Scaffold Suite: Loading Search Engine Results

This document describes how to load a number of different search engine results into Scaffold. It also provides suggestions on how to set certain search parameters so that loading the resulting data into Scaffold becomes seamless.

External Document Resources

- Files Scaffold Reads
- System Requirements
- Installation Guide

Contents

Scaffold suite: Compatibility Matrix

		Label-free quantitation								
		Spectrum Counting methods		Total Ion Current TIC	Precursor Intensity	Retention Time feature	Isobaric Tags		Stable Isotope Labeling (Multiplex)	
Supported Input Data file	Scaffold Suite Application	Total Spectra / Weighted Spectra / exFDR / NSAF					iTRAQ	TMT	SILAC	Dimethyl
SEQUEST® *.OUT (*.DTA must be in same directory) *.SQT (*.MS2 must be in same directory) *.SRF	Scaffold	X		X						
	Scaffold Q+	X		X			X	X		
	Scaffold Q+S	X		X			X	X		
Mascot® (Matrix Science) = *.DAT	Scaffold	X		X						
	Scaffold Q+	X		X			X	X		
	Scaffold Q+S	X		X			X	X		
X! Tandem® = *.XML	Scaffold	X		X						
	Scaffold Q+	X		X			X	X		
	Scaffold Q+S	X		X			X	X		
Phenyx (GeneBio) = *.XML [NOTE: GeneBio discontinued this product]	Scaffold	X		X						
	Scaffold Q+	X		X			X	X		
	Scaffold Q+S	X		X			X	X		
OMSSA = *.OMX	Scaffold	X		X						
	Scaffold Q+	X		X			X	X		
	Scaffold Q+S	X		X			X	X		
Waters® IdentityE = *.XML	Scaffold	IdentityE								
	Scaffold Q+	IdentityE					X	X		
	Scaffold Q+S	IdentityE					X	X		
Agilent Spectrum Mill = Individual Results Directory	Scaffold	X		X	X	X				
	Scaffold Q+	X		X	X	X	X	X		
	Scaffold Q+S	X		X	X	X	X	X	X	
Mascot Distiller® (Matrix Science) = *.XML (the *.ROV file, which points to the *.DAT file, must be in the same directory as the *.XML file).	Scaffold	X		X	X	X				
	Scaffold Q+	X		X	X	X	X	X		
	Scaffold Q+S	X		X	X	X	X	X	X	X
Proteome Discoverer® (Thermo Scientific) = *.MSF	Scaffold	X		X	X	X				
	Scaffold Q+	X		X	X	X	X	X		
	Scaffold Q+S	X		X	X	X	X	X	X	X
MS Amanda (in PD 1.4)	Scaffold	X		X	X	X				
	Scaffold Q+	X		X	X	X	X	X		
	Scaffold Q+S	X		X	X	X	X	X	X	X
MaxQuant/Andromeda = Individual Results Directory (versions 1.4 and lower)	Scaffold	X		X	X	X				
	Scaffold Q+	X		X	X	X	X	X		
	Scaffold Q+S	X		X	X	X	X	X	X	X
mzIdentML format = *.MZID / *.MZID.GZ: Any search results exported in mzIdentML format can be loaded in Scaffold.	Scaffold	X								
	Scaffold Q+	X					X	X		
	Scaffold Q+S	X					X	X		

Notes
These files must be created from Scaffold plugin installed in Waters PLGS(v2.4 or higher)
Scaffold only supports modifications that are reported by Spectrum Mill as having single sites with a single formula
If the *.DAT file is inaccessible from the server, it should also be included in the same directory as the *.XML and *.ROV files Quantitative data generated by the Mascot Daemon using Distiller libraries not supported.
Run PD with appropriate templates for precursor intensity.
Run PD with appropriate Isobaric tags templates.
Run PD with appropriate templates for Stable Isotope labeling.
This search engine is available in PD 1.4 as Sequence Database Search node, it is supported in Scaffold 4.0 and higher. Run PD with appropriate templates for precursor intensity.
Run PD with appropriate Isobaric tags templates..
Run PD with appropriate templates for Stable Isotope labeling .
MaxQuant supported versions 1.2.2.5 and higher. Precursor intensity is supported only for 1 raw file in MaxQuant version 1.4 when selecting the Label Free Quantitation Option (LFQ)
Examples include Peaks, Byonic, MyriMatch, SQID, MS-GF+, including unknown search engines.

For more detailed information: http: //www.proteomesoftware.com/pdf/file_compatibility_matrix.pdf

Scaffold PTM and Scaffold perSPECtives: Compatibility Matrix

	Supported Input Data file	Scaffold Suite Application	Label-free quantitation: Spectrum Counting methods: Spectral counts	Label-free quantitation: Spectrum Counting methods: Weighted Spectra	Label-free quantitation: Spectrum Counting methods: emPAI	Label-free quantitation: Spectrum Counting methods: NSAF	Label-free quantitation: Total Ion Current TIC	Label-free quantitation: Precursor Intensity	Retention Time feature	Isobaric Tags: iTRAQ	Isobaric Tags: TMT	Stable Isotope Labeling (Multiplex): SILAC	Stable Isotope Labeling (Multiplex): Dimethyl	Notes
mzIdentML	Proteome Software Scaffold = *.mzid (*.mgf from Scaffold must be in same directory)	Scaffold PTM	X							x	x	x	x	Label quantitation is included only when a ScaffoldQuantML report (*.sqml file) exported from Scaffold Q+ or Q+S is loaded in Scaffold PTM
		Scaffold perSPECtives	X											
mzIdentML	Matrix Science Mascot® = *.mzid (*.mgf from Mascot must be in same directory)	Scaffold PTM	X											
		Scaffold perSPECtives	X											
mzIdentML	Bioinformatics Solutions PEAKS = peptides_1_0_0.mzid (*.mgf from PEAKS must be in same directory)	Scaffold PTM	X											
		Scaffold perSPECtives	x											
mzIdentML	Any application that produces a valid version 1.0 or 1.1 mzIdentML = *.mzid (with some exceptions – see notes)	Scaffold perSPECtives	x											Examples of incompatible tools are SpectraST (through pepXML file conversion), MyriMatch, and Pepitome. For more info see Scaffold perSPECtives User's manual
For more detailed information: http://www.proteomesoftware.com/pdf/file_compatibility_matrix.pdf														

Scaffold Suite: Search Engines Supported Versions

Search Engine	File extensions	Supported Version	Notes
SEQUEST®	*.out (*.dta must be in same directory) *.sqt (*.ms2 must be in same directory) *.srf	Latest version supported by Thermo and SageN Research	
COMET®	*.sqt (*.ms2 must be in same directory)	2014.01 rev. 0 (2014.01.0)	Appear as if SEQUEST files were loaded
Mascot® Server (Matrix Science)	*.dat	2.4 .1	
X! Tandem® =	*.xml	SLEDGEHAMMER (2013.09.01)	Version run by Scaffold: CYCLONE (2010.12.01.1)
Phenyx (GeneBio)	*.xml	2.6	GeneBio discontinued this product
OMSSA	*.omx		NCBI Discontinued this search engine in 2010
Waters® IdentityE	*.xml	3.0 and lower	
Agilent Spectrum Mill	Individual Results Directory	B.04.00	
Mascot Distiller® (Matrix Science)	*.xml	2.5.1.0 MDRO 2.5.0.0	The *.rov file, which points to the *.dat file, must be in the same directory as the *.xml file)
Proteome Discoverer® (PD) (Thermo Scientific)	*.msf	1.4 and older	
MS Amanda (in PD 1.4)	*.msf	PD 1.4	
Byonic (in PD 1.4)	*.msf	PD 1.4	Not supported yet. Feature request in the works
MaxQuant/Andromeda	Individual Results Directory	1.4 and older	
PEAKS	mzIdentML (*.mzid, *.mgf)	7 and older	
MyriMatch	mzIdentML (*.mzid, *.mgf)	2.1	
MS-GF+	mzIdentML (*.mzid, *.mgf)	2014.03.26	
Byonic	mzIdentML (*.mzid, *.mgf)	1.4-76	
For more detailed information: http: //www.proteomesoftware.com/pdf/file_compatibility_matrix.pdf			

Loading Data in Scaffold perSPECtives

Scaffold perSPECtives allows inspection and analysis of peptide and protein identification data from many sources, including:

- Scaffold
- Mascot
- IDPeaker
- PEAKS
- Byonic

and other sequence database and spectral library search applications. The application creates experiments by loading *.mzid or *.mzid.gz files version 1.1.0 and higher. It does not require loading of the corresponding peak list files, *.mgf files, but if they are included, spectra will be available for inspection. Loading only *.mzid files reduces considerably the size of the Scaffold perSPECtives experiment file.

Scaffold perSPECtives is not designed to perform any protein assembly or protein scoring. This means that MZID files from applications that provide only peptide information cannot be loaded directly into perSPECtives. These files may be processed first by Scaffold and the resulting MZID may be loaded into perSPECtives. Some examples of incompatible tools are:

- SpectraST (through pepXML file conversion)
- MyriMatch
- Pepitome.

For a more comprehensive explanation of mzIdentML support in perSPECtives see the www.proteomesoftware.com/pdf/scaffold_perspectives_users_guide.pdf

Loading quantitative data in Scaffold, Scaffold Q+ and Scaffold Q+S

Mascot Distiller

Scaffold, Scaffold Q+ and Scaffold Q+S load Precursor intensity analysis results, while Scaffold Q+S also loads Stable Isotope Labeling data. Both types of analysis can be run using Mascot Distiller. Find below instructions on how to set up Mascot Distiller for analyzing these types of quantitations and how to load the results into Scaffold.

Stable Isotope labeling quantitative data

SILAC quantitation

Follow these instructions to run Mascot Distiller analysis on SILAC data

1. Open Distiller and select RAW files to process, by going to File > New > Project.
2. Either choose Thermo XCalibur one file or open a **Multi-File Project**.
3. When RAW files are ready, choose **Processing > Process and Search**. This will open the Mascot server window.
4. Follow directions in the **Mascot Search Dialog** to set up a search. Be sure to select the correct SILAC quantitation method for your experiment.
5. Choose FASTA database file and quantitative settings, if applicable. Be sure to set any other settings here as well, like variable and fixed modifications.Then click Search.
6. Once the search is completed save the Mascot Distiller project by going to **File > Save Project As...**This will create a *.rov type of file.
7. To compute the quantitative values select **Choose Analysis > Quantitate.** You may need to adjust settings here, eg, All Families.
8. Then, for a Scaffold compatible export, choose **Analysis > Quantitative Report > Save as XML**.
9. When you load the data in Scaffold you will need the ROV file as well present in the same directory that contains the *xml file

Dimethyl Labeling-based Quantitation

The work-flow for this type of quantitation is the same as described in the SILAC quantitation section with the exception that when setting up the Mascot search, you must select Dimethylation [MD] as the quantitation method.

Precursor Intensity (AUC Integration)

The work-flow for this type of quantitation is the same as described in the SILAC

quantitation section with the exception that when setting up the Mascot search, you must select Average [MD] as the quantitation method.Then, when the search is completed do the following:

1. Select the menu option **Analysis > Calculate XIC**
2. Choose **Analysis > Quantitate** to export the XML file.

Loading Mascot Distiller results into Scaffold

When loading Mascot Distiller XML through the Scaffold Wizard, select the quantitative technique corresponding to the type of quantitation searched in the XML files you want to load.

- • For SILAC and Dimethyl-based quantitation select **Stable Isotope Labeing (Multiplex)**
- For simple Precursor Intensity select **Precursor Intensity (Standard)**

To load Stable Isotope labeling or Precursor Intensity data (AUC Integration) into Scaffold, select the XML file. However, the ROV project file (created using the command **File > Save Project As…**) must be in the same directory as the XML file. The ROV file contains information that helps Scaffold trace back to the DAT file on the Mascot server.

If the DAT file cannot be accessed on the Mascot server, then you must copy it to the same directory as the XML and ROV files. In general it is better to have all three files, XML, ROV and DAT, saved in the same location.

Thermo Proteome Discoverer

Scaffold, Scaffold Q+ and Scaffold Q+S load Precursor intensity analysis results, while Scaffold Q+S also loads Stable Isotope Labeling data. Both types of analysis can be run using Proteome Discoverer 1.3 and higher. Find below instructions on how to set up PD for analyzing these types of quantitations and how to load the results into Scaffold.

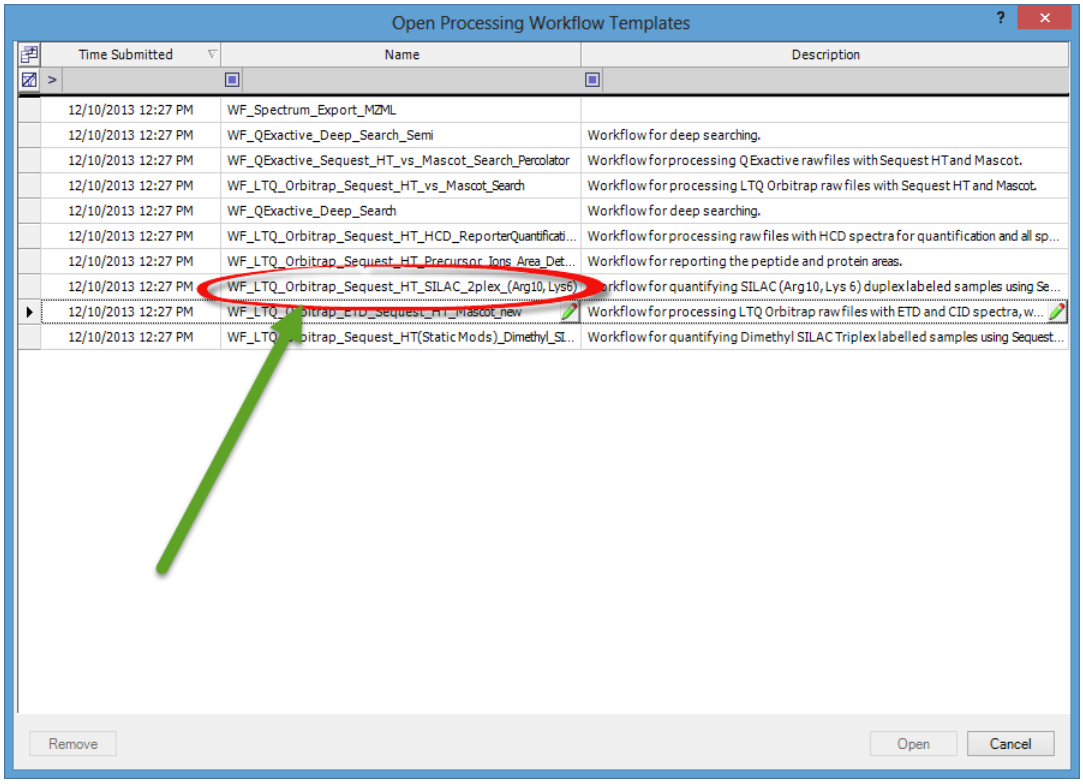
Stable Isotope labeling quantitative data

SILAC quantitation

To set up a search for SILAC quantitation in Proteome Discoverer (PD) 1.3 and higher start from the standard work-flow templates available in PD.

1. Start PD, select **Workflow Editor > Open From Template...** from the main menu options.
2. From the list of work-flow templates appearing in the dialog select WF_LTQ_Orbitrap_Sequest_HT_SILAC_2plex(Arg10,Lys6), see Figure 1.

Figure 1: Proteome Discoverer: work-flow templates

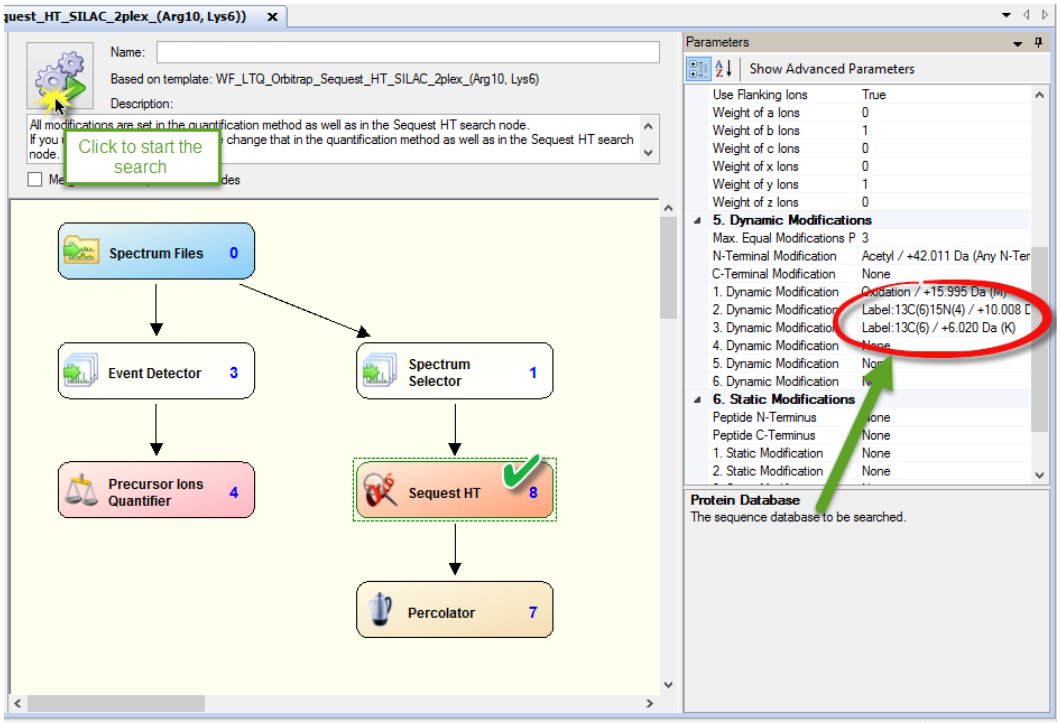


3. This particular workflow uses Sequest HT as the search engine, but you can substitute this node or add another search engine node like Mascot, regular Sequest or other search engine nodes available in your copy of PD.

4. Select the Sequest HT node and check the type of modifications added for SILAC quantification. The isotopic labels are added as variable modifications, see Figure 2. SILAC experiments can use a variety of heavy isotopic labels and combine them in duplex or triplex type of experiments. When setting up the search adjust the variable modifications accordingly. For more information consult the chapter Quantification in the Proteome Discoverer User Guide.

5. To optimize Scaffold analysis of PD search results we also advise the User to adjust a number of default parameters appearing in the various Sequence Database Search nodes. For more information please check the Scaffold User's Manual Configuring Proteome Discoverer Sequest, Sequest HT and Mascot nodes

Figure 2: Proteome Discoverer: SILAC workflow template

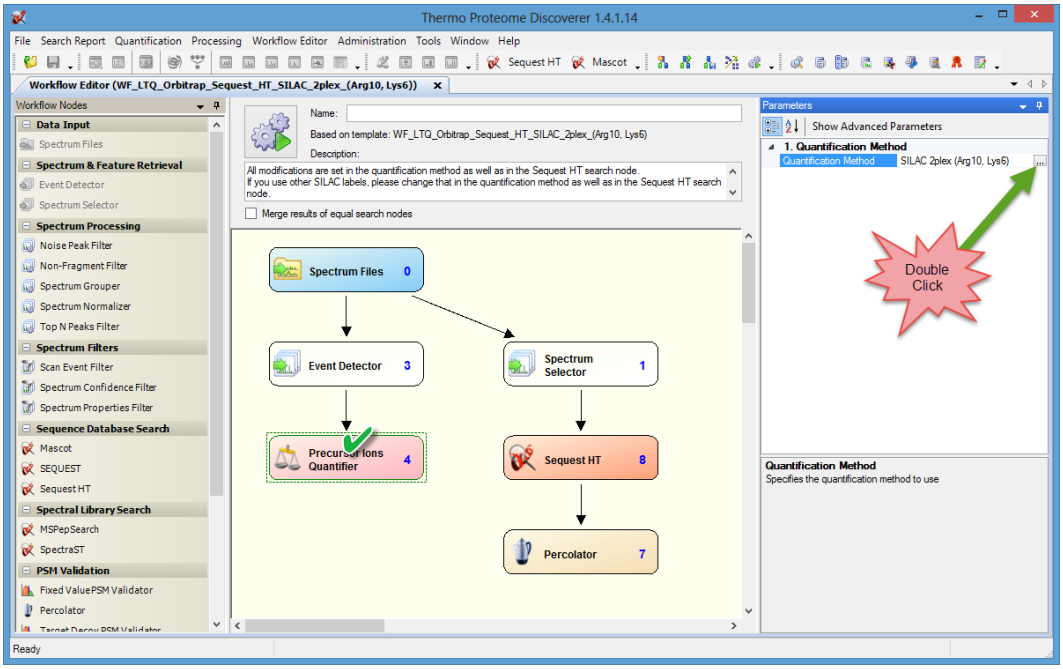


6. In this particular workflow Percolator is selected as the PSM validation node. Note that Scaffold will not read Percolator results but will use the generated decoys to calculate the protein and peptide FDR values shown in the Samples View.

7. Now select the Precursor Ion Quantifier node, see Figure 3. The parameter pane shows the Quantification method selected for this particular workflow.

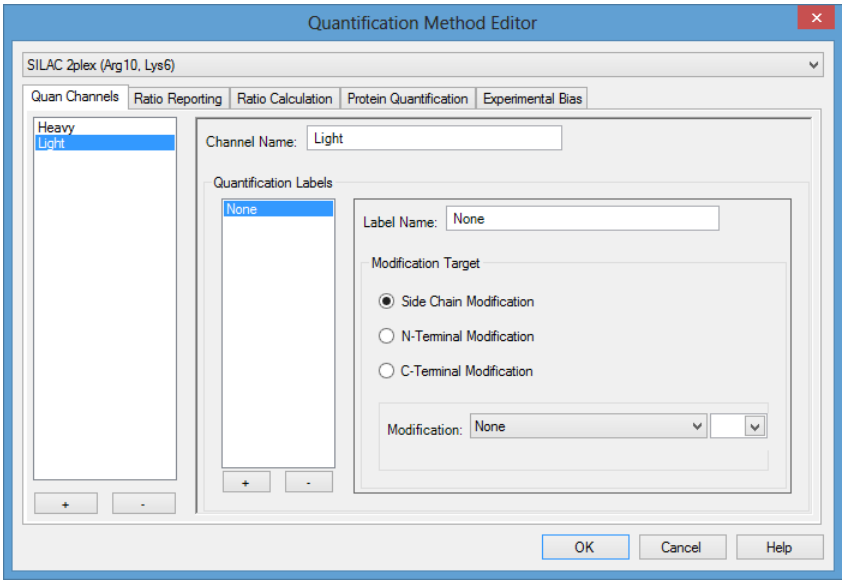
8. If you want to change or adjust the labels used in your SILAC experiment double click on the little square containing dots and the Quantification method Editor opens.

Figure 3: Proteome Discoverer: Precursor Ion Quantifier



9. In the Quantification Method Editor, see Figure 4, you can either select the Quantification method you need through the pull down list at the top of the dialog or adjust the parameters for the current one as you wish. Make sure that the modifications labels you set up correspond to the ones you have added in the search node, see Step 4

Figure 4: Proteome Discoverer: Quantification Method Editor



10. Once the parameters are properly selected, name the search and run it by clicking the button shown in Figure 3. Once the search is completed the results will be saved in a *.msf file.

Dimethyl Labeling-based Quantitation

Proteome Discoverer supports the dimethylation 3plex method to compare up to three samples. The User cannot apply labels to the C terminus, nor to arginine.

Setting up this type of quantitation search in PD works exactly like the SILAC quantitation set up described above. You just need to adjust the variable modifications added in the search engine node accordingly, see Step 4 and select the Dimethylation 3plex Quantification method from the pull down list in Step 9.

Precursor Intensity (AUC Integration)

Proteome Discoverer provides a workflow template for computing precursor intensity values. As shown for SILAC Quantitation Step 2, select the template **WF_LTQ_Orbitrap_Sequest_Precursor_ions_Area_Detector** for precursor intensity label free quantitation. The template can be used as a starting point, and the search engine choice or instrument settings may be changed. Scaffold reads the precursor intensities from the MSF file.

Loading Proteome Discoverer results into Scaffold

When loading Proteome Discoverer quantitative results into Scaffold select from the Loading Wizard the quantitative technique corresponding to the type of quantitation searched in the MSF files you want to load.

- For SILAC and Dimethyl-based quantitation select **Stable Isotope Labeling (Multiplex)**
- For simple Precursor Intensity select **Precursor Intensity (Standard)**

When asked to select files for loading point Scaffold to the MSF files you want to load.

PEAKS

The Scaffold suite of programs can now load mzIdentML export files from a variety of different search engines, including PEAKS.

Running Peaks

For information on how to install and run PEAKS, please go to the Bioinformatics Solutions inc website.

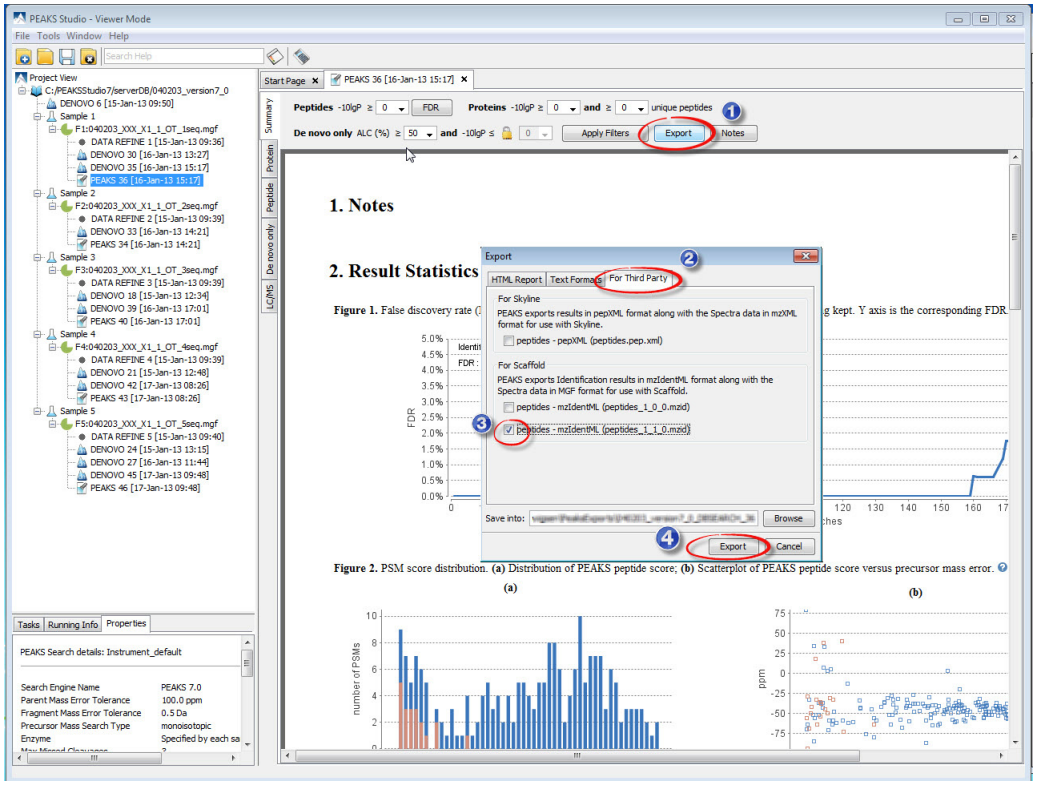
Exporting mzIdentML files from PEAKS

PEAKS 7

This version of PEAKS provides a third party export feature designed to export mzIdentML and MGF files for loading into Scaffold:

1. Click the Export button in the title bar of the search Summary view panel, see Figure 5(1). This opens an export dialog.

Figure 5: PEAKS7: Exporting mzIdentML files



2. Select the **For Third Party** tab, see Figure 5(2).

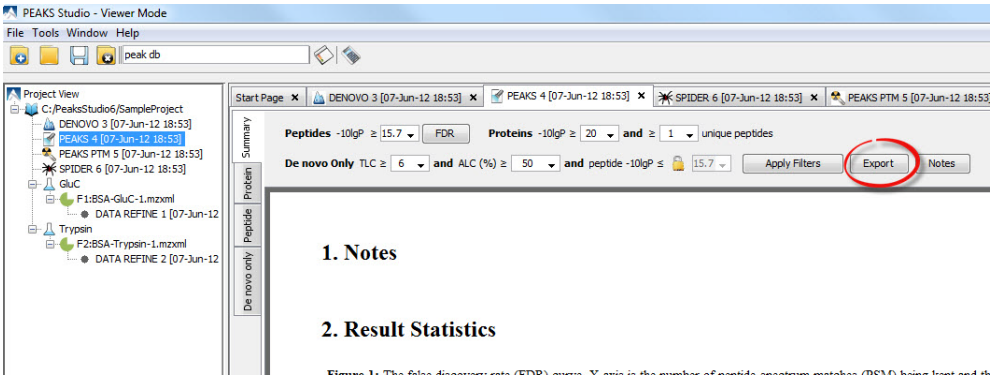
3. Choose one of the mzIdentML format available, see Figure 5(3). Scaffold can load either the 1.0.0 or the 1.1.0 format.
4. Select a location for saving and click the Export button, see Figure 5(4). A folder will be created which contains the MZID and MGF files.
5. Load the MZID into Scaffold. Scaffold will read the spectra from the corresponding MGF file because it is located in the same folder.

PEAKS 6

For PEAKS 6 follow the steps described below to export mzIdentML files:

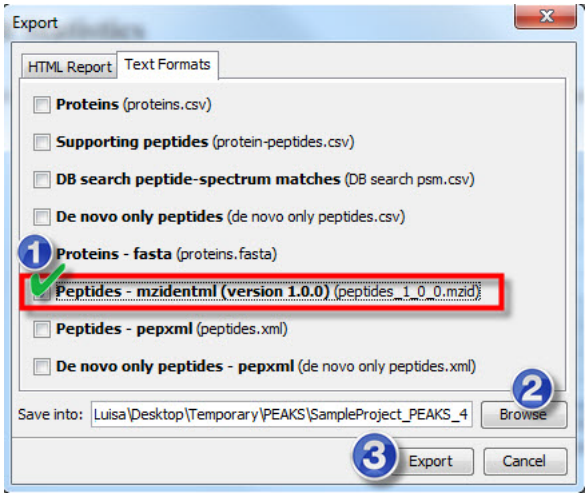
1. Starting on the PEAK DB Summary View, click the Export button, see Figure 6.

Figure 6: PEAKS DB Summary view



2. When the export dialog opens, click on the Text Formats tab and Check the box Peptides - mzidentml (version 1.0.0), see Figure 7(1). If you do not have to export any other file, leave all the other boxes unchecked

Figure 7: PEAKS Export dialog

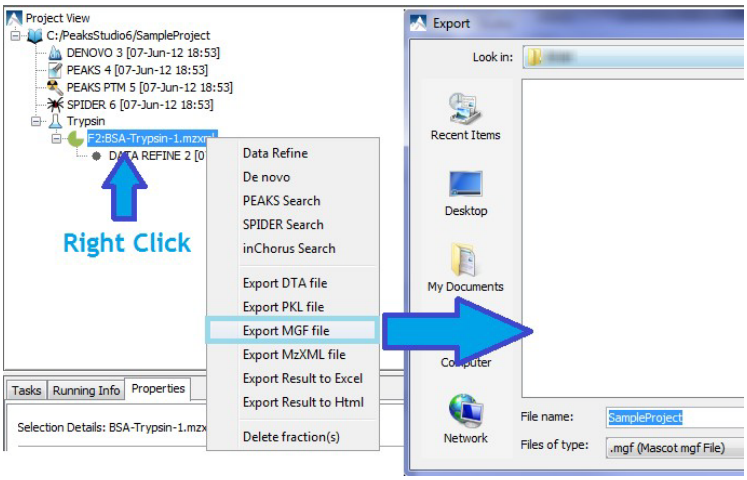


3. Browse to the location where you wish to save the export file, see Figure 7(2).
4. Click Export, see Figure 7(3)
5. The exported file (peptides_1_0_0.mzid) will appear in the specified location in a folder called PEAKS Identification_Sample Name.

Most of the programs in the Scaffold Suite require both the MZID and the accompanying MGF peak list file. Here is the procedure for exporting an MGF file from PEAKS:

1. Right click on the fraction node: the green icon (same procedure if User selects Project Node, blue book icon).

Figure 8: PEAKS: Exporting MGF files



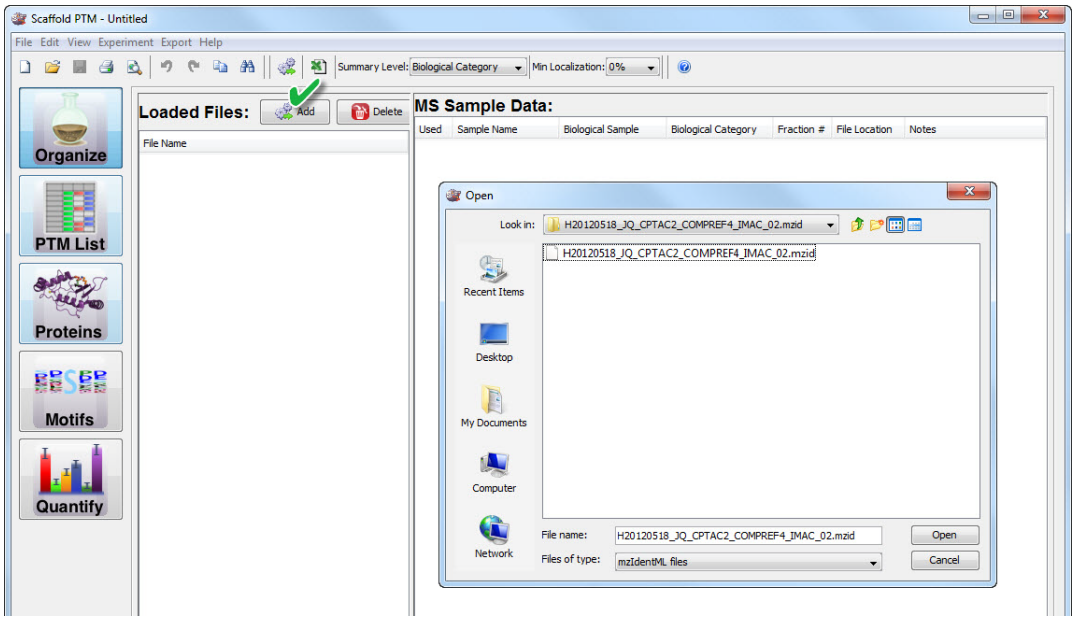
2. Choose Export MGF File, see Figure 8.
3. The export browser opens. Assign a name and specify the location where you wish to store the MGF file. We typically suggest to save the file in the same directory where the *.mzid file was saved.

Loading PEAKS results into Scaffold PTM

To load PEAKS MZID files into Scaffold PTM follow these instructions:

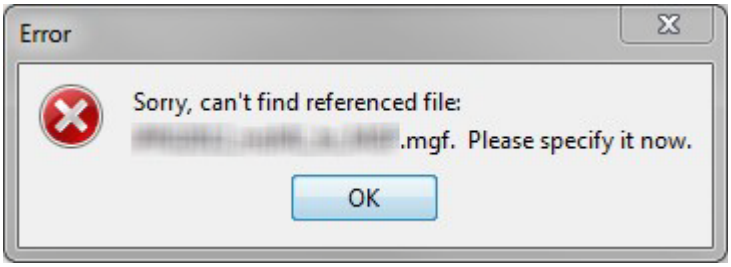
1. Open PTM
2. Select **New** experiment.
3. Click the **Add** icon in Scaffold PTM **Loaded Files** pane, see Figure 9.

*Figure 9: Scaffold PTM: load *.mzid file*



4. When the file browser appears locate and select the MZID file you exported from PEAKS.
5. Click Open and the MZID file will be listed in the Scaffold PTM Queue Data Files dialog, click Load.
6. If you did not copy the MGF adjacent to the MZID and name it the same as the MZID, Scaffold PTM will prompt you for the MGF file when it needs it.

Figure 10: Scaffold PTM warning for locating MGF files



Loading PEAKS Results into Scaffold

When selecting Queue Files for loading a browser appears. Locate and select the MZID files you exported from PEAKS.

Note that Scaffold, when loading MZID files, is not going to alert you if it dos not find the related MGF files. It will simply not report the spectrum for the peptide shown in the proteins view.

Spectrum Mill

Hereafter we provide a few suggestions on how to set up Spectrum Mill runs to simplify loading of its search results into Scaffold, Scaffold Q+, Scaffold Q+S.

Running Spectrum Mill

For Spectrum Mill server/client setup instructions and Quick Start Guide, please see the Spectrum Mill Documentation: www.chem.agilent.com/Library/usermanuals/Public/G2721-90036_SpectrumMill_QuickStart.pdf.

1. Start by opening the Spectrum Mill program to configure the FASTA databases and extractor settings.
   - Be sure to add any FASTA databases in the Protein Databases Utilities.
2. Next, browse to the directory where you will be storing your Spectrum Mill results. Ultimately, the data you will load into Scaffold will be a directory and this directory contains all results data Scaffold needs to process Spectrum Mill results.

   When Spectrum Mill is installed and configured, it creates a directory called **Spectrum Mill\msdataSM**. We recommend creating a directory, within **msdataSM**, for each search/analysis you will be doing. Name the directory with something that is descriptive for archiving, like:

   <date>_<descriptor>_<descriptors-params>_<further-descriptors-params>

   For example the name of the directory containing the Spectrum Mill search results could be something like:

   2012-0415_spectrum-mill_bob-jones_phase-1
3. The next step is to copy the raw data file(s) you wish to analyze into your newly created search directory.

   At this point you have one directory with one or more raw data files inside.
4. Spectrum Mill has the option to create Work-flows that streamlines your extraction, search and summary. The next step is to either build a work-flow or run each step individually.
5. Depending on how you are licensed, you can load and extract a variety of files.Since there are very many settings in Spectrum Mill, covering them all here is beyond the scope of this document.

   You should, at the minimum, check the tolerances, instrument type, modifications, search mode and validation in the summary. Finally, be sure to confirm labels and modifications.

Loading Spectrum Mill Results Directory into Scaffold

Scaffold locates files you want to load through the **Queue Files For Loading...** command

which can be selected from the following locations in the program:

- The Experiment menu, **Experiment > Queue Files For Loading...**
- The Load data View, clicking the button **Queue Files For Loading**
- The **Queue files for loading** page in the Wizard

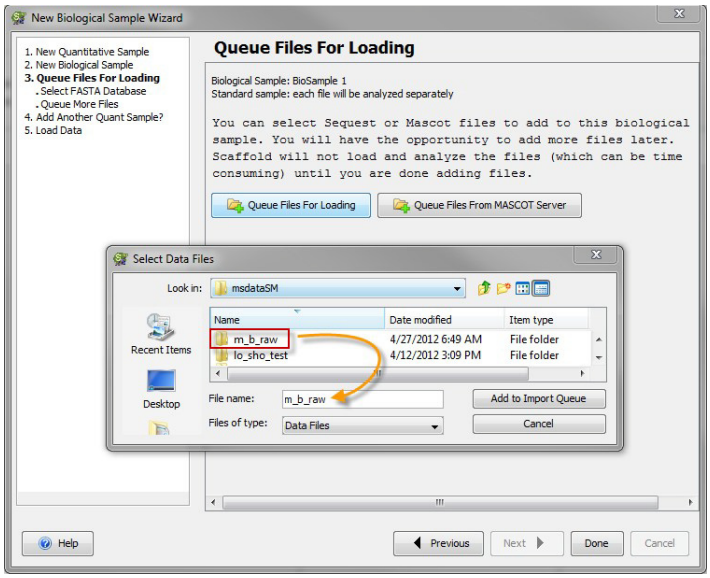
The command opens a browser that allows you to locate the files you want to load.

When prompted to load data, choose the directory that you created following the instructions provided in Running Spectrum Mill; as of the example provided you would choose: 20120415_spectrum-mill_bob-jones_phase-1.

There are two ways to select the Results Directory

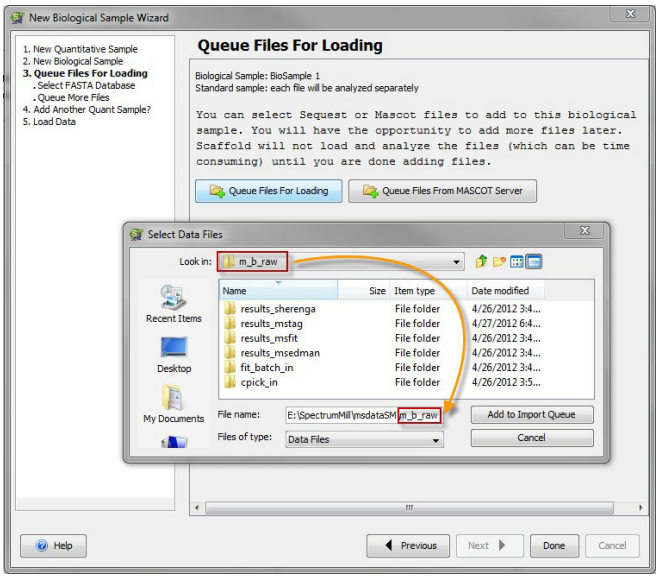
1. You can simply navigate to the directory of interest and left click it once, so that it is highlighted, see Figure 11.

Figure 11: Selecting directory in Scaffold with left click



2. Navigate to the directory of interest and double click it so that the contents are visible in the dialog box, see Figure 12.

Figure 12: Double Clicking directory in Scaffold and contents are visible



Both options will give the same results. In either case, you are selecting the directory that now contains the Spectrum Mill results you want to load into Scaffold.

Next, click Done and follow the directions in the Scaffold Load Data Wizard. When finished, you should see the results in Scaffold.

MaxQuant

Running MaxQuant

Hereafter we provide a few suggestions on how to set up MaxQuant (MQ) runs to simplify the loading of its search results into Scaffold, Scaffold Q+, Scaffold Q+S. MaxQuant versions 1.2.2.5 and above are currently supported.

Stable Isotope Quantitative data

SILAC Quantitation

1. Open the AndromedaConfig.exe program to configure Andromeda search engine. Be sure to add any FASTA databases, proteases or additional modifications in this window.
2. Save the configuration by going to **File > Save > all** (or modifications, proteases, or databases).
3. Browse to the directory where you will be storing your MaxQuant results.

   Ultimately, the data you will load into Scaffold will be a directory and this directory contains all results data Scaffold needs to process MaxQuant results.
4. Outside of MQ create the following directories: (please note that if you have an existing workflow, this step can be skipped if you feel comfortable loading MaxQuant data into Scaffold already).
   - •Create a directory for each search/analysis you will be doing that is named in a descriptive fashion for archiving purposes:

     <date>_<descriptor>_<descriptors-params>_<further-descriptors-params>

     Like for example:

     2012-0415_maxquant_3-plex_bob-jones_phase-1
5. Copy the RAW file(s) you wish to analyze. At this point you have one directory with one or more RAW files inside.
6. Open MaxQuant.exe and load the RAW file(s) you added to the directory described above.

While setting up the MaxQuant analysis, keep the following in mind (Please note that these are only recommendations, not requirements):

- Check the FDR settings. If the settings are too low, some peptides may be missed.Try setting the peptide and protein FDR values to 1.0 if in doubt.
- Check the Keep low-scoring versions of identified peptides drop-down and consider running with the Also between parameter groups for more matches.
- Finally, be sure to confirm labels and modifications. There are three locations where modifications can be set. Evaluate all of them carefully.

Precursor Intensity (AUC Integration)

Follow the basic workflow described in SILAC Quantitation but with the following comments in mind.

Precursor intensity may be computed when analyzing a single raw file in MQ 1.4 (as opposed to MQ 1.3) if the user selects the Label Free Quantitation option. Individual results may then be loaded into separate BioSamples in Scaffold and used for Precursor Intensity Quantitation in either Scaffold, Scaffold Q+ or Scaffold Q+S.

If two or more raw files are analyzed together in MQ1.4 with the LFQ option selected, they form a single combined folder which loads into Scaffold as a single BioSample. In this case, Scaffold, Scaffold Q+ or Scaffold Q+S are unable to perform Precursor Intensity Quantitation.

In MQ 1.3, a multi-raw-file run created an experiment file, and since, at the time, this was the only method of running LFQ, Scaffold has a special dialog that opens when the program recognizes the presence of an experiment file. This dialog asks which file should be loaded into the current BioSample, allowing the User to load each experiment into its own BioSample and thus to perform precursor intensity quantitation.

It is possible, although not required, in MQ 1.4 to create an experiment file. The experiments can be named through the MQ 1.4 GUI, and then an experiment file can be exported by right-clicking and choosing Export. The user should name the file “Experiment.txt” and then Scaffold will pick it up and loading can proceed as it did for MQ 1.3 files.

Loading MaxQuant Results Directory into Scaffold

Scaffold locates files to load through the **Queue Files For Loading...** command which can be selected from the following locations in the program:

- The Experiment menu, **Experiment > Queue Files For Loading...**
- The Load data View, clicking the button **Queue Files For Loading**
- The **Queue files for loading** page in the Wizard

The command opens a browser that allows you to locate the files you want to load.

When loading MaxQuant quantitative results into Scaffold select from the Loading Wizard the quantitative technique corresponding to the type of quantitation searched in the MaxQuant results you want to load.

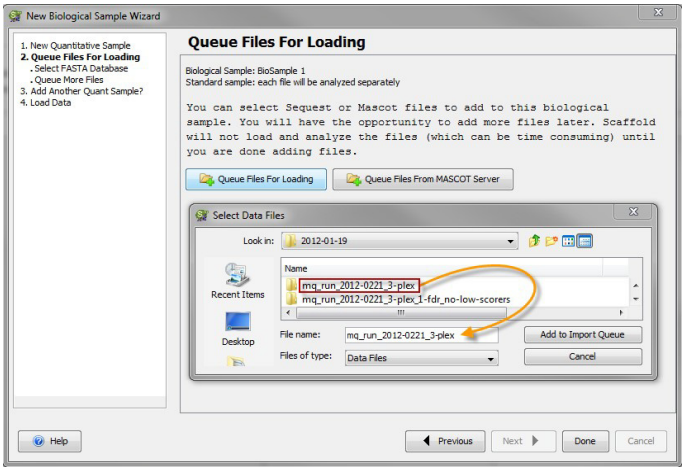
- For SILAC and Dimethyl-based quantitation select **Stable Isotope Labeling (Multiplex)**
- For Precursor Intensity select **Precursor Intensity (Standard)**

When prompted to load data, choose the directory that you created following the instructions provided in SILAC Quantitation; in the example provided you would choose: 20120415_maxquant_3-plex_bob-jones_phase-1.

There are two ways to **select** the **Combined Results Directory**:

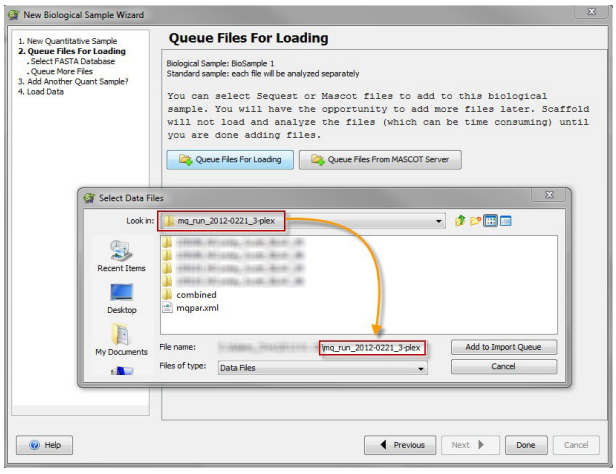
1. Simply navigate to the directory of interest and left click it once, so that it is highlighted, see Figure 13.

Figure 13: Selecting directory in Scaffold with left click



2. Navigate to the directory of interest and double click it, so that the contents are visible in the dialog box, see Figure 14.

Figure 14: Double Clicking directory in Scaffold and contents are visible



Both options give the same results. In either case, you are selecting the directory that now contains the MaxQuant results you want to load into Scaffold.

Next, click Done and follow the directions in the Scaffold Load Data Wizard. When finished, you should see the results in Scaffold.

Some Comments on Required MaxQuant Files for Scaffold

The User may realize that dealing with the entire MaxQuant results directory can be cumbersome; some of these directories get to be large: 8 GB or more. Furthermore, Scaffold does not require the entire results directory. In fact, it only requires a few files inside the whole directory.

For *MaxQuant version 1.3*, the files that Scaffold reads are:

- All APL files: **combined/*.apl**
- **combined/txt/msms.txt**
- **combined/txt/evidence.txt**
- **combined/txt/summary.txt**
- **combined/txt/parameters.txt**

For *MaxQuant version 1.4*, the files that Scaffold reads are: (note the andromeda directory inside combined)

- All APL files: **combined/andromeda/*.apl**
- **combined/txt/msms.txt**
- **combined/txt/evidence.txt**
- **combined/txt/summary.txt**
- **combined/txt/parameters.txt**

Figure 15: MaxQuant: Parent directory showing combined directory

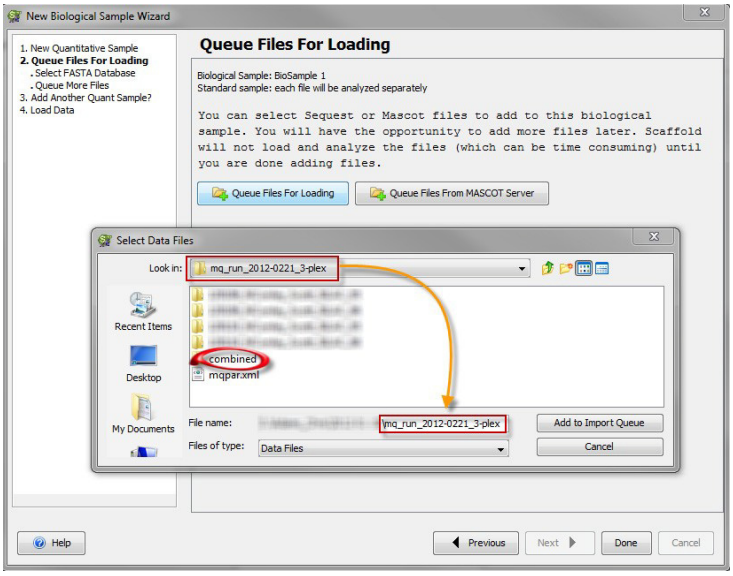
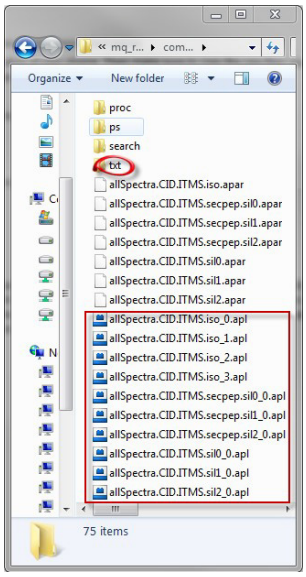


Figure 16: APL files and txt directory



Some Comments

On FASTA Database Files

- As is the case for all search engines and Scaffold, be sure to use exactly the same FASTA file as you did in the search engine or proteomics application described in Loading quantitative data in Scaffold, Scaffold Q+ and Scaffold Q+S.
- When you search your data make sure you choose the parse rules that best suit the type of FASTA you are using, eg, UniProt, NCBI, IPI, etc.
- When adding the database in Scaffold to obtain the optimal parsing and get molecular weights and peptide sequences, the best option is given by **Auto Parse**.
- When Scaffold needs more directions in identifying the correct parsing rules, the option **Use Regular Expressions** provides a variety of tools that help optimize the selection of th proper parsing rules.
- Please note that **Spectrum Mill** uses a modified parse rule for NCBI accessions, eg, instead of gi|123456, only the number is used: 123456. Scaffold has a specific parse rule to match this feature when selecting the option **Use Regular Expressions**, so when loading your database in Scaffold, keep this in mind.

Loading PLGS results in Scaffold

Check the following document for detailed information on how to load PLGS data in Scaffold: http://proteome-software.wikispaces.com/file/view/White_paper_scaffold_4_PLGS_3_plugin.pdf/.

Release Information

The following release information applies to this version of the *White Paper:scaffold loading search egine results*. This document is applicable for Scaffold, Release 4.0 or greater, and is current until replaced.

Document Version Number	*Scaffold 4.0-Load_search_results_rev_9*
Document Status	*Released*
Document Release Date	*April 23, 2014*

Customer Support

Customer support is available to organizations that purchase *Scaffold, Scaffold Q+* or Scaffold Q+S and that have an annual support agreement. Contact Proteome Software at:

Proteome Software, Inc.
1340 SW Bertha Blvd, Suite 10
Portland, OR 97219
1-800-944-6027 (Toll Free)
1-503-245-4910 (Fax)
www.proteomesoftware.com