Compound Discoverer 3.1 Tutorial for E & L Studies

To familiarize yourself with the Thermo Compound Discoverer™ 3.1 application and its features that address the data analysis requirements of the Extractables & Leachables market, follow the topics in this tutorial to set up a study and an analysis, process a set of example Xcalibur™ RAW files, review the result file produced by the analysis, and create a report.

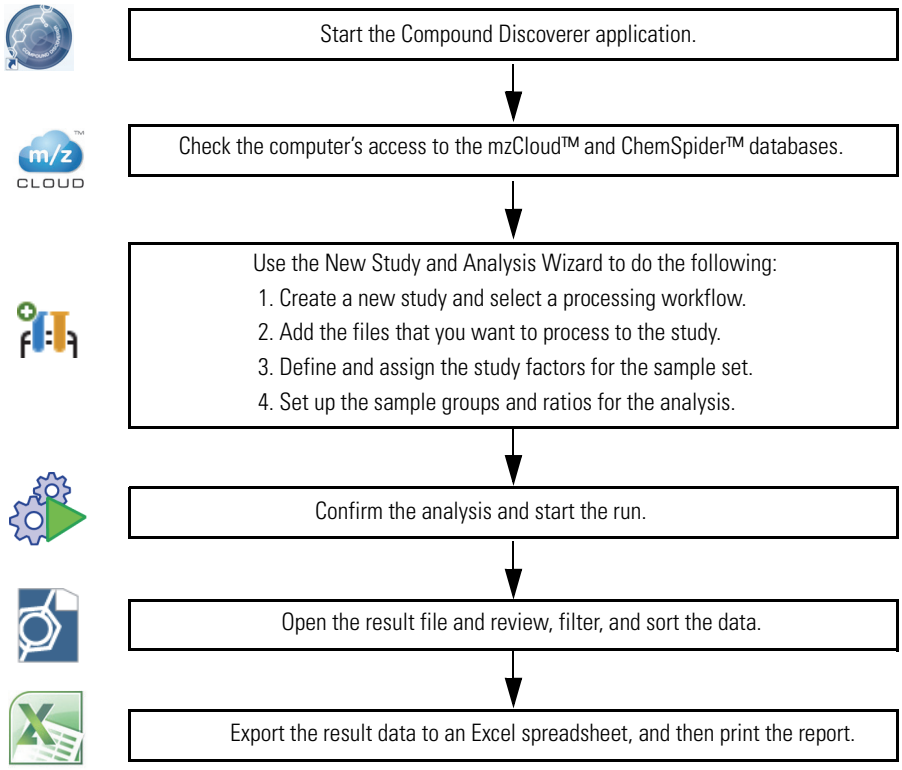
Contents

Overview

To create a practice study, use the example Xcalibur RAW files. These files are provided on the key-shaped USB key in the software media kit. You can find these files in the following folder:

Example Studies\E & L O-ring Study\

Copy the E & L O-ring Study folder to your data processing computer. This figure shows the tutorial's workflow.



Revision A XCALI-98116

thermoscientific

The E & L O-ring Study folder contains the following files.

Figure 1. Example files for E&L studies

Name	Type
Red_H2O_2.raw	RAW File
Red_H2O_1.raw	RAW File
Red_100EtOH_2.raw	RAW File
Red_100EtOH_1.raw	RAW File
Red_50EtOH_2.raw	RAW File
Red_50EtOH_1.raw	RAW File
H2O_2.raw	RAW File
H2O_1.raw	RAW File
EandL O-ring Example.cdStudy	CDSTUDY File
EandL Example.cdResultView	CDRESULTVIEW File
EandL Example.cdResult	CDRESULT File

Name	Type
Brown_H2O_2.raw	RAW File
Brown_H2O_1.raw	RAW File
Brown_100EtOH_2.raw	RAW File
Brown_100EtOH_1.raw	RAW File
Brown_50EtOH_2.raw	RAW File
Brown_50EtOH_1.raw	RAW File
100EtOH_2.raw	RAW File
100EtOH_1.raw	RAW File
50EtOH_3.raw	RAW File
50EtOH_2.raw	RAW File
50EtOH_1.raw	RAW File

Starting the Application

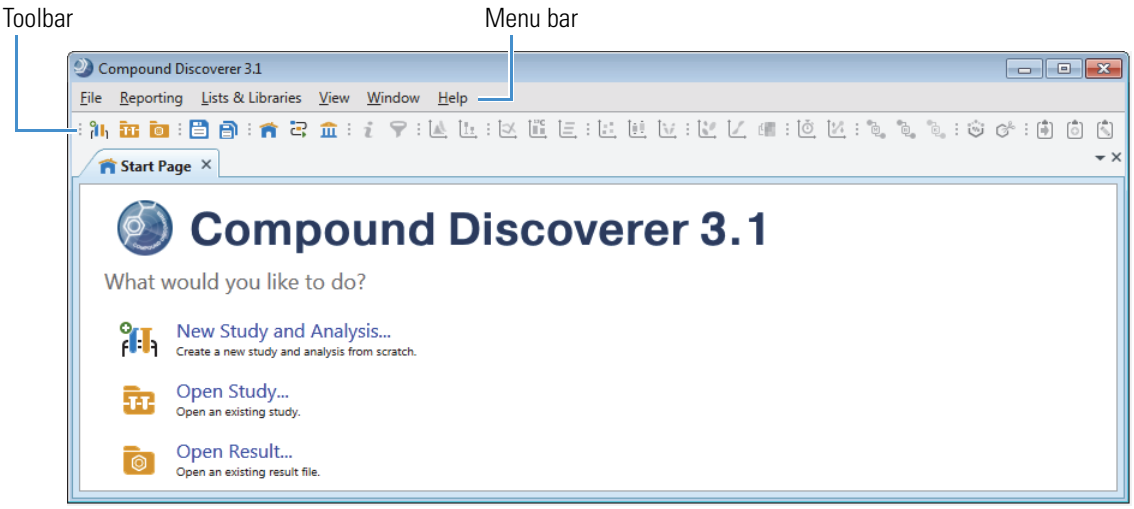
❖ To start the application

- From the taskbar, choose **Start > All Programs** (or **Programs**) **> Thermo Compound Discoverer 3.1**.

–or–

- From the computer desktop, double-click the **Compound Discoverer** icon,  .

The application opens to the Start Page.



Accessing Help

The application provides Help for the views, tabbed pages, and dialog boxes.

❖ To open the Help topic for a specific view, tabbed page, or dialog box

1. Open the view, tabbed page, or dialog box.
2. On the computer keyboard, press the F1 key. 

The application also provides a user guide, four tutorials by study field, and a quick start guide as PDF files.

❖ To access the user documentation

1. From the menu bar, choose **Help > Manuals**.
2. Select from the following documents.

Compound Discoverer Metabolism Tutorial
Compound Discoverer Metabolomics Tutorial
Compound Discoverer E&L Tutorial
Compound Discoverer Stable Isotope Labeling Tutorial
Compound Discoverer Reporting Quick Start
Compound Discoverer User Guide

Checking the Computer's Access to the External Databases

This tutorial uses a processing workflow that searches the online mzCloud and ChemSpider databases and the local Extractables and Leachables mass list that comes with the application. To run the online searches, your processing computer must have unblocked access to the databases on the Internet.

❖ **To verify that your computer has access to the online databases**

1. From the menu bar, choose **Help > Communication Tests**.
2. Run the mzCloud communication tests as follows:
   a. Click the **mzCloud** tab.
   b. Click **Run Tests**. When the tests are complete, go to the next step.
3. Run the ChemSpider communication test as follows:
   a. Click the **ChemSpider** tab.
   b. Click **Run Tests**.

If your computer has an Internet connection and these tests fail, leave the Communication Tests dialog box open and press the F1 key to open the Help to the testing communication topic. Then, follow the instructions to troubleshoot the communication failure. If your computer does not have an Internet connection, you can complete this tutorial by modifying the processing workflow (see "To remove the search nodes that require Internet access" on page 15).

Go to the next topic "Setting Up a New Study and a New Analysis."

Setting Up a New Study and a New Analysis

Make sure to copy the Xcalibur RAW files from the USB flash drive (in the software media kit) to an appropriate folder on your processing computer. See "Overview" on page 1.

Follow these topics to create a new study and a new analysis:

Opening the New Study and Analysis Wizard

To set up a new study, use the New Study and Analysis Wizard.

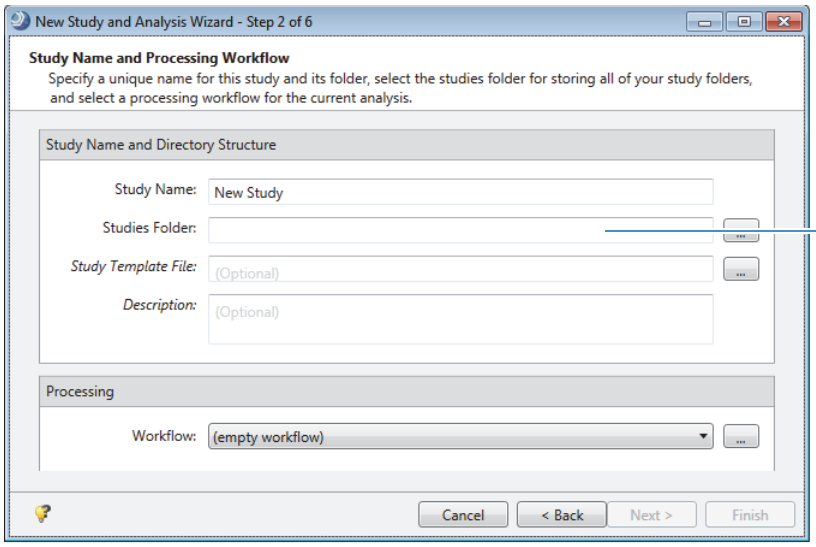
❖ **To open the New Study and Analysis Wizard**

From the menu bar, choose **File > New Study and Analysis**.

The New Study and Analysis Wizard opens.

Click **Next** to open the Study Name and Processing Workflow page.

Use the second page of the New Study and Analysis Wizard to name the study, select the studies folder for storing all or some of your studies, and select a processing workflow for the current analysis.

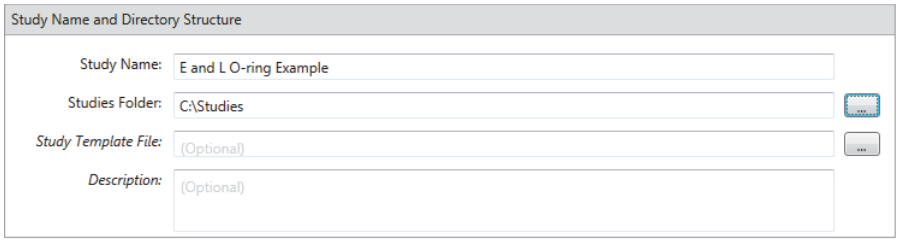


Setting Up the Study Folders

Each time you create a new study, the application creates a new study folder with the same name and stores the study file (.cdStudy) in the new folder. When you first install the Compound Discoverer application, you must set up a top-level folder for the study folders.

❖ To name the new study and set up the top-level folder

1. On the Study Name and Processing Workflow page, in the Study Name box, name the study.
   - For an extractables and leachables study with the o-ring example files, type **E and L Example**.
   - For a study that includes your own Xcalibur RAW files, type an appropriate name.
2. Select the folder where you want to store your Compound Discoverer study folders as follows:
   a. Click the browse icon, ..., next to the Studies Folder box.
   b. Browse to your local disk drive or a location on your local network.
   c. Click **New Folder** to create a new folder and name the folder **Studies**.



After you select or create a top-level folder, stay on this wizard page and go to the next topic, "Selecting the Processing Workflow."

When you complete the wizard, the application creates the E and L Example.cdStudy file, stores the study file in the E and L O-ring Example folder, and stores the E and L O-ring Example folder in the Studies folder. When you run an analysis, the application stores the result files (.cdResult) in the E and L O-ring Example folder.

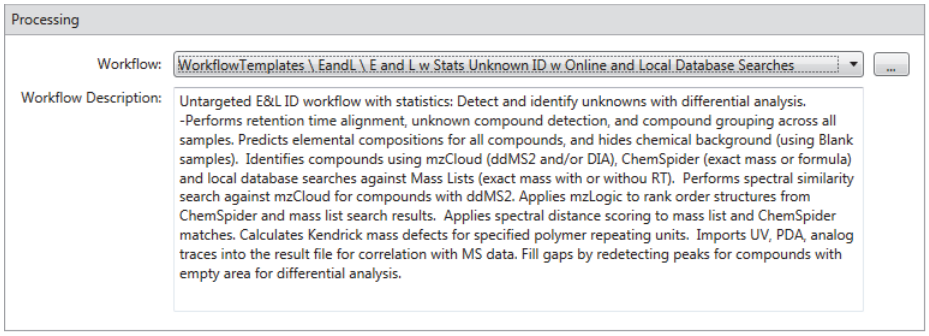
Selecting the Processing Workflow

In the Compound Discoverer application, the processing method that interprets the raw data is called a processing workflow (.cdProcessingWF). The application comes with defined processing workflows for several fields of study including the E & L field. This tutorial uses a defined processing workflow that searches the mzCloud and ChemSpider databases, a local mzVault library, and the Extractables and Leachables mass list to identify the compounds detected in a set of input files (Xcalibur RAW files).

❖ To select the processing workflow

1. Under Processing, select the following processing workflow from the Workflow list:

   Workflow Templates\EandL\E and L w Stats Unknown ID w Online and Local Database Searches



2. Click **Next** to open the Input File Selection page of the wizard.

Adding the Input Files to the Study

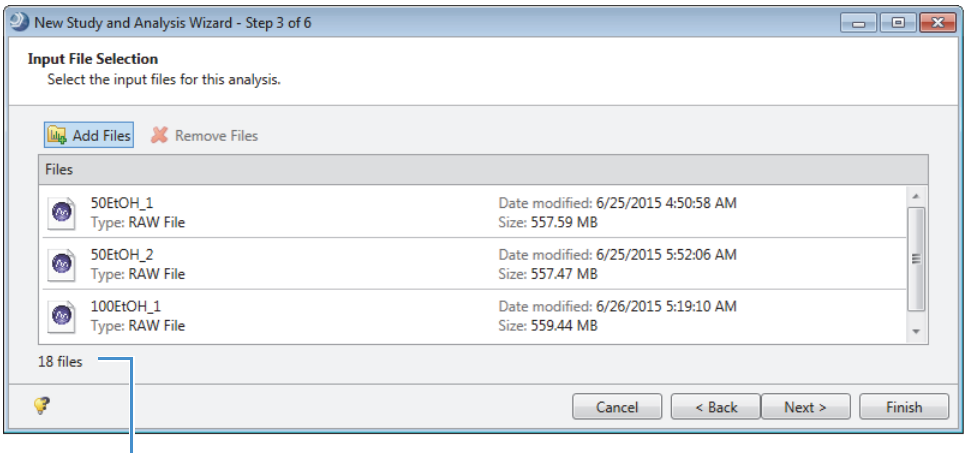
❖ To add input files to the study

1. On the Input File Selection page of the wizard, click **Add Files** in the command bar.

   The Add Files dialog box opens.

2. Browse to the folder where you copied the Xcalibur RAW files.
3. Select all 18 of the Xcalibur RAW files in this folder and click **Open**.

   The file names of the selected files appear in the Files box on the Input File Selection page.



Number of selected data files

Click **Next** to open the Input File Characterization page of the wizard.

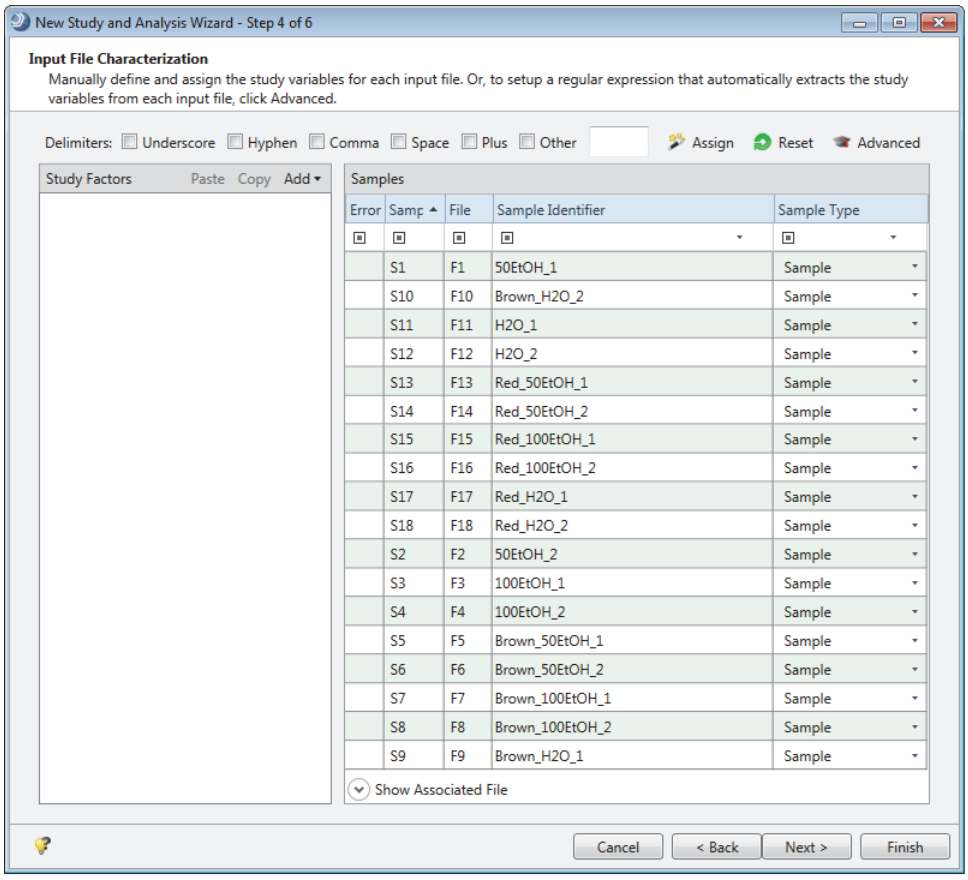
Defining the Study Variables

By default, the application assigns Sample as the Sample Type to new samples. To group the samples according to your experimental design, you must define the study variables, which include the sample types and the study factor values. In this tutorial, you are comparing the compounds extracted from two types of o-rings by three different solvents. The o-rings are identified by their colors—Red or Brown. The three solvents are water, 50% ethanol, and 100% ethanol. The sample types are Sample and Blank.

Follow these topics:

- Defining and Assigning the Study Factors
- Manually Selecting the Blank Sample Type

This figure shows the Input File Characterization page with the newly imported input files.



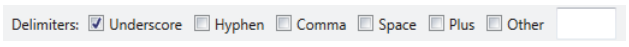
Defining and Assigning the Study Factors

To define and assign the study factors, follow these procedures:

1. To select the delimiters that separate the factors in the file names
2. To add a categorical study factor for the solvent
3. To add a categorical study factor for the o-ring color
4. To assign the study factor values to the samples

❖ To select the delimiters that separate the factors in the file names

Select the **Underscore** check box.

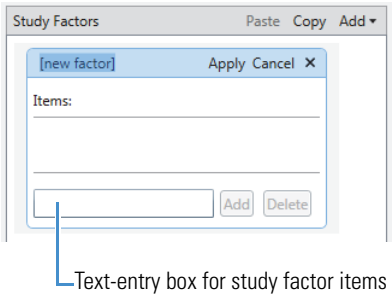


The study factors are o-ring (red or brown) and solvent (water, 50% ethanol, or 100% ethanol).

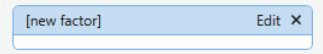
❖ To add a categorical study factor for the solvent

1. In the Study Factors area of the Input File Characterization page, choose **Add > Categorical Factor**.

   The categorical study factor editor opens in the Study Factors area with the [new factor] box selected.



2. Type the factor name: **Solvent**.
3. If the editor closes before you type o-ring click **Edit** to reopen it. Select **[new factor]** and type **Solvent**.



4. Add the solvents types to the Items list by entering text in the box next to the Add button:

   Note When the file names include study factors that are separated by the selected delimiters, the application automatically parses the file names and enters the study factor values in the box next to the Add button as you type.

   a. Begin typing H2O and click **Add**.
   b. Begin typing 50EtOH and click **Add**.
   c. Begin typing 100EtOH and click **Add**.

5. Click **Apply**.

❖ To add a categorical study factor for the o-ring color

1. In the Study Factors area of the Input File Characterization page, choose **Add > Categorical Factor**.
2. Type the factor name: **O-ring**.
3. Add the o-ring colors to the Items list by entering text in the box next to the Add button:
   a. Begin typing Red and click **Add**.
   b. Begin typing Brown and click **Add**.
4. Click **Apply**.

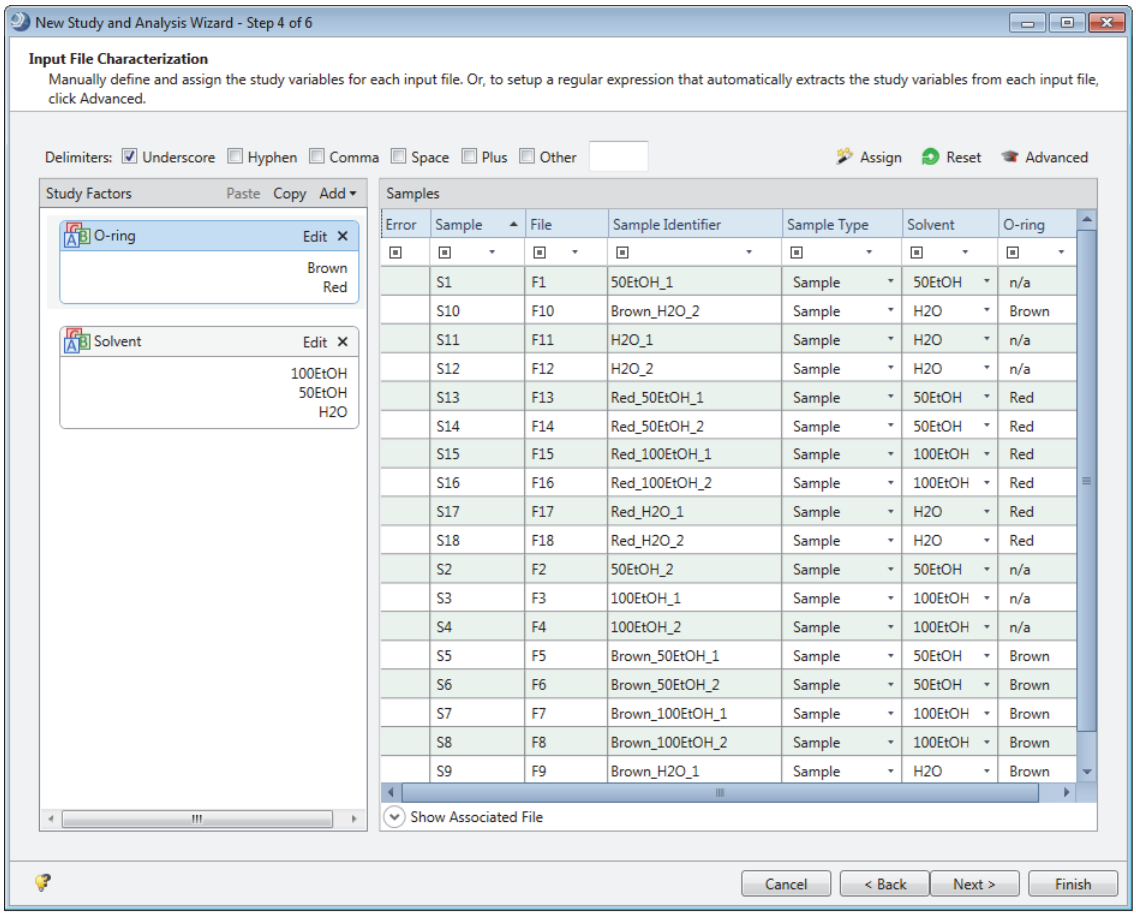
   The editor closes and the Edit button replaces the Apply and Cancel buttons.

❖ To assign the study factor values to the samples

In the command bar next to the delimiters, click **Assign**.

The application assigns the study factor items to the samples.

This figure shows the study factor assignments.



Manually Selecting the Blank Sample Type

Because the file names for these samples do not include the text *Blank*, you must manually select the sample type.

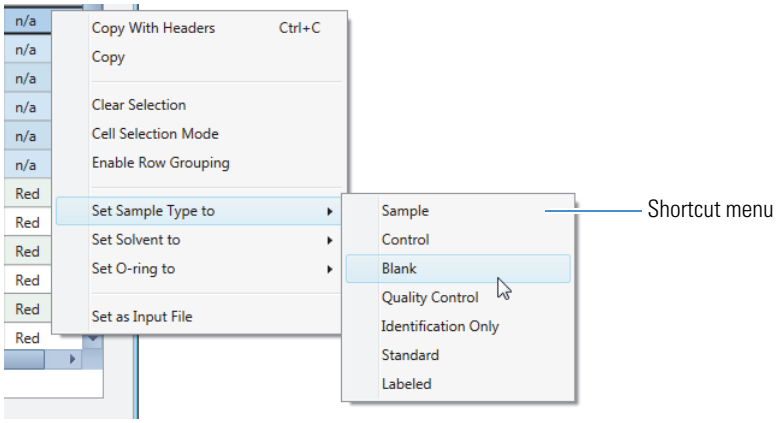
❖ To select the sample type for the solvent blanks

1. To sort the samples by the O-ring column, click the O-ring column header.
2. To select the solvent blanks, use the SHIFT key and select the samples with the n/a assignment in the O-ring column.

> **Tip** To select a table row, click a cell in the row that does not have a drop-down list.

3. Right-click the selection and choose **Set Sample Type To > Blank**.

   This figure shows the selected sample rows that are highlighted in blue and the shortcut menu.



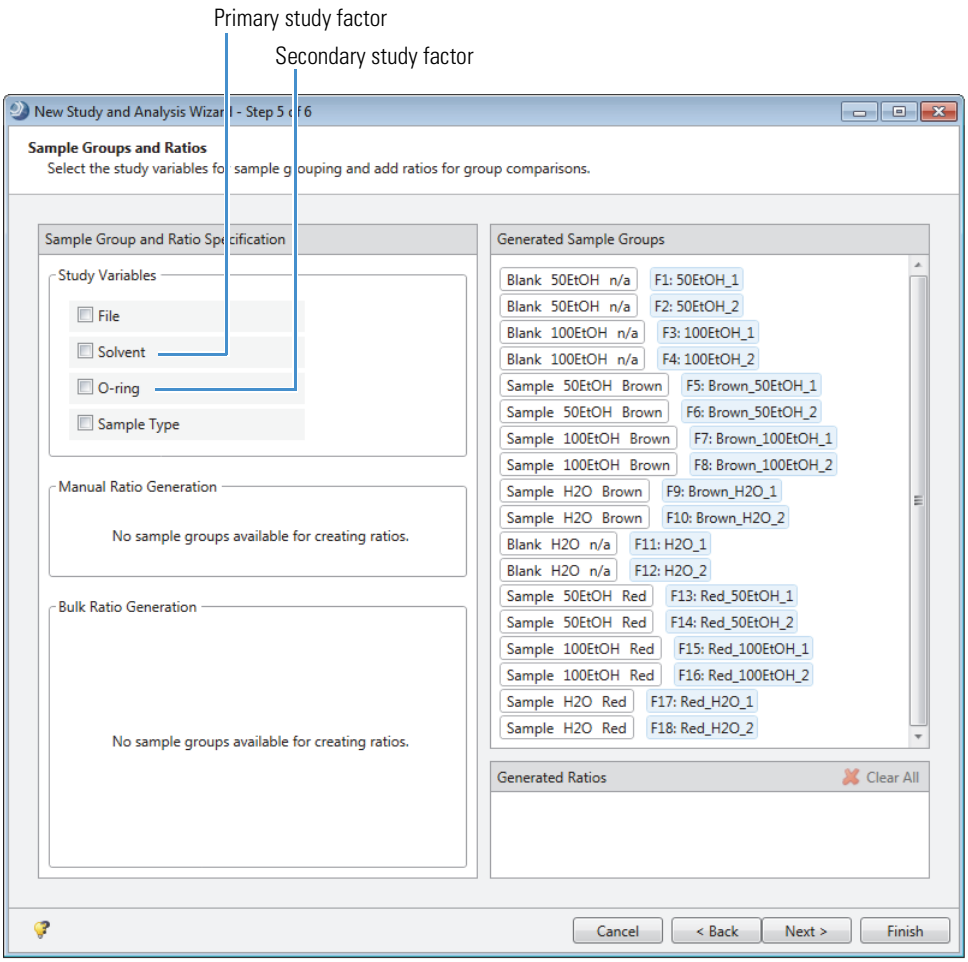
Setting Up the Sample Groups and Ratios

Click **Next** to open the Sample Groups and Ratios page of the wizard.

Use the Sample Groups and Ratios page of the wizard to set up three differential analyses that compare the compounds extracted from the two o-ring types by the three solvents:

- Red versus Brown in water
- Red versus Brown in 50% ethanol
- Red versus Brown in 100% ethanol

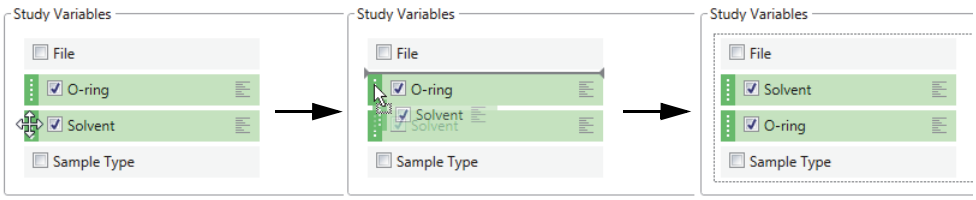
This figure shows the study variables on the left and the ungrouped samples on the right. If you set up the study factors as described in "Defining and Assigning the Study Factors" on page 6, solvent is the primary study factor, and o-ring is the secondary study factor.



❖ To make Solvent the primary study factor

Note If Solvent is not the primary study factor, change its hierarchy.

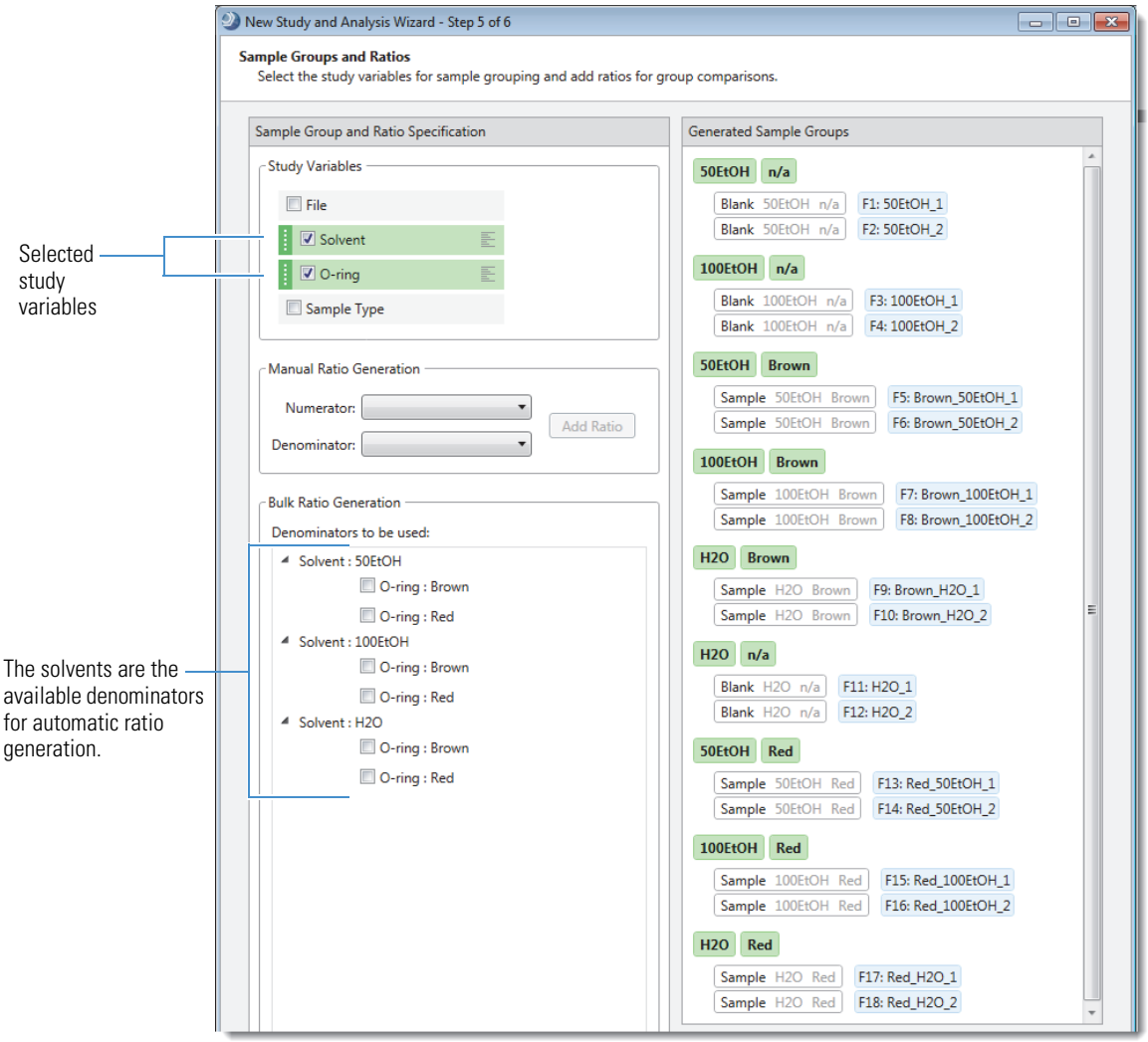
1. Place the pointer over the handle ( ) to the left of **Solvent**.
2. When the move cursor ( ) appears, drag the variable up in the list.



❖ To set up the sample groups

In the Study Variables area, select the **Solvent** and **O-ring** check boxes.

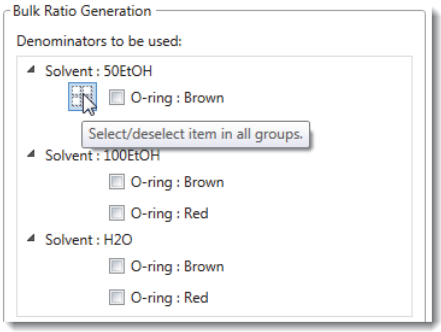
The application groups the samples by solvent, o-ring color, and solvent blanks (n/a). Nine sample groups with two samples each appear in the Generated Sample Groups area. Three groups, one for each solvent, appear in the Bulk Ratio Generation area.



❖ To set up the ratios with the Bulk Ratio Generator

1. In the Bulk Ratio Generation area, select the brown O-ring as the denominator for each ratio as follows:
   a. Place the pointer to the left of any of the O-ring color: Brown check boxes.

      The Select/Deselect Item in All Groups icon,  , appears.



   b. Click the **Select Item in All Groups** icon,  .

2. Click **Add Ratios**.

   The three ratios, one for each extraction solvent, appear in the Generated Ratios area.

   Figure 2. Ratios for a differential analysis

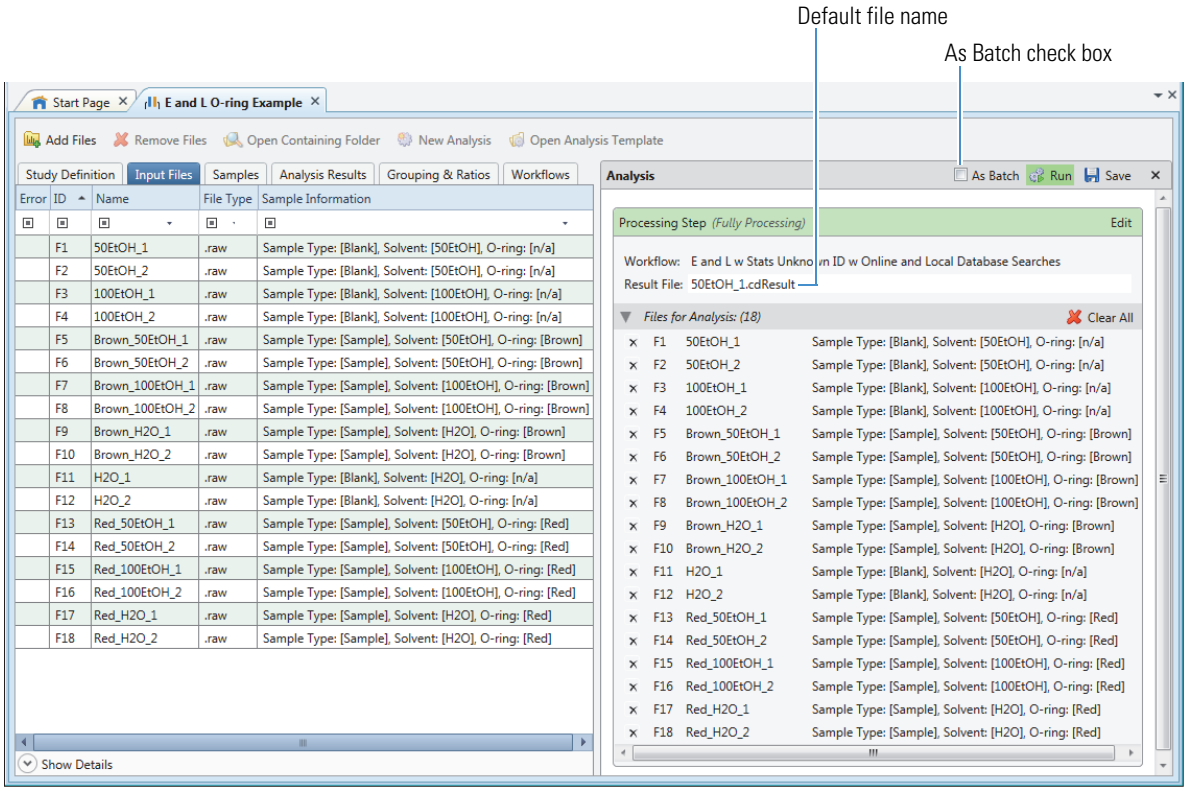
Closing the Wizard

❖ **To close the wizard and save the study**

Click **Finish**.

The E and L O-ring Example tab, the two analysis page tabs (Grouping & Ratios and Workflows), and the Analysis pane appear.

The Analysis pane lists the 18 input files. The As Batch check box is not selected, so the analysis will generate one result file. The Workflow box contains the name of the selected processing workflow and the Result File box contains the file name of the first file in the Files for Analysis list.



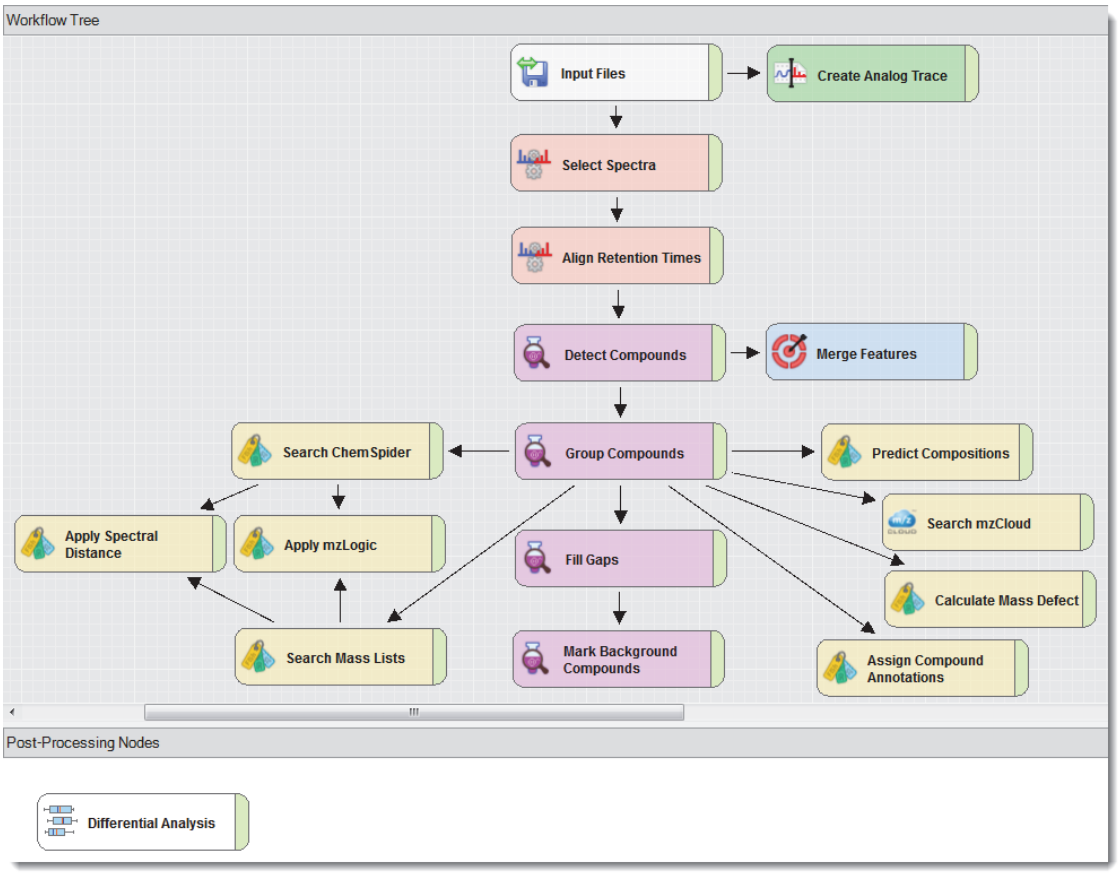
Reviewing the Processing Workflow

Before submitting the analysis to the job queue, review the processing workflow and make changes as needed. If your processing computer does not have Internet access, remove the Search mzCloud node and the Search ChemSpider node (see "To remove the search nodes that require Internet access" on page 15).

❖ To review the processing workflow

1. Click the **Workflows** tab to open the Workflows page.

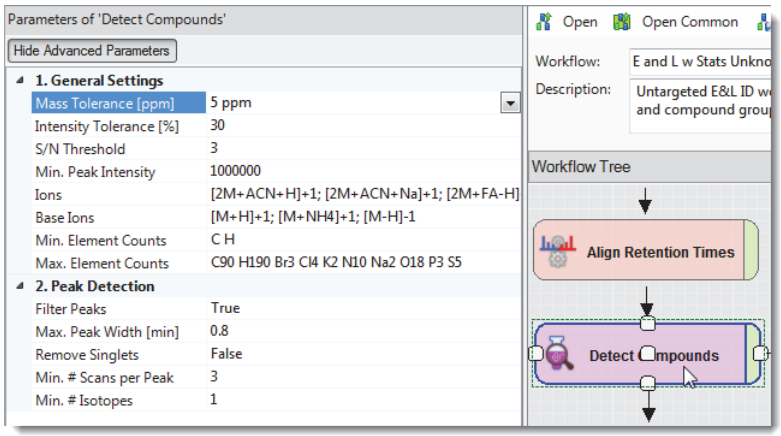
   Figure 3. Processing workflow for a metabolomics study



   Review the settings for the following workflow nodes, but do not change the settings for this tutorial: Detect Compounds, Group Compounds, Search ChemSpider, Search mzCloud, Calculate Mass Defect, and Search Mass Lists.

   To check the settings for a node, select the node in the Workflow Tree area and review its parameter settings on the Parameters page to the left.

2. Select the **Detect Compounds** node and check the parameter settings.



For the Detect Compounds node, do the following:

- Check the Min. Peak Intensity setting against the suggested setting for your data set.

  The minimum peak intensity setting defines the base peak intensity for the unknown compound detection. For this tutorial, keep the setting of 1 000 000.

  Table 1 lists the recommended range for the minimum peak intensity parameter. The optimal setting depends on the sensitivity of the mass spectrometer.

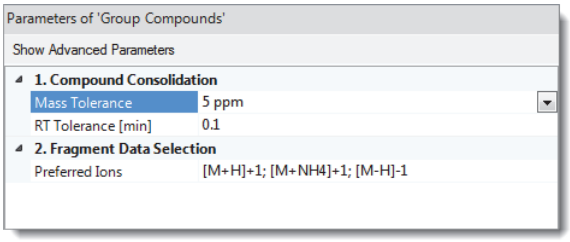
  Table 1. Recommended minimum peak intensity range

Mass spectrometer	Minimum peak intensity (chromatographic peak height)
Q Exactive™, Q Exactive Plus™, Q Exactive HF	500 000 to 1 000 000
Exactive™, Exactive Plus™, Orbitrap Elite™, Orbitrap Velos Pro™	100 000 to 500 000
Orbitrap Fusion™, Orbitrap Lumos, Orbitrap ID-X	50 000 to 100 000
LTQ Orbitrap XL™, LTQ Orbitrap Velos™	25 000 to 100 000

- Check the Base Ions list. For this tutorial, do not change the selection.

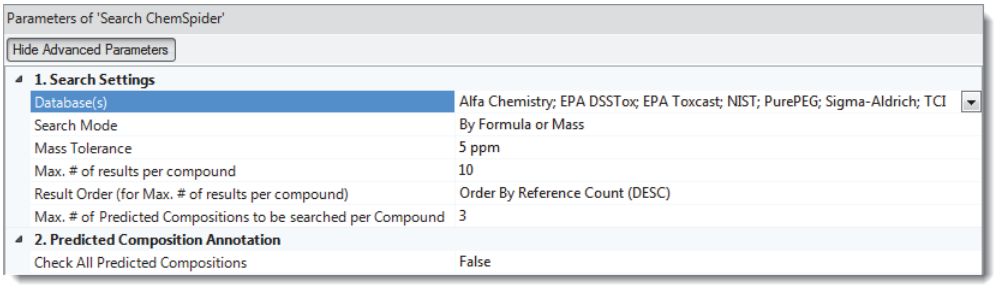
> **Tip** Mobile phase additives can have a significant effect on the base ions (adduct ions with the highest intensity) in the full scan data for an LC/MS experiment. To ensure that the application interprets the isotopic ion clusters correctly, select the predominant adduct ions for the chromatographic analysis from the Base Ions list. For example, if the mobile phase contains a significant amount of ammonium acetate, consider adding the ammonium adduct, [M+NH4]+1, to the list. All the processing workflows in the E and L folder include [M+NH4]+1 in the Base Ions list.

3. Select the **Group Compounds** node and review the parameter settings. Check the Preferred Ions list. The list should include the base ions selected in the Detect Compounds node. Do not change the settings for this tutorial.



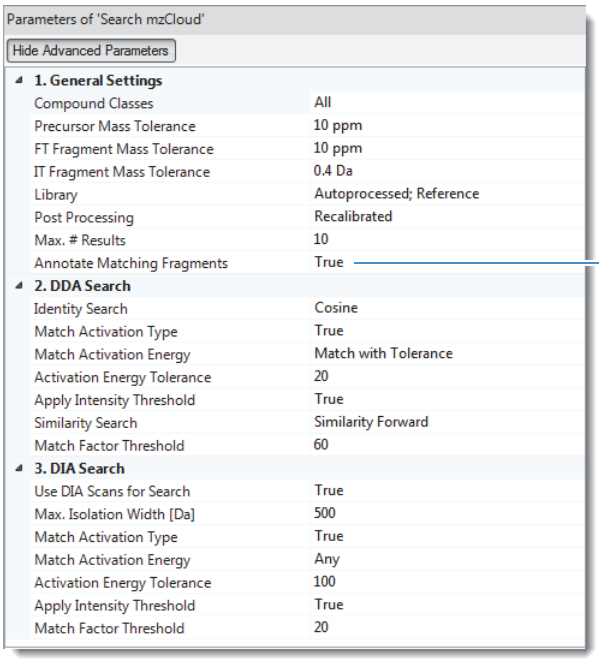
4. Select the **Search ChemSpider** node and review its parameter settings. For other analyses, select the databases that you want the analysis to search and the maximum number of results per compound to return.

   In the selected processing workflow template, the Search ChemSpider node searches 7 out of 363 databases. Do not change the settings for this tutorial.



5. Select the **Search mzCloud** node and review its parameter settings.

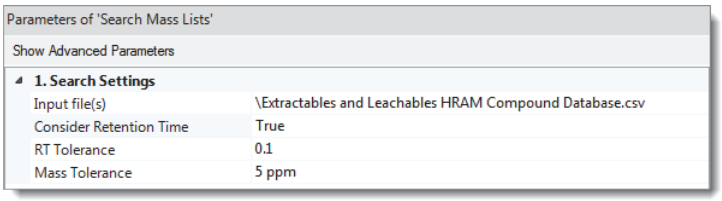
   With the current settings, the node runs an identity search and a similarity search. For this tutorial, do not change the settings.



This feature is enabled (set to True) in all the processing workflow templates provided with the application.

6. Select the **Search Mass Lists** node and review its parameter settings.

   For this tutorial, do not change the settings.

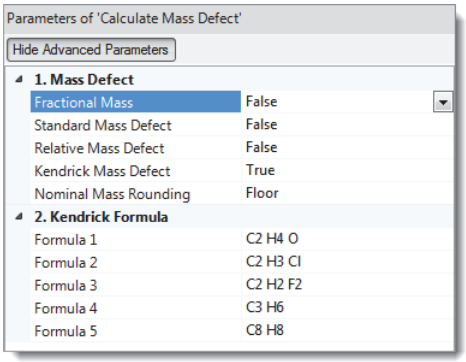


7. Select the **Calculate Mass Defect** node and review its parameter settings.

   The node is set to calculate the mass defect for polymers with the following repeating units:

   - –[C2H4O]– polyethylene glycol
   - –[C2H3Cl]–
   - –[C2H2F2]–
   - –[C3H6]– propyl repeating units
   - –[C8H8]– styrene repeating units

   For this tutorial, do not change the settings.



❖ **To remove the search nodes that require Internet access**

> **Note** Only follow this procedure if your computer does not have Internet access.

1. If the Workflows page is closed, open it.
2. In the Workflow Tree area (Figure 3 on page 12), do the following:
   - Right-click the Search mzCloud node and choose **Cut**.
   - Right-click the Search ChemSpider node and choose **Cut**.

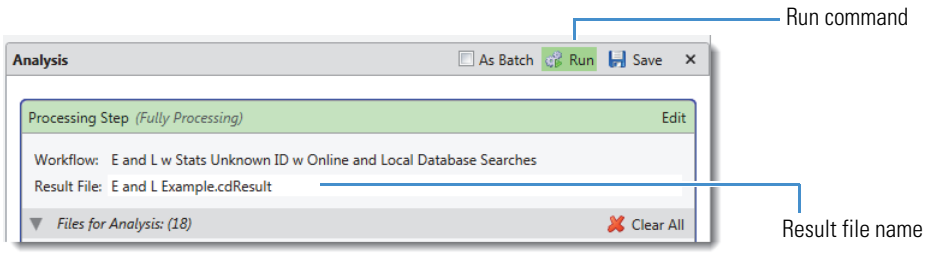
Submitting the Analysis to the Job Queue

In the E and L w Stats and Unknown ID w Online and Local Database Searches processing workflow, none of the workflow nodes require customization.

❖ **To submit the analysis to the job queue**

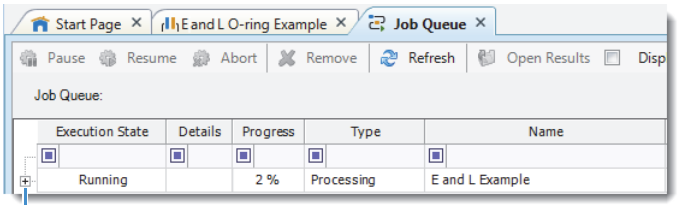
1. To create one result file for the input file set, leave the **As Batch** check box clear.

   By default, the application uses the name of the first input file as the result file name.
2. In the Result File box, rename the result file **E and L Example**.



3. Make sure that the As Batch check box is not selected, and click **Run** to submit the analysis to the job queue.

   The Job Queue page opens.
4. To view the processing messages, click the expand icon, ⊞, to the left of the job row.



> **Note** During the run, the analysis generates several warning messages, which you can ignore. Warning messages have a yellow background.

Leave the Job Queue page open and go to "Reviewing the Analysis Results."

Reviewing the Analysis Results

Follow these topics to review the analysis results:

- Opening the Result File
- Default Layout for a Result Page
- Working with the Result Tables
- Working with the Chromatograms View
- Working with the Mass Spectrum View
- Reviewing the Mass List Search Results
- Reviewing the mzLogic Analysis for a Compound
- Viewing a Mass Defect Plot
- Applying the Statistics Layout
- Working with the Volcano Plot for the Differential Analysis
- Viewing a Hierarchical Cluster Analysis
- Viewing the Principal Component Analysis
- Working with the Partial Least Squares Discriminate Analysis View
- Making Structure Proposals
- Using the Result Filters View

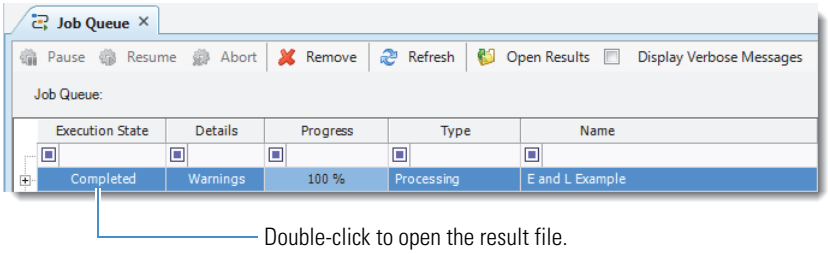
Opening the Result File

You can open a result file from multiple locations: the Job Queue page, the Analysis Results page of a study, the Compound Discoverer Start Page, or the menu bar.

> **Tip** If you did not reprocess the example data set, open the result file provided on the Compound Discoverer 3.1 USB key—**Example Studies\E & L O-ring Study\EandL Example.cdResult**. Then, from the menu bar, choose **Window > Reset Layout**.

❖ **To open the result file from the Job Queue page**

1. If you closed the Job Queue page, from the application menu bar, choose **View > Job Queue**.
2. On the Job Queue page, double-click the run.



Default Layout for a Result Page

The factory default layout for a result file includes the following items:

- A page tab with the result file name.
- A Chromatograms view on the top left that is populated with XIC traces for the compound in the first row of the Compounds table. The view automatically zooms in to the start and end points of the chromatographic peak for the compound.
- A Mass Spectrum view on the top right that is populated with the MS1 scan (for a preferred ion) that is closest to the chromatographic peak apex for the selected compound across the input files. The spectrum tree to the left includes the MS1 scans and the fragmentation scans for the preferred ions that were acquired within the following retention time window:
  - Peak apex (RT) ± the peak's full width at half maximum (FWHM)

  –or–
  - Start and end points of the chromatographic peak, as determined by the peak detection algorithm

Note If the data set does not include data-dependent MS2 scans within the retention time window but does include DIA scans within this window, the spectrum tree includes the DIA scans.

DIA scans include all-ion fragmentation scans (AIF) and data-independent scans (DIA) from your Thermo Scientific mass spectrometer.

- A set of tabbed main tables below the two graphical views.
- A collapsed area for the related tables below the main tables.

Figure 4 shows the factory default layout for the E and L Example.cdResult file.

Figure 4. Factory default layout for the example result file

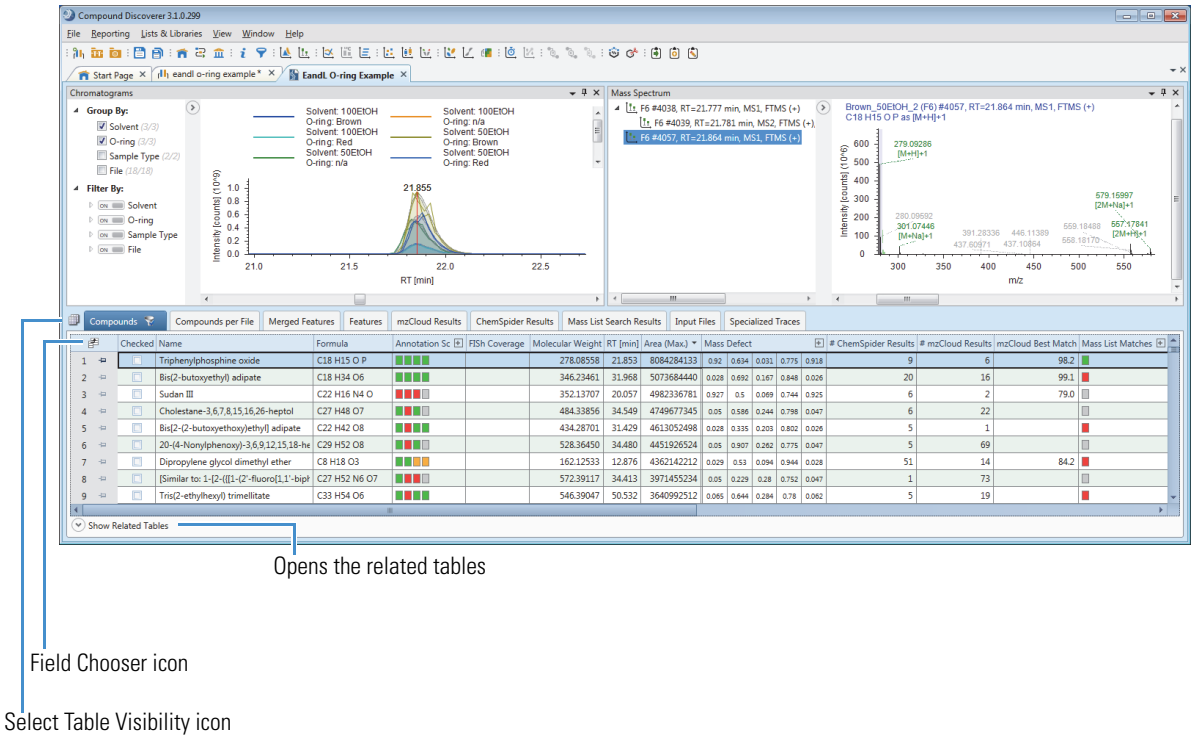


Table 2 describes the main result tables that the selected processing workflow produces. The Compounds table is the active table and is sorted by the Area (Max.) column. The first row displays the compound with the largest chromatographic peak area (found in one of the sample files). Because the selected processing workflow includes the Mark Background Compounds node, the Compounds tab has a filter icon with a check mark ( ). The compounds that the analysis identified as background compounds are marked as background compounds in both the blank and non-blank samples and are hidden from the table.

Table 2. Main tables and some of the related tables for the selected processing workflow (Sheet 1 of 2)

Result table	Description
Visible main (top-level) tables	
Compounds	Displays all the compounds that the analysis detected grouped by their molecular weight and retention time (MW×RT) dimensions across all the sample files.
Compounds per File	Displays all the compounds that the analysis detected across all the sample files on a per file basis. Does not list compounds that the Fill Gaps node detected by filling a full gap.
Merged Features	Displays all the features (ions with the same *m/z* ×RT dimensions) across the sample files in descending order by the Max. Area column.
Features	Displays all the features (ions with the same *m/z* ×RT dimensions) that the analysis detected across all the sample files on a per file basis in descending order by the Area column.

Table 2. Main tables and some of the related tables for the selected processing workflow (Sheet 2 of 2)

Result table	Description
mzCloud Results	Lists the mzCloud search results across all the sample files.
ChemSpider Results	Lists the ChemSpider search results across all the sample files.
Mass List Search Results	Lists the mass list search results across all the sample files in descending order by molecular weight.
Input Files	Describes the input files that the application processed to create the result file.
Specialized Traces	Lists the specialized traces. For this analysis, lists the analog traces that the Create Analog Trace node extracted across the sample files.
Visible tables related to the Compounds table	
Structure Proposals	Displays your structure proposals for the selected compound selected in the Compounds table. Initially, this table is empty.
Compounds per File	Displays information about the selected compound on a per file basis.
Predicted Compositions	Displays the predicted compositions for the selected compound. The predicted compositions are based on the molecular weight of the neutral compound.
Merged Features	Displays the features detected across the sample files for the selected compound.
mzCloud Results	Displays the mzCloud results for the selected compound.
ChemSpider Results	Displays the ChemSpider results for the selected compound.
Mass List Search Results	Displays the mass list search results for the selected compound.
Visible tables related to the Features table	
Compound per File	Displays the compound detected for the selected feature.
Chromatogram Peaks	Describes the chromatographic peaks for the selected feature.
Visible tables related to the Input Files table	
File Alignments	Describes the alignment for the input file selected in the Input File table.
Hidden tables	
Adducts (main table)	Lists the adducts in the Adducts library.
Filled Gaps (related table)	Provides information about the imputed chromatographic peak areas (when the processing workflow includes the Fill Gaps node). Imputed peak areas improve the statistical analysis results.

Working with the Result Tables

The following topics describe how to modify the layout of the result tables:

- Showing or Hiding Result Tables and Viewing the Related Tables
- Showing or Hiding Table Columns
- Sorting Table Rows

Showing or Hiding Result Tables and Viewing the Related Tables

The default layout does not include all the results tables that the selected processing workflow produces. By default, the following tables are hidden: the Filled Gaps related table, which is created by the Fill Gaps node, and the Adducts table.

❖ **To show or hide result tables**

1. Click the **Select Table Visibility** icon,  , to the left of the result table tabs.

   The Select Visible Tables dialog box opens.
2. To display a table, select its check box. To hide a table, clear its check box.

   For this tutorial, do not change the selections.
3. Click **OK**.

❖ To view the tables related to the Compounds table

With the main Compounds table selected as the active table, click **Show Related Tables** at the bottom of the window to display the tables related to the Compounds table.

Showing or Hiding Table Columns

❖ To display or hide a table column

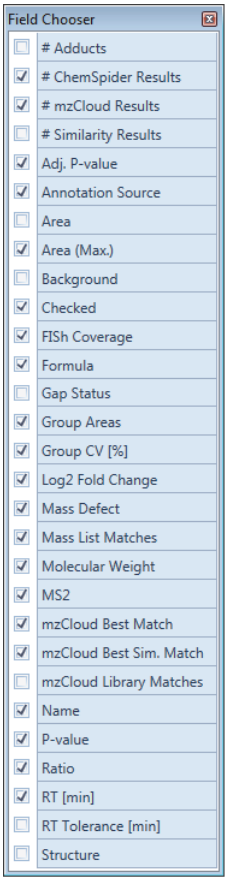
1. Click the **Field Chooser** icon, , to open the Field Chooser box.

   In the Compounds table, the following columns are hidden, by default:

   - #Adducts
   - #Similarity Results
   - Area (for compound in each sample file)
   - Background
   - Gap Status
   - RT Tolerance [min]
   - Structure

2. Do one of the following:

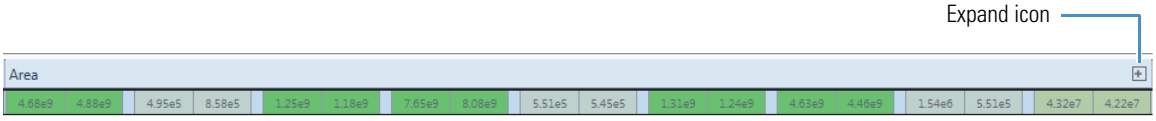
   - To display a column, select its check box.
   - To hide a column, clear its check box.



❖ To show the Area column in the Compounds table

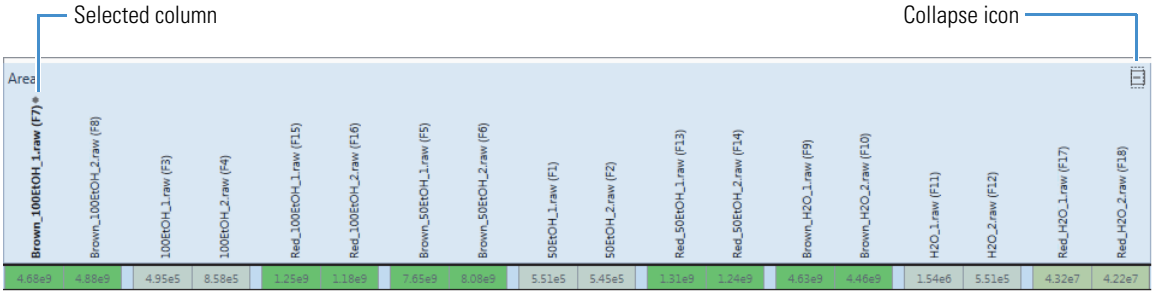
1. Click the **Field Chooser** icon, .
2. Select the **Area** check box.

   The Area column appears in the Compounds table to the left of the Group Areas column. The Area column contains 18 subcolumns, one for each input file (samples and blanks).



3. To close the Field Chooser box, click its close icon, .
4. To view the sample names for the area columns, click the expand icon, , to the right of the column name.

   The names of the input files appear vertically above the individual area columns. The name of the first input file, 50EtOH_1.raw (F1), appears in bold font and has an asterisk indicating that it is the selected area column.



5. Click the collapse icon, , to hide the sample names.
6. To hide the Area column, open the Field Chooser box and clear the **Area** check box.

Sorting Table Rows

❖ To sort the table rows based on the contents of the table columns

1. Click a column header to sort the rows between ascending order (A, B, C …) and descending order (Z, Y, X …), based on the contents of the column.

   Note The application treats formulas the same as text strings and sorts them by the order of the characters in the formula string, not by the actual number of elements in the formula.

2. To sort the data by a second column, hold down the CTRL key and click the second column heading.

❖ To sort the table rows by a column that contains a distribution map

1. Click the expand icon to display the vertical headings of the subordinate columns.
2. Select the heading of the subordinate column that you want to sort by.

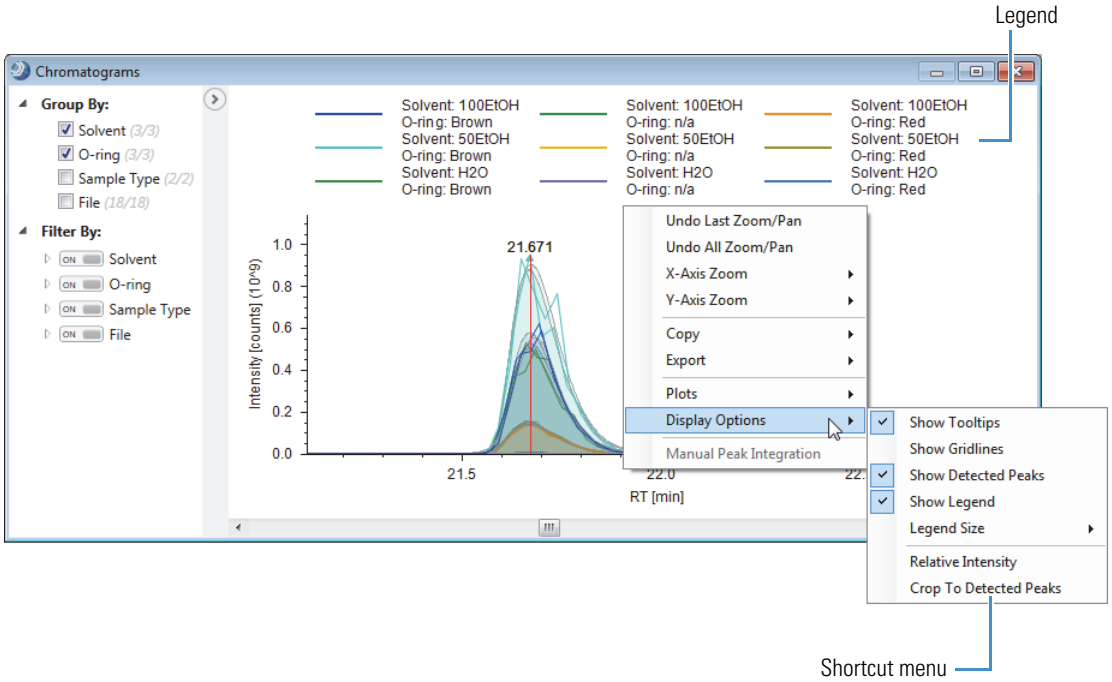
   The selected subordinate column heading appears in bold text.
3. Click the column heading to sort the table rows.

❖ To view a mass chromatogram

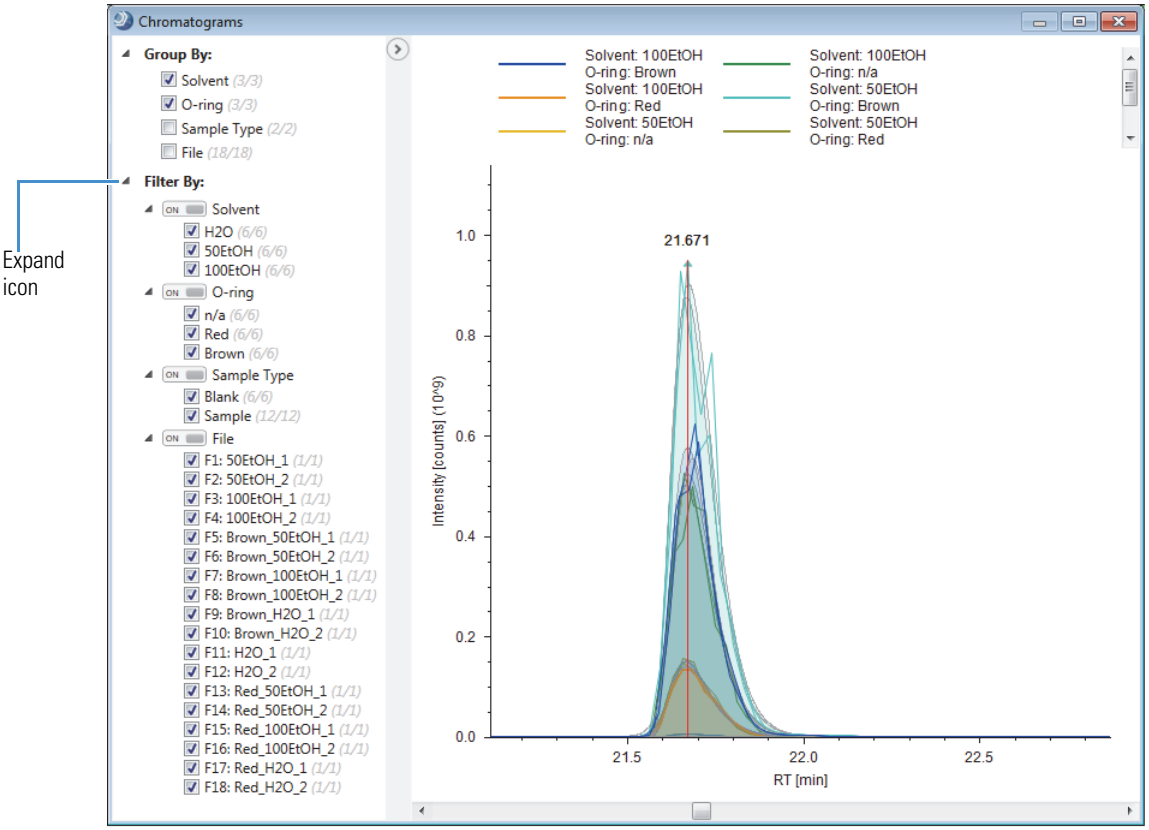
1. In the Compounds table, click the **Area (Max.)** column heading to sort the table by Area (Max.) in descending order (▼).
2. To display the overlaid traces for the compound with the largest peak area, select **row 1 (TriPhenylphosphine oxide)** (Figure 4 on page 17).

   By default the display options for the Chromatograms view are set to Show ToolTips, Show Detected Peaks, Show Legend, and Zoom to Detected Peaks. The *x*-axis zoom is set to the width of the detected peaks, and the *y*-axis zoom is set to auto scale.

   This figure displays the overlaid traces for the input files that contain the selected compound and the legend and shortcut menu for the Chromatograms view. (The Mass Spectrum view is closed.)



3. Under Filter By, expand the all the filters by clicking their expand icons, ◢.



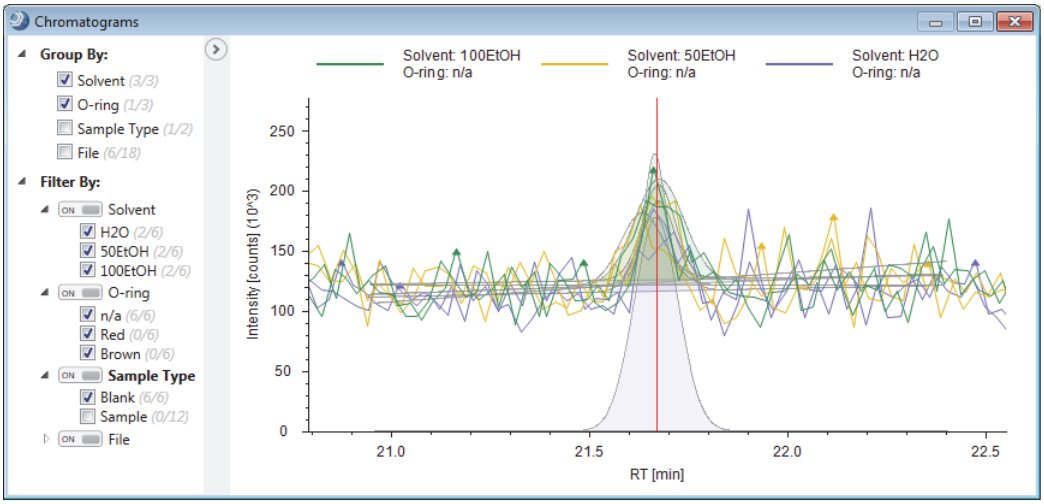
Expand icon

4. Review the information in the collapsible left pane:
   - Under Group By, the counts for the O-ring and Solvent study factors are 3 out of 3 (3/3), the count for Sample Type is 2 out of 2 (2/2), and the count for File is 18 out of 18 (18/18). These counts mean that the analysis detected the selected compound in all the input files.

     Tip Because the processing workflow included the Fill Gaps node, the analysis detected the selected compound in all the input files, including the solvent blanks.

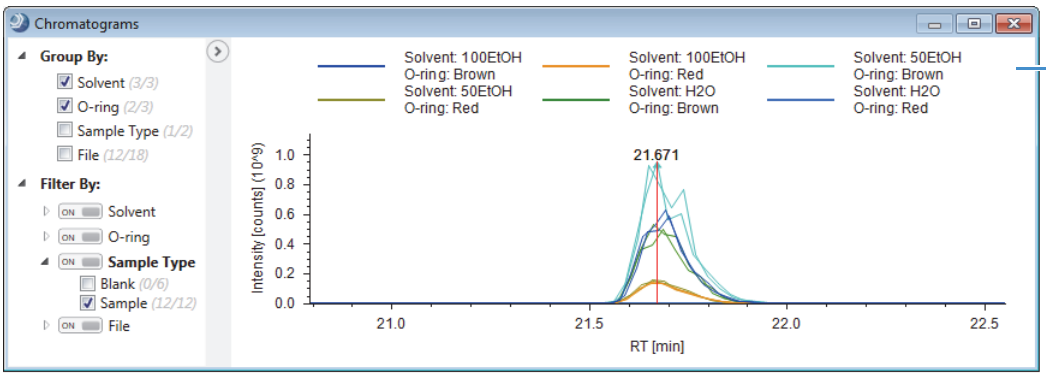
   - Under Filter By, all the check boxes are selected. This means that the XIC traces for all the files are displayed in the Chromatograms view.
5. To view the traces for the solvent blanks, under Filter By > Sample Type, clear the **Sample** check box.

   This figure shows the traces for the solvent blanks.



6. To view the chromatographic traces without the shading that can mask underlying peaks, right-click the Chromatograms view and choose **Display Options > Show Detected Peaks**.
7. To view the traces for the o-ring samples, under Filter By > Sample Type, select the **Sample** check box and clear the **Blank** check box.

   This figure shows the overlaid traces for the 12 samples that you are comparing. The application displays the six sample groups in different colors. The legend shows the display colors for the groups.



Working with the Mass Spectrum View

Follow these procedures to review the mass spectra for a compound.

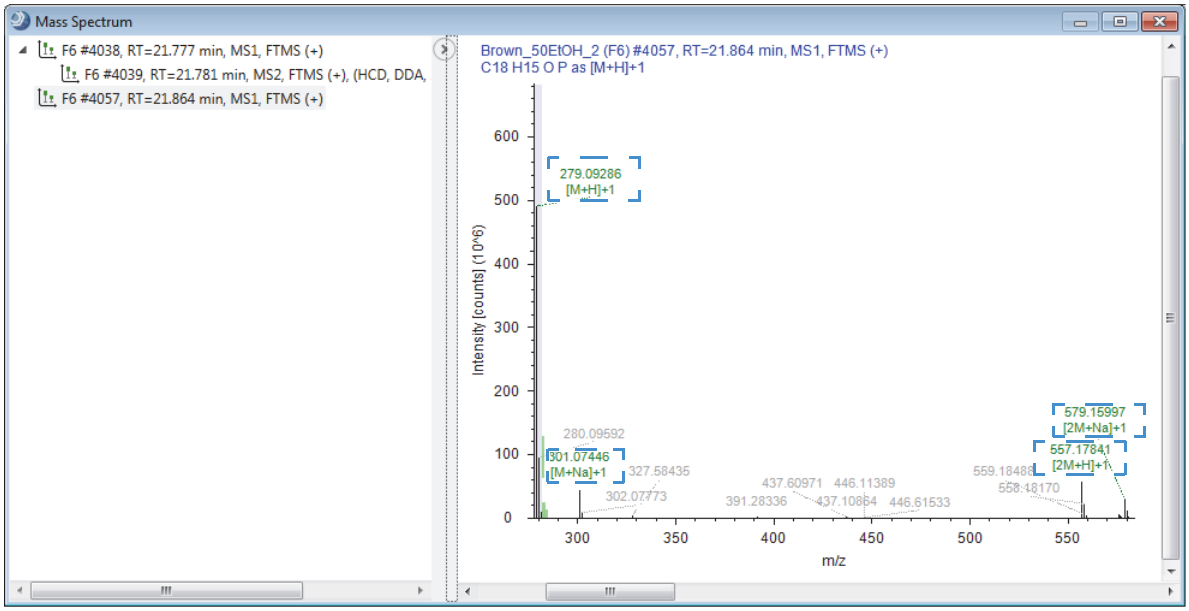
- To view the MS1 spectrum
- To inspect the isotopic cluster for the most intense adduct ion
- To view the results of an mzCloud search for a compound and display a mirror plot

❖ To view the MS1 spectrum

1. Sort the Compounds table by the **Area (Max.)** column in descending order (▼).
2. Select **row 1**: **Triphenylphosphine oxide** (Figure 4 on page 17).
3. Close the Chromatograms view by clicking the **Close** icon on its title bar.
4. To expand the Mass Spectrum view to the full screen width, drag the view by its title bar to a second monitor.

   This figure shows the Mass Spectrum view as a floating window. The spectrum tree on the left lists the MS1 scans and the fragmentation scans for the preferred ions that are within the search window.

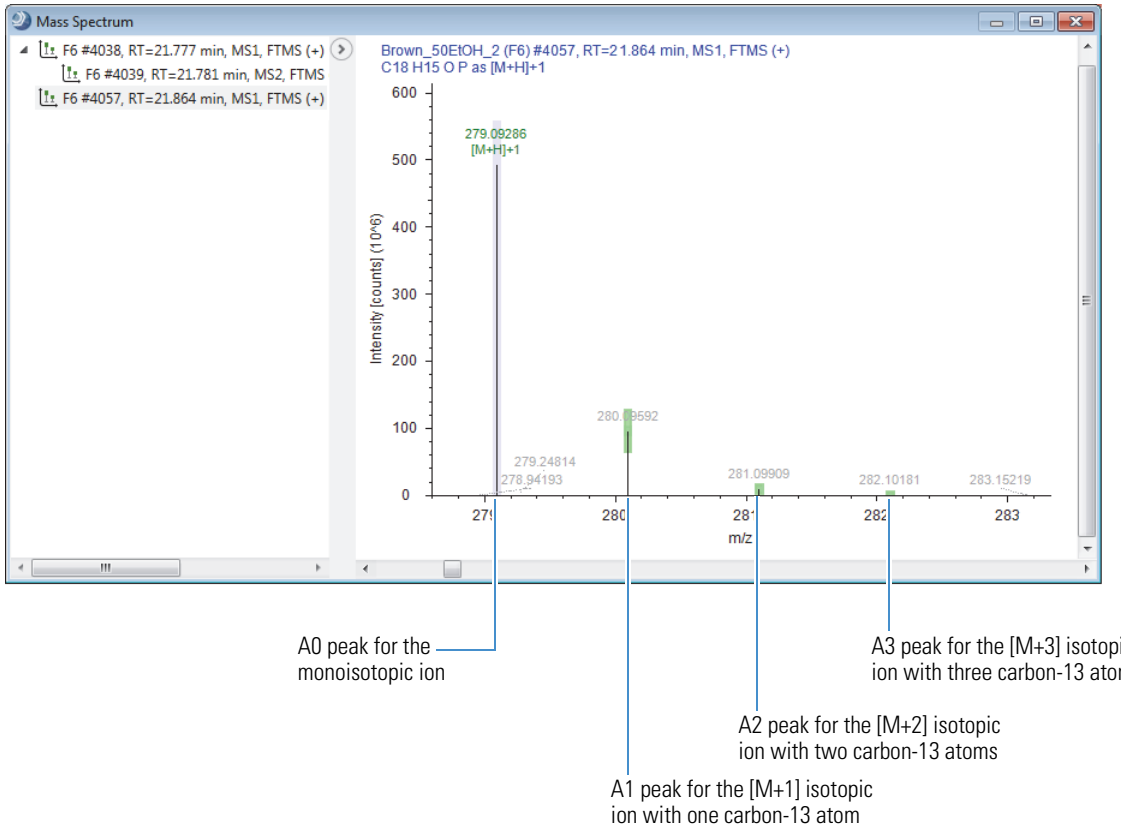
   The *x* axis expands to display the detected adduct ions: [M+H]+1, [2M+H]+1, [M+Na}+1, and [2M+Na]+1. The analysis detected these adduct ions across the input files. In input file F6, the [2M + Na]+1 ion is from a gap-filled chromatographic peak.



❖ To inspect the isotopic cluster for the most intense adduct ion

1. In the Compounds table, select **row 1**: **Triphenylphosphine oxide** (Figure 4 on page 17).
2. To zoom in on the isotopic cluster for the [M+1]+1 adduct ion, drag the pointer horizontally across the mass spectrum from *m/z* 279 to *m/z* 282.

   Because the selected compound has a predicted formula, the MS1 spectrum includes color-coded centroids.



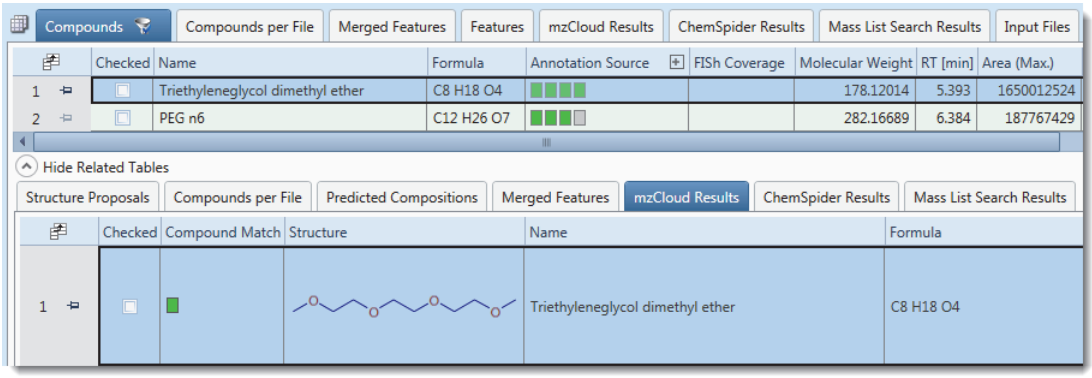
   The centroids in the MS1 scan are color-coded as follows:

   - Lavender bars indicate the A0 peak for the monoisotopic ion. The *x*-axis position and the width of the bar reflect the expected *m/z* value of the centroid and the user-specified mass tolerance, respectively.
   - Green rectangles indicate matching centroids for isotopic ions. When you zoom in on the matching centroid, the *x*-axis position and width of the rectangle reflect the expected *m/z* value of the centroid and the user-specified mass tolerance, respectively. The *y*-axis position and height of the rectangle reflect the expected relative intensity of the centroid and the user-specified intensity tolerance, respectively.
   - Red rectangles indicate centroids that are missing from the expected isotopic pattern.
   - Blue rectangles indicate centroids that are missing from the expected isotopic pattern but that are also expected to have an intensity below the measured baseline noise (determined by the Fourier transform mass spectrometry (FTMS) mass analyzer).

❖ To view the results of an mzCloud search for a compound and display a mirror plot

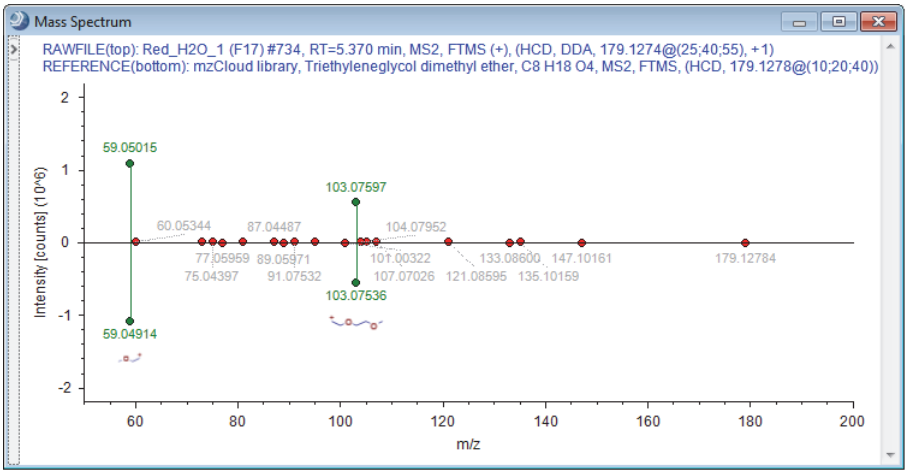
1. Sort the Compounds table by the **mzCloud Best Match** column in descending order (▼).
2. In the Compounds table, select **row 1(Triethyleneglycol dimethyl ether)**.
3. To display the related tables below the main tables, click **Show Related Tables**.

4. Click the **mzCloud Results** tab and select **row 1(Triethyleneglycol dimethyl ether)**.



An annotated mirror plot appears in the Mass Spectrum view with scan #734 from input file F17 on top and the mzCloud reference spectrum on the bottom.

The centroids for the matching fragments are displayed as green sticks with a green circle at the end. Red circles on the *x*-axis indicate the *m/z* values of the missing fragments.

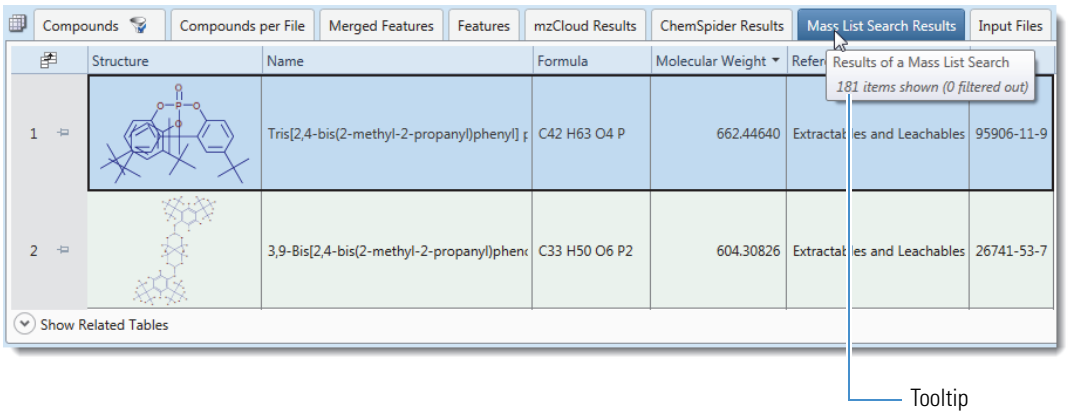


Reviewing the Mass List Search Results

❖ **To review the mass list search results for the analysis**

1. In the main tables, click the **Mass List Search Results** tab.
2. To display a tooltip that lists the number of compounds found across the input files, point to the Mass List Search Results tab.

   This figure shows that the mass list search found 181 matching entries.

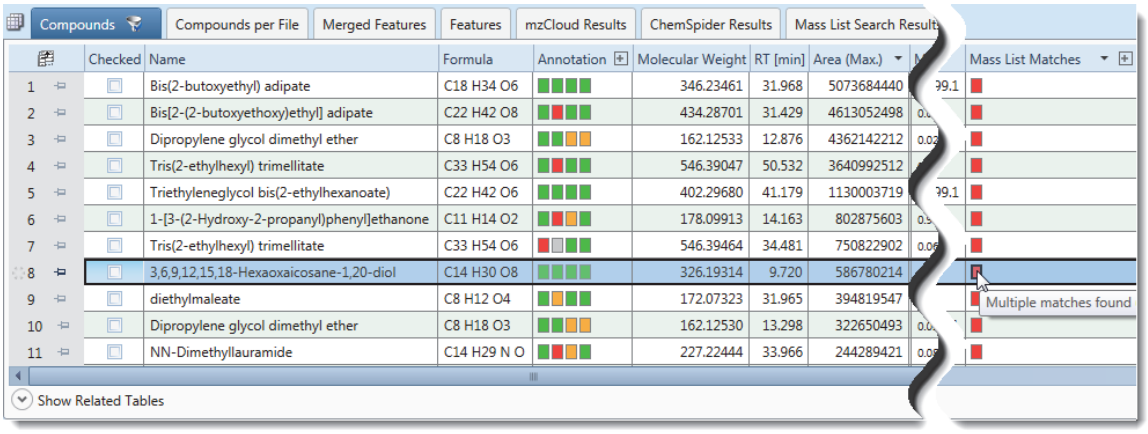


3. In the main tables, click the **Compounds** tab.
4. Sort the Compounds table in descending order (▼) by the **Mass List Matches** column. Then, hold down the CTRL key and sort the table in descending order (▼) by the **Area (Max)** column.

   The color-coded rectangles indicate whether the mass list search found matching compounds.

   - (■) Red—Multiple matches found
   - (■) Green—Single match found
   - (■) Gray—No matches found
5. In the sorted Compounds table, select **row 8** (3,6,7,12,15,18-Hexaoxaicosane-1,20-diol).

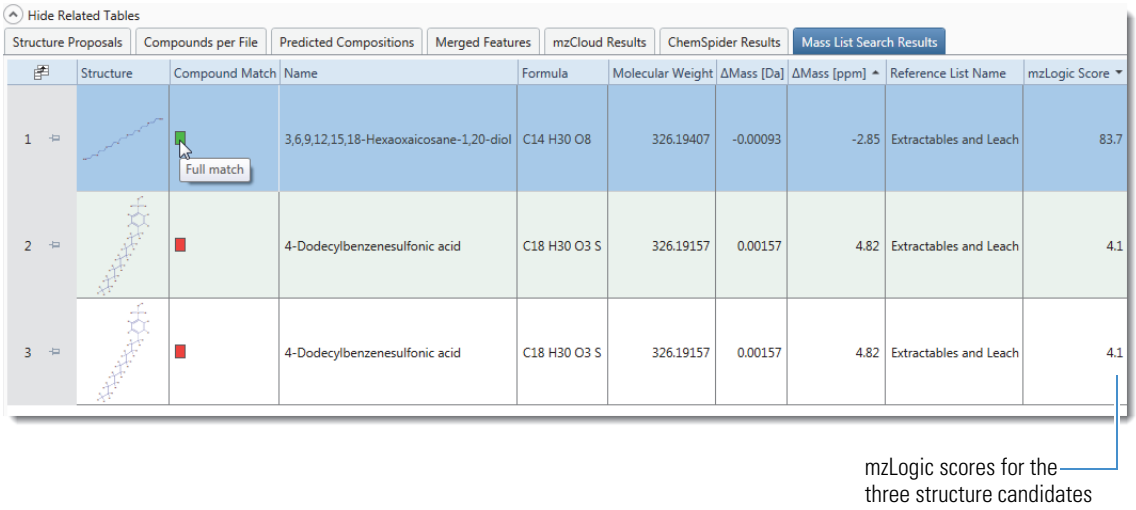
Figure 5. Compounds table sorted by the Mass List Matches column and the Area (Max) column in descending order



6. Open the related tables and click the **Mass List Search Results** tab.
7. Review the mass list search results for the selected compound.

   This figure shows the mass list search results for 3,6,7,12,15,18-Hexaoxaicosane-1,20-diol. In the Mass List Matches column of the Compounds table, the red status (■) indicates that the analysis found multiple matches. The related Mass List Search Results table describes the three matches. The first compound is a full match (■) and has an mzLogic score of 83.7. The formulas of the second and third compounds do not match the formula of the selected compound (■ indicates No match). Each of these compounds has an mzLogic score of 4.1.

Figure 6. Mass list search results for 3,6,7,12,15,18-Hexaoxaicosane-1,20-diol



Reviewing the mzLogic Analysis for a Compound

When an mzCloud search yields no identity matches for an unknown compound, use the mzLogic™ scores for the search hits to determine the best explanation for a compound.

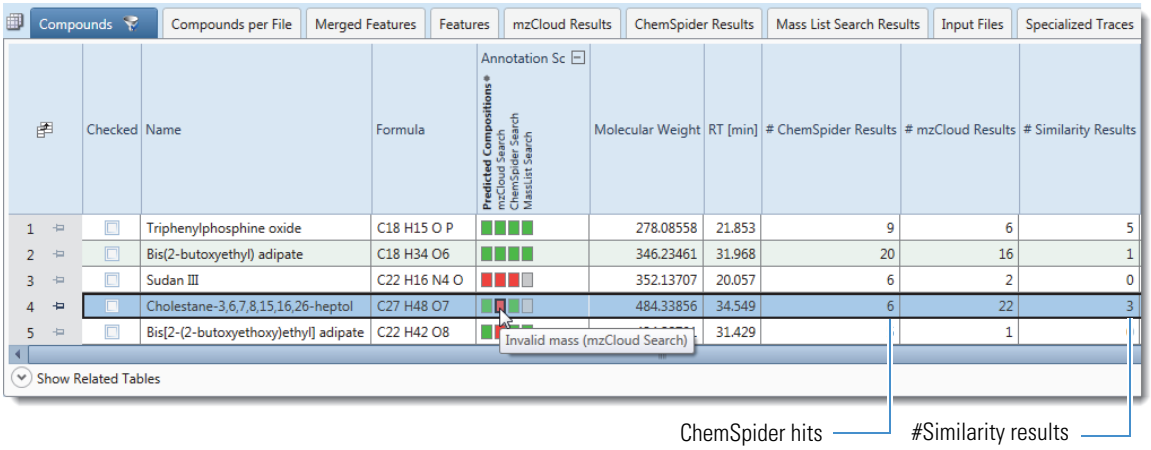
An mzLogic analysis requires candidate structures from the ChemSpider database, a Metabolika pathway, or a mass list with structures. It also requires spectral similarity hits from an mzCloud similarity search. The processing workflow for this E&L tutorial (Figure 3 on page 12) includes two nodes that provide structures—the Search ChemSpider node and the Search Mass Lists node.

❖ To review the results of an mzLogic analysis for structure candidates

1. From the application menu bar, choose **Window > Reset Layout**.
2. Sort the Compounds table in descending order by Area (Max.).
3. Open the Compound table's Field Chooser dialog box (see "To display or hide a table column" on page 19) and select the **#Similarity Results** check box.
4. In the Compounds table, expand the **Annotation Sources** column.

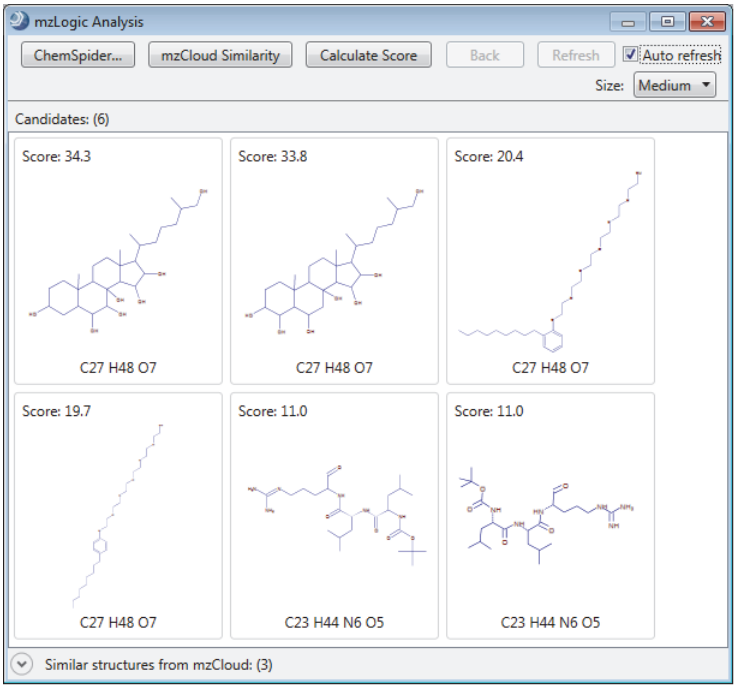
   For cholestane-3,6,7,8,15,16,26-heptol (row 4), the analysis found no matching structures from the mzCloud identity search. But it did find six ChemSpider hits and three structures from the mzCloud similarity search.

   Note The FISh Coverage, Area (Max.), and Mass Defect columns are hidden in the following figure.



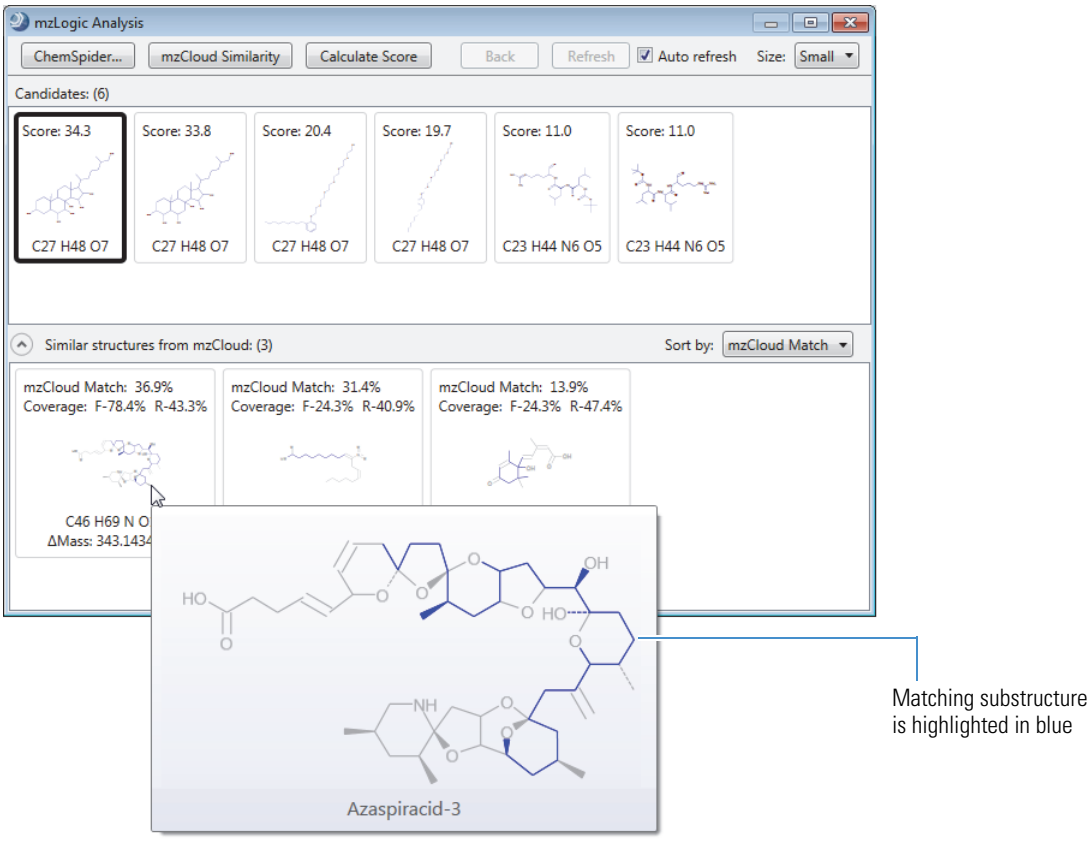
5. In the Compounds table, select **row 4** (Cholestane-3,6,7,8,15,16,26-heptol).
6. From the application menu bar, choose **View > mzLogic Analysis**.

   The mzLogic Analysis view opens to the right of the result tables.



7. Below the Candidates pane, click **Similar Structures from mzCloud**.

   The application automatically selects the highest ranking candidate in the Candidates pane and highlights the matching substructures in blue in the Similar Structures from mzCloud pane. If you select another candidate, the application determines the structural coverage for that candidate.



Viewing a Mass Defect Plot

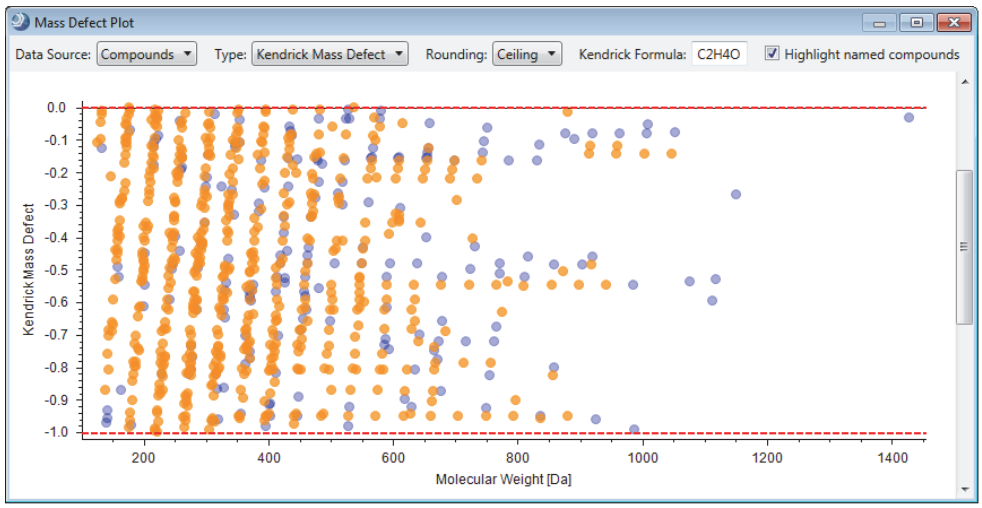
You can view a mass defect plot for the compounds in the Compounds table. Follow this procedure to set up a mass defect plot that highlights polyethylene glycols (PEG).

❖ To set up a mass defect plot for a Kendrick formula

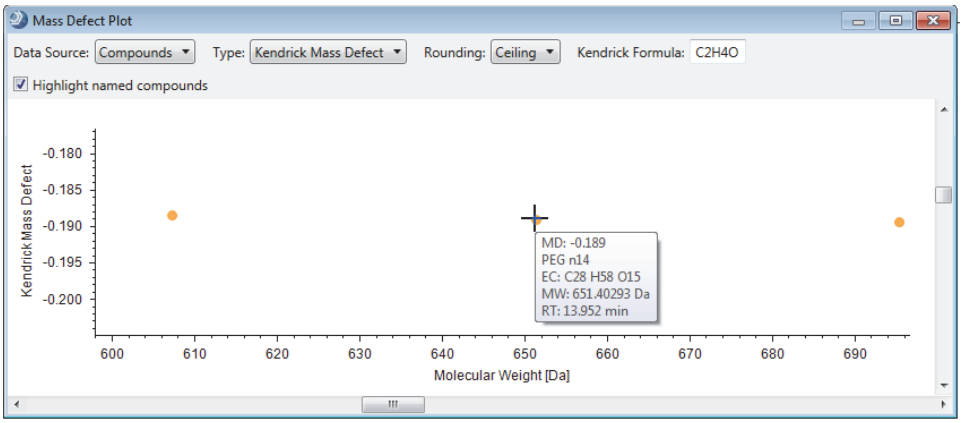
1. From the application menu bar, choose **View > Mass Defect Plot**.

   The Mass Defect Plot opens to the right of the result tables.

2. In the Kendrick Formula box, type C2H4O.
3. Select the **Highlight Named Compounds** check box.



4. Zoom in on named compounds that form a horizontal line.



5. Right-click the plot and choose **Check All Visible Points**.
6. Sort the Compounds table by the Name column in descending order.
7. Notice that the three PEG compounds are checked. Clear these check boxes.

Applying the Statistics Layout

The application comes with the factory default layout and four named layouts: Identification, Quantification, Stable Isotope Labeling, and Statistics. When running statistical analyses, for ease of use, apply the Statistics layout.

❖ To apply the Statistics layout

1. If the result file is not the active page in the application window, click the result file tab.
2. From the menu bar, choose **Window > Reset Layout**, and then choose **Window > Apply Layout > Statistics**.

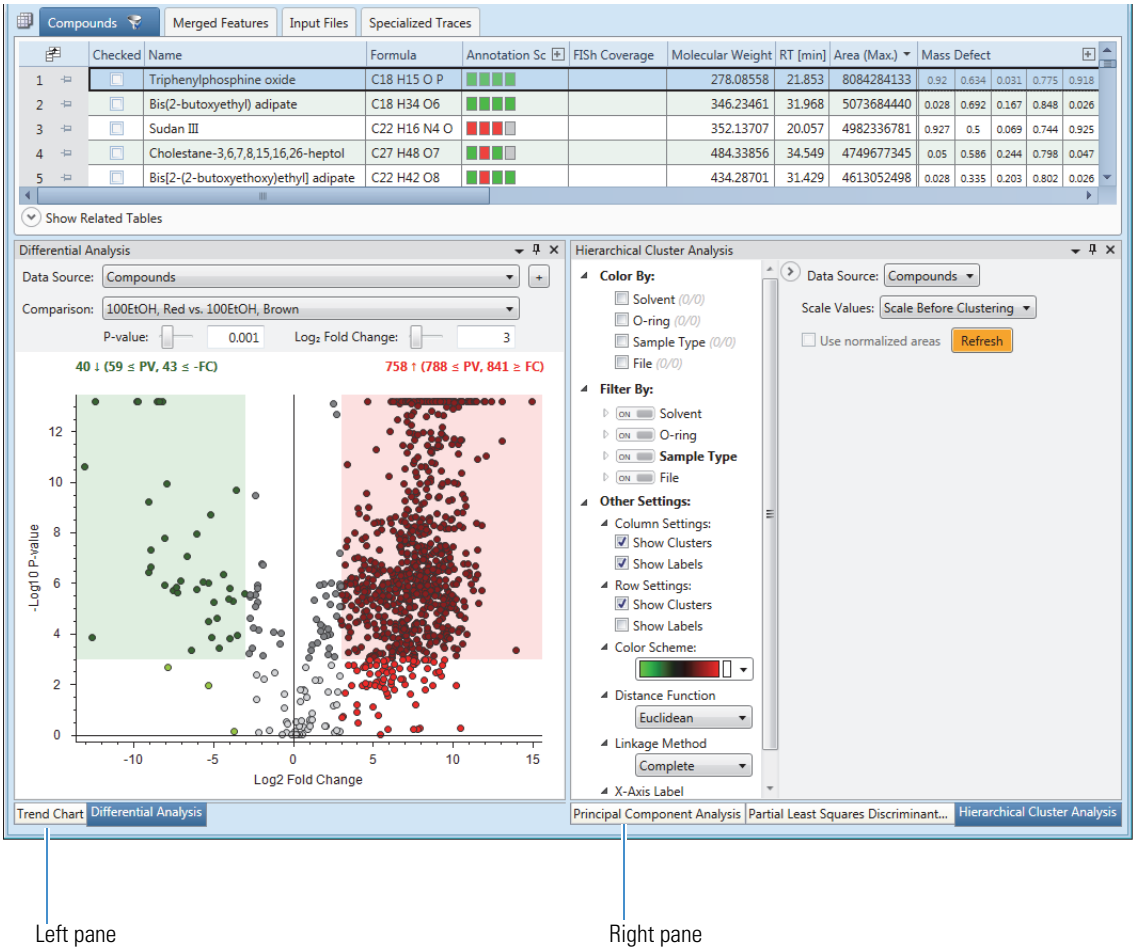
   For the example result file, the Statistics layout does the following (Figure 7):

   - Moves the result tables to the top of the page and hides most of the main tables, except for the Compounds table Merged Features table, Specialized Traces table, and Input Files table.
   - Opens the Differential Analysis and Trend Chart views as tabbed views on the bottom left of the page with the Differential Analysis view selected as the active tab. For the differential analysis, sets the P-value parameter to 0.001 and the Log_2 Fold Change parameter to 3.
   - Opens the Principal Component Analysis, Partial Least Squares Discriminant Analysis, and Hierarchical Cluster Analysis views as tabbed views on the bottom right of the page.

Note The Compounds table, the Partial Least Squares Discriminant Analysis view, the Differential Analysis view, the Hierarchical Cluster Analysis view, and the Loadings Plot page of the Principal Component Analysis view are interactive—that is, checking the points in any of these views checks the points in all of these views.

In the Hierarchical Cluster Analysis view, a segmented bar—red for checked and gray for unchecked—to the right of the heat map indicates whether a compound is checked. In the other views, checked data points turn blue.

Figure 7. Statistics layout for a result file



Working with the Volcano Plot for the Differential Analysis

The processing workflow for this tutorial included the Differential Analysis post-processing node. This node runs a differential analysis on the ratios that you set up and displays the results of the analysis as a volcano plot when you open the Differential Analysis view.

To familiarize yourself with viewing and modifying a differential analysis, follow these procedures:

- To view the volcano plot for the current sample comparison
- To change the p-value for the analysis
- To change the log $_2$ fold change value for the analysis
- To view a different comparison
- To automatically check the up-regulated or down-regulated compounds in the Compounds table
- To uncheck selected compounds
- To view information about a specific data point (compound) in the other open views

❖ To view the volcano plot for the current sample comparison

1. Apply the Statistics layout (see "Applying the Statistics Layout" on page 28).

   The Differential Analysis view displays a volcano plot—a plot of the p-value, the result of a significance test, on the *y* axis versus the $\log_2$ fold change between two sample groups on the *x* axis.

   The *y*-axis scale is the $-\log_{10}$ of the p-value. As the p-value increases from 0 to 1, the $-\log_{10}$ of the p-value decreases from infinity to 0.

The shaded regions of the volcano plot contain compounds where the relative difference between the comparison groups was statistically significant for the specified p-value and outside the specified upper and lower fold change thresholds:

- The pink region contains up-regulated compounds; that is, compounds where the amount extracted from the red o-rings was significantly different from the brown o-rings and greater than the upper fold-change threshold.
- The green region contains down-regulated compounds; that is, compounds where the amount extracted from the red o-rings was significantly different from the brown o-rings and less than the lower fold-change threshold.

2. To change the pane on the bottom left to a floating window, drag it away from the application window.
3. To display the Differential Analysis view's legend, right-click the view and choose **Show Legend** (Figure 8 on page 30).

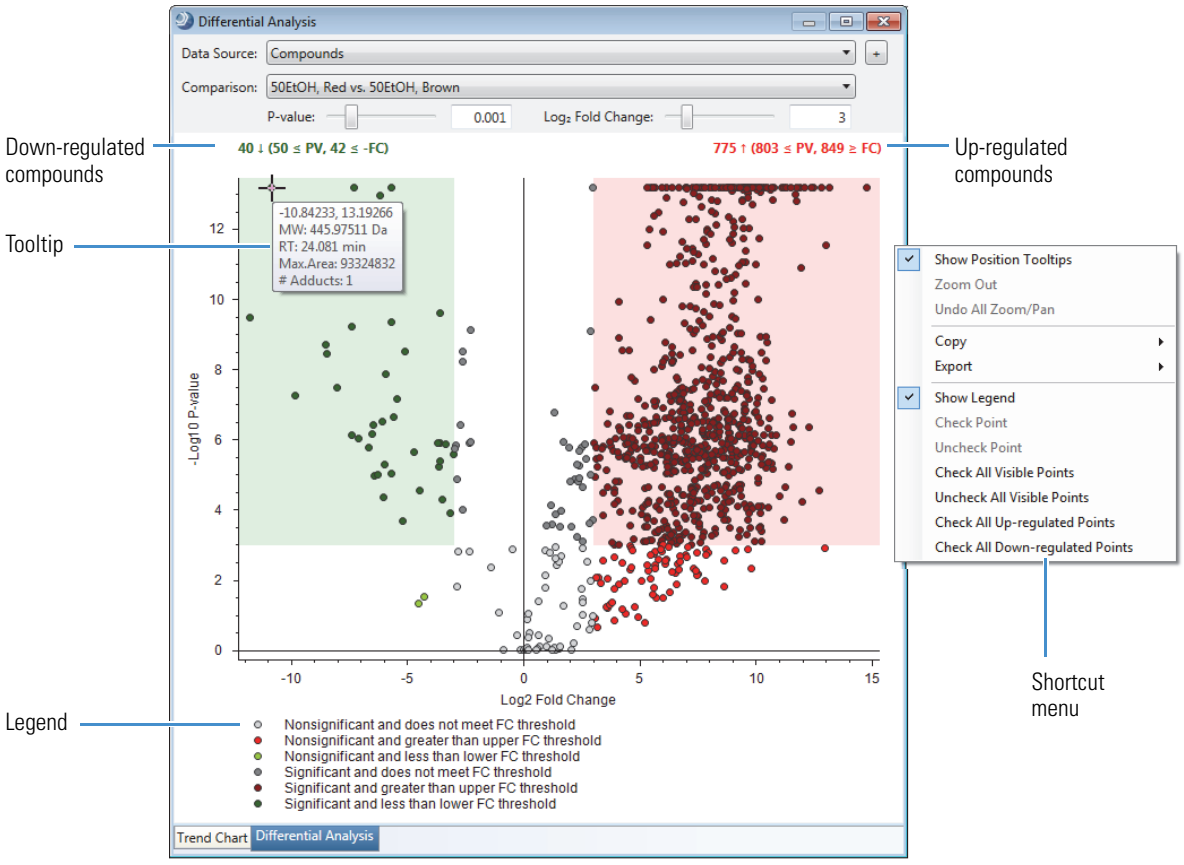
   The legend appears at the bottom of the page with details about the color-coded data points.
4. To display information about a data point, point to it.

   The ToolTip displays the coordinates, MW, RT, maximum area, and number of adducts for the selected compound.

Note You might need to click the view to activate the tooltips.

Figure 8 shows the differential analysis for the comparison between the relative amount of each compound (by chromatographic peak area) extracted from the red o-rings versus the brown o-rings with a solution of 50:50 ethanol/water.

Figure 8. Differential analysis for the compounds extracted from red o-rings or brown o-rings with 50% ethanol/water



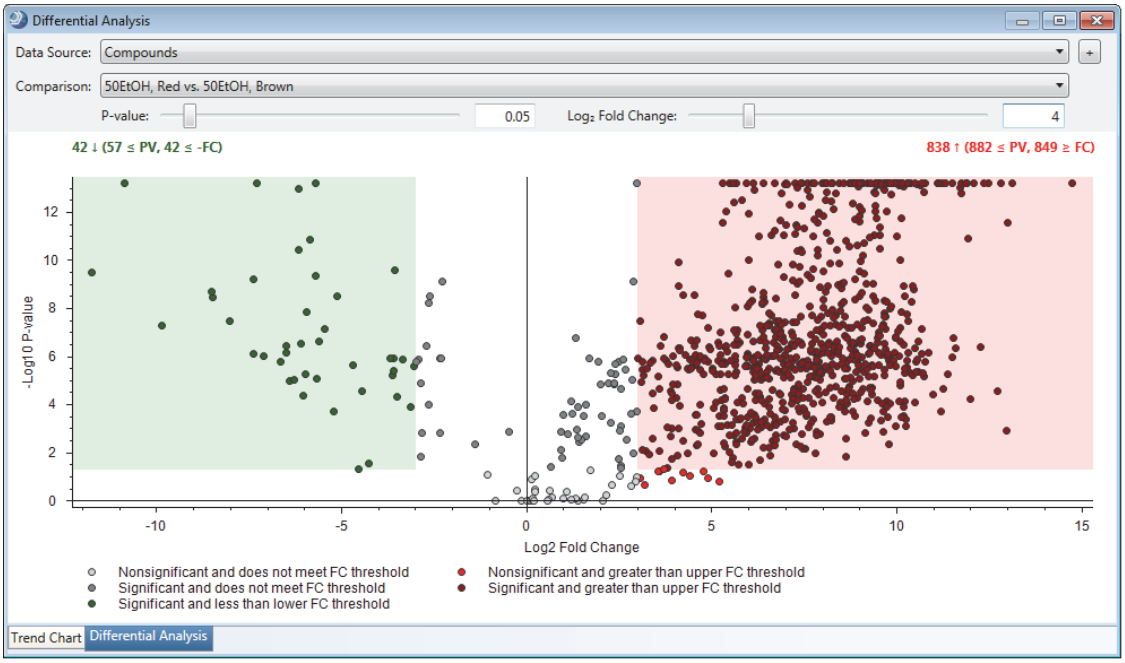
❖ To change the p-value for the analysis

Move the P-value slider to the left or right. Or, type a value in the box next to the slider.

❖ To change the log $_2$ fold change value for the analysis

Move the Log_2 Fold change slider to the left or right. Or, type a value in the box next to the slider.

This figure shows a volcano plot for a p-value setting of 0.05 and a fold change of 4 (log_2 4 = 2).



❖ To view a different comparison

In the Comparison list, select another ratio.

❖ To automatically check the up-regulated or down-regulated compounds in the Compounds table

Do one or both of the following:

- Right-click the view and choose **Check All Up-Regulated Points**.
- Right-click the view and choose **Check All Down-Regulated Points**.

The color of the checked points changes to blue and the application selects the check boxes for the selected compounds in the Compounds table. The application also selects the same data points in other open statistical views.

This figure shows the selection of the up-regulated data points.

❖ To view information about a specific data point (compound) in the other open views

Double-click the compound in the volcano plot to view it in the Compounds table and other open views.

Double-clicking a data point highlights the row for the corresponding compound in the Compounds table; it does not select the corresponding check box.

❖ To uncheck selected compounds

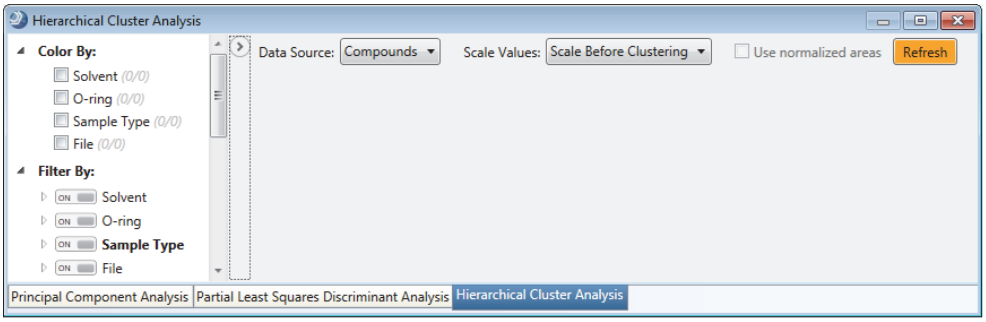
- To uncheck a single data point, right-click it and choose **Uncheck Point**.
- To uncheck all the selected data points, right-click the view and choose **Uncheck All Visible Points**.

Viewing a Hierarchical Cluster Analysis

Use the Hierarchical Cluster Analysis view to visualize the correlation between detected compounds and selected samples in a two-dimensional array of color-coded rectangles (heat map) where each rectangle represents the relative amount (by area) of a specific compound in a specific sample.

❖ To view a hierarchical cluster analysis for the compounds in the Compounds table

1. Apply the Statistics layout (see "Applying the Statistics Layout" on page 28).
2. To change the bottom right pane to a floating window, drag it away from the application window.



3. In the Hierarchical Cluster Analysis view, click **Refresh**.

   A heat map and two dendograms for the cluster analysis appear.
4. In the left pane, under Color By, select the **Solvent** and **O-ring** check boxes.

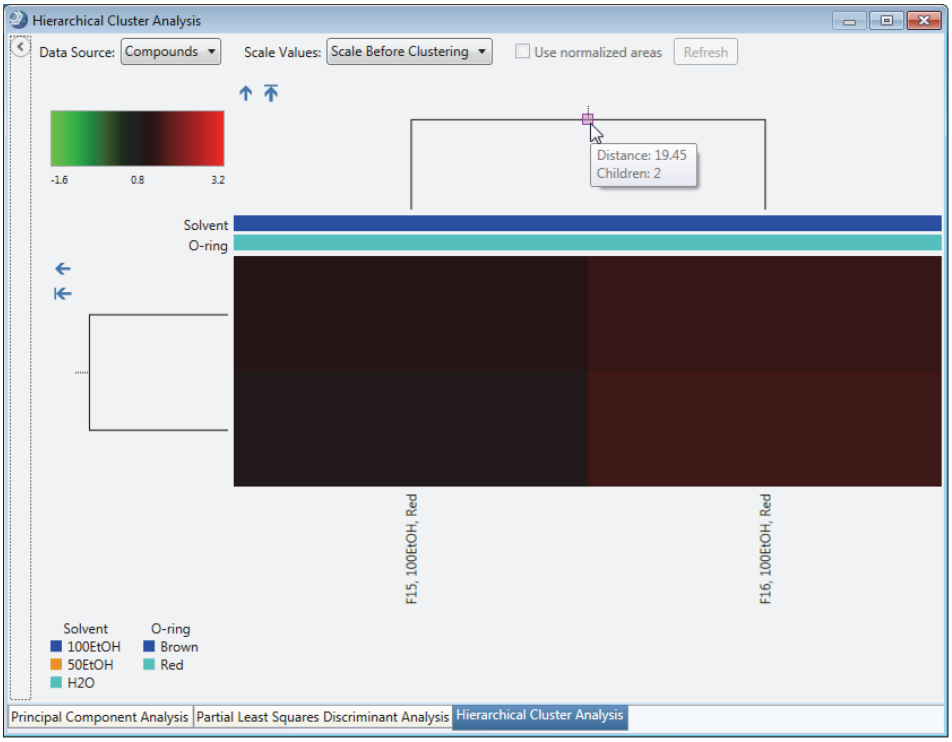
   Color bars appear above the heat map to visually differentiate the samples by their study factors variables.

   This figure shows a hierarchical cluster analysis for the compounds in the Compounds table. Pointing to a cluster's node displays a tooltip.

5. To zoom in on a specific area of the heat map, drag the pointer across the rectangular area of interest.
6. To view information for a cell in the heat map, point to it.

   The tooltip displays the row and column coordinates. The row coordinates are the compound's name, MW, and RT. The column coordinates are the file ID and study factor values.

   This figure shows a magnified view of four cells in the lower right corner of the heat map and their corresponding dendograms.



Go to the next topic "Viewing the Principal Component Analysis."

Viewing the Principal Component Analysis

Use the Principal Component Analysis view to display the results of the principal component analysis. The principal component analysis reduces the dimensionality of the data set to a set of principal components, PC1, PC2, and so on, where PC1 is the principal component with the most variance.

❖ To view the Principal Component Analysis

1. Apply the Statistics layout (see "Applying the Statistics Layout" on page 28).
2. Click the Principal Component Analysis tab.

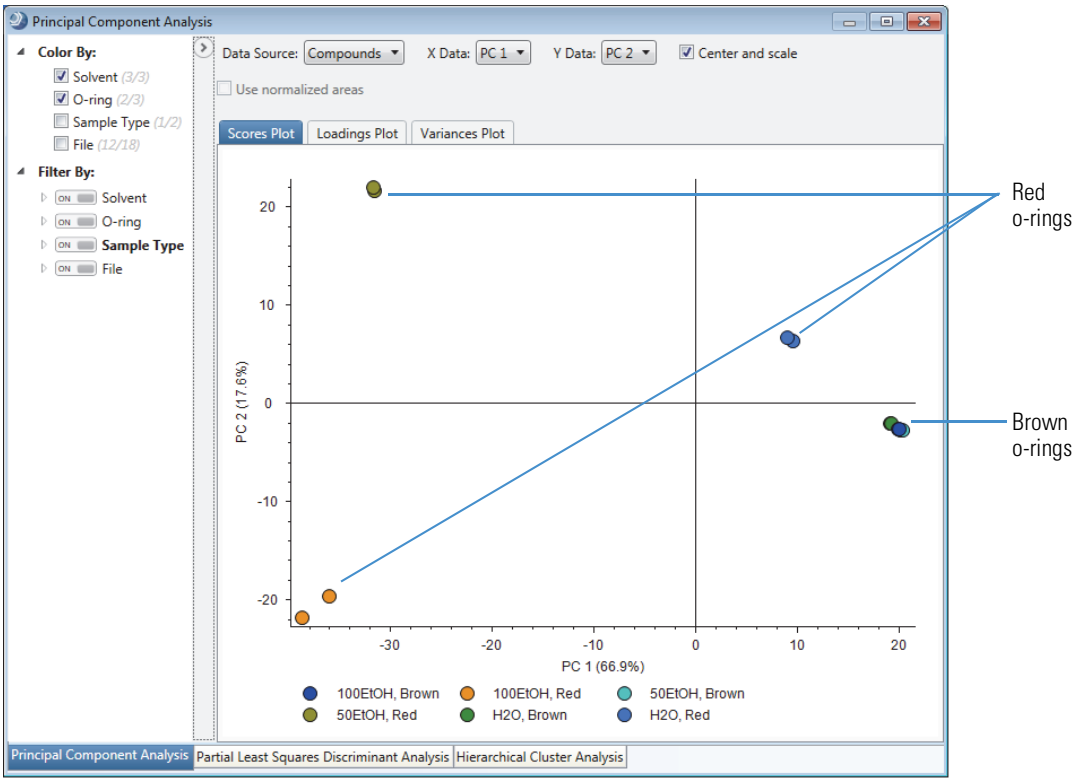
   The Principal Component Analysis view opens to the Scores Plot page (Figure 9 on page 35).

   Use the scores plot to interpret the relationship among the sample groups. Sample groups that are near each other are similar.

The data for the brown o-rings is clustered in the lower-right quadrant, which means that the choice of extraction solvent made little difference.

The data for the red o-rings is divided into three separate quadrants, which means that the choice of extraction solvent did make a difference.

Figure 9. Principal component analysis of compounds extracted from two types of o-rings with three different solvents



Go to the next topic "Working with the Partial Least Squares Discriminate Analysis View."

Working with the Partial Least Squares Discriminate Analysis View

Use the Partial Least Squares-Discriminant Analysis view to find a set of compounds that you can use to discriminate between sample groups.

To familiarize yourself with running and reviewing a partial least squares discriminant analysis, follow these procedures:

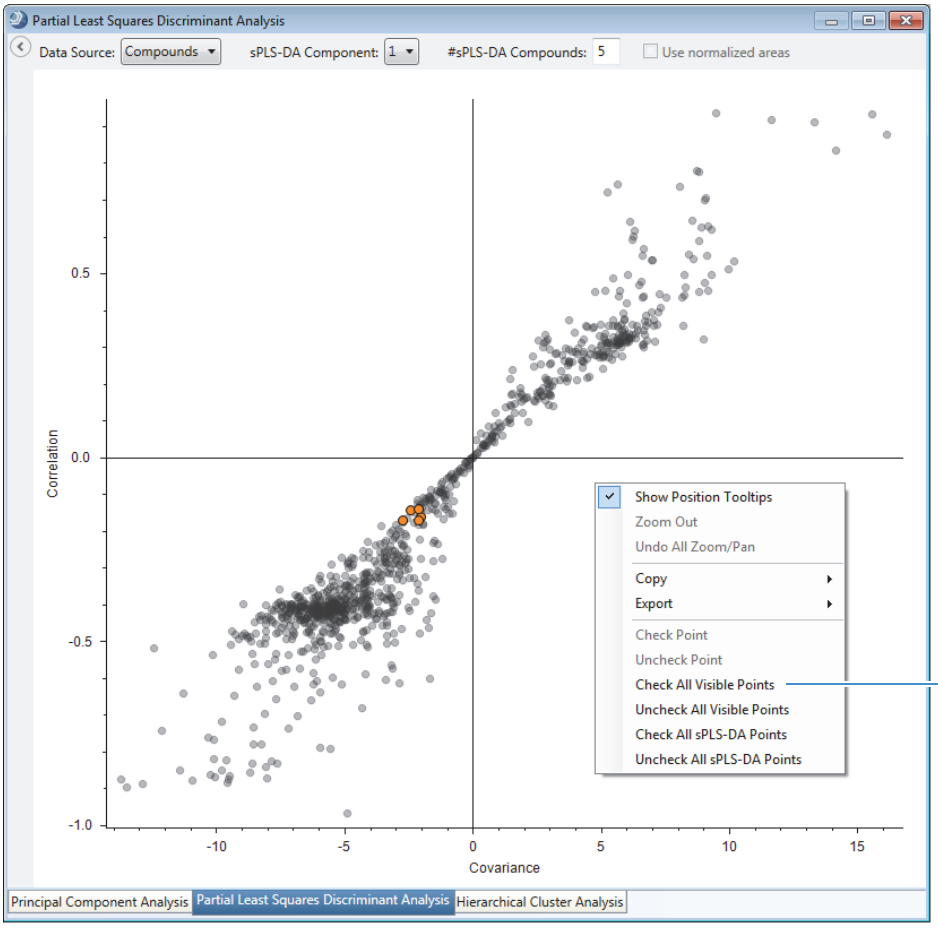
- To make the Partial Least Squares Discriminant Analysis view active
- To rerun the analysis on a different number of discriminating compounds
- To automatically select the check boxes for the discriminating compounds in the Compounds table

❖ To make the Partial Least Squares Discriminant Analysis view active

1. Apply the Statistics layout (see "Applying the Statistics Layout" on page 28).
2. Click the **Partial Least Squares Discriminant Analysis** tab.
3. If the graph includes checked data points, right-click the view and choose **Uncheck All Visible Points**.

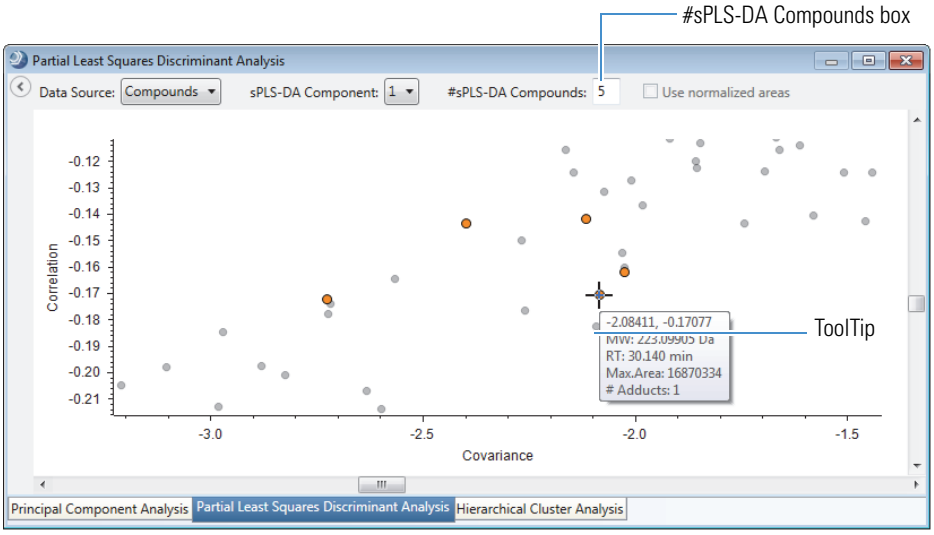
   Figure 10 shows the Partial Least Squares Discriminant Analysis view with the default settings for the Statistics layout. The orange circles represent the five discriminating compounds.

Figure 10. Default settings for the PLS-DA view



4. To zoom in on the points of interest, drag the pointer across the region with the orange circles.

   This figure shows the five orange circles that represent the discriminating compounds. To display a ToolTip, place the crosshair cursor over a circle.



❖ To rerun the analysis on a different number of discriminating compounds

1. Type the number of discriminating compounds in the **#sPLS-DA Compounds** box.
2. To start the analysis, click the view.

❖ To automatically select the check boxes for the discriminating compounds in the Compounds table

Right-click the view and choose **Check All sPLS-DA Points**.

The color of the circles changes to blue and the check boxes for the associated compounds in the Compounds table are selected. If the Differential Analysis view is open, the same compounds are selected.

> **Note** If the discriminating compounds are checked, right-click the view and choose **Uncheck All sPLS-DA Points** before changing the number of discriminating compounds.

Making Structure Proposals

You can add your own custom structure proposals to the result file.

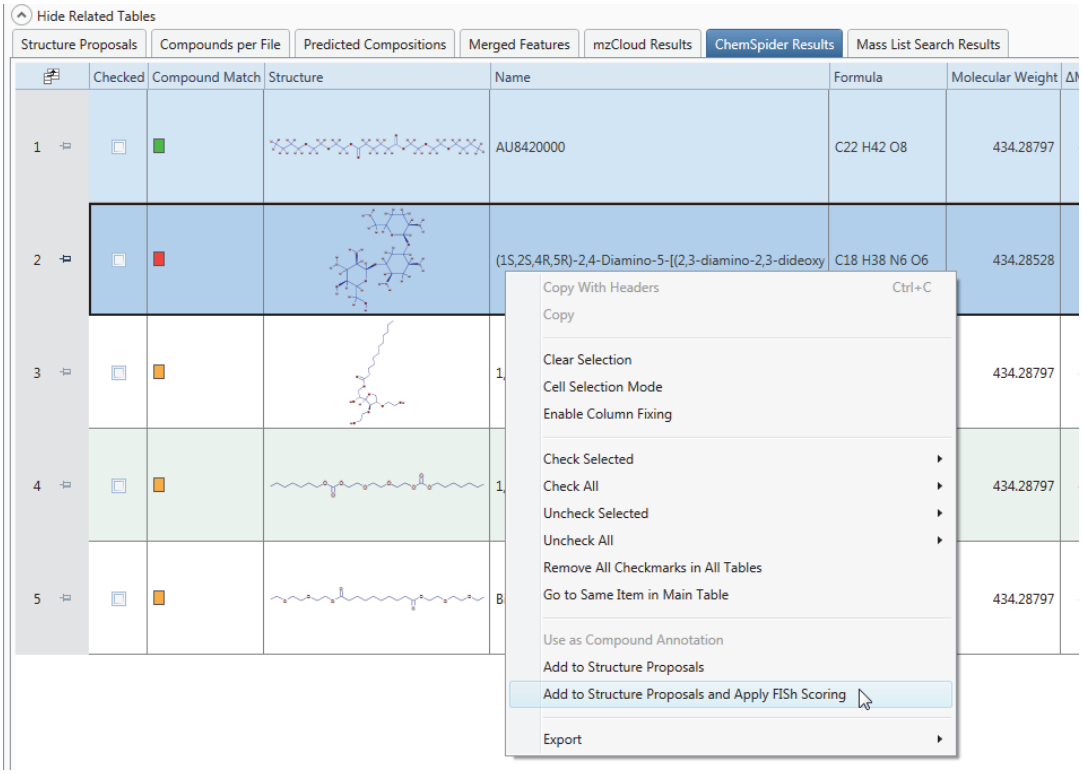
❖ To add a structure proposal for a compound

1. To reset the result file's layout, from the menu bar, choose **Window > Reset layout**.
2. With the default layout, the Compounds table is sorted in descending order by the Area (Max.) column.
3. Select **row 5** (Bis[2-(2-butoxyethoxy)ethyl] adipate).
4. Click **Show Related Tables** below the main tables.
5. Click the **ChemSpider Results** tab.

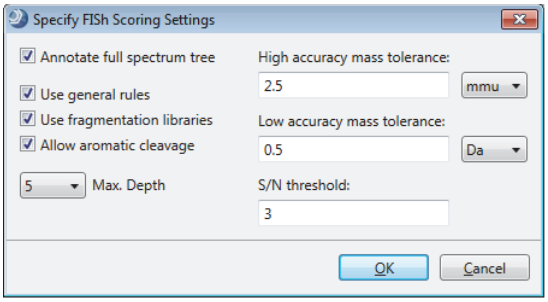
   The ChemSpider Results table for the selected compound lists five structures.
6. In the ChemSpider Results table, select the first two rows. Then, right click the table and choose **Add to Structure Proposals and Apply FISh Scoring**.

> **Tip** To submit multiple entries to the FISh Scoring Queue, hold down the CTRL key, select the structures of interest, right-click and choose **Add to Structure Proposals and Apply FISh Scoring**.

Figure 11. ChemSpider Results table for Bis[2-(2-butoxyethoxy)ethyl] adipate

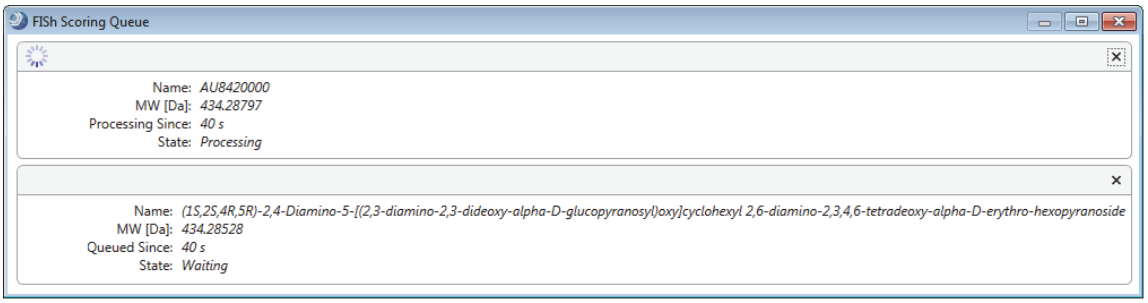


The Specify FISh Scoring Settings dialog box opens. You can change the settings. For example, you can increase the depth of the reaction pathways to a maximum of 20. However, increasing the maximum depth also increases the processing time. If you want the scoring algorithm to ignore small fragments, raise the S/N threshold to 20.



7. Click **OK**.

   The FISh Scoring Queue opens to the right of the result tables.

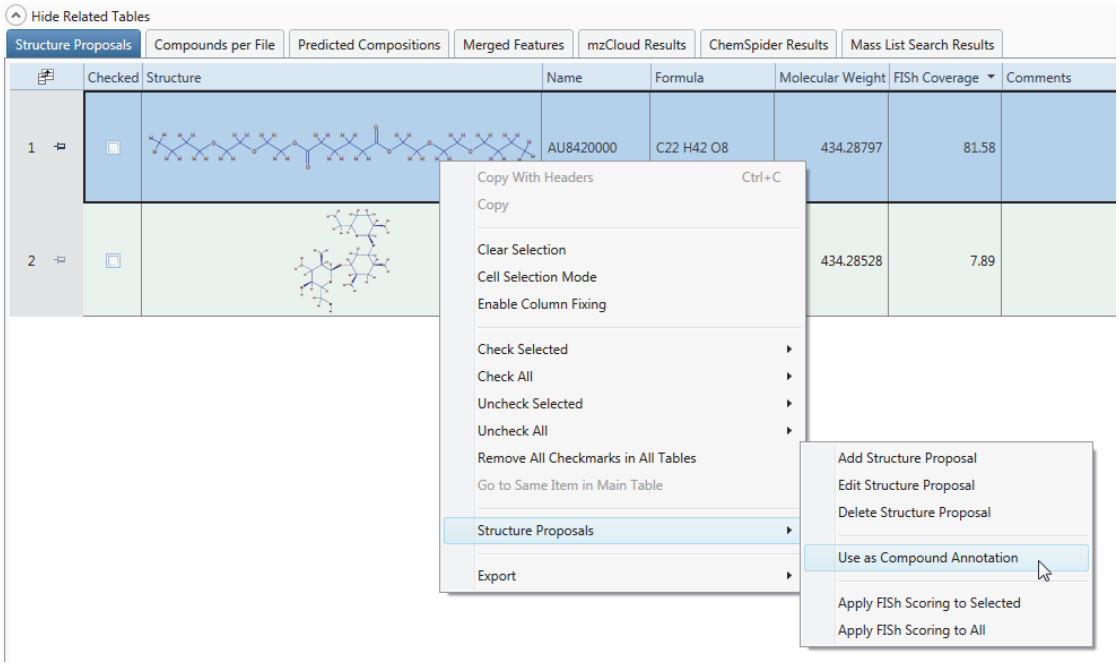


   When the processing is complete, a custom structure proposal and a FISh coverage score appear in the Structure Proposals table for the selected compound.

❖ To use a custom annotation for the compound

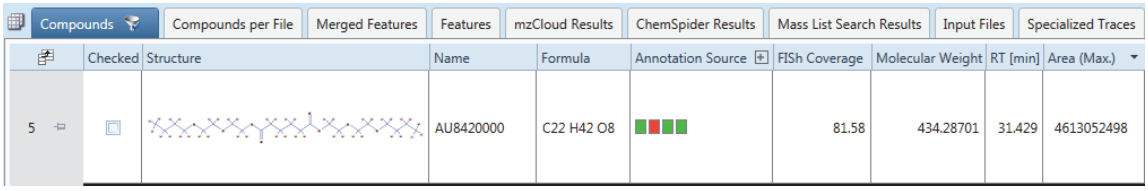
1. With **row 5** (Bis[2-(2-butoxyethoxy)ethyl] adipate) selected in the Compounds table, click the **Structure Proposals** tab in the related tables.
2. In the Structure Proposals table for Bis[2-(2-butoxyethoxy)ethyl] adipate, right-click **row 1** and choose **Structure Proposals > Use as Compound Annotation**.

   This figure shows the shortcut menu for the Structure Proposals table.



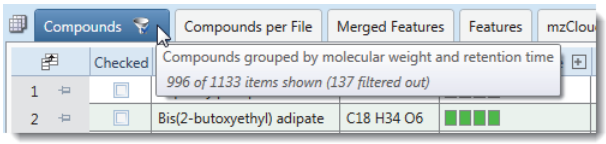
The FISh Coverage column appears in the Compounds table. The custom annotation includes the new name, structure, formula, and FISh coverage score for the compound.

3. To view display the Structure column, open the Field Chooser dialog box and select the **Structure** check box.



Using the Result Filters View

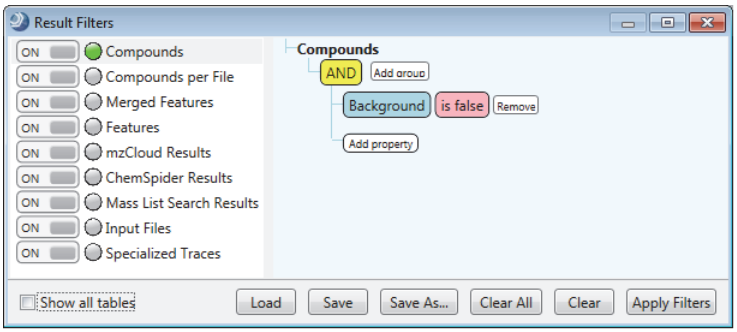
The analysis found a total of 1133 compounds, including 137 background compounds. So the Compounds table displays 996 compounds. To reduce the number of compounds to review, filter the table.



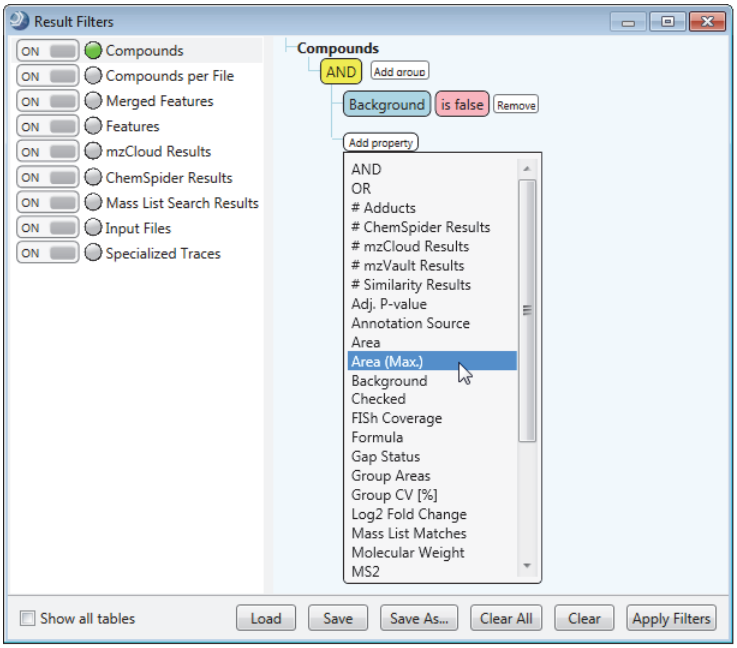
❖ To reduce the number of Compounds to review by the Area (Max.) column

1. Click the **Compounds** tab in the set of main tables.
2. From the menu bar, choose **View > Result Filters**.

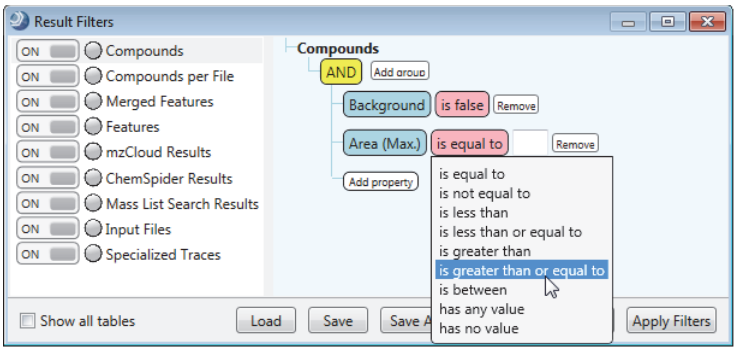
   The Result Filters view opens as a floating window. Because the processing workflow included the Mark Background Compounds node, the filter for the Compounds table already includes a Background filter and the table does not display the background compounds.



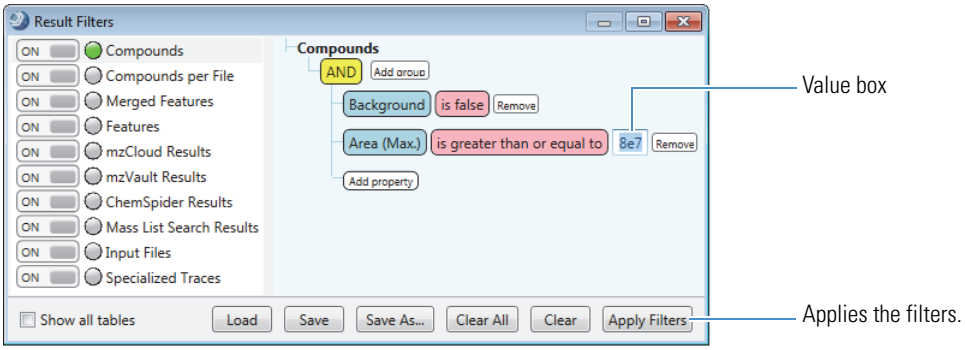
3. Click **Add Property** and select **Area (Max.)**.



4. In the relation list, select **Is Greater Than or Equal To**.



5. In the value box, type **4e9**.



6. Click **Apply Filters**.

   The Compounds table now displays only seven compounds.

❖ To display only the checked compounds

1. Click the **Compounds** tab in the set of main tables.
2. From the menu bar, choose **View > Result Filters**.
3. Click **Add Property** and select **Checked**.
4. Click **Apply Filters**.

Exporting the Results to an Excel Spreadsheet

Before exporting the results to a spreadsheet, filter the results table or select the check boxes for the compounds of interest (see "Using the Result Filters View" on page 39).

❖ To check the number of table rows

Point to the Compounds tab.

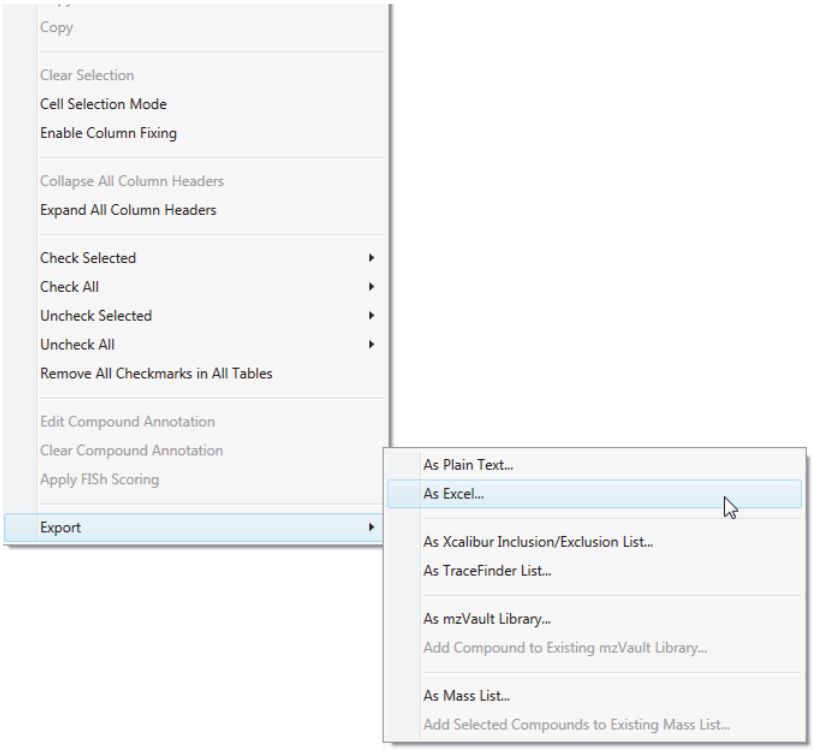
A tooltip appears with the number of displayed compounds.

❖ To select the columns that you want to export

1. Click the **Field Chooser** icon, , for the Compounds table.
2. For this tutorial, clear all the check boxes, except for the following:
   - Area (Max.)
   - Formula
   - Mass List Matches
   - MW
   - Name
   - RT

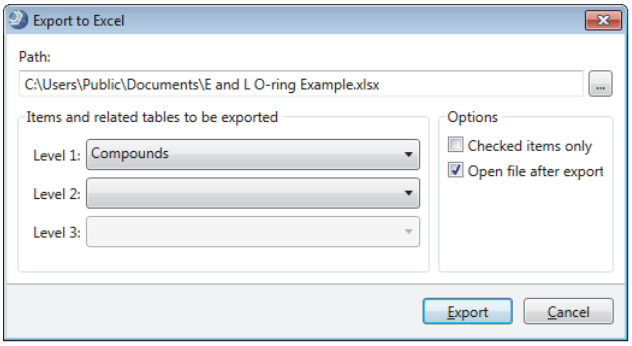
❖ To export the filtered and sorted results to an Excel spreadsheet

1. Right-click the Compounds table and choose **Export > As Excel**.



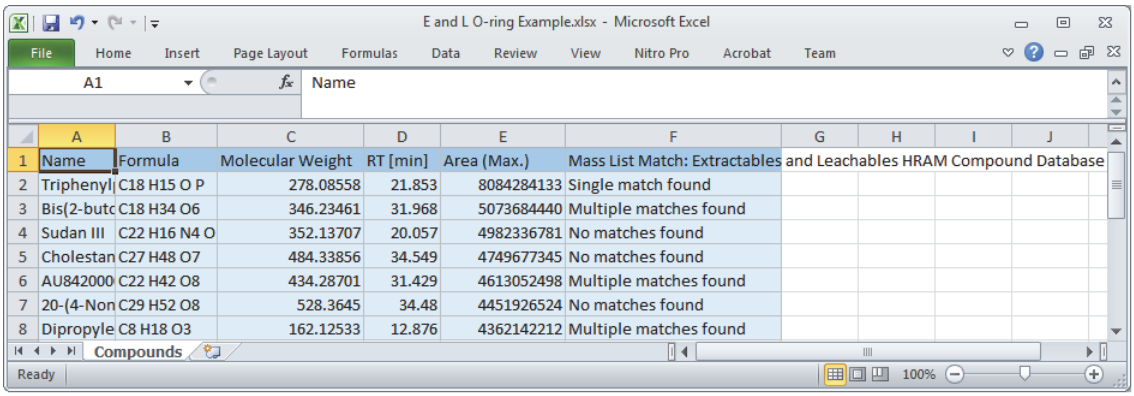
   The Export to Excel dialog box opens.

2. Check the file name and location in the Path box. Change the file name and folder as appropriate.
3. In the Options area, select the **Open File After Export** check box.



4. Click **Export**.

   The Excel spreadsheet opens.



	A	B	C	D	E	F	G	H	I	J
1	Name	Formula	Molecular Weight	RT [min]	Area (Max.)	Mass List Match: Extractables	and Leachables HRAM Compound Database			
2	Triphenyl	C18 H15 O P	278.08558	21.853	8084284133	Single match found				
3	Bis(2-butc	C18 H34 O6	346.23461	31.968	5073684440	Multiple matches found				
4	Sudan III	C22 H16 N4 O	352.13707	20.057	4982336781	No matches found				
5	Cholestan	C27 H48 O7	484.33856	34.549	4749677345	No matches found				
6	AU842000	C22 H42 O8	434.28701	31.429	4613052498	Multiple matches found				
7	20-(4-Non	C29 H52 O8	528.3645	34.48	4451926524	No matches found				
8	Dipropyle	C8 H18 O3	162.12533	12.876	4362142212	Multiple matches found				

Trademarks

Compound Discoverer, Exactive Plus, mzCloud, mzLogic, and Q Exactive are trademarks and Exactive, Orbitrap, Orbitrap Fusion, and Xcalibur are registered trademarks of Thermo Fisher Scientific Inc. in the United States.

Microsoft and Excel are registered trademarks of Microsoft Corporation in the United States and other countries.

All other trademarks are the property of Thermo Fisher Scientific Inc. and its subsidiaries.