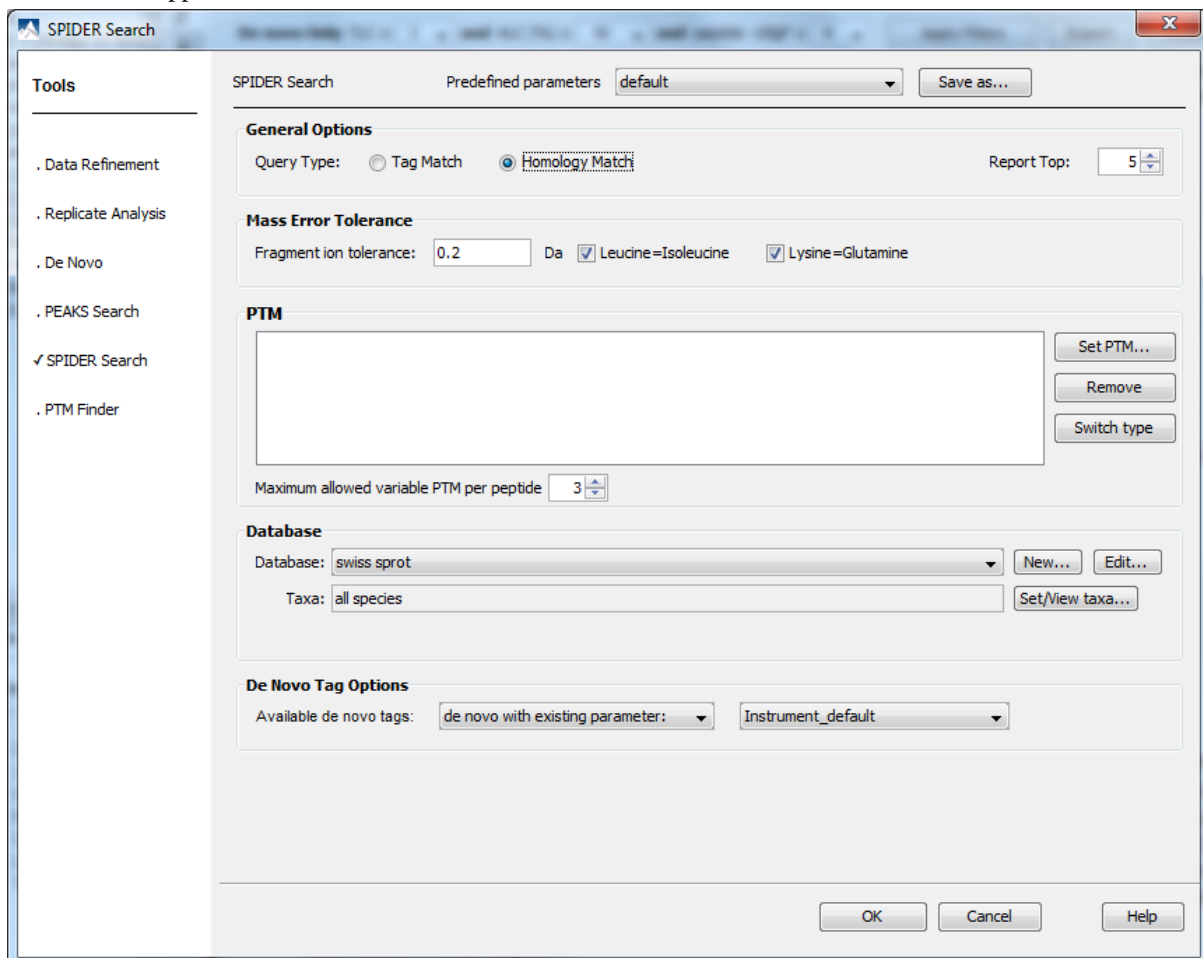
Mass Error Tolerance. Enter the range of fragment mass error in Daltons that PEAKS will allow for when determining the peptide sequences. The two additional checkboxes let SPIDER match the two pairs of amino acids without any penalty as they have identical or very similar mass values.

PTM. Clicking the "Set PTM" button will bring up a separate window for PTM configuration. The PTM configuration is the same as it is in *de novo* sequencing (Section 2.3, "Fixed and Variable PTMs").

Filter. The filter option asks for the minimum *de novo* tag score (ALC) and maximum database search score ($-10\lg P$) for a spectrum to be used. If the ALC is too small, then the spectrum is unlikely to provide a significant hit. If the database search score is very high, then the spectrum was already assigned confidently in the PEAKS DB search step.

1.2. Run SPIDER Independently

The configuration panel of an independent SPIDER Search will be invoked by selecting either a *de novo* result or a data file and clicking the SPIDER icon on the toolbar  or choosing SPIDER Search from the Tools menu. Running SPIDER independently allows you to find protein matches from a homologous database even when there were no results in a conventional database search or when a database search wasn't run. The configuration panel in this case appears as follows:



The majority of options are the same as the previous section. The differences are:

- **Database** - select the sequence database for the search. If a taxa is specified, then the search is limited only to the proteins that belong to that taxa.
- **De Novo Tag** - you can either run SPIDER based on an existing *de novo* sequencing result or create a new one.

2. Understand SPIDER Results

Clicking on the “Peptide View” tab will display results that look very much like the results for PEAKS DB.

2.1. SPIDER Peptide View

Click on the "SPIDER Peptide" then the "Protein" tab to see the SPIDER matches shown in red. Note that only SPIDER homology search results will display a reconstructed or "real" sequence. If the search was done against a PEAKS DB result, the "PEAKS DB Peptide" tab will display hits from the PEAKS DB search.

SPIDER 36 [11-May-11 16:13]

1 - 1000 of 7295

scan = search

Peptide	SPIDER Score	RSD	Mass	ppm	m/z	RT	Scan	#Spec	Accession	PTM
1 KMFNVGGYLQAVLDR	40.50	0.07	1709.8923	-4.0	855.9500	61.621	F3:4342	12	P00489 PYGM_RABIT	
2 LWMNTEVLAMPYDTPVPTFK	38.13	0.38	2451.2217	17.8	1226.6400	56.166	F1:3698	1	Q8C194 PYGB_MOUSE	
3 KWCKTQVLAMPYDTPVPGYR	37.82	0.29	2451.2441	8.7	1226.6400	56.345	F3:3800	2	P00489 PYGM_RABIT	
4 KLTLALGDVNHDPVVGDR	37.04	0.47	2017.0957	-15.0	1009.5400	67.850	F1:4804	15	P00489 PYGM_RABIT	
5 RHMLGDVNHDPVVGDR	36.71	0.29	1914.9482	19.4	958.5000	60.316	F3:4212	3	P00489 PYGM_RABIT	
6 RHMLGDVNHDPVVGDR	36.71	0.29	1914.9482	19.4	958.5000	60.316	F3:4212	2	Q8C194 PYGB_MOUSE	
7 KWCKTQVLAMPYDTPVPGYR	35.82	0.33	2451.2441	8.7	1226.6400	56.345	F3:3800	2	P11216 PYGB_HUMAN	
8 KEGEYGFQNALVR	35.32	0.07	1622.8416	27.0	812.4500	96.016	F3:7021	121	P49064 ALBU_FELCA	
9 QQNVGGYLQAVLDR	35.22	0.20	1709.8735	7.0	855.9500	62.634	F3:4443	10	Q5M1B6 PYGB_SHEEP	
10 VKQKTRVSTPLVEVR	34.50	0.24	1927.1577	-4.9	643.3900	74.775	F3:5550	58	P49064 ALBU_FELCA	
11 KLTLALGDVNHDPVVGDR	34.04	0.47	2017.0957	-15.0	1009.5400	67.850	F1:4804	15	Q9VUB3 PYGM_MOUSE	
12 TDSGAHWRNVSVSEALEASTR	33.34	0.39	2456.2043	21.9	819.7600	90.107	F1:6429	2	P00330 ADH1_YEAST	

Peptide Spectrum Match Protein

P00489|PYGM_RABIT

Denovo KMFNVGGYLQAVLDR

Real KMFNVGGYLQAVLDR

Database KMFNVGGYLQAVLDR

RSD (h=0,1,2,3):
0.0667 0.0667 0.0667 0.0667

SPIDER matches shown in red

MSRPLSDQEK RKQISVRGLA GVENVTELKK NFNRLHFTL VKDRNVATPR DYYFALAHTV

1 RDHLVGRWIR TOQHYEKP KRIYLSLEF YMGRTLQNTM VNLALENACD EATYQLGLDM

EELLEEIEEDA GLNGGLGRL AACFLDSMAT LGLAAYGYI RYEFIFNOK ICGGWOMEA

61 DDWLRYGNPW EKARPEFTLP VFHYGRVEHT SQGAKWVDTO VVLAMPYDTP VEGYRNNVNV

TMRLWSAKAP NDFNLKDFNV GGYIQAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV

121 VAATLQDLIR RFKSSKFGCR DPVRTNFDAF PDKVAIQND THPSLAPEL MRVLVDLERL

DWDKAWETV KTCAYTNHTV LPEALERWPF HLETLLEPRH LQIYEINQR FLNRVAAFP

181 GDVDRLRMS LVEEGAVKRI NMAHLCIAGS HAVNGVARIH SEILKKTIFK DFELEPHKF

QNKNTGITPR RWLVCNPGI AEIIAERIGE EYISDLQDLR KLLSYDDEA FIRDVARVKQ

241 ENKLKFAAYL EREYKVHIMP NSLFDVQVQR IHEYKQQLN CLHVITLYNR IKKEPNKFPV

FRTVMIGGKA APGYHMAKMI IKLITAIGDV VNHDVPVVGDR LRVIFFENYR VSLAEKVIPA

301 ADLSEQISTA GTEASGTGNN KFMLNGALTI GTMDGANVEM AEEAGEENFF IFGMREVDVD

RLDQRGYNQA EYDRIPELR QIIEQLSSGF FSPKQPDLPK DIVNMLMHHH RFKVFADYEE

361 YVKQOERVSA LYKNPREWTR MVIRNIATSG KFSSDRITAO YAREINGVEP SRQRLPAPDE


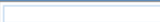
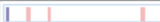

KIP

For Homology Match results, the above display contains additional information in the form of an alignment. Letters on a green background with vertical bars, indicate agreement. Letters on a red background indicate sequencing error. Colour codes on the *de novo* sequence letters still indicate positional confidence. Letters on a blue background indicate uncertainty or mutation. "+" signs represent more likely mutations. [Square brackets] indicate an equal mass substitution, common non-critical *de novo* errors. While <Angle brackets> indicate an equal mass substitution and a mutation.

When simply identifying exact peptides from the database using PEAKS DB, SPIDER tag match or inChorus, there is no need to reconstruct the 'real' sequence.

2.2. SPIDER Protein View

Clicking on "Protein View" will yield a similar display as was seen for PEAKS DB. When running SPIDER on its own, all peptides in the "Coverage" column will appear red. However, when SPIDER is searched based on a PEAKS DB result, the column will hold additional information. A red region indicates areas of homology and potential mutation identified by SPIDER. Blue regions indicate areas of homology from the database search, and purple indicates that the spectrum was identified by both SPIDER and database search.

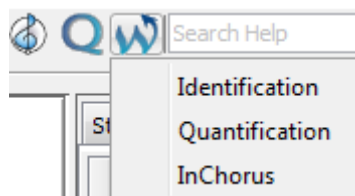
Accession	-10lgP	Coverage	#SPIDER Pe...	#SPIDE...	#DB P...	#DB U...	Mass	Description	...
Q9WUB3 P...	63.95	 29%	30	4	2	2	97286...	Glycogen phosphoryl...	<input checked="" type="checkbox"/>
Q8WZ42 TI...	62.40	 5%	117	73	32	18	381608...	Titin OS=Homo sapie...	<input checked="" type="checkbox"/>
Q8BPX9 S1...	60.82	 10%	4	4	1	1	64050...	Solute carrier family ...	<input checked="" type="checkbox"/>
Q6N8V0 DN...	49.18	 9%	5	5	3	3	13339...	Error-prone DNA pol...	<input checked="" type="checkbox"/>

Click on the “Coverage” tab. The highlighted colors are the same as those shown in the “Display” column above.

SPIDER Peptides	PEAKS DB Peptides	Coverage	Tool Box
<pre>>sp Q8BPX9 S15A3_MOUSE Solute carrier family 15 member 3 OS= Mus musculus GN=Slc15a3 PE=2 SV=1</pre>			
1	MSAPRAEEQP SR	SGERQPLV ARGPRGPRRW RRTAAAVLL VQMLERAAFF GVTSNLVLYL	
61	NSLNFNWDGQ QASRATLLFL GAS	SYLLAPVG GWLADVYLGR FLTISLSLLL YLAASGLLLT	
121	TITNDGRRSF CGEMPPLPLE PACPSSSCQG SWSSPYCATT	LYLVLLLLAL AASSVRSTLT	
181	SFGADQVMDL GRDATRRFFN WFYWSINLGA ILSLLVVAFI EQNISFLWGY SIIVGLVGLA		
241	FFIFLFATPV FITKPPTGSQ VSSMLKLAFO NCCPCRRSSS RDESAAHLLP DQRSNQPGPS		
301	PQEDMANFQV LVKILPVMVT LVPYWMVYFQ MQSTYVLQGL HLHIPNIFRT NPNISLLLRS		
361	DSSNYRIPEA WLLLANVAVI LILIPVKDHL IDPLLLRCKL LPSSLQKMAL GMFFGFTSII		
421	VAGVLEKERL OYIAANOTVP OLIGKDLYYA APLSIWWOIP OYLLIGVSEI FASIPGLEFA		


Chapter 16. Creating a High-Throughput Workflow

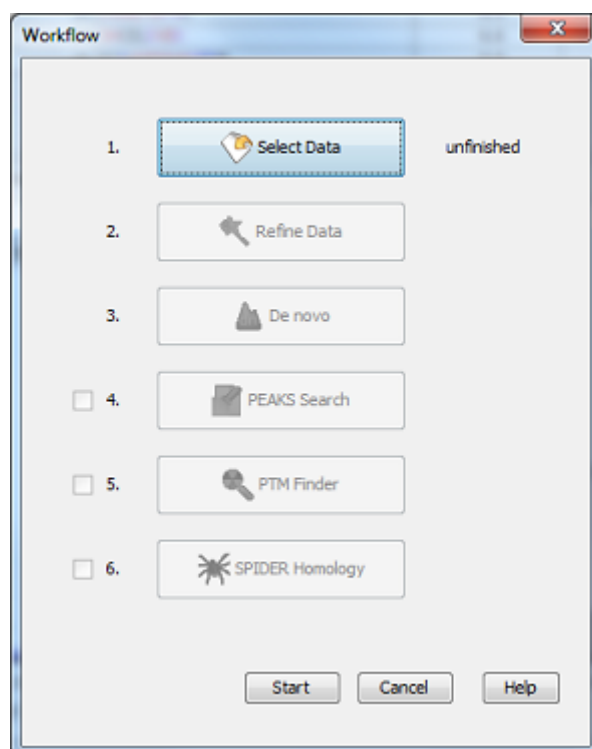
For your convenience, PEAKS provides workflows for protein identification, quantification and inChorus search (multi-engine protein ID).



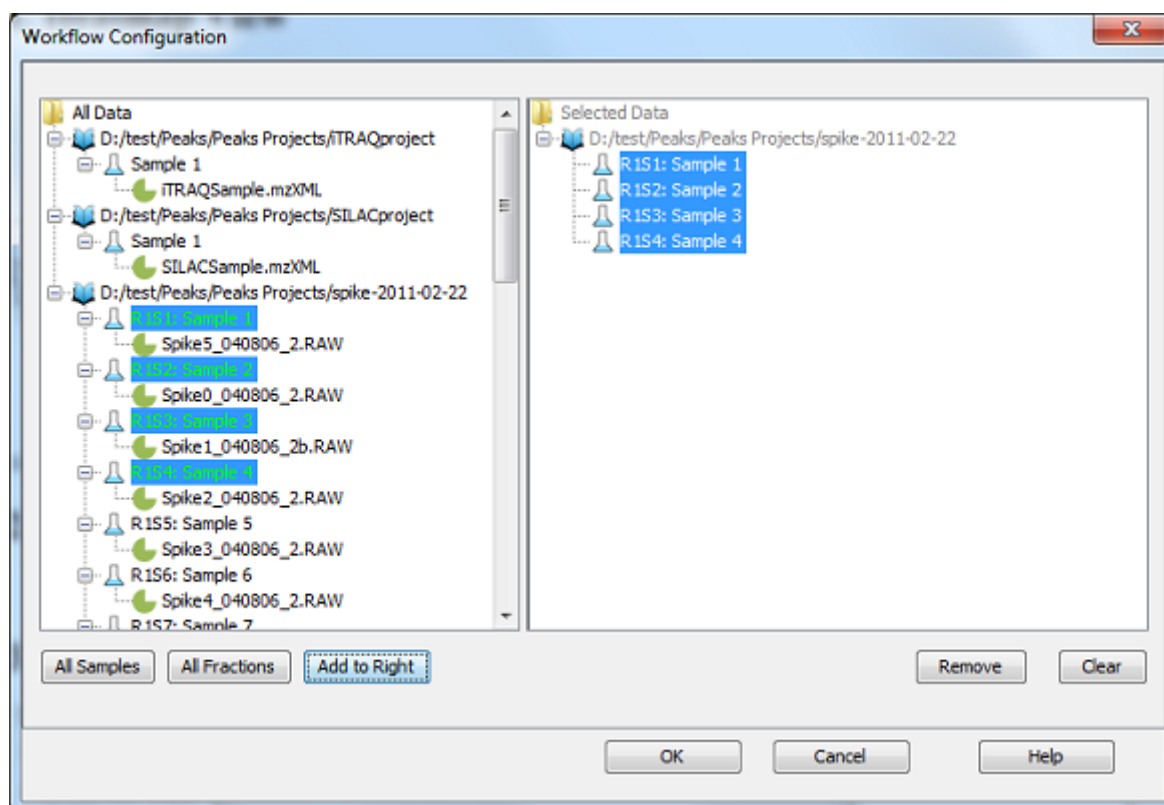
Once a specific workflow is selected, a dialogue pops up to allow you to specify the analysis steps and the parameters to use in each step.

1. Identification Workflow

Click the workflow icon on the toolbar  and select "Identification", the identification workflow setup window will appear:



Click "Select Data" to open the "Workflow Configuration" dialogue where you can select the data you wish to perform identification analyze. Only projects that are open in the "Project View" panel can be selected for analysis. To select which files/samples you would like to analyze, either select the individual file/sample, click the "All Samples" or the "All Fractions" buttons and then click the "Add to Right" to transfer the samples/files to the "Selected Data" list on the right hand side. Use the "Remove" and "Clear" buttons to remove selected files/samples or all files/samples respectively from the "Selected Data" list. Click "OK" to proceed to the next step.



Note

All files loaded in a single workflow will be processed in exactly the same way, using exactly the same parameters. If you want to do some differently than others, you must set up separate workflows.

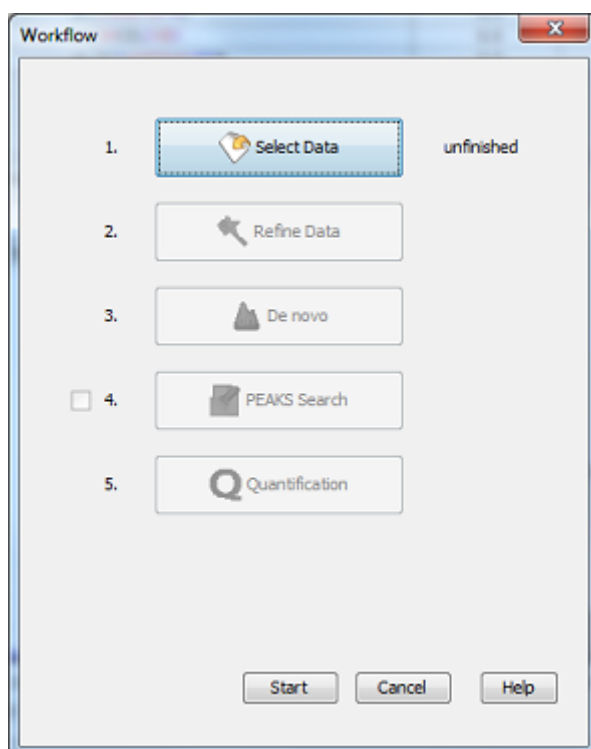
Once the data is selected, you can specify parameters for the downstream analysis steps one by one by clicking the other buttons in the workflow dialogue. Please refer to the chapters on each individual function if you require more details on setting up the parameters.

Note

PEAKS DB search, PTM Finder and SPIDER are the optional functions in Identification Workflow. You can uncheck them if you do not want to perform those functions.

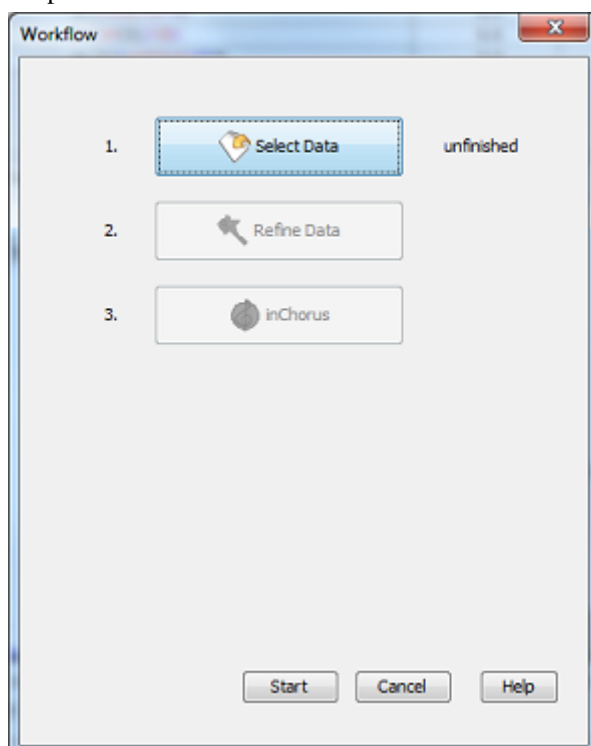
2. Quantification Workflow

The quantification workflow is similar to the identification workflow except, after an optional PEAKS DB, you can define quantification parameters to perform labeled or label free quantification.



3. inChorus Workflow

The inChorus workflow is similar to the identification workflow except that after “Select Data” and “Refine Data” steps, you can specify inChorus parameters to start an inChorus search.

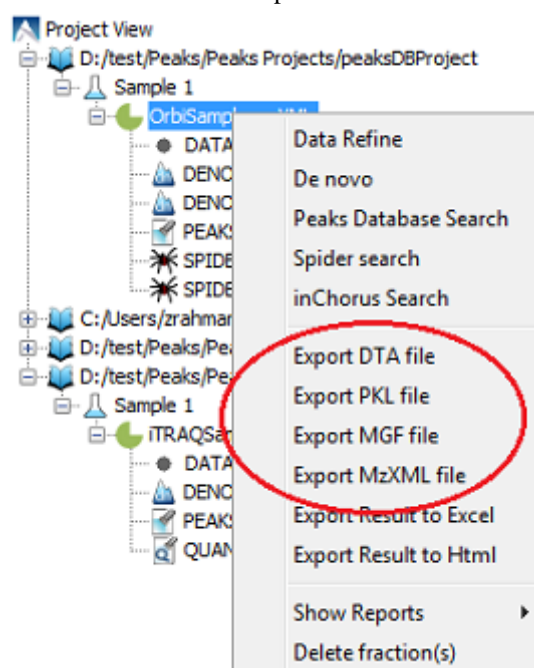


Chapter 17. Exporting Data/Reports and Printing

PEAKS 5.3 offers a rich collection of exporting functions to allow users to create reports to share the analysis results with collaborators, colleagues and clients. The supported formats include HTML, Comma Separated Values (CSV), pepXML, and various image formats for image exporting. Labs with in-house software can easily make use of the CSV files in their own analysis workflow. The exported results in HTML can be viewed with a web browser. The entire exported result directory can be zipped and emailed to colleagues, or the whole directory can be put directly on a website.

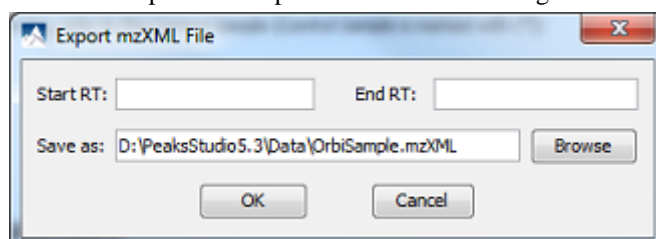
1. Export Data

Data can be exported in a number of file formats including .mzxml, .mgf, DTA, MGF and PKL. To do so, right click on the sample node or the data file that is to be exported and select the desired export format.



Clicking “Export DTA file” will open a dialog prompting for the folder name and location to which DTA files will be exported. For MGF and PKL, the dialog will ask for a name and a location for the file.

Clicking “Export MzXML File” will open the “Export mzXML File” dialog.



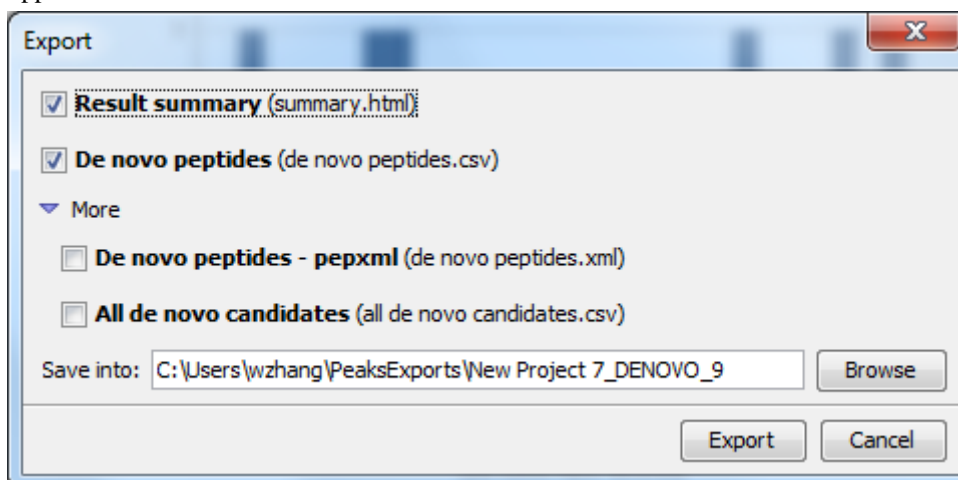
Enter the starting and ending retention times in the appropriate boxes. Then click the “Browse” button to select a destination to save your file.

2. Export *De Novo* Result

PEAKS *de novo* sequencing result can be exported to other supported formats. All export functions are available through “Summary” view panel.

2.1. Export Summary and Peptides

To export results, press the “Export” button in the title bar of the “Summary” view panel. The following export dialog will appear.



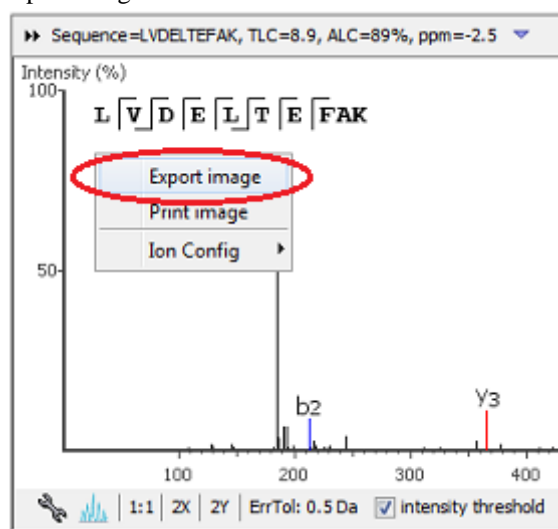
The following exporting options are available:

- **Result summary:** The “Summary” view page will be saved in “summary.html” file in HTML format in the specified location.
- **De novo peptides:** The peptides identified by *de novo* sequencing will be saved in “peptides.csv” file in Comma-separated Values (CSV) format in the specified folder.
- **De novo peptides - pepXML:** In addition to CSV format, the peptides will be saved in pepXML format.

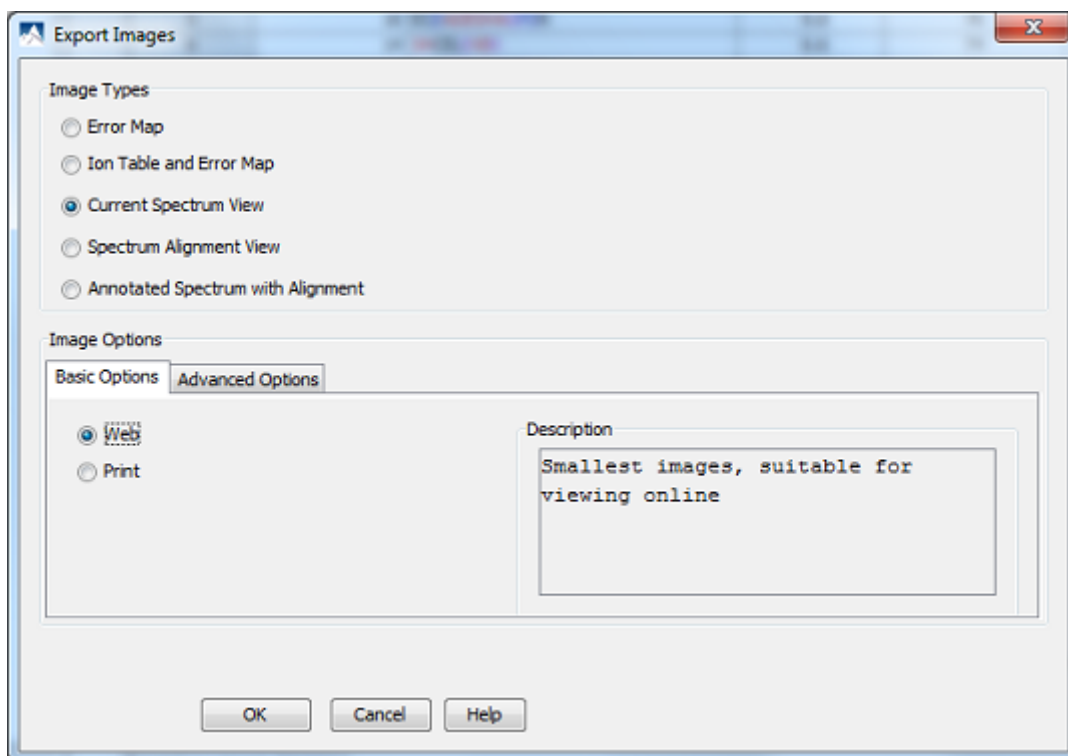
Click the “Export” button to save the selected result components to the specified location.

2.2. Export Images

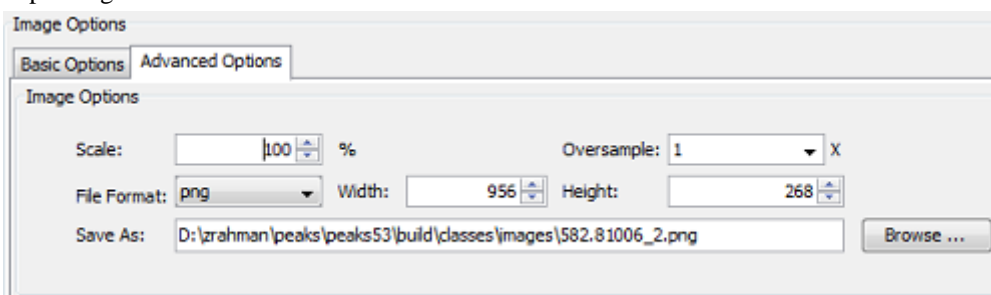
The annotated Spectrum, Ion Match table, Error Map, and Spectrum Alignment can all be exported to image files. To do so, position the cursor on any of those items in the result panel and click the right mouse button to view the pop-up menu and select the “Export Image” command from the menu.



This will display the “Export Images” dialog for selecting the result items to export.



Select the desired result elements from the “Image Types” list. The “Basic Options” panel offers two choices for image size: “Web” - small images that are suitable for online viewing and “Print” - oversampled images that are suitable for printing.



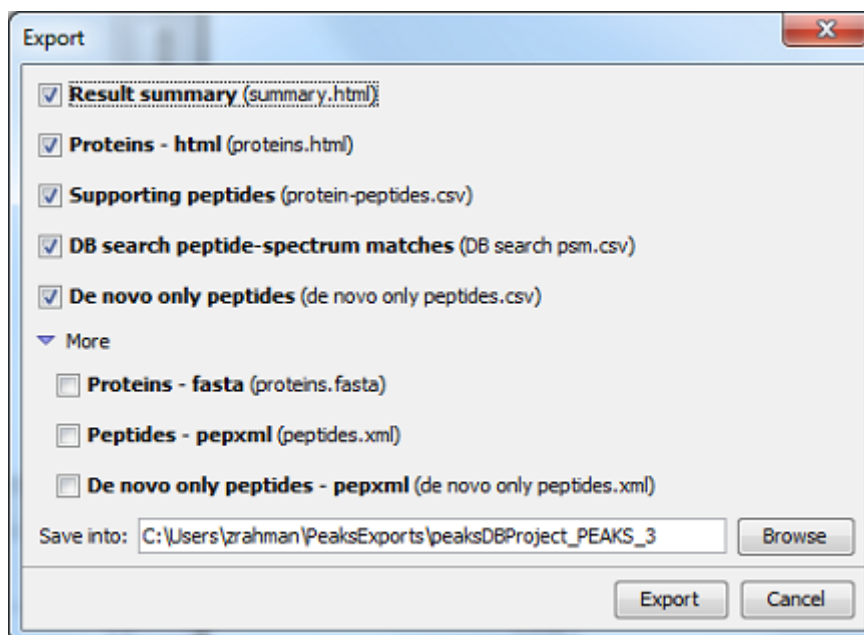
The “Advanced Options” panel offers choices for scaling factor, file format, resolution, oversample factor and location to save the images. PEAKS supports BMP, GIF, JPEG, PNG, and SVG image formats. After setting all parameters, click the “OK” button to export the selected result item to an image.

3. Export PEAKS DB Result

PEAKS DB results and PTM Finder results can be exported to other supported formats. All export functions are available through “Summary” view panel.

3.1. Export Summary, Proteins and Peptides

To export the result, press the “Export” button in the title bar of the “Summary” view panel. The following export dialog will appear.



The following exporting options are available:

- **Result summary:** The “Summary” view page will be saved in “summary.html” file in HTML format in the specified location.
- **Proteins - html:** A list of protein identifications will be saved in “proteins.html” file in HTML format in the specified location.
- **Supporting peptides:** A list of supporting peptides for each protein identification will be saved in “protein-peptides.csv” file in Comma Separated Values (CSV) format in the specified folder.
- **DB search peptide-spectrum matches:** A list of peptide-spectrum matches with scores greater than the threshold will be saved in “DB search psm.csv” file in CSV format in the specified location.
- **De novo only peptides:** A list of good *de novo* sequences that do not have good or no database matches will be saved in “*de novo* only peptides.csv” file in CSV format in the specified location.
- **Proteins - fasta:** A list of protein identifications will be saved in “proteins.fasta” file in FASTA format in the specified location.
- **Peptides - pepxml:** A list of peptide-spectrum matches will be saved in “peptides.xml” file in pepXML format in the specified location.
- **De novo only peptides - pepxml:** A list of good *de novo* sequences that do not have good or no database matches will be saved in “*de novo* only peptides.xml” file in pepXML format in the specified location.

Click the “Export” button to save the selected result components to the specified location.

3.2. Export Images

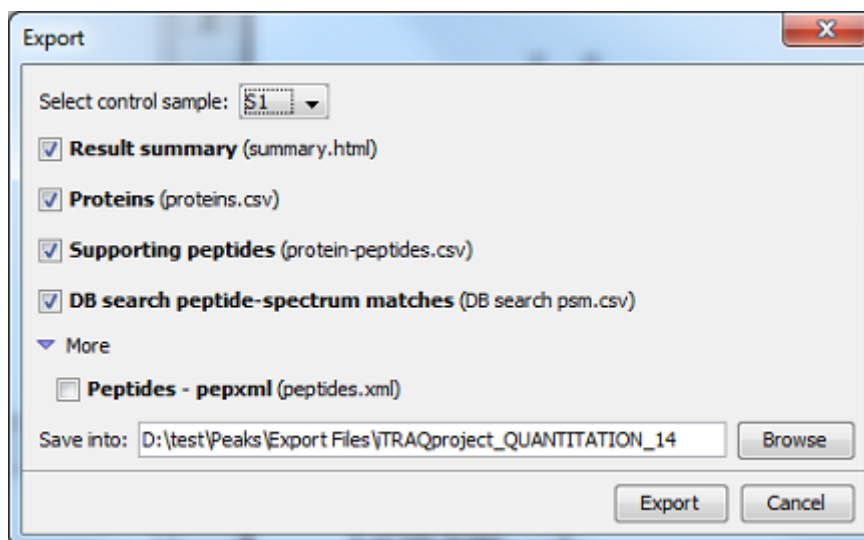
From the “Peptide” view and the “*De novo* only” view, the Annotated Spectrum, Ion Match table, Error Map, or Spectrum Alignment can be exported to an image file. To do so, position the cursor on any of those items in the result panel, right click and select the “Export Image” command from the menu. Refer to Section 2.2, “Export Images” for details.

4. Export Quantification Results

PEAKS Q(labeled and label free quantification) results can be exported to other supported formats. All export functions are available through “Summary” view panel.

4.1. Export Labeled Quantification Results

To export the result, press the “Export” button in the title bar of the “Summary” view panel. The following export dialog will appear.



The following exporting options are available:

- **Result summary:** The “Summary” view page will be saved in “summary.html” file in HTML format in the specified location.
- **Proteins:** A list of protein identifications will be saved in “proteins.csv” file in Comma-Separated Values (CSV) format in the specified location.
- **Supporting peptides:** A list of supporting peptides for each protein identification will be saved in “protein-peptides.csv” file in CSV format in the specified folder.
- **DB search peptide-spectrum matches:** A list of peptide-spectrum matches will be saved in “db search psm.csv” file in CSV format in the specified location.
- **Peptides - pepxml:** A list of peptide-spectrum matches will be saved in “peptides.xml” file in pepXML format in the specified location.

Click the “Export” button to save the selected result components to the specified location.

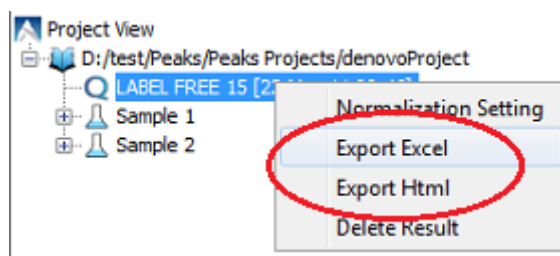
From the “Peptide” view, the Annotated Spectrum, Ion Match table, Error Map, or Spectrum Alignment can be exported to an image file. To do so, position the cursor on any of those items in the result panel, right click and select the “Export Image” command from the menu. Refer to Section 2.2, “Export Images” for details.

4.2. Export Label-Free Quantification Results

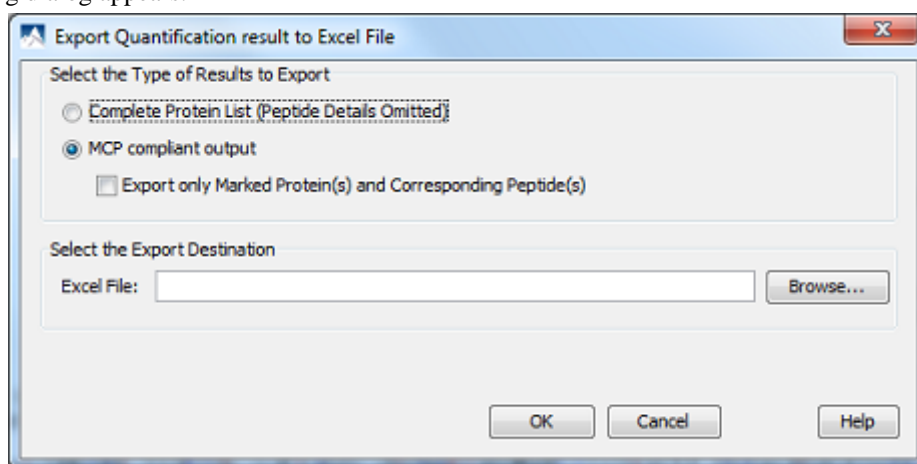
The label-free quantification results can be exported in Excel (.xls) or HTML (.html) format.

4.2.1. Export Result in Excel or HTML

To export the label free quantification results, right click on a label-free result node and choose “Export HTML” or “Export Excel” command from the pop-up menu.



The following dialog appears.



PEAKS provides two types of exporting functions: “complete protein list” without peptide details, or “MCP compliant output”. When you select MCP compliant output, you can check the “Export only Marked Protein(s) and Corresponding Peptides(s)” if you are only interested in some proteins and previously marked them in the result table.

The output of “Complete Protein List” consists of two major sections: one is the representations table which displays a representative protein for each cluster; the other is the whole protein table, which lists all the clustered proteins.

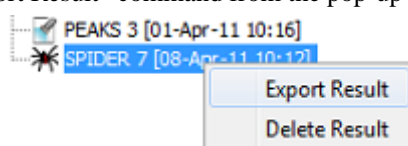
The MCP compliant output contains the two tables described above; however, it also provides more information than that in the whole protein table. These additions include all of the supporting peptides and their coverage within the protein. False discovery rate (FDR) estimation is also displayed if PEAKS DB was run with a decoy database. The results also include the “Single-Peptide Based Protein” table which contains all the proteins with only one supporting peptide detected.

4.2.2. Export Summary Page

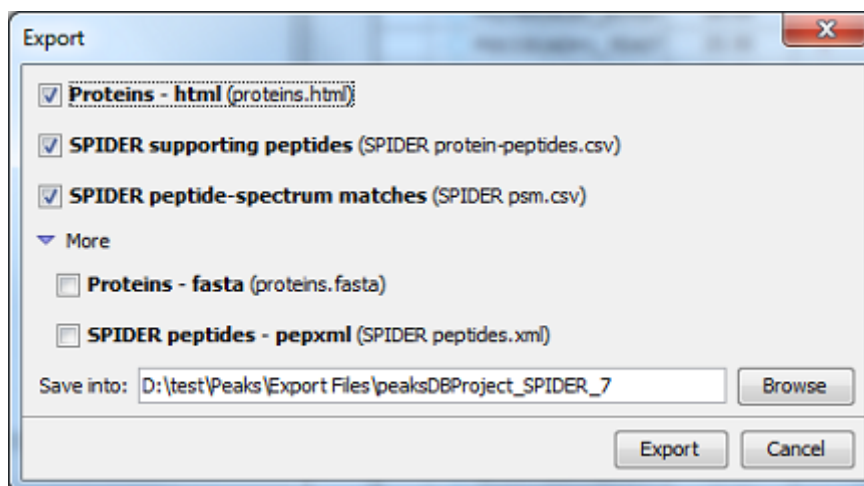
The “Summary” view page can be exported in HTML format. To export the summary page, click the “Export” button in the “Summary” view. This opens a “Save” dialog box for specifying the filename and location of the result.

5. Export SPIDER Result

PEAKS' SPIDER search result can be exported to other supported formats. To do so, select the SPIDER result node, right click and select the “Export Result” command from the pop-up menu.



The following export dialog will appear with the available exporting options listed below:



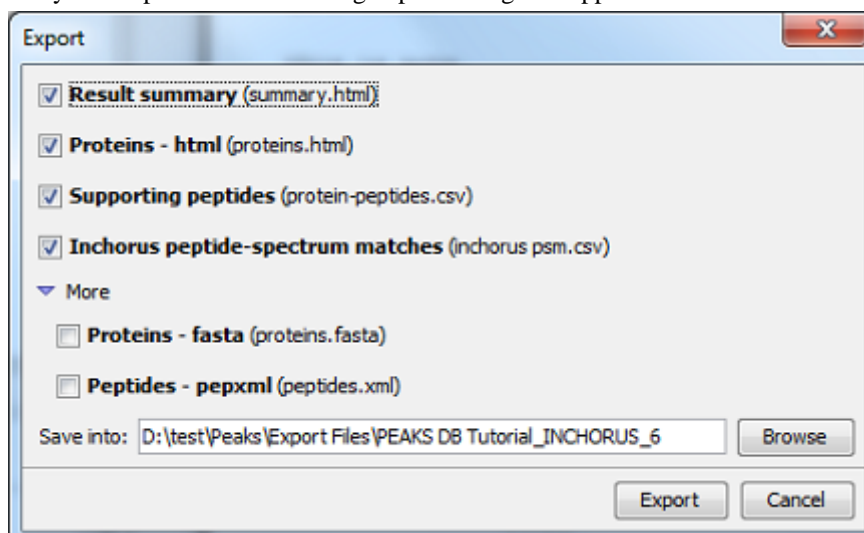
- **Proteins - html:** A list of protein identifications using SPIDER search will be saved in “proteins.html” file in HTML format in the specified location.
- **SPIDER supporting peptides:** A list of supporting peptides for each protein identification will be saved in “SPIDER protein-peptides.csv” file in Comma Separated Values (CSV) format in the specified folder.
- **SPIDER peptide-spectrum matches:** A list of peptide-spectrum matches will be saved in “SPIDER psm.csv” file in CSV format in the specified location.
- **Proteins - fasta:** A list of protein identifications will be saved in “proteins.fasta” file in FASTA format in the specified location.
- **SPIDER Peptides - pepxml:** A list of peptide-spectrum matches will be saved in “peptides.xml” file in pepXML format in the specified location.

Click the “Export” button to save the selected result components to the specified location.

From the “Peptide” view, the Annotated Spectrum, Ion Match table, Error Map, or Spectrum Alignment can be exported to an image file. To do so, position the cursor on any of those items in the result panel, right click and select the “Export Image” command from the menu. Refer to Section 2.2, “Export Images” for details.

6. Export inChorus Result

The inChorus result can be exported to other supported formats. To do so, click the “Export” button in the title bar of the “Summary” view panel. The following export dialog will appear.



The following exporting options are available:

- **Result summary:** The “Summary” view page will be saved in “summary.html” file in HTML format in the specified location.
- **Proteins - html:** A list of protein identifications will be saved in “proteins.html” file in HTML format in the specified location.
- **Supporting peptides:** A list of supporting peptides for each protein identification will be saved in “protein-peptides.csv” file in Comma Separated Values (CSV) format in the specified folder.
- **inChorus peptide-spectrum matches:** A list of peptide-spectrum matches will be saved in “inchorus psm.csv” file in CSV format in the specified location.
- **Proteins - fasta:** A list of protein identifications will be saved in “proteins.fasta” file in FASTA format in the specified location.
- **Peptides - pepxml:** A list of peptide-spectrum matches will be saved in “peptides.xml” file in pepXML format in the specified location.


Click the “Export” button to save the selected result components to the specified location.

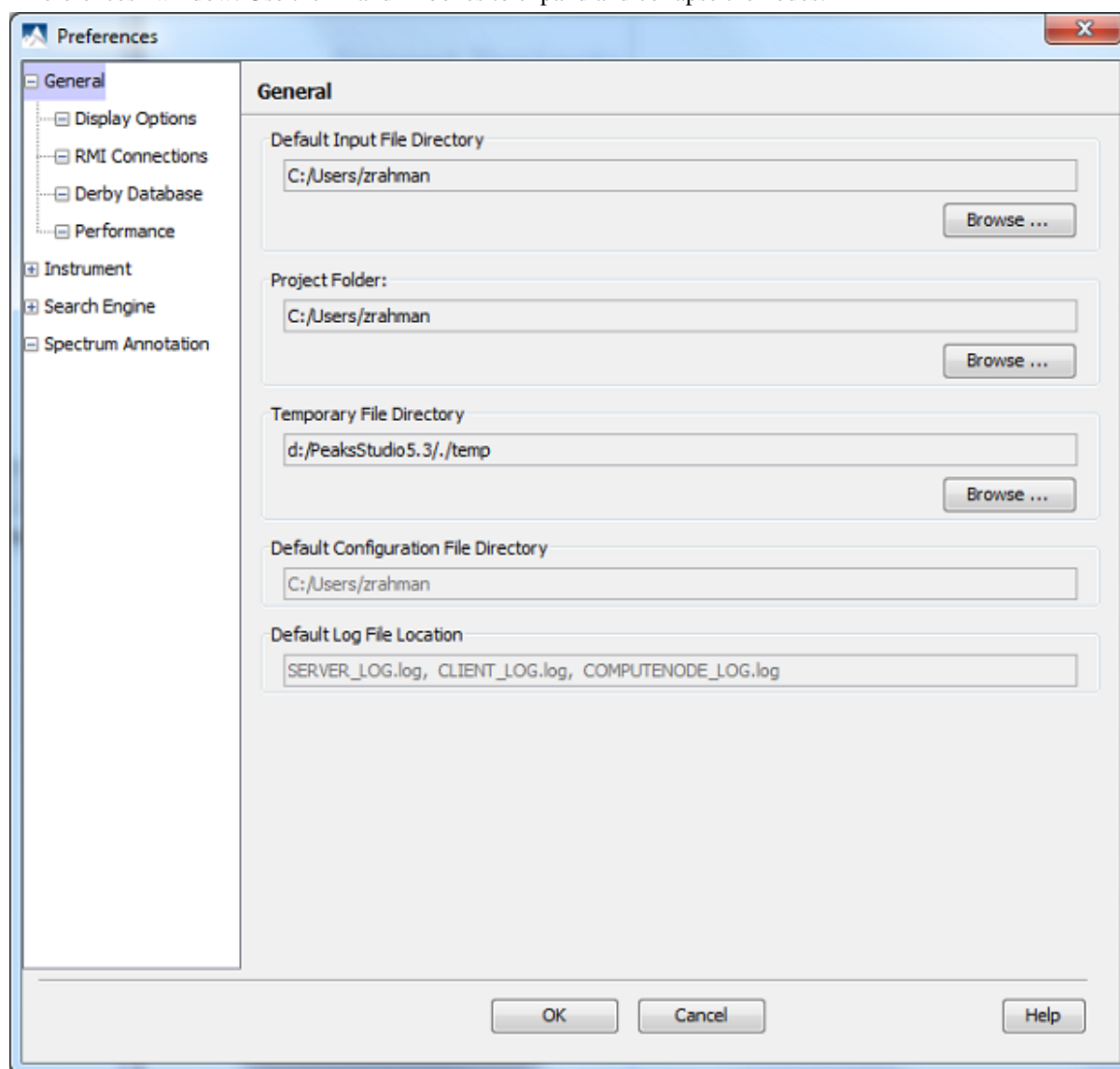
From the “Peptide” view, the Annotated Spectrum, Ion Match table, Error Map, or Spectrum Alignment can be exported to an image file. To do so, position the cursor on any of those items in the result panel, right click and select the “Export Image” command from the menu. Refer to Section 2.2, “Export Images” for details.

Chapter 18. Advanced Configuration and Environment Preferences

1. PEAKS Environment Preferences

This section will describe the settings of the environmental preferences including general, instrument, search engine and ion editor configurations.

To begin, click the “Preferences” toolbar icon  or select “Preferences” from the “Window” menu to open the “Preferences” window. Use the '+' and '-' boxes to expand and collapse the nodes.



1.1. General Preferences

Default Input File Directory- Select the “Browse” button to change the directory that will appear when adding data to a project.

Project folder- PEAKS uses [USER_HOME] as the default output folder for project files, where USER_HOME is the user home directory in your system. Select the “Browse” button to change this location.

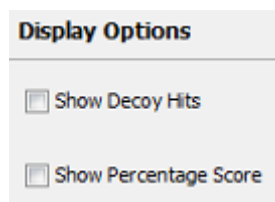
Temporary File Directory- PEAKS uses [PEAKS_HOME_DIRECTORY]/temp as the default temporary file output directory, where PEAKS_HOME_DIRECTORY is the location where PEAKS is installed. Select the “Browse” button to change this location.

Default Configuration File Directory- Configuration files for PEAKS can be found at [USER_HOME]/.peaks by default. These files' locations cannot be changed.

Default Log File Location- Log files for PEAKS can be found at [USER_HOME]/.peaks by default. These files' locations cannot be changed.

1.1.1. Display Options

Clicking on “Display Options” on the menu on the left hand side will display interface preferences on the right hand side.

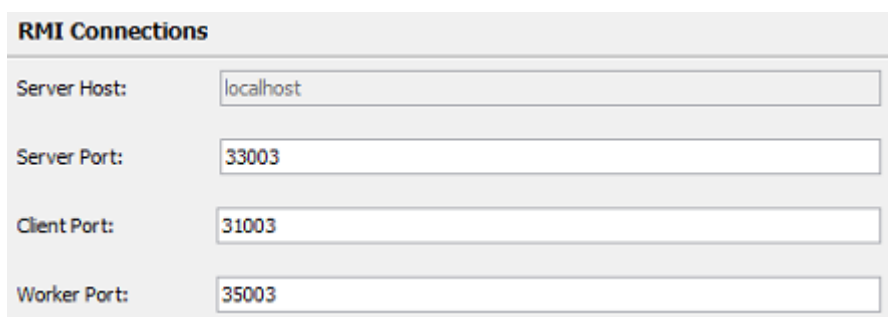


Check “Show Decoy Hits” to display protein and peptide hits from decoy database in PEAKS DB results.

PEAKS uses -10lgP to display its results by default. To view the percentage score along with -10lgP in peptide and protein view as well as the exporting files for PEAKS DB result, check "Show Percentage Score".

1.1.2. RMI Connections

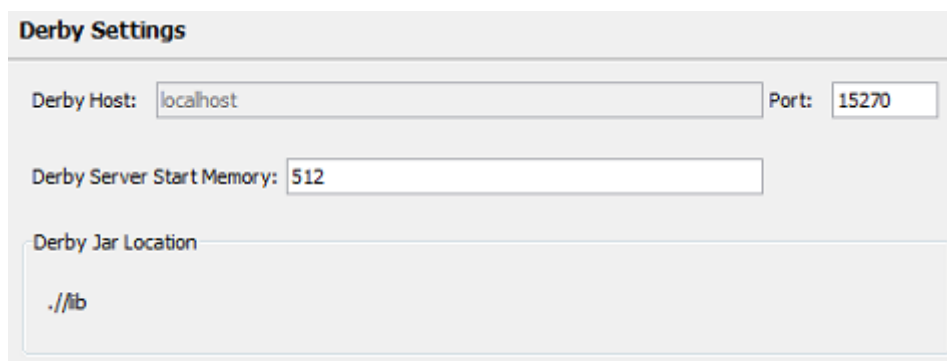
Clicking on “RMI Connections” on the menu on the left hand side will show the RMI connections preferences on the right hand side.



The default port numbers for the Server, Client and Worker will appear. The port numbers can be changed if conflicts arise. Contact technical support at BSI for more information.

1.1.3. Derby Database

Clicking on “Derby Database” on the menu on the left hand will show the derby database preferences on the right hand side.



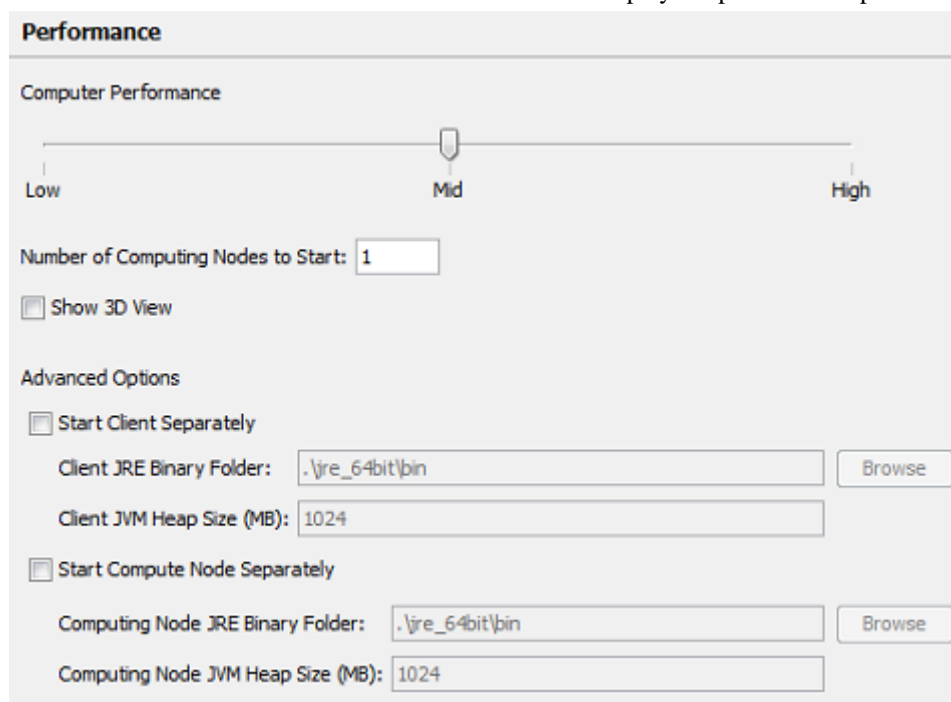
Derby Host. The name of the “Derby Host” as well as the “Port” number will come up by default. The port number can be changed.

Derby Server Start Memory. The amount of memory used to start the derby server will also come up by default but can be changed if more memory is available; however, it is not recommended to change this from the default setting. To increase performance, use the performance settings or the memory utility (see Section 5, “Adjusting PEAKS’ Memory Usage”).

Derby Jar Location. The “Derby Jar Location” panel will list the location of the derby jar file by default. This is displayed to find its location. Its location cannot be changed.

1.1.4. Performance

Clicking on “Performance” on the menu on the left-hand side will display the performance preferences.



The screenshot shows the 'Performance' settings window. At the top is a 'Computer Performance' slider with 'Low', 'Mid', and 'High' markers; the slider is currently positioned at 'Mid'. Below this is a 'Number of Computing Nodes to Start:' field with the value '1'. There is a checkbox for 'Show 3D View' which is currently unchecked. Under the 'Advanced Options' section, there are two checkboxes: 'Start Client Separately' and 'Start Compute Node Separately', both of which are unchecked. For 'Start Client Separately', there is a text field for 'Client JRE Binary Folder:' containing '.\jre_64bit\bin' and a 'Browse' button. Below that is a text field for 'Client JVM Heap Size (MB):' containing '1024'. For 'Start Compute Node Separately', there is a text field for 'Computing Node JRE Binary Folder:' containing '.\jre_64bit\bin' and a 'Browse' button. Below that is a text field for 'Computing Node JVM Heap Size (MB):' containing '1024'.

Computer Performance. For 512 MB to 1 GB of RAM, select low. Select medium for 1 to 2 GB of RAM. Select high for 2 GB or more.

3D View. PEAKS can display a 3D view for quantification results. Check the “Show 3D View” box to enable this function. PEAKS comes with the Java3D program to support the viewing of 3D images. If this feature is not required, deselect it to increase performance.

Advanced Options. These options are present for users who want to take full advantage of their 64 bit Windows operating system.

- **Start Client Separately:** Select this option to load the client JRE separately. Select “Browse” to choose the location of the “Client JRE Binary folder”. It is recommended that the client JVM be a 32 bit JRE. The default setting is the 32 bit JRE which comes with PEAKS.
- **Start Compute Node Separately:** Select this option to load the compute node separately. It is possible to choose a 64 bit JRE downloaded from Java Sun by clicking the “Browse” button. Select the location of the JRE binary folder. With a 64 bit JRE, the Heap Space can be set to any amount within the maximum physical memory of the computer.

1.2. Instrument Preferences

This section allows users to change preferences for the following instruments: ABI, Bruker, Shimadzu and Varian.

1.2.1. ABI (.wiff)

Clicking on “Instrument” and then “ABI (.wiff)” on the menu on the left hand side will show the preferences for ABI instrument.

Note

Refer to Section 3.4.1, “QSTAR or QTRAP” for details on ABI(.wiff) preferences.

1.2.2. Bruker (.yep/baf, fid)

Clicking on “Instrument” and then “Bruker (.yep/baf, fid)” in the menu on the left hand side will display the Bruker instrument preferences.

Note

Refer to Section 3.5, “Bruker Data” for details on Bruker instrument preferences.

1.2.3. Shimadzu Axima (.run)

Clicking on “Instrument” and then “Shimadzu AXIMA (.run)” in the menu on the left hand side will show the Shimadzu instrument preferences.

Note

Refer to Section 3.6, “Shimadzu Data” for details on Shimadzu instrument preferences.

1.2.4. Varian (.xms)

Clicking on “Instrument” and then “Varian (.xms)” in the menu on the left hand side will display the Varian instrument preferences.

Note

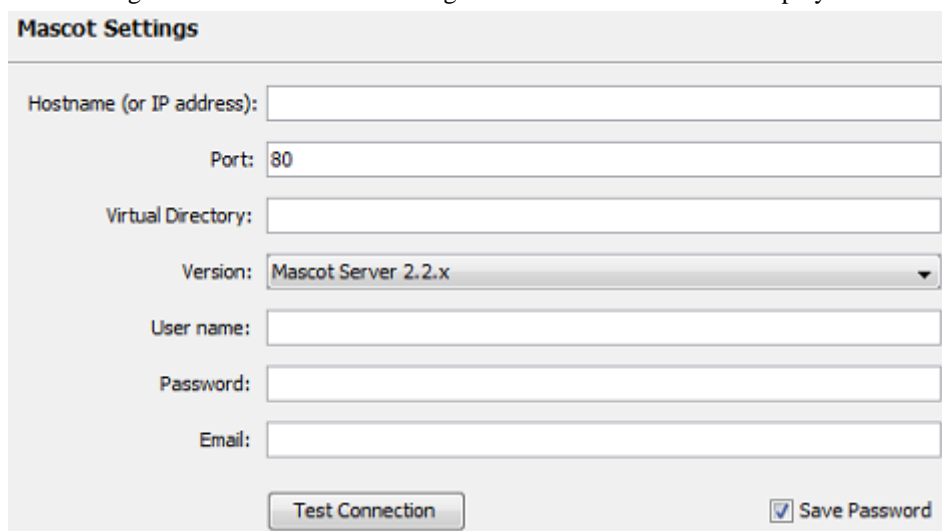
Refer to Section 3.7, “Varian” for details on Varian instrument preferences.

1.3. Search Engine Preferences

This section allows users to configure preferences for the following search engines: Mascot, X!Tandem, OMSSA, and Sequest.

1.3.1. Mascot Settings

Clicking on “Search Engine” and then “Mascot Settings” on the left-hand side will display the Mascot preferences.



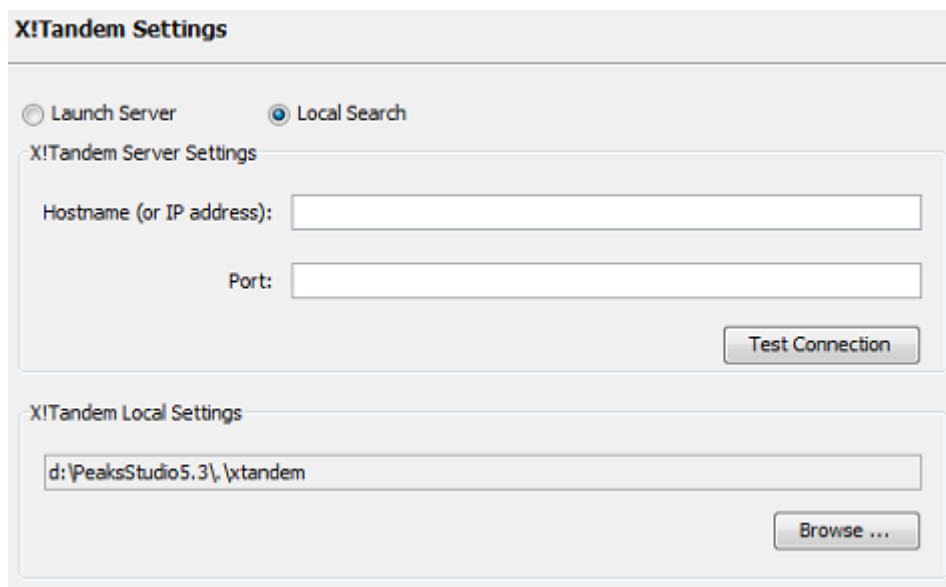
The screenshot shows the "Mascot Settings" window with the following fields and controls:

- Hostname (or IP address):
- Port:
- Virtual Directory:
- Version:
- User name:
- Password:
- Email:
- Test Connection:
- Save Password: ☒ Save Password

The settings parameters specify how PEAKS accesses the Mascot server (if applicable). Enter the hostname (or an IP address), port, virtual directory, Mascot server version as well as your username, password and email address. To make sure that everything is entered correctly and that the server is working, click the “Test Connection” button. To save the password and avoid entering it every time, check the “Save Password” box.

1.3.2. X!Tandem Settings

Clicking on “Search Engine” and then “XTandem Settings” on the left-hand will display the X!Tandem preferences.



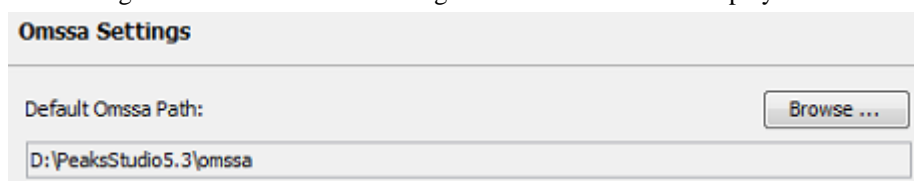
The screenshot shows the "X!Tandem Settings" dialog box. At the top, there are two radio buttons: "Launch Server" (unselected) and "Local Search" (selected). Below this, there are two sections. The first section, "X!Tandem Server Settings", contains two text input fields: "Hostname (or IP address):" and "Port:". To the right of these fields is a "Test Connection" button. The second section, "X!Tandem Local Settings", contains a text input field with the default path "d:\PeaksStudio5.3\xtandem". To the right of this field is a "Browse ..." button.

Start by selecting whether PEAKS should access a server or local version of X!Tandem. For the server version, enter the hostname or IP address as well as the port. To make sure that everything is entered correctly and that the server is working, click the “Test Connection” button.

As PEAKS provides a local copy of X!Tandem, upon installation, a default path will appear in the Local Settings section. To use another license/location for X!Tandem, click the “Browse” button to tell PEAKS where to find the search engine.

1.3.3. OMSSA Settings

Clicking on “Search Engine” and then “Omssa Settings” on the left hand will display the OMSSA preferences.




The screenshot shows the "Omssa Settings" dialog box. It contains a single section with a text input field labeled "Default Omssa Path:" which contains the default path "D:\PeaksStudio5.3\omssa". To the right of this field is a "Browse ..." button.

As PEAKS provides a local copy, upon installation, a default path will appear here. To use another license/location for OMSSA, click the “Browse” button to tell PEAKS where to find the desired search engine.

1.3.4. Sequest Settings

Clicking on “Search Engine” and then “Sequest Settings” on the left-hand will display Sequest preferences.



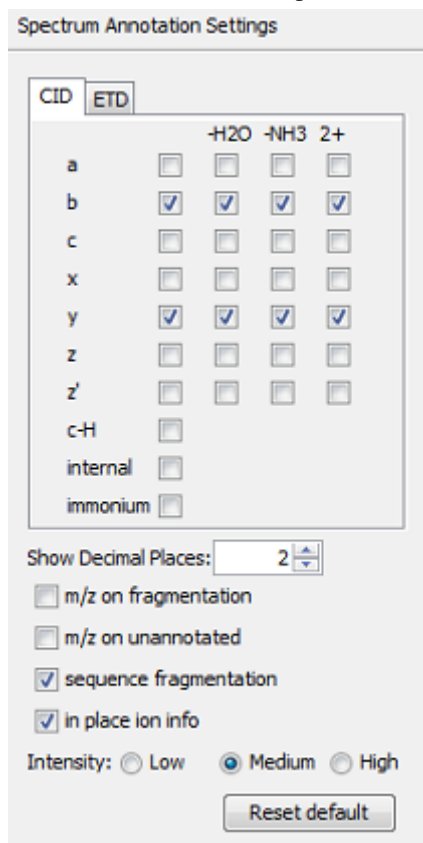
The Sequest Settings dialog box contains three rows, each with a text field and a 'Browse ...' button:

- Sequest Location: [text field] [Browse ...]
- Default Sequest Parameter File (.params): [text field] [Browse ...]
- Sequest Result Output Folder: [text field] [Browse ...]

To use Sequest, click the “Browse” button to tell PEAKS where to find the search engine. Make sure to specify the location of the “Default Sequest Parameter File (.params)” and the “Sequest Result Output Folder”.

1.4. Spectrum Annotation Preferences

Clicking on “Spectrum Annotation” on the left-hand side will open the following window:



The Spectrum Annotation Settings dialog box has two tabs: CID and ETD. The ETD tab is selected. Below the tabs is a table of ion types with checkboxes for selection. The table has columns for -H2O, -NH3, and 2+.

	-H2O	-NH3	2+
a	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
c	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
x	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
y	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
z	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
z'	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c-H	<input type="checkbox"/>		
internal	<input type="checkbox"/>		
immonium	<input type="checkbox"/>		

Below the table are several options:

- Show Decimal Places: 2
- ☐ m/z on fragmentation
- ☐ m/z on unannotated
- ☒ sequence fragmentation
- ☒ in place ion info
- Intensity: ☐ Low ☒ Medium ☐ High
- [Reset default]

The annotated spectrums in PEAKS results can be annotated by the selected ion types from a big collection of ions that PEAKS offers. The selected ion types will be displayed in the “Ion Match” table as well. It is possible to annotate the spectrum with various ions for both CID and ETD. By default, y-ion, y-H2O, y-NH3, y(2+), b-ion, b-H2O, b-NH3, b(2+) are selected.

Note

If you are upgrading from an earlier version of PEAKS, or simply wish to reset the settings, use the “Reset default” button to update to the 5.3 defaults.

Show Decimal places. Select the number of decimal places that will appear in the ion table and spectrum view. The default is set to two decimal places.

m/z on fragmentation. Select this to display m/z value on top of the annotated ions.


m/z on unannotated. Select this to display m/z value on top of the peaks without ions.

sequence fragmentation. Select this to display sequence fragmentation on top left corner of the “Spectrum Annotation” view.

in place ion info. Ion information, m/z value and relative intensity are displayed in a pop-up in the “Spectrum Annotation” view when this option is checked and the cursor is placed on a peak.

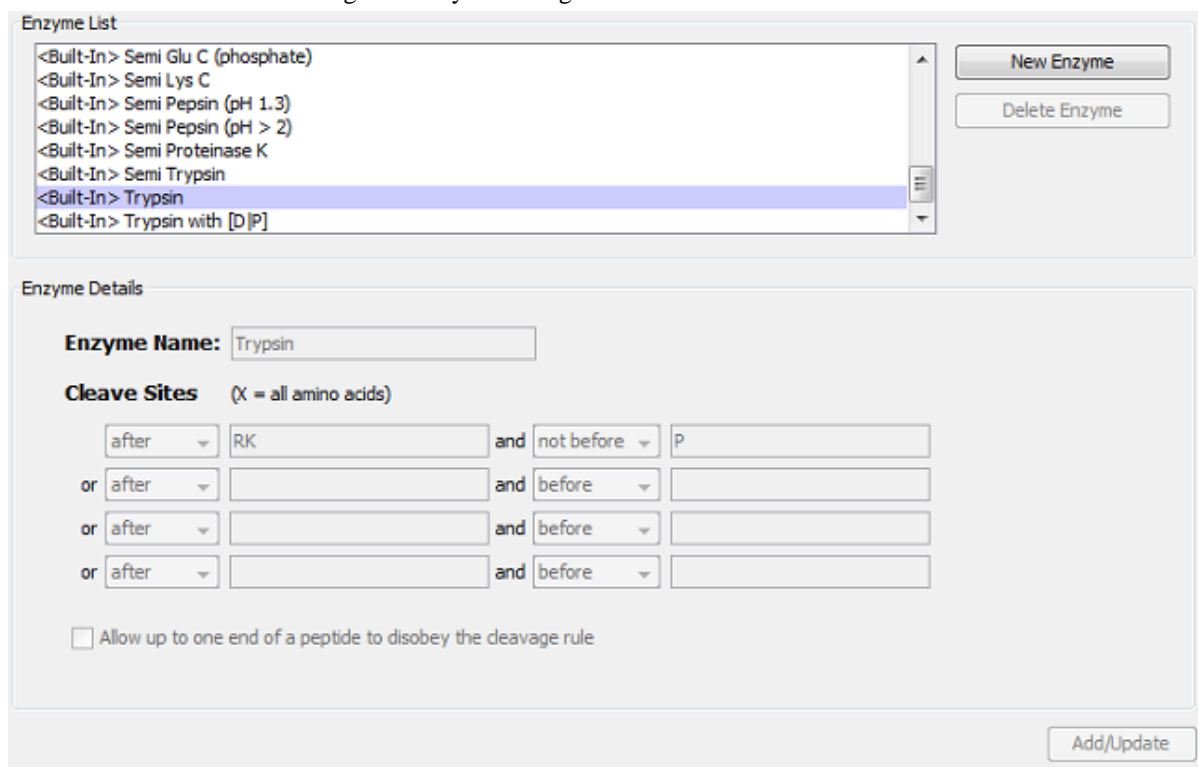
Intensity. You can set the intensity threshold for spectrum annotation to low (2%), medium (5%), or high (10%). To apply this intensity threshold, you have to select the “intensity threshold” checkbox in the “Spectrum Annotation” view.

2. PEAKS Configuration

This step includes the configuration of enzymes, PTMs, databases, instruments, and parameters. To begin click the Configuration toolbar icon  or select “Configuration” from the “Window” menu.

2.1. Enzyme Configuration

PEAKS can use almost any enzyme or combination of enzymes in your analysis. Select built-in enzymes from the extensive list provided in PEAKS or define a new one. From the “Configuration” window select “Enzyme” from the left-hand side menu to change the enzyme configuration.



Built-in enzymes. All of the built-in enzymes within PEAKS are listed in the “Enzyme List”. Clicking on one of these built-in enzymes will display the information about that enzyme in the “Enzyme Details” panel.

Note

A built-in enzyme cannot be deleted or edited and therefore the Delete Enzyme button will be disabled.

Create a new enzyme. Click on the “New Enzyme” button to create a new enzyme. Specify how the custom enzyme will cleave the protein between two amino acids to create peptides in the “Enzyme Details” panel. The letter X denotes *any amino acid in this position*, while {set brackets} indicate any amino acid except the one in the brackets. Choose where the cleave sites are by selecting *after* or *not after* and *before* or *not before* to specify the range. There is also an option to “Allow up to one end of a peptide to disobey the cleavage rule”.

Click the “Add/Update” button to save the changes. The new enzyme will now appear in the “Enzyme List” where it can be accessed later. To delete an enzyme that was created, select the appropriate enzyme and click the “Delete Enzyme” button.

Note

For information on defining new enzymes “on the fly” for PEAKS *de novo* or PEAKS DB, refer to sections Section 2.2, “Enzyme Specificity”.

2.2. PTM Configuration

From the “Configuration” window select “PTM” from the left-hand side menu to change the PTM configuration.

The screenshot shows the PTM Configuration window. The top panel, titled "PTM List", contains a scrollable list of PTMs. The first item, "<Built-In> 4-hydroxynonenal (HNE)", is selected and highlighted in blue. To the right of the list are two buttons: "New PTM" and "Delete PTM". Below the list is a checkbox labeled "Show unimod". The bottom panel, titled "PTM Details", contains several input fields: "PTM name:" with the value "4-hydroxynonenal (HNE)", "Mass (Monoisotopic):" with the value "156.11504", "Neutral loss mass (Monoisotopic):" with the value "0.0", "Residues that can be modified:" with a dropdown menu showing "CHK" and radio buttons for "Anywhere", "Middle Only", "N-term", and "C-term", "Formula:" with the value "H(16) C(9) O(2)", and "Rule:" which is empty. An "Add/Update" button is located at the bottom right of the window.

Built-in PTMs. The built-in PTMs within PEAKS are listed in the “PTM List”. To see additional built-in PTMs from the *Unimod* library, click the “Show Unimod” box. Clicking on one of these built-in PTMs will display the information listed about that PTMs in the “PTM Details” panel.

Note

A built-in PTM can not be deleted or edited and therefore the Delete PTM button and the PTM Details panel will be grayed out.

Create a new PTM. Click on the “New PTM” button. Now simply enter the information about your PTM in the “PTM Details” panel.

- **PTM Name:** this name will appear in the PTM list for future use after it is saved.
- **Monoisotopic mass:** the mass that the residue gains or losses as a result of the PTM.
- **Neutral loss mass:** the mass that the modified residue losses as a result of fragmentation in Daltons. For example, 28 would signify a loss of 28 Daltons.
- **Residues that can be modified:** enter residues that can be modified anywhere, residues that can only be modified if they are at the N-or C-terminus or in the middle only.
- **Chemical formula:** the chemical formula of the PTM. This should correspond to the mass listed above.
- **Rule:** you can enter a comment for your reference.

Click the “Add/Update” button to save the changes. The new PTM will now appear in the “PTM List” where it can be accessed later. To delete a PTM that was created, select the appropriate PTM and click the “Delete PTM” button.

Note

For information on defining new PTMs “on the fly” for PEAKS *de novo* or PEAKS DB, refer to section Section 2.3, “Fixed and Variable PTMs”.

2.3. Database Configuration

To use the PEAKS DB function to search through a database to identify proteins, PEAKS must have access to a protein or EST database in FASTA format or an EST database of DNA sequences. Point PEAKS to an existing database on the system, or download one. Additionally, it is possible to specify taxonomy with certain databases.

From the “Configuration” window select “Database” from the left-hand side menu to change the database configuration. Refer to Chapter 6, *Adding a Sequence Database* for details on configuring a new database

Delete a previously saved database. To delete a database file, select the database to be deleted from the “Database List” and click on the “Delete Database” button to the right of the “Database List”.

Setting the default database. To set a database as the default, select the database from the “Database List” and click the “Set As Default” button. This database will now be used by PEAKS when PEAKS DB is run.

Moving/Updating a database. To move a database to another directory, the location must be updated in PEAKS. Select the database, and then specify the new location using the “Browse” button next to the “Path” field. Then click “Add/Update” to save the new settings. If the database location is invalid, the database name will appear in *red* in the list of databases and any protein identification using that database will fail. If an update is made to the database file (perhaps by downloading the latest database file and overwriting the old database file), PEAKS will show the database information in *light gray*. A light grey color could also mean that the database does not have header information.

Configure databases for use with other search engines in PEAKS inChorus function. The database configured here will also be used in PEAKS inChorus function with other search engines. Refer to Section 2, “Databases to be Used in PEAKS' inChorus Function” for details on configuring databases for use with other search engines.

2.4. Instrument Configuration

From the “Configuration” window select “Instrument” from the left-hand side menu to change the instrument configuration.

The screenshot shows a software interface with two main panels. The top panel, titled "Instrument List", contains a "Manufacturer:" dropdown menu set to "_General", a list of instrument names (FT-trap, FT-trap (ecd-cid), FT-trap (etd), FT-trap (pqd), FTMS, FTMS (ecd), FTMS (ecd-cid)), and buttons for "New Instrument" and "Delete Instrument". The bottom panel, titled "Instrument Details", shows the "Instrument Name:" field set to "FT-trap". It has two sub-sections: "Basic Options" with dropdowns for "Manufacturer:" (_General), "Ion Source:" (ESI(nano-spray)), "MS Precursor Scan:" (FT-ICR/Orbitrap), "Fragmentation Type:" (CID, CAD, IRMPD (y and b ions)), and "MSn Product Scan:" (Linear Ion Trap); and "Advanced Options" with radio buttons for "Precursor mass search type:" (Monoisotopic selected, Average unselected), and input fields for "Parent mass error tolerance:" (0.01 Da) and "Fragment mass error tolerance:" (0.5 Da). An "Add/Update" button is at the bottom right.

Built-in instruments. Select the manufacturer of the instrument from the “Manufacturer” drop-down list. The names of the instruments will then appear in their vendor specific formats. Select an instrument to view the detail instrument information in the “Instrument details” panel below. Select “General” in the manufacturer list and the instruments will be listed in a general format.

Note

The details of a built-in instrument cannot be deleted or edited and therefore the Delete PTM button and the Instrument Details panel will be grayed out.

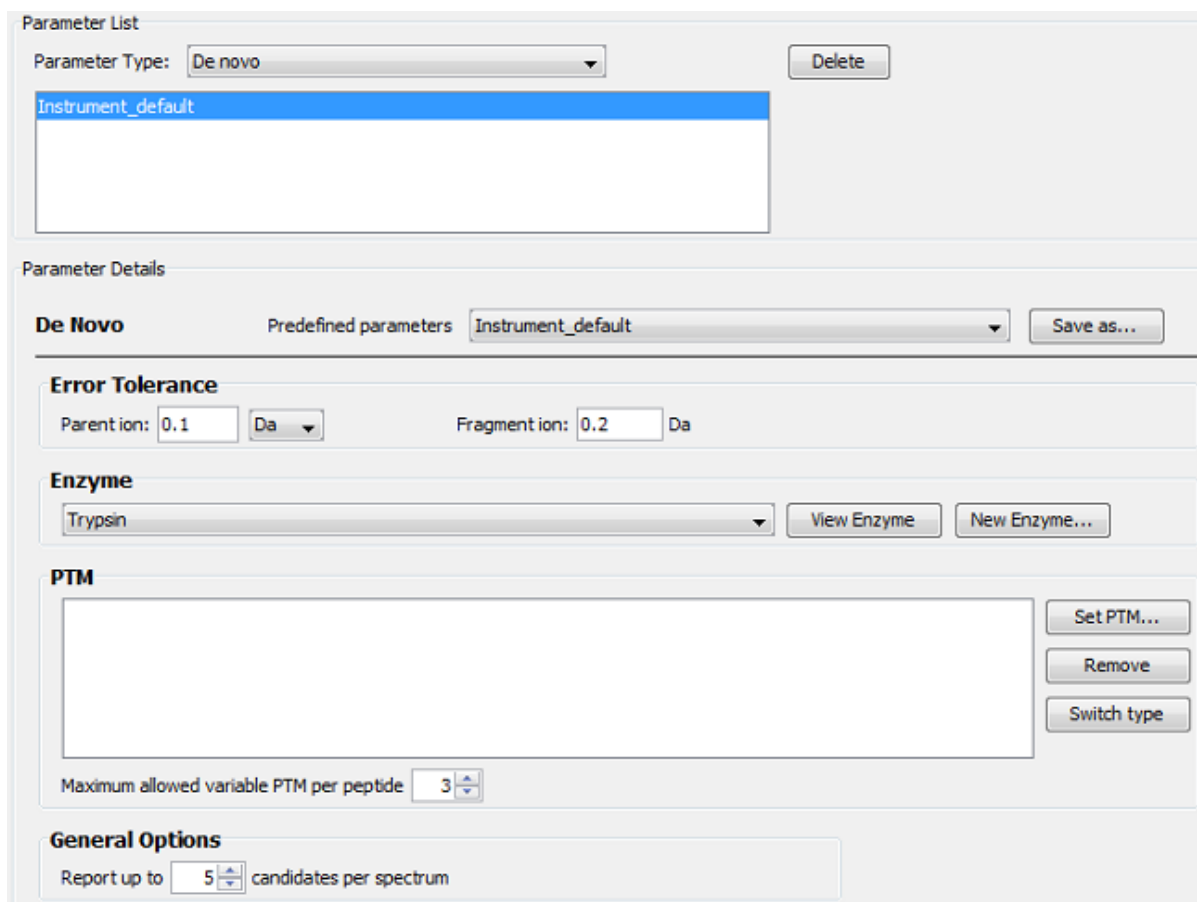
Create a new instrument.

1. Click on the “New Instrument” button.
2. In the “Instrument Details” panel, provide a name for the instrument.
3. Next, fill in the details in the “Basic Options” panel. In the “Manufacturer” drop-down list, select a specific vendor or “General”.
4. Use the “Ion Source” drop-down list to select the ion source that was used: MALDI/SELDI or ESI (nano-spray). This will help the PEAKS Data Refine tool to decide the charge of the ions.

5. Use the “MS Precursor Scan” drop-down list to select the type of MS scan that was performed. This selection will tell the PEAKS Data Refine tool whether the survey scan is of sufficient resolution to determine the charge and the monoisotopic peak from the examination of the survey scan.
6. Use the “Fragmentation Type” drop-down list to select the method of fragmentation used. This selection will tell PEAKS the type of ion-series to expect for PEAKS auto *de novo* sequencing and PEAKS DB search. Select CID/ECD if alternating fragmentation was used to allow the algorithm to determine the type of fragmentation from each scan header.
7. Use the “MSn Product Scan” drop-down list to select the type of MSn scan that was performed. This selection will help PEAKS decide which internal parameters (for weighing fragments and amount of noise) to use during PEAKS auto *de novo* sequencing and PEAKS DB search. Select LIT/FT if alternating high-res/low-res modes were used. This will allow the algorithm to determine the mass analyzer from the scan header.
8. Use the “Advanced Options” to specify additional parameters.
9. Select “Monoisotopic” or “Average” as “Precursor Mass Search Type”. For ion-trap instruments, it is usually beneficial to allow the PEAKS DB search to use an average mass.
10. Specify the values for “Parent mass error tolerance” and “Fragment mass error tolerance” in Daltons or ppm. These will appear on the PEAKS *de novo* and PEAKS DB options screens when the instrument is selected.
11. Click the “Add/Update” button to save the changes. The new instrument will appear in the “Instrument List” where it can be accessed when creating a new project file. To delete an instrument that was created, select the appropriate instrument from the “Instrument List” and click the “Delete Instrument” button.

2.5. Parameter Configuration

From the “Configuration” window select “Parameters” from the left-hand side menu to change the parameter configurations. Please note that parameters can only be viewed and deleted from within this parameter window. From the “Parameter Type” drop-down list at the top of the window, select *de novo*, PEAKS parameters, SPIDER parameters or other parameter categories. The parameters that have been saved within the selected category will be displayed in the list below. Select the desired parameter set from the list to view the parameter details.



The screenshot displays two windows from a software interface. The top window, titled "Parameter List", has a "Parameter Type:" dropdown set to "De novo" and a "Delete" button. Below is a list box containing "Instrument_default", which is highlighted in blue. The bottom window, titled "Parameter Details", shows the configuration for the selected parameter set. It includes a "De Novo" section with "Predefined parameters" set to "Instrument_default" and a "Save as..." button. The "Error Tolerance" section has "Parent ion:" set to "0.1" Da and "Fragment ion:" set to "0.2" Da. The "Enzyme" section has a dropdown set to "Trypsin" and buttons for "View Enzyme" and "New Enzyme...". The "PTM" section features a large empty list box, buttons for "Set PTM...", "Remove", and "Switch type", and a "Maximum allowed variable PTM per peptide" spinner set to "3". The "General Options" section has a "Report up to" spinner set to "5" and the text "candidates per spectrum".

Creating a new parameter set. Create and save new parameters during or before setting up auto *de novo* sequencing (see Section 2.5, “Saving the Parameters for Future Use”), PEAKS DB (see Section 2, “Set PEAKS DB Parameters”) or SPIDER (see Section 1.2, “Run SPIDER Independently”). These references will provide explanations of all of the parameters.

Deleting a previously saved parameter set. To delete a parameter set, select the parameter set and click the “Delete” button.

Viewing a previously saved parameter set. Selecting a parameter set will display the details of that set below. For an explanation of the parameters, please see the references listed in the “Creating a new parameter set” section above.