

Compound Discoverer 3.3 SP3 GC PCI/ MS/MS Tutorial

Thermo Scientific™ Compound Discoverer is a small molecule identification application for GC Orbitrap and high-resolution accurate-mass LC-MS/MS data.

This tutorial guides you through the process of setting up a study and an analysis, submitting the analysis to the job queue, reviewing the results of the analysis, and creating a report. The raw data that you are processing was acquired with a GC Orbitrap™ mass spectrometer operated in the positive chemical ionization (PCI) mode.

Contents

- [Overview](#)
- [Start the application](#)
- [Check the computer's access to external databases](#)
- [Set up the study and analysis](#)
- [Review the processing workflow](#)
- [Submit the job to the job queue](#)
- [Open the result file and review the default layout](#)
- [Modify the mass options for a compound](#)
- [Add downstream nodes and reprocess the analysis](#)
- [Open the reprocessed result file and review the compound annotations](#)
- [Make structure proposals for a compound and run FISH Scoring](#)
- [Filter the data to reduce the number of compounds to report](#)
- [Export the spectral information for various compounds to a NIST MSP file](#)
- [Print a report by using a defined template](#)

Overview

To get started, see these topics:

- [Copy the example files to your processing computer](#)
- [Summary of the tutorial workflow](#)
- [The Help system](#)

Copy the example files to your processing computer

The example files for this tutorial are provided on the key-shaped USB drive that comes in the software media kit. You can find these files in the following folder:

Example Studies\GC\PCI\Pesticides

Tip You can also download these example files from the Compound Discoverer 3.3 SP3 Product Download page of the LSMS Software Download and Licensing Portal website.

To access this site, go to the following URL (case-sensitive):

thermo.flexnetoperations.com

v **To copy the example files to your processing computer**

1. Locate the example files on the USB drive that Thermo Fisher Scientific provides in the application media kit or download the files from the LSMS Software Download and Licensing Portal website.
2. Copy the files to a folder on your processing computer.

Example files	Description
Pesticides_standard.raw	Xcalibur RAW data file from a pesticides sample. The data file was acquired at a resolution of 120k.
Pesticides.cdStudy	This study file includes the location and sample type of the input file. The example study does not include the analysis settings.
Pesticides core.cdResult	This example result file includes the results of the core analysis described in this tutorial. The core analysis only includes compound detection; it does not include compound identification.
Pesticides with ID.cdResult	This example result file includes the identification results for the pesticides detected in the sample file.

IMPORTANT For optimal performance, store all the Compound Discoverer study files (.cdStudy) and result files (.cdResult) on a local hard drive, ideally a solid state drive (SSD). Latency, read- and write speeds of external USB-connected hard drives and network drives are typically much slower than internal hard drives. Because the result files are continuously accessed throughout the entire data processing workflow, processing times can be significantly longer when using external drives. Unlike result files, raw data files (input files) are read only once, at the very beginning of the processing workflow. So, you can store them on an internal or external drive, without significantly affecting the processing time of your analyses.

The flowchart in [Figure 1](#) shows the workflow for creating, running, and reviewing an analysis of GC PCI data where you identify the compounds in a single sample.

Tip The analysis described in this tutorial does not require access to online data bases or custom mass lists. For information about checking your computer's access to online databases or adding custom mass lists to the application, refer to the online Help.

In addition to guiding you through the basic workflow, the advanced workflow shown in [Figure 1](#) guides you through the process of modifying the mass options for compounds of interest.

In this tutorial, you do the following:

1. Run the core processing workflow that deconvolves the spectra and detects and groups the chromatographic peaks.

Note The Compound Discoverer application comes with multiple processing workflow templates for GC data. Two of these templates include only the core processing workflow nodes. For GC PCI data, the name of the core processing workflow template is Core GC PCI - Before Reprocessing. [Figure 8](#) on [page 11](#) shows the workflow nodes in this core processing workflow.

**Summary of the
tutorial
workflow**

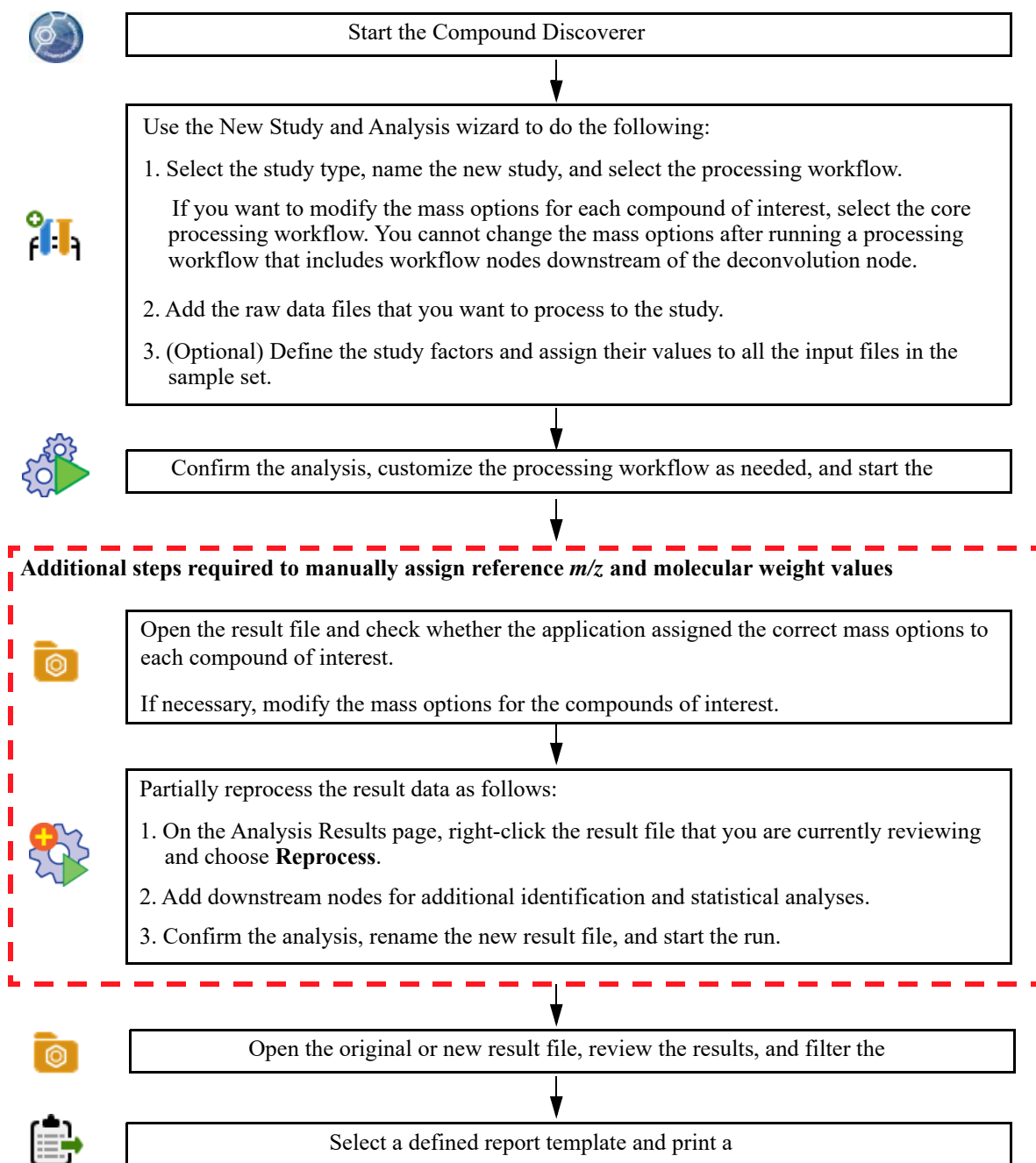
2. Open the result file and manually change the mass options for a compound of interest.
3. Add downstream nodes to the processing workflow and partially reprocess the data.

Partial reprocessing is faster than full processing because it uses the results from the initial deconvolution process.

Note If the deconvolution node typically assigns the correct mass options (reference m/z and molecular weight values) to your compounds of interest, select a processing workflow that includes downstream nodes instead of the core processing workflow and skip the intermediate data review and partial reprocessing steps for your analyses.

The following flowchart shows the typical workflow for the analysis of GC/MS data with the Compound Discoverer application.

Figure 1. Advanced workflow that includes two processing steps



Note In this tutorial, there is only one sample type, and you do not need to customize any of the lists and libraries provided with the application.

The Help system

The Compound Discoverer 3.3 SP3 application provides Help for the views, pages, and dialog boxes.

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

1. Open the view, page, or dialog box

Start the application


1. On the computer keyboard, press the **F1** key, or equivalent (Fn + F1 keys).

v To start the Compound Discoverer application

Do one of the following:

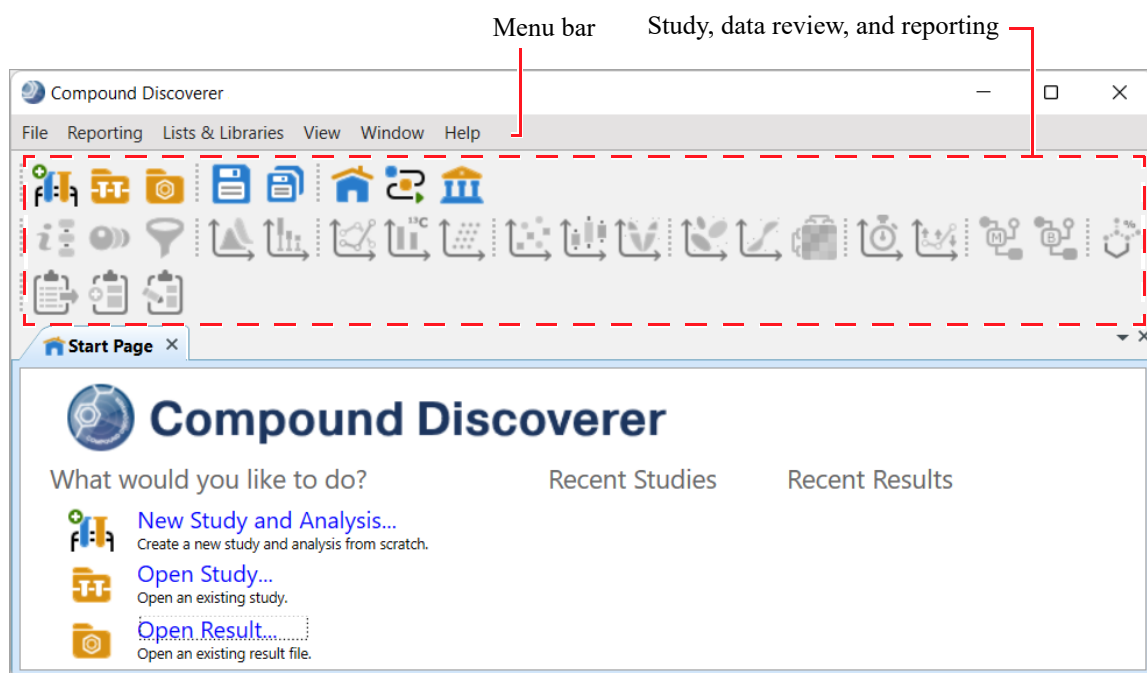
- From the Windows taskbar, choose **Start > All Programs (or All Apps) > Thermo Compound Discoverer 3.3**.

—or—

- From the computer desktop, double-click the **Compound Discoverer 3.3** icon, .

The Compound Discoverer window opens with the Start Page displayed as a tabbed document (see Figure 2). As you create studies and process data, the application creates and populates recent file lists to the right of the What Would You Like to Do? hyperlinks.

Figure 2. Application window with the initial Start Page and large toolbar icons



Check the computer's access to external databases

The processing workflow that you select for this tutorial uses the ChemSpider database to identify unknown compounds. To use any of the processing workflows that use online databases, such as ChemSpider, your processing computer must have unblocked access to these databases on the Internet.


To test the communication to the online databases that you use in this tutorial, do the following:

1. From the menu bar, choose **Help > Communication Tests**.
2. In the Communication Tests dialog box, click the **ChemSpider** tab and click **Run Tests**.

If your computer has an Internet connection, but these tests fail, leave the Communication Test dialog box open and press the **F1** key to open the Help. Then, follow the instructions to troubleshoot the communication failure.

v To set up the new study and analysis

1. Do one of the following:

- From the menu bar, choose **File > New Study and Analysis**.
- In the toolbar, click the **New Study and Analysis** icon, .
- In the What Would You Like to Do? area on the Start Page, click **New Study and Analysis**.

The New Study and Analysis wizard opens to the Study Name and Processing Workflow page.

The first time you open the wizard, the Studies Folder parameter is undefined (Figure 3). After you select the directory folder for your Compound Discoverer studies, the application remembers the location. You can store all your studies in one directory or create new directory folders as needed.



Tip To display instructions for each page of the wizard, click the **Show Description** icon, , in the lower-left corner.

Figure 3. First page of the wizard

Initially, the studies directory is undefined.

 Opens instructions to the left of the page for the current page of the wizard.

2. On the Study Name and Processing Workflow page, do the following in any order:

- In the Study Type area, select the **GC** option.
- In the Study Name and Directory Structure area, do the following:
 - In the Study Name box, type **Pesticides**.

Note The application creates a study file and a study folder with this name. The application stores the study file (CDSTUDY) and all the result files (CDRESULT) that you generate by running analyses within the study in the study folder.

- Click the browse icon to the right of the Studies Folder box. Then, create a new folder for storing all your Compound Discoverer studies (or a subset of studies) and name this folder **Studies**.

Note You can store one or more of your studies in this top-level folder—that is, you do not need to create a new top-level *Studies* folder each time you create a new study. When you create a study, the application creates a study folder with the same name as the study and stores this folder in the top-level *Studies* folder.

- In the Processing area, open the Workflow list and select the following processing workflow:

Core GC PCI - Before Reprocessing

Figure 4 shows the selections and entries for this tutorial.

Figure 4. New Study and Analysis wizard—Study Name and Processing Workflow page

New Study and Analysis Wizard - Step 1 of 5

Study Name and Processing Workflow
Specify a unique name for this study and its folder, select the studies folder for storing all of your study folders, and select a processing workflow for the current analysis.

Study Type

GC LC

Study Name and Directory Structure

Study Name: Pesticides

Studies Folder: C:\Studies

Study Template File: (Optional)

Description: (Optional)

Processing

Workflow: WorkflowTemplates \ GC \ GC PCI \ Core GC PCI Workflow - Before Reprocessing

Workflow Description: For Methane PCI

Cancel < Back Next > Finish

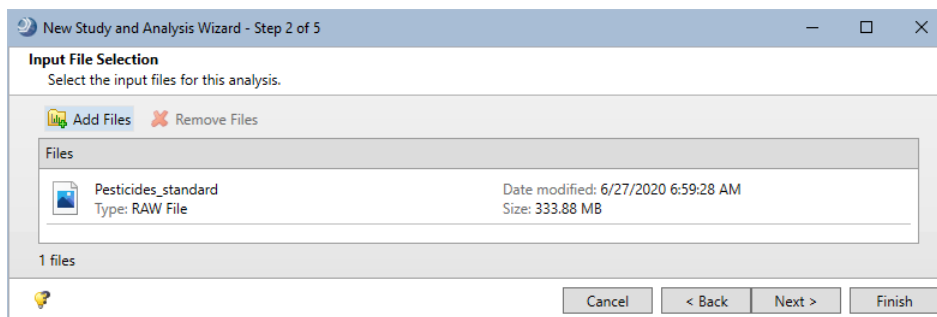
3. Click **Next**.
4. On the Input File Selection page, do the following:
 - a. Click **Add Files**.

The Add Files dialog box opens.

- b. Select the example data file, **pesticides_standard.raw**, and click **Open**.

The file name of the selected raw data file appears in the Files box and the Finish button becomes available (Figure 5).

Figure 5. Input File Selection page with the selection of one file



c. Click **Finish**.

Note There are no study factors for this study. You are processing a single raw data file to detect and identify compounds in the sample.

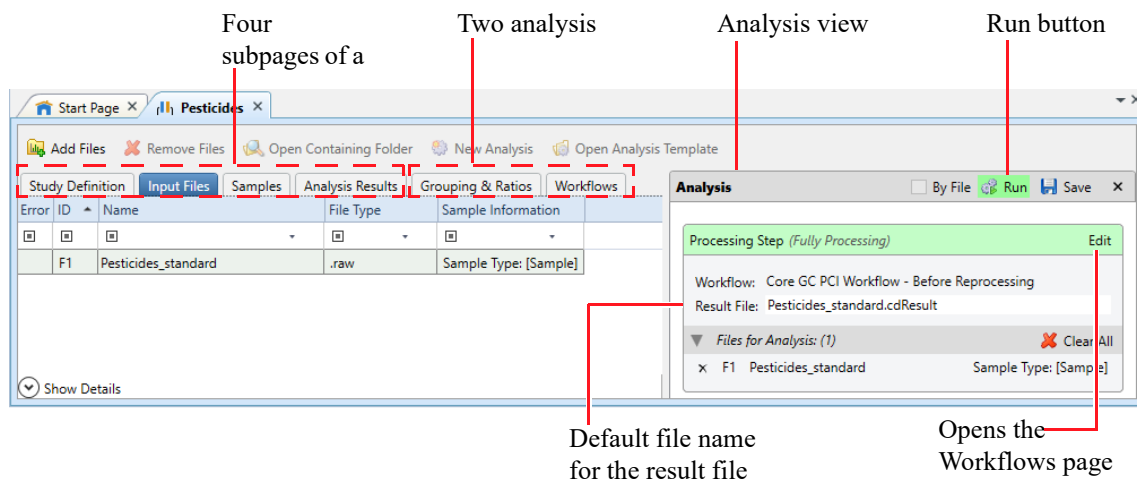
The study opens as a tabbed document at the left of the application window, and the Analysis view opens to the right of the study.

The study consists of four subpages (from left to right)— Study Definition, Input Files, Samples, and Analysis Results.

The two pages (Grouping & Ratios and Workflows) to the right of the study pages are part of the analysis—that is, if you close the Analysis view, these pages disappear.

The Analysis view lists the name of the processing workflow, the name of the result file, and the raw data files that you selected. If the analysis is valid, the Run button is green. If the analysis is invalid, the Run button is unavailable.

Figure 6. Pesticides study page and Analysis view



Note By default, the application uses the name of the first input file as the result file name. You can edit the file name.

Tip If you close the Analysis view, the Grouping & Ratios and Workflows pages also close.

To recreate the analysis, do the following:

1. In the command bar on the study page, click **New Analysis**.

The Analysis view opens. A caution symbol appears to the right of the Edit button because the Files for Analysis area is empty.

2. To repopulate the Files for Analysis area, drag the input files from the Input Files page or the Samples page of the study to the Files for Analysis area of the Analysis view.
3. To select the processing workflow, do the following:
 - a. Open the Workflows page by clicking the **Workflows** tab.
 - b. In the Workflows command bar, click **Open Common**.
 - c. Browse to and select a processing workflow.
 - d. Click **Open**.

The processing workflow appears in the Workflow Tree pane.

Go the next topic to [“Review the processing workflow,”](#) for the core analysis for GC PCI data.

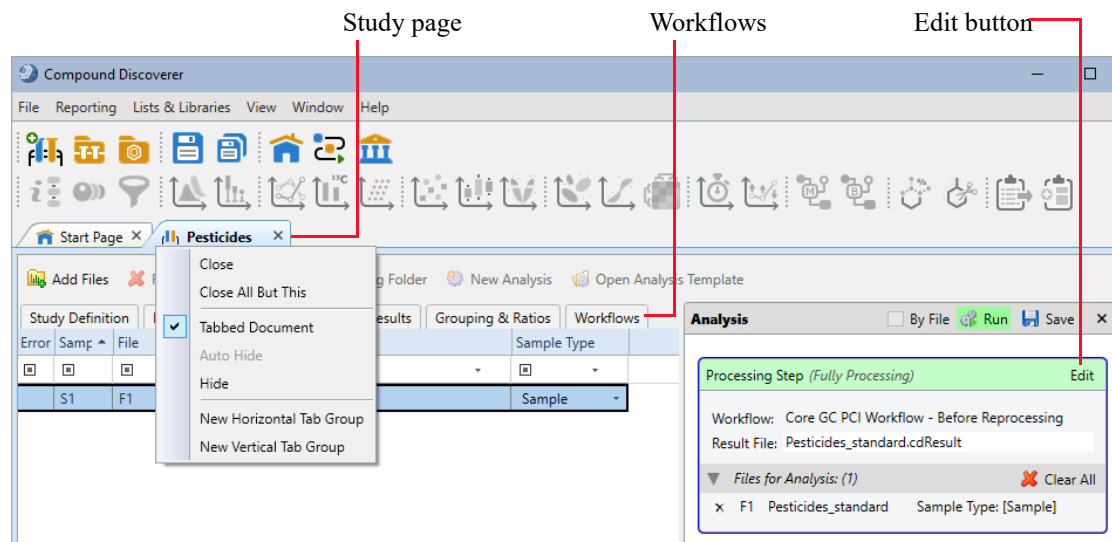
Review the processing workflow

For this tutorial, you use the default parameter settings in the Core GC PCI - Before Reprocessing workflow template. However, to familiarize yourself with the parameters that you might want to customize for your own analyses, review the parameter settings in the GC CI Deconvolution node as described in the following procedure.

v To review the parameter settings in the GC CI Deconvolution node

1. If you have not already created a new study and a new analysis by using the New Study and Analysis wizard, create one as described in [“Set up the study and analysis”](#) on page 6. Or create a new study and analysis with your own GC PCI data files.

When you finish the wizard, the study opens as a tabbed document in the application window.



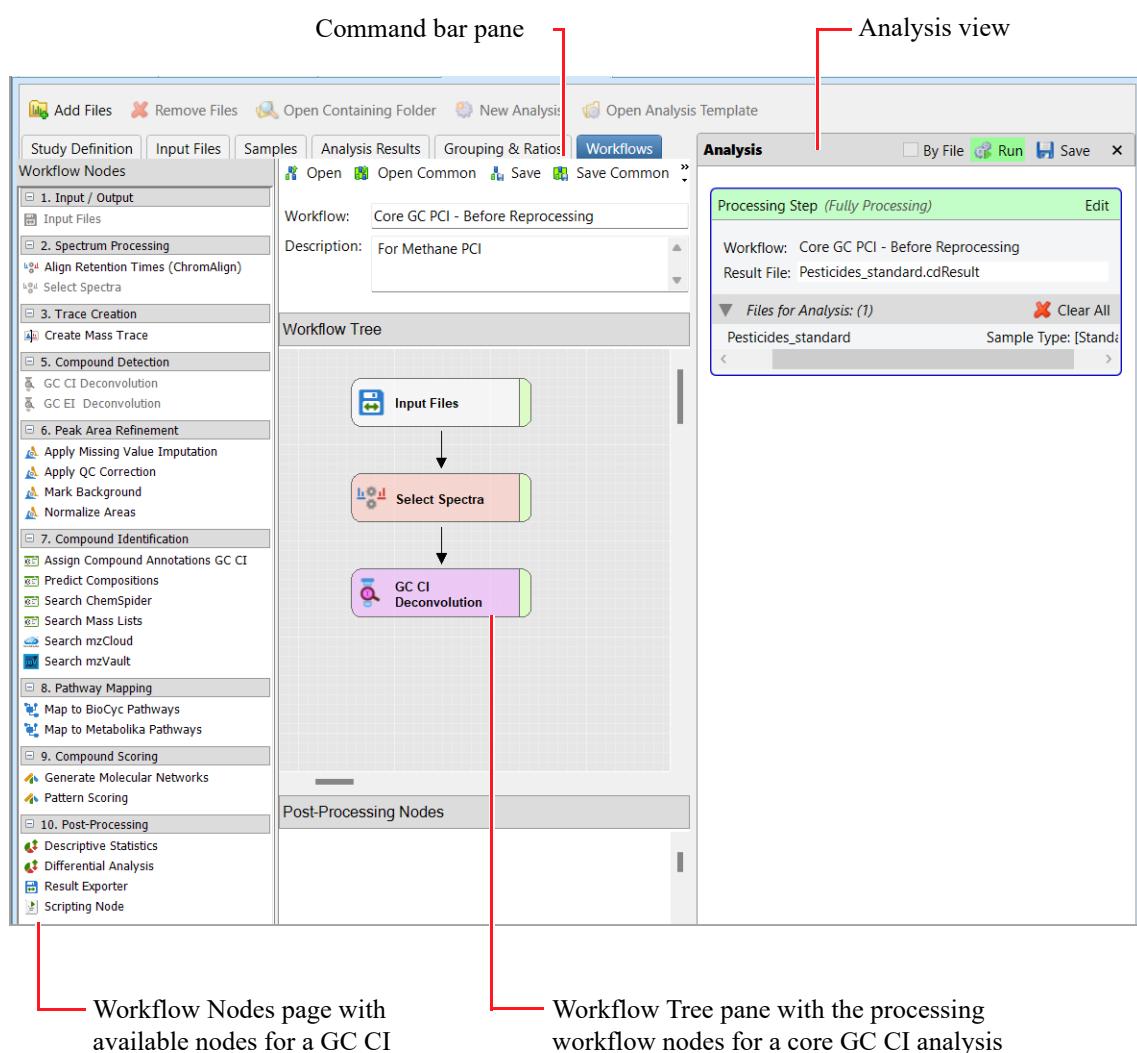
2. Open the Workflows page of the analysis by clicking the **Workflows** tab or by clicking **Edit** in the Analysis view.

The Workflows page opens to the left of the Analysis view.

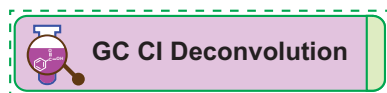
The Workflows page includes four panes that are always visible—the command bar pane for opening, renaming, and saving processing workflow templates; the Workflow Tree and Post Processing Nodes panes where you can add and delete workflow nodes; and the workflow nodes and parameters pane, which includes two pages—the Workflow Nodes page and the parameters page for the workflow node that is currently selected in the Workflow Tree pane or the Post-Processing Nodes pane.

The Workflows page also includes the Current Workflow Issues pane, which opens when the processing workflow contains missing parameter settings or missing connections between the workflow nodes and closes after you fix all the issues.

Figure 7. Workflow Nodes page at the left, Workflow Tree pane in the middle, and Analysis view at the right



- In the Workflow Tree pane of the Workflows page, click the **GC CI Deconvolution** node to select it. A dashed green border appears around the GC CI Deconvolution node.



And, the parameters page for the GC CI Deconvolution node opens to the left of the Workflow Tree pane. Most of the parameters in the node are set to their default values, but Thermo Fisher Scientific customized a few of the parameters settings for the processing workflow template. See [Figure 8](#).

When you select a parameter in the parameters pane, information about the parameter appears at the bottom of the pane. For this tutorial, do NOT modify the settings in the processing workflow template.

Figure 8. Parameter settings for the GC CI Deconvolution node in processing workflow template

The screenshot shows the 'Parameters of GC CI Deconvolution' pane on the left, which is organized into four sections:

- 1. Peak Detection Settings:** Mass Tolerance (5 ppm), Spectral S/N Threshold (3), Peak S/N Threshold (3), Smoothing (9), TIC Threshold (1000), Ion Overlap Window [%] (98), CI Gas Type (Methane), Include Reference and Exception Peaks (False).
- 2. RI Settings:** Calculate RIs (False), Column Type (Semi std non polar), n-Alkane Reference Items.
- 3. Group Compounds Settings:** Group Compounds Across Samples by Spectrum Dot Product and Retention Time (True), RT Tolerance [sec] (5), Dot Product Threshold (500), Composition Threshold [%] (0).
- 4. Library Search Settings:** Do Library Search (False), Library Search Type (Normal), Search Libraries, SI/RSI Threshold (500), Use Reverse Search (False), Use RI Delta Filter (False), Maximum RI Delta (100), Use RI Diff[%] Filter (False), Maximum RI Diff [%] (10), Use Unspecified Column Type (False).

The 'Workflow Tree' pane on the right shows a workflow with three nodes: 'Input Files', 'Select Spectra', and 'GC CI Deconvolution'. The 'GC CI Deconvolution' node is highlighted with a dashed green border. The 'Post-Processing Nodes' pane is currently empty.

Note For your own analyses, modify the processing workflow template as follows:

- Peak Detection Settings area** Select the CI gas type for your GC/MS system: Ammonia or Methane. Take care when modifying any of the other parameter settings for peak detection. Decreasing the Ion Overlap Window [%] from the default value of 98% increases the probability of grouping coeluting ions together as one compound.
- RI Settings area** When you have the retention indexing information for your GC/MS system, enable the Calculate RIs feature and enter the n-alkane reference items. To limit the library hits for compounds with RI data to those with matching RI values within a user-specified tolerance, enable the Use RI Delta Filter or the Use RI Diff [%] filter.

- **Group Compounds Settings area** Keep True as the setting for the Group Compounds Across Samples By Retention Time and Spectrum Dot Product parameter even when the analysis includes only one input file, as this parameter must be set to True to send data about the detected compounds to the downstream nodes. When you reprocess and analysis, adding downstream nodes does not trigger full processing, but resetting any parameter in the deconvolution node does trigger full reprocessing.

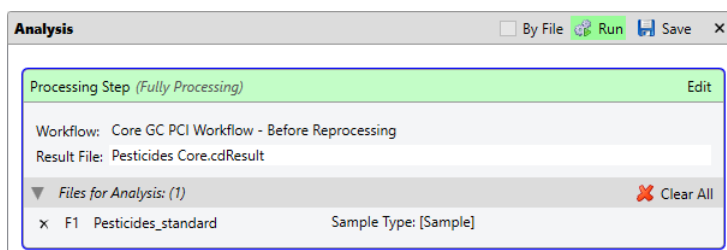
Consider increasing the Composition Threshold. For example, if the analysis includes multiple input files and you want the application to ignore compounds unless they are present in more than one of the input files or all the input files, increase this percentage.

Go to the next topic to “[Submit the job to the job queue.](#)”

Submit the job to the job queue

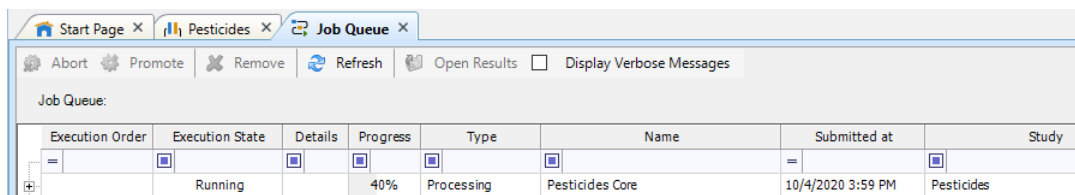
v To submit the job to the job queue

1. In the Result File box of the Analysis view, change the name to **Pesticides Core**.



2. In the Analysis view, click **Run**.

The Job Queue page opens.



Processing for this single file is complete within a few minutes.

Go to the next topic to “[Open the result file and review the default layout.](#)”


Open the result file and review the default layout

To open the result file and familiarize yourself with the default layout, see these topics:

- [Open the result file](#)
- [Default result page layout](#)
- [Check how many compounds the analysis detected](#)
- [Layout modifications](#)

v To open the result file when the current analysis is completed

Double-click the completed run on the Job Queue page.

Tip If the Job Queue page is closed, open it by choosing **View > Job Queue** from the menu bar or clicking the **Show Processing Job Queue** icon, , in the toolbar.

Open the result file

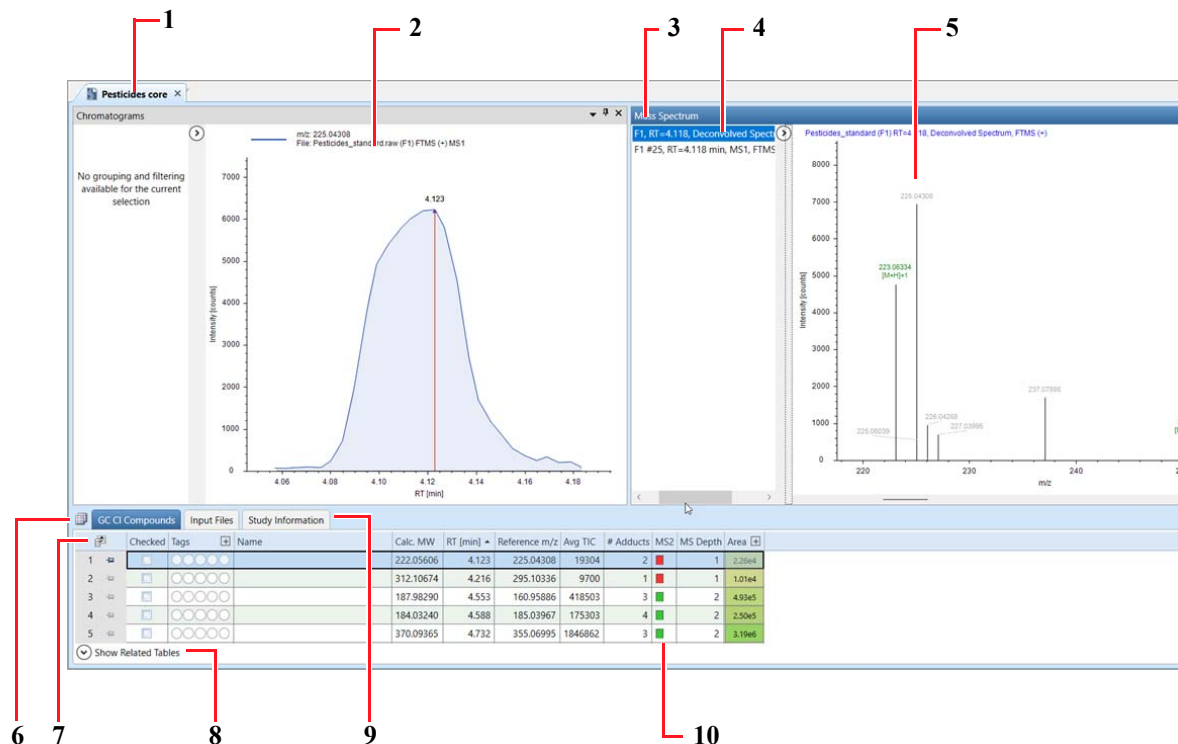
Default result page layout

The result file (analysis result) opens as a tabbed document in the application window. The next topic, “[Default result page layout](#),” describes the default layout for a result page from a core GC CI analysis.

Figure 9 shows the factory default layout for the Pesticides Core.cdResult file. In the Chromatograms and Mass Spectrum views, the application automatically displays the chromatogram and the deconvolved spectrum for the compound in the first row of the GC CI Compounds table.

When methane is used as the chemical ionization gas, the GC CI Deconvolution node labels the following adduct ions (in green) in the MS1 spectrum, when it detects them: $[M+H]^+$ (pseudo molecular ion), $[M-e]^+$, $[M-Hydrate]^+$, $[M+C_2H_5]^+$, and $[M+C_3H_5]^+$.

Figure 9. Factory default layout for the core processing workflow (numbered from left to right)



No. Description

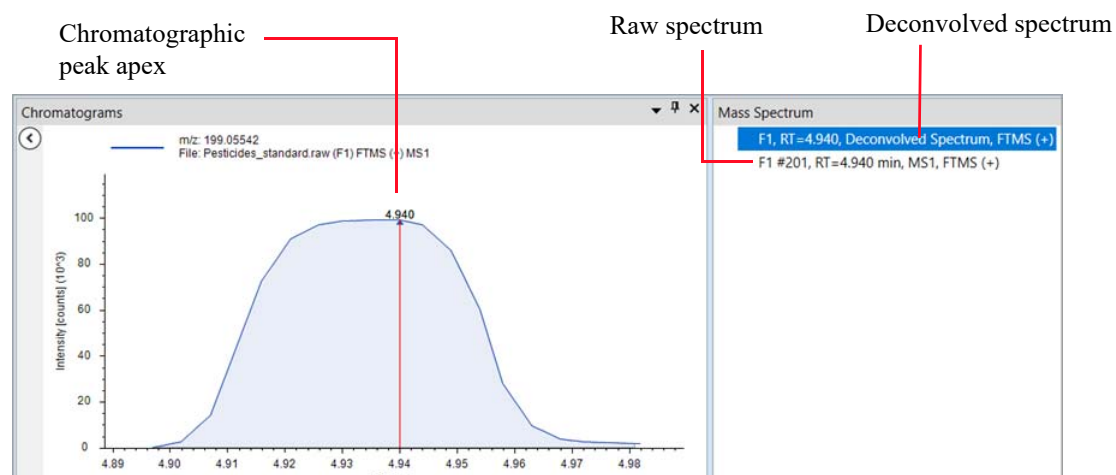
- 1 Tab with the result file name.
- 2 Chromatograms view on the top left populated with XIC traces for the compound in the first row of the GC CI Compounds table, with one trace for each sample where the analysis detected the compound. The application crops the extracted ion current (XIC) traces to the start and end points of the compound's chromatographic peak before writing the data to the result file.
The analysis for this tutorial included only one sample.
- 3 Mass Spectrum view on the top right populated with the deconvolved MS1 spectrum for the base compound.
- 4 Spectrum tree to the left of the spectrum plot (see [Figure 10](#), [Figure 11](#), and [Figure 12](#)).
- 5 When the spectrum for the base compound includes the molecular ion, the monoisotopic molecular ion is annotated (in green) with its m/z value and formula.
- 6 Opens the Select Visible Tables dialog box.
- 7 Opens the Field Chooser for selecting the visible columns in the active table.

No.	Description
8	Collapsed area for the related tables (Related tables are tables related to the item selected in a main table).
9	Main tables below the two graphical views. The GE CI Compounds table is the active table and is sorted by RT in ascending order (primary) and average total ion current (Avg TIC) in descending order (secondary). The application groups compounds across the input files by their RT × spectrum dot product dimensions. The base compound is the compound with the largest total ion current from the spectra across its chromatographic peak.
10	The green and red squares in the MS2 column indicate the presence or absence, respectively, of MS2 scans for a compound.

The spectrum tree to the left of the spectrum plot includes the following nodes:

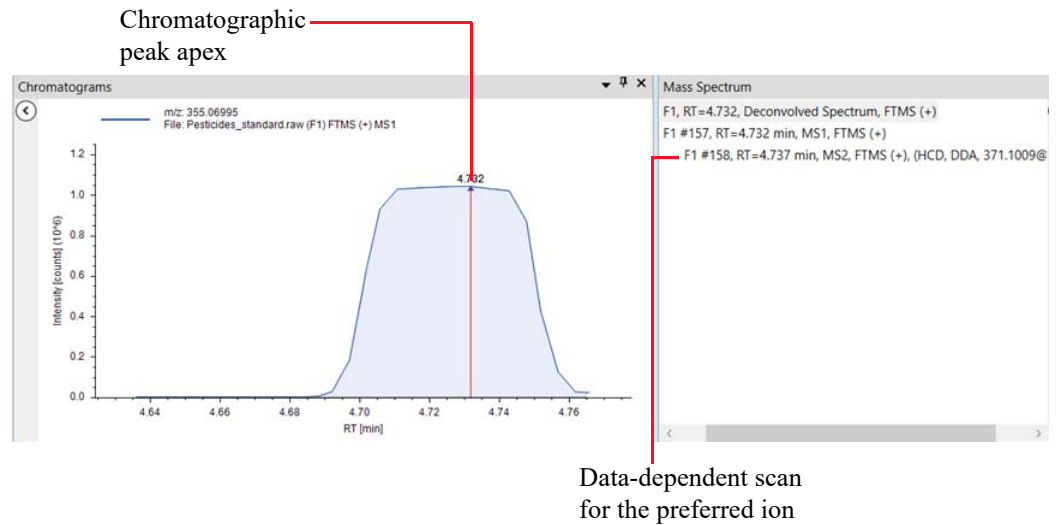
- First node—MS1 deconvolved spectrum
- Second node—Raw MS1 spectrum from the apex of the compound's chromatographic peak

Figure 10. Spectrum tree with no data-dependent fragmentation scans



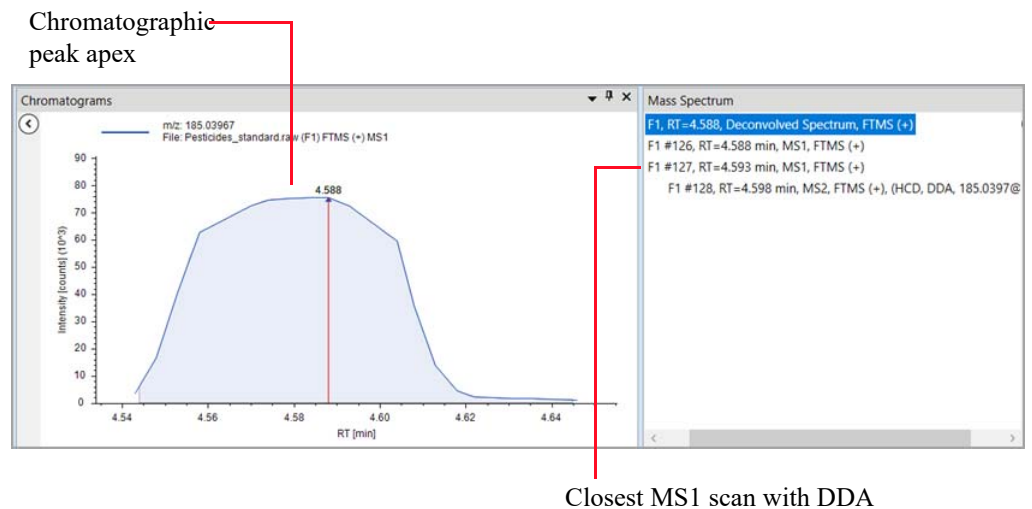
- Subsequent nodes:
 - If the MS1 spectrum at the peak apex has a data-dependent fragmentation scan for the preferred ion, this data-dependent fragmentation scan is listed next.

Figure 11. Spectrum tree with data-dependent (DDA) scans for the preferred ion



- If the MS1 spectrum at the peak apex does not have any data-dependent scans, the next closest MS1 scan with data-dependent (DDA) scans is listed next, followed by its data-dependent scan.

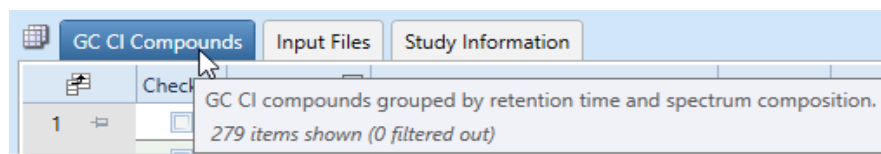
Figure 12. Spectrum tree where the MS1 spectrum at the peak apex does not have any DDA scans



Check how many compounds the analysis detected

- v **To check how many compounds the analysis detected**

Point to the GC CI Compounds tab to display a tooltip.







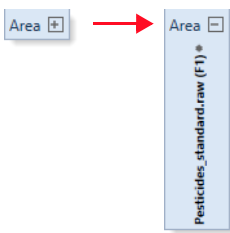






The analysis detected 279 compounds. Because the processing workflow did not include the Mark Background Compounds node or blank samples, no compounds were filtered out.

Layout modifications

You can change which columns, tables, and views are visible or hidden. To display the subcolumn headings for columns that contain multiple subcolumns, you must expand the column headers.

The Area column is the only column in the analysis result for a core GC CI analysis that can include multiple columns—for example, when you process multiple input files. In this tutorial, you process only one input file, so the Area column includes only one column.

Table 1. Common layout modifications

To do this	Do the following
Show or hide a table column	Open the Field Chooser for a table by clicking the icon,  , in the upper-left corner of the table. To display a column, select its check box. To hide a column, clear its check box.
Show or hide a table	Open the Select Visible Tables dialog box by clicking the icon,  , at the left of the table tabs.
Close a view	Click the close icon,  , in the upper-right corner of the view.
Open a view	In the application menu bar, choose View > Specific View . Or, in the toolbar, click the icon for the view.
Expand the header for a column with multiple subcolumns	Click the expand icon,  , to the right of the heading. 
Freeze a column to the left side of the table	Right-click the table and choose Enable Column Fixing . Then, click the pin icon to the right of the column heading.
Pin a row to the top of the result table.	Click the pin icon to the right of the row number (unpinned,  , or pinned, )
Sort a result table by a column with numeric or text information.	Click a column header once or twice to sort the rows in ascending order () or descending order (), based on the contents of the column.
Sort a result table by multiple columns.	Click the column header of the primary sort column once or twice to sort the rows in ascending order () or descending order (), based on the contents of the column. Hold down the CTRL key and click the column header of the secondary sort column once or twice to set the sort order.

You can also change the relative location of the table columns and views. Or, float the views and drag them to another monitor. For more details, refer to the following Help topics:

“Modify the result page layout” and “Common operations for manipulating data tables”

Modify the mass options for a compound

The application does not always identify the ideal reference mass, which it uses to calculate the chromatographic peak area for a compound, or the correct pseudo molecular ion, which it uses to calculate the molecular weight for a compound. So in some cases, you might need to manually modify these values.

You cannot change the molecular weight or reference mass value for a compound in a result file from a processing workflow with downstream nodes—that is, from a processing workflow that included any of the identification, peak area refinement, mapping, or post-processing nodes.

Note In this tutorial, you are working with a sample that includes pesticide standards in a simple solvent matrix. For this sample, the application makes the correct assignments for the reference mass and pseudo molecular ion for each detected compound.

Follow this procedure to learn how to change the mass options for your analytes when applicable.

v To change the molecular weight for a compound (and reverse the change when applicable)

1. Open the result file from the analysis where you selected the core processing workflow. See “[Open the result file and review the default layout](#)” on [page 12](#).

Note If you did not process the example data file, open the **Pesticides Core.cdResult** file in the Pesticides example folder that you copied from the USB drive in the media kit or downloaded from the Life Sciences Mass Spectrometry Software Download and Licensing Portal.

2. Make sure that the GC CI Compounds table is sorted by RT [min] in ascending order.

To sort the table by RT in ascending order click the **RT [min]** column heading until the arrow to the right of the heading points up. Or, choose **Window > Reset Layout** from the application menu bar.

3. In the main GC CI Compounds, select **row 93** (RT 14.062 min).

4. Open the Features table for this compound as follows:

- a. Click **Show Related Tables** at the bottom left of the result page.
- b. In the related tables pane, click the **GC CI Compound per File** tab.
- c. Select **row 1** in the GC CI Compound per File table.
- d. Click **Show Related Tables** at the bottom left of the result page.

The related tables pane opens with the Features per File table displayed.

5. For this tutorial, sort the **Is Ref Mass** column of the Features per File table in descending order (▼). Then, hold down the CTRL key and sort the **Ion** column in descending order (▼).

In [Figure 13](#), the Features per File table is sorted by the Is Ref Mass column and the Ion column.

Figure 13. Original analysis results for the compound at RT 14.062 minutes (in the GC CI Compounds table)



No. Description

- 1 Original calculated MW for the compound in row 93 = 320.89501
- 2 Original fragment assignment for the pseudo molecular ion

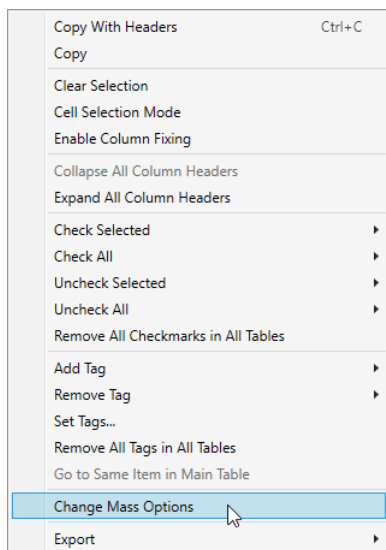
Note You can use the delta mass columns in the Features per File table to identify plausible alternative selections for the representative pseudo-molecular ion for the compound.

Each delta mass column heading indicates the relationship to the measured m/z value if the feature was selected as the $[M+H]^+$ ion. For example, the ΔC_2H_4 m/z column for a feature lists the detected mass that would represent the $[M+C_2H_5]^+$ adduct if the measured m/z value for the feature was selected as the $[M+H]^+$ adduct, the ΔC_3H_4 m/z column lists the detected mass that would represent the $[M+C_3H_5]^+$ adduct, and so on.

The presence of multiple delta masses for a feature indicates that the feature could be the $[M+H]^+$ ion for the compound because its surrounding spectral pattern at least partially fits the pattern typical of a pseudo-molecular ion cluster.

6. In the Features per File table, right-click **row 1** (Measured m/z 323.89926 and Is Ref Mass = True) and choose **Change Mass Options**.

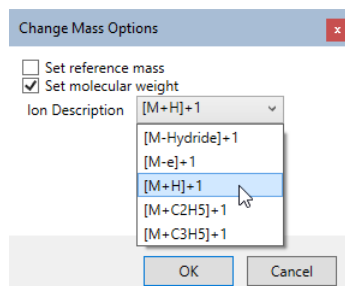
Figure 14. Shortcut menu for the Features per File table



The Change Mass Options dialog box opens (Figure 15).

7. In the Change Mass Options dialog box, do the following:
 - a. Select the **Set Molecular Weight** check box.
The Ion Description list appears.
 - b. Select the **[M+H]⁺+1** adduct ion from the Ion Description list.

Figure 15. Change Mass Options dialog box

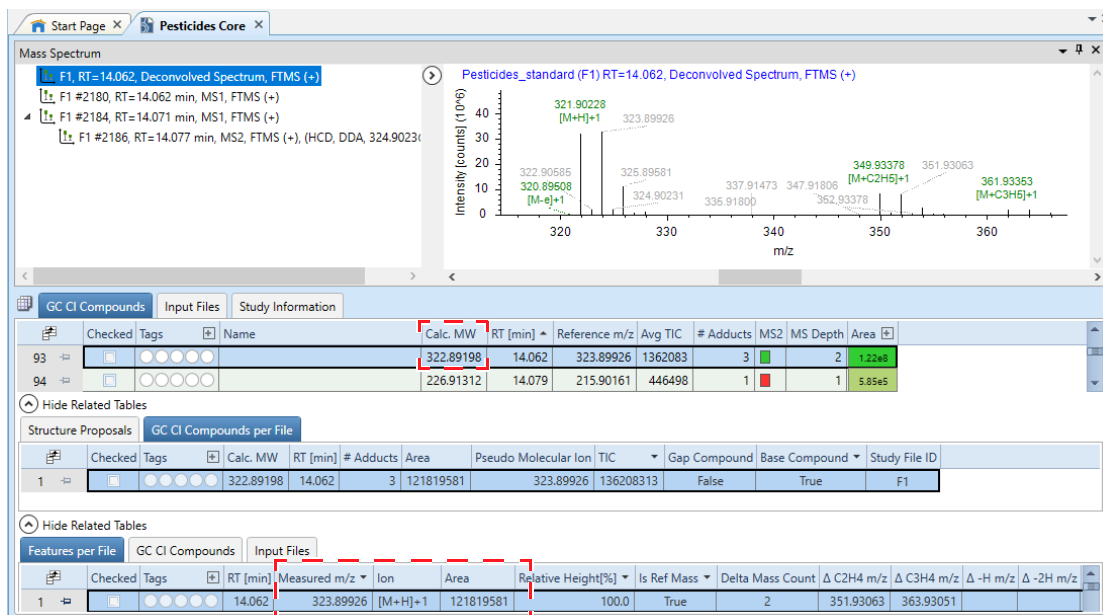


- c. Click **OK**.
8. Check the Calc. MW column for the compound in row 93 of the GC CI Compounds table.

Figure 16 shows the recalculated molecular weight for the compound in row 93 (RT = 14.062 min).

Column	Original	Revised
Calc. MW	320.89501	322.89198 Da
Reference m/z	323.89926	no change
#Adducts	5	3
Area	1.22 e ⁸	no change

Figure 16. Recalculated molecular weight for the compound at RT = 14.062 minutes



9. Undo the ion selection for the compound in row 93 as follows:
 - a. In the Features per File table, right-click the original measured m/z assigned as the pseudomolecular ion—321.90228—and choose **Change Mass Options**.
 - b. In the Change Mass Options dialog box, select the **Set Molecular Weight** check box. Then, select **[M+H]⁺1**.

The calculated molecular weight for the compound in row 93 returns to 320.89501 Da.

Go to the next topic to “[Add downstream nodes and reprocess the analysis.](#)”

After you update the mass options for the compounds of interest, add identification nodes to the processing workflow and reprocess the analysis.

Note In the previous procedure “[Modify the mass options for a compound](#)” on page 17, you changed the molecular weight for the compound in row 93 (sorted by RT in ascending order) and then reversed the change.

To add downstream nodes to the processing workflow and reprocess the analysis, follow these steps:

1. [Open the core analysis](#)
2. [Add downstream nodes to the processing workflow](#)
3. [Select the databases for the Search ChemSpider node](#)
4. [Submit the new analysis to the job queue](#)

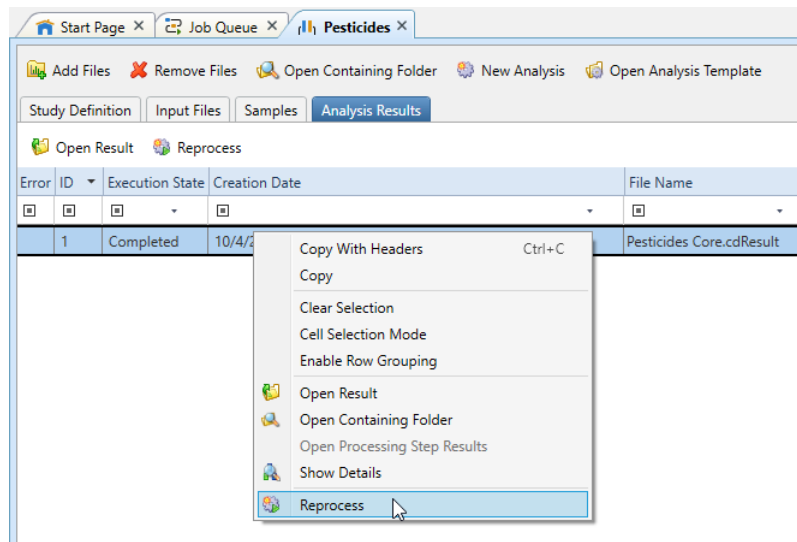
**Add
downstream
nodes and
reprocess the
analysis**

Open the core analysis

v To open an analysis

1. Open the Pesticides study page by doing one of the following:
 - If the Pesticides study is open, but it is not the active page, click the **Pesticides** tab.
 - If the Pesticides study is closed, doing one of the following to open it:
 - Click the **Pesticides** link under Recent Studies on the Start Page.
 - Choose **File > Open Study** from the application menu bar, locate the Pesticides folder in the Studies directory, select the **Pesticides.cdStudy** file, and click **Open**.
2. Click the **Analysis Results** tab.
3. On the Analysis Result page, right-click the **Pesticides Core** analysis and choose **Reprocess** (Figure 17).

Figure 17. Shortcut menu for the Analysis Results page

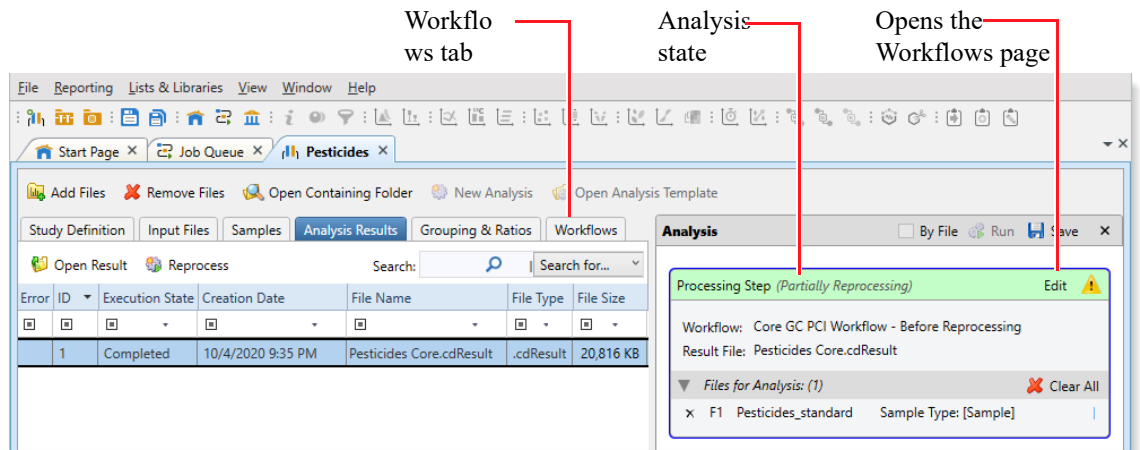


Clicking Reprocess after selecting a file on the Analysis Results page does the following:

- Reopens the Analysis view if it was closed and sets the status of the analysis to Partially Reprocessing and ready to be edited.
- Populates the Files for Analysis area with the input files from the selected analysis.
- Populates the Workflows page with the processing workflow from the selected analysis.

The Run button remains unavailable until you edit the analysis. If you change any of the parameter settings in the core workflow nodes, remove or add input files in the Files for Analysis area, or both, the analysis state changes to Fully Reprocessing.

Figure 18. Analysis set to the partially reprocessing state



Add downstream nodes to the processing workflow

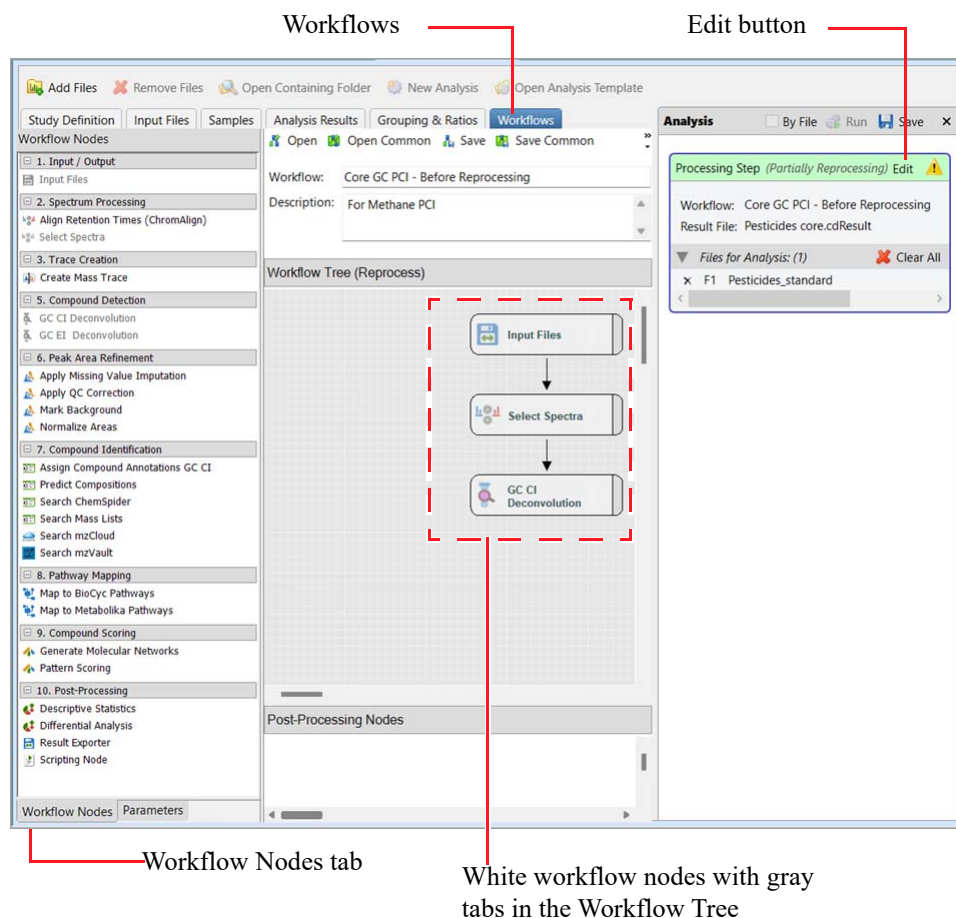
v **To add downstream nodes to the core processing workflow**

1. Open the Workflows page by doing one of the following:
 - Click **Edit** in the Processing Step status bar of the Analysis view.
 - Click the **Workflows** tab.

The processing workflow appears in the Workflow Tree (Reprocess) area. All the nodes are white with a gray tab, which indicates that they are not set for reprocessing.

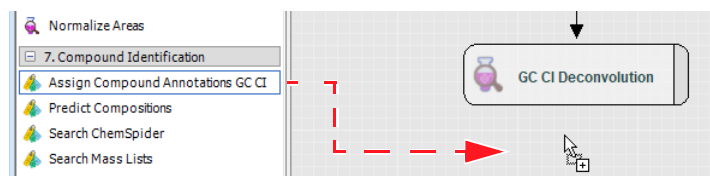
2. If the Workflow Nodes pane is not displayed, click the **Workflow Nodes** tab at the bottom left of the Workflows page to display it.

Figure 19. Unedited analysis set for reprocessing



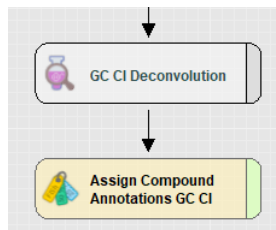
3. Add the Assign Compound Annotations GC CI node to the processing workflow as follows:
 - a. In the Workflow Nodes pane, select the **Assign Compound Annotations GC CI** node in the Compound Identification area and drag it to the Workflow Tree (Reprocess) pane.

Figure 20. Dragging the node from the Workflow Nodes pane to the Workflow Tree pane



- b. Release the mouse button.

The GC CI Deconvolution node automatically connects to the Assign Compound Annotation GC CI node.



Tip If you click a node to select it after you release the mouse button, the Workflow Nodes page closes, and the parameters page for the selected node opens. To reopen the Workflow Nodes page, click the **Workflow Nodes** tab.

4. Add the **Predict Compositions** node and the **Search ChemSpider** node to the processing workflow by dragging them one-by-one to the Workflow Tree (Reprocess) pane.

The GC CI Deconvolution node automatically connects to these nodes.

5. In the Workflows page command bar, click **Auto Layout**.

The application optimizes the layout of the workflow nodes in the Workflow Tree (Reprocess) pane.

Note Adding workflow nodes to the processing workflow for the current analysis does not affect the original processing workflow template. For information about creating and saving your own custom processing workflow templates, refer to the Help or the *Compound Discoverer User Guide for GC Studies*.

Select the databases for the Search ChemSpider node

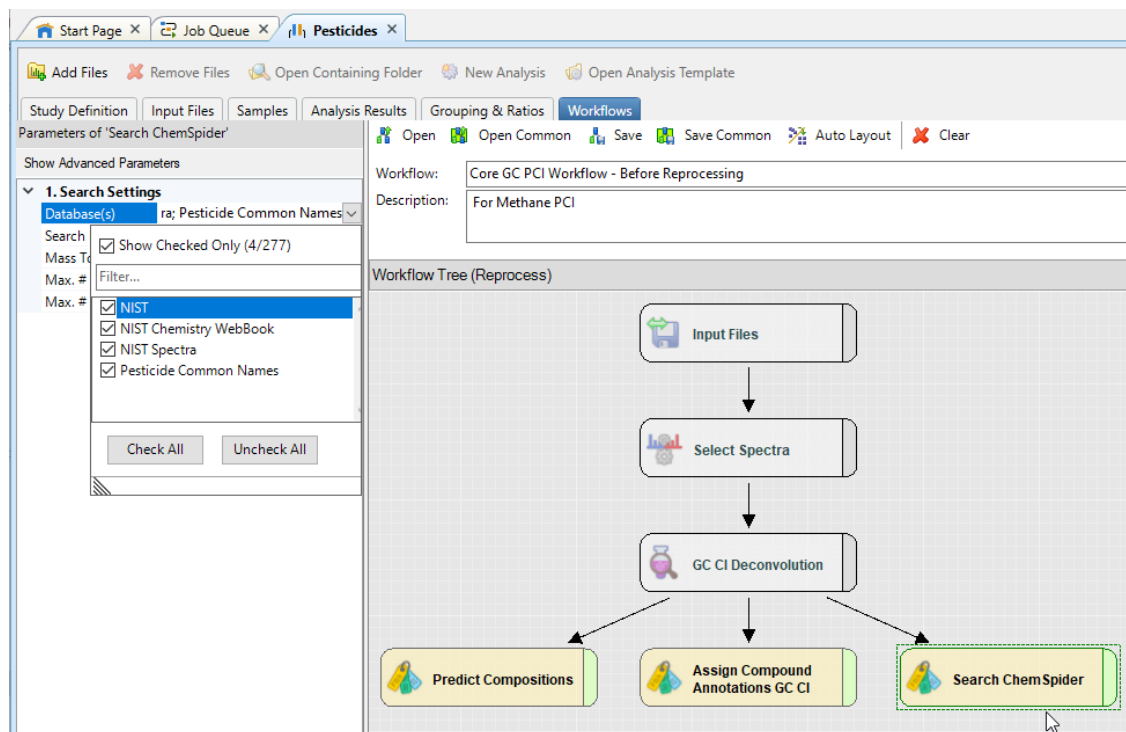
- v **To select the databases to search in the ChemSpider node**

1. In the Workflow Tree (Reprocess) pane, select the **Search ChemSpider** node.

The Parameters of Search ChemSpider pane appears at the left. By default, the KEGG database is selected.

2. Click the **Databases** box to make the down arrow available. Then, click the down arrow to open the database list (Figure 21).
3. Clear the check box for the KEGG database and select the check boxes for the following databases:
 - NIST
 - NIST Spectra
 - NIST Chemistry WebBook
 - Pesticide Common Names
4. Select the **Show Checked Only** check box and verify that only the four databases shown in Figure 21 are selected.

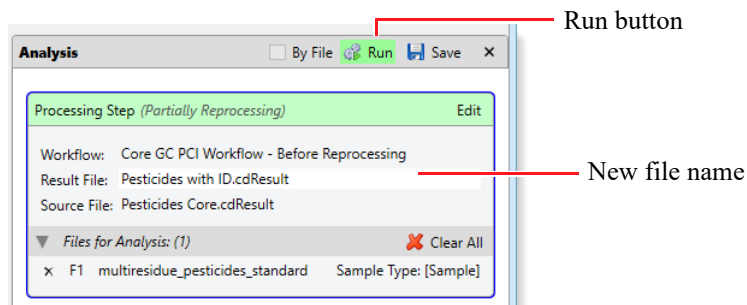
Figure 21. Modified processing workflow and selected ChemSpider databases



Submit the new analysis to the job queue

v To submit the new analysis to the job queue

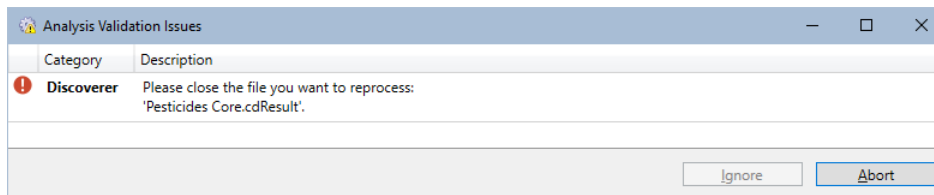
1. In the Analysis view, rename the new result file to **Pesticides with ID**.



2. Close the Pesticides Core.cdResult file if it is open.
3. Click **Run**.

If you did not close the original result file before clicking Run, the application prompts you to close it.

4. If the Analysis Validation Issues dialog box opens, click **Abort**, close the original result file, and click **Run** again.



The Job Queue opens, and the analysis takes less than two minutes to run to completion.

Go to the next topic to “[Open the reprocessed result file and review the compound annotations.](#)”

Open the reprocessed result file and review the compound annotations

Open the reprocessed result file

Review the compound annotations

To open the reprocessed result file and review the compound annotations, follow these steps:

1. [Open the reprocessed result file](#)
2. [Review the compound annotations](#)

Note The new annotations include the predicted formulas and any hits from the selected ChemSpider databases.

v To open the new result file after you reprocess the initial analysis

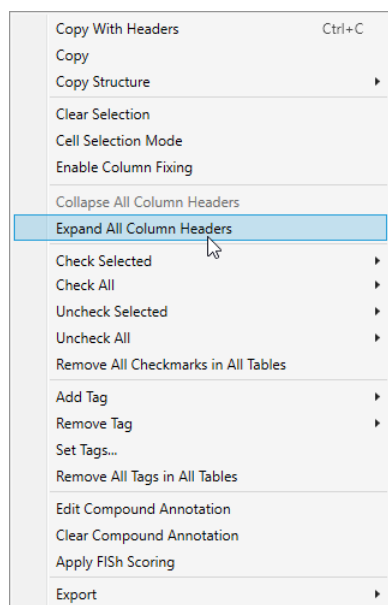
On the Job Queue, double-click the completed run for the Pesticides with ID.cdResult file. Or, open the Pesticides with ID.cdResult file provided in the example study folder.

Tip To make sure that the result page is set to the default layout, choose **Window > Reset Layout** from the application menu bar.

v To review the compound annotations

1. On the result page, close the Chromatograms view by clicking the close icon in its upper-right corner. The Mass Spectrum view expands to the full width of the result page.
2. Right-click the GC CI Compounds table and choose **Expand All Column Headers**.

Figure 22. Shortcut menu for the GC CI Compound table



The column headings for columns with subcolumns expand to display the vertical subcolumn headings. See [Figure 23](#).

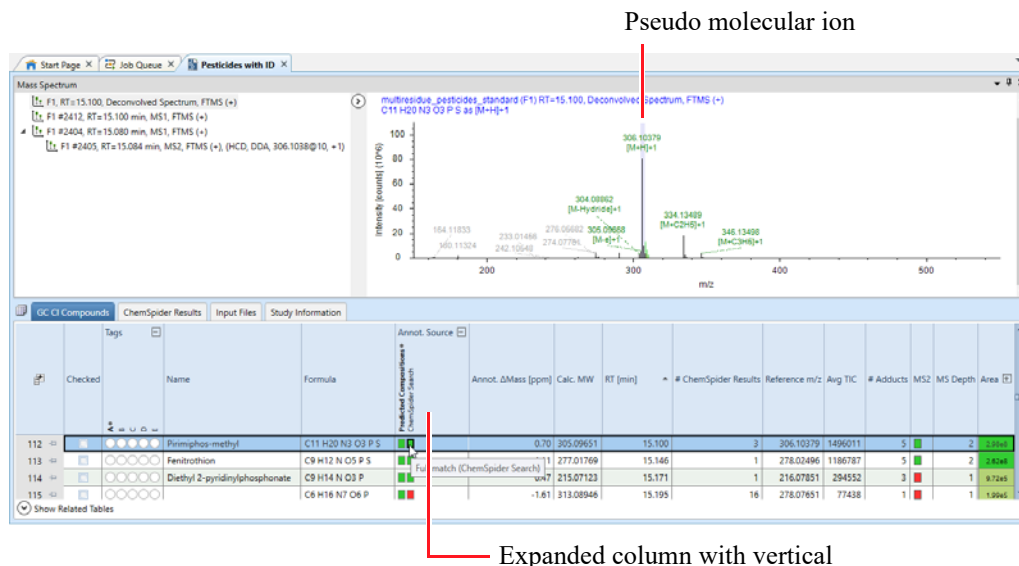
The Annotation Source column contains two subcolumns—Predicted Compositions and ChemSpider Search.

The status boxes indicate the following:

- Gray—No results
- Red—No match
- Green—Full match

3. Select **row 112** (RT = 15.100 min, identified as pirimiphose-methyl). See [Figure 23](#).

Figure 23. Mass Spectrum view showing the deconvolved spectrum for pirimiphos-methyl in row 112



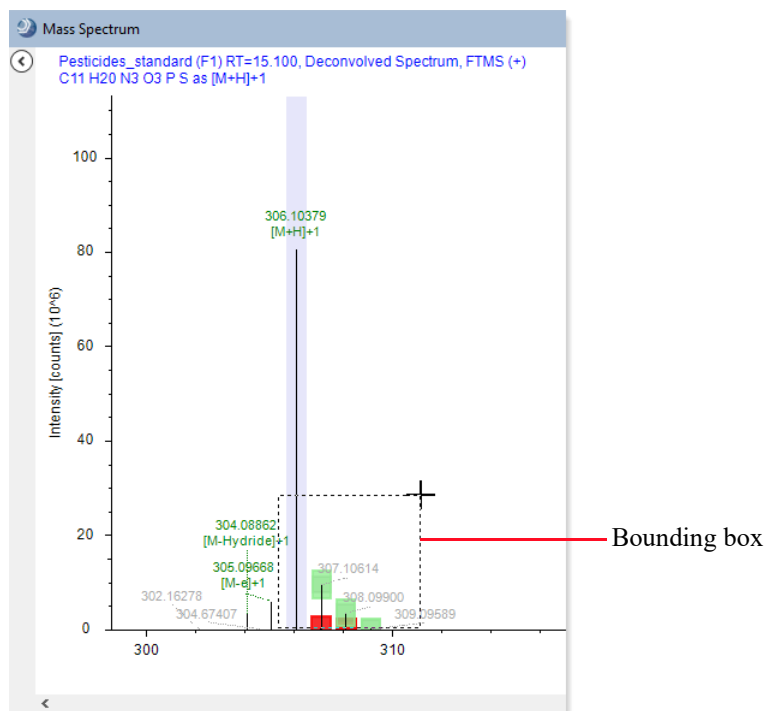
- In the Mass Spectrum view, review the annotated spectral peaks and zoom in on the annotated ion cluster for the pseudo molecular ion (m/z 306 to m/z 310).

Tip To zoom in on a section of the spectrum plot, do the following:

- Drag the cursor diagonally across the section.

A dashed bounding box appears.

- Release the mouse button.



The application annotates the mass spectrum with rectangles that indicate the theoretical isotope pattern for the compound's predicted formula (C11 H20 N3 O3 P S). See [Table 2](#) and [Figure 24](#).

Table 2. Color coding for the centroids in an MS1 spectrum for a compound with a formula annotation

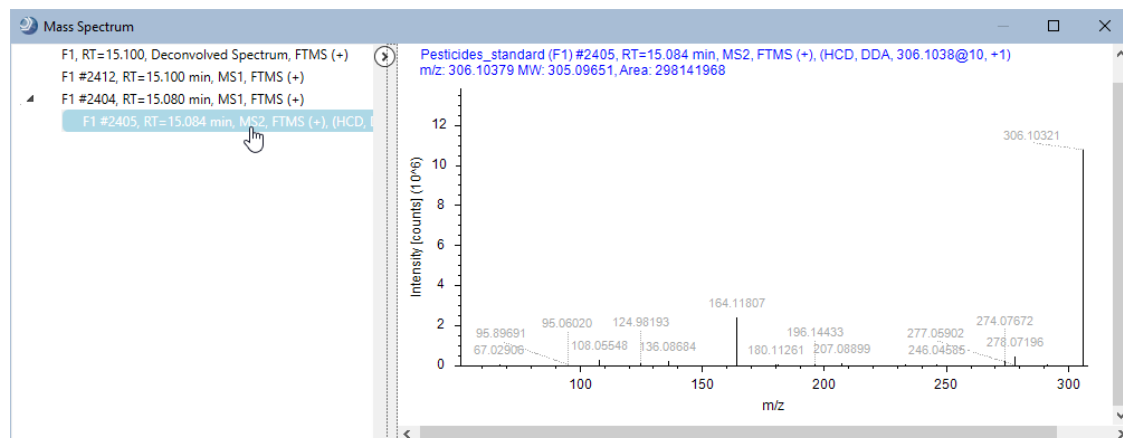
Color	Meaning
(L) Lavender	The labeled centroid matches the highest intensity peak in the isotope pattern for the adduct ion's predicted formula. Typically, the peak with the highest intensity in the isotope pattern is the monoisotopic mass for the adduct ion.
(G) Green	The labeled centroid matches the delta mass and the relative intensity of the theoretical isotope pattern within the specified tolerances.
(R) Red	The expected centroid for this m/z value is missing or its intensity does not fall within the tolerance range for the theoretical isotope pattern.

Figure 24. Deconvolved spectrum for the compound in row 112



3. To reset the zoom level, right-click the view and choose **Undo All Zoom/Pan**.
4. In the spectral tree pane, select the MS2 spectrum for the compound in row 112.

The current annotations include only the m/z values of the centroids.



Make structure proposals for a compound and run FISH Scoring

You can make structure proposals for any compound and run FISH scoring on any compound with MS2 data and a proposed structure.

The following error messages appear in the FISH Scoring Queue view, when you submit compounds with insufficient data.

Incorrect action	Error message
Submit a compound without structure information.	Compound does not provide required structure property.
Submit a compound without MS2 data.	Cannot assign FISH annotation because no fragmentation scans are available.

v To make structure proposals for a compound and run FISH scoring on these proposals

- Sort the GC CI Compounds table by the MS Depth column in descending order. Then, hold down the CTRL key and sort the #ChemSpider Results column in descending order.
- Scroll back to the top of the table.
Pirimiphos-methyl has moved up to row 5 from row 112.
- Select **row 5** (pirimiphos-methyl).
- Click **Show Related Tables** at the bottom left of the result page.
- Click the **ChemSpider Results** tab.
- Select all three rows in the ChemSpider Results table.

Figure 25 shows the GC CI Compounds table sorted by the MS Depth and #ChemSpider Results columns.

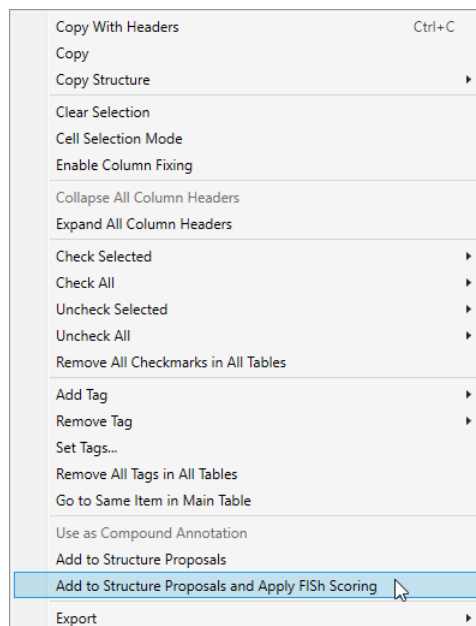
Note To display the columns of interest at the left of the table, column fixing was enabled.

Figure 25. GC CI Compounds table sorted by the MS Depth and #ChemSpider Results columns

GC CI Compounds									
Checked	Tags	Name	Formula	Annot. Source	# ChemSpider Results	MS Depth	RT [min]		
<input type="checkbox"/>	○ ○ ○ ○ ○	Diazinon	C12 H21 N2 O3 P S	■ ■	3	2	12.501		
<input type="checkbox"/>	○ ○ ○ ○ ○	Pirimiphos-methyl	C11 H20 N3 O3 P S	■ ■	3	2	15.100		
<input type="checkbox"/>	○ ○ ○ ○ ○	Pirimiphos-ethyl	C13 H24 N3 O3 P S	■ ■	3	2	16.519		
<input type="checkbox"/>	○ ○ ○ ○ ○	Quinalphos	C12 H15 N2 O3 P S	■ ■	3	2	17.451		

ChemSpider Results							
Checked	Tags	Compound Match	Structure	Name	Formula	Molecular Weight	
<input type="checkbox"/>	○ ○ ○ ○ ○	■		Pirimiphos-methyl	C11 H20 N3 O3 P S	305.09630	
<input type="checkbox"/>	○ ○ ○ ○ ○	■		pyrimitate	C11 H20 N3 O3 P S	305.09630	
<input type="checkbox"/>	○ ○ ○ ○ ○	■		Diisopropyl ((2E)-2-((2E)-1,4-dithiazepan-5-ylidene)acetate)	C11 H20 N3 O3 P S	305.09630	

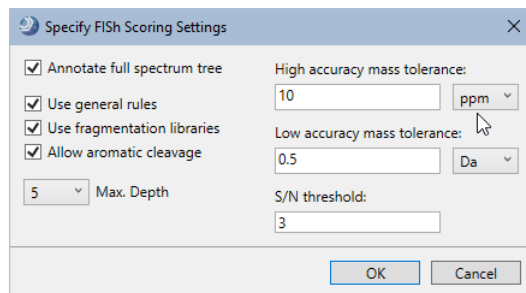
7. To add the ChemSpider results for pirimiphos-methyl to its Structure Proposals table and run FISH scoring on these proposals, do the following:
- With all three rows selected, right-click the table and choose **Add to Structure Proposals and Apply FISH Scoring**.



The Specify FISH Scoring Settings dialog box opens.

- To change the High Accuracy Mass Tolerance setting to 10 ppm, type **10** in the box and select **ppm** from the list at the right. See [Figure 26](#).

Figure 26. Specify FISH Scoring Settings dialog box with a High Accuracy Mass Tolerance setting of 10 ppm



- Click **OK** to submit the structures to the FISH Scoring Queue.

The FISH Scoring Queue opens to the right of the tabbed pages. See [Figure 27](#).

Note The processing time for FISH scoring increases when you do the following:

- Submit multiple compounds for processing.
- Increase the maximum depth of the analysis. The maximum depth is the maximum number of fragmentation reactions to consider in the fragmentation pathway.
- Submit compounds with complex structures for processing.
- Select the Use Fragmentation Libraries check box for the analysis.

Figure 27. FISh Scoring Queue with three compounds

The screenshot shows a window titled "FISh Scoring Queue" with three entries:

- Entry 1:** Name: Pirimiphos-methyl, MW [Da]: 305.09630, Processing Since: 5 s, State: Processing. The chemical structure is a pyrimidine ring with a dimethyl phosphorothioate group at the 2-position.
- Entry 2:** Name: pyrimitate, MW [Da]: 305.09630, Queued Since: 5 s, State: Waiting. The chemical structure is a pyrimidine ring with a dimethyl phosphorothioate group at the 2-position and a methyl group at the 4-position.
- Entry 3:** Name: Diisopropyl ((2E)-2-((2E)-1,3-thiazol-5-ylidene)acetate), MW [Da]: 305.09630, Queued Since: 5 s, State: Waiting. The chemical structure is a thiazole ring with a dimethyl phosphorothioate group at the 5-position and a diisopropylamino group at the 2-position.

- d. Click the **Structure Proposals** tab and review the FISh Coverage scores for the three ChemSpider results.

Figure 28. Structure Proposals table for pirimiphos-methyl

Hide Related Tables

Structure Proposals | GC CI Compounds per File | Predicted Compositions | ChemSpider Results

	Checked	Tags	Structure	Name	Formula	Molecular Weight	FISh Coverage
1	<input type="checkbox"/>	○○○○○		Pirimiphos-methyl	C11 H20 N3 O3 P S	305.09630	66.67
2	<input type="checkbox"/>	○○○○○		pyrimitate	C11 H20 N3 O3 P S	305.09630	44.44
3	<input type="checkbox"/>	○○○○○		Diisopropyl ((2E)-2-((2E)-1,3-thiazol-5-ylidene)acetate)	C11 H20 N3 O3 P S	305.09630	11.11


FISh Coverage

8. In the Structure Proposals table, right-click row 1 (pirimiphos-methyl) and choose **Structure Proposals > Use As Compound Annotation**.

The screenshot shows a context menu for the Structure Proposals table. The menu items are:

- Structure Proposals
- Export
- Use as Compound Annotation (highlighted)
- Add Structure Proposal
- Edit Structure Proposal
- Delete Structure Proposal
- Apply FISh Scoring to Selected
- Apply FISh Scoring to All

The application annotates pirimiphos-methyl in the GC CI Compounds table with the FISh Coverage score, but you cannot see the score because the FISh Coverage column is hidden.

9. To view the FISH Coverage score in the GC CI Compounds table, do the following:
- Click the **Field Chooser** icon, , at the left of the table above the row numbers.


The Field Chooser dialog box for the GC CI Compounds table opens. You do not need to close the Field Chooser dialog box to continue working in other areas of the user interface.



- Select the **Fish Coverage** check box.

The FISH Coverage column with the score for pirimiphos-methyl appears in the GC CI Compounds table.

Tip The Field Chooser dialog box automatically closes when you make a different tabbed document, other than the current result file, the active document in the application window. Examples of other tabbed documents are the study page, the Start Page, other result files, and so on.

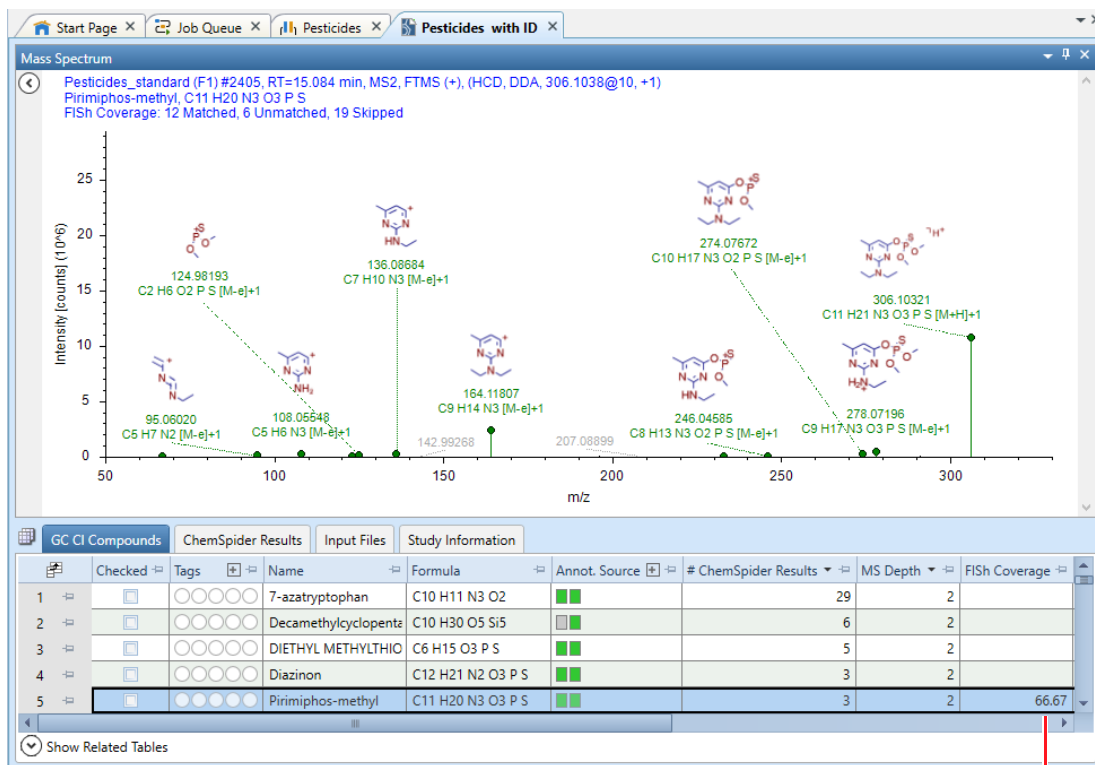
To close the Field Chooser dialog box without leaving the current result page, click the **Close** icon, , in the upper-right corner of the dialog box.

10. In the mass spectrum view, view the annotated FISH fragments in the MS2 spectrum for pirimiphos-methyl. See [Figure 29](#).

The FISH scoring algorithm was able to match 12 of the 18 fragments it expected to find, yielding a score of 66.67. It skipped 19 theoretical fragments that it expected to be below the noise threshold.

[Figure 29](#) shows the annotated MS2 spectrum and the FISH coverage score for pirimiphos-methyl in the GC CI Compounds table.

Figure 29. Annotated MS2 spectrum for pirimiphos-methyl

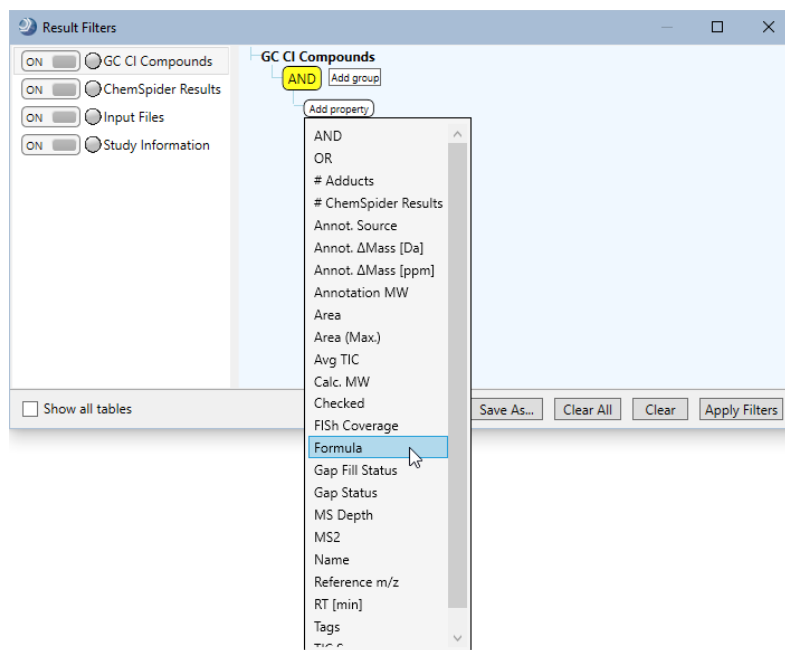


FISH coverage score

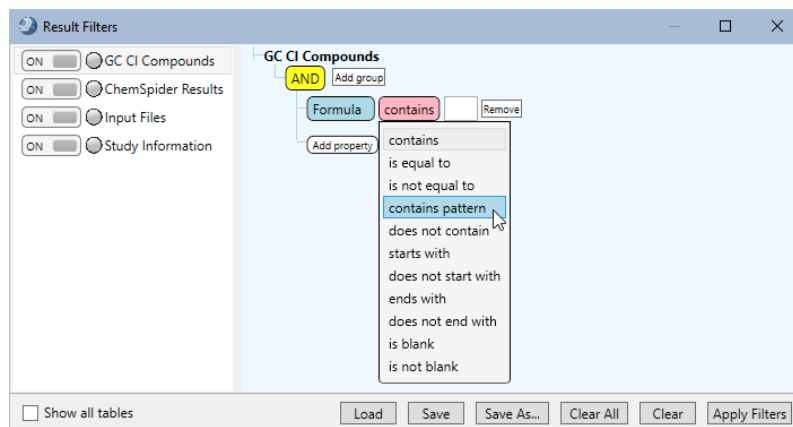
Filter the data to reduce the number of compounds to report

v To reduce the number of compounds to report in the GC CI Compounds table

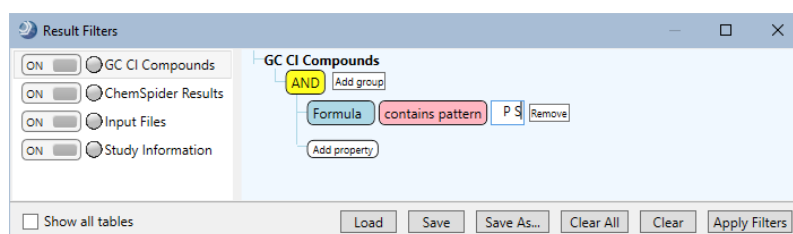
1. From the menu bar, choose **View > Result Filters**.
The Result Filters view opens as a floating window.
2. Add a formula filter for the P S pattern as follows:
 - a. Click **Add Property** and select **Formula** from the dropdown column list.



- b. Click the pink box and select **Contains Pattern** from the dropdown operator list.

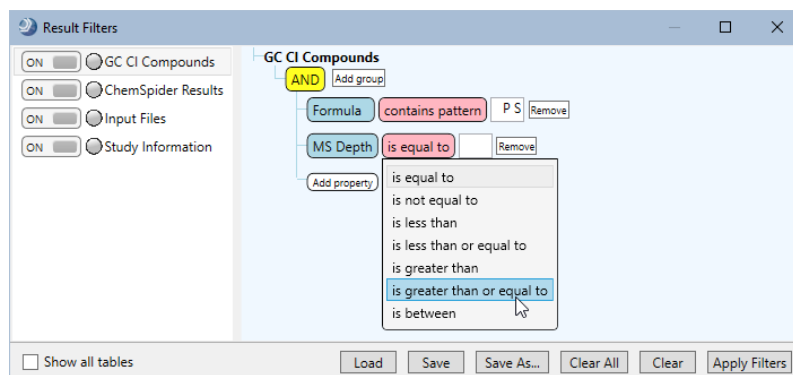


- c. Type **P S** in the value box, with a space between P and S.



3. Add an MS Depth filter as follows:

- a. Click **Add Property** and select **MS Depth** from the dropdown column list.

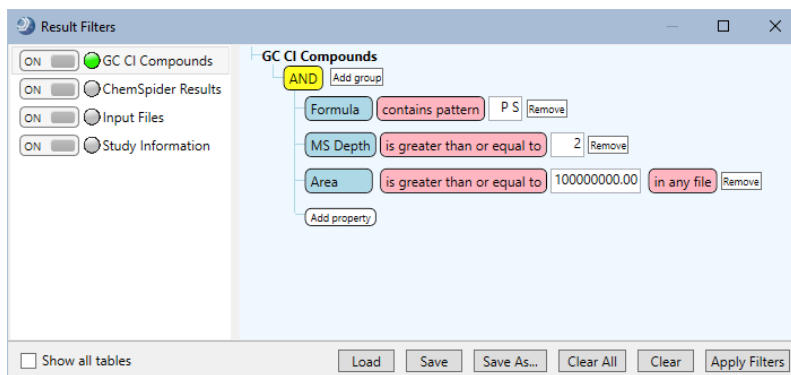


- b. In the value box to the right of the operator, type **2**.

4. Add a filter for the minimum chromatographic peak area for the compound as follows:

- Click **Add Property** and select **Area** from the dropdown column list.
- Click the pink box and select **Is Greater Than or Equal To** from the dropdown operator list.
- Type 1e8 in the value box to the right of the operator.
- Click the pink box to the right of the value box and select **In Any File** from the list.

5. To apply the filters to the GC CI Compounds table, click **Apply Filters**.



Eight compounds remain.

6. Sort the table in descending order by the Area column.

Sorted in descending order by the area

	Checked	Tags	Name	Formula	Annot. Source	MS Depth	FISH Coverage	Area	RT [min]
1	<input type="checkbox"/>	○○○○○	Pirimiphos-methyl	C11 H20 N3 O3 P S	■ ■	2	66.67	2.99e8	15.100
2	<input type="checkbox"/>	○○○○○	Fenitrothion	C9 H12 N O5 P S	■ ■	2		2.62e8	15.146
3	<input type="checkbox"/>	○○○○○	Diazinon	C12 H21 N2 O3 P S	■ ■	2		2.59e8	12.501
4	<input type="checkbox"/>	○○○○○	Pirimiphos-ethyl	C13 H24 N3 O3 P S	■ ■	2		2.18e8	16.519
5	<input type="checkbox"/>	○○○○○	Quinalphos	C12 H15 N2 O3 P S	■ ■	2		2.15e8	17.451
6	<input type="checkbox"/>	○○○○○	8145294	C9 H17 Cl N3 O3 P S	■ ■	2		1.99e8	12.875
7	<input type="checkbox"/>	○○○○○	577209	C14 H20 N3 O5 P S	■ ■	2		1.24e8	26.219
8	<input type="checkbox"/>	○○○○○	Chlorpyrifos	C7 H7 Cl3 N O3 P S	■ ■	2		1.22e8	14.062

7. Close the Result Filters view by clicking the close icon in the upper-right corner.

Go to the next topic to export the deconvoluted spectra for these compounds to an MSP file.

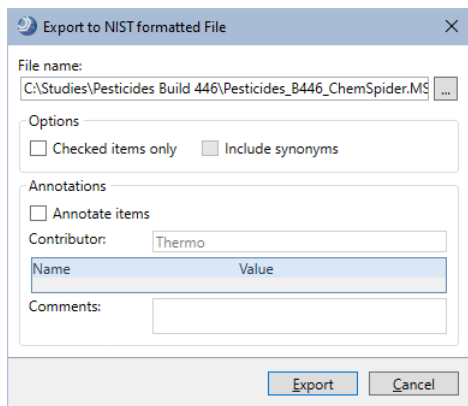
You can export the deconvoluted spectra for compounds in the GC CI Compounds table to an MSP file that you can then open in the NIST MS Search application for curation.

v **To export the deconvoluted spectra for compounds to an MSP file**

1. Filter the compounds table to remove compounds that are not of interest or select the compounds of interest by selecting their check boxes.
2. Right-click the GC CI Compounds table and choose **Export > As NIST MSP File**.

Export the spectral information for various compounds to a NIST MSP file

The Export to NIST Formatted File dialog box opens. The Include Synonyms check box is unavailable for GC CI data.




3. Specify the location and file name for the MSP file and the information you want to export to the MSP file as follows:
 - Click the browse icon to the right of the File Name box. Then, in the Export to NIST MSP dialog box, browse to and select the folder where you want to store the NIST MSP file, name the file, and click **Save**.
 - In the Options area of the Export to NIST Formatted File dialog box, do one of the following:
 - If you filtered the GC CI Compounds table as described in “[Filter the data to reduce the number of compounds to report](#)” on [page 33](#), do not select the Checked items Only check box.
 - If you did not filter the GC CI Compounds table, but you did select compounds in the GC CI Compounds table (by selecting their check boxes in the Checked column, select the Checked Items Only check box.
 - In the Annotations area of the Export to NIST Formatted File dialog box, select the **Annotate** items check box. Then, in the Comments box, type a brief comment, for example, **Testing 123**.
4. Click **Export**.

The application generates a NIST MSP file. You can open the file in the NIST MS Search application.

If you are not working with the Pesticides with ID.cdResult file, filter the GC CI Compounds table to reduce the number of compounds to report.

v **To print a report using a defined report template**

1. Filter the GC CI Compounds table as described in “[Filter the data to reduce the number of compounds to report](#),” if you have not already done so.
2. Sort the table by the RT [min] column in ascending order.
3. From the application menu bar, choose **Reporting > Create Report**. Or, click the **Create Report** icon, , in the application toolbar.

The Open Report Design Template dialog box opens.

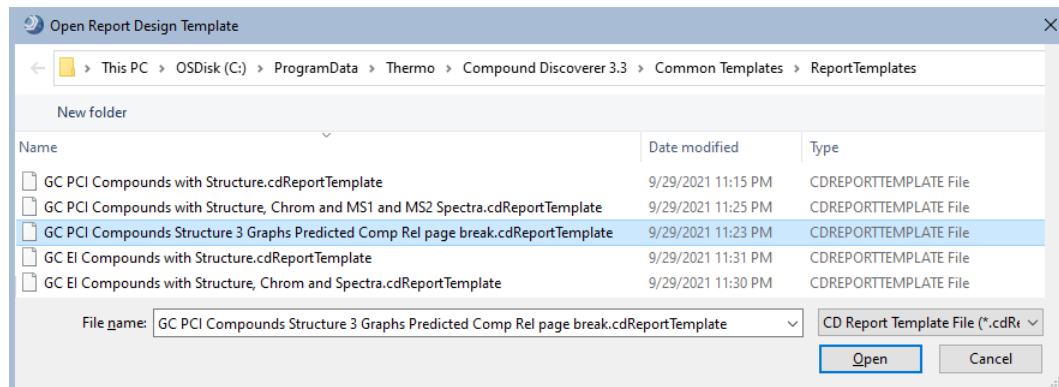
4. Select the following report template:

GC CI Compounds Structure 3 Graphs Predicted Comp Rel Page Break

**Print a
report by
using a
defined
template**

Note The application resolves the data in the result file only when the result file includes a Predicted Compositions table.

Figure 30. Report templates



5. Click **Open**.

The report resolution page opens and the application resolves the data with the template.

Figure 31. Resolved report

Print icon Create Report

Structure	Name	Formula	Calc. MW	RT [min]	# Adducts	MS2	Area
	Pirimiphos-methyl	C11H20N3O3P5	305.09651	15.100	5		

Compound Match	Formula	ΔMass [ppm]	SFit [%]	Pattern Cov. [%]	MS Cov. [%]	MSMS Cov. [%]
	C11H20N3O3P5	0.70	30	97.75	99.25	98.70

6. On the report resolution page, click the **Print** icon, , in the toolbar to print the report.
The Print dialog box opens.

7. In the Print dialog box, do the following:
 - a. Select the appropriate printer and the page range that you want to print.

Note The report templates that come with the application default to printing on A4 paper.

- b. If you are not printing on A4 paper, change the printer setting.

Note If you are printing on letter size paper, which is shorter than A4 paper, and you do not change the printer setting, the report prints on two pages per compound.

- c. Click **OK** to print the report.

Trademarks

All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified.

NIST is a registered trademark of the National Institute of Standards and Technology in the United States.

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