

Agilent G2723AA Spectrum Mill Data Extractor for Generic Peak List Files

Quick Start Guide

What is the Spectrum Mill Generic Data Extractor?

The Agilent Spectrum Mill MS Proteomics Workbench is a collection of tools for high-throughput processing of MS and MS/MS spectra to provide protein and peptide identifications and relative quantitation. The Agilent G2723AA Spectrum Mill Data Extractor for Generic Peak List Files enables use of the time-saving Spectrum Mill workbench with peak list files, such as those exported from Micromass Q-ToF using the ProteinLynx package. The software handles individual *.pkl and *.dta spectral files, or appended *.pkl files that contain multiple spectra. It also processes most *.mgf files, such as those from MALDI-TOF-TOF. The Spectrum Mill Generic Data Extractor prepares the peak list files for further Spectrum Mill processing. It does the following:

- Calculates spectral features used in other Spectrum Mill programs
- Filters MS/MS spectra by quality
- Generates individual spectral files from appended *.pkl or *.mgf files

Because this extractor processes exported rather than raw files, it retains less functionality than the raw file extractors and provides less accurate quantitation. With the exception of $^{14}\text{N}/^{15}\text{N}$ and iTRAQ, quantitation is not supported for labelled peptides.

Installation

To activate the license for this software, see the *Spectrum Mill MS Proteomics Workbench Installation Guide*.



Agilent Technologies

Spectrum Mill Help - core product documents and online help

See the following to learn the core Spectrum Mill workbench.

Scientists **Quick Start Guide** Get a quick overview of the Spectrum Mill workbench.

Familiarization Guide Follow step-by-step instructions to process example Agilent ion trap, Q-TOF, and TOF data.

Application Guide Learn step-by-step details to use all functions of the software.

Online Help Consult the online help for in-depth information not given in the *Application Guide*. Display online help in one of three ways:

- Click help links on the left-hand side of the Spectrum Mill home page.
- Click the **Help** button at the top of a Spectrum Mill form to get complete instructions for that form.
- Click links on the blue dividing bars in the forms to get field-level help.

Quick Reference Card After you are familiar with the software, consult this card for an overview of the steps to process MS/MS data.

System Administrators **Installation Guide** Use this guide to install the Spectrum Mill workbench, and to activate the Generic Data Extractor license.

Application Guide See the following chapters:

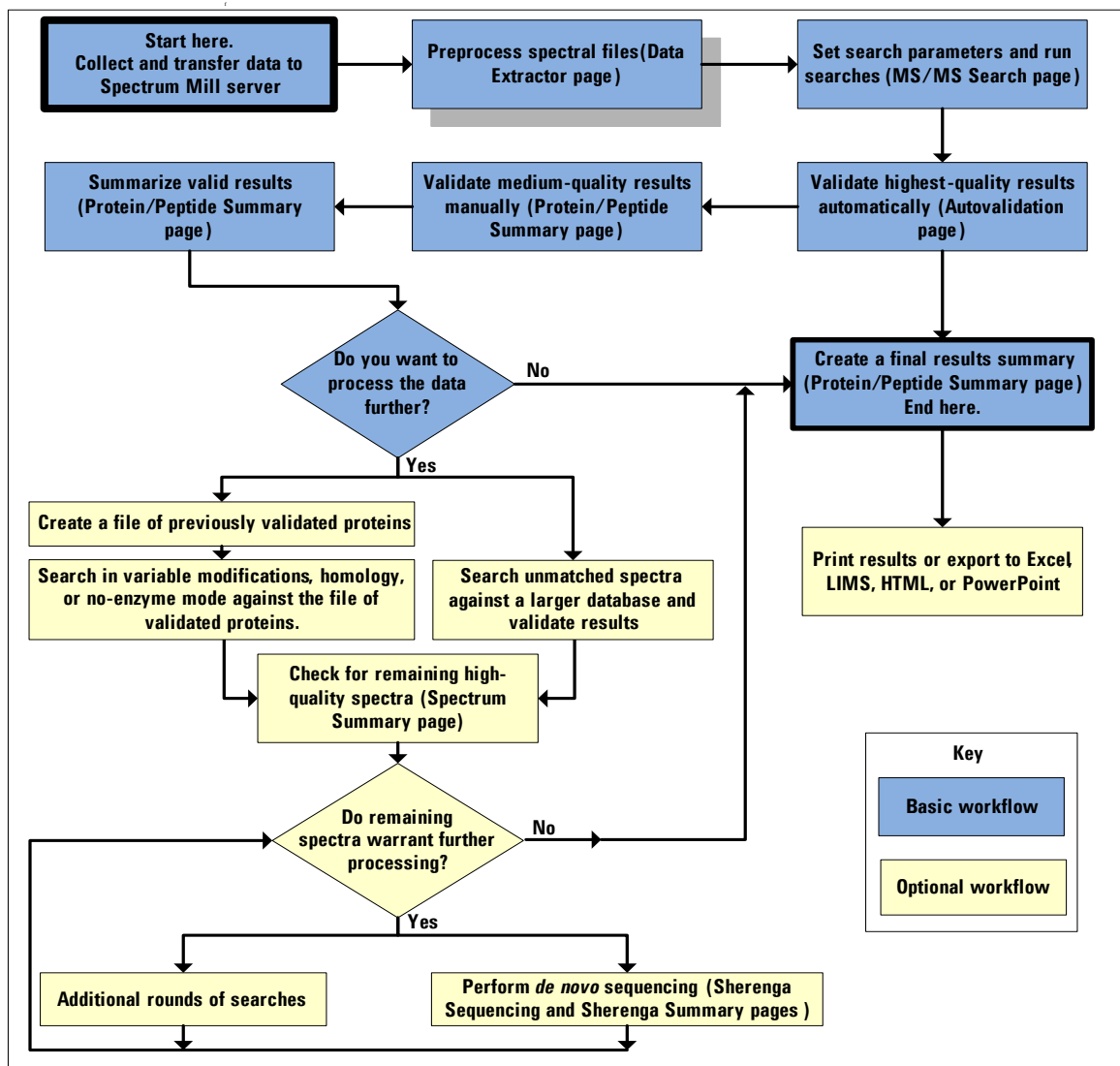
- **Chapter 8: System Administration**
Get an overview to install databases and perform other system administration tasks.
- **Chapter 9: Files Created during Spectrum Mill Data Processing**
Refer to this chapter to troubleshoot data processing, to selectively remove parts of the processing, or to decide which files to archive.

Online Help From any **Help** page, click links under **For System Administrators**:

- **Protein Databases** (link to `millhtml\SM_instruct\faman.htm`)
Learn details to install databases, create indices, and create subset databases.
- **Server Administration** (link to `millhtml\SM_instruct\servadmn.htm`)
Learn details to perform other system administration tasks.

Roadmap for MS/MS data processing

This diagram shows the overall Spectrum Mill work flow for MS/MS data. The Generic Data Extractor accomplishes the part of the work flow with the shadow.



Familiarization tutorial

Exercise 1. (Optional) Transfer appended peak list file to the Spectrum Mill server

To process peak list files with the Spectrum Mill workbench, you first move or copy the files into the appropriate directory on the Spectrum Mill server.

To make it easy to compare data sets, it is important that you set up the appropriate directory structure for your files on the Spectrum Mill server. Whenever you want to compare samples in a set, you need to set up a subdirectory for each sample. This directory may contain data files from multiple sample fractions or gel slices.

NOTE

The generic example file may have been installed on the server at the time the Spectrum Mill software was installed. If so, you may either skip this exercise or (if you want the practice) delete the example data from the server and install it again with this exercise.



Steps	Detailed Instructions	Comments
1 Copy the example data file, X:\ExampleData\msdataSM\ExampleData\Generic\Generic_Mix.pkl , from the <i>Spectrum Mill Example Data CD</i> onto your client PC.	<ul style="list-style-type: none">Copy Generic_Mix.pkl from the <i>Spectrum Mill Example Data CD</i> to any directory on your client PC.	<ul style="list-style-type: none">You do this exercise outside the Spectrum Mill workbench. Use your normal file management utilities.This exercise simulates the likely laboratory scenario of file transfer from a client or instrument PC to the server; you may transfer files directly from the CD to the server if you prefer.


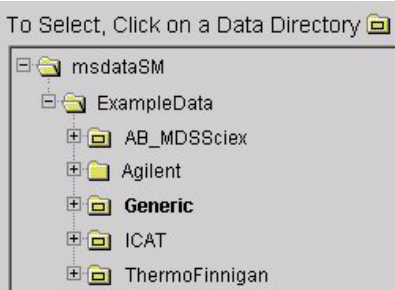
Steps	Detailed Instructions	Comments
<p>2 Create the msdataSM\ExampleData\Generic folder on the Spectrum Mill server.</p>	<p>a On your server, find the SpectrumMill folder.</p> <p>b In this folder, click to open the folder msdataSM.</p> <p>c In msdataSM, create ExampleData\Generic. You need only one folder because this exercise uses a single sample.</p>	<ul style="list-style-type: none"> • If you don't know how to find your Spectrum Mill file system, ask the person who installed your software. • Do not include spaces or parentheses in your directory name. • When you process your own samples, remember to set up a separate folder for each sample. Each folder should contain all sample fractions. • You may create up to ten folders between msdataSM and your data files, but shorter path lengths reduce memory usage, especially for large data sets.
<p>3 Copy the generic example file to the new directory.</p>	<ul style="list-style-type: none"> • Copy or move Generic_Mix.pkl from your client PC to the new folder on the server PC. 	<ul style="list-style-type: none"> • When you process your own data: <ul style="list-style-type: none"> • If you have individual *.pkl or *.dta files, create a subdirectory called cpick_in under your new data directory, and move your files to cpick_in. • If you have a *.mgf file with multiple spectra (for example, from a MALDI-TOF-TOF analysis), copy it to the root data directory, the same as in this exercise. • Remove any spaces or parentheses in the data file names. For example, change the file my sample.d to my_sample.d.


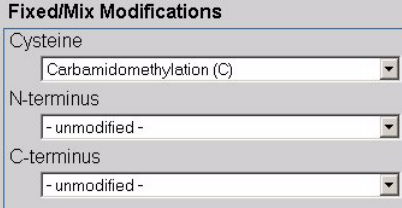
Steps	Detailed Instructions	Comments
<p>4 Make sure you have both read and write permissions for the data folder you just created on the server.</p>	<ul style="list-style-type: none"> a Right-click the Generic folder and select Properties. b Clear the Read-only check box if it is marked. c In the Confirm Attribute Changes dialog, click Apply changes to this folder, subfolders, and files. d Click OK. e If necessary, repeat step a. f Click the Security tab. g Make sure all user groups have full permissions. h Click OK. 	<p>If the Spectrum Mill workbench cannot write to the folders that contain your data files, you may encounter errors.</p>


Exercise 2. Run the Generic Data Extractor

The Spectrum Mill Data Extractor preprocesses data files to extract high-quality spectra for database searches. The Data Extractor automatically detects which type of file you have and then invokes the appropriate extraction program. For peak list data, it invokes the Generic Data Extractor. This program calculates spectral features and filters MS/MS spectra by quality. When you apply this extractor to appended peak list files, it also extracts the individual spectra.

Steps	Detailed Instructions		Comments
1 Start the Spectrum Mill workbench on your PC.	<ul style="list-style-type: none">• Double-click the desktop icon to launch the Spectrum Mill workbench.		You will see the Spectrum Mill home page.
2 If you don't see the Spectrum Mill icon, launch your web browser and type the URL for the Spectrum Mill home page.	<ul style="list-style-type: none">a Click your Internet Explorer icon.b In the web browser window, type the URL for the Spectrum Mill server.		If you don't know the URL, ask the person who installed the software.
3 Check that you see the Spectrum Mill home page.			
4 If you don't, check that the server is booted up and that your URL is correct.			
5 Navigate to the Data Extractor page.	<ul style="list-style-type: none">• From the Spectrum Mill home page, click the Data Extractor link.		

Steps	Detailed Instructions	Comments
<p>6 Select the msdataSM\ExampleData\Generic data directory.</p>	<p>a Click the Select... button near the middle of the form.</p>  <p>b Expand the directory tree and click the word Generic to select that data file directory.</p>  <p>c Make sure that the name of the Generic directory changes to a bold font. This indicates that it has been selected.</p> <p>d Click OK. The data directory appears as in Figure 1 on page 10.</p>	<ul style="list-style-type: none"> • Directories are identified by different types of icons: • Plain folders indicate directories that do not have data files directly beneath them. • Folders with rectangles indicate data directories. • Folders with line spectra (bar graphs) indicate data files. • Try clicking the names of each type to see which turn bold, indicating that they are selectable. • The software remembers your data file selection when you go to other Spectrum Mill forms. • If you mark the Make Default check box in this dialog, the software remembers your data directory even after you close your web browser.
<p>7 Make sure Instrument is set to the default of ESI QTOF Generic.</p>		<ul style="list-style-type: none"> • When you process your own data, if you have a *.mgf file, select the appropriate instrument (for example, MALDI-TOF-TOF). If you do not see your instrument on the list, alert your system administrator to set up an instrument per instructions in the online help for system administrators.

Steps	Detailed Instructions	Comments
<p>8 Ensure that you will generate individual extracted files.</p>	<ul style="list-style-type: none"> Make sure the check box for Generate spectral features file only is <i>not</i> marked. (Use the default setting.) 	<ul style="list-style-type: none"> When you have *.dta files, or *.pkl files that represent individual spectra, you put your files in the cpick_in folder, and then you <i>must</i> mark this check box. For this exercise, you use an appended *.pkl file that contains multiple spectra, so you do not mark the check box. If you have a *.mgf file with multiple spectra, you also do not mark the check box.
<p>9 Choose the appropriate cysteine modification.</p>	<p>a Click the Choose... button near the middle of the form.</p>  <p>b Under the Cysteine heading, select carbamidomethylation.</p>  <p>c Click OK. The name of the modification appears as in Figure 1 on page 10.</p>	<ul style="list-style-type: none"> Carbamidomethylation is the default when you first install the Spectrum Mill workbench, so you may be able to skip this step. To view details about the modifications that are currently available on your server, click the Details button at the bottom right of the Choose Modifications dialog. For more information about choosing modifications, see the online help. Your system administrator can configure custom modifications.
<p>10 Set other parameters as shown in Figure 1 on page 10.</p>	<p>a Keep the defaults whenever appropriate. This example uses all default settings, except for Data Directory and Instrument.</p> <p>b Examine the items in red text carefully, since these are the ones you may need to change when you process your own samples.</p> <p>c Click a blue section divider bar to display help for that section of the form.</p>	<ul style="list-style-type: none"> You use the Sequence tag length to filter out noisy spectra. This is the longest sequence of amino acids that is represented by the fragments in a spectrum. If you set this to >1, you ensure that all possible good spectra are extracted. You have opportunities to set more stringent requirements later when you perform the database search.

Steps	Detailed Instructions	Comments
11 Start the extraction.	<p>a Click the Extract button. </p> <p>b View extraction progress in the Results area to the right of the Data Extractor form.</p> <p>c Scroll to the <i>top</i> of the Results area to see the message that indicates that extraction is finished.</p>	<ul style="list-style-type: none"> • The Data Extractor processes all files in the directory. • Extraction time varies depending on the number and size of the files. • You can use your client PC for other tasks while the extraction is taking place. • If you want to stop the extraction, click the red Stop Extraction PID: xxx link at the top of the Results section. Then see the Tool Belt chapter in the <i>Spectrum Mill MS Proteomics Workbench Application Guide</i> for further instructions.
12 (Optional) View the cpick_in subdirectory to see the files Data Extractor has created.	<p>a Navigate to the folder SpectrumMill\msdataSM\ExampleData\Generic\cpick_in.</p> <p>b Notice the new files created there. Each one represents an extracted spectrum from your appended peak list file.</p>	<p>File names are in the format: Data_File_Name.aaaa.bbbb.c.pkl, where</p> <ul style="list-style-type: none"> • aaaa = first merged scan • bbbb = last merged scan • c = assigned precursor charge (0 means charge was ambiguous)

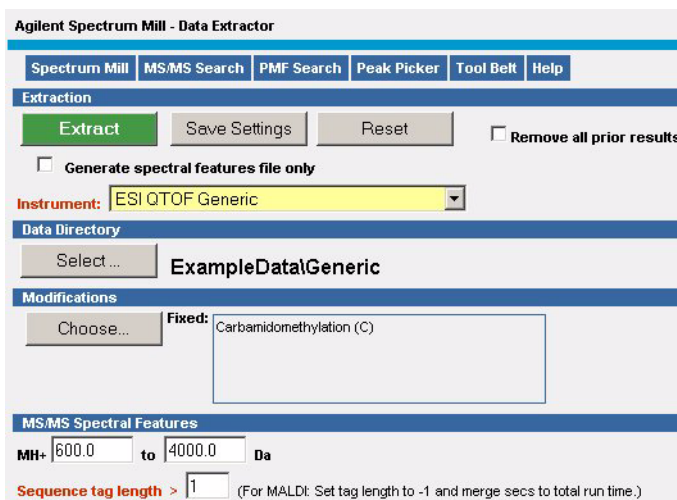
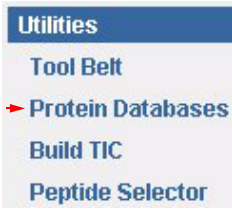


Figure 1 Settings for Data Extractor

Exercise 3. Create indices for user-created database

The Spectrum Mill workbench allows you to create and search user databases. In this exercise, you create indices for a user database so you can later search it.

Steps	Detailed Instructions	Comments
1	<p>If necessary, copy the user database NCBInr.stdmix to your Spectrum Mill server.</p> <p>a If the NCBInr.stdmix database is already installed in your Spectrum Mill database directory, skip to step 2.</p> <p>b If not, find the file X:\Example Databases\NCBInr.stdmix on your <i>Spectrum Mill Installation CD</i>.</p> <p>c Copy it to your database directory on your Spectrum Mill server.</p>	<ul style="list-style-type: none"> The database directory likely has a name like X:\seqDB. If you are not sure, ask your system administrator.
2	<p>Make sure security permissions are set properly for the folder that contains the file dbname.js.</p> <p>a Right-click the SpectrumMill\millhtml\SM_js folder and select Properties.</p> <p>b Click the Security tab.</p> <p>c Make sure all user groups have full permissions.</p> <p>d Click OK.</p>	<p>If the Spectrum Mill workbench cannot write to this folder, you may encounter errors.</p>
3	<p>Navigate to the Protein Databases page.</p> <p>• From the Spectrum Mill home page, click the Protein Databases link.</p>  <p>The screenshot shows a vertical menu with the following items: Utilities (highlighted in blue), Tool Belt, Protein Databases (with a red arrow pointing to it), Build TIC, and Peptide Selector.</p>	
4	<p>Fill in the form as shown in Figure 2.</p> <p>a On the Protein Databases page, click the Create indices for new database option. (This is the default.)</p> <p>b Set Newly downloaded database to NCBInr.stdmix.</p>	<p>For Newly downloaded database, you type the exact database file name.</p>
5	<p>Click the Create Indices button.</p> <p>• After a short wait, check the bottom of the page for text that indicates that the software is creating the database indices.</p>	<p>These indices bear no resemblance to those used by another popular database search program.</p>

Agilent Spectrum Mill - Protein Database Utilities

Spectrum Mill | **MS/MS Search** | **PMF Search** | **Tool Belt** | **Help**

Create indices for new database
 Create species subset database
 Create subset with indices from saved hits
 Create or append user database
 Database summary report

(After creating a database, click the "Update Database List" button to see the database listed)

Create Indices



Newly downloaded database:


Existing databases:

Figure 2 Create indices for user database

Exercise 4. Run database searches

After you have extracted your spectra and created indices for your user database, you are ready to search each spectrum against a protein or DNA database. As you process data with the Spectrum Mill workbench, you may iterate through multiple rounds of database search and results validation, with the goal of identifying as many spectra as possible. These exercises illustrate only a single identity mode search. In identity mode, the spectra must be consistent with the database sequence. When you process your own data, you might also search in variable modifications or homology mode, where spectra may show modifications relative to the database sequence.

Steps	Detailed Instructions	Comments
<p>1 Navigate to the MS/MS Search page.</p>	<ul style="list-style-type: none"> Do one of the following: <ul style="list-style-type: none"> From the Protein Databases page, click the MS/MS Search button.  <ul style="list-style-type: none"> From the Spectrum Mill home page, click the link to MS/MS Search. 	<p>Since Spectrum Mill pages have buttons to take you to the next processing tasks, you can navigate directly from the Protein Databases page to the MS/MS Search page.</p>
<p>2 Check that your Data Directory is set to Generic.</p>	<ul style="list-style-type: none"> If you have just performed the data extraction, your Generic data directory should already be set correctly. If not, click the Select... button to select the Generic folder. 	
<p>3 Select the “standard mix” database.</p>	<ul style="list-style-type: none"> For Database (under Search Parameters) select the NCBIInr.stdmix database for which you just created indices. 	<p>If the database name does not appear, reload or refresh the MS/MS search form.</p>

Steps	Detailed Instructions	Comments
<p>4 Choose the appropriate cysteine modification.</p>	<p>a Click the Choose... button near the middle of the form.</p>  <p>b Under the Cysteine heading, select carbamidomethylation</p> <p>c Click OK. The name of the modification appears as in Figure 3 on page 17.</p>	<ul style="list-style-type: none"> • The modification is likely already set. • Three types of modifications are available for MS/MS Search: <ul style="list-style-type: none"> • The fixed modifications are assumed to apply universally and are searched in a single search cycle. • The mix modifications trigger cyclic MS/MS searches, where a different form of the modification is searched in each cycle. • The variable modifications allow for both modified and unmodified forms within a peptide. All variable modifications are searched in each cycle. • For more information about choosing modifications, see the online help. • Your system administrator can configure custom modifications.
<p>5 Make sure Instrument is set to the default of to ESI QTOF Generic.</p>		<ul style="list-style-type: none"> • When you process your own data, select the appropriate instrument. • If your spectra contain many noise peaks (for example, noisy spectra from a MALDI analysis), reduce the value for Minimum matched peak intensity.

Steps	Detailed Instructions	Comments
6 Set up the Search Mode .	<p>a Make sure that the check box for Calculate reversed database scores is marked. (This is the default setting.).</p> <p>b Mark the check box for Dynamic peak thresholding.</p> <p>c Verify that the Search mode is set to Identity.</p>	<ul style="list-style-type: none"> • When you Calculate reversed database scores, you search against peptide sequences in their forward and inverted directions. If you obtain similar scores for both searches, you likely have a false positive. • Dynamic peak thresholding is a scoring enhancement that enables identification of more low-abundance and short-chain peptides. For each extracted spectrum, the software calculates the search scores as the number of spectral peaks varies from n=4 up to the maximum set by the variable peakLimitCount in instrument.txt. It then displays the best score from the set.

Steps	Detailed Instructions	Comments
<p>7 Make sure other parameters are set as shown in Figure 3 on page 17.</p>	<p>a Keep the defaults whenever appropriate.</p> <p>b Examine the items in red text carefully, since these are the ones you may need to change when you process your own samples.</p> <p>c Click a blue section divider bar to display help for that section of the form.</p>	
<p>8 Start the search.</p>	<p>a Click the Start Search button</p> <p>b View search progress in the Results area to the right of the MS/MS Search form.</p> <p>c Scroll to the <i>top</i> of the Results area to see the message that indicates that the search is finished.</p>	<ul style="list-style-type: none"> • MS/MS Search processes all spectral files in the directory. • Search time varies depending on the size of the database. This search goes fairly quickly because you search a small user database. • You can use your client PC for other tasks while the search is taking place. • If you want to stop the search, click the red Stop Search PID: xxx link at the top of the Results section. Then see the Tool Belt chapter in the <i>Spectrum Mill MS Proteomics Workbench Application Guide</i> for further instructions.

Agilent Spectrum Mill - MS/MS Search

Spectrum Mill Easy MS/MS Autovalidation Protein/Peptide Summary Extractor Databases Tool Belt Help

Search

Start Search Save Settings Reset Remove all prior MS/MS Search results

Data Directory

Select... ExampleData\Generic

Search Parameters

Validation filter: spectrum-not-marked-sequence-not-validated Batch size: 81

Search previous hits Max reported hits: 5

Database: NCBI\ncr.stdmix Digest: Trypsin

Species: All Maximum # missed cleavages: 2

Protein pI: from 3.0 to 10.0 All Required AAs: Disallowed AAs:

Modifications

Choose... Fixed: Carbamidomethylation (C) Variable:

Search Criteria

Matching Tolerances

Minimum scored peak intensity: 50 %

Instrument: ESI QTOF Generic

Masses are: Monoisotopic no e- correction

Precursor mass tolerance: +/- 100 ppm

Product mass tolerance: +/- 500

Maximum ambiguous precursor charge: 3

Spectral Quality

Sequence tag length: > 3

Minimum detected peaks: 4

Search Mode

Calculate reversed database scores

Proton mobility scoring

Dynamic peak thresholding

Search mode: Identity

Data Files

Spectrum files (./cpick_in/):

*.pk1
*.dta

Figure 3 MS/MS Search settings

Exercise 5. Run Autovalidation

After you have completed a database search, you validate the good results. Validation means that you accept that the matches are correct.

The Spectrum Mill workbench provides a means for segregating search results that contain a valid interpretation of an MS/MS spectrum from those that do not. Results that are *not* validated can then be subjected to subsequent rounds of searches (against other databases or in variable modifications mode, for example). Results that *are* validated can be summarized in a results table.

The Spectrum Mill workbench provides two ways of validating results. One way uses the Autovalidation page, and is totally automated. You use this method only to validate the highest-scoring results—those that do not require manual review. The other method of validating uses the Protein/Peptide Summary page for manual review and validation. This exercise describes autovalidation. To learn more about manual validation, see the *Spectrum Mill MS Proteomics Workbench Application Guide*.

Steps	Detailed Instructions	Comments
1 Navigate to the Autovalidation page.	<ul style="list-style-type: none">• Navigate to this page from one of two other pages:<ul style="list-style-type: none">• MS/MS Search• Protein/Peptide Summary	You will see the form shown in Figure 4 on page 20.
2 Check that your Data Directory is set to Generic .	<ul style="list-style-type: none">• If you have just performed data extraction and MS/MS search, your data directory should already be set correctly. If not, click the Select... button to select the Generic folder.	
3 Validate first in the Protein details mode.	<ul style="list-style-type: none">a For Mode, keep the default of Protein details.b Keep the default scoring presets.c Click the Validate Files button.d Watch for a Spectrum Mill Validation Summary that lists the hits and spectra that have been validated.	<ul style="list-style-type: none">• In this mode, the software summarizes results by protein, and considers all the peptides that belong to a given protein.• Using the default scoring, individual peptides must have scores greater than 6 to 12 (depending on charge state), and the cumulative protein score must be greater than 20.

Steps	Detailed Instructions	Comments
<p>4 Validate second in the Peptide mode.</p>	<p>a For Mode, select Peptide. b Keep the default scoring presets. c Click the Validate Files button. d Watch for a Spectrum Mill Validation Summary that lists the hits and spectra that have been validated.</p>	<ul style="list-style-type: none"> • In this mode, the software summarizes results by peptide. Even if it finds only a single peptide corresponding to a protein, it validates the corresponding search results provided that the peptide score is high enough. • Using the default scoring, individual peptides must have scores greater than 11 to 15 (depending on charge state). This score threshold is higher than in the Protein details mode, where you have the additional assurance of knowing you have identified more than one peptide per protein.
<p>5 Close the Autovalidation page.</p>		

Agilent Spectrum Mill - MS/MS Autovalidation

[Help](#)

Automatic Validation

Validate Files

Mode: Filter proteins by score: Group proteins across all directories

Filter by peptide pI: Low High

Data Directories

ExampleData\Generic

Search result files:

Protein Rules

Rule	Precuror Charge	Score Threshold	% SPI Threshold	Prod - Rev Score Threshold	Rank 1-2 score Threshold
1.	<input type="text" value="2"/>	<input type="text" value="6.0"/>	<input type="text" value="60.0"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
2.	<input type="text" value="1"/>	<input type="text" value="6.0"/>	<input type="text" value="70.0"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	<input type="text" value="3"/>	<input type="text" value="8.0"/>	<input type="text" value="70.0"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.	<input type="text" value="4"/>	<input type="text" value="8.0"/>	<input type="text" value="70.0"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.	<input type="text" value="5"/>	<input type="text" value="12.0"/>	<input type="text" value="70.0"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.	<input type="text" value="2"/>	<input type="text" value="6.0"/>	<input type="text" value="90.0"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 4 Autovalidation form

Exercise 6. Display valid database search results

In this exercise, you summarize the results that you just automatically validated.

Steps	Detailed Instructions	Comments
1 Navigate to the Protein/Peptide Summary page.	<ul style="list-style-type: none"> If the Protein/Peptide Summary page is already open, do nothing. If not, do one of the following: <ul style="list-style-type: none"> From the MS/MS Search page, click the Protein/Peptide Summary button. From the Spectrum Mill home page, click the link to Protein/Peptide Summary. 	There are links to this page from many other pages.
2 Check that your Data Directory is set to Generic .	<ul style="list-style-type: none"> If you have just performed autovalidation, your data directory should already be set correctly. If not, click the Select... button to select the Generic folder. 	
3 Set the Mode to Protein Summary .		Notice that the summary input form changes to correspond with the new display mode.
4 Set the Filter results by to valid .		<ul style="list-style-type: none"> The Filter results by setting selects from your data only the results that match your setting. In this case, it selects only the results that match the valid setting (which was designated during autovalidation).
5 Leave Sort proteins by set to the default of Score .		<ul style="list-style-type: none"> The Sort parameters determine how the data are sorted in the results display. When you process your own data, select the setting that is most helpful to you.
6 Set the score and % SPI filters to > 0 .	<p>a Set Filter by protein score to > 0.</p> <p>b For Filter peptides by, set both Score and % SPI to > 0.</p>	When you set this to > 0 , you are sure to display all the valid results.

Steps	Detailed Instructions	Comments
7 Set the Review Fields .	<ul style="list-style-type: none"> a Mark the Protein MW check box. b Keep the other default settings. c Check to see that your form looks like that in Figure 5 on page 23. Note that changes from defaults are highlighted in yellow. 	<ul style="list-style-type: none"> • Review Fields determine what information you see in the final results summary. • The default settings are shown in the online help. Click the Review Fields blue dividing bar to access the online help.
8 Click the Summarize button.		
9 Examine the overall summary report.	<ul style="list-style-type: none"> • Check that your results are similar to those in Figure 6 on page 23. 	<p>Note the colored cells in the summary report.</p> <ul style="list-style-type: none"> • The color-code indicates relative quantities. Dark red is highest, orange is intermediate, and yellow is lowest. • The Distinct Peptides column gives the number of peptides detected for each protein. • The Mean Peptide Spectral Intensity is an average of the intensities for all peptides detected for that protein. These intensities are calculated from extracted ion chromatograms from the precursor ions. • For more accurate quantitation, under Review Fields, for Intensity, select Total rather than Mean. • These intensity results are sufficient for studies where you are interested in differences of two-fold or more.

Figure 5 Settings to summarize valid results - Protein Summary mode

Agilent Spectrum Mill - Protein/Peptide Summary									
Spectrum Mill	Summary Settings	Autovalidation	Easy MS/MS	MS/MS Search	Spectrum Summary	Build TIC	Tool Belt	Help	
Results Shown Filtered by Validation Category: valid									
Data Directory: msdataSMExampleData\Generic									
hit table read - SpecFeatures read									
valid hits read from tagSummary file - Files: 153 Hits: 153									
beginning to assemble proteins proteins assembled 0.023002 sec									
proteins filtered by unique peptides 0.016584 sec									
proteins filtered by score									
calculated protein coverage maps 0.171662 sec									
beginning to roll up proteins into groups proteins rolled up into groups 0.02476 sec									
protein groups ready for display									
proteinGroupingMethod: oneSharedPeptide 10 Proteins listed									
Group (#)	Spectra (#)	Distinct Peptides (#)	Distinct Summed MS/MS Search Score	% AA Coverage	Mean Peptide Spectral Intensity	Protein MW (Da)	Database Accession #	Protein Name	
1	72	40	623.50	66	1.01e+004	69293.9	1351907	bovine serum albumin	
2	28	25	365.03	43	5.66e+003	77050.4	4557871	transferrin	
3	19	17	240.87	28	7.37e+003	90569.6	4505881	plasminogen	
4	13	13	180.80	15	4.13e+003	116483.5	114939	beta-galactosidase	
5	7	7	111.16	17	3.34e+003	53354.2	71826	fibrinogen beta chain	
6	5	4	66.82	36	7.98e+003	16950.6	70561	Myoglobin	
7	5	4	62.84	21	1.35e+004	24529.1	115646	Alpha-S1 Casein precursor	
8	2	2	29.91	8	6.94e+003	26018.8	115654	Alpha casein S2	
9	1	1	17.36	17	1.05e+004	13821.6	2136813	Ribonuclease A	
10	1	1	12.69	13	2.99e+003	10902.8	115114	Aprotinin	
Totals:	153	114							

Figure 6 Protein summary of valid search results

www.agilent.com

In this Book

The *Quick Start Guide* presents first steps to use the Spectrum Mill Data Extractor for Generic Peak List Files.

© Agilent Technologies, Inc. 2007

Printed in USA
Fifth edition, June 2007



G2723-90005



Agilent Technologies