

MassHunter Profinder Software

Quick Start Guide

For Research Use Only. Not for use in diagnostic procedures.

Getting Started	11
User Interface	12
Basic Profinder Workflow	22
Add sample files to your project	24
Align chromatograms	28
Edit and run your feature extraction algorithm method	31
Export your Profinder features	33
Save your Profinder method	36
Save your Profinder project	37
Feature Extraction Workflow Algorithms	38
Monitoring your computer while running Profinder	45
Profinder Installation	46

What is MassHunter Profinder?

Agilent MassHunter Profinder software is a stand-alone program, optimized for batch feature extraction from TOF and Q-TOF based data files and GC/MSD EI data files. Profinder is also part of the MassHunter VistaFlux software. Feature extraction, combined with chromatographic alignment across multiple data files, is a critical step in the peak finding and data reduction workflow. The appearance of both false positive and false negative features is minimized by “binning” the features in the chromatographic time domain. Profinder reduces your acquired data size and complexity through the removal of redundant and non-specific information by identifying the



What's new in B.08.00 Service Pack 2?

important variables (features) associated with the data. Careful feature extraction yields a smaller data set that is more easily processed without any compromise in the information quality.

Profinder is optimized to not only extract features from large data sets but also provides you with an intuitive user interface to inspect and review each feature across the files associated with your data set. With Profinder you can review and compare extracted ion chromatograms, mass spectral data, and isotopologues associated with each feature. You can save your feature extraction progress to a project to review, adjust, and reprocess at a later time.

The feature files generated by Profinder (PFA files and CEF files) can be imported into Agilent Mass Profiler Professional (MPP), a chemometrics platform used in a differential analysis to determine relationships among two or more sample groups and variables. Profinder Archive files (PFA files) containing extracted compound features and sample group information are imported into Omix Premium, a visualization platform, to view your data in combination with biochemical networks, including isotopologue results as part of a metabolic flux analysis.

What's new in B.08.00 Service Pack 2?

- The time alignment algorithm can correct for retention time drift across several data files. All .d files in a given Profinder project can be time-aligned to either a reference data file (*.d) or a database of retention time standard compounds. You cannot use time alignment on GC/MS data files.
- Multiple sample groups are now supported. The sample groupings can then be transferred directly to Mass Profiler Professional (MPP) by exporting a PFA file.
- You can import and export Sample Grouping Files in the **Add/Remove Sample Files** dialog box.
- Windows 10 is supported.

What's new in B.08.00?

- Isotopologue feature extraction support is added for metabolic flux analysis in MassHunter VistaFlux Software.
- Profinder Archive export provides a single-file transfer of project data to Mass Profiler Professional (MPP) and Omix Premium.
- Untargeted and recursive feature extraction is supported for GC/Q-TOF and GC/MSD.
- Targeted All Ions fragment confirmation is supported for LC/TOF, LC/Q-TOF and GC/Q-TOF data files.
- Additional feature extraction algorithms are multi-threaded and memory optimized to support large projects and datasets.

Features of Profinder

- Open all of the files associated with your Agilent TOF and Q-TOF MS data set (sample files) in a single selection process.
- Create methods for your feature extraction and chromatographic alignment by following a guided step-by-step wizard for the selected algorithm.
- Align chromatograms without standards (by specifying a reference data file) or with standards (by specifying a CSV file).
- Choose from five feature extraction algorithms:
 - Batch Molecular Feature Extraction
 - Batch Recursive Feature Extraction (small molecules/peptides)
 - Batch Recursive Feature Extraction (large molecules)
 - Batch Targeted Feature Extraction
 - Batch Isotopologue Extraction
- Extract the features from all of the data files in your data set by running a single algorithm.

Where is Profinder used in your experiment?

- Automatically generate an extracted ion chromatogram (EIC) and extracted compound chromatogram (ECC) for each feature found in each data file.
- Review the chromatographic integration/alignment, mass spectra, and isotopologues using an intuitive graphical interface.
- Curate the features in your data post-analysis.
- Export your extracted features via three file formats:
 - PFA and compound exchange format (CEF) are file formats that are used to exchange data between Agilent MassHunter software. PFA files, for example, are imported into MassHunter Mass Profiler Professional for statistical analysis.
 - Profinder Archive (PFA) is a binary file format used to transfer your features to Omix Premium for metabolic flux analysis and visualization within the context of biochemical networks.
 - Comma separated values (CSV) is an ASCII text file that is readily imported into text editors and spreadsheets.
- Save your feature extraction parameters as Profinder methods.
- Save your feature extraction progress as a Profinder project file and open the project at a later time to review, adjust, and reprocess your data files.

Where is Profinder used in your experiment?

Profinder is used to extract the features from LC/TOF, LC/Q-TOF, GC/Q-TOF, and GC/MSD data files you acquired from your experimental samples. An untargeted differential analysis experiment may include the steps shown below.

- 1 Prepare for your experiment
- 2 Acquire your data - Data Acquisition
- 3 Extract the spectral features - **Profinder**
- 4 Perform a differential analysis - Mass Profiler Professional
- 5 Identify features in your data - ID Browser

- 6 Perform advanced analysis operations - Mass Profiler Professional
- 7 Visualize your Profinder results within biochemical networks - Omix Premium

Figure 1 shows the steps and Agilent tools that are used in an untargeted differential analysis and visualization in biochemical networks. Some of the feature finding steps performed by Profinder can alternatively be accomplished using a sequence of steps involving MassHunter Qualitative Analysis, DA Reprocessor, and MPP as described in the *Agilent Metabolomics Workflow - Discovery Workflow Guide*.

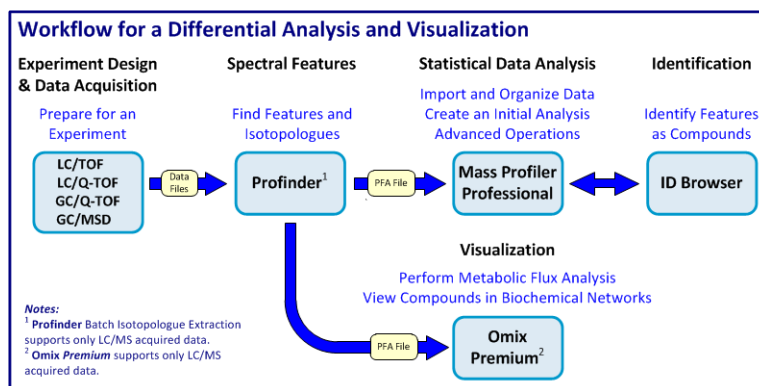


Figure 1 MassHunter Profinder is an optimized feature extraction tool.

Time Alignment Algorithms Overview

Introduction

Chemometrics is the use of statistical methods to perform differential analysis on analytical chemistry data, to correlate significant changes in chemical component signals with one or more experimental parameters. Chemometrics on mass spectrometry data requires that the chemical signal be extracted from raw data through a process called feature extraction. Such feature extraction can be performed as a targeted extraction (for instance, the ion signature for a given

compound is known and the data is mined specifically for that signature) or an untargeted extraction (i.e. ion signals that correlate via the dimensions of retention time and neutral mass (LC/MS and CE/MS) or spectral similarity (GC/MS) are assumed to be from the same compound or compounds). In turn, these compound features are then compared across multiple samples for statistical comparison of compound feature abundances.

Whether the compound feature extraction process is targeted or untargeted, chemometric analysis requires that compound features be binned or grouped together with equivalent compound features in other samples. In targeted analysis where the compound signal is uniquely identified, the compound identity itself is sufficient to bin compounds across multiple samples. However, if the feature extraction is untargeted, binning must be performed within an acceptable error tolerance in measured mass and retention time or spectral similarity (if chromatography is used as in LC/MS or GC/MS).

Binning of mass spec data coupled with liquid chromatography (LC/MS) or capillary electrophoresis (CE/MS) can be performed by binning features by observed neutral mass and retention time (or migration time in CE) with error tolerances in both dimensions. For GC/MS data generated by electron ionization (EI) or other hard ionization methods, dot products are commonly used to bin similar features in the spectral dimension.

With modern mass spectrometers, drift in the mass precision is usually well controlled and the expected instrumentation error model is well characterized, such that it is straightforward to specify error tolerances for mass measurement. However, the chromatographic dimension is much less controlled. For example, LC retention time is dependent on sample complexity and the robustness of the separation method. For example, retention time (RT) shifts can happen as separation columns are aging or need to be switched. Also, the commonly used hydrophilic interaction chromatography (HILIC) separation method is prone to retention time drift. Additionally, CE migration times will drift dramatically due to variations in field strength, migration length and/or electroosmotic flow. [3] The

poor predictability of the retention time shifts in both methods leads to difficulty in specifying appropriate error tolerances for feature binning, especially with isomers that closely coelute (such as leucine/isoleucine). If the tolerance is set too tightly, features outside of the bin window will not be binned together. On the other hand, if the tolerance is set too loosely, isomers may be binned together mistakenly, introducing downstream statistical error during profiling.

While some improvement in run-to-run reproducibility can be made by optimization of the LC or CE separation itself, the raw mass spec data can also be transformed as a post-processing step, in a procedure commonly known as "chromatogram warping" or time alignment.

Time Alignment Methods

Several common approaches are available to achieve retention or migration time alignment through data transformation. One such commonly applied approach is the ordered bijective interpolated warping algorithm (commonly known as "obiwarp"). In this approach, dynamic programming methods are used to calculate a distance score between each mass spectrum across pairs of sample data files. An optimal path is identified that represents the time correction function. However, the approach can be quite computationally intensive and require optimization of parameters that do not translate well to physicochemical properties or instrument effects.

Two alternative methods are implemented in the Agilent MassHunter Profinder feature extraction software (referred to as "Profinder"). One is a Legendre polynomial curve fitting, which can be applied generically to any separation method, including LC and CE. The polynomial method uses a 5th order Legendre polynomial, with a damping matrix applied to the 2nd through 4th order terms. The resulting separation time curve is estimated by least squares approximation using the difference between retention/migration time values from a list of reference compound features and their corresponding measured retention/migration time values.

The second method is a CE specific method based on first principles of capillary electrophoresis. Both methods can be used in untargeted and targeted approaches.

Determination of the Time Correction Function

Profinder has two ways to specify the compound feature list for curve fitting. The list of retention/migration times can be provided as a database containing a list of known retention/migration time standards, which are subsequently feature extracted. Alternatively, untargeted feature extraction can be performed on the samples and a reference data file, which can be a separate file or one of the samples in the batch. Subsequently, a list of untargeted features that are commonly found across all data files can serve as the reference compound features for curve fitting. In either approach, the time correction function models the continuous retention/migration time function for a given data file.

The time correction function is first determined for each data file. In the case where a reference data file is specified, the polynomial or CE specific correction function is fit to the data. In the case where a reference database is specified, the list of data points to be used in the curve fit is restricted to those found in the reference database. In both cases, the points used for curve fitting are limited to those that are present in both the reference and in the data file to be corrected.

When using time alignment standards in a reference database, the following requirements should be fulfilled for best results:

- All time alignment standards should be detectable in all data files.
- Time alignment standards should not have isomers in the data files.
- Time alignment standards should cover the entire time range intended for correction.

For more information, see “Pherogram normalization in capillary electrophoresis and micellar electrokinetic chromatography analyses in cases of sample matrix-induced migration time shifts” in the Journal of Chromatography B, vol. 770 (2002): 45-51.

Where to Find More Information

Online Help

Press F1 To get more information about a window or dialog box, place the cursor on the window or dialog box of interest and press **F1**.

Help Menu Click **Help > Contents** to access the contents of the Profinder Help including wizards, basic tasks, user interface, and reference information.



Click **Help** for information specific to wizards.

Documents

Mass Profiler Professional

- Agilent G3835AA MassHunter Mass Profiler Professional Software - Quick Start Guide
- Agilent G3835AA MassHunter Mass Profiler Professional Software - Familiarization Guide
- Agilent G3835AA MassHunter Mass Profiler Professional Software - Application Guide
- Agilent MassHunter Mass Profiler Professional User Manual

Workflow Guides and Overviews

- MassHunter VistaFlux Software - Workflow Guide
- Agilent Metabolomics Workflow - Discovery Workflow Guide
- Agilent Metabolomics Workflow - Discovery Workflow
- Integrated Biology with Agilent Mass Profiler Professional - Workflow Guide
- Integrated Biology with Agilent Mass Profiler Professional - Workflow Guide Overview
- Class Prediction with Agilent Mass Profiler Professional Workflow Guide

Where to Find More Information

- Class Prediction with Agilent Mass Profiler Professional Workflow Overview

MassHunter Qualitative and Quantitative Analysis

- MassHunter Qualitative Analysis Familiarization Guide for LC/MS
- MassHunter Qualitative Analysis eFamiliarization Guide for Q-TOF
- MassHunter Quantitative Analysis Familiarization Guide
- MassHunter Quantitative Analysis eFamiliarization Guide
- MassHunter Data Analysis Resource DVD

Training


Road Show Slide Presentation Metabolome Analysis - From Sample Prep Through Data Analysis (Metabolomics Road Show, March 2011)

Training Courses Visit www.chem.agilent.com to view a listing of training courses for MPP.

Getting Started

How do I get started?

This *Quick Start Guide* helps you install Profinder, review the Profinder user interface, and perform feature extraction using your data.

- 1 Install Profinder. Follow the instructions in “[Profinder Installation](#)” on page 46.
- 2 Launch Profinder. Double-click the **Profinder** icon  located on your desktop, or click **Start > All Programs > Agilent > MassHunter Workstation > Profinder B.08.00**.
- 3 Review the following three sections in this *Quick Start Guide*:
 - “[User Interface](#)” on page 12
 - “[Basic Profinder Workflow](#)” on page 22
 - “[Feature Extraction Workflow Algorithms](#)” on page 38
- 4 Review the *Agilent Metabolomics Workflow - Discovery Workflow Guide*. This guide helps you improve your results by covering topics that help you:
 - Prepare for an experiment
 - Explain the parameters related to finding features
 - Perform an initial differential analysis using MPP
- 5 Extract the features in your samples following the example Basic Profinder Workflow (see “[Basic Profinder Workflow](#)” on page 22).

User Interface

A feature may be referred to as a compound, molecular feature, element, or entity during the various steps of analysis using Agilent MassHunter software.

Help and detailed information regarding the various parameters and statistical treatments are available when you press **F1** or click **Help > Contents** from the menu bar.

Main Functional Areas

The main Profinder window consists of three parts: (1) the Menu Bar, (2) the Toolbar, and (3) the Main Window. The main functional areas are shown in Figure 2.

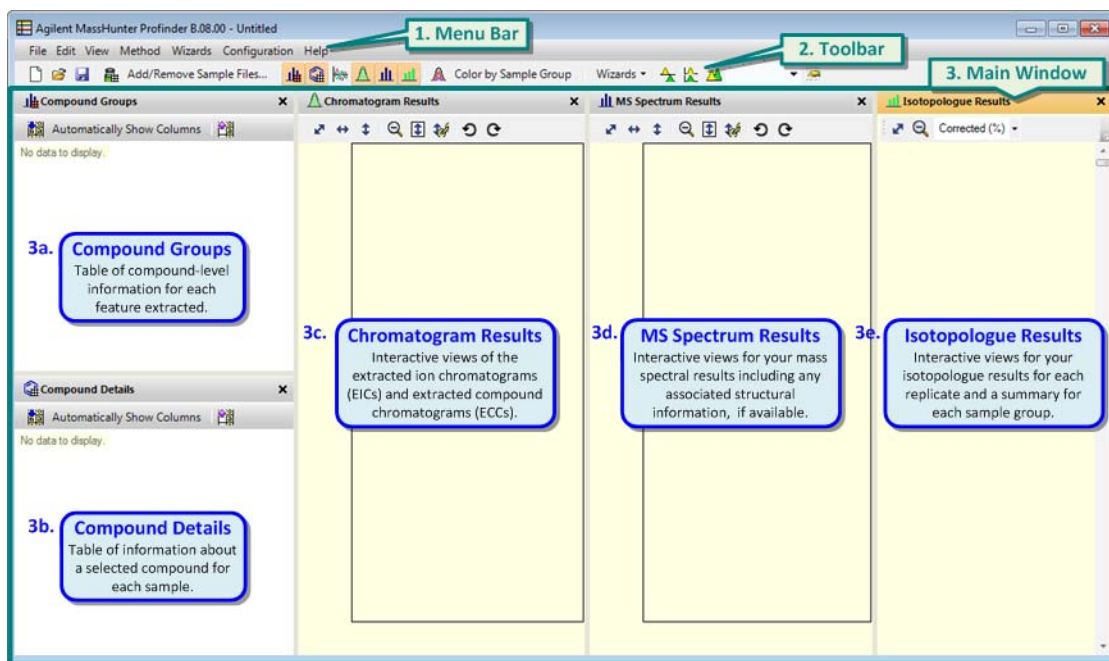


Figure 2 The main functional areas of Profinder as viewed before you begin a project.

1. Menu Bar

The menu bar (Figure 3) provides actions that are used for managing your projects, methods, display, and extracting features.

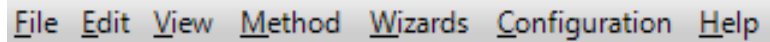





Figure 3 Menu bar


2. Toolbar

The toolbar is located below the menu bar and contains five groups of buttons for commonly performed tasks:







Project New project, Open project, and Save project

Button	Equivalent Command
	File > New Project
	File > Open Project
	File > Save Project


Samples Add sample files to your new or current project.

Button	Equivalent Command
	File > Add/Remove Sample Files


Main Window Display or hide the various tables and results generated by Profinder, so you can increase the available display area for your review.

Button	Equivalent Command
	View > Compound Groups
	View > Compound Details
	View > Sample Chromatograms
	View > Chromatogram Results
	View > MS Spectrum Results
	View > Isotopologue Results






Feature Coloring Toggle the feature coloring by sample group.

Button	Equivalent Command
	Edit > Color by Sample Group

Extraction Select the feature selection algorithm you want to review, edit, and apply to your data set.

Button	Equivalent Command
	Wizards > Batch Molecular Feature Extraction Wizards > Batch Recursive Feature Extraction (small molecules / peptides) Wizards > Batch Recursive Feature Extraction (large molecules) Wizards > Batch Target Feature Extraction Wizards > Batch Isotopologue Extraction

Results Modes Select the display mode and display options to use in your results windows.

Button	Equivalent Command
	List mode, no equivalent command
	Sample group mode, no equivalent command
	Overlaid mode, no equivalent command
	Maximum number of chromatograms or spectra to display, no equivalent command
	Configuration > Plot Display Options

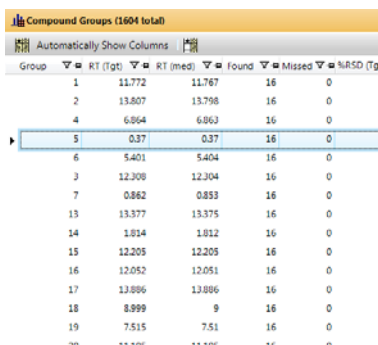
3. Main Window

The main window, see [Figure 2](#) on page 12, is further divided into up to five windows – (3a) Compound Groups, (3b) Compound Details, (3c) Chromatogram Results, (3d) MS Spectrum Results, and (3e) Isotopologue Results that are used to review the results from applying the feature extraction method to your data set. Each window can be floated independently to any location and size on your computer display or arranged to your preference within the main window. The various windows are described in the following pages.

The Sample Chromatograms window is used to align the chromatograms.

User Interface

Main Functional Areas



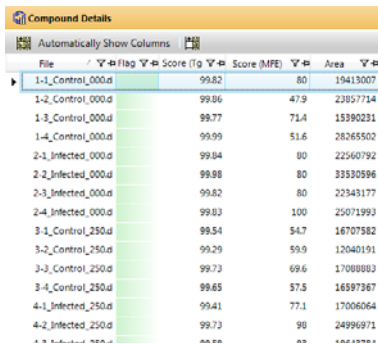
Group	RT (tgt)	RT (med)	Found	Missed	%MSD (Tg)
1	11.772	11.767	16	0	
2	13.807	13.798	16	0	
4	6.864	6.863	16	0	
5	0.37	0.37	16	0	
6	5.401	5.404	16	0	
3	12.308	12.304	16	0	
7	0.862	0.853	16	0	
13	13.377	13.375	16	0	
14	1.814	1.812	16	0	
15	12.205	12.205	16	0	
16	12.052	12.051	16	0	
17	13.886	13.886	16	0	
18	8.999	9	16	0	
19	7.515	7.51	16	0	

Compound Groups The data presented in Compound Groups is organized as a list of all of your extracted feature data averaged and summarized across all of the data files in your project.

The Compound Groups window shows a table of compound-level information for each feature extracted from at least one data file, if the data was extracted using Batch Molecular Feature Extraction, or for all targeted features, if the data was extracted using Batch Targeted Feature Extraction or Batch Isotopologue Extraction. Semi-quantitative information is shown as the average value for the feature across all of the files where the feature was found.

A *compound group* is a single compound (feature) found in any one or more of the data files in a project. For example, if the first data file in the project yields 35 compounds, then at least 35 compound groups are in the project. If additional unique compounds are found in the other data files, then additional compound groups are created.

Information regarding the available columns are found in the online Help in the topic “Compound Groups Columns.” A list of the available columns is displayed when you right-click within the Compound Groups table, and then click **Add/Remove Columns** (see Figure 4 on page 20).



File	Flag	Score (Tg)	Score (MFE)	Area
1-1_Control_000.d		99.82	80	19413007
1-2_Control_000.d		99.86	47.9	23857714
1-3_Control_000.d		99.77	71.4	15390231
1-4_Control_000.d		99.99	51.6	28265502
2-1_Infected_000.d		99.84	80	22560792
2-2_Infected_000.d		99.88	80	39330596
2-3_Infected_000.d		99.82	80	22348177
2-4_Infected_000.d		99.83	100	25071993
3-1_Control_250.d		99.54	54.7	16707582
3-2_Control_250.d		99.29	59.9	12040191
3-3_Control_250.d		99.73	69.6	17088803
3-4_Control_250.d		99.65	37.5	16597367
4-1_Infected_250.d		99.41	77.1	17006064
4-2_Infected_250.d		99.73	98	24996971

Compound Details The data presented in Compound Details is organized as a list of the appearance of a selected feature (compound) in all of the data files in your project - *feature information by data file*.

The Compound Details window shows a table of compound-level information for a single feature selected in the Compound Groups window. The quantitative information is shown for the selected feature as it is found in each data file in your project.

Information regarding the available columns are found in the online Help “Compound Details Columns.” A list of the available columns is displayed when you right-click within the Compound Details table, and then click **Add/Remove Columns** (see Figure 4 on page 20).

The *Flags (Tgt)* column, highlighted, provides a feature extraction summary for each isotopologue for the selected compound, for each sample data, file based on warning, or error flags, reported from the Find by Formula algorithm.

? = no errors (qualified)

The isotopologue qualified and was used in the calculation.

M = m/z tolerance (disqualified)

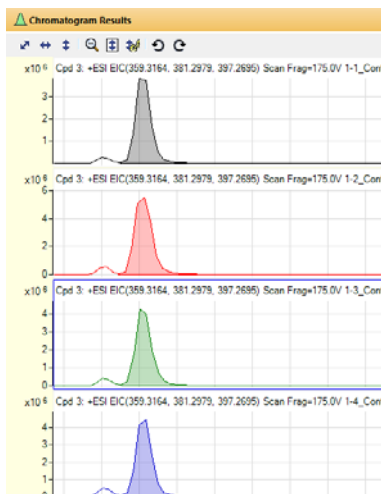
The isotopologue was disqualified for not meeting the m/z tolerance.

T = time correlation (disqualified)

The isotopologue was disqualified for not coeluting.

_ = no peak (disqualified)

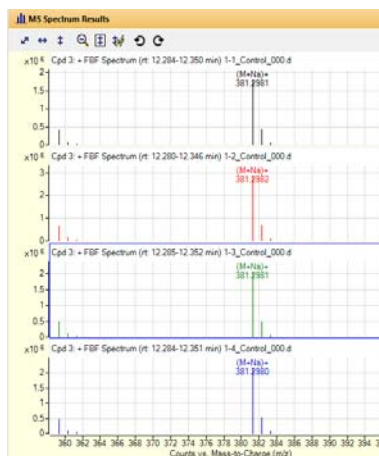
The isotopologue was disqualified because its EIC did not have enough contiguous non-zero points.



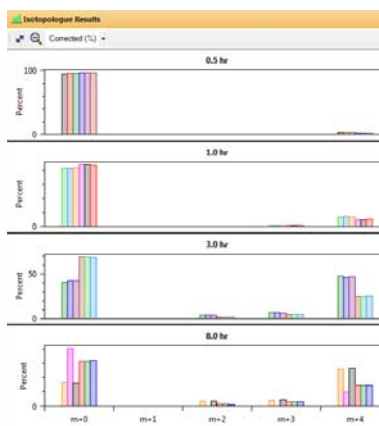
Chromatogram Results Chromatogram Results presents the extracted ion chromatogram (EIC) for each feature and, for isotopologue extraction results, the sum of the EICs for all of the isotopologues. For non-targeted feature extraction, the extracted compound chromatogram (ECC) is displayed for the ions contained in the molecular feature of the feature selected in the Compound Groups window. An EIC/ECC set is displayed for each data file. By default the chromatograms are displayed in an alternating cycle of ten colors to help you review the data for a particular data file as you select different features. **Color by Sample Group** displays the samples in an alternating cycle of colors based on the sample group assignment.

User Interface

Main Functional Areas



MS Spectrum Results MS Spectrum Results presents the averaged mass spectrum (MS) across the integrated ECC for the feature selected in the Compound Groups window for each data file. For isotopologue extraction results the mass spectra are presented across the isotopologue extraction region. By default the MS is displayed in an alternating cycle of ten colors, matched with the Chromatogram Results, to help you review the MS data for a particular data file as you select different features. **Color by Sample Group** displays the samples in an alternating cycle of colors based on the sample group assignment.



Isotopologue Results Isotopologue Results presents a sequence of charts, or a single chart, depending on the results mode selected from the toolbar (see “Results Modes” on page 15).

List mode displays the isotopologue results for each sample file. The isotopologue charts are arranged in the order of your sample groups with each sample replicate displayed in an alternating cycle of ten colors, matched with the Chromatogram Results and MS Spectrum Results. The coloring can be changed to represent the sample groups by selecting **Color by Sample Group** from the toolbar.

Sample group mode displays the isotopologue results for each sample group. Each sample replicate is displayed within each group chart in an alternating cycle of ten colors, matched with the Chromatogram Results and MS Spectrum Results. The coloring can be changed to represent the sample groups by selecting **Color by Sample Group** from the toolbar.

Overlaid mode displays a single summary chart of the isotopologue results. The summary chart contains the average and standard error for each isotopologue per sample group presented in gray scale. The sample groups can be viewed in color by selecting **Color by Sample Group** from the toolbar.

The order of the appearance of the samples and sample groups is set in the **Add/Remove Sample Files** dialog box. The compound containing the isotopologues is selected in the Compound Groups window.

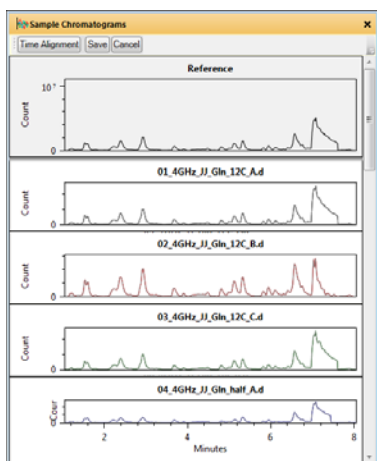
The chart y-axes can be scaled to raw abundances (Raw), raw abundances normalized to 100% (Raw (%)), natural isotope abundance corrected abundances (Corrected), and natural isotope abundance corrected and normalized to 100% (Corrected (%)).

Raw: The actual abundances of each isotopologue by sample data file, or average abundance of each isotopologue when the data is viewed in the summary chart.

Corrected: The abundance for each isotopologue is corrected to remove the natural isotopic contributions so that the abundance is due to the isotopic enrichment from the qualitative flux analysis.

Raw (%): The actual relative abundances of each isotopologue by sample data file, or average of each isotopologue when the data is viewed in the summary chart.

Corrected (%): The relative abundance for each isotopologue after the abundance for each isotopologue is corrected to remove the natural isotopic contributions so that the abundance is due to the isotopic enrichment from the qualitative flux analysis.



Sample Chromatograms After you have added sample files to your project, the chromatograms are shown in the Sample Chromatograms window.

When you click **Time Alignment** in the toolbar for this window, the Time Alignment dialog box is opened. You can select the following options:

- Without standards (you specify a reference data file)
- With standards (you specify a CSV file containing the retention time and mass for the standards)
- Polynomial interpolation
- CE specific (Reijenga et al. 2002)

Shortcut menu commands

A right-click in any of the main window shows you a menu of options (Figure 4 on page 20) applicable to where you click and to your current selection. Options available allow you to customize the columns in the table views, copy to the system clipboard, and delete compound groups or individual compounds.

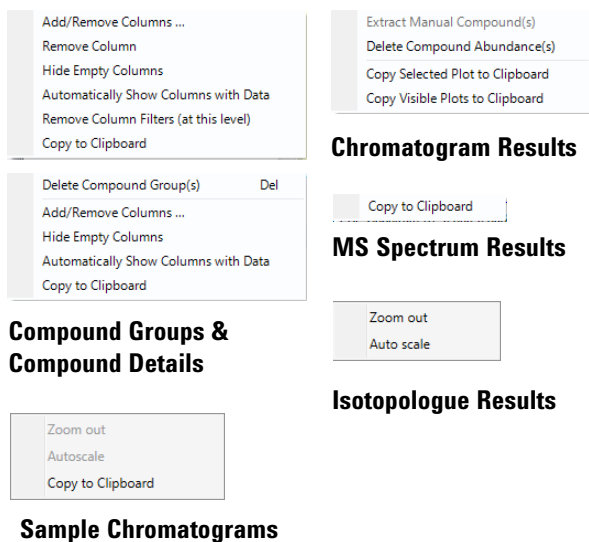


Figure 4 The shortcut menu commands available within the main window automatically adjust based on the active window.

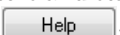
Definitions

Algorithm An algorithm is a set of automated, sequential mathematical tasks performed to find, filter, align, and extract features from your chromatographic/mass spectral data sets.


Feature A feature is synonymous with compound. A feature is referred to interchangeably with compound, molecular feature, element, or entity during the various steps of analysis using Agilent MassHunter software.

Isotopologue	Molecules that contain the same molecular formula and structure but differ in their isotopic composition through the substitution of one or more atoms with a different isotope.
Isotopomer	Molecules that contain the same molecular formula, structure, and number of isotopes but differ in the location of the isotopes in the molecular structure.
Method	A method is a set of parameters that are associated with the feature extraction algorithms used by Profinder. Methods containing the parameters for the algorithms can be saved using unique file names.
Wizard	A wizard is a sequence of interactive steps used by Agilent MassHunter software to guide you through the steps necessary to complete an analytical task. Profinder uses a set of wizards to guide you through the parameters associated with each feature extraction algorithm.
Workflow	A workflow is an Agilent document or a graphical overview that captures a sequence of steps to guide you through an analytical task. A workflow may cover more than one wizard and may include steps performed by more than one Agilent MassHunter software program.

Help

Online Help and detailed information regarding the various parameters and statistical treatments are available by pressing **F1**, clicking , or clicking **Help > Contents** from the menu.

Unsaved parameter changes

When you make a change to a parameter in Profinder, the software automatically places a change icon  (a blue triangle shape) in the wizard tab and next to the value containing the parameter where you made a change. This icon indicates that you have unsaved parameters changes and helps you remember to save the changes you have made to the method. The original parameter value may be viewed by placing your pointer over the change icon. When you save your method, the change icon(s) disappear.

Basic Profinder Workflow

The basic Profinder workflow guides you through the steps necessary to extract features from your sample files. The steps shown in [Figure 5](#) are described in the following sections:

- “Add sample files to your project” on page 24
- “Edit and run your feature extraction algorithm method” on page 31
- “Export your Profinder features” on page 33
- “Save your Profinder method” on page 36
- “Save your Profinder project” on page 37

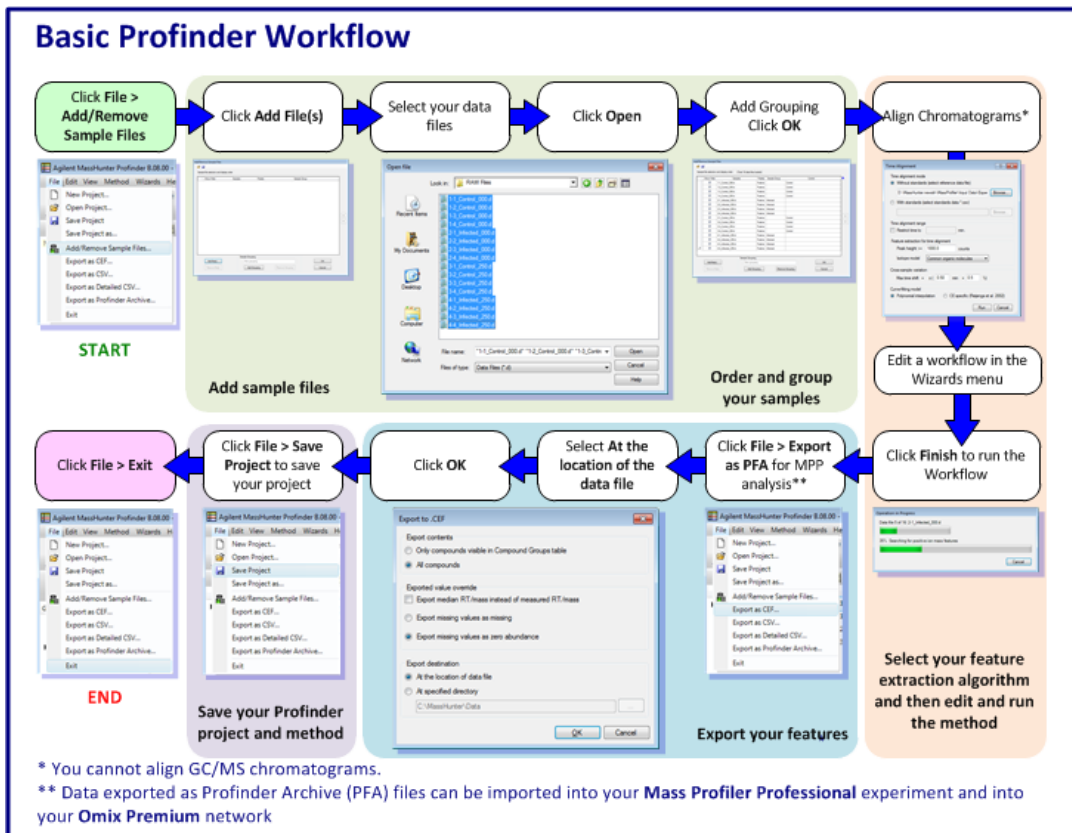


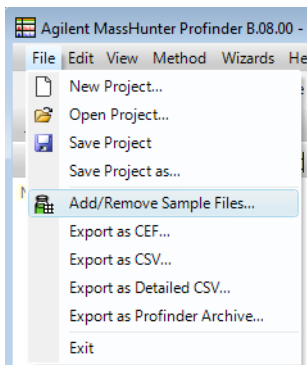


Figure 5 The basic Profinder workflow (see the *MassHunter VistaFlux Software - Workflow Guide* for the Profinder Batch Isotopologue Extraction workflow)

Add sample files to your project

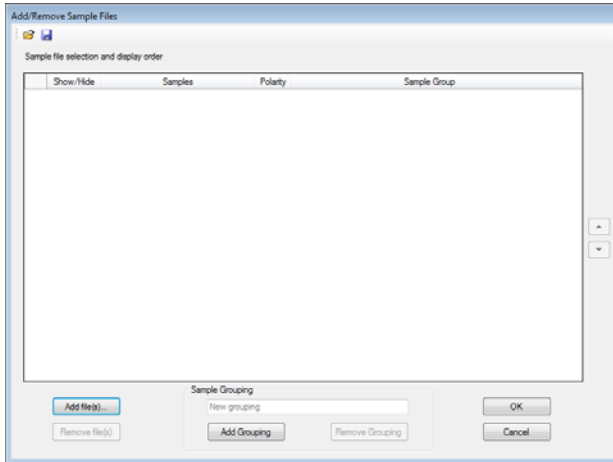
In this task, you launch Profinder, select your sample files, and add the sample files to your Profinder project.

Steps	Detailed Instructions	Comments
1 Start Profinder.	<ul style="list-style-type: none">Click the Profinder icon  on your desktop, or click Start > All Programs > Agilent > MassHunter Workstation > Profinder B.08.00.	<ul style="list-style-type: none">The first time you start Profinder it displays in full screen. Subsequent program launches remember the screen size from your prior session.
2 Begin your Profinder project by adding sample files.	<ul style="list-style-type: none">Click File > Add/Remove Sample Files, or click  Add/Remove Sample Files... on the toolbar.	<ul style="list-style-type: none">If you click File > New Project after starting Profinder, you are automatically prompted to add sample files.When you click Add/Remove Sample Files after launching Profinder, Profinder starts a new project as if you clicked File > New Project.

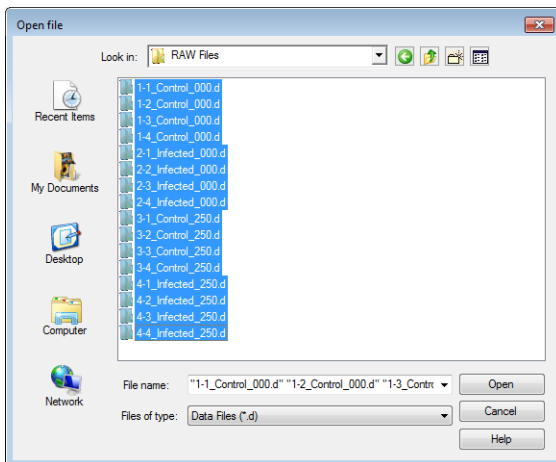


Steps	Detailed Instructions	Comments
-------	-----------------------	----------

- | | | |
|--|--|--|
| <p>3 Add files to the Add Sample Files dialog box.</p> | <ul style="list-style-type: none"> Click Add file(s) to display the Open file dialog box. | |
|--|--|--|



- | | | |
|---|--|--|
| <p>4 Select all of the sample data files in the Open file dialog box.</p> | <ul style="list-style-type: none"> a Browse to the folder containing your sample files. b Select all of the sample data files for your Profinder project. c Click Open. | |
|---|--|--|



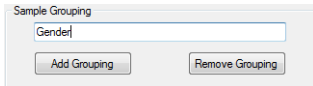
- Selected sample data files are highlighted in the **Open file** dialog box.
- A Profinder project contains all of the sample data files from your experiment - replicates samples for each condition (group) in your experiment.
- You can click **Add Sample Files**, to add additional sample data files to your project and rerun feature extraction.

Basic Profinder Workflow

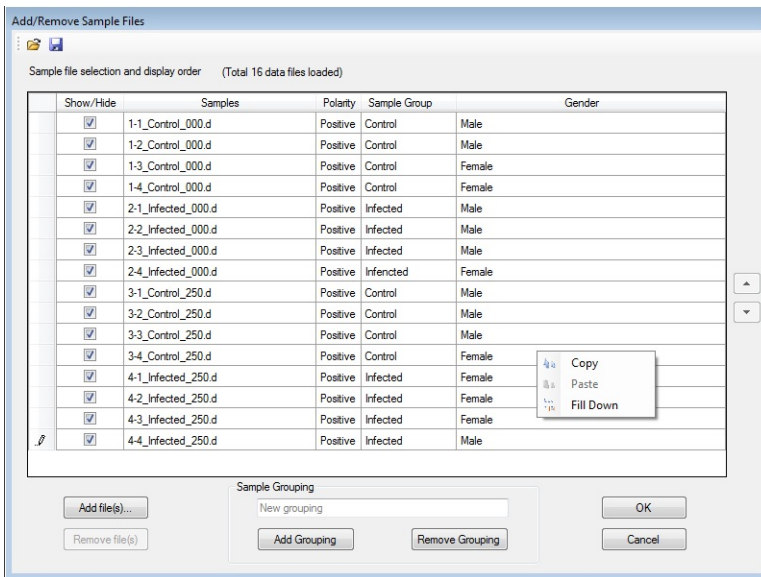
Add sample files to your project

Steps	Detailed Instructions	Comments
5 (optional) Specify Sample Group for each sample in the project.	<p>a Select the first cell in the Sample Group column.</p> <p>b Type either <i>Control</i> or <i>Infected</i>, depending on the sample.</p> <p>c Fill in the rest of the table.</p>	<ul style="list-style-type: none"> You can also right-click the Group column and click Remove Grouping.

6 Add groups to the project.	<p>a At the bottom of the Add/Remove Sample Files dialog box, enter another possible grouping.</p> <p>b Click Add Grouping.</p>	<ul style="list-style-type: none"> You can add multiple groups to a project. Additional groups could include gender, drug dose, cancer stage, etc. Samples can belong to multiple groups.
------------------------------	---	--



7 Enter Sample Group values for the sample files in the Add/Remove Sample Files dialog box.	<p>a Click the data entry cell under the Sample Group column, next to the sample file name.</p> <p>b Enter the sample group identification text.</p> <p>c Add <i>Male</i> or <i>Female</i> in the <i>Gender</i> column.</p> <p>d Repeat for each sample file.</p>	<ul style="list-style-type: none"> When a data entry cell is selected it is highlighted using a background color. When entering group values, the entries must use identical letters, numbers, punctuation, and case in order for the grouping functions to perform properly.
--	--	---



- Use the **Copy**, **Paste**, and **Fill Down** shortcut commands (click the right mouse button) to help assign **Sample Grouping** values to each sample.

Steps	Detailed Instructions	Comments
8 Mark the samples to add to your project.	a Mark the samples to add to your project in the Show/Hide column. b Click OK .	<ul style="list-style-type: none">• All of the Samples in the same project must have the same polarity. If some of the Samples have a Positive Polarity and some have a Negative Polarity an error message is shown when you click OK.

Your new Profinder project is created.

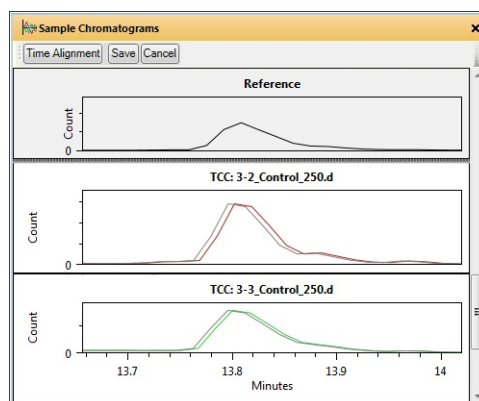
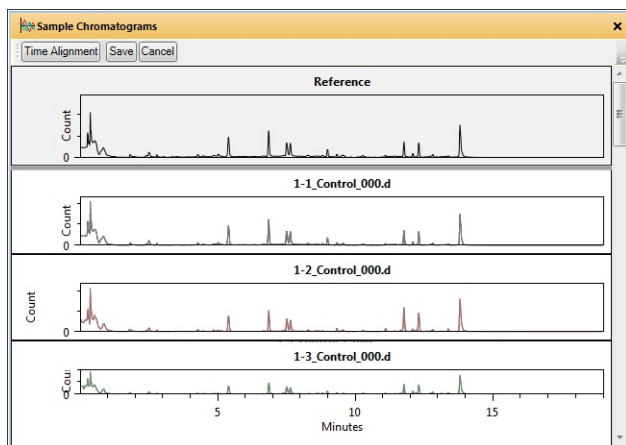
The Sample Chromatograms window shows a chromatogram for each sample. You can now do time alignment or run a wizard.

Align chromatograms

Two algorithms have been implemented to align chromatograms. See “[Time Alignment Algorithms Overview](#)” on page 5.

Steps	Detailed Instructions	Comments
1 Edit the Time Alignment parameters.	<p>a Click Time Alignment in the toolbar in the Sample Chromatograms window.</p> <p>b Click either Without standards or With standards for the Time alignment mode.</p> <p>c Click the Browse button and select either a reference data file (*.d) or a Standards database file (*.cdb, *.csv).</p> <p>d Click a Curve fitting model.</p> <p>e Review the other parameters.</p> <p>f Click Run.</p>	<ul style="list-style-type: none"> The online Help has information on the Curve fitting models. The CE specific (Reijenga et al. 2002) curve-fitting model is optimized for Capillary Electrophoresis data. The Standards database file must have a time column (with the header RT, Time, or Retention Time) and a Mass column. Mass is the value of the neutral mass of the first isotope. Agilent recommends that you set the same time range in the Wizard for feature extraction.

Steps	Detailed Instructions	Comments
2 Review the time alignment.	<p>a Select the maximum number of list panes in the main toolbar.</p> <p>b Hover the mouse over a peak in the Reference chromatogram. A tooltip shows the Minutes and Counts for that peak.</p> <p>c Hover the mouse over a peak in one of the Sample data files. A tooltip shows the Minutes and Counts for that peak.</p> <p>d Compare these values.</p> <p>e Continue comparing other peaks.</p> <p>f If needed, edit the parameters in the Time Alignment dialog box and run the algorithm again.</p>	<ul style="list-style-type: none"> The maximum number of list panes does not include the Reference data file if you clicked Without standards as the Time alignment mode. After you run the time alignment algorithm, the original chromatogram is shown in gray. The chromatogram that has been time aligned is shown in color.

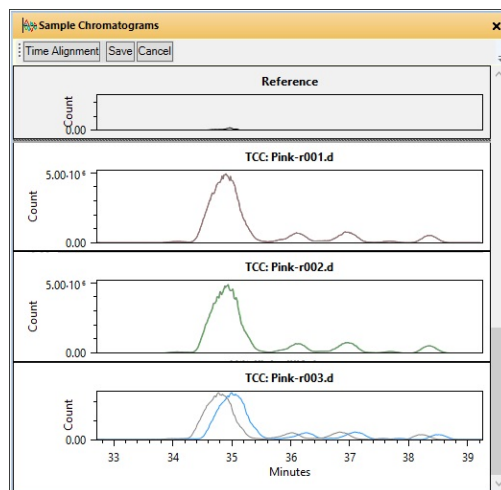
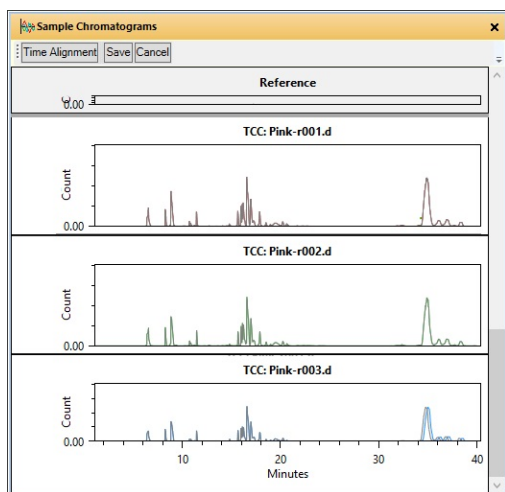


3 Save the time alignment.	<p>a Click Save in the toolbar in the Sample Chromatograms window.</p> <p>b Click either Yes or No in the Replace existing time alignment message box.</p>	<ul style="list-style-type: none"> You do not need to run Time Alignment before running a wizard. If you do run Time Alignment, you have to either save or discard your time alignment results before running any of the wizards.
----------------------------	---	---

Basic Profinder Workflow

Align chromatograms

Steps	Detailed Instructions	Comments
<ul style="list-style-type: none">Here is an example of using Time Alignment on CE/MS data.	<ol style="list-style-type: none">Select the maximum number of list panes in the main toolbar.Hover the mouse over a peak in the Reference chromatogram. A tooltip shows the Minutes and Counts for that peak.Hover the mouse over a peak in one of the Sample data files. A tooltip shows the Minutes and Counts for that peak.Compare these values.Continue comparing other peaks.If needed, edit the parameters in the Time Alignment dialog box and run the algorithm again.	<ul style="list-style-type: none">The maximum number of list panes does not include the Reference data file if you clicked Without standards as the Time alignment mode.After you run the time alignment algorithm, the original chromatogram is shown in gray. The chromatogram that has been time aligned is shown in color.

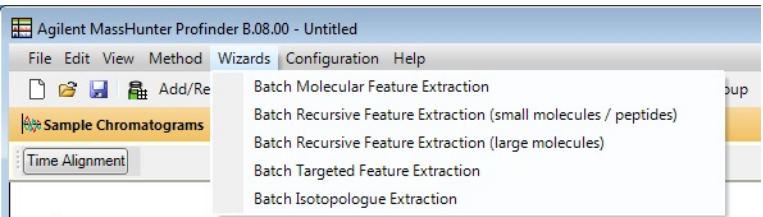


Edit and run your feature extraction algorithm method

Batch feature extraction performed by Profinder is part of a workflow for statistical data analysis, qualitative flux analysis, and data visualization in biochemical networks.

- Batch Molecular Feature Extraction
- Batch Recursive Feature Extraction (small molecules / peptides)
- Batch Recursive Feature Extraction (large molecules)
- Batch Targeted Feature Extraction
- Batch Isotopologue Extraction

An overview of each extraction algorithm is found in the section “[Feature Extraction Workflow Algorithms](#)” on page 38. Batch Isotopologue Extraction is optimized to extract isotopologues for targeted features for qualitative flux analysis.

Steps	Detailed Instructions	Comments
1 Select the feature extraction algorithm to apply to your samples.	<p>a Click one of the algorithms in the Wizards menu.</p> 	<ul style="list-style-type: none"> • Molecular Feature Extraction performs chromatographic deconvolution and aligns the features across the selected sample files using mass and RT. See “Batch Molecular Feature Extraction” on page 38 for the wizard steps. • Recursive Feature Extraction performs MFE and then uses the mass and RT of each feature to automatically perform a targeted feature extraction. See “Batch Recursive Feature Extraction” on page 39 for the wizard steps. • Target Feature Extraction uses an input formula source to extract features. See “Batch Targeted Feature Extraction” on page 41 for the wizard steps. • Isotopologue Extraction locates possible isotopologues of target compounds. See “Batch Isotopologue Extraction” on page 42 for the wizard steps.

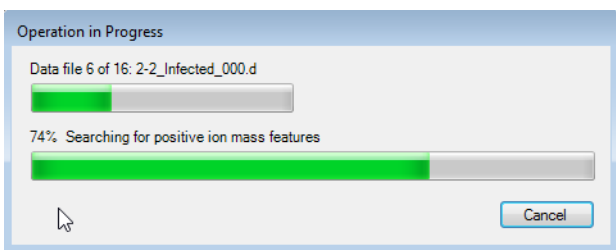
Basic Profinder Workflow

Edit and run your feature extraction algorithm method

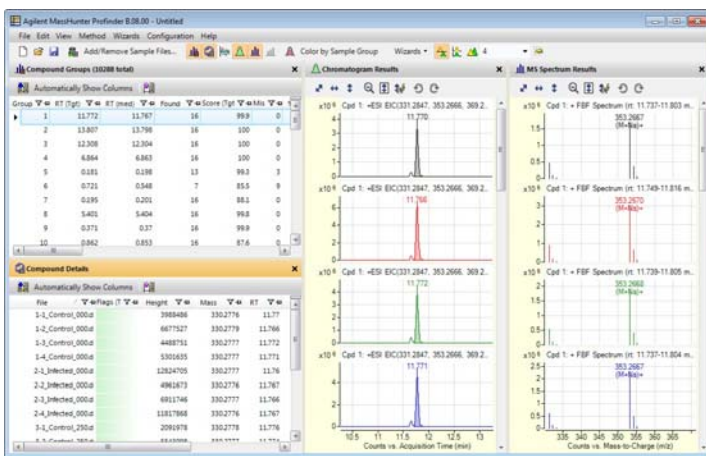
Steps	Detailed Instructions	Comments
2 Review and edit the parameters for the feature extraction algorithm.	<ul style="list-style-type: none">a Make changes to the parameters on each wizard tab for each step.b Click Next to go to the next step of the wizard.c Click Previous to return to an earlier step.	<ul style="list-style-type: none">• Click Help to open online Help specific for the current tab in the wizard.• Click Cancel to stop editing the algorithm parameters. Any changes made to the algorithm parameters are not saved.



3 Run the feature extraction algorithm.	<ul style="list-style-type: none">a Click Finish to run the extraction algorithm on your sample data.b An Operation in Progress dialog box is displayed while the extraction process is running.	<ul style="list-style-type: none">• The feature extraction process can take a long time. To significantly improve performance, increase the amount of memory and use a faster processor.
---	---	--



4 Review your results.	<ul style="list-style-type: none">• Review your results and then export the results and save your method.	<ul style="list-style-type: none">• Using the Compound Groups, Compound Details, Chromatogram Results, and MS Spectrum Results windows, you can edit your results: delete a compound group, delete abundance from compound details window or chromatogram results window, and integrate a peak manually. More information is available in the online Help.• The next steps in this <i>Quick Start Guide</i> help you export your results and save your method.
------------------------	---	---



Export your Profinder features

You save your Profinder results as a Profinder Archive (PFA) file to import your molecular features into Mass Profiler Professional (MPP) and to visualize your results in the context of biochemical networks in Omix Premium.

When a given compound in a compound group is not integrated (either because no detectable abundance was found for the compound in a given sample or because of a chromatographic peak integration issue), you have the opportunity to choose how such missing values are handled in a statistics/chemometrics software like Mass Profiler Professional.

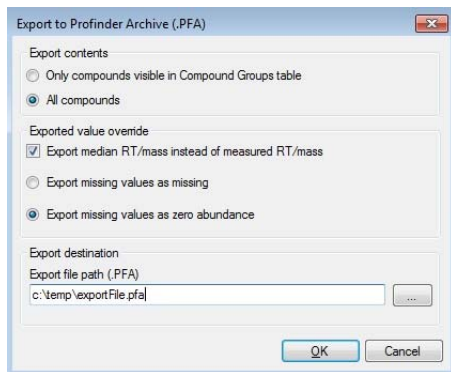
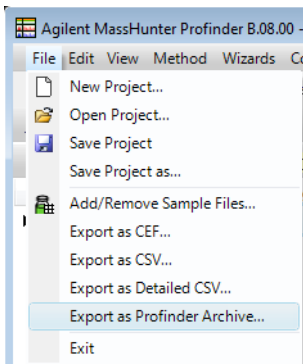
If missing values are exported as “missing,” then these values are ignored in calculations of averages, in t-test/ANOVA statistical tests and in fold change calculations between sample groups. If you have run a basic Batch Molecular Feature Extraction without recursion or if any integration issues remain after running Batch Recursive or Batch Targeted Feature Extraction, it is preferable to export missing values as “missing.”

If missing values are exported as “zero abundance,” then the value 0.0 is used and included in calculation of statistics. If the values are truly zero (if you have reviewed all compound peak integrations and any empty compounds are truly not measured), then this is the preferred mode of export.

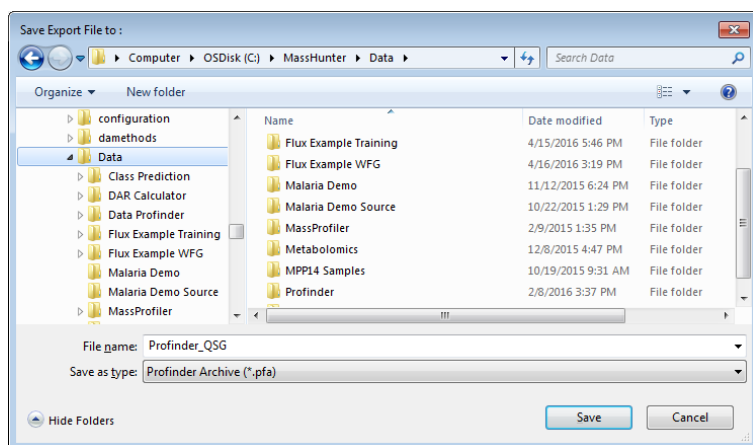
Basic Profinder Workflow

Export your Profinder features

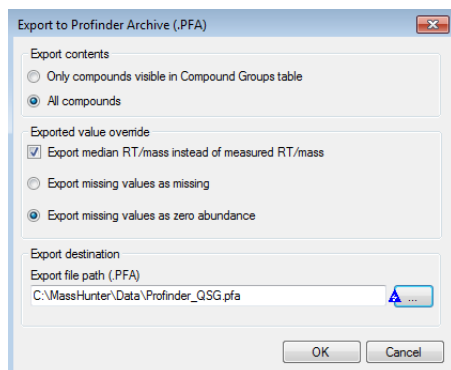
Steps	Detailed Instructions	Comments
1 Run Export as Profinder Archive .	<ul style="list-style-type: none">Click File > Export as Profinder Archive.	<ul style="list-style-type: none">Alternate export options include:<ul style="list-style-type: none">Export as CEF - features saved to a file format used to exchange data between Agilent software.Export as CSV - average Mass and RT with actual feature abundance for each sampleExport as Detailed CSV - average Mass and RT with all actual feature values for each sample
2 Select your export contents and destination.	<ol style="list-style-type: none">Select All compounds in the Export to Profinder Archive (.PFA) dialog box.Mark or clear the Export median RT/mass instead of measured RT/mass check box.Click either Export missing values as missing or Export missing values as zero abundance.Click the <input type="button" value="..."/> button to select your export destination folder, and enter your file name.	



Steps	Detailed Instructions	Comments
3	<p>Navigate to the folder and enter your file name.</p> <p>a Navigate to the folder to save your project in the Save Export File to dialog box.</p> <p>b Type a name for File name.</p> <p>c Click Save in the Save Export File to dialog box.</p>	<ul style="list-style-type: none"> The Profinder Archive is not saved at this time; it is saved from the Export to Profinder Archive (.PFA) dialog box.



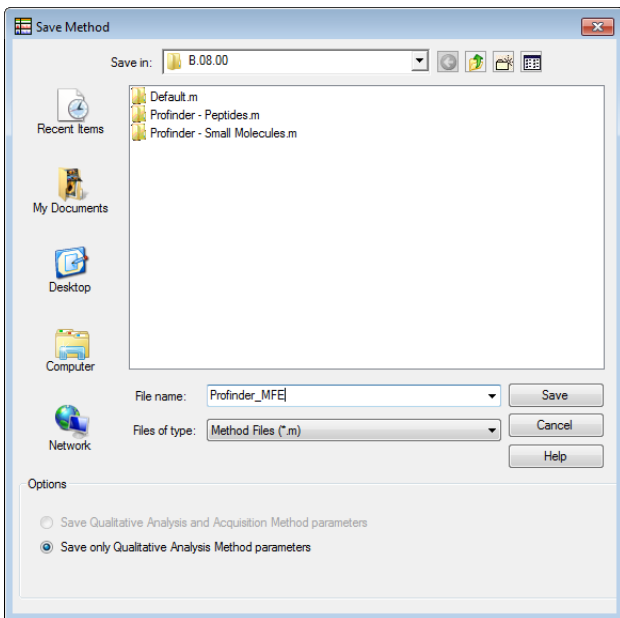
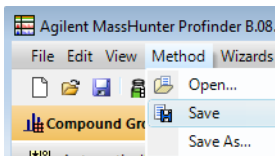
4	<p>Save your Profinder archive file.</p> <p>a Review your selection, path and file name in the Export to Profinder Archive (.PFA) dialog box.</p> <p>b Click OK to save your Profinder archive.</p>	<ul style="list-style-type: none"> Your Profinder Archive (PFA) file is imported into Omix Premium where you can visualize your results in the context of biochemical networks.
---	---	--



Save your Profinder method

You save your Profinder feature algorithm method using a unique name.

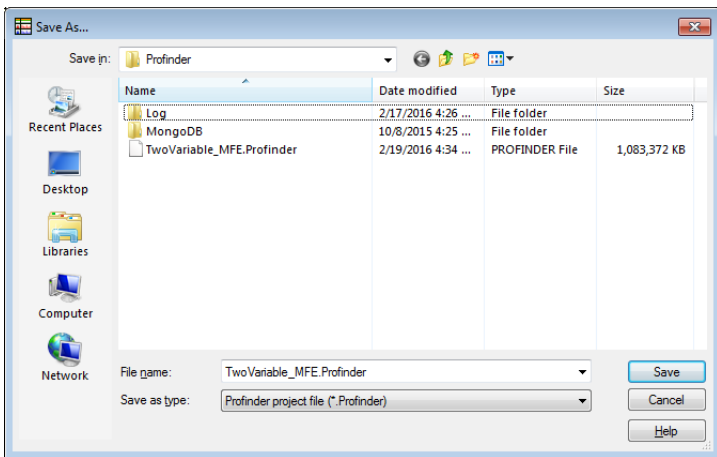
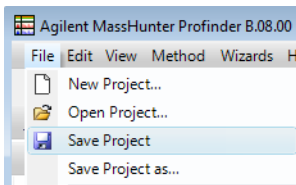
Steps	Detailed Instructions	Comments
1 Select Save from the menu bar.	<ul style="list-style-type: none">Click Method > Save.	<ul style="list-style-type: none">If you have not previously saved your method, Save is the same as Save As and prompts you for a name to save your method.
2 Type the file name and save your method.	<ul style="list-style-type: none">a Type the name to use for your method file in the Save Method dialog box.b Click Save.	<ul style="list-style-type: none">The Save Method dialog box does not appear when you click Save for your method if you previously saved your method.Click Method > Save As if you want to save your method using a new file name.If acquisition parameters exist in the current method, you can select whether to also save the acquisition method parameters to the new method.



Save your Profinder project

This is an optional step. You can save your Profinder project, method, and the current sample data file extraction results so that you can continue reviewing your results and the extraction method at a later time.

Steps	Detailed Instructions	Comments
<p>1 Select Save Project from the menu bar.</p>	<ul style="list-style-type: none"> Click File > Save Project. 	<ul style="list-style-type: none"> If you have not previously saved your project, Save Project is the same as Save Project as and prompts you for a name to save your project.
<p>2 Type the file name and save your project.</p>	<ol style="list-style-type: none"> Type the name to use for your project file in the Save as dialog box. Click Save. 	<ul style="list-style-type: none"> The Save As dialog box does not appear when you click Save Project if you previously saved your project. Click File > Save Project as if you want to save your project using a new file name. Note: Profinder project files can be one GB or more in file size. Note: Remember to include the original sample data when you share a Profinder project.



Feature Extraction Workflow Algorithms

Five feature extraction algorithms are available in Profinder. Each extraction algorithm is designed to efficiently extract the features (compounds) in your sample data files. The resulting compound groups are imported into your MPP experiment or Omix Premium as part of your qualitative flux analysis.

When you begin a new analysis, start with permissive feature extraction algorithm parameters and review the results for false positive and false negative features. Then adjust the algorithm parameters to reduce the appearance of false positive features without increasing the loss of actual features - false negative features.

Batch Molecular Feature Extraction

Molecular feature extraction (MFE) performs chromatographic deconvolution to find the features in your samples. After finding the features, the features (compounds) are aligned across all of the selected sample files using mass and retention time (RT). After performing MFE, Profinder employs an algorithm to re-assign ion species using cross-file analysis to generate compound groups.

The wizard workflow is shown in [Figure 6](#).

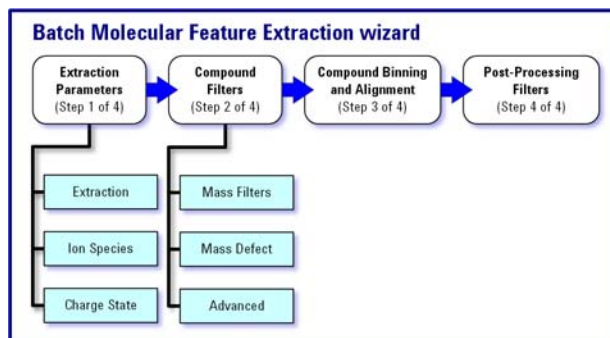


Figure 6 Four steps are presented in the Batch Molecular Feature Extraction wizard.

Batch Recursive Feature Extraction

Recursive Feature Extraction (RFE) performs MFE, as described above, and then uses mass and retention time of the results to perform a targeted feature extraction referred to as Find by Ion (FbI). Using the untargeted MFE results for a targeted FbI feature extraction is referred to as *recursive feature extraction*.

After calculating the median mass, median RT, and composite spectrum from the aligned features found using MFE, FbI uses the median values to perform a targeted extraction to improve the reliability in finding the features in your data. This recursive feature finding, combined with replicate samples for each of your experiment conditions, improves the statistical accuracy of your analysis and reduces the potential for obtaining a false positive or false negative answer to your hypothesis.

The recursive feature extraction workflow performed by Profinder is similar to that outlined in the *Agilent Metabolomics Workflow - Discovery Workflow Guide*. The difference is that the feature extraction process happens in a single step, in the same application, without moving CEF files back and forth between MassHunter Qualitative Analysis, DA Reprocessor, and Mass Profiler Professional as illustrated in [Figure 1](#) on page 5. For a comparison of feature finding between Profinder and MassHunter Qualitative Analysis, refer to *Class Prediction with Agilent Mass Profiler Professional Workflow Guide*.

The wizard workflow for small molecules is shown in [Figure 7](#) on page 40.

The wizard workflow for large molecules is shown in [Figure 8](#) on page 40.

Feature Extraction Workflow Algorithms

Batch Recursive Feature Extraction

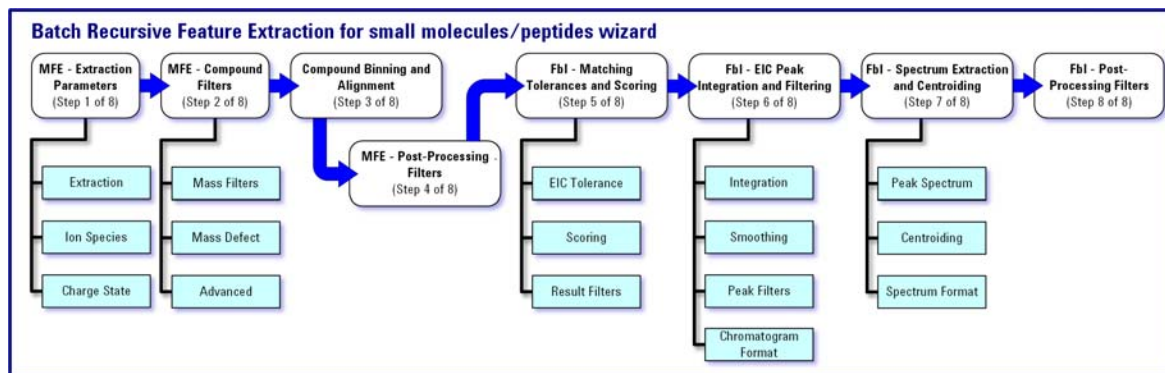


Figure 7 Eight steps are presented in the Batch Recursive Feature Extraction (small molecules / peptides) wizard.

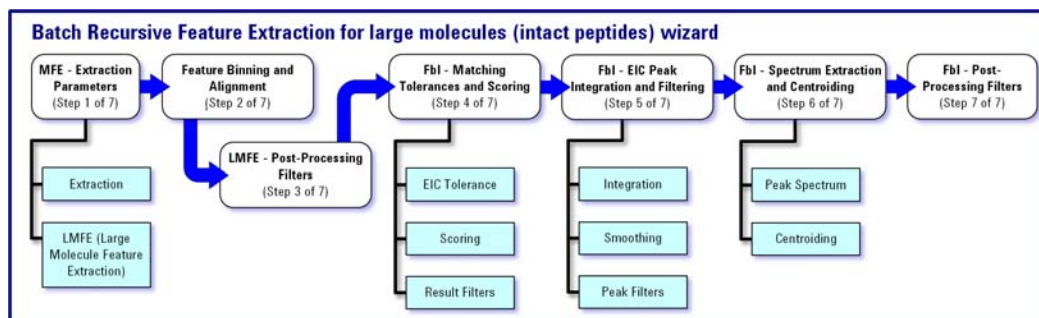


Figure 8 Seven steps are presented in the Batch Recursive Feature Extraction (large molecules) wizard.

Batch Targeted Feature Extraction

Targeted feature extraction (TFE) uses an input formula source, CEF file, or database containing molecular formulas, mass, and/or retention time information to extract features from your data using a process referred to as Find Compounds by Formula. After feature extraction is performed, compound alignment produces a result such that one compound target can only have one compound extracted per data file.

Find Compounds by Formula typically uses molecular formula information, like a PCDL, to calculate the ions and isotope patterns derived from the formula as the basis to find features in the sample data file. When the input molecular features contain mass and retention time, as found in CEF file, instead of just the molecular formula, TFE calculates reasonable isotope patterns and uses these patterns with retention time tolerances to find the target features in the sample data files. When the input molecular features are filtered from a find process that was previously untargeted, the molecular features found using Find Compounds by Formula is also referred to as recursive finding.

The wizard workflow is shown in [Figure 9](#).

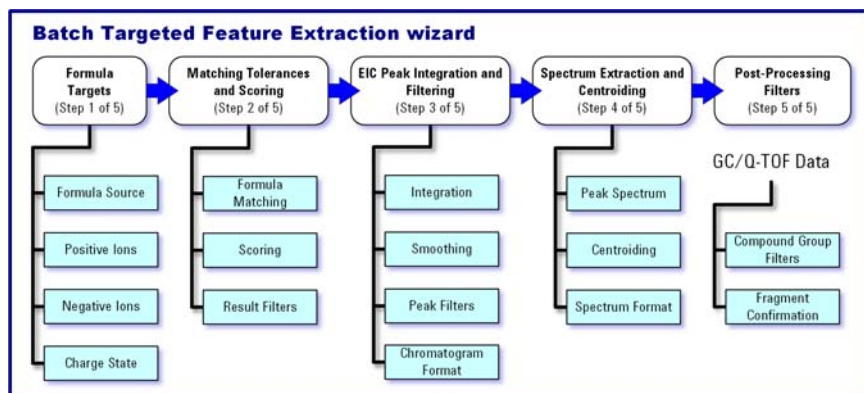


Figure 9 Five steps are presented in the Batch Targeted Feature Extraction wizard.

Batch Isotopologue Extraction

Batch isotopologue extraction supports only LC/MS acquired data. Unlike the other batch feature extraction wizards, target retention times are required for this workflow.

Isotopologue extraction (IE) uses an input CSV file or compound database file, PCD/PCDL, containing the target feature molecular formulas, mass, and/or retention time information, and anticipates that the target compound may have undergone some degree of isotope labeling, to extract features from your data using a process referred to as Find Compounds by Formula. After feature extraction is performed, the extraction algorithm determines which of the possible isotopologues are actually present, measures the raw abundances of the isotopologues, and corrects the isotopologues abundances for the natural occurrence of the unlabeled ions.

See the *MassHunter VistaFlux Software - Quick Start Guide* and the *MassHunter VistaFlux Software - Workflow Guide* for additional information.

Isotopes, Isotopomers, Isotopologues, and Mass Spectra

Isotopologues are molecules that contain the same molecular formula and structure but differ in their isotopic composition through the substitution of one or more atoms with a different isotope. The exact location of the isotope in the molecule, while important chemically, is not important in flux analysis, just the number of isotopes in the molecule. Isotopologues can be identified using single-stage MS.

Isotopomers are molecules that contain the same molecular formula, structure, and number of isotopes but differ in the location of the isotopes in the molecular structure. Isotopomers can be identified using advanced MS/MS techniques.

Note: Profinder finds and extracts isotopologues in your sample data; it does not find or extract isotopomers.

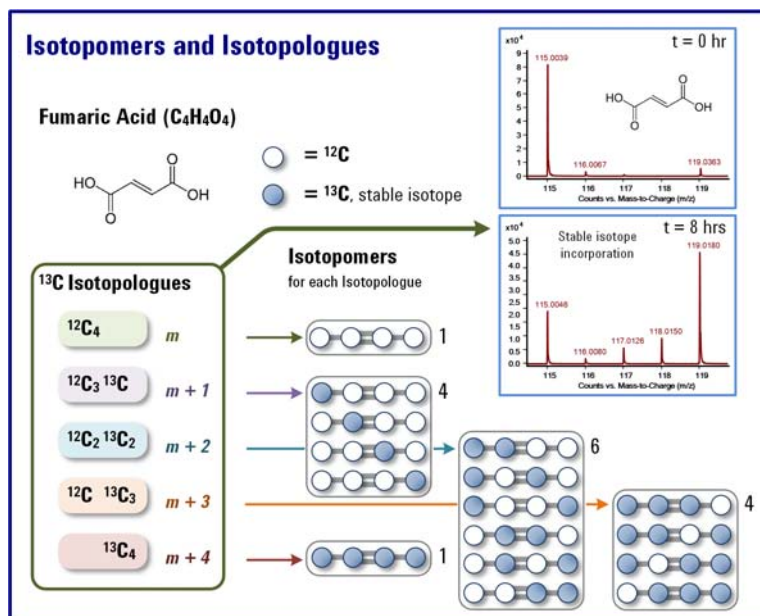


Figure 10 An illustration of how two stable carbon isotopes in a four-carbon molecule relate to isotopomers and isotopologues. Isotopologues are viewed in mass spectra during flux analyses.

A simple four carbon molecule, fumaric acid ($C_4H_4O_4$), is used to explain the relationship of isotopes to isotopomers, isotopologues, and mass spectra. The most abundant isotope of carbon is ${}^{12}C$. ${}^{13}C$, also stable, is not nearly as naturally abundant as ${}^{12}C$; ${}^{13}C$ has a natural occurrence of 1.1% of ${}^{12}C$. For simplicity, naturally occurring ${}^{13}C$ is considered to be negligible; therefore, the mass of the naturally occurring four- ${}^{12}C$ molecule is m (represented as ${}^{12}C_4$), and no positional differences exist among the isotopes of the carbon atoms.

When a single ${}^{13}C$ atom is substituted for one ${}^{12}C$ atom, four locations are possible where the ${}^{13}C$ atom can be placed (isotopomers as shown in Figure 10), and each isotopomer has a mass of $m + 1$ (${}^{12}C_3 {}^{13}C$). When two of the ${}^{12}C$ atoms are replaced with ${}^{13}C$ atoms, six isotopomers are possible, and each of the doubly substituted molecules has a mass of $m + 2$ (${}^{12}C_2 {}^{13}C_2$). When three of the ${}^{12}C$ atoms are replaced with ${}^{13}C$

atoms, four isotopomers are possible, and each isotopomer has a mass of $m + 3$ ($^{12}\text{C} \ ^{13}\text{C}_3$). Finally, when all four of the ^{12}C atoms are replaced with ^{13}C atoms, only a single arrangement with a mass of $m + 4$ ($^{13}\text{C}_4$) exists. The five different masses m , $m + 1$, $m + 2$, $m + 3$, and $m + 4$ represent the masses of the five isotopologues visible in the resulting mass spectra.

Isotopologue mining

Profinder performs isotopologue mining in two stages, an initial screening followed by refinement. The initial screening stage extracts isotopologue EICs around the target retention time range and then evaluates peak mass spectral data to find ions that match the predicted list of possible isotopologues. The refinement stage uses a self-optimizing peak finder to refine the m/z assignment from the profile data and then re-extracts the EICs using the new isotopologue m/z values, refines the start and end retention time bounds on the newly extracted EICs, and then reports both EIC peak area and summed isotopologue peak heights as the compound abundances.

The wizard workflow is shown in Figure 11.

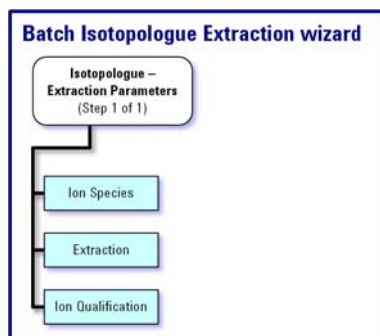


Figure 11 One step is presented in the Batch Isotopologue Extraction wizard.

Monitoring your computer while running Profinder

You can monitor your computer's resources, especially memory, while running Profinder. Right-click on the computer taskbar and then click **Start Task Manager** to launch the Windows Task Manager. Select the **Performance** tab to monitor your CPU and memory usage.

For additional details regarding your computer's memory usage click **Resource Monitor** and select the **Memory** tab. See [Figure 12](#). If your memory usage approaches the amount of physical RAM, adding additional RAM increases performance.

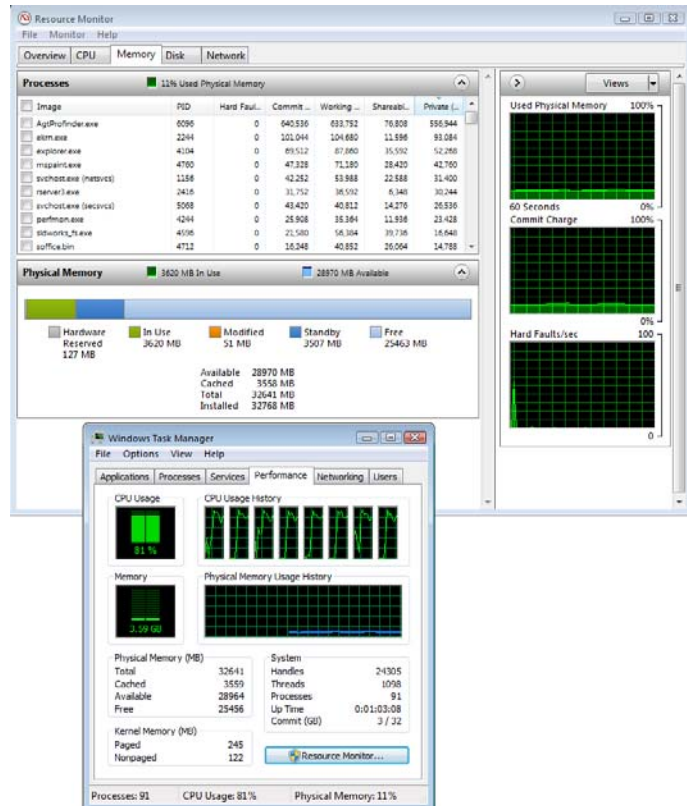


Figure 12 Monitoring the system and memory resources on your computer.

Profinder Installation

Standalone MassHunter Profinder Profinder is installed from a Setup Wizard, which you run from the main installation program. If you have a prior version, uninstall Profinder before installing this newer version (see “Uninstall a prior version” on page 46).

Right-click **ProfinderSetup.exe**, and then click **Run as administrator**.

Install Profinder on the highest performing PC you have available to reduce the time it takes to extract the features from your sample files. Profinder requires a PC running Windows 7 (64-bit) with at least 8GB of RAM and at least 30GB of available disk space.

Note: A PC with 16GB or more of RAM and a solid-state drive will see significantly improved Profinder performance and reduction in the time it takes to extract features from large data sets.

As part of MassHunter VistaFlux Software Profinder is also part of the MassHunter VistaFlux Software suite; for Profinder installation instructions refer to the software installation instructions found in the *MassHunter VistaFlux Software - Quick Start Guide*.

Uninstall a prior version If you have a prior version of Profinder installed, you must delete the prior version before installing this new version.

- 1 Click **Start > Control Panel**.
- 2 Click **Programs and Features**.
- 3 Click **Agilent MassHunter Workstation Profinder Software**.
- 4 Click **Uninstall/Change** to uninstall Profinder.

This page intentionally left blank.

www.agilent.com

In this book

The MassHunter Profinder Software Quick Start Guide presents the first steps to use the MassHunter Profinder Software. Profinder is part of the MassHunter VistaFlux software.

This Quick Start Guide applies to MassHunter Profinder B.08.00 SP2 and later until superseded, as part of the MassHunter VistaFlux Software.

For research use only. Not for use in diagnostic procedures.

© Agilent Technologies, Inc. 2017

Revision B, September 2017



G3835-90034



Agilent Technologies