Xcalibur™

Mass Frontier™ 5.0
User's Guide

XCALI-97181 Rev A July 2006

Analyze • Detect • Measure • Control™

Thermo
ELECTRON CORPORATION

Printing history: Revision A printed July 2006

Written by Robert Mistrik.

Contents

Preface

About This Guide

Welcome to the Mass Frontier™ software, which is part of the Thermo Electron Xcalibur® mass spectrometry data system. Mass Frontier 5.0 provides tools for the management, evaluation, and interpretation of mass spectra.

This guide describes how to use Mass Frontier for mass spectral interpretation.

Related Documentation

In addition to this guide, you can use the Help available from within the Mass Frontier software.

Special Notices

This guide contains special notices in the text, which can include the following:

IMPORTANT Highlights information necessary to avoid damage to software, loss of data, invalid test results, or information critical for optimal performance of the system.

Note Highlights information of general interest.

Tip Helpful information that can make a task easier.

Contacting Us

There are several ways to contact Thermo Electron Corporation.

Assistance

For new product updates, technical support, and ordering information, contact us in one of the following ways:

Visit Us on the Web

www.thermo.com/finnigan

Contact Technical Support

Phone: 1-800-685-9535
Fax: 1-561-688-8736
E-mail: techsupport.finnigan@thermo.com

Find software updates and utilities to download at http://mssupport.thermo.com.

Contact Customer Service

In the US and Canada for ordering information:
Phone: 1-800-532-4752
Fax: 1-561-688-8731
Web site: www.thermo.com/finnigan

Changes to the Manual and Online Help

To suggest changes to this guide or to the online Help, use either of the following methods:

- Fill out a reader survey online at www.thermo.com/lcms-techpubs
- Send an e-mail message to the Technical Publications Editor at techpubs.finnigan-lcms@thermo.com

Chapter 1 Introducing Mass Frontier

The Mass Frontier 5.0 software can help you manage, evaluate, and interpret mass spectra. The software provides a number of tools for processing and organizing mass spectral and chromatographic data. Because of the large volume of information a mass spectrometer produces, management capabilities are essential for mastering your analytical workloads. In contrast to other software systems, Mass Frontier offers features for the interpretation of mass spectra when an unknown is not contained in your libraries.

This chapter contains the following topics:

General Information

Mass Frontier is based on ten modules, which are accessible as windows. All the modules are seamlessly integrated within an intuitive multidocument interface (MDI). All the windows are located in your program desktop. However, four of the fourteen modules (Fragments & Mechanisms, Neural Networks, Spectra Projector, and Components Editor) cannot be directly opened from the program desktop by clicking a button or menu item; they must be generated from user-supplied input.

Although each of the modules is independent, the program automatically establishes a link between several modules. For example, the program can make a link between the Database Manager window and the Fragments & Mechanisms module. In this case, the peaks in a mass spectrum, displayed in the Database Manager module, are linked with mass-to-charge ratios in the Fragments & Mechanisms module. Records in Database Manager can also be linked with objects in the Spectra Classifier and Spectra Projector modules. Double-click these objects to get the spectra (Spectra Classifier) or spectrum (Spectra Projector) that are linked with a particular record in Database Manager. In addition, spectra, structures, and library entries can be exchanged between modules by using the Copy and Paste commands.

System Requirements

At a minimum, Mass Frontier requires the following:

- Microsoft® Windows® 2000 SP 3 or XP SP 1
- Computer with Pentium® 255 MHz processor or higher
- 128 MB of RAM (512 MB recommended)
- SVGA monitor
- 2 GB available hard disk space
- Microsoft Office XP
- Microsoft Internet Explorer 4.01 SP 1 or higher
- Xcalibur 2.0 installed with local user access

Installation

Before you install Mass Frontier software, make sure the Xcalibur data system is installed on your system with local user access. Mass Frontier cannot be activated without it.

Installing Xcalibur

To install the Xcalibur software

1. Insert the Xcalibur Core Data System Software CD into your CD-ROM drive.

2. Choose **Start > Run** from your Windows desktop.

3. Type **D:\setup.exe**, where D is the letter of the CD-ROM drive. Click **OK** to begin the installation.

4. Follow the instructions on the screen until you reach **Finish**.

Installing SQL Server and Mass Frontier

Before running the Mass Frontier installation, install the Microsoft SQL Server desktop engine and run the automated configuration procedure from the Mass Frontier **Setup Launcher**.

Installation of the Microsoft SQL Server desktop engine is a complex procedure, which depends on the system setup. If the installation fails, make a note of the error message and contact the HighChem database group at *database@highchem.com*.

To run Mass Frontier, you must have Full Control permission on the Mass Frontier application directory (by default, X:\Program Files\HighChem\Mass Frontier 5.0). Windows XP does not have this permission set for Limited Account users. If you are a Limited Account user, ask your system administrator to grant Full Control permission on the Mass Frontier directory for your account.

To install the MSDE and Mass Frontier software

1. Insert the Mass Frontier CD in the CD-ROM drive. The Setup Launcher starts automatically.

   If the Setup Launcher does not start automatically, choose **Start > Run** from your Windows desktop. Type **D:\setup.exe**, where D is the letter of the CD-ROM drive. Click **OK** to begin the installation.

2. Click the **Install MSDE** button to start the installation of the Microsoft SQL Server desktop engine. The installation takes several minutes.

3. Click the **Configure MSDE** button to start the automated configuration of the Microsoft SQL Server desktop engine.

4. If the installation and configuration of the Microsoft SQL Server desktop engine are successful, the installation of Mass Frontier starts automatically. Follow the instructions that appear on the screen.

   If the Microsoft SQL Server desktop engine cannot be properly installed or configured, the installation ends. If you cannot resolve the problem, contact the HighChem database group at *database@highchem.com*. After the problems have been resolved, click the **Install Mass Frontier** button to continue with the Mass Frontier installation.

Unless you specify a different location, Mass Frontier is installed at

X:\Program Files\HighChem\Mass Frontier 5.0\MassFrontier.exe

where, X is the CD-ROM drive on your computer.

Activating Mass Frontier

Mass Frontier requires an activation key for each system where you install the software. The activation keys are not transferable from one system to another. If you have purchased multiple copies of Mass Frontier, perform the following procedure for each copy of Mass Frontier that you install.

When you first start Mass Frontier, the License page appears (Figure 1) and prompts you for an activation key.

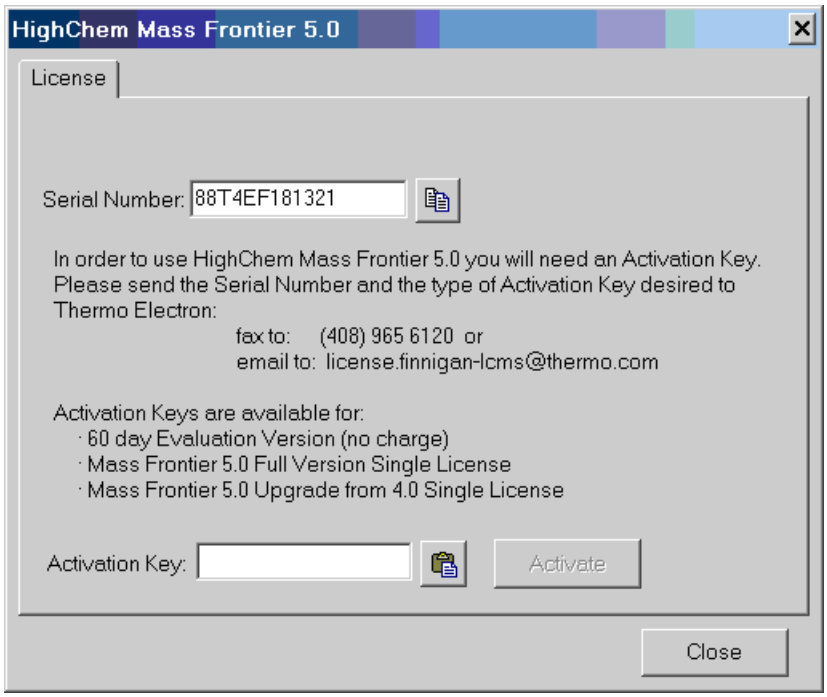


Figure 1. License page, showing the serial number and the types of activation keys available

1. On the **License** page, highlight the text that appears in the **Serial Number** box, and then press CTRL+C to copy the serial number to the Windows clipboard.

2. Obtain an activation key by e-mail or fax using the following procedures.

3. On the **License** page, paste the new activation key in the **Activation Key** box, and then click **Activate**.

Note The serial number begins with a number, whereas the activation key begins with a letter. Both use only capital letters. The number "0" and the letters "O" and "I" are not used.

To obtain an activation key by e-mail

Send an e-mail message containing the following information to *licenses.ms@thermo.com*:

1. Type **license request** in the subject line of the e-mail, and paste the text from the License page Serial Number box into the body of the e-mail.

2. Locate the bar code on the back of the Mass Frontier CD case. Type the serial number that appears below the bar code into the body of the e-mail message.

3. In the body of the e-mail message, include your name, company name, and phone number, and provide the software version, for example:

   Mass Frontier 5.0
   Full Version Single License

To obtain an activation key by fax

1. On the Windows desktop, double-click the **Xcalibur** shortcut icon to open the Xcalibur Home Page.

2. Choose **Tools > Configuration** to open the Xcalibur Configuration dialog box.

3. On the **Customer Information** page, enter all the information that a Technical Support representative might need to contact you.

4. Place the pointer below the last line of text in the **Address** box, and then press CTRL+V to paste the serial number from the Windows clipboard into the Address box.

5. Below the serial number, specify the product name and license version that you want, for example:

   Mass Frontier 5.0
   Full Version Single License

6. Click **Print User Info** to automatically create the form you need to fax or mail to Technical Support. Click **OK**.

7. Send the printed request form to Thermo Electron using one of the following fax numbers:

   USA or Canada: 1-408-965-6120
   International: international access + 1-408-965-6120

Starting Mass Frontier

To start Mass Frontier, double-click the **Mass Frontier** shortcut icon on the Windows desktop, or choose
Start > All Programs > HighChem Mass Frontier 5.0 > Mass Frontier

To add Mass Frontier to the Xcalibur Home Page Window Tool menu or the Qual Browser window Tool menu, choose **Tools > Add Tools** from the appropriate Xcalibur window. The Add Programs to Tool Menu dialog box appears.

Program Limitations

Mass Frontier offers a number of features. However, there are some limitations.

- Mass Frontier deals primarily with small organic structures rather than peptides and other biologically-related molecules. The Structure Editor and other modules dealing with structures have a limit of 199 non-hydrogen atoms per structure. If you try to exceed this number, a message box appears, reminding you of this limit.
- Mass Frontier uses pure substances only. Mixtures are not accepted. The program considers a mixture to be two or more structures, depicted in the same window, that are not connected by a bond (represented as a line). If you try to generate fragments and mechanisms from a mixture, the message box alerts you that this action is not permitted. The Check Structures option also detects mixtures as an error. However, library utilities support mixtures, to assure backward compatibility with commercial libraries. Mixtures can also be added to a user library. Fragmentation Library does not support mixtures.
- The automated generation of fragments and mechanisms has been greatly extended by numerous new aspects, but some restrictions still remain. Mass Frontier can select reagent gases for Chemical Ionization. However, the relative ionization potentials of reagent gases cannot be modified. Negative ionization (deprotonation) is only supported using the fragmentation library. There are no general rules for negative ionization. Because soft ionization techniques are mainly low-energetic experiments, which often yield complex skeletal and random rearrangements, the predictability of these fragmentation and rearrangements processes is not as high as that attained by electron impact ionization. Improve the degree of predictability by using compound-specific fragmentation mechanisms in the Fragmentation Library module.
- Use Mass Frontier with neutral and single charged molecules. As a consequence, you can attach the charge symbol (+ or -) to only one atom. If the charge multiplicity is described in general. use the unspecified charge location option in Structure Editor ($[M+2H]^{2+}$, $[M+3H]^{3+}$). Biradicals are not supported by any module.
- Mass Frontier supports high resolution mass spectra with a *m/z* range of 1-3000 mass units. Mass spectra containing peaks with a mass-to-charge ratio higher then 3000 are not displayed. Classification modules allow only 800 peaks per spectrum. If a spectrum contains more than 800 peaks, the classification procedure selects the 800 most prominent peaks.

- Classification methods supports only low resolution spectra. If you try to classify high-resolution spectra, the program automatically converts them into low-resolution spectra.

New Features in Mass Frontier 5.0

- Automated report creation in the Report Creator module
- Fragmentation Library in SQL Server format
- Retention time search in Database Manager module
- Component detection on chromatograms with multiple experimental settings (polarity switching, IT/FT, direct infusion MS^n)
- Three new component detection algorithms
- Advance chromatogram filtering (four algorithms)
- Chromatogram baseline correction (five algorithms)
- Chromatogram smoothing (three algorithms)
- Chromatogram three-dimensional view
- Automated spectra annotation capabilities with predicted fragments
- Formula Generator module
- Full support of Thermo Electron MS data, including data from the OrbiTrap™ and LTQ FT™ mass spectrometers
- Unspecified bond location in Structure Editor
- Spectral and fragmentation libraries automatic backup and restore procedures
- Fragmentation Library™ with more than 100 000 mechanisms
- Spectral Tree™ MS^n library of human and veterinary pharmaceuticals, endogenous metabolites, drugs of abuse and doping agents (ESI +/-)
- OCX version (optional)

Modules Overview

Mass Frontier 5.0 consists of the following modules:

- Structure Editor
- Database Manager
- Chromatogram Processor
- Spectra Classifier
- Fragments Comparator
- Fragmentation Library
- Isotope Pattern
- Periodic Table
- Formula Generator
- Report Creator
- Spectra Projector
- Neural Networks
- Fragments & Mechanisms

The Fragments & Mechanisms, Spectra Projector, Neural Networks, Fuzzy Clustering, Components Editor, and Hit Selector modules cannot be opened directly from the program desktop. These modules must be generated from data you supply.

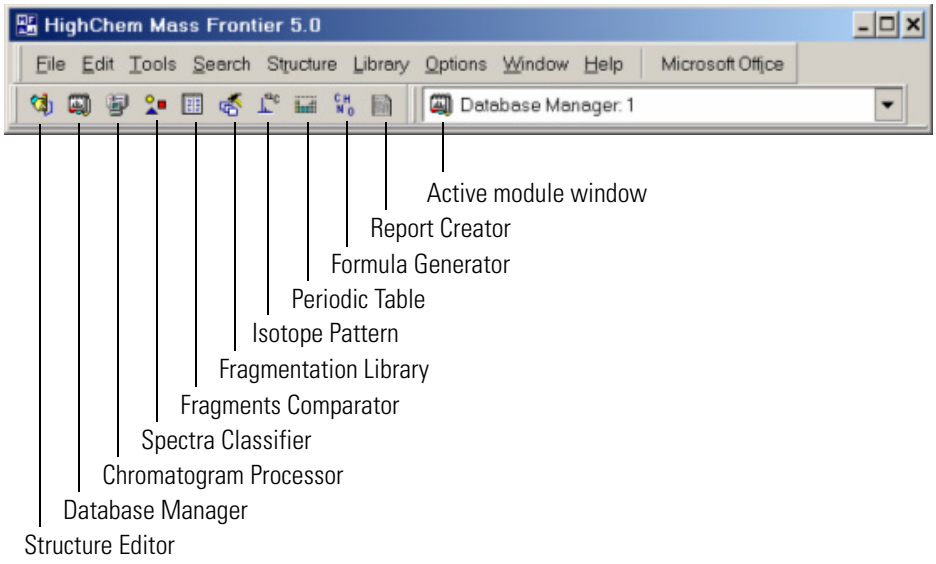


Figure 2. Mass Frontier modules toolbar

Structure Editor (Chapter 2) is a structure drawing tool that automatically calculates the mass of a selected fragment and the corresponding loss. Many chemical structures created in this module are used throughout the program.

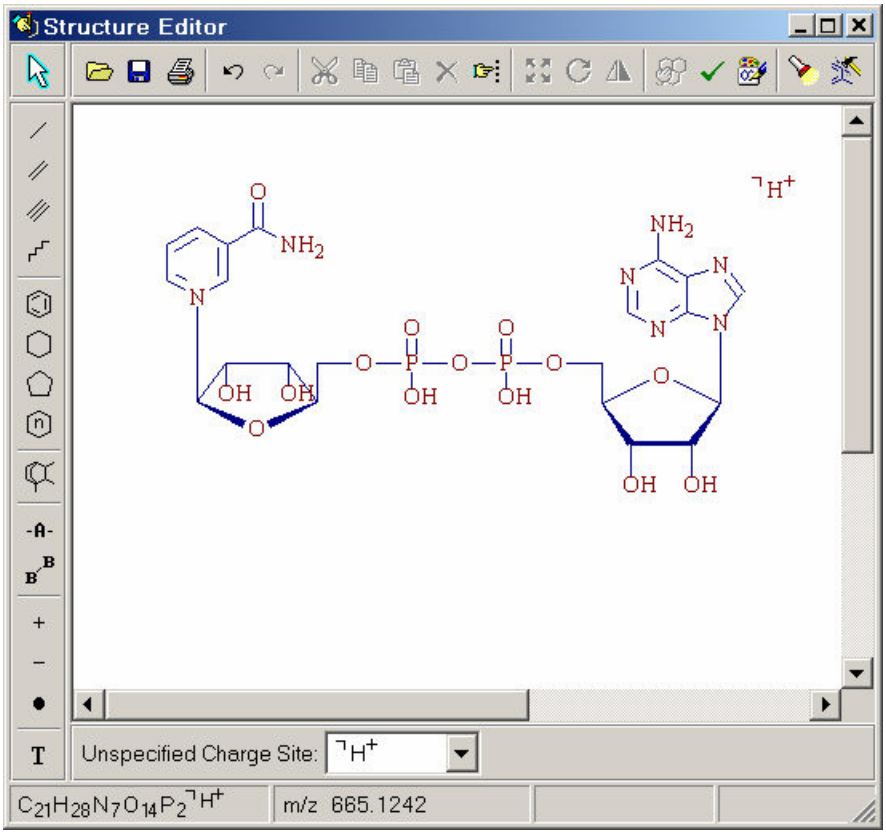


Figure 3. Structure Editor window

Database Manager (Chapter 4) provides a number of ways for organizing and processing mass spectra, chemical structures and libraries. A spreadsheet format is provided for data handling. Advanced database query and search features give you instant access to the information needed for rapid compound identification. User libraries containing spectra, MS^n trees, chromatograms, chemical structures and extensive compound characteristics can be created with a mouse click. This module supports data exchange with Microsoft Excel®.

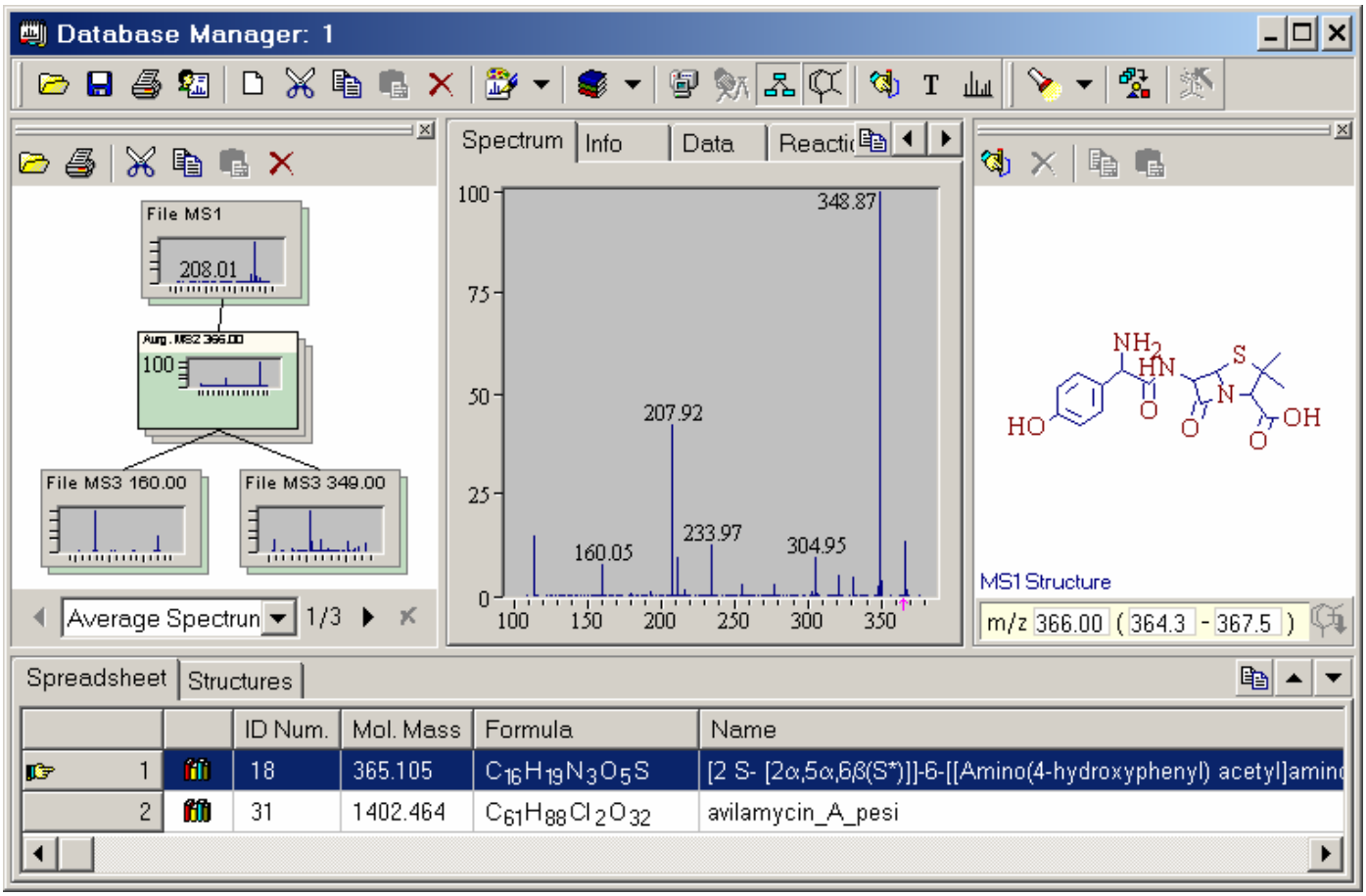


Figure 4. Database Manager window

Use the ***Chromatogram Processor*** (Chapter 13) to extract and process mass spectral scans of hyphenated chromatographic techniques such as GC/MS or LC/MS. Component detection and spectra deconvolution algorithms enable the automated extraction of individual spectra or MS^n trees from complex chromatographic data files. This module also provides visual tools for annotating particular components or chromatographic peaks.

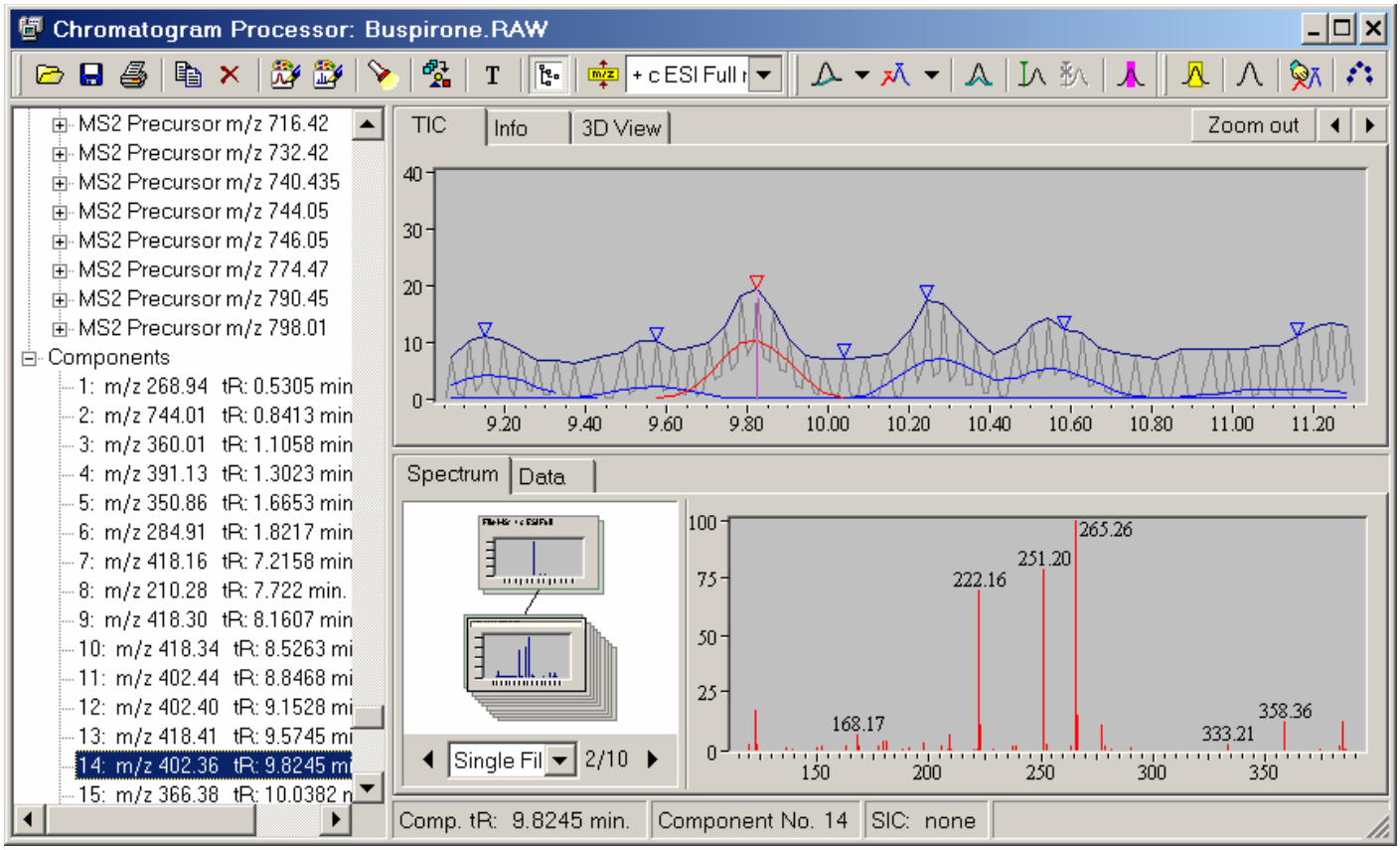


Figure 5. Chromatogram Processor window

Use the ***Spectra Classifier*** (Chapter 10) to retrieve and organize spectra intended for classification. Because spectra can be classified according to various criteria (structural, physical and other properties), it is useful to organize the spectra into different groups. Such groups of spectra can be visually represented in different ways (using colors, symbols, and numbers) to highlight similarities or dissimilarities among the spectral groups you choose.

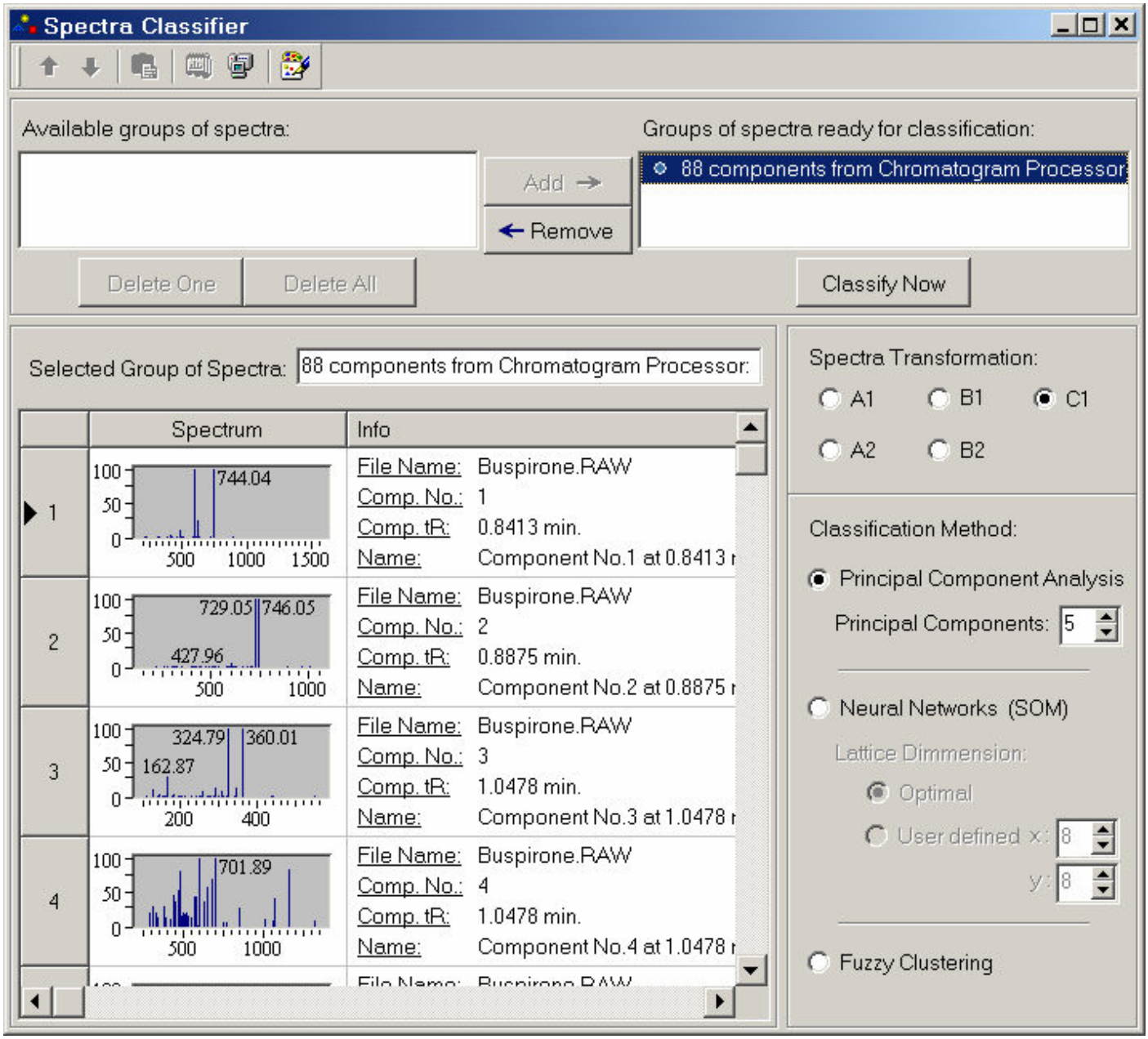


Figure 6. Spectra Classifier window

Fragments Comparator (Chapter 8) displays a series of fragments in a table format. Columns are made up of individual compounds and rows show either mass-to-charge ratios or the structures of fragments. Use this module to compare the product ions of analogous molecules.

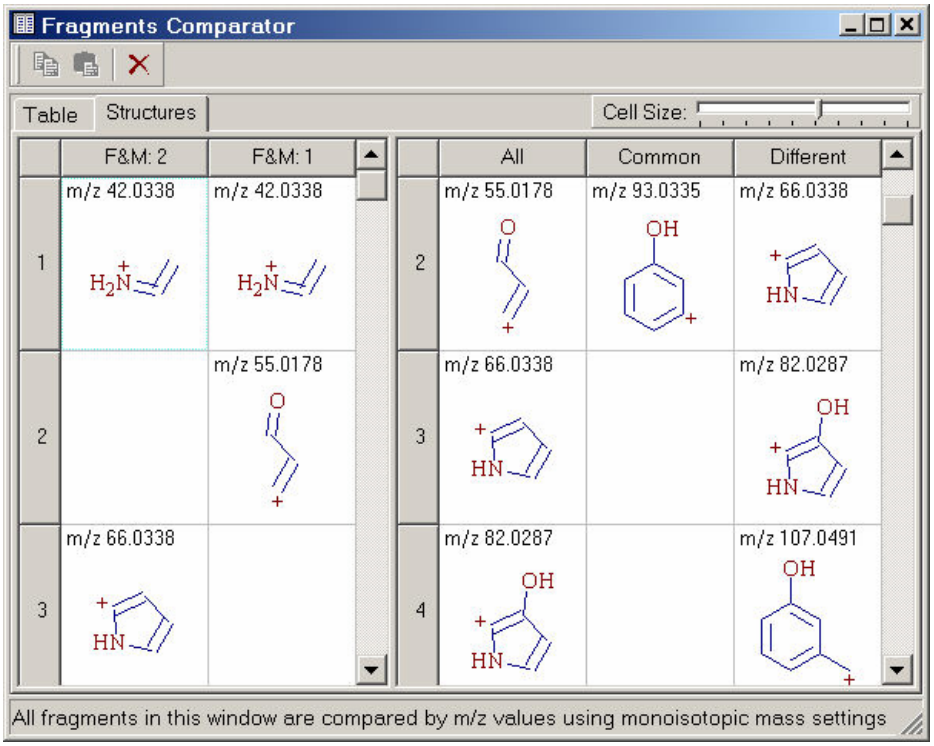


Figure 7. Fragments Comparator window

Isotope Pattern displays the isotopic profile whenever a structure or fragment is selected in the program. Isotope patterns can also be calculated from a molecular formula you supply.

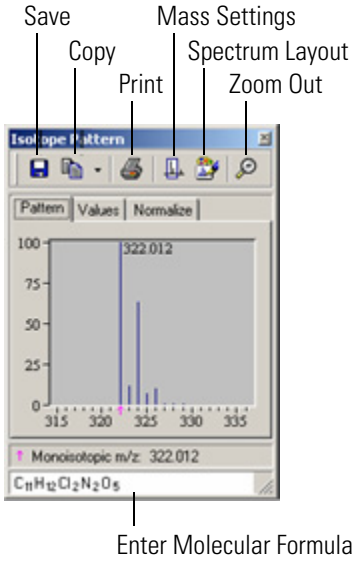


Figure 8. Isotope Pattern window

Use the ***Periodic Table*** to display the terrestrial isotopic abundance of elements and their multi-atomic isotopic profiles.

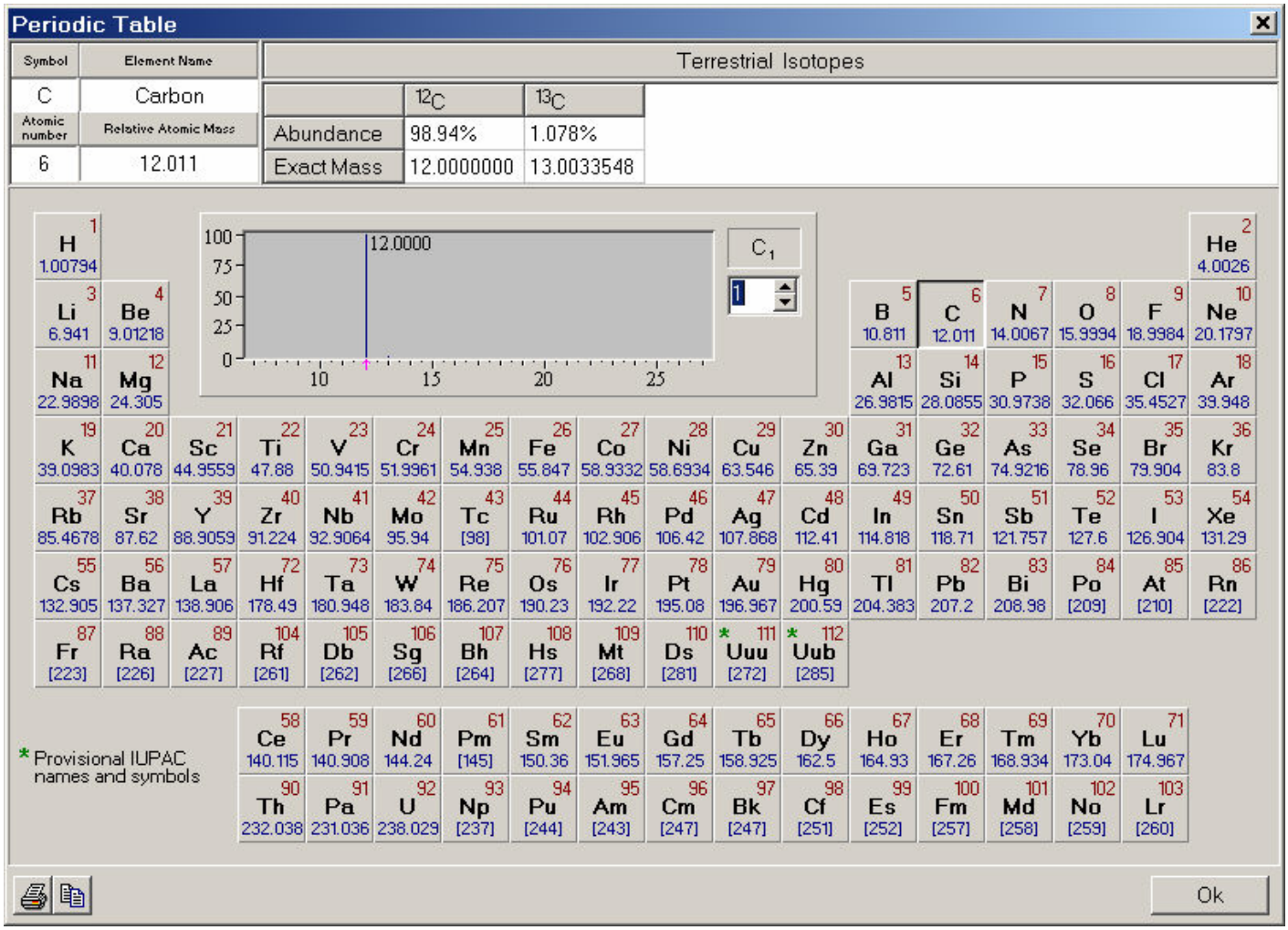


Figure 9. Periodic Table window

Use the ***Formula Generator*** (Chapter 17) to a calculate list of theoretical molecular formulas that best fit an *m/z* value.

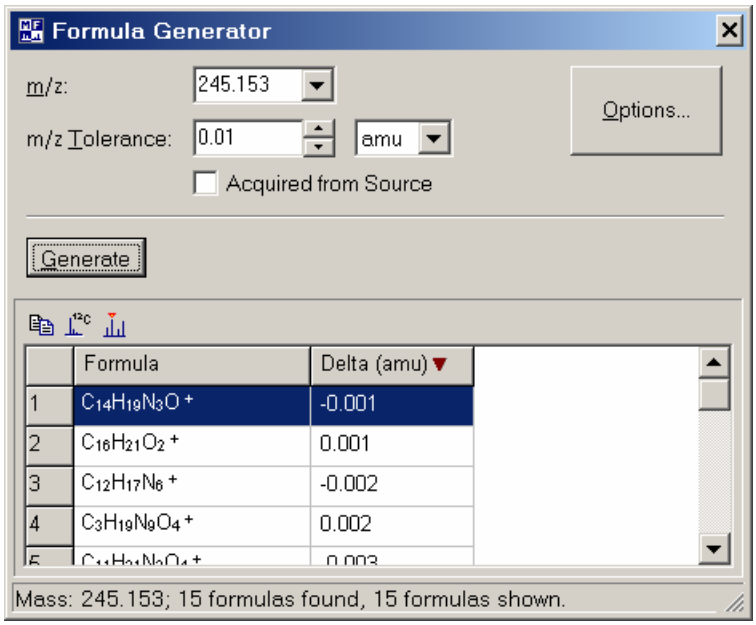


Figure 10. Formula Generator window

Use the ***Report Creator*** (Chapter 18) to create customizable reports from modules displayed on the screen. Reports can be printed or exported as PDF files.

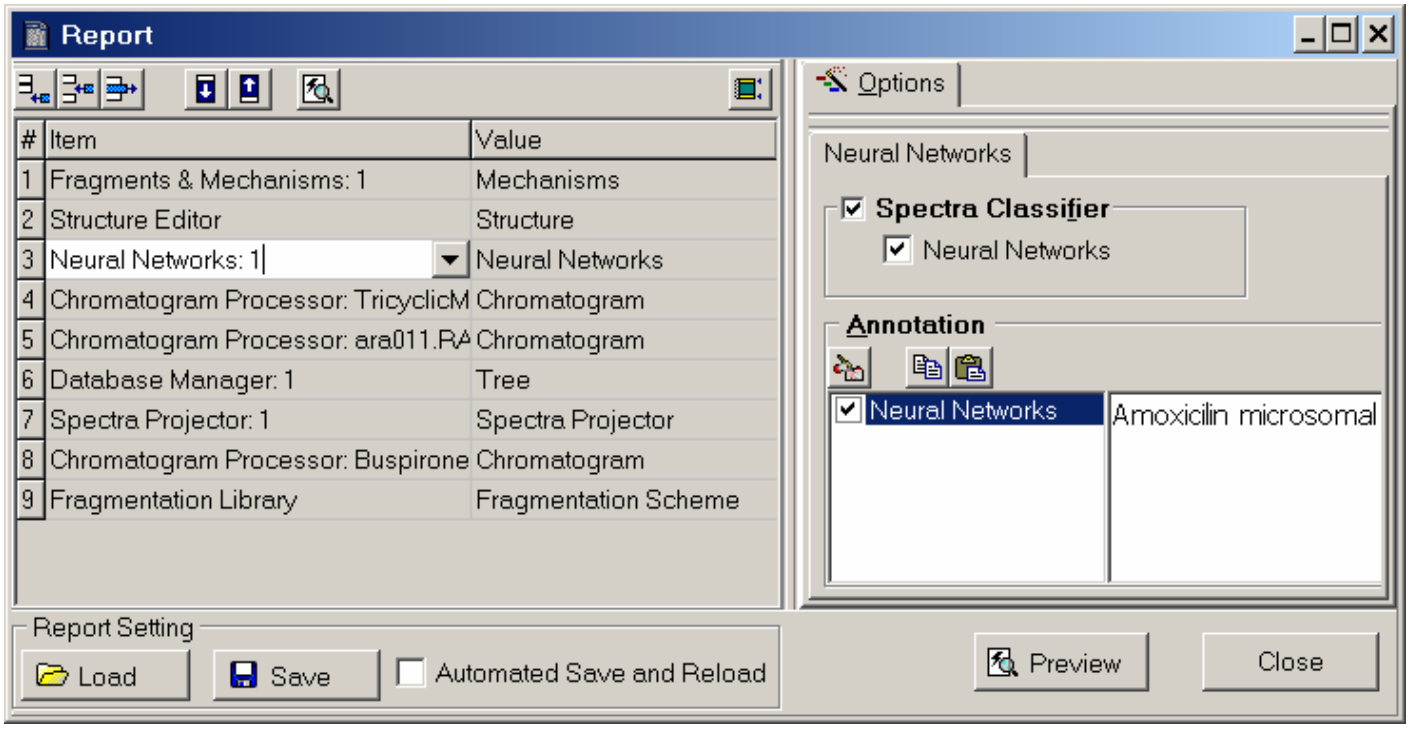


Figure 11. Report Creator window

Spectra Projector (Chapter 11) displays the results of Principal Component Analysis and Fuzzy Clustering classification methods. Mass spectral data can be classified using two-dimensional (2-D) or two-dimensional (3-D) projections in which each point represents a spectrum. If the class membership of an unknown spectrum needs to be determined, open or paste it into an existing projection.

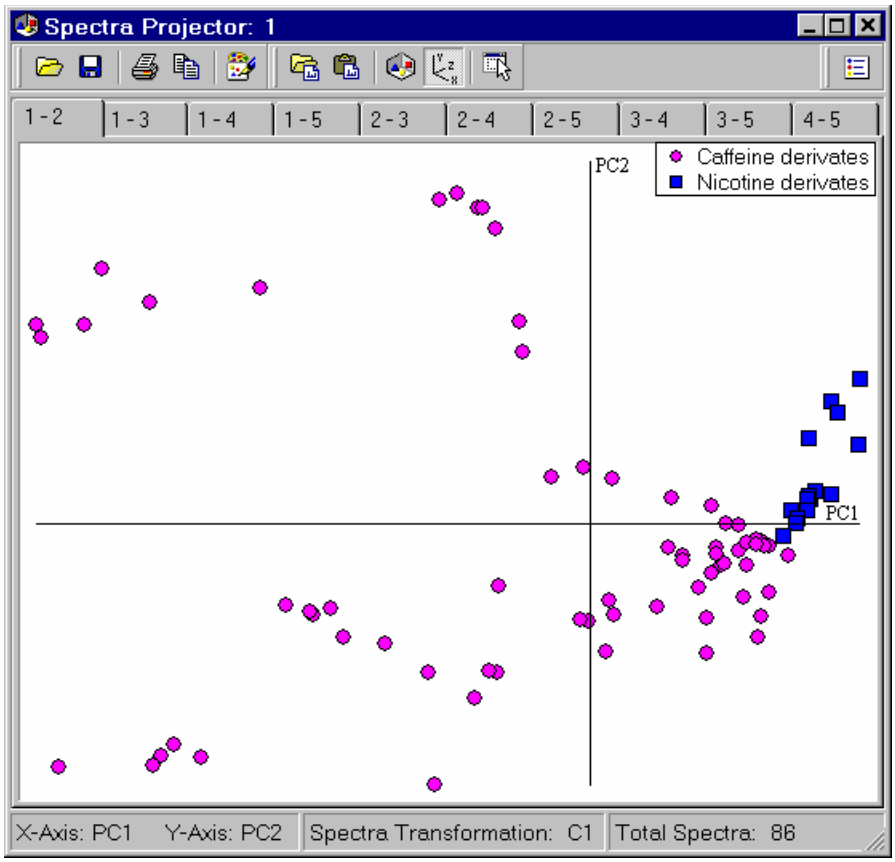


Figure 12. Spectra Projector window

Neural Networks (Chapter 12) is an additional classification strategy in Mass Frontier. Mass spectra are classified using a powerful method called Self-Organizing Maps (SOM), which are a special class of neural network. If two or more spectra activate the same neuron, the corresponding compounds will exhibit similar physical or chemical properties, or biological activities.

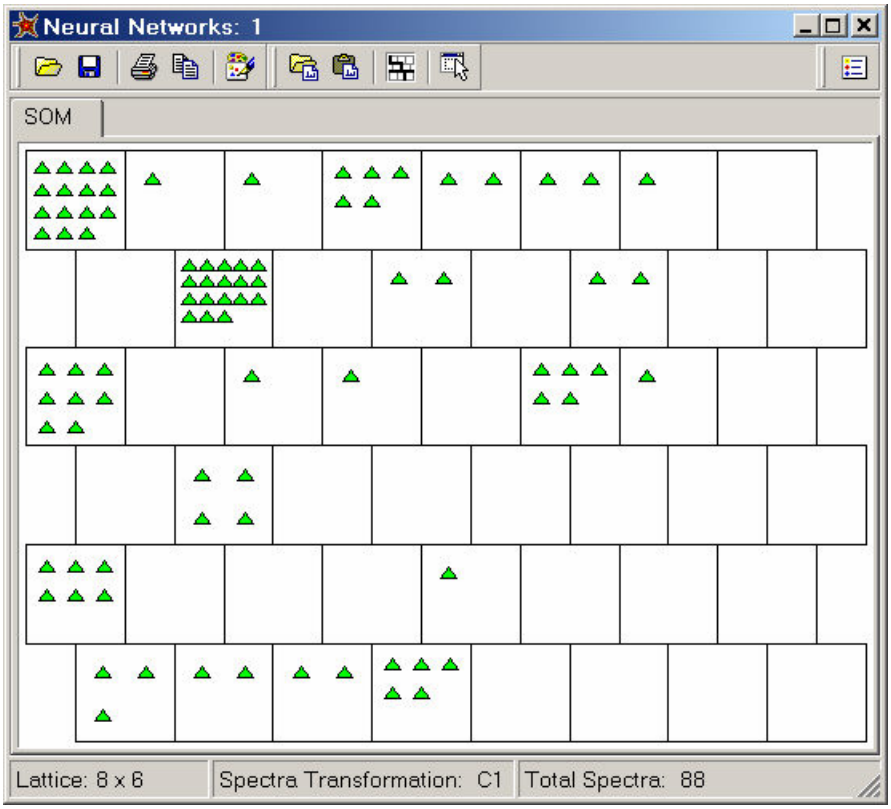


Figure 13. Neural Networks Window

The ***Fragments & Mechanisms*** (Chapter 6) module is an expert system for automated generation of fragments and detailed fragmentation and rearrangement mechanisms from a chemical structure you supply. This module consists of a system that includes a set of known general reaction mechanisms and comprehensive collection of library mechanisms which enable automated prediction at an expert level. This module can be generated either from the Structure Editor or the Database Manager module.

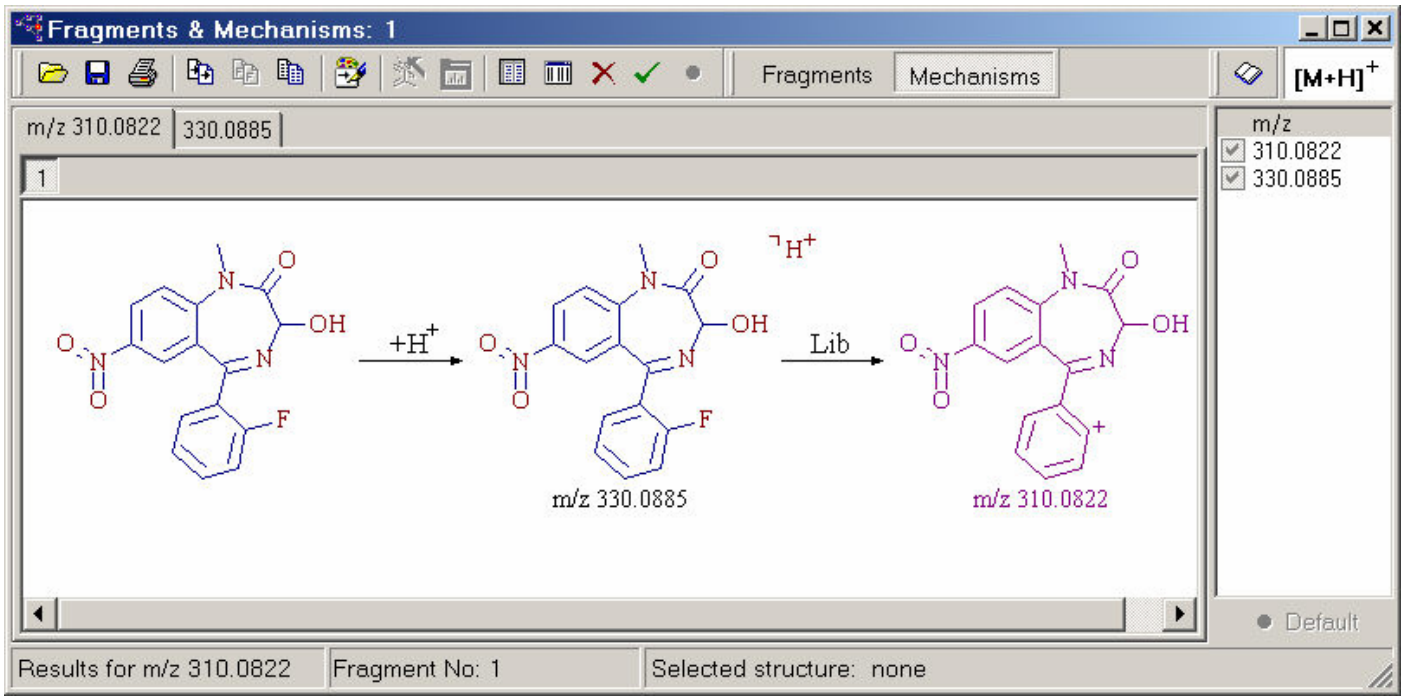


Figure 14. Fragments & Mechanisms window

Use the ***Fragmentation Library*** (Chapter 7) module for the creation and management of fragmentation mechanism databases. This module contains an expert system that automatically extracts a decomposition mechanism for each fragmentation reaction in the database and determines the compound class range that the mechanism can be applied to. Mass Frontier uses this expert system to apply database mechanisms to a user provided structure and automatically predicts the fragmentation reactions for a given compound.

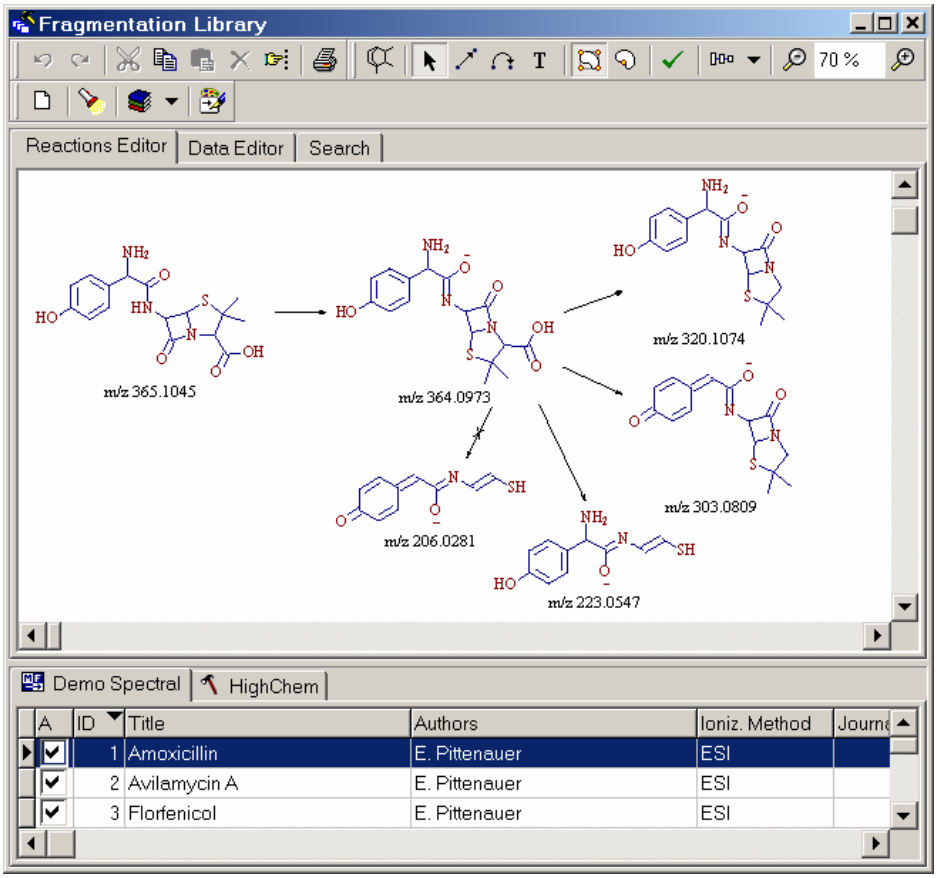


Figure 15. Fragmentation Library window

Chapter 2 Structure Editor

Mass spectra reflect the structural features of molecules which are essential for the interpretation and the investigation of structure-spectra relationships. Mass Frontier incorporates the Structure Editor structure drawing tool, which enables the interactive handling of all kinds of structural information. Use the Structure Editor for editing, importing, exporting and checking chemical structures. The Structure Editor is the gateway to four other modules in this program: Database Manager, Fragments & Mechanisms, Fragmentation Library, and Isotope Pattern.

This chapter contains the following topics:

Structure Editor Window

The Structure Editor Window is shown in Figure 16.

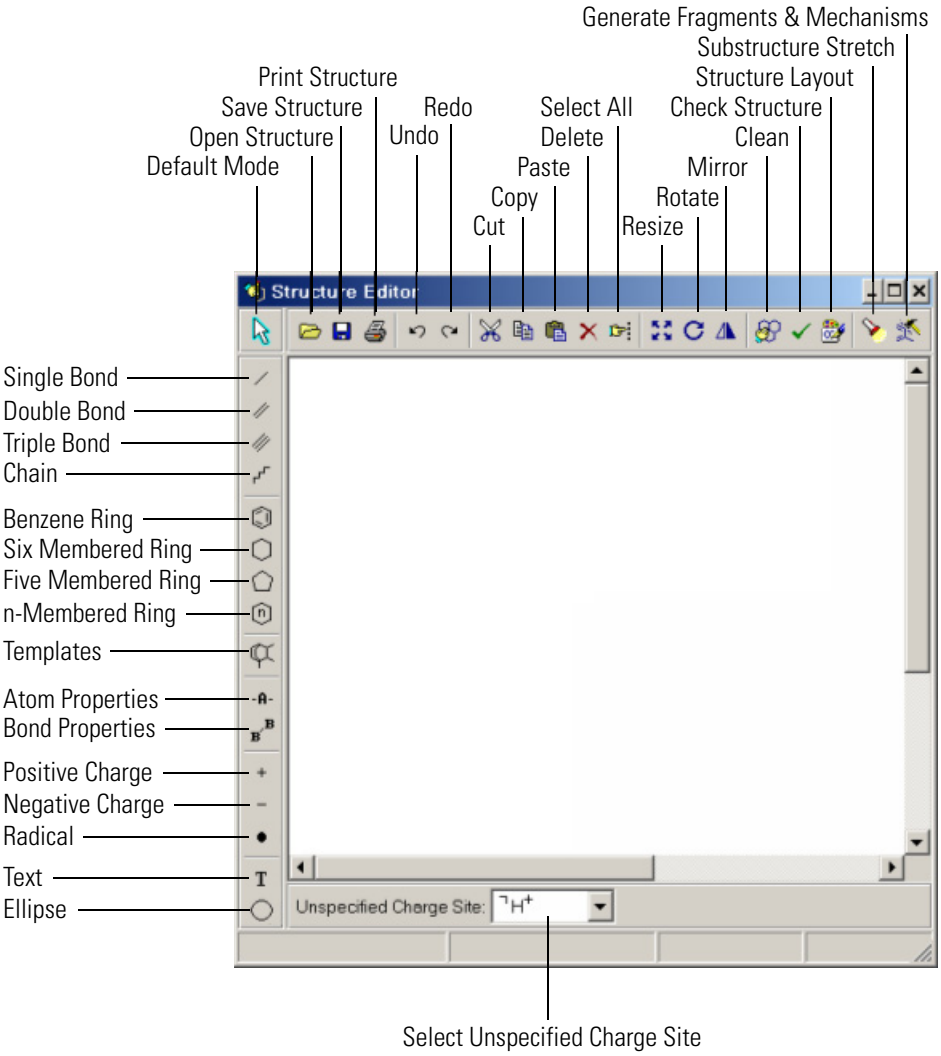


Figure 16. Structure Editor window showing names of icon commands

To open the Structure Editor window

- Click the **Structure Editor** button on the main tool bar.
- Or choose **Tools > Structure Editor**.
- Open a structure by choosing **File > Open > Structure**. The Structure Editor starts automatically.

Note Only one Structure Editor window can be open at any one time in the program. If you click the Structure Editor button, or choose **Tools > Structure Editor** and the Structure Editor is already open, this window becomes active.

To begin drawing a chemical structure in Structure Editor, click a button on the vertical bar. When you click one of these buttons, the shape of the cursor changes to visually represent the engaged drawing mode. The vertical buttons, in contrast to the horizontal buttons, are not represented in the menu. If the function of a button is not apparent from its appearance, move the cursor over the button and a hint appears.

Restoring Defaults

In the Structure Editor's default state, all buttons are switched off and no atom or bond is selected. The plain cursor indicates that the Structure Editor is in default state.

To restore the default state of the editor

1. Click the **Default mode** button in the upper left corner in the Structure Editor window.
2. Right-click and switch off the activated button and deselect all atoms.

Opening and Saving Structures

To open or save a structure

1. Click the **Open Structure** or **Save Structure** button in the Structure Editor window.
2. Choose **File > Open > Structure** or **File > Save > Structure**.

If you are opening a file which contains more then one structure (.sdf file), only the first structure in the file is loaded into Structure Editor.

Mass Frontier is a 32-bit application, enabling you to use long names to save structures. You can also save structures by their actual names (for example, 1-Amino-2-hydroxyindane.mol).

Structure Data Formats

Structure Editor supports two kinds of structure formats: MDL MOL files, (SDF files), with the .mol (.sdf) extension, and HighChem MCS format (Maximal Compressed Structure), with the .mcs extension. These formats are also supported in the Database Manager module. Templates are stored in MCS format, using the .tml extension.

Mass Frontier features the ability to restrict a search by a set of structural constraints called the Good-Bad list. For example, you can instruct the program to conduct a library search comparing an unknown spectrum only with the spectra of ketones. This feature provides an endless range of possibilities to target your search results with. The Good-Bad structures are stored in the directory …\Constraints, and the structures are saved in MSC format with the .mcs extension. The program automatically retrieves all MCS structures from the ..\Constraints directory and puts them in a Good-Bad box in the Constraints dialog window.

Structure Layout

With Mass Frontier. you can change almost anything for structures, as well as for other objects. Every layout setting change also affects printing and copying to the Clipboard, except background color, which affects only screen display in Mass Frontier. Use the various layout items to tailor the graphics to your individual report or publication needs.

By default, the symbols for hydrogen atoms attached to carbon atoms are not displayed (example a). To display them, select **Show Carbon Symbols** (example b) in the Structure Layout dialog window. See Figure 17.

To open Structure Layout dialog window

1. Click the **Structure Layout** button in the Structure Editor window.

2. Choose **Options > Structure > Layout**.

Hydrogen atoms are only displayed for carbons if the Show Carbon Symbols box is selected. Otherwise, corresponding hydrogens are displayed for heteroatoms only.

Note If you draw nonisotopic explicit hydrogen atoms (see example c in Figure 17) these are removed in the Fragments & Mechanisms window because they can make the mechanism network unclear, especially for complex hydrogen rearrangement steps.

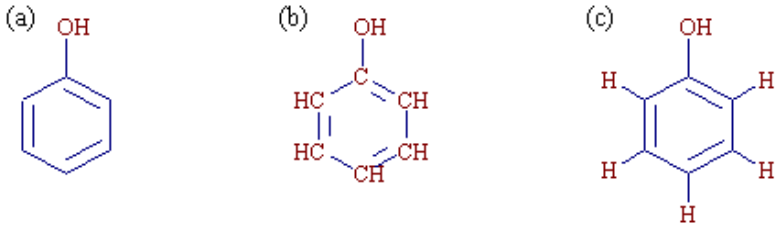


Figure 17. Structures showing different displays of hydrogens

The structure layout settings apply to all structures in Mass Frontier simultaneously. This means that if you change a structure layout item, all structures in the Structure Editor, Database Manager, Fragments & Mechanisms, and Fragments Comparator modules are affected.

Note If you are printing in black and white and have set bright colors for bonds or atoms, the lines and fonts might appear indistinct. To avoid this, specify darker colors for all structural items, including spectra, chromatograms, and mechanisms.

Text

Structure Editor offers the possibility of labeling a structure or displaying a text note on the screen or on the printout.

To enter a text note

1. Click the **Text** T button in the Structure Editor window.
2. Click anywhere in the drawing area to place the text.
3. Type the desired text.
4. Confirm the text by clicking outside the text area or click any button in Structure Editor.

You can create up to 127 separate text notes. If you want to change the font, color, size, or background of the text notes, use the Structure Layout window on the Text tab.

Note Text notes are not associated with structures. As a result, the Open, Save, Copy, and Paste actions apply only to structures. When these actions are applied, the text notes are ignored even when a structure is selected together with a text. Additional structure handling routines such as resizing, rotating, or mirroring can be performed only on structures.

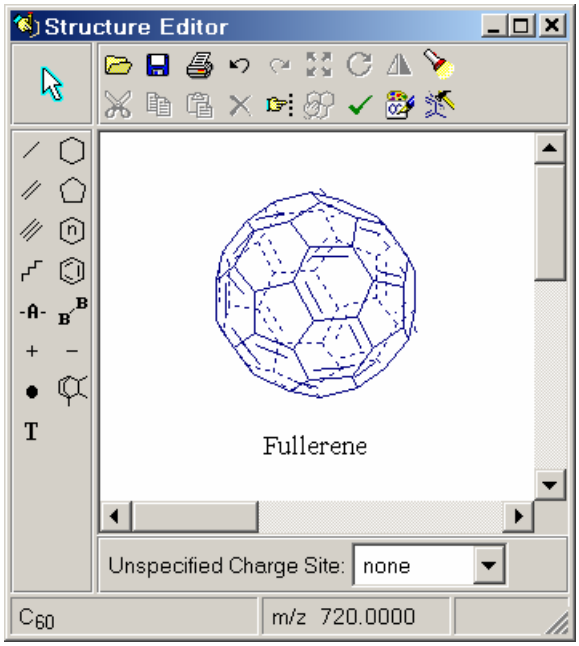


Figure 18. Structure labeled with its chemical name

Template Structures

When you click the **Templates** button in the Structure Editor window, the Template dialog window appears. Mass Frontier comes with more then 200 predefined templates.

To insert a template structure into Structure Editor

1. Select a group of templates in the directory list box by using the arrow keys on the keyboard or click the appropriate name of the group.
2. Click any atom or bond, depending on whether you want to attach the template to an atom or a bond of a structure in Structure Editor.
3. The Template dialog window disappears and you can place or attach the chosen template.
4. Switch off the template button or restore the default state of the Editor.

You can create your own group of templates or add a structure to an existing group. The templates are organized by directory. The template root directory is …\Templates. Every group of templates is stored in a separate subdirectory of the template root directory. Subdirectories are named after compound groups (for example, Steroids). The files within each subdirectory are named after actual structures using the .tml extension (for example, Cholesterol.tml). When you save a structure for template purposes select the Template format with the .tml extension in the Save Structure dialog window.

To build your own templates

1. Draw a template structure.
2. Click the **Save Structure** button in the Structure Editor window or choose **File > Save > Structure**.
3. Choose **Template format** in the Files of type: box in the Save Structure dialog window.

Selecting Atoms and Bonds

Any modification that you make to a structure applies only to the selected atoms or bonds In addition, when a substructure search is initiated, the program automatically uses the selected substructure in Structure. Before you select one or more atoms, restore the default state of the editor.

To select a group of atoms that are next to each other

While pressing the mouse button, drag a rectangle around the atoms you want to select.

The Windows convention for selecting multiple items applies. To select atoms at different locations, use the keyboard Shift key. You can select a group of atoms that are not adjacent in one of two ways:

- Click the atoms you want to select while holding down the Shift key.
- Or hold down the mouse button and drag a rectangle around the atoms you want to select while holding down the Shift key.

To select all of the atoms and bonds in the structure:

- Click the **Select All**  button in the Structure Editor window or choose **Structure > Select All**.
- Or double-click anywhere in the draw area within Structure Editor, except on atoms and bonds.

Structure Editor offers two selection modes: Rectangle Selection and Lasso Selection.

To choose the selection mode

- Right-click the **Default Mode** button and choose the appropriate selection mode from the popup menu.
- Or right-click anywhere in the draw area within Structure Editor, and choose **Rectangle** or **Lasso Selection** from the pop-up menu that appears in the draw area.

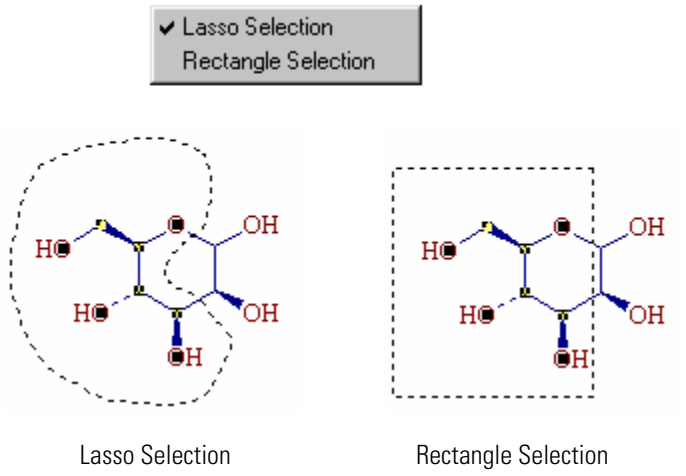


Figure 19. Lasso and rectangle selection in Structure Editor

Atom Properties

Use the Atom Properties dialog window to change the charge state or isotope of an atom, or to change the element entirely.

To open the Atom Properties window

- Click the **Atom Properties** -A- button in the Structure Editor window and then click the atom you want to change.
- Or restore defaults, and then double-click the atom you want to change.

In the Atom Properties dialog window, make changes by clicking the appropriate element button, charge and radical box, or nucleon number box.

Note All changes carried out in the Atom Properties dialog window affect only a single atom.

To change an element that has a single character symbol, such as C, H, N, O, B, F, K, P, S, I, V, W, Y, U and R; select all the atoms that you want to change and click the appropriate key on the keyboard. All the selected atoms transform into the element you have chosen.

You can set chlorine (Cl) or bromine (Br) atoms by selecting all the atoms that you want to change and click either the C (for chlorine) key or B (for bromine) key on the keyboard, while holding down the Shift key.

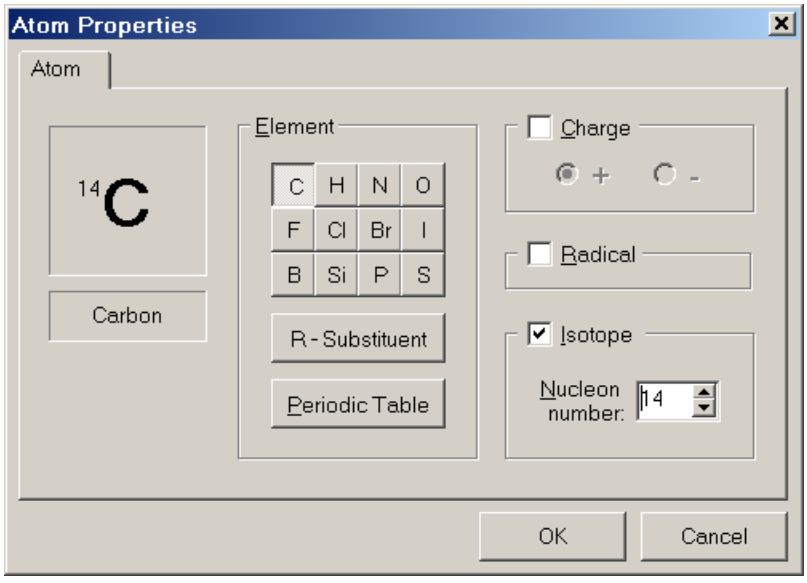


Figure 20. Atom Properties window showing ^{14}C atom

You can use a substituent instead of a specific element. A substituent is any atom, functional group, or substructure substituted for another, or entering a structure in place of some other part which is removed. The symbol ”R” represents a substituent. A substituent can be with or without index. Substructure search and fragment-generation algorithms consider substituents with identical indexes as equal and substituents with different indexes as not equal.

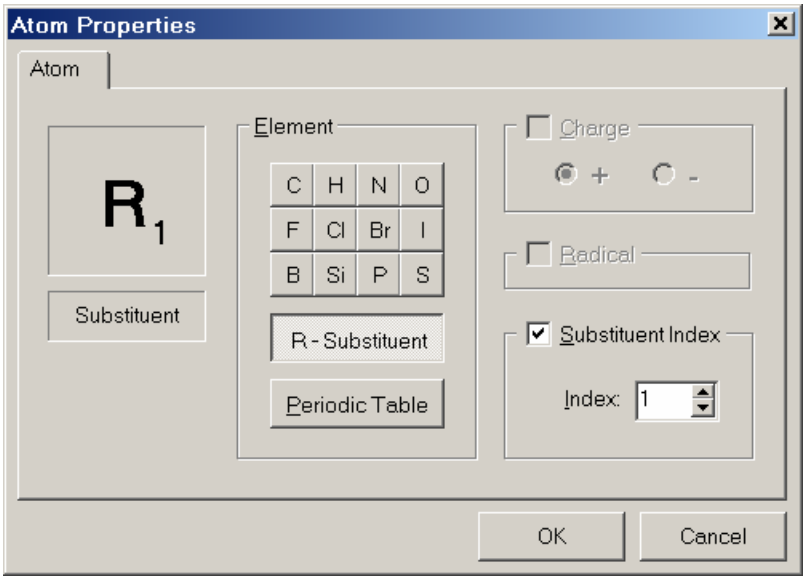


Figure 21. Atom Properties window showing substituent R_1

Bond Properties

The Bond Properties include bond multiplicity, bond style, and bond color.

To change the multiplicity of a bond, click the , , or button in the Structure Editor window and then click the bond you want to change.

To change the color or optical orientation of a bond, use the Bond Properties dialog window.

To open the Bond Properties window

- Click the **Bond Properties** button in Structure Editor and then click the bond you want to change. The Bond Properties dialog window appears.
- Or restore defaults, and then double-click the bond you want to change. The Bond Properties dialog window appears.

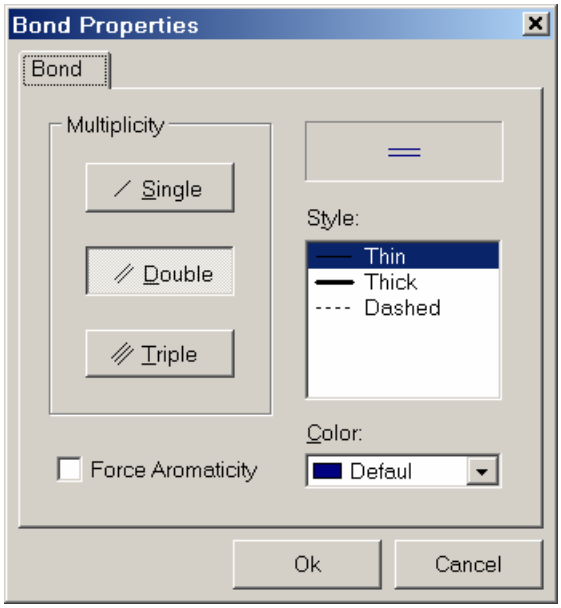


Figure 22. Bond Properties window

Mass Frontier automatically recognizes aromatic bonds in an appropriate six-membered ring or in polyaromatic structures. However, if unusual semiaromatic or aromatic resonance structures are required, the aromatic bond can be forced to select bonds by selecting Force Aromaticity.

Copying Structures

Mass Frontier supports extensive use of the Windows Clipboard for the exchange of structural information between modules. In addition, copy and paste functions can be used inside Structure editor. To draw larger structures efficiently, use the copy and paste functions.

To copy a structure or part of a structure to the Clipboard

1. Select the structure or part of the structure you want to copy.

2. Click the **Copy** button in the Structure Editor, or choose **Edit > Copy**.

Note Only the selected atoms and their associated bonds are copied.

In addition to structure exchange between modules, Mass Frontier allows structure export to other programs that deal with structural information. When you copy a structure, Mass Frontier automatically copies two different formats to the Clipboard: structural information in MOL format and graphics in Windows metafile format. If you paste a structure into the structure editing software, the MOL format is used. If you paste the structure into any text editor, spreadsheet or program that works with graphics, the graphical information is used. All these actions occur automatically.

Pasting Structures

If you copy a structure or fragment anywhere in the program or in a third party structure drawing tool, you can paste it to Structure Editor. If necessary, the structure can be changed or corrected and then returned to where it originated which is useful for structure elucidation. For example, you can copy a structure from the Database Manager window, paste it to Structure Editor, make appropriate changes, and then move it back to Database Manager. If the spectrum and the structural proposal are not consistent, repeat the process.

To paste a structure to Structure Editor

Click the **Paste** button in the Structure Editor. Or, choose **Edit > Paste**.

If you have copied a structure or fragment in a program other than Mass Frontier, you can only paste this structure if the external structure drawing software supports MOL format and this format is activated. The majority of structure drawing tools support MOL format and have this format activated by default. If you paste a structure from an external source, it might appear larger in Mass Frontier than in the original software. If this occurs, make the structure smaller by using the Resize tool.

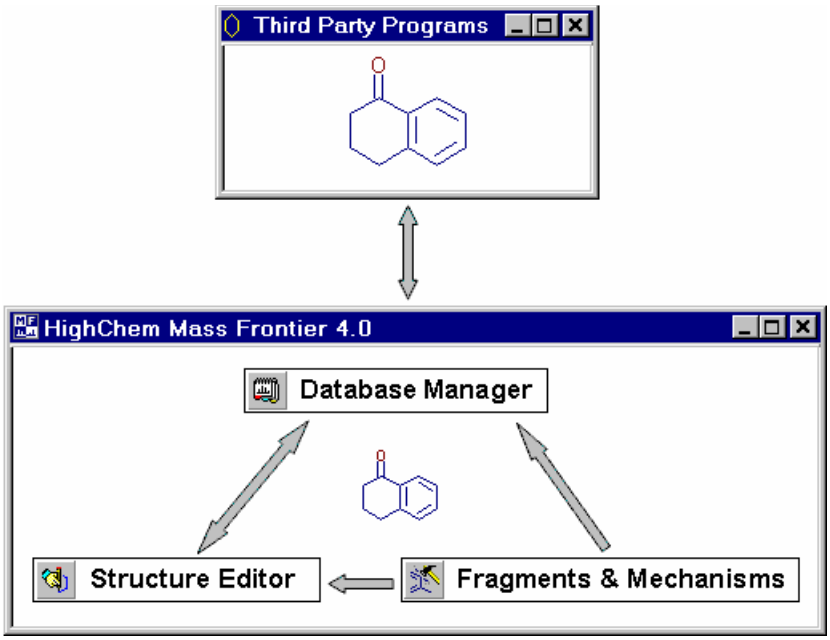


Figure 23. Structure Exchange by using the Copy and Paste commands

Moving, Resizing, Rotating, and Mirroring Structures

To move a structure in the Structure Editor

1. Select the atoms or bonds to move.
2. Point the cursor at any selected atom or bond.
3. Hold down the mouse button and drag the selected structure to the new location.
4. Release the mouse button to drop the selected structure at the new location.

To resize a structure in the Structure Editor

1. Select the structure or part of the structure to resize.
2. Click the **Resize** button, or choose **Structure > Resize**.
3. Drag one of the small rectangles on the structure edge until the new size is achieved.
4. Release the mouse button.

> **Note** If you drag one of the diagonal rectangles, the aspect ratio is kept constant during structure resizing.

Use the Rotate Structure option to twist a structure in any direction. The center of rotation, indicated by a small circle with a cross in the middle, can be moved to any location.

To rotate a structure in the Structure Editor

1. Select the structure or part of the structure you want to rotate.
2. Click the **Rotate** button, or choose **Structure > Rotate**.
3. Move the center of rotation to the desired position by dragging the circle with a cross.
4. Drag any of the small rectangles on the structure edge to achieve the new angular position.

To make a mirror image of your structure in the Structure Editor

1. Select the structure or part of the structure you want to mirror.
2. Click the **Mirror**  button, or choose **Structure > Mirror**.
3. Click one of the small rectangles on the structure edge.

Cleaning Structures

Use the Clean function to achieve a professional look for your structures. With Mass Frontier, you can clean up an individual part of a structure. For example, you can restrict cleaning to certain functional groups, while the main skeleton remains intact. However, the algorithm of cleaning 2-D structures is a particularly difficult mathematical problem and has yet to be completely solved. As a result, this function might, in some complicated cases, lead to unsatisfactory structures. If this occurs, use the Undo function.

To clean a structure

1. Select the structure or part of the structure to clean up.
2. Click the **Clean** button, or choose **Structure > Clean**.

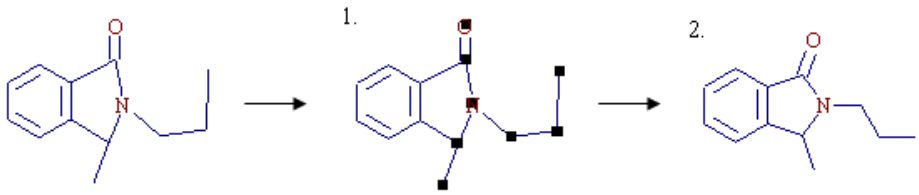


Figure 24. Structures showing the results of using the Clean command

Note If you want to clean only part of a structure, the selected atoms must be connected or a message box appears to remind you that only connected atoms can be cleaned.

Checking Structures

Structure Editor comes complete with a function for checking chemical structures. Structure Checker searches for formal errors and unusual structural features. If a structure is formally incorrect, or Structure Checker considers there is some doubt about its validity, a Structure Check Results window appears with a list of errors and warnings. When this window is closed the program automatically selects the atoms and bonds, which are considered incorrect. As mentioned in "Program Limitations" on page 9, structures that are not connected are considered to be mixtures, which are reported as errors.

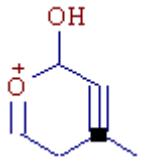


Figure 25. Structure showing the result of the Check Structure command

To check a structure

1. Click the **Check Structure** button in the Structure Editor
2. Choose **Structure > Check Structure**.

Note This option does not perform quantum mechanical or thermodynamical calculations concerning possible structure stability.

After finishing a structure drawing, always check it for errors before proceeding with any other procedure. Once fragments and mechanisms generation is initiated, a structure is automatically checked for errors. If any error is discovered, the program prevents you from continuing with the generation.

Before running Generation of Fragments and Mechanisms and after finishing structure drawing in Fragmentation Library window structure check is automatically performed.

MS Calculations

When you select a part of a structure, the Structure Editor automatically displays the molecular formula and molecular mass of the selected atoms (F:) in the status bar of the Structure Editor, together with corresponding loss (L:). See Figure 26. Use this information for simple consistency checking of mass spectrum and chemical structure.

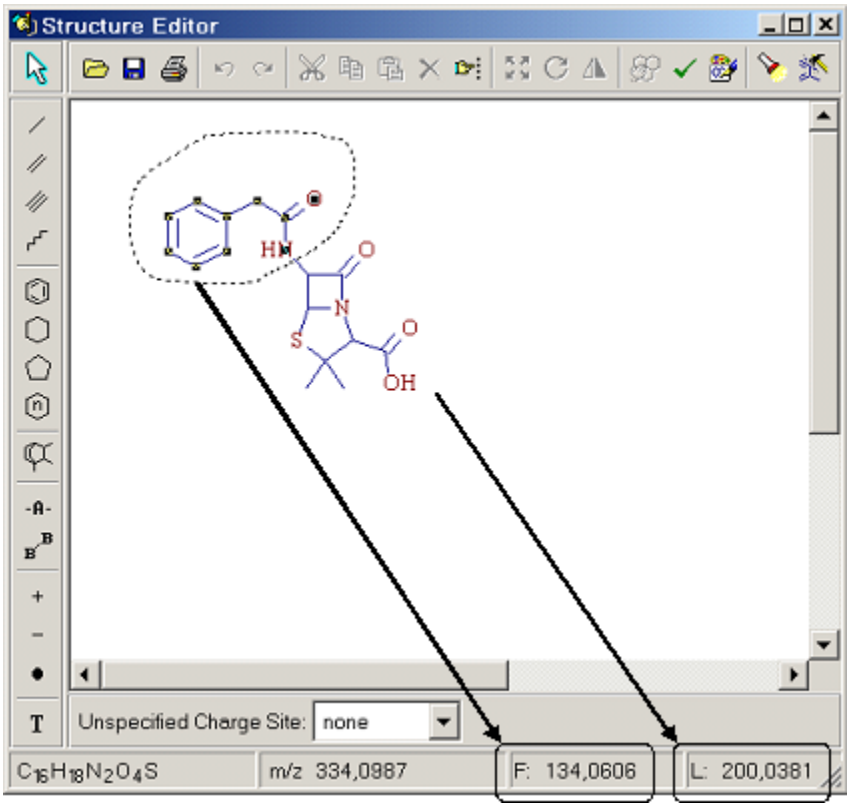


Figure 26. Structure Editor showing the result of selecting part of a structure

Unspecified Bond Location

Mass spectrometry is often unable to provide information about the exact position of a functional group on a chain or ring portion of a structure. In a number of application areas, even incompletely characterized molecules can be sufficient to study a particular phenomenon. A typical example is metabolite characterization in an early drug discovery process, where knowledge of the precise location of biotransformation action is of less importance.

To display and calculate monoisotopic mass or the isotopic pattern of a fragment or molecule with an unspecified bond location, the Structure Editor provides Ellipse as a graphical tool. Ellipse visually defines the region on a structure where a functional group could potentially be attached. In order for the software to correctly interpret an unspecified bond location, the ellipse must enclose one or more atoms of the core structure and a single atom of a functional group that is not attached to the core structure. The ellipse is green for better visual orientation. See Figure 27.

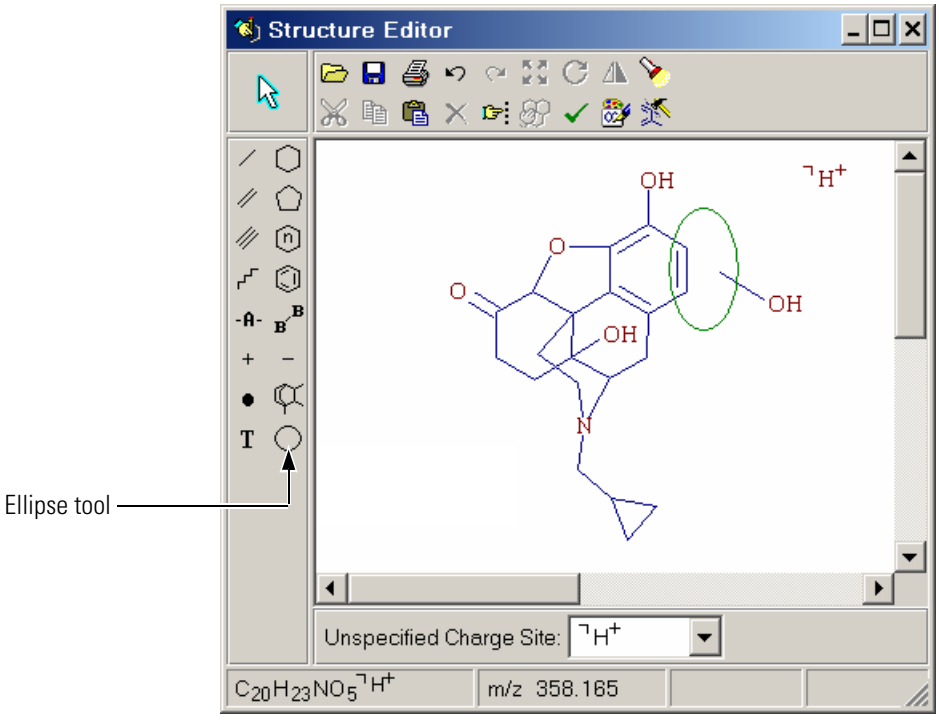


Figure 27. Structure Editor with unspecified bond location

Note The Ellipse tool is available only in the Structure Editor module, You cannot use a structure with an unspecified bond location in other modules. In addition, the generation of fragments and mechanisms from such structures is not supported.

Unspecified Charge Site

Use Mass Frontier to process structures where the charge site is not specified. This kind of molecule representation is important when working with ionic structures, because the favored ionization site or the explicit charge site during fragmentation reactions is often unclear. Several ion types can be chosen in the program.

To create an ionic structure with an unspecified charge site, select one of the ion types from the box at the bottom of Structure Editor. Every structure with an unspecified charge site has a specific symbol on the upper left part of the structure.

none	Unspecified charge location inactive
⌝ $+\bullet$	Radical cation
⌝ H^+	Protonation
⌝ $-H^+$	Deprotonation
⌝ $\bullet-$	Radical anion
⌝ NH_4^+	Ammonium cluster
⌝ H_3O^+	Hydronium cluster
⌝ Li^+	Lithium adduct
⌝ Na^+	Sodium adduct
⌝ K^+	Potassium adduct
⌝ CH_3^+	CI reagent gas CH_4 adduct ion
⌝ $i\text{-}C_4H_9^+$	CI reagent gas C_4H_{10} adduct ion
⌝ NH_2^+	CI reagent gas NH_3 adduct ion
⌝ OH^+	CI reagent gas H_2O adduct ion
⌝ NO^+	CI reagent gas NO adduct ion
⌝ $[M+2H]^{2+}$	Protonation doubly charge ion
⌝ $[M+3H]^{3+}$	Protonation triply charge ion
⌝ $[2M+H]^+$	Protonation dimer

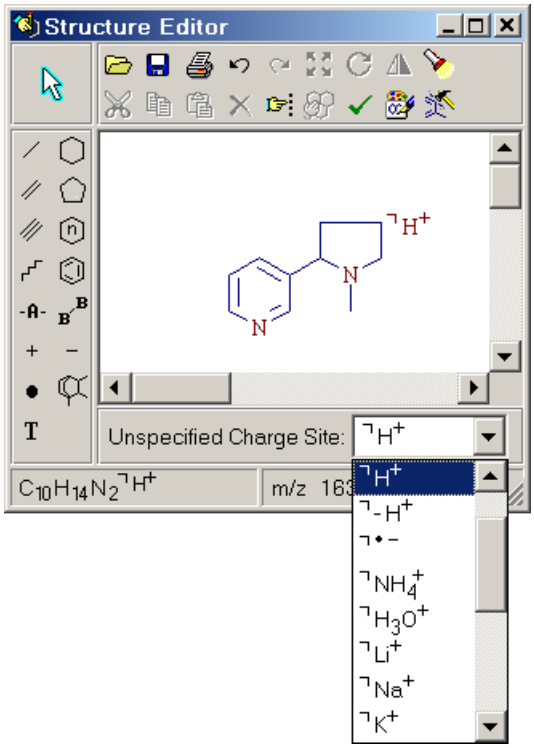


Figure 28. Structure Editor window showing Unspecified Charge Site option

Chapter 3 Spectral Trees

Mass Frontier uses tree representation for MS^n spectra. The tree structure best reflects the hierarchical spectra dependencies in tandem experiments. Mass Frontier provides a graphical user interface for the management and processing of spectral trees. Spectral trees can be reconstructed from data files, automatically extracted from data dependent chromatographic components, or manually created by the user. Trees are supported by all the spectral modules except Spectra Classifier, which uses total composite spectra generated from trees. You can store and search trees in libraries, annotate every node spectrum, or create chromatographic libraries with spectral tree components. Exchange trees between modules by using the copy and paste commands just as for single spectra. In addition, spectral trees can be exported or imported from or to Excel in text format.
Mass Frontier uses a newly developed algorithm, which is integrated in the database search procedures, for the comparison of spectral trees.

Note Most management and processing actions distinguish between a single spectrum and a tree. Therefore, when dealing with data that contains trees, determine whether a particular action needs to be applied to the entire tree or the selected single spectrum. The displayed information is also associated either with the tree or the spectrum, or both depending on the information type.

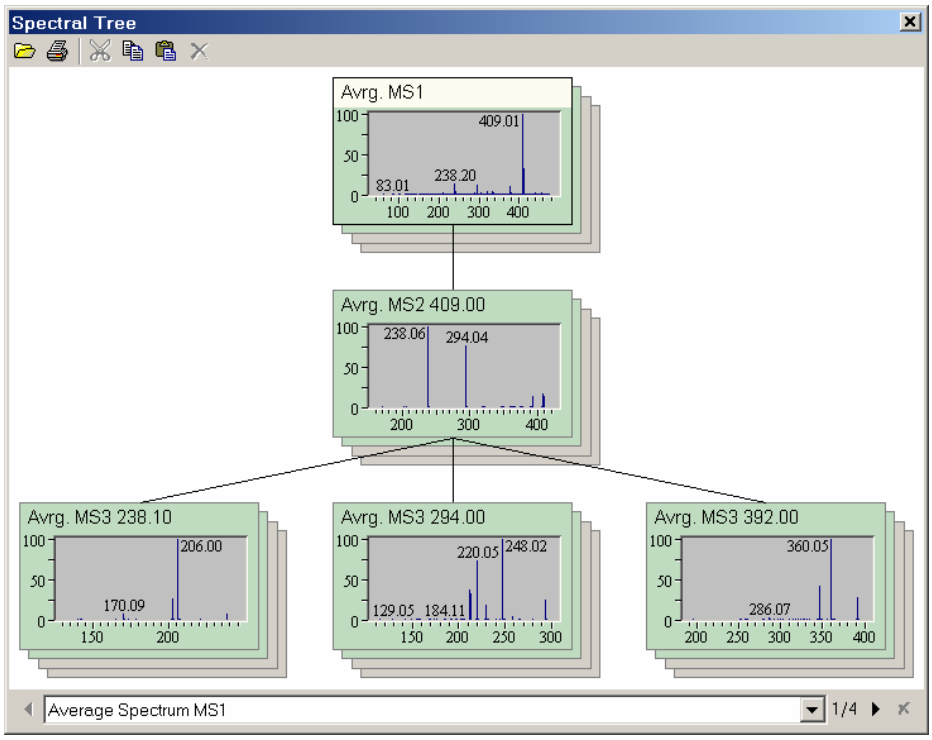


Figure 29. Spectral Tree window showing MS^3 spectra representation

This chapter contains the following sections:

- Tree Arrangement
- Tree Node Items
- Tree Layout
- Tree Generation
- Manual Creation and Editing of Trees
- Copying and Pasting Trees
- Tree Chromatograms

Tree Arrangement

Mass Frontier uses tree representation for MS^n spectra. A spectral tree consists of levels, nodes, node connectors, and node items. The levels symbolize MS^n stages starting at n=1. The node connector is a graphical symbol of the precursor *m/z* value, or the precursor *m/z* range if the isolation width is included. The node is a holder of node items. The node item stands for the product or calculated spectra of identical precursor *m/z* value or *m/z* range or for a chromatogram. Node product spectra, or parallel spectra, represent spectra acquired at various collision energies and isolation widths, or using wide band activation or they can be zoom spectra, source CID spectra, or any other spectra that enhance reproducibility in compound identification. If a node contains more than two parallel spectra, the average and composite spectra are automatically calculated. In addition to spectra, each node can contain a chromatogram.

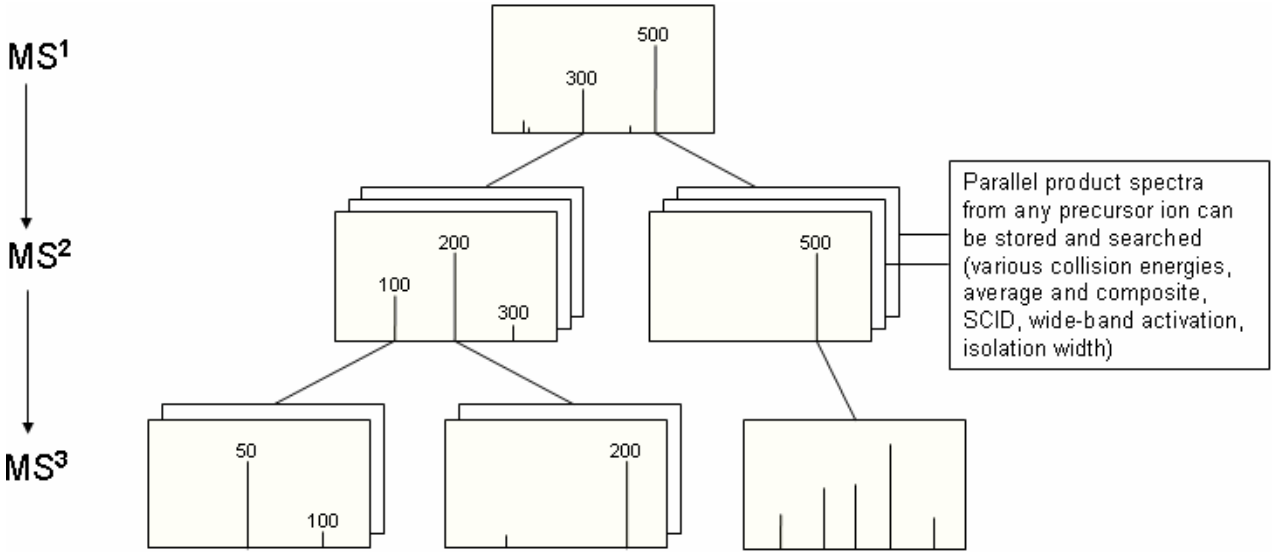


Figure 30. MS^n Data Tree Structure

Note Chemical structures are associated with nodes and not with entire trees so that you can assign fragments for product spectra. The top-level node MS^1 holds the structure of the neutral compound.

Complete trees can be stored in a library and updated at any time. Any complementary information associated with a single stage spectrum or a chromatogram can be associated with a node spectrum or node chromatogram.

Note To display or edit complementary information for a particular node spectrum, first select it.

Tree Node Items

Some mass spectrometric techniques generate spectra whose appearance is dependent upon the experimental conditions and sample preparation. To allow the management and search of diverse product spectra with an identical precursor ion for a single chemical entity, Mass Frontier uses spectral trees that can contain nodes with several node items. The node item stands for any product or calculated spectrum of an identical precursor *m/z* value or *m/z* range (node spectra) or for a chromatogram.

Node product spectra represent spectra acquired at various collision energies and isolation widths; by using wide band activation; or they can be zoom spectra, source CID spectra, or any other spectra that enhance reproducibility in compound identification. If a node contains more than two parallel spectra, the average and composite spectra are automatically calculated. In addition to spectra, each node can contain a chromatogram.

There are several reasons for using node spectra. They can significantly contribute to correct compound identification in a tree library search, they allow the study of fragmentation processes by changing the experimental conditions, and they permit the efficient organization of product spectra. The spectral node strategy strengthens the robustness of all the mathematical processing methods and, compared to simple spectra averaging, does not distort the highly nonlinear peak ratio progress.

Node spectra are easily accessible from a library and can be created or edited by using the graphical interface. Every tree item has its own annotation caption that can be edited by double-clicking it.

Select node items by clicking the mouse on the edge of the spectral or chromatographic node item, or by browsing in the box displayed below the tree. All processing actions are accessible from the pop-up menu that appears when you double-click the node.

Tree Layout

A spectral tree consists of levels, nodes, node connectors, and node items. Node items are divided into five groups that are differentiated by the displayed color: Single, Average, Composite, and Source CID spectrum, and Chromatogram. Change color settings in the **Options > Spectrum Layout** dialog window in the **MS Tree** page. Because display and editing actions affect the selected node item, the item selection is also distinguished by color.

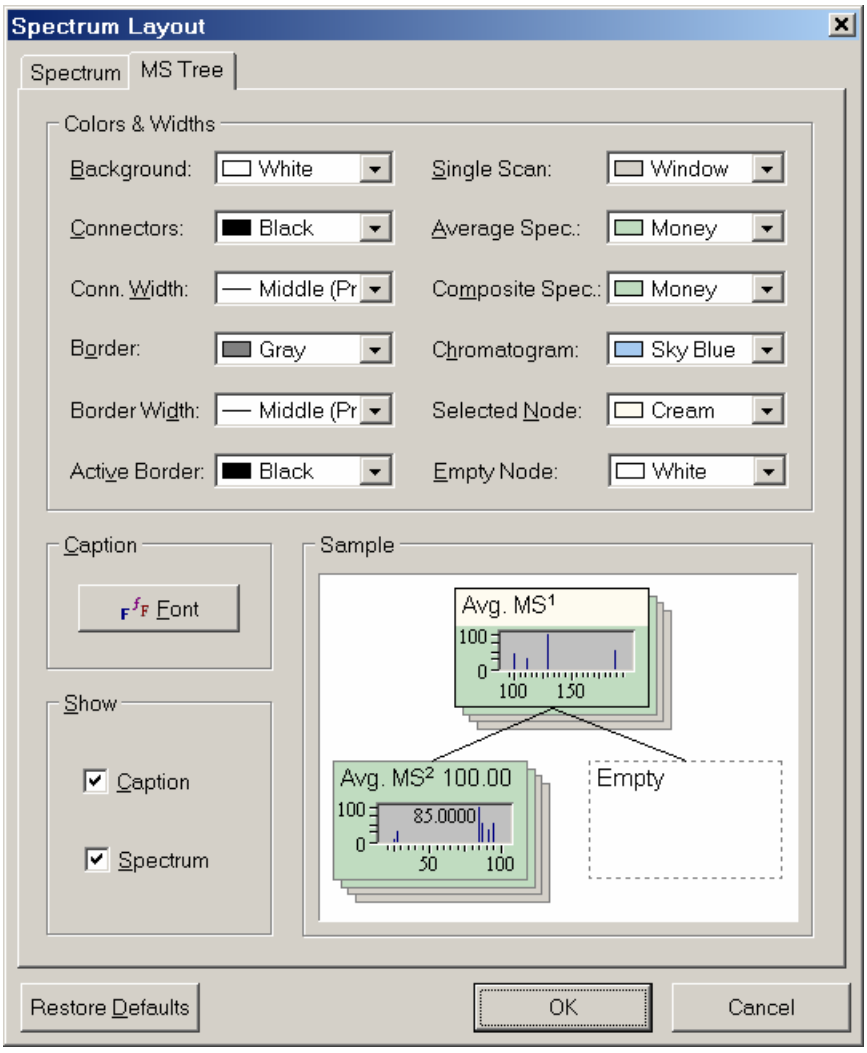


Figure 31. MS Tree page showing customizable spectral tree layouts

Tree Generation

A spectral tree can be created in the following four ways:

- Reconstruction from raw files where each file contains a single MS^n experiment (direct infusion).
- Reconstruction from a single raw file that contains various MS^n experiments in one run (direct infusion).
- Spectral tree deconvolution from a data-dependent MS^n chromatogram. See Chapter 13, "Chromatogram Processor."
- Manual creation using tree editing utilities.

Reconstruction from raw files where each file contains a single MS^n experiment (direct infusion)

The spectral tree reconstruction feature reads Xcalibur .raw files stored in a directory and automatically creates a tree according to the precursor *m/z* and isolation width values. All files (spectra) in a directory need to be from an identical chemical entity (compound or chromatographic component) with one directory per tree per compound. This feature works on Xcalibur .raw files acquired using direct injection and a single MS^n stage. Each file in a directory can be acquired separately at a specific collision energy; a wide band activation spectrum, zoom or source CID spectrum; or the sample preparation (pH, concentration, buffer, and so on) can be diverse. Example files are installed automatically into the \Mass Frontier 50\Chromatograms\Cloramphenicol directory.

To import trees using the tree reconstruction feature

1. Click the **Import** button in the Tree pane in Database Manager
2. Choose **Import Tree.**
3. Select the directories to import.
4. Click **Add** and then click **OK**.

Use this procedure to create an MS^n library.

Reconstruction from a single raw file that contains various MS^n experiments in one run (direct infusion)

You can create a spectral tree of a reference compound by reading various MS^n spectra acquired in one run using direct infusion and stored in a single raw file. To do this, you must use the Component Detection & Spectra Deconvolution feature in the Chromatogram Processor module.

To create a spectral tree from a single run using the tree reconstruction feature

1. Click the **Chromatogram Processor** button in the main Mass Frontier menu bar.

2. Select the direct infusion file to process and click the **Open** button. The file opens in the Chromatogram Processor window.

3. Click the **Component Detection & Spectra Deconvolution** button and select the Direct Infusion button from the pop-up menu.

4. If required, change the preset parameters and click **Calculate**.

5. Click the component triangle on the left side of the TIC pane and the reconstructed spectral tree appears below.

Use this procedure to create an MS^n library.

Spectral tree deconvolution from a Data Dependent MS^n chromatogram

Spectral trees can also be generated from chromatographic components. Chromatographic data must be processed using a different feature. To create a tree from a chromatogram, use the Component Detection & Spectra Deconvolution feature in the Chromatogram Processor module. For more information, see Chapter 13, "Chromatogram Processor."

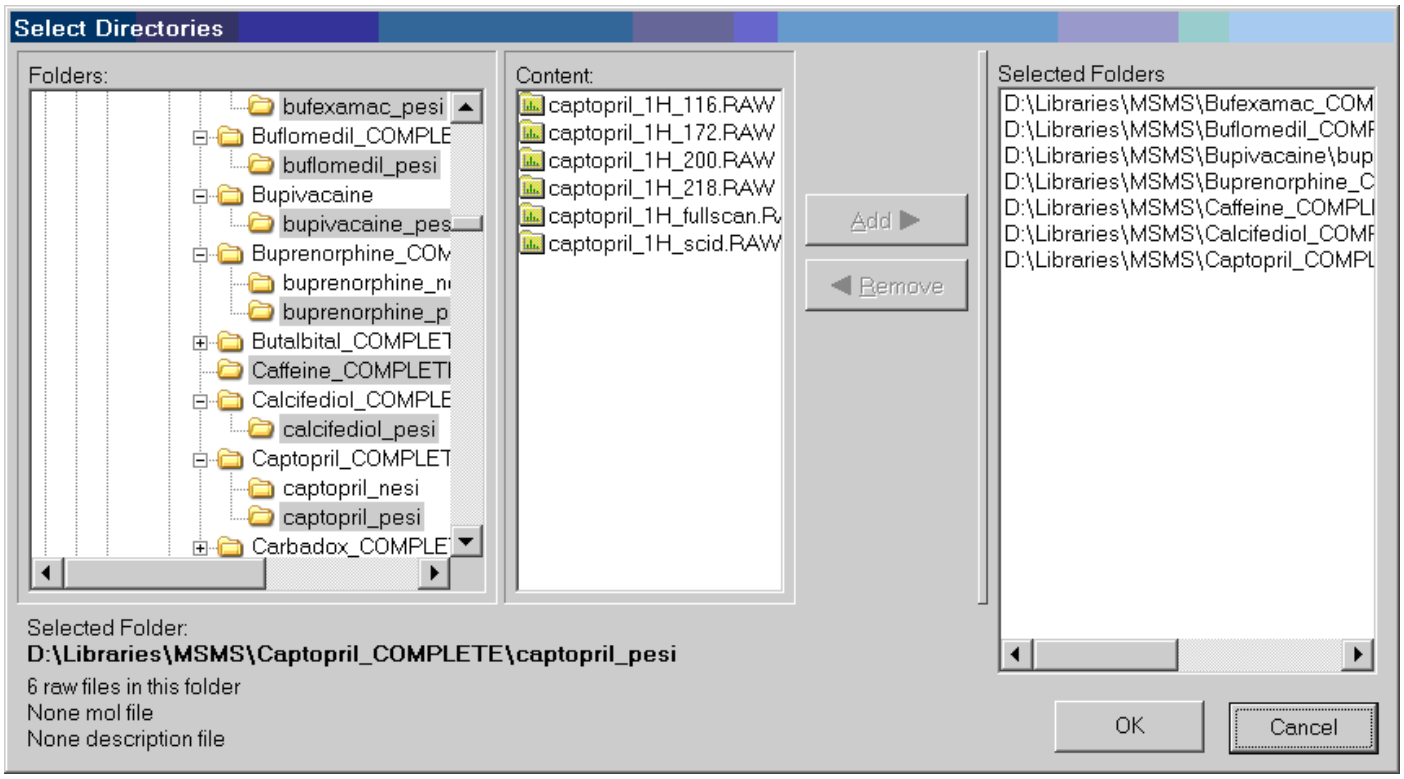


Figure 32. Select Directories window for tree reconstruction from raw files

Manual creation using tree editing utilities

The tree reconstruction automatically assigns node spectra to specific nodes utilities in addition to creating levels, nodes and node connections.

Manual Creation and Editing of Trees

Spectral trees can be manually created and edited in Database Manager in the Spectral Tree pane. All editing actions are accessible from the pane buttons and from the pop-up menu that appears when you right-click.

Note All actions described below are applied only to the selected node or spectrum.

You might want to construct a tree from scratch. You can start creating an empty tree framework by adding new nodes and then successively pasting and importing spectra into the nodes. To add a new node, right-click the parent node and choose **Add > Add Node**. An empty node can be filled with a spectrum from the Clipboard or imported from a file. Right-click the node and choose **Paste > Paste Parallel Spectrum**. This item is displayed if the Clipboard contains a spectrum. To add another empty node product spectrum (parallel spectrum), right-click the node and choose **Add > Add Parallel Spectrum**. A node can contain only one empty parallel spectrum. If you add more than one parallel spectrum to the node, the software automatically generates the average and composite parallel spectra, which are differentiated by color. If you import a .raw file into the node, the program calculates the average spectrum from all scans. Do not import a chromatogram .raw file into a spectrum node. To create a node with a chromatogram, you must paste a chromatogram from the **Chromatogram Processor** window.

A tree would be incomplete without precursor *m/z* values and the isolation width of the product nodes. There are two ways to define the precursor *m/z* value. You can either manually type the value in the box in the Structure pane or paste or draw the ionic product fragment in the Structure pane and click the **Restore Precursor** *m/z* button in the bottom right corner of the Structure pane. The default value of the isolation width is determined automatically, but can be manually changed in the box in brackets.

Edit the node caption of every parallel spectrum or chromatogram by double-clicking the caption and typing the new formatted text into the annotation dialog window.

Copying and Pasting Trees

Use Mass Frontier for the exchange of trees between windows, records, chromatographic components or programs (for example, Excel) by using the copy and paste commands. When using the paste command, you must distinguish between a single spectrum and a tree. Be sure to use the correct paste option when exchanging trees.

Trees can be pasted into Excel in numerical format and then edited and copied back into Mass Frontier. The tree numerical format supports every types of spectral trees and parallel spectra, except chromatograms. You can change the precursor *m/z* value, modify spectra, and edit the node caption by using Excel.

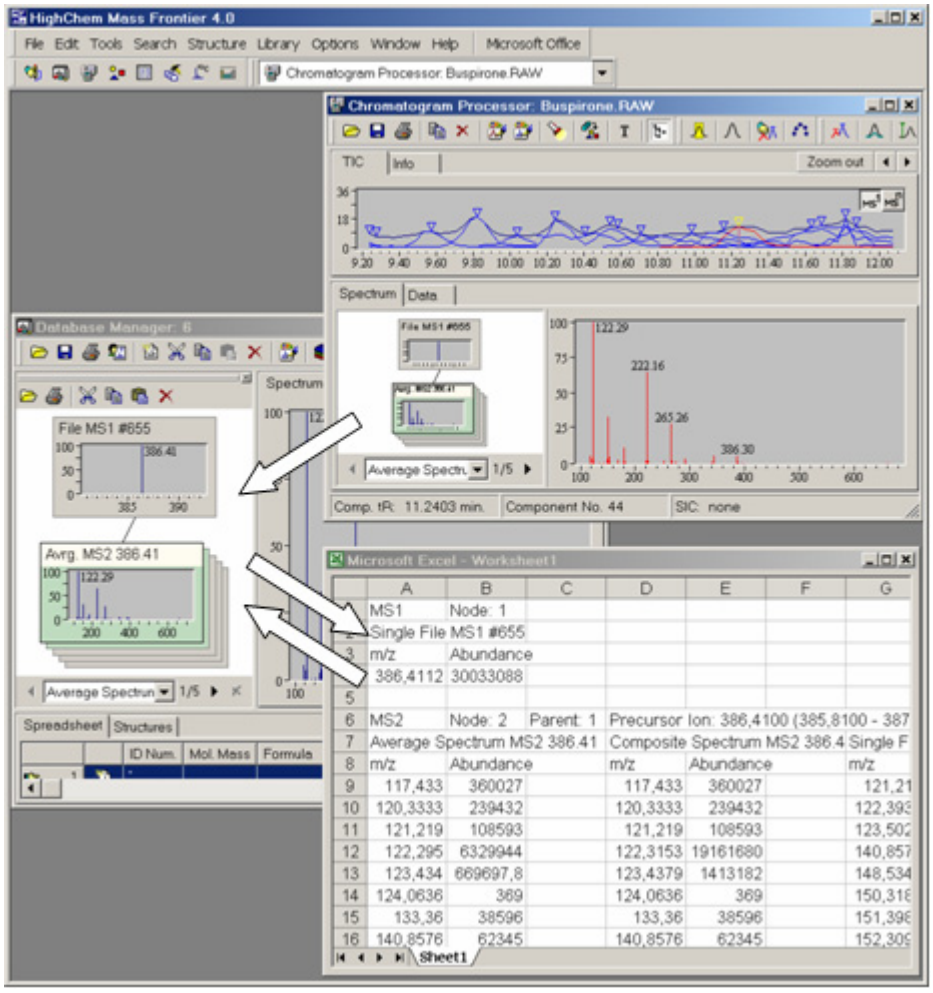


Figure 33. Exchanging spectral trees between modules

Tree Chromatograms

Mass Frontier supports spectral trees that contain nodes with data-reduced chromatograms. Chromatograms can be assigned to every node. You can store product ion chromatograms for a particular precursor *m/z* value. In contrast to spectra, only one chromatogram per node is allowed. Tree chromatograms with components and selected scans are fully searchable. A node chromatogram can be reviewed or reloaded from the original data file in Chromatogram Database Viewer, which is launched from the button of the same name in Database Manager. Review node chromatogram components or selected scans in the Chromatogram Database Viewer and the Components Editor window.

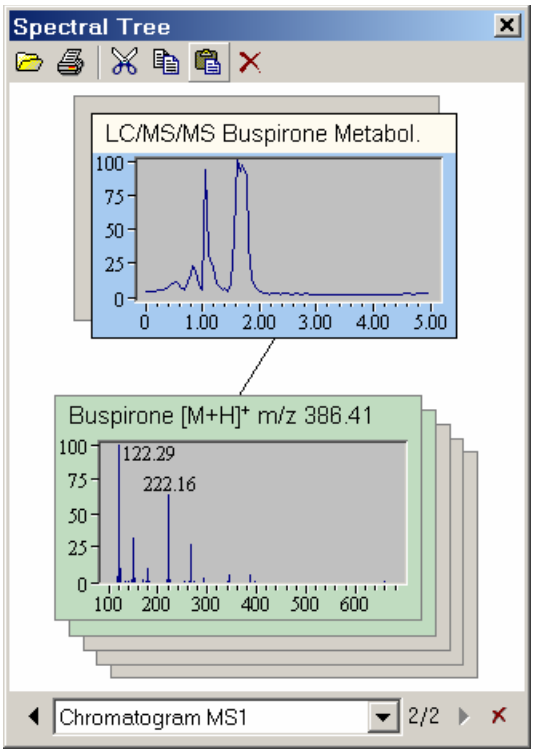


Figure 34. Spectral Tree window showing LC/MS/MS chromatogram assigned to the top level (full scan) node

Chapter 4 Database Manager

Use the Database Manager to manage spectral and structural information in the SQL desktop engine database. This module provides library maintenance utilities that enable you to create and organize spectral and chromatographic libraries with chemical structures. In addition, because the program supports ion structures and tree spectra representation, you can also create true MS^n libraries. Advanced library query and search features provide access to the information needed for compound identification and to help interpret unknown spectra. A flexible set of search restrictions is available to target your search results, useful when using large libraries.

Structural and spectral data are organized in spreadsheet-like windows, together with a variety of supplementary information. The number of Database Manager windows that can be opened simultaneously is limited only by system resources. This means unconstrained flexibility in handling spectral and structural data. It is easy to move spectra-structure pairs from one Database Manager window to another, enabling you to organize your libraries, search results and any other data.

The Database Manager provides a customizable tool for creating reports, which can be either printed directly or copied and pasted to a word processor for more advanced reports.

For each record in the Database Manager you can view the mass spectrum; a list of peaks; the compound identification information and the neutral losses spectrum, if the molecular mass is available; or compare two spectra. View this information by selecting the appropriate tab.

This chapter contains the following sections:

- Search Utilities
- Mass Spectral Data Exchange Between Modules
- Fragment Assignment to Spectral Peaks

Open a Database Manager Window

To open a Database Manager window

Do one of the following:

- Click the **Database Manager**  button on the tool bar in the main window.
- Choose **Tools > Database Manager**. An empty Database Manager window opens.
- Start any search from the Search menu and, if a spectrum or structure is found, a Database Manager window opens containing search results.

If an empty Database Manager window has already been opened, you cannot open another. Each search dialog window has a box at the bottom, with the caption *Merge Results into Active Database Manager*. As the caption indicates, when you activate this box, the search results are merged at the end of the Spreadsheet in the active Database Manager window. If unchecked, a new Database Manager window opens and the search result records are added to it.

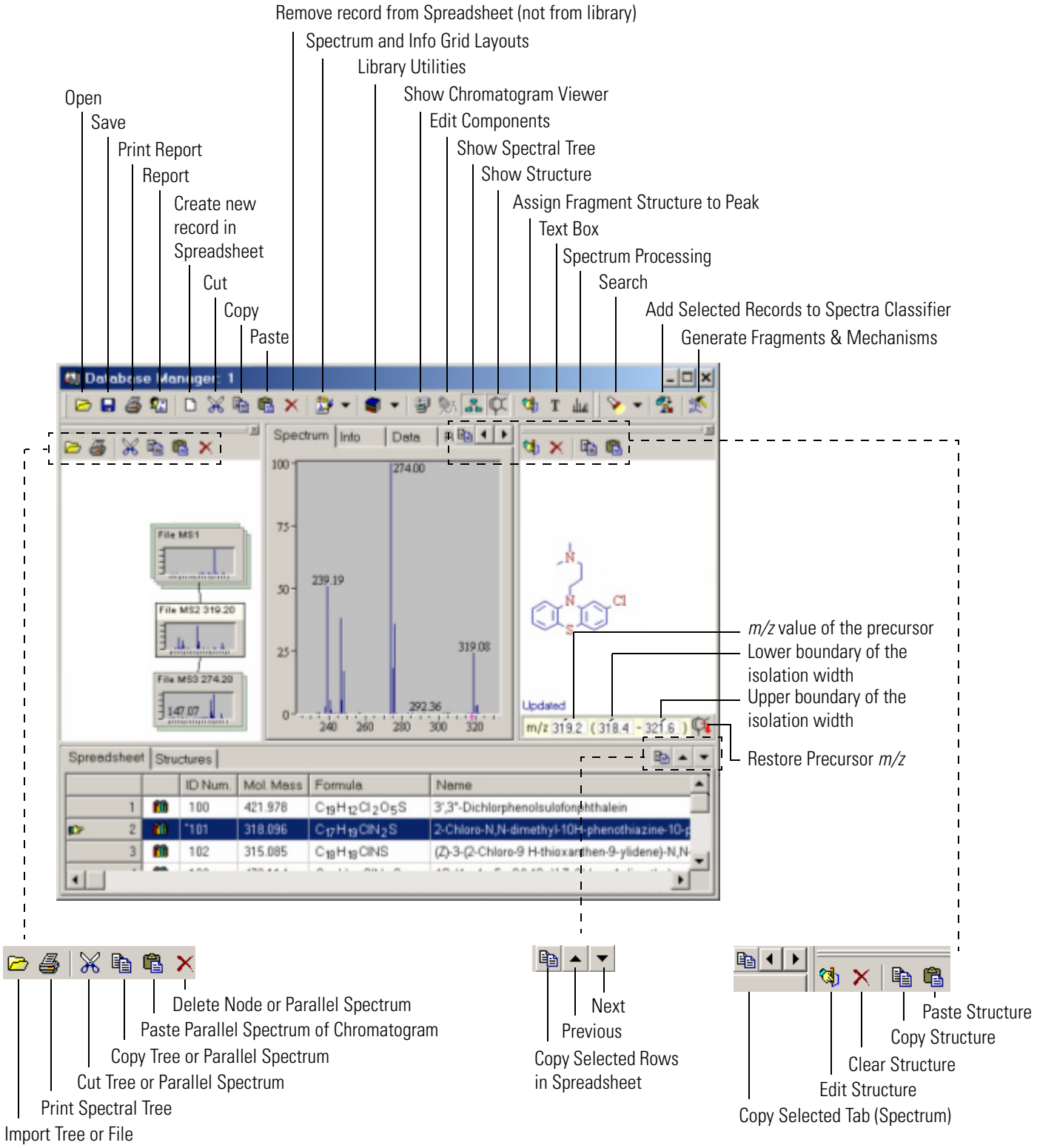


Figure 35. Database Manager window

Records in the Database Manager

A single record in the Database Manager contains one spectrum with a structure (if available) or one spectral tree with associated structures (if available). and complementary information associated with the record. Each record is visually represented as a single line in the spreadsheet. The hand icon ☞ always points to the active record. The active record is also highlighted on the Spreadsheet. The structure, spectrum, spectral tree and any other available information are displayed for the active record only. The data associated with the active record are displayed in the upper half of the window.

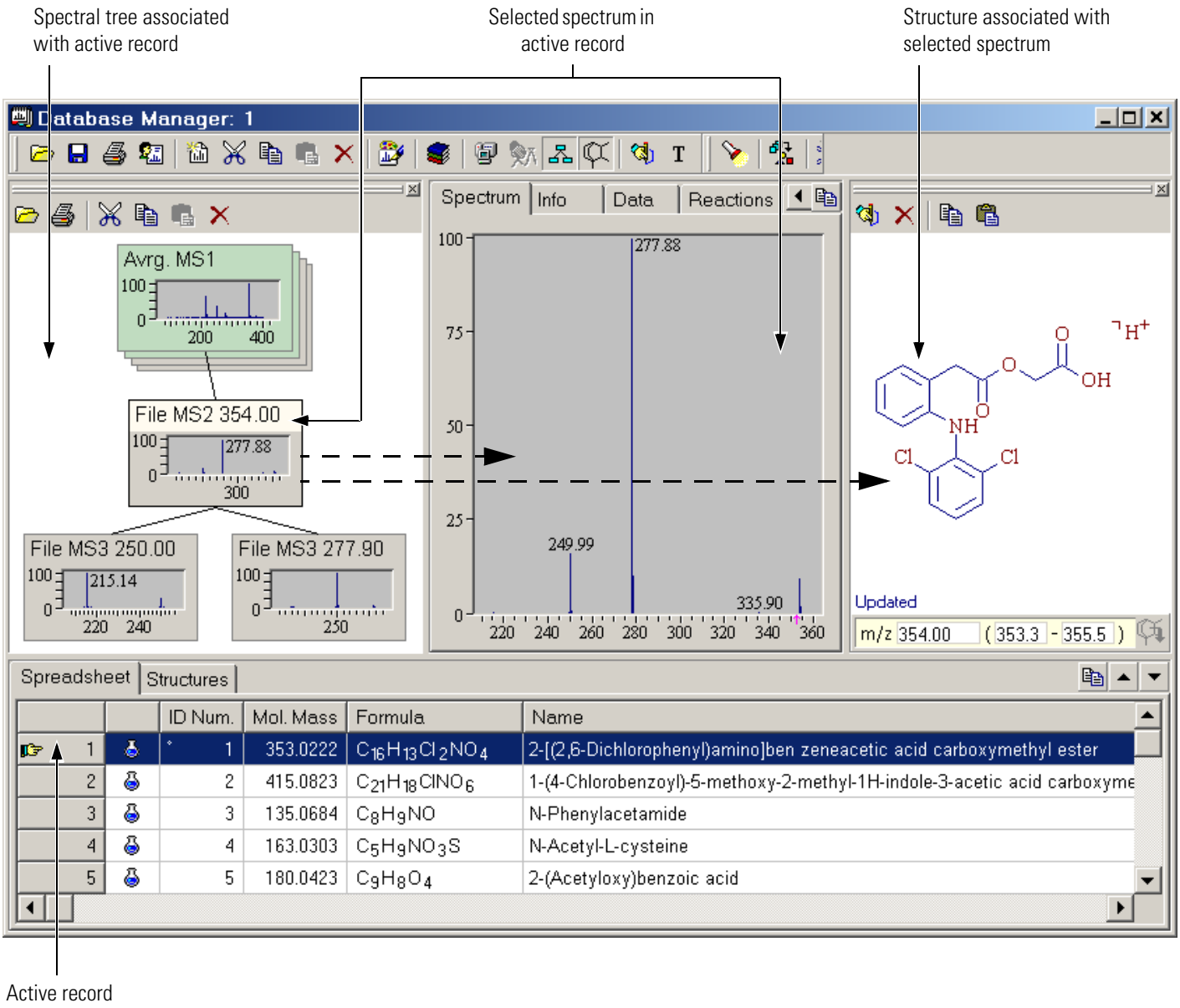


Figure 36. Database Manager window showing the selected tree items in active record

Additional Information Associated with a Record

In addition to the mass spectrum or spectral tree and the chemical structure, each record includes number of compound identifications, property and origin information. This additional information, which is associated with the record, is found in the Info tab in the Database Manager window. The Info tab contains editable fields organized by group and subgroups. Fields with information that is automatically calculated from a chemical structure (formula, molecular mass, and so on) or acquisition settings adopted from a file cannot be edited. Text in uneditable fields is unavailable. The information in the Info tab is divided into three groups:

- Main (compound identification, compound characteristic, biochemical information, sample preparation, comments, instrument, contributor)
- Chromatographic Info
- Spectral Info (mass resolution, MS^n info, scan, file)

It is important to understand the connection between data in the Info page and the spectral tree hierarchy. The Main group is associated with the entire tree, the Chromatographic Info is associated with a tree node (one chromatogram per tree node is allowed) and Spectrum Info is associated with the selected parallel spectrum. Every parallel spectrum contains its own data in the Spectrum Info group. This is because the parallel spectra can be acquired under different experimental conditions (collision energy, isolation width, and so on) that must be stored independently.

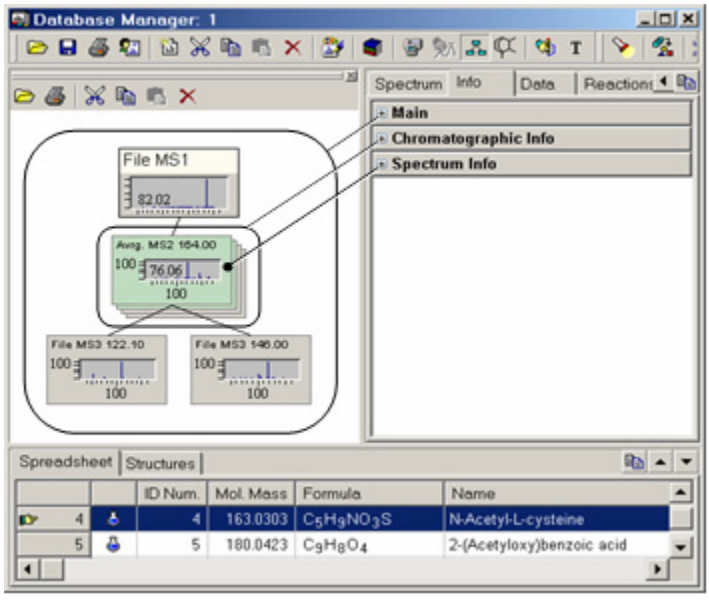


Figure 37. The relationship between tree hierarchy and Info page

If a record is saved in a library, all data in the Info tab is automatically stored in the SQL database. To permanently store any change, update, or deletion in the Info tab, you must save the corresponding record in the library.

The Main group contains the Name subgroup with three fields: IUPAC, Synonyms and Commercial Product. The library name search covers all three fields. The Spreadsheet column Name is associated with the IUPAC field in the Info tab.

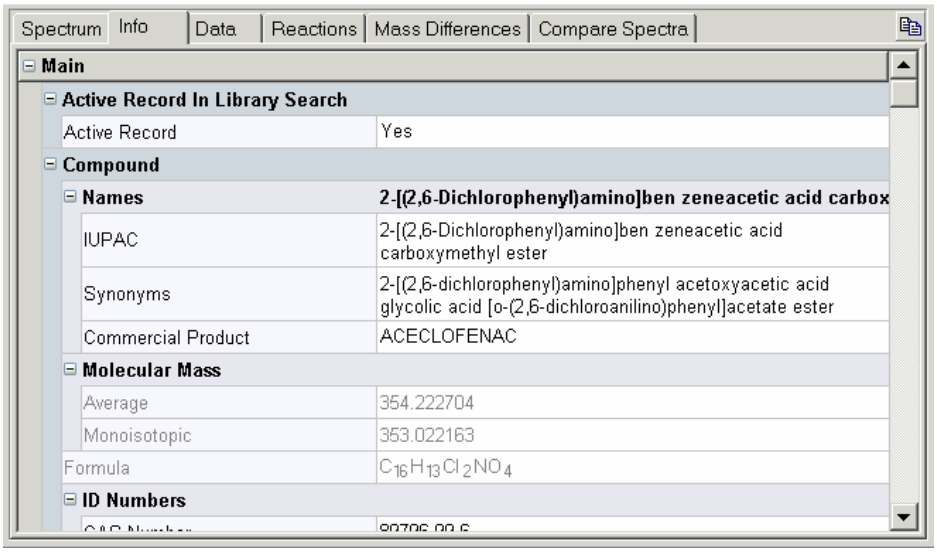


Figure 38. Database Manager window showing the Info page

The predefined list of additional information associated with a record might contain a large number of unused items, Mass Frontier allows items to be selected which appear in the Info tab for each library separately.

To exclude or include items from the predefined list of additional information

Click the **Layout** button in the Database Manager window and choose **Info Grid Layout**. The Info Grid Layout dialog window appears.

Mass Differences

Use the Mass Differences page in the Database Manager to select any peak and see the mass differences to the right and left (possible parent-daughter ions). Use the track bar at the top of the page to move the *m/z* scale. If molecular mass information is available, the zero value of the *m/z* scale automatically starts at the molecular mass; that is, a standard neutral loss spectrum is displayed. In this case, a blue selection bar in the track bar marks the shift of the *m/z* scale with respect to molecular mass. Enlarge the graphic to see the number for less prominent peaks when a large number of peaks are close to each other.

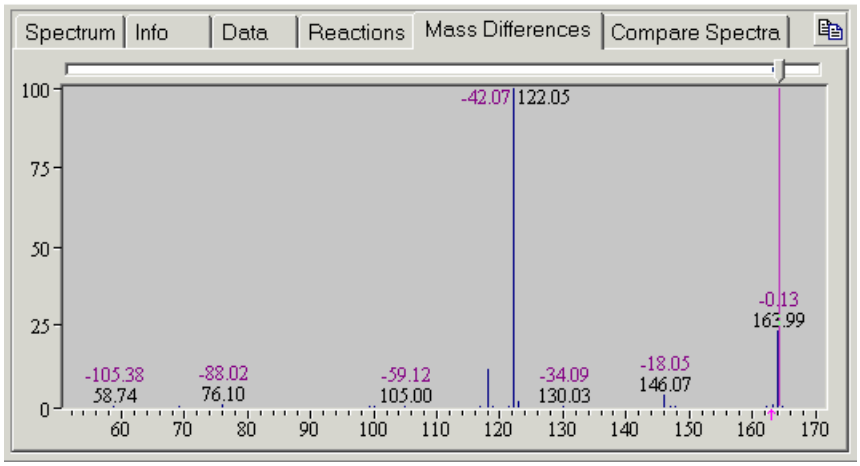


Figure 39. Database Manager window showing the Mass Differences page

Comparing Spectra

Use the Compare Spectra tab in the Database Manager to compare two spectra. The bottom spectrum is from the active record in the Spreadsheet. The middle spectrum displays the difference between the top and bottom spectrum. The top spectrum can be added from an active record by clicking the **Add** button or pasted from the Clipboard by clicking the **Paste** button. The peaks in the top and bottom spectrum have different colors; thus, the color of a difference peak in the middle spectrum is taken from the spectrum that has a more abundant peak at a particular *m/z* value.

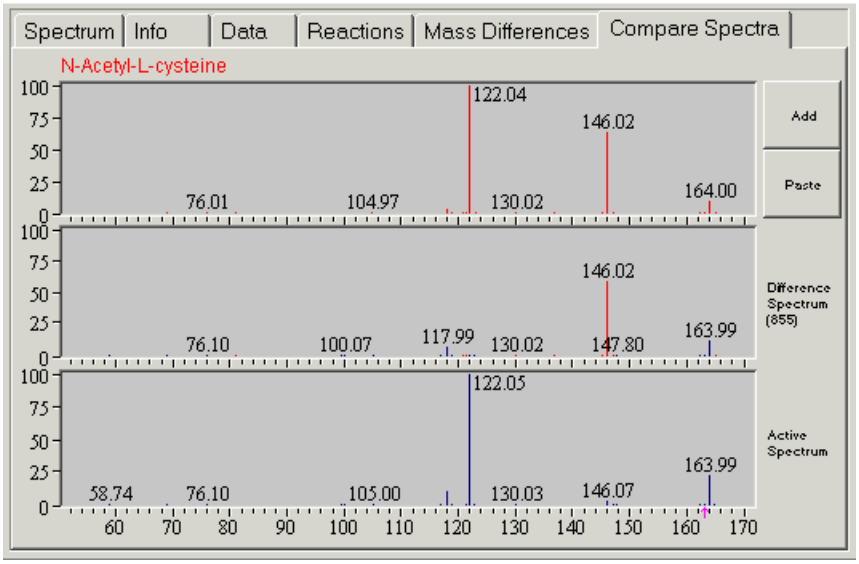


Figure 40. Database Manager window showing the Compare Spectra page

Note After a spectrum search has been completed, the query spectrum is automatically pasted to the top spectrum in the Compare Spectra tab where you can view of the peak differences of spectra in the match list and query spectrum.

Cutting, Copying, and Pasting Records

Use the cut, copy and paste functions to copy or move records from one Database Manager window to another. Records can also be copied or moved to different locations within a spreadsheet in the same Database Manager window. This permits clear organization and maintenance of your experimental data or search results.

Mass spectra can be exchanged between the Database Manager and Excel by using the import and export features.

The cut and copy functions apply only to selected records. With the spreadsheet, you can select multiple contiguous records to copy or move many records at once.

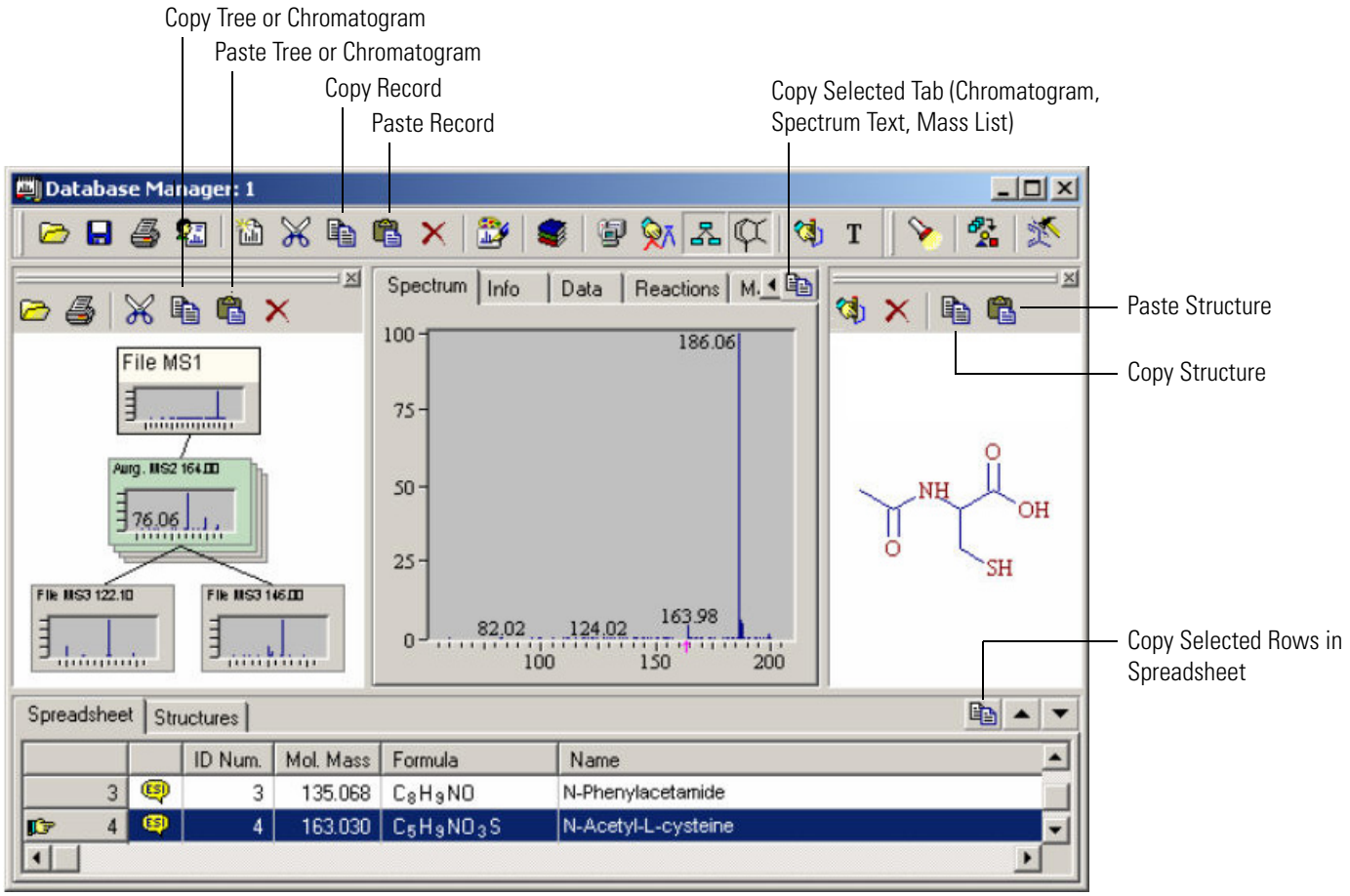


Figure 41. Database Manager window showing the diverse Copy and Paste buttons

To cut or copy records

1. Select one or more records in the Spreadsheet.

2. If the Database Manager is an active window, click the **Cut** or **Copy** button in the Database Manager or choose **Edit > Cut** or **Edit > Copy** in the main menu.

If you cut or copy a record the selected mass spectrum, in graphic format, is simultaneously copied to the Windows Clipboard, in addition to the record. If you have selected more then one record, only the graphic of the spectrum in the first record is copied to the Clipboard. This graphic spectrum representation can then be pasted to any Windows application. Graphics are copied to the Clipboard in 32-bit format (enhanced Windows metafile), which is sometimes not supported by older 16-bit applications.

To paste records

1. Select one record in the Spreadsheet where you want to paste records.

2. Click **Paste** in any Database Manager window or choose **Edit > Paste** in the main menu.

A single spectrum copied in the Chromatographic Processor can be pasted to the Database Manager to extract spectral scans and move them to the Database Manager for further processing.

Structures in Database Manager

Chemical structures are used at every stage in the Database Manager module and can be added to corresponding records. You can create libraries with structures including isotopes, ions, radicals and optically active compounds. Structures in the Database Manager can be used in connection with the Fragments & Mechanisms module to check the consistency of a mass spectrum and chemical structure. Use the Database Manager to perform structure elucidation through modification of input structure and regeneration of product fragments.

When working with spectral trees structures are associated with nodes and not with an entire tree, these structures can be added to the Database Manager record in one of three ways:

- Opening the Structure Editor directly from structure pane and drawing a new structure.
- Pasting the structures.
- Loading the structures from an external file.

To add structure to a record that contains a tree, select the appropriate node by clicking any spectrum in that node. You can assign fragments for product spectra for elucidation and database maintenance of CID spectra. The top-level node MS^1 holds structural information about neutral compound. Structure displayed in the Database Manager is associated with selected node. Once again, structure you see or edit in the structure pane in the Database Manager is associated with selected node.

You can add structure not only to a spectrum or tree node but also to any peak by clicking the **Assign Structure to Any Peak** button. A New Structure Editor window opens and you can draw a fragment. After you finish drawing, the program tries to automatically connect your structure with a spectral peak according the fragment's *m/z* value. If you need to connect a drawn fragment to a different peak, drag the connecting circle to move it to any other peak. The drawn fragment can be resized. Note that structures assigned to spectral peaks are not searchable.

You can copy a structure you want to paste into the Database Manager from anywhere in the program. To paste a structure into the Database Manager, use the **Paste Structure** button in the top-right corner. The **Copy** and **Paste** buttons on the button bar are intended for records, not single structures.

To paste a structure to a record

1. Activate the record you want to paste the structure into.
2. If processing a tree, select node you want to paste the structure into.
3. Paste the structure by clicking the **Paste Structure** button.

If a structure has been added to a record or an existing structure has been replaced, the word Updated appears at the bottom of the structure pane. If anything is changed in the record, including the structure, a small circle is displayed in the ID Num. column in the spreadsheet. After a structure is added or changed, the molecular formula and molecular mass are automatically calculated and updated.

If a structure has been added to a tree node or an existing structure has been replaced, the **Restore Precursor m/z** button at the bottom of the structure pane becomes active if the structure molecular mass differs from the existing precursor *m/z* value. Use this button to set the precursor *m/z* value of the active product spectrum according to the molecular mass of the drawn structure.

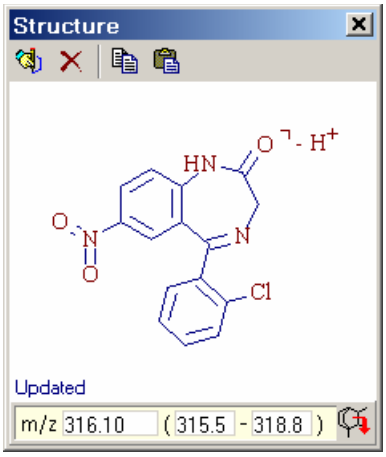


Figure 42. Detachable Structure window showing the precursor ion structure, the precursor *m/z* value (316.1) and the isolation width *m/z* range (314.5-318.8)

Import structures to the Database Manager by loading them from an external .mol file or .sdf file. In contrast to the pasting of a structure, when you load from an external file, you can add structures to more than one record at a time. You can, for example, assign a large number of structures to library spectra when importing an external library to Mass Frontier.

To import structures from a file

1. Select the records to which you want to add structures.

2. Load the structures by choosing **File > Open > Structure.**

When adding more than one structure to your records, the structures are added in the same order as they are present in the file, from the first to the last selected record. If the number of structures in an external file is greater than the number of selected records, structures are added only to the selected records. If the number of selected records is greater than the number of structures in the file, all the structures are added, and some records remain without structures.

Chemical structures can be added not only to a spectrum or tree node but also to any peak. To add a chemical structure to a peak, click the **Assign Structure to Any Peak** button. A New Structure Editor opens and you can draw a fragment. After you finish drawing, the program automatically attempts to connect your structure with a spectral peak according to the fragment's *m/z* value. If, for whatever reason, you need to connect the drawn fragment to a different peak, drag the connecting circle with the mouse to the required peak. The drawn fragment can be resized.

A structure can also be assigned to a peak by using the Data tab, where all peaks in the spectrum are listed. Click on the line where the peak is listed, then click the ... button in the Fragment column and the Structure Editor opens.

Note Structures assigned to spectral peaks are not searchable.

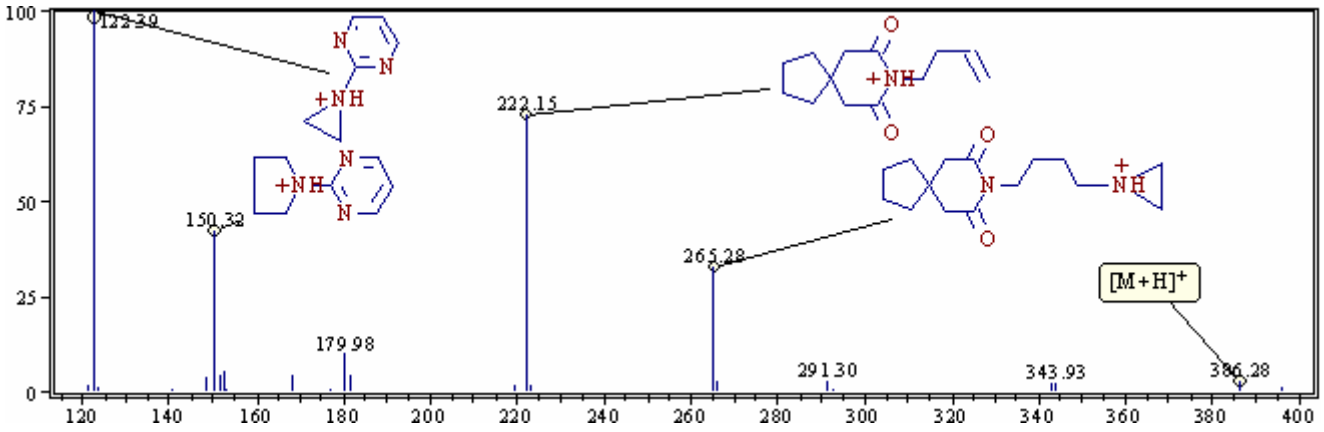


Figure 43. The spectral peaks with assigned fragments

Working with Spreadsheets

The spreadsheet is organized by rows where each record is represented by a single row. The columns contain supplementary record information. In the spreadsheet, move columns by dragging the appropriate column header to a new position. If you want to move a row, use the cut and paste functions as dragging is not supported.

When you open a Database Manager window, the spreadsheet is empty. You can add records to the spreadsheet by conducting a search, opening spectra or references, or by pasting records or stand-alone spectra. For an active record, the associated spectrum and structure (optional) appear in the upper half of the window. The hand-shaped cursor in the first column indicates which record is active. You can select more then one record, but the row with the hand icon is always the active one.

A Database Manager window can contain 999 records, but you can open as many Database Manager windows as your system allows.

Record data from the spreadsheet can be exported to Excel in text format.

Sometimes you might need to simultaneously preview all structures rather than record information in spreadsheet. To display a grid of structures, select the Structure tab next to the Spreadsheet tab. In this window, the structures are organized in small cells. If structures need to be enlarged, resize any structure cell by dragging the edges of the column or row headers.

Note In the Structure view option, you can select only a single record.

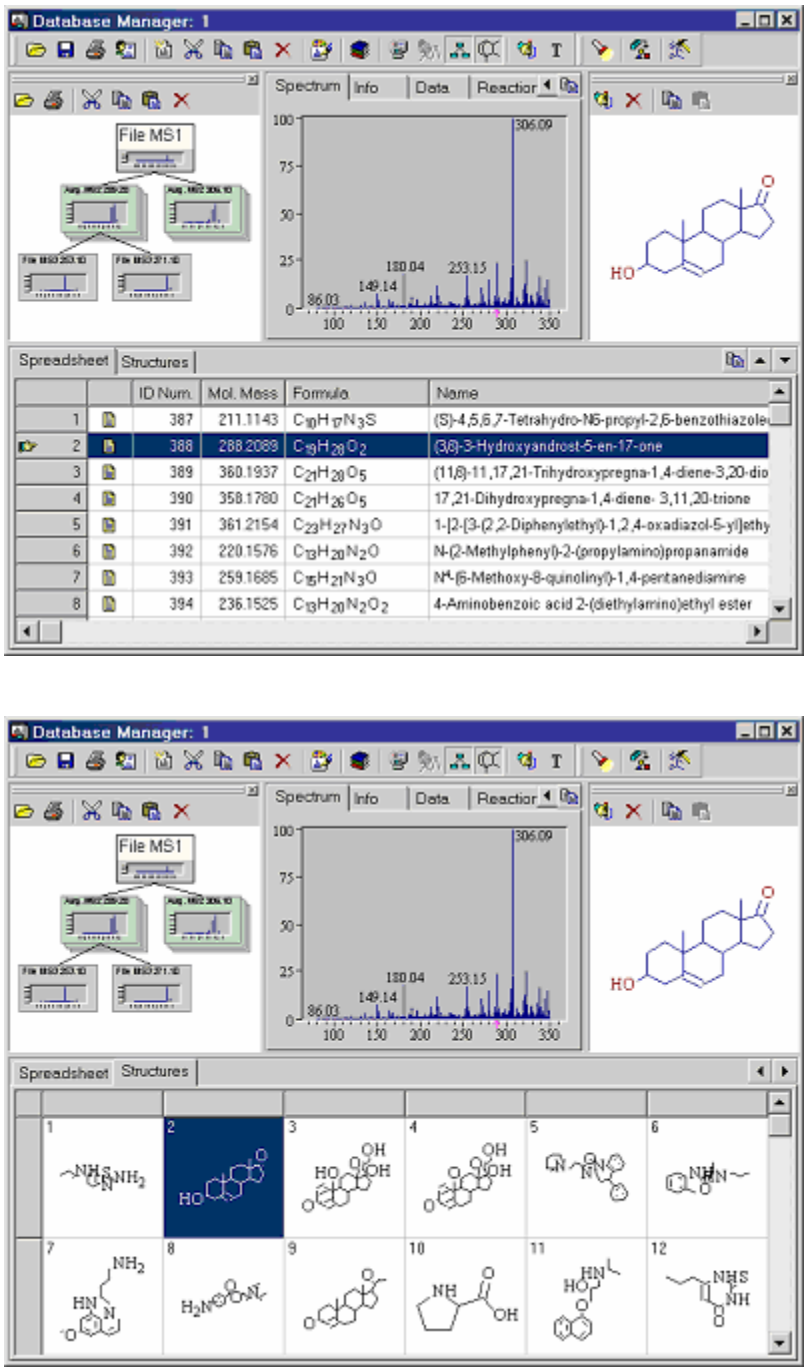


Figure 44. Database Manager windows showing Spreadsheet page and Structure page

Search Utilities

For retrieving spectra and structures from libraries, Mass Frontier has several query and search features. You can search every library simultaneously. The following search options are available:

- Spectrum Search

  Searches for the library spectra most closely matching an unknown spectrum.

- Tree Search

  Searches for the library spectral tree most closely matching an unknown spectral tree or subtree.

- Substructure Search

  Searches for an exact match for the query structure (structure search) or searches for an exact match for the structure subset (substructure search).

- Name Search

  Incremental name search.

- Molecular Mass Search

  Searches for compounds with a given molecular mass (unit resolution only).

- Formula Search

  Searches for compounds with a given molecular formula.

- ID Number Search

  Searches for library entries with a given ID number.

- CAS Number Search

  Searches for compounds with a given Chemical Abstract Service registry number

- Retention Time Search

  Searches for library entries with a range of given retention times.

Every search dialog window contains a box with the caption *Merge results into Active Database Manager Window*. If the box is clicked, the search results beyond the last record into the Spreadsheet are stored in an active Database Manager window. If you uncheck this box, a new Database Manager window opens and the search results are put into it. If this box is disabled, there is no active Database Manager window and a new window opens.

To start a search

1. Click **Search** in the main menu.
2. Choose the specific search you want to conduct.
3. When the dialog windows appears, select the active library or libraries to be searched.

If the search is successful, the results are stored in the Spreadsheet in the Database Manager window.

Spectrum Search

Mass Frontier uses a search algorithm developed by HighChem. This algorithm is based on the optimized dot-product function together with an additional term, based on ratios of peak intensities.

The query spectrum can originate from the Database Manager, Component Editor, or from the Chromatogram Processor by selecting a spectral scan in a chromatogram. In addition, the query spectrum can be pasted from the Clipboard to the search dialog window. In this case, the spectrum can be copied to only Mass Frontier.

Use the Search Spectrum option to choose between Identity or Similarity searches. Use identity searching to locate a library spectrum that closely matches an unknown. Use similarity searching to retrieve spectra library entries of similar compounds when the unknown is either not in the library or its spectrum is so badly distorted that a reliable match is not possible.

After a spectrum search has been performed, the search results (or match list) are stored in the Spreadsheet. The usual result of the spectrum search is a match list of 100 records. After a search has been completed, a match factor column, with the caption *Match* is automatically added to the Spreadsheet. The match factor is a number from 1 to 999 that specifies the measure of similarity between the query spectrum and the library reference spectrum. A Match factor of 999 means a perfect match. To draw your attention to a match greater than 930, a lightning-shaped icon is displayed in the Match column in the spreadsheet.

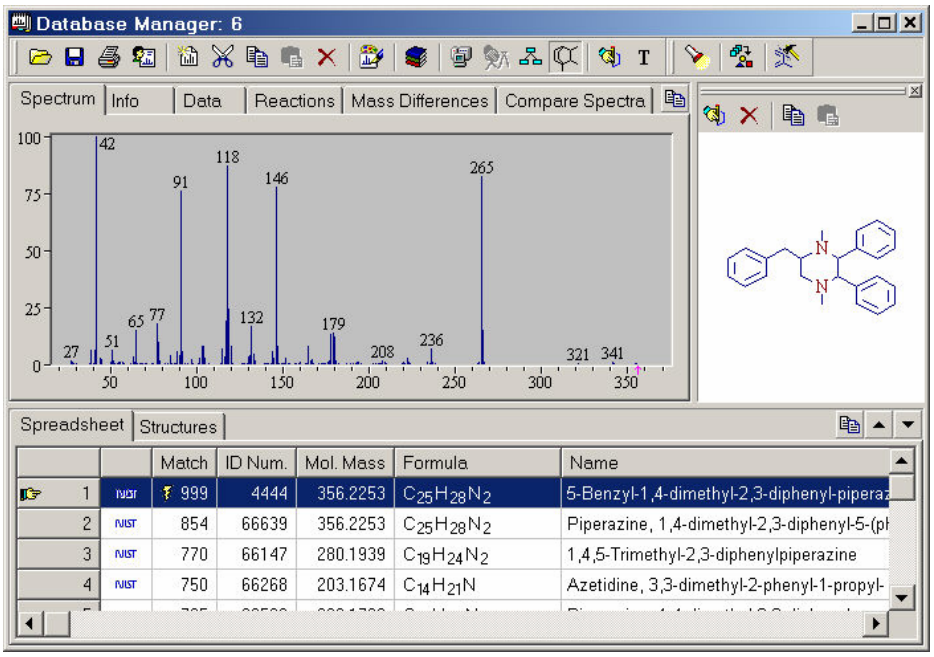


Figure 45. Spreadsheet page of the Database Manager window showing spectra search results with Match factors

If an Identity search does not provide an acceptable match (for example, the unknown has not been positively identified), use a Similarity Search. In this case, the algorithm does not use the high mass peak index, but instead uses wider abundance ranges. The match list resulting from a similarity search can be valuable in deducing structure, especially in establishing a structural proposal for an unknown spectrum.

To establish a structural proposal for an unknown, switch to the Structures tab. From the match list of similar compounds, you might recognize some common structural features which are displayed in the structures grid. You can copy the structures to the Structure Editor and put the pieces of the structural puzzle together and so create an initial structure. This structure can be pasted back into a Database Manager window to the record holding the unknown spectrum. After this, a comparison of the peaks in the spectrum with generated fragments can provide valuable information about the consistency of the proposed structure and unknown spectrum. If the *m/z* values of the fragments do not match the spectrum, modify the structure and repeat this procedure. However, if the structures in the hit list are highly diverse and dissimilar, combine this approach with other methods.

Tree Search

Mass Frontier allows the processing of MS^n spectra in hierarchically consistent trees that can be searched in spectral or chromatographic libraries using various options that can be found in the **Advance** tab in the Spectra or Tree Search dialog window. Each node in spectral trees can consist of four types of spectra (average, composite, parallel and source collision induced dissociation) and you can specify in which type spectra is searched.

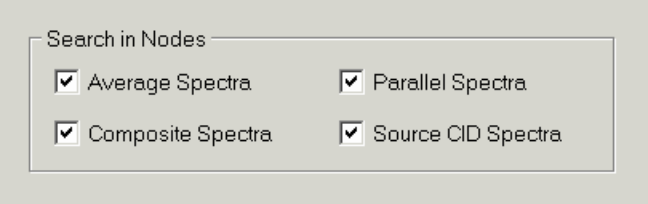


Figure 46. Spectra type options for searching in tree nodes

This option is useful when dealing with source collision induced dissociation (CID) spectra. If you have a library which consists exclusively of source CID spectra and your unknown spectrum is also a source CID, exclude other spectra types from a search. If no such library is available, you can search source CID spectra in product CID spectra, but be careful regarding the search results. Source CID spectra might contain fragmentation products from all the ions present in the source including adduct or cluster ions, while product CID spectra are preferably generated from protonated or deprotonated ions.

There are two combinations relating to the tree search:

- Search single spectrum in library trees
- Search tree in library trees

If you search a single spectrum in library trees, the spectrum is compared to every spectrum in the tree hierarchy and the match factor is individually calculated. A single spectrum can be searched only in the top level (full scan, source CID), everywhere except for the top level (first stage), or everywhere in the tree.

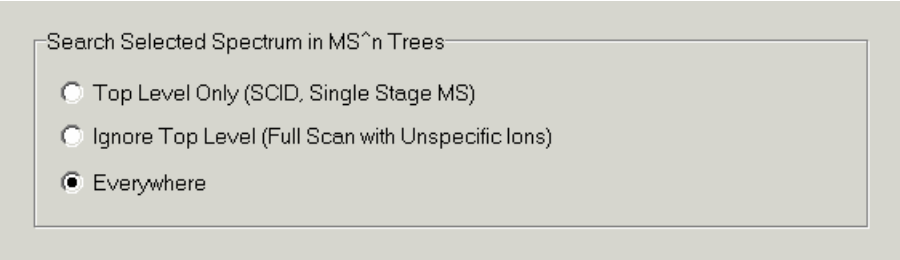


Figure 47. Search options for searching single spectrum in trees

If you search a tree (unknown) against spectral trees in a library, the spectrum is compared according to a special logic. The corresponding spectra on an identical level (MS^n stage) with a common precursor *m/z* are compared using an algorithm based on the optimized dot-product. If a spectrum only appears on one side, it is ignored and does not negatively effect the search result. If there are several corresponding spectra (such as single node with average, composite, parallel, or other spectra), the best match is accepted (optimistic approach). The total match factor is calculated from all non-zero match factors.

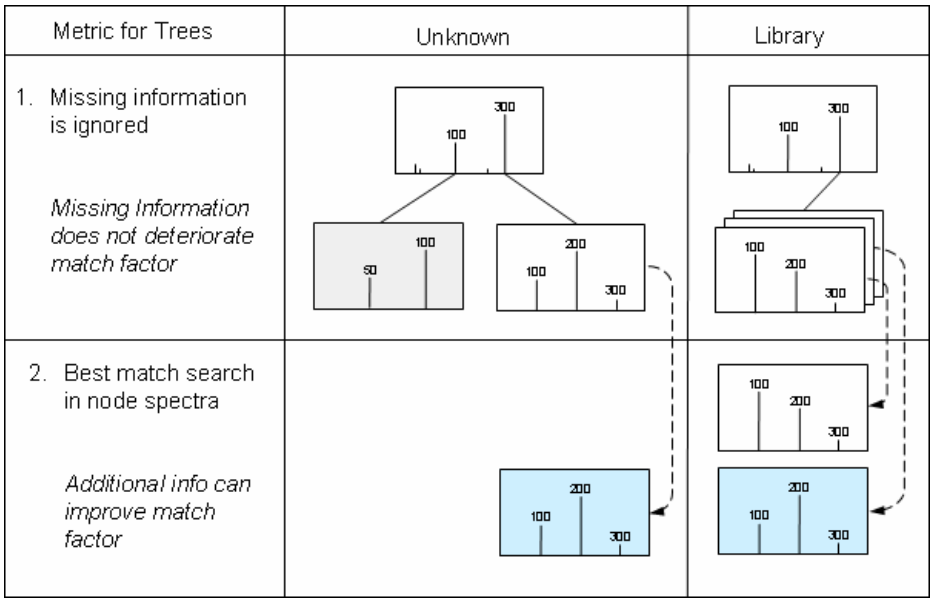


Figure 48. Metric principles of tree comparison in library search

A spectral tree can be searched using two options. The top tree level (full scan, source CID) can be included or excluded from a search and the MS^n stage of the tree spectra can be identical (identity search) or not identical (subtree search). See Figure 49.

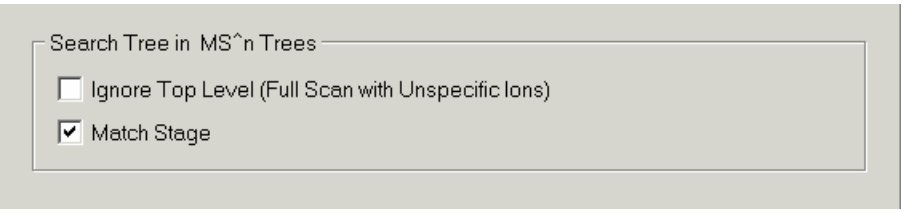


Figure 49. Search options for spectra tree search

Substructure Search

A structure or substructure query can be taken from the Structure Editor, Database Manager or a fragment copied in a Fragments & Mechanisms window.

Note When initiating a substructure search from the Structure Editor, the (sub)structure query must be selected.

Structure and substructure searches are important for retrieving library entries. The structure search is the most straightforward method for finding compounds in a library. Because the rules of systematic nomenclature do not necessarily lead to a unique name for each compound, name search can be, in many cases, ineffective. It is easier to draw or import a structure query than to type a complicated name.

While a structure search provides an exact match of query and library structure, a substructure search retrieves compounds that contain a common structural subset, called substructure. The exact substructure must be embedded in each molecule retrieved. The exact match, in structure and substructure searches, has a notable exception—that stereo bonds are ignored, because optical activity does not play a significant role in mass spectrometry. All other structural features such as bond multiplicity, atom state and skeletal arrangement must match exactly. Charge, radical and unspecified charge site can be optionally ignored and isotopes can be optionally disregarded.

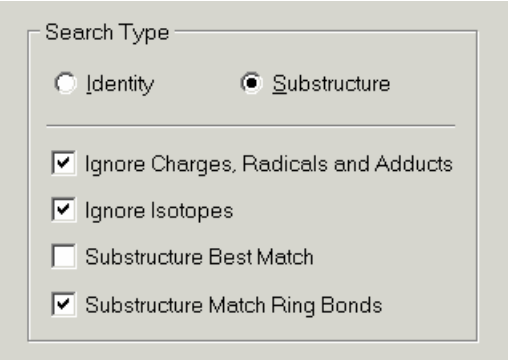


Figure 50. Substructure search options

Substructure search offers two additional search options: Substructure Best Match and Substructure Match Ring Bonds. A substructure can sometimes fit at several locations of a larger structure. The Substructure Best Match option that represents the closest match found and appears in red in the Structure tab. Note that this option can lengthen calculation times. Because fragmentation mechanisms on rings significantly differ from acyclic moieties, search substructures that exactly match the ring membership for each bond. Set this option in the Search Type box. Note that a substructure search can also lengthen calculation times.

Use a Substructure search for a variety of purposes such as to study mass spectra of structurally related molecules or to positively identify an unknown. If you need help in interpreting a mass spectrum, compounds retrieved by substructure search can provide analogies to the fragmentation processes in the spectrum under consideration.

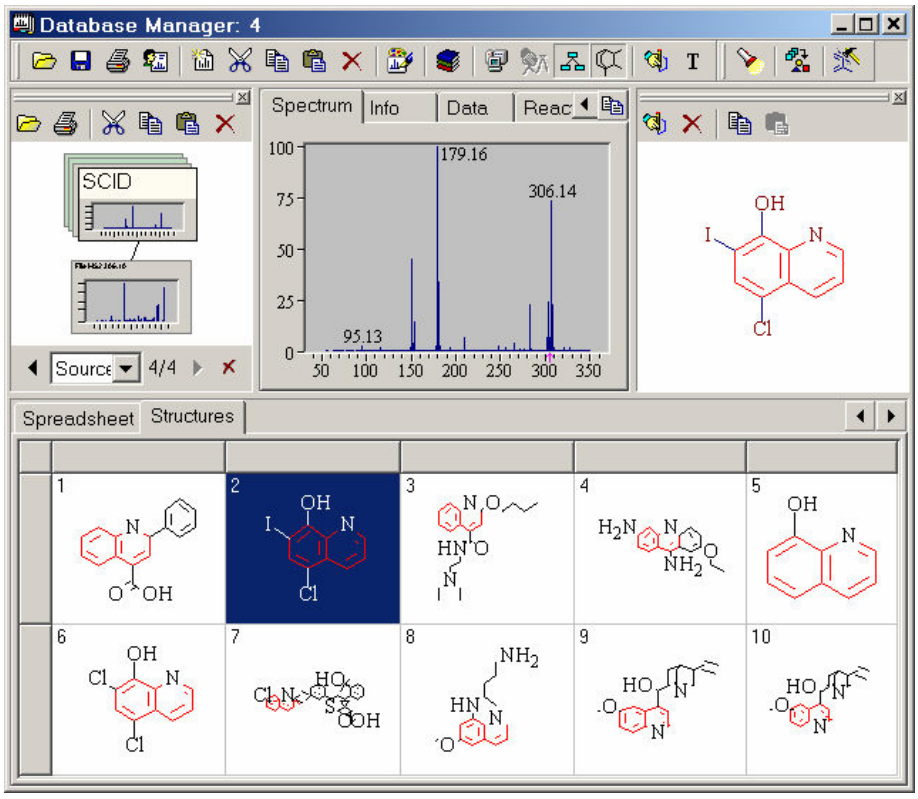
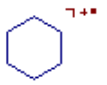
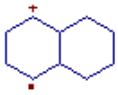
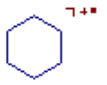
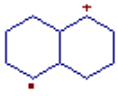
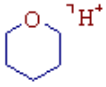
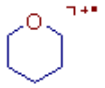
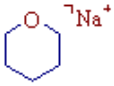
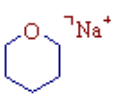


Figure 51. Structures page of the Spectra Manager window showing the results of a substructure search

Substructure Search Rules

When searching structures with an unspecified charge site or substituents, the search rules in Figure 52 apply.

Search Type	Query	Library	Search Result
S			True
S			False
I			True
I			True
I			False
S			True
S			True

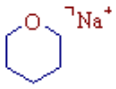
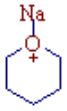
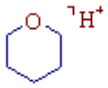
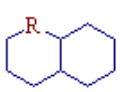
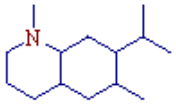
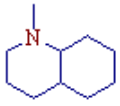
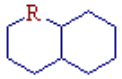
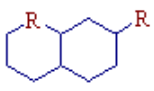
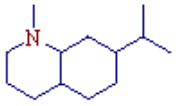
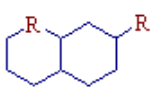
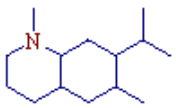
Search Type	Query	Library	Search Result
I			True
S			False
I			True
S			True
S			False
I			True
I			False

Figure 52. (Substructure Search Rules

A structure similarity search can be performed using substituents.

Note A substructure search ignores the number and position of hydrogen atoms.

Name Search

The name search option provides incremental search capabilities. As you start typing a portion of a name, a list of compounds that are spelled similarly is displayed with a slight delay. The chemical structure is displayed for the highlighted name. Use the up and down keys to browse the displayed names. The library name search covers all three name types: IUPAC, Synonyms, and Commercial Product.

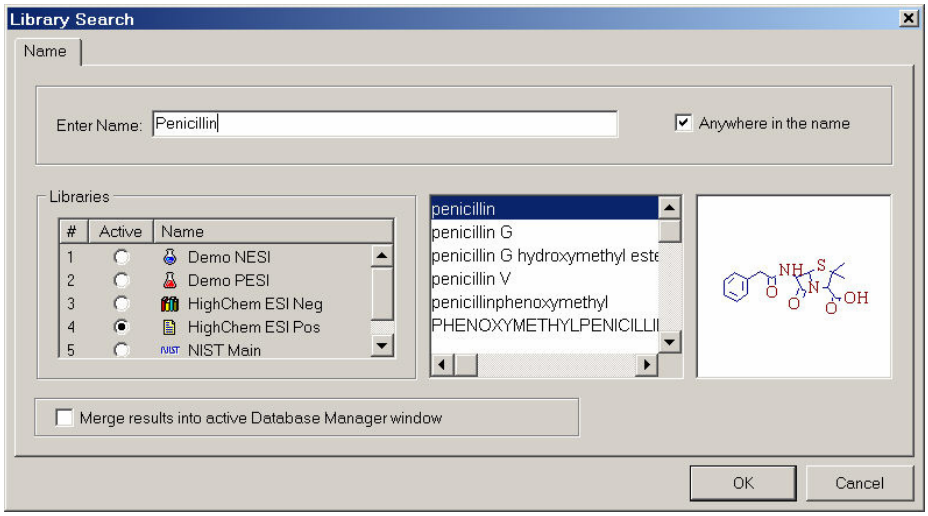


Figure 53. Dialog window showing interactive name search

Molecular Formula Search

Use the molecular formula search option to search for all compounds with a specific molecular formula. You can use lowercase letters when typing the formula unless this would lead to an ambiguous query; for example, *si* could be interpreted as *Si* or *SI*). In this case, be careful to avoid misinterpretation.

Molecular Mass Search

In this search, the term molecular mass is used to mean monoisotopic molecular mass. The monoisotopic mass of an ion is the mass of the isotopic peak which is composed of the most abundant isotopes of its elements.

ID Number Search

Each library entry is individually numbered. In this mode, you can search for a single ID number or for a range of ID numbers. The ID search dialog window contains two boxes (From and To) for you to type the ID range. If you want to retrieve a single ID number, leave the To box blank. The maximal ID number found is displayed in the Max. ID box.

Note If you delete a record from a library, the ID numbers of records with a higher ID than the deleted record are decremented. This means that if you delete one or more records, there are no ID number gaps in the library.

CAS Number Search

In contrast to earlier versions of Mass Frontier, the Chemical Abstract Service (CAS) registry number search option is now available for all libraries. Do not use dashes when typing CAS numbers.

Retention Time Search

Each spectrum of a tree or stand-alone spectrum of a record can include a retention time value (Spectrum Info/Scan Info/RT item in Info tab). Using a retention time search, you can retrieve spectra whose retention time values fit into a range of retention values you have defined.

Search Constraints

All searches, except Name Search, can be restricted by a set of constraints. The dialog windows of these searches contain a panel with buttons for activating constraints and for editing search constraints. When you click the **Edit** button, a dialog window appears to enable you to search selected libraries for matches with a specific set of criteria. Four constraint types are available: Molecular Mass range, Number of Atoms range, Allowed Elements and Good-Bad list. Searches conducted with activated constraints can be time consuming because each library entry is examined for matching criteria. Use Search Constraints when you are working with large libraries and you want to retrieve matches that are of interest for your specific problem.

Use the Good list (required substructures) to focus your search results on the particular compound classes you are most interested in. The Bad list (forbidden substructures) eliminates all structures containing unwanted functional groups from a match list. For example, the Good-Bad List can be used in a search of acids with a specific molecular formula, or you can search for spectra similar to an unknown, with the requirement that ketones not appear in the match list. Figure 54 depicts a substructure search using the Search Constraints option. The figure shows a Structure Constraints dialog window set for C, N, and O as allowed elements. The Good-Bad list is set

for esters as required (good) and for tertiary amines as a forbidden (bad) functional group. The substructure query is the triphenyl group and the search results, with the described restrictions, are also shown.

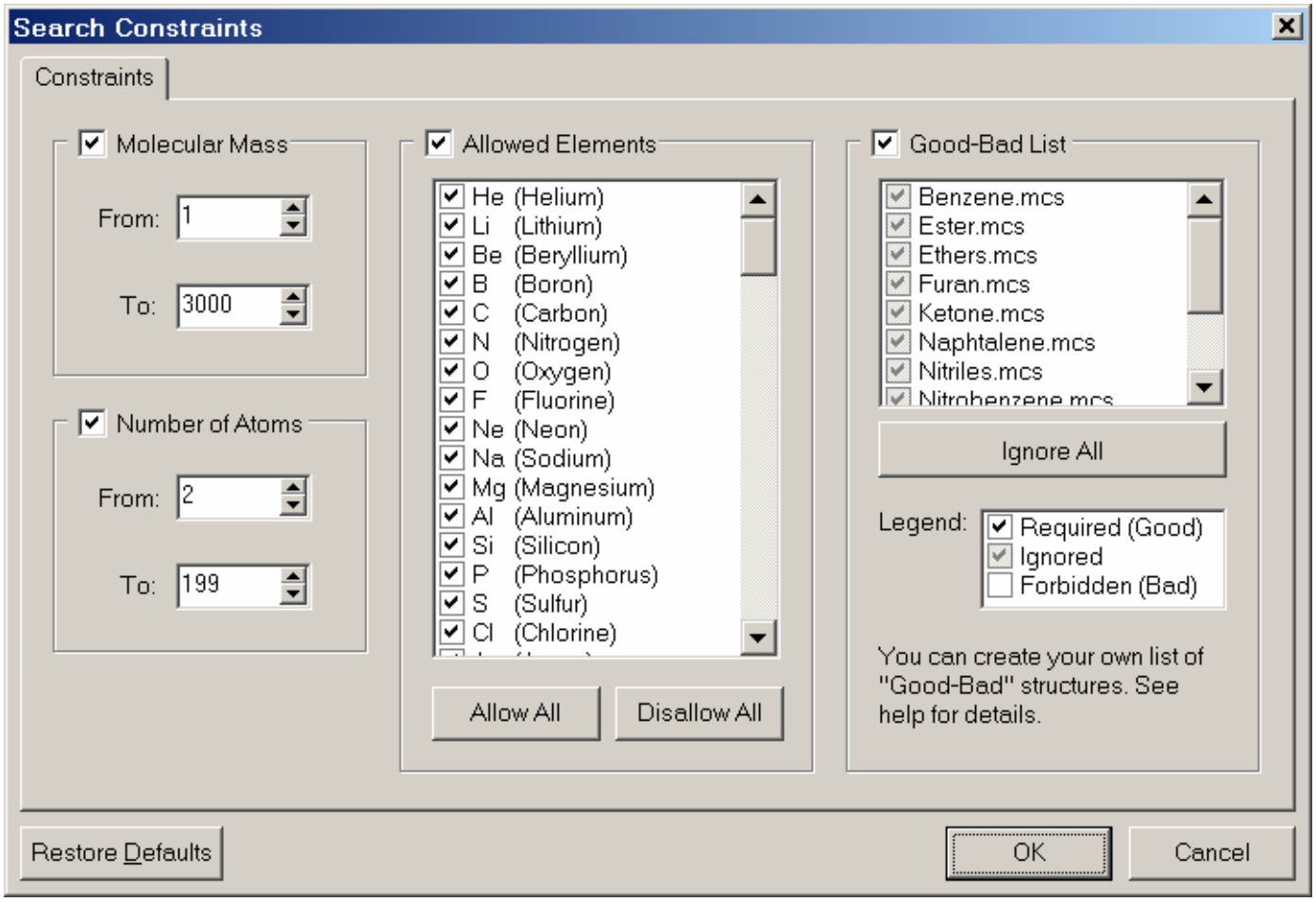


Figure 54. Search Constraints dialog window showing allowed elements and functional groups

Mass Spectral Data Exchange Between Modules

Database Manager serves as a gateway for several modules. Mass spectra can be imported directly from a file, the Chromatographic Processor, Excel, or Xcalibur. Spectra can also be exchanged between the Database Manager modules. In addition, Mass Frontier is able to automatically establish a link between Database Manager records and other modules that need to access spectral or structural data. With the ability to link equivalent spectral information, you can access the original data that was supplied as input information for one of the many interpretation methods available in Mass Frontier. The direct feedback between source (records and mass spectra) and results (mechanisms, bar code spectra, classes and projections) helps you to keep track of all the modules that originate from a single source. This feature makes it easier to use more modules simultaneously, thus enabling additional problem-solving approaches.

Fragment Assignment to Spectral Peaks

Mass Frontier enables the automated prediction of fragments from a structure you provide. If you start the generation of fragments from the Database Manager and the active record contains a structure (an MS^1 structure for Spectral Tree), the generated fragments are automatically linked with peaks in the spectrum (spectra in the Spectral Tree) according to their *m/z* values. After a generation, highlighted (explained) peaks are displayed in a different color (the default is red) in the original mass spectrum. Clicking a highlighted peak reveals all the mechanisms leading to it. In addition, generated fragments (a corresponding Fragments & Mechanisms window must be open) can be assigned automatically to peaks in a spectrum by clicking the **Spectra Processing** button and choosing **Auto Fragment Annotation of Peaks**.

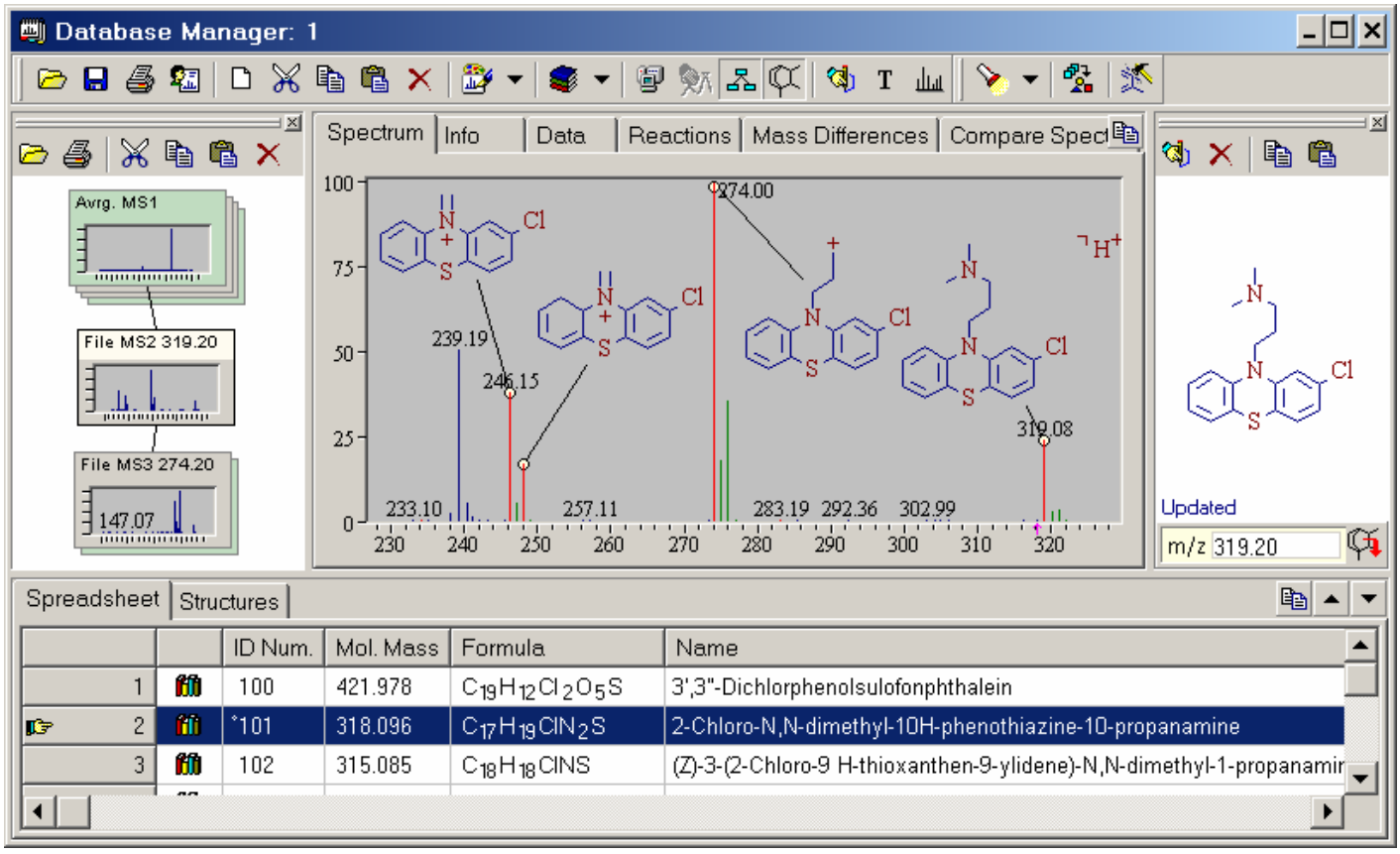


Figure 55. Database Manager window

In addition to automated peak annotation, you can also manually assign a fragment structure to any peak by clicking the **Assign Fragment Structure to Peak** button.

Chapter 5 Library Utilities

Mass Frontier provides a number of methods for creating and maintaining mass spectral and chromatographic libraries with chemical structures. To help you visually distinguish between libraries, each library has its own icon. You can select an icon for your libraries. National Institute of Standards and Technology (NIST) main and replicate and HighChem MS^n libraries icons are assigned automatically by the program and cannot be changed. Up to 255 libraries can be installed at once.

Both low- and high-resolution mass spectra are supported. Mass spectra can be stored as individual single spectra or as MS^n trees with multiple node spectra. Data-reduced chromatograms with selected scans and components can be stored and searched. Any information in a library record can be updated, except the NIST database which is read-only.

This chapter contains the following sections:

Microsoft SQL Server 2000 Desktop Engine

Mass Frontier library utilities are based on the Microsoft SQL Server 2000 desktop engine. The SQL Server 2000 desktop engine is a data engine built and based on core SQL server technology. It is a reliable storage engine and query processor for desktop applications. The SQL Server 2000 desktop engine is a royalty-free database engine that is fully compatible with the SQL Server. The SQL Server 2000 desktop engine is included in the Mass Frontier installation process.

The SQL Server 2000 desktop engine is Mass Frontier's background application. Library utilities are seamlessly integrated in a graphical interface and you do not need to directly interact with the database engine. The Mass Frontier database concept offers a broad spectrum of functionality and provides flexible features, but these are connected with an overhead problem that affects all modern database systems. Therefore, complex searches in large amounts of data might require longer search times than in previous versions of Mass Frontier.

There is a restriction in the SQL Server 2000 desktop engine that allows the creation of databases with a maximum file size of 2 GB. Therefore, libraries created in Mass Frontier cannot exceed this file size. If you try to store data above this limit, Mass Frontier informs you that this action cannot be completed and to store the extra data in a new library. This problem occurs more often with chromatographic libraries than with spectral libraries. For example, the NIST 2002 database with 175000 spectra takes up 800 MB in the SQL Server 2000 desktop engine format. However, if you need to create large spectral or chromatographic databases, the SQL Server 2000 desktop engine can be upgraded to later SQL Server editions (Enterprise, Standard, or Personal) that have terrabyte capacity. Mass Frontier libraries are fully compatible with other editions of the Microsoft SQL Server. Later SQL Server editions must be setup individually by a database expert. If you require additional information on how to upgrade the SQL Server 2000 desktop engine, contact Microsoft or the HighChem, Ltd. database group.

NIST/EPA/NIH Mass Spectral Database

Mass Frontier can work as a stand-alone application, or combined with the NIST/EPA/NIH Mass Spectral Database 2002 (NIST 2002 Library). If the NIST 2002 Library has not been simultaneously installed with Mass Frontier, the library can be either installed, if the library is in Mass Frontier format, or imported from NIST file format. During import of a NIST library, the data is converted to the SQL Server database format used by Mass Frontier. This process can take several hours. If you select the NIST Main library, the NIST Replicate library is installed automatically. The NIST Replicate library cannot be installed as a stand-alone library without the Main library.

To import NIST library or any other library in NIST format, choose **Library > Spectral Library > Import NIST Library**.

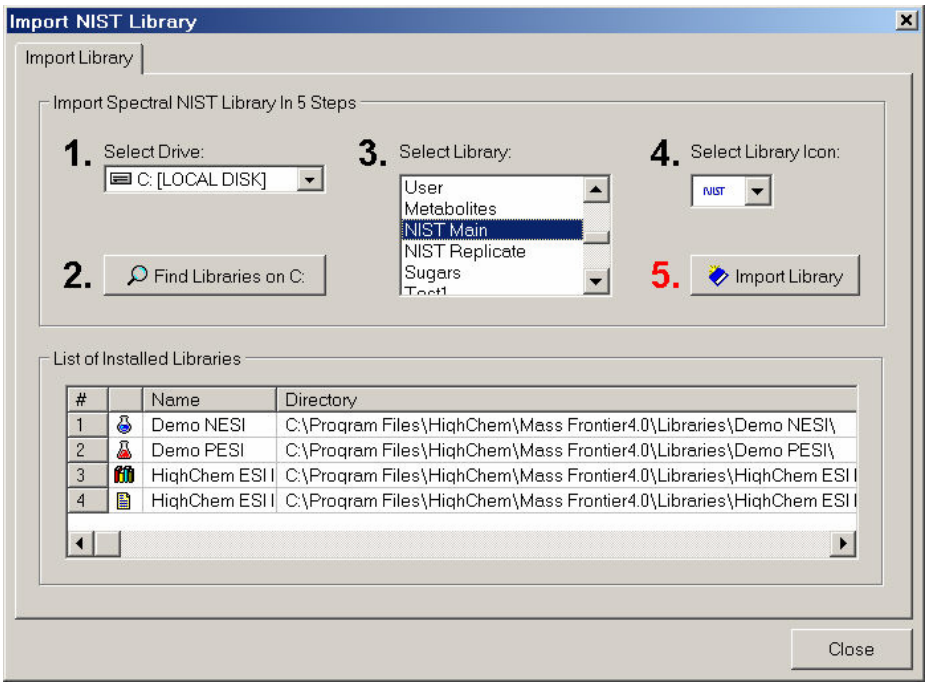


Figure 56. Import NIST or Mass Frontier 5.0 User library dialog window

Library Installation

To install a user, NIST, or HighChem MS^n library into Mass Frontier, choose **Library > Spectral Library > Install Library** from the main menu. When the Install Library dialog window appears, follow these installation steps:

1. Select the drive where the library is located.
2. Click the **Find Library on X:** button (X is the driver you have selected in step 1). The program scans the drive for libraries. If a library is found, proceed with step 3.
3. Select the library you want to install.
4. Choose an icon for the library.
5. Click **Install Library**.

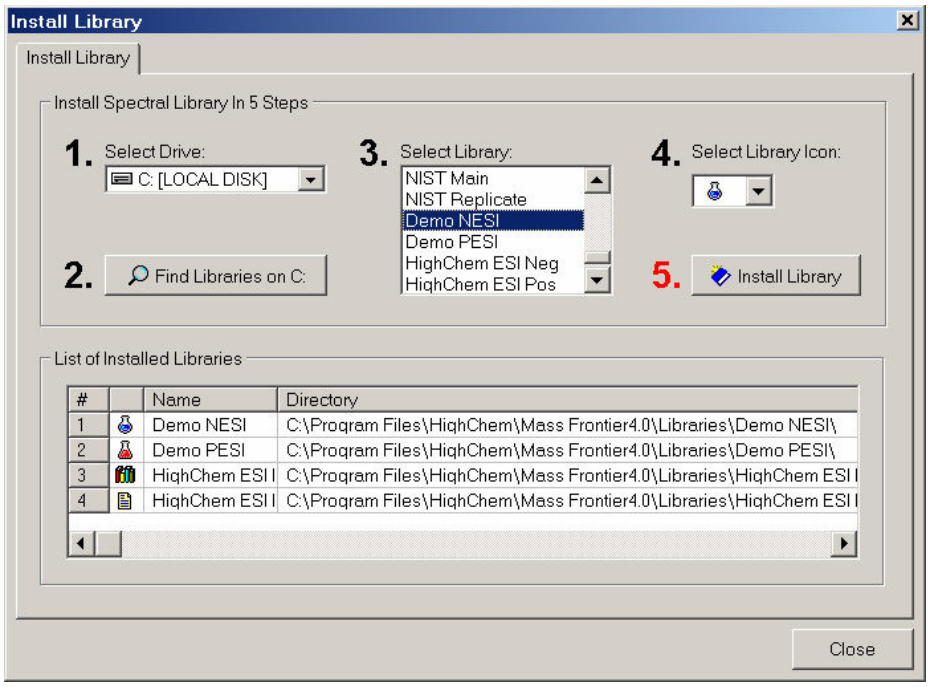


Figure 57. Install Library dialog window

If the library has been successfully installed, the icon, name, and path of the library are displayed in the List of Installed Libraries grid in the Install Library dialog window.

Note A library can be installed in Mass Frontier only if the data is in the Mass Frontier SQL Server database format. For libraries in NIST format (NIST Library, user libraries from Mass Frontier 3.0 and earlier versions) use the Import NIST Library option instead. Choose **Library > Spectral Library > Import NIST Library**.

Creating User Libraries

Experimental data, as well as reference spectra, can be stored in user libraries to organize your mass spectral, chromatographic and structural data. With Mass Frontier, you can create highly annotated user libraries with structures. Both low- and high-resolution mass spectra are supported. Mass spectra can be stored as individual single spectra or as MS^n trees with multiple node spectra. Data reduced chromatograms with selected scans and components can be stored in user libraries and searched.

To create a user library, choose the **Library > Spectral Library > Create User Library** item in the main menu. When the Create User Library dialog window appears, follow these steps:

1. Enter or select the directory for the user library. We recommend accepting the default directory.

2. Enter the library name. Only the characters a-z, A-Z, space and 0-9 can be used.

3. Select the library icon.

4. Click the Create & Install User Library icon.

If the library has been successfully created the icon, name and path of the library are displayed in the List of Installed Libraries grid. The program creates an individual subdirectory for each new library, complete with the name of the library. All the library data are stored in an .mdf file. The full library path is displayed in the Directory column.

An unlimited number of user libraries can be created, but no more then 255 libraries can be installed in Mass Frontier at the same time. If you require more installed libraries, you can temporarily uninstall a library you do not need and create or install the library you want to work with.

Note A higher number of spectral libraries slows the program start, consume system resources and potentially lengthen the search processes.

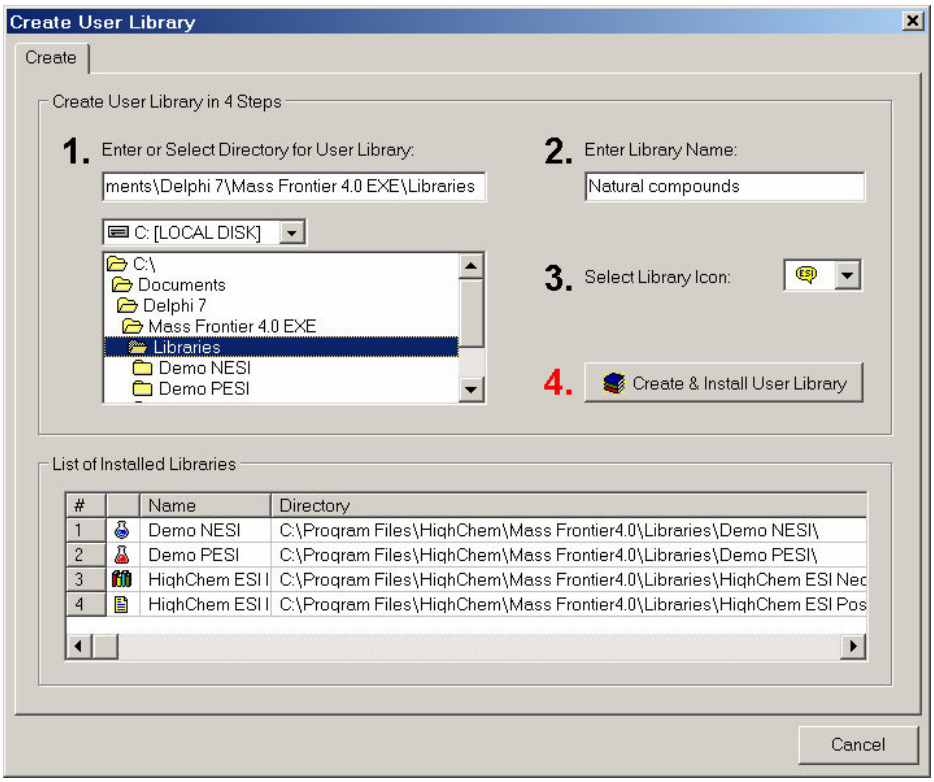


Figure 58. Create User Library dialog window

Uninstall A Library

Because database files can be copied or deleted only if the library is uninstalled, you might need to uninstall a library. In addition, a large number of libraries can slow the overall performance of Mass Frontier so the uninstallation of unused libraries can be beneficial. If you uninstall a library, the library is not deleted. The library reference is merely removed from the program without the loss of spectral or structural information. An uninstalled library can be reinstalled at any time. With Mass Frontier, you can install or uninstall a library as many times as you want.

To uninstall a library

1. Choose **Library > Uninstall Library.** The Uninstall Library window appears. See Figure 59.

2. Select the library to uninstall.

3. Click **Uninstall**.

Note Database files with the .mdf extension cannot be copied or deleted when the library is installed. If you need to make a backup copy or you want to permanently delete a library from your hard drive, you must uninstall it first.

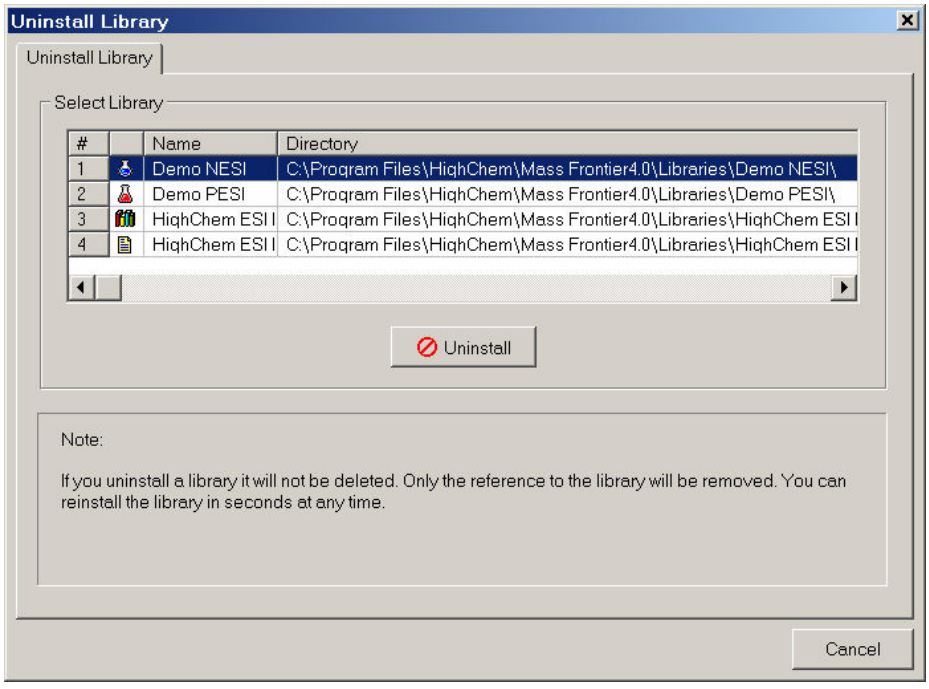


Figure 59. Uninstall Library dialog window

Adding Records to a User Library

Mass Frontier provides a way of adding new records (spectra, trees, chromatograms, structures and compound identification information) to a user library. This option creates new library entry with a new ID. To update an existing record, use the Save Changes in Library option.

To add one or more records from the Database Manager to a user library

1. Select the records you want to add to the user library from the spreadsheet in the Database Manager.
2. Choose **Library > Add To Library**.
3. When the Add To Library dialog window appears, select the library you want to add to the records by choosing the appropriate button.
4. Click **Add**.

Note Only selected records in a Database Manager Spreadsheet can be added to a user library.

A spectrum you intend to add to a library must not contain more than 3000 peaks and the *m/z* value might not be greater then 3000. A single structure accompanying a spectrum in a user library must not contain more then 199 non-hydrogen atoms.

Check whether the structure associated with the spectrum you want to add to a user library is already present in the library. Perform redundancy checks by clicking the **Check** button in the Check Structure Redundancy window. When this option is selected, the program compares each structure to be added to the library with structures in the selected library. Use this option if the structure is present, otherwise this option is ignored.

Deleting Library Entries

Deleting a library record causes the information to be irreversibly lost. There is no undo function for deleted library records. You cannot delete library entries from a NIST library.

To delete one or more records

1. Retrieve the records you want to delete using any search.
2. Select the records you want to delete in the Database Manager Spreadsheet.
3. Choose **Library > Delete** from the Library menu.
4. When the Delete From Library window appears, review the spectra and structures carefully.
5. If you want to delete all records at once that are displayed in Preview, choose **Delete without confirmation**. If you want to delete record by record, select **Confirm Each Spectra/Structure Pair Before Deletion**.
6. Click **Delete**.

Note If you delete a record from a library, the ID numbers of records with a higher ID than the deleted record are decremented. This means that if you delete one or more records, there are no ID number gaps in the library.

Saving Changes in Libraries

Any record information (spectra, trees, chromatograms, structures and compound identification information) can be changed and saved in a library. If anything is changed in the record, a small circle is displayed in the Database Manager in the ID Num. column in the spreadsheet. Changes in records are not automatically updated in the database unless you save them.

To save changes in library, choose **Library > Spectral Library > Save Changes in Library**.

If you try to save a newly created record, you are redirected to the Add New Records to Library option.

Note If you try to close a Database Manager window, or Mass Frontier with changed unsaved records, you are prompted to save them.

SQL Server and Library Tools

To open the SQL Server & Library Tools dialog window, choose **Library > Library Tools** from the main menu tool bar. The SQL Server & Library Tools dialog window contains three tabs:

- Spectral Libraries
- Fragmentation Libraries
- SQL Server

Spectral Libraries

The Info grid lists all the attached spectral and chromatographic libraries to the SQL Server. Libraries incompatible with the current version of Mass Frontier can be detached or removed from your system by using the context menu which you access by right-clicking.

In the Update group, you can compress or re-index selected libraries to shrink their sizes, or defragment indexes if the performance decreases. Note that these actions might require additional time to perform, depending on the amount of stored data.

Note Make backup copies of your libraries before reannexing and compressing spectral libraries.

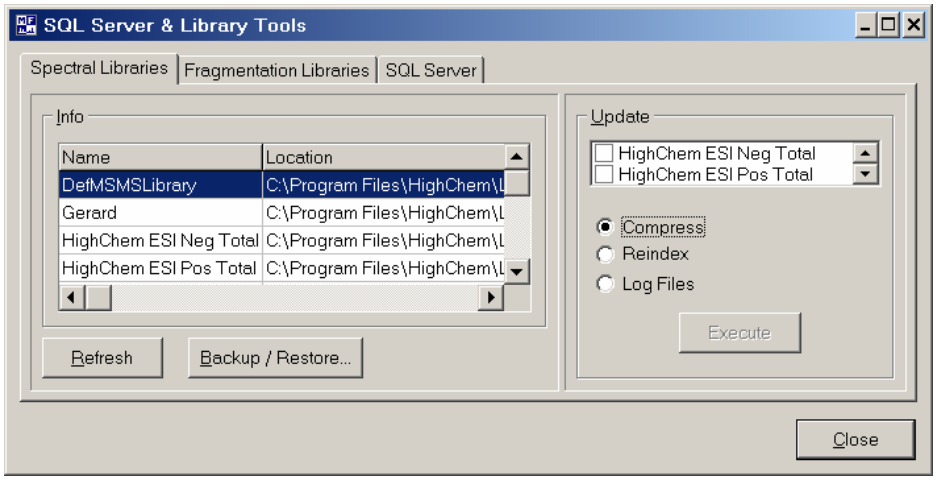


Figure 60. Spectral Libraries page of the SQL Server & Library Tools dialog window

Fragmentation Libraries

The **Info** grid lists all the spectral and chromatographic libraries attached to the SQL Server with their paths. Libraries incompatible with the current version of Mass Frontier can be detached or removed from your system by using the context menu which you access by right-clicking.

In the Update group, you can compress or re-index selected libraries to shrink their sizes, or defragment indexes if the performance decreases. Be aware that these actions might require additional time to perform, depending on the amount of stored data. If you build large fragmentation libraries, perform these actions regularly.

Note Make backup copies of your libraries before re-indexing and compressing fragmentation libraries.

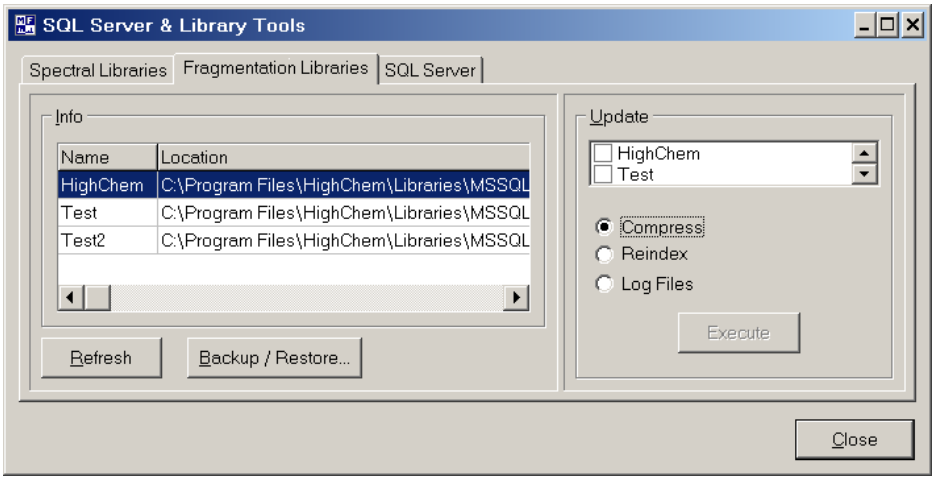


Figure 61. Fragmentation Libraries page of the SQL Server & Library Tools dialog window

SQL Server

Note Use this page only if you are familiar with SQL Server database technology.

The Monitor box displays information about the SQL Server status and contains buttons for starting, running and stopping the SQL Server service.

Use the Extended Stored Procedures box in the Reload box to redefine the extended stored procedures used by the SQL Server. These procedures are essential for spectra search functionality. Use this action if a library search does not work properly.

The Default Library box in the Reload box re-creates damaged internal database files. Use this action if the message *Cannot open library DefMSMSLibrary!* appears.

The Connected Libraries box in the Reload box reconnects all the spectral libraries. Use this action if the SQL Server was stopped and restarted outside of Mass Frontier; for example, in the task bar notification area.

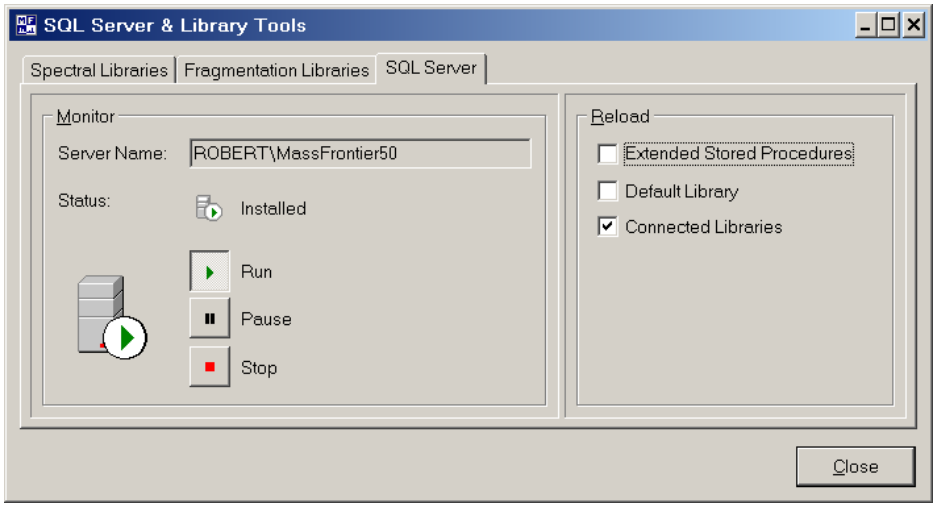


Figure 62. SQL Server page of the SQL Server & Library Tools dialog window

Backup and Restore Libraries

It is not advisable to back-up and restore spectral or fragmentation libraries by a library files copy and paste operation. Use the backup and restore procedures in Mass Frontier.

To backup a library

1. Open the SQL Server & Library Tools dialog window by choosing **Library > Library Tools**.

2. Select the library type by clicking the Spectral Library or the Fragmentation Library tab.

3. Click the **Backup/Restore** button and the dialog window appears.

4. Specify the library you want to back up and type the name, description and destination of the backup files.

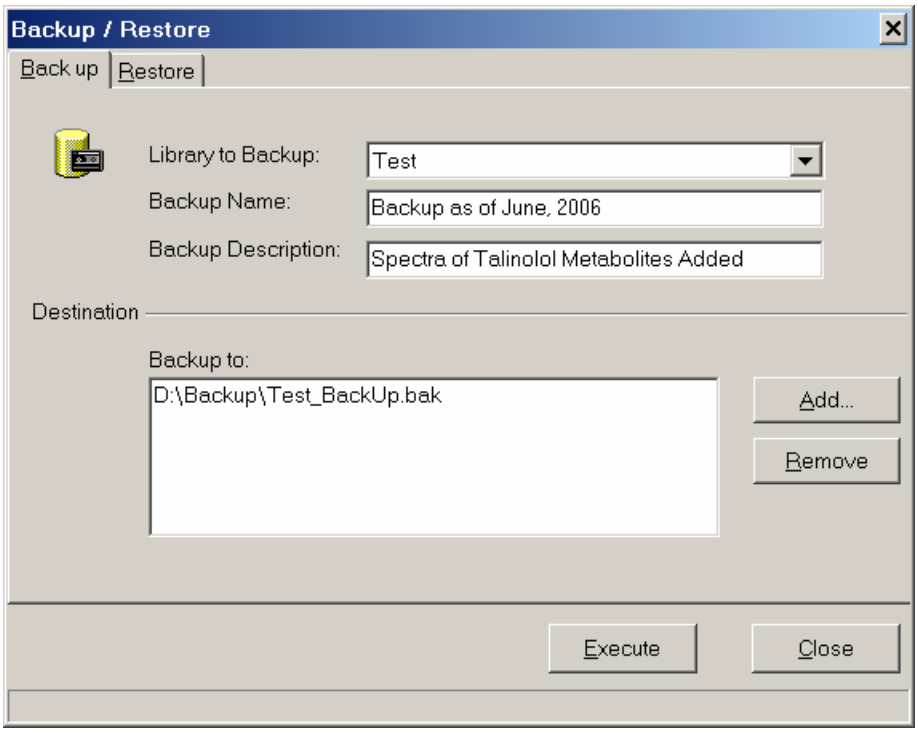


Figure 63. Backup/Restore dialog window showing Backup page

You can restore library data from previous backups into a new library or into an existing library.

Note If you restore library data into an existing library, its data is irreversibly overwritten. Thermo recommends that you restore data into a new library with a unique name.

To restore a library from backup files

1. Open the SQL Server & Library Tools dialog window by choosing **Library > Library Tools**.
2. Select the library type by clicking the Spectral Library or the Fragmentation Library tab.
3. Click the **Backup/Restore** button and the dialog window opens.
4. Specify the library name in the Library to Backup box where the data is to be restored. Type a new library name here.
5. Select the library data to restore from the list of backups.

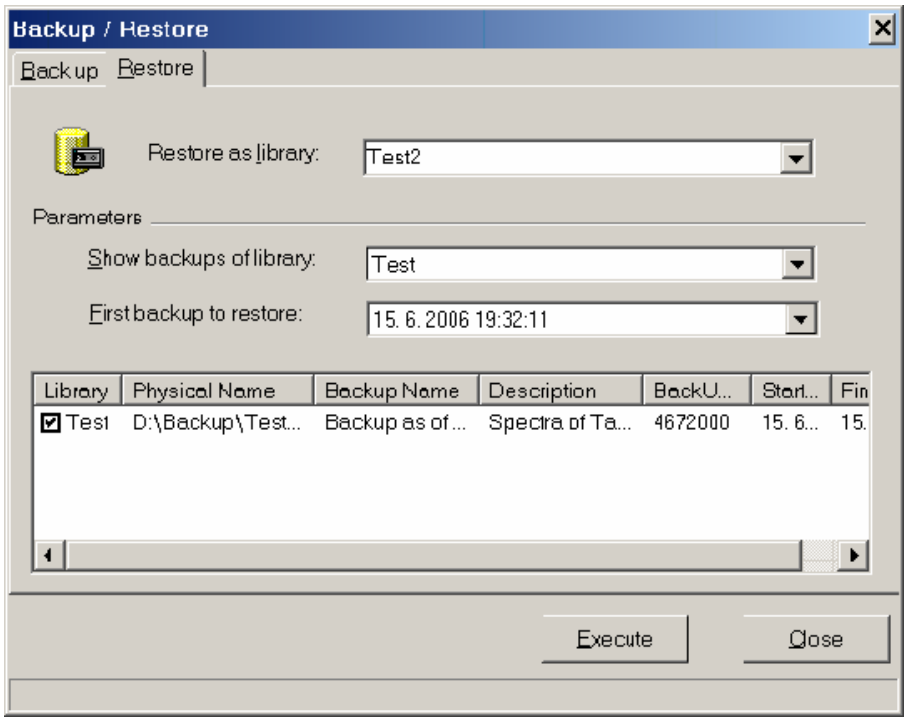


Figure 64. Backup/Restore window showing Restore page

Chapter 6 Fragments and Mechanisms

One of Mass Frontier´s features is the automated generation of possible fragments at an expert level, including complete fragmentation and rearrangement mechanisms, starting from a chemical structure you supply. The Fragments & Mechanisms module provides information about basic fragmentation and rearrangement processes that might occur in a mass spectrometer.

Use generated fragments and corresponding mechanisms for:

- Checking consistency between a chemical structure and its mass spectrum.
- Confirming library search identifications
- Recognizing the structural differences between spectra of closely related compounds
- Interpreting the spectra of isotopically-labeled compounds
- Illustrating the structure-spectra relationship for educational purposes

This chapter contains the following sections:

Features

The Fragments & Mechanisms module is based on an expert system that uses a mathematical approach for the simulation of unimolecular ion-decomposition reactions. It is important for you to understand what the system is and is not capable of, and what you can and cannot expect from this module.

The expert system, which generates possible fragmentation and rearrangement pathways, is based on the following assumptions.

General Fragmentation and Rearrangement Rules

The system optionally predicts reaction pathways that are based on general fragmentation and rearrangement rules. Compound-specific mechanisms that cannot be applied generally are not included in this feature. Use this feature in combination with a substructure search for identifying specific compound classes.

Fragmentation Library Mechanisms

The system optionally accesses an intelligent fragmentation mechanism knowledge base for the prediction of unimolecular decomposition reactions. HighChem Fragmentation Library currently contains around 19000 individual mechanisms. User mechanisms can be included in fragmentation prediction as well.

Charge Localization Concept

Every ion-decomposition reaction that is generated is based on the charge localization concept. The location of the charge site in all precursor and product ions is exactly determined by the program. Mass Frontier internally generates resonance reactions which, by default, are not displayed. Note that charge sites can be moved to distant locations by these reactions and in some complicated structures it might appear that the charge localization concept has been violated. If a reaction step is not clear, you can instruct the program to display mechanisms along with resonance reactions. It is, however, possible to use an unspecified charge location that is internally transformed to all combinatorial structures with a localized charge.

Unimolecular Linear Reaction Mechanisms

Mass Frontier generates only unimolecular reactions. The reaction pathways are displayed as linear reaction mechanisms which incorporate one intermediate left-hand site and one intermediate right-hand site for each reaction step. Thus, only ionic products are included in reaction pathways and neutral fragments are not displayed.

Even- Electron Rule

Reactions mechanisms are generated in accordance with the Even-Electron rule. The Even-Electron rule states that the homolytic bond cleavages of an even electron ion are energetically unfavorable. Therefore, Mass Frontier never generates radical cations from an even-electron ion.

Bond Cleavages Only

Using the General Fragmentation and Rearrangement Rules option, fragments can only be generated from bond cleavages. Bond creation is not supported, with the exception of the creation of an H-X bond (where X=any element) in hydrogen rearrangements. Thus, for this option the expert system does not include ring contractions, cyclizations, or skeletal and non-hydrogen rearrangements. The fragmentation library option supports bond creation with all rearrangements and ring transformations.

Ionization Methods

Mass Frontier supports Electron Impact, protonation, deprotonation, cluster ion formation, alkali metal adducts, and chemical ionization methods.

Formally Possible Solutions

The mechanisms generated by Mass Frontier contain formally possible reaction steps. The determination of the stability of product ions from thermodynamic data or rates of reaction is not performed. When evaluating generated mechanisms, remember that short and uncomplicated reaction pathways are more favorable than complex mechanisms involving complicated, multistep hydrogen rearrangements.

Fragmentation, Rearrangement and Resonance Reactions

Mass Frontier utilizes general unimolecular reactions for prediction of fragmentation, rearrangement and resonance mechanisms.

To preview unimolecular reactions used by the program

Click the **Reaction Mechanisms Overview** button in any Fragments & Mechanisms window. The Reaction Mechanisms Overview window appears. See Figure 65.

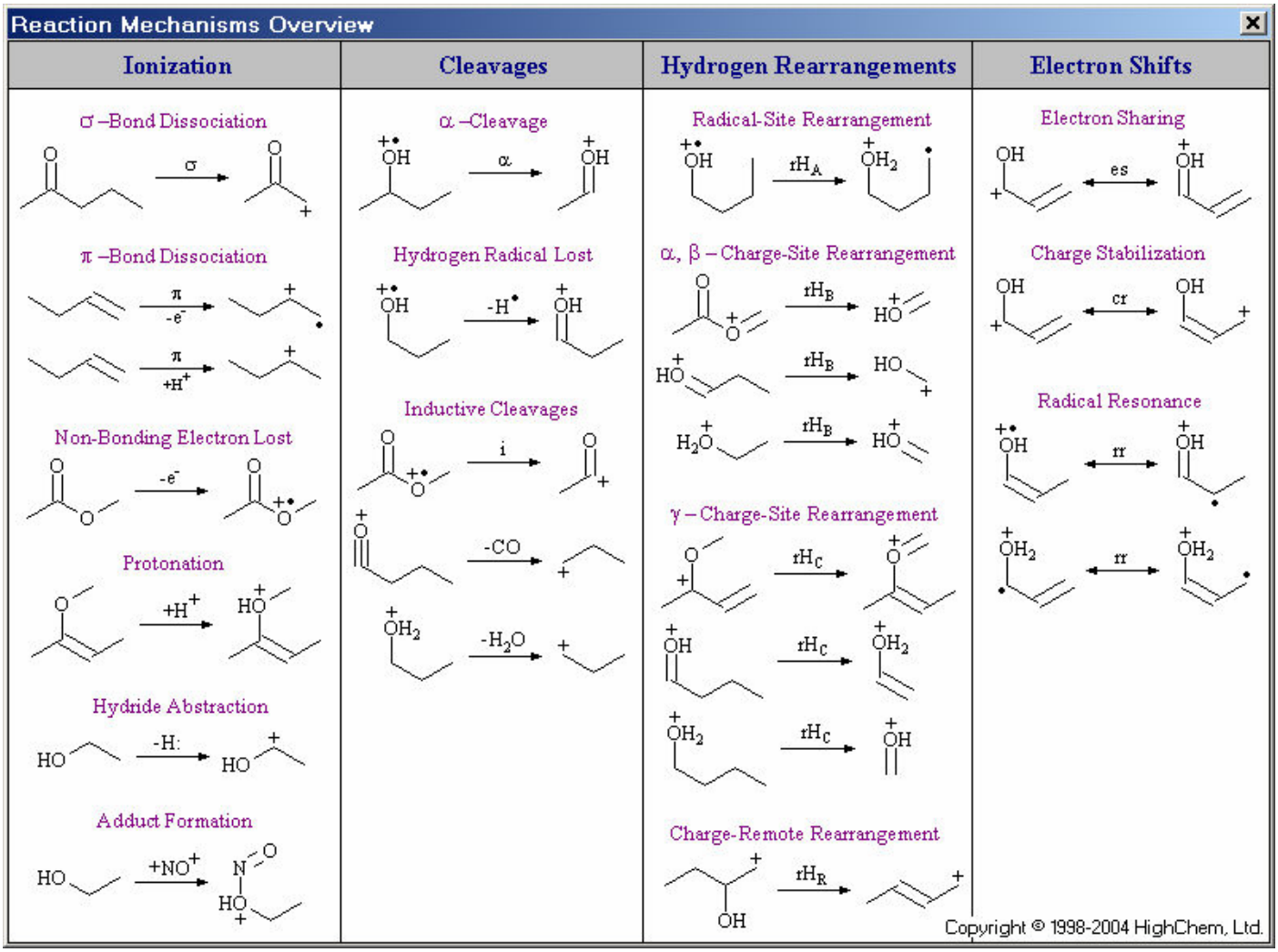


Figure 65. Fragmentation, rearrangement, and resonance reactions

Table 1. Reaction formalism used in Mass Frontier

α	α - cleavage
π	π electron ionization
σ	Sigma electron ionization
cr	Charge stabilization
$-e^-$	Non-bonding electron ionization
es	Electron sharing
i	Inductive cleavage
$+H^+$	Protonation
$-H^\bullet$	Hydrogen radical loss
-H:	Hydride abstraction
rH_A	Radical site rearrangement
rH_B	Charge site rearrangement (α, β)
rH_C	Charge site rearrangement (γ)
rH_R	Charge-remote rearrangement
$rH_{1,2}$	Hydrogen shift to adjacent position
rr	Radical resonance
Lib	Fragmentation Library Mechanisms

Starting Generation

Fragmentation and rearrangement pathways can be generated from any structure you supply, including ions and isotopically-labeled compounds. The structure can originate from the Structure Editor, Database Manager, or from a Fragments & Mechanisms window. Before the generation is performed, the program checks the input structure for errors. If any errors are found, a message box appears that informs you about these errors and the generation ends.

When you start a generation from the Structure Editor, in contrast to the copy function or substructure search, the structure does not have to be selected. The program assumes that the complete structure is intended as input for the generation.

If a generation is started from the Database Manager window, the program automatically links the generated fragments in the Fragments & Mechanisms window with the corresponding spectrum in the Database Manager window. Peaks having the same *m/z* value as the generated fragments are highlighted in the color set in the Spectra Layout dialog window (by default, red). Selecting a highlighted peak reveals all the pathways leading to it.

To start a generation of possible fragmentation and rearrangement pathways

1. Click the  button in the Structure Editor, Database Manager, or Fragments & Mechanisms window.

2. Or make the Structure Editor or Database Manager active and then choose **Tools > Fragments & Mechanisms**.

When you start generation, a Reaction Restriction window appears to specify the type of knowledge used for fragmentation.

During the generation of possible fragmentation and rearrangement pathways, the Generation of Fragments & Mechanisms window appears. See Figure 66.

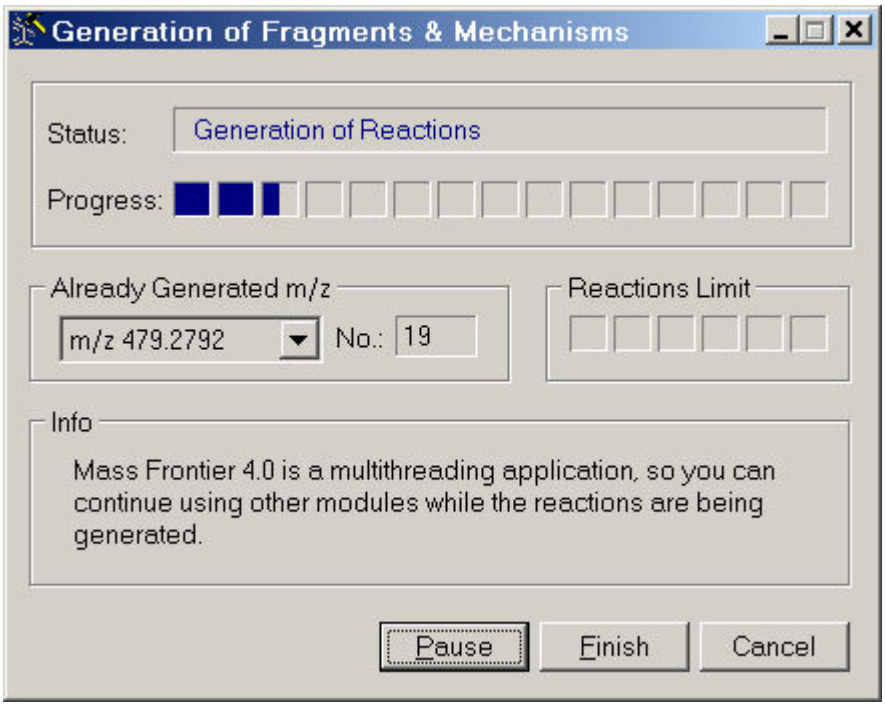


Figure 66. Generation of Fragments & Mechanisms dialog window

The Already Generated m/z box that stores *m/z* values of ions that have already been generated. Click the down arrow to display a list of ions that have been generated. The total number of ions that have already been generated appears next to the box.

While a generation is in progress, the Reactions Limit bar provides an approximate indication of how many temporary internal reactions have been generated from a particular structure. Large and structurally complicated molecules can produce a large number of reactions. Because the generated reactions consume a significant amount of memory, there is a limit to the number of temporarily-generated reactions. If the reactions limit is reached, the generation stops and a message box appears to remind you of this. Even if the generation is stopped, the fragments and mechanisms generated up to that point are displayed. The fragments and mechanisms that the system generates first are the most important. Therefore, when a generation is stopped, the most important fragments have likely already been generated. However, if you are missing an important fragment and you assume that it is because the generation was interrupted, you can increase the number of internal reactions in the Reaction Restriction window in the Size tab.

The message *Mass Frontier is a multithreading application, so you can use this program while reactions are being generated* appears in the Info window. The multithreading strategy allows more than one task to be performed simultaneously. For example, while reactions are being generated, you can search libraries or analyze your spectra. You can even run two or more generations of fragments and mechanisms at the same time.

Click the **Pause** button to temporarily interrupt generation to redirect processing power to other processes that might be running simultaneously in Mass Frontier or to other Windows applications.

Click the **Finish** button to stop generation. Fragments, along with their production mechanisms generated up to that point, appear.

Click the **Cancel** button to end a generation. It can take several seconds for the window to disappear after the generation has ended.

Fragments & Mechanisms Window

Once reaction-generating process has finished, the Fragments & Mechanisms window appears. See Figure 67.

A Fragments & Mechanisms window contains complete mechanisms of ion-decomposition reactions, or the resulting fragments only. To display the mechanism or fragment for a particular *m/z* value, choose the appropriate *m/z* tab. If you do not like the tab control, replace it with a box in the Options | Mechanism Layout dialog window.

A Fragments & Mechanisms window contains two **Copy** buttons. Use the first one to copy mechanisms in the Enhanced Windows Metafile format .emf. Use the second one o copy a selected fragment. If you copy a fragment, you can paste it to the Structure Editor, Database Manager in Mass Frontier, or to any Windows application.

Note Only a selected fragment can be copied to the Clipboard.

Several possible isobaric fragments can be generated for a particular *m/z* value. If more then one fragment is generated for a *m/z* value, a hand-shaped pointer with the caption *Select possible fragments with m/z xy* moves from right to left to draw your attention to this. Click the numbered buttons next to the hand pointer to display the isobaric fragments and their corresponding mechanisms.

The fragments are sorted according to the simplicity of their production mechanism. Thus, fragment number 1 is produced by the simplest (shortest) mechanism. The isobaric fragments are usually isomers of the same fragments with a different charge, radical or, p-bond location.

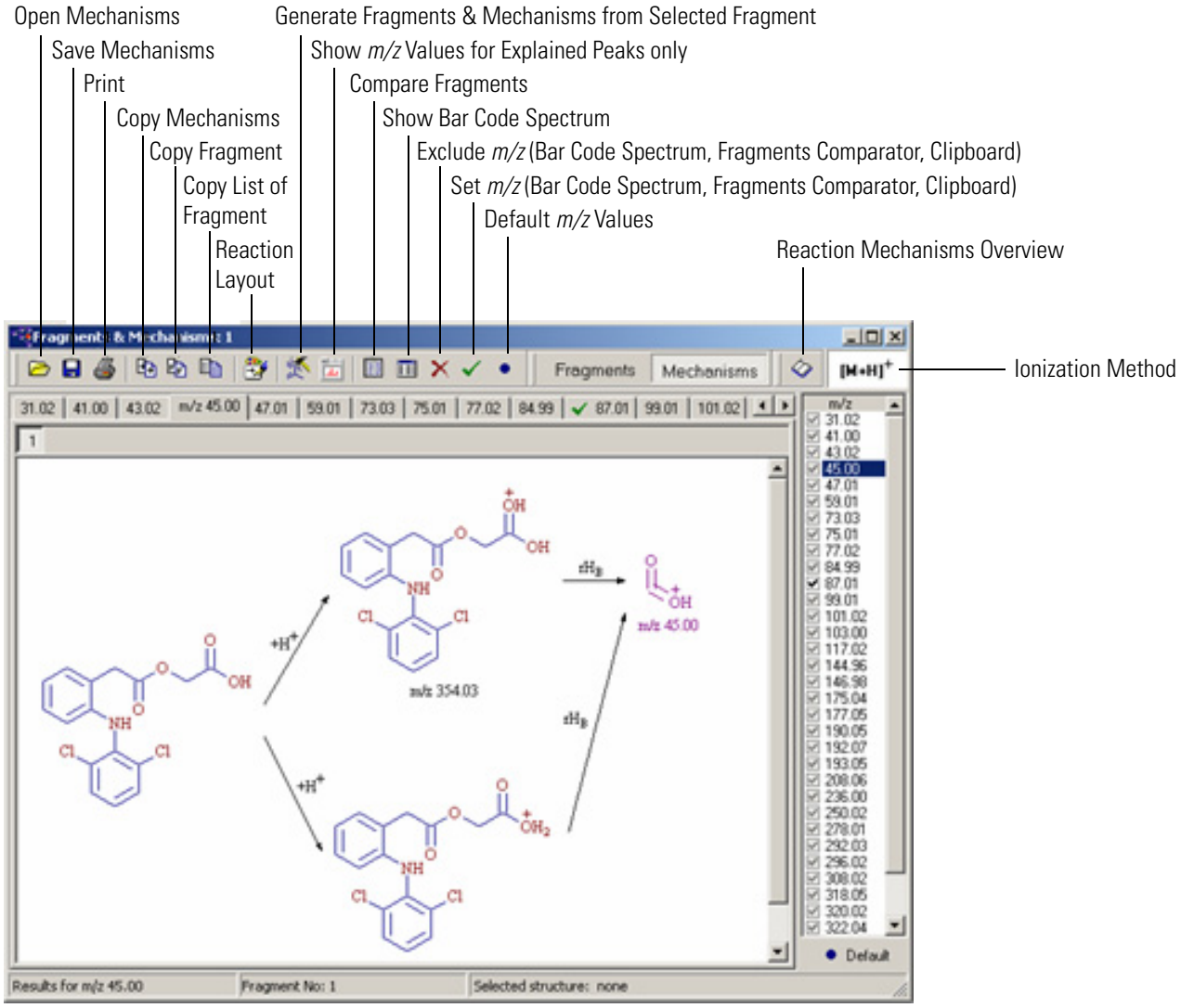


Figure 67. Fragments & Mechanisms window

Reaction Restrictions

Use Reaction Restrictions to change the default settings of the ionization method and of basic fragmentation and rearrangement mechanisms. To set the *m/z* precision of calculated fragments in the Mass Settings dialog window, choose **Options > Mass Settings**.

To Restrict Restrictions

Choose **Options > Reaction Restrictions**.

The Reaction Restrictions window contains six tabs: Knowledge Base, Ionization & Cleavage, H-Rearrangement, Resonance, Additional, and Sizes.

If you have established restriction settings that you want to apply frequently to a specific compound class, save the current reaction restrictions to a file with the .rrs file extension by clicking the **Save** button in the Reaction Restrictions dialog window. When you open a reaction restriction file by clicking the Open button, the dialog window adopts the restrictions saved in the file. You must click **OK** to make these restrictions active in Mass Frontier.

Note All changes in the Reaction Restrictions window take effect after the regeneration of fragments and mechanisms. The existing Fragments & Mechanisms windows are not affected by changes in Reaction Restrictions. In addition, the changes in the Reaction Restrictions window affect all subsequent generations, unless you restore default settings. Therefore, when changing the settings, remember to restore the default reaction restrictions when the changes are no longer needed.

Knowledge Base Page

You can specify the type of knowledge base the expert system uses for the prediction of fragmentation pathways in the Knowledge Base box in the Reaction Restrictions window. The options are General Fragmentation Rules, Fragmentation Library, or both.

If you choose the Fragmentation Library, select which library to use from the list of libraries. The preinstalled HighChem Fragmentation Library™ contains around 19000 mechanisms, and so calculation times are significantly longer when it is used.

There are four options connected with the Fragmentation Library knowledge base. Because some mechanisms of small fragments can fit to virtually every user-provided structure and can generate many useless fragments, there is an option to deactivate a record in the Fragmentation

Library window. To use this option, select the Active Record Only option in the Reaction Restriction window. If you do not, all mechanisms from the selected libraries are considered by the expert system.

If a neutral molecule is provided, Mass Frontier automatically simulates the ionization process according to the selected ionization mode in the Ionization & Cleavage page. Click Library Ionization Only to disable this automatic simulation and use a different ionization site by adding an ionization mechanism to the library.

Because some fragmentation library mechanisms follow general fragmentation rules, you can speed up the generation by enabling General Fragmentation Rules in the Knowledge Base group and disabling general fragmentation rules in Fragmentation Library Options group in the Ignore General Frag. Rules in the Library Reaction box.

With Mass Frontier, you can work with structures with an unspecified charge site in the fragmentation library or in the starting structure but it dramatically slows the generation process. If an unspecified charge site is used, the expert system must consider a large number of combinations for every step. Therefore, avoid using unspecified charge sites in the fragmentation library, if possible. To exclude mechanisms in the fragmentation library that contain unspecified charge sites from the generation process, click Charge Localization Concept Only.

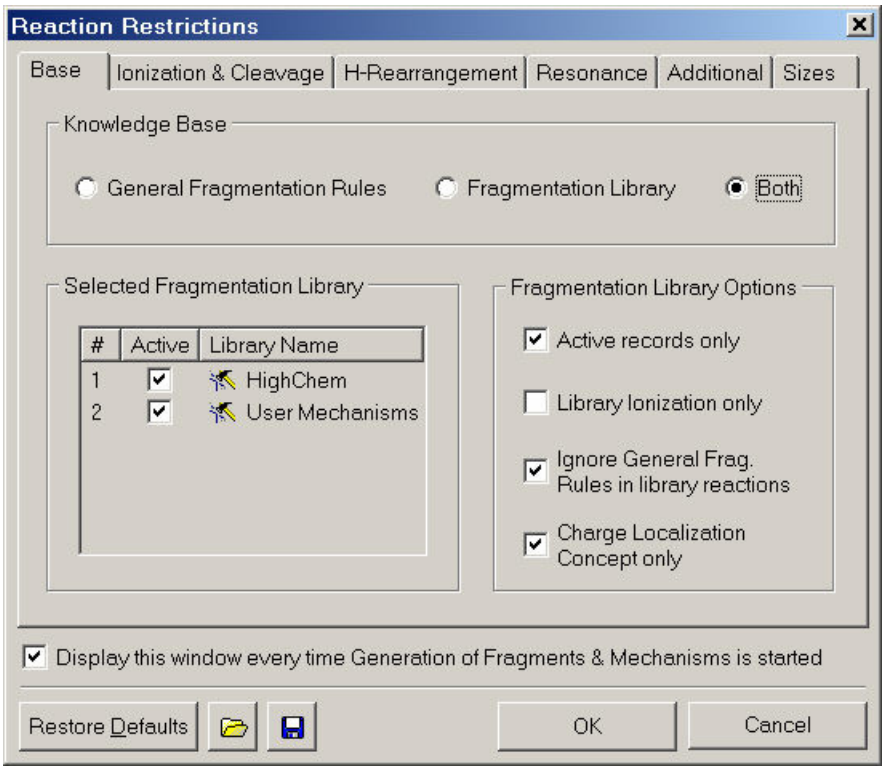


Figure 68. Reaction Restriction dialog window showing the Base page

Ionization & Cleavage Page

In the Ionization & Cleavage tab in the Reaction Restrictions window, choose among the ionization mode Electron Impact (EI) that produces M^+ ions, the protonation mode that produces $[M+H]^+$ ions, deprotonation - negative ionization ($[M-H]^-$), cluster ion formation ($[M+NH_4]^+$, $[M+H_3O]^+$), alkali metal adducts ($[M+Li]^+$, $[M+Na]^+$, $[M+K]^+$), and Chemical Ionization (CI). The protonation and deprotonation mode represents soft ionization techniques such as Electrospray Ionization (ESI), Atmospheric Pressure Chemical Ionization (APCI), Fast Atom Bombardment (FAB) and others. The chemical ionization option offers three basic ionization reactions: protonation, hydride abstraction and adduct formation. In addition, you can select one of six most common reaction gases: methane (CH_4), hydrogen (H_2), isobutane ($i\text{-}C_4H_{10}$), ammonia (NH_3), water (H_2O) and nitrogen monoxide (NO).

Note The deprotonation option does not support General Fragmentation Rules in the Knowledge Base. Deprotonation can be used only in connection with the Fragmentation Library.

When comparing generated fragments and mechanisms with a mass spectrum, always choose the correct ionization method. A warning message appears if the reaction restrictions are set for protonation techniques or chemical ionization and you are attempting to compare generated fragments with a spectrum from the NIST library which contains EI spectra only. However, if the spectrum is from a file, the mass spectrum type and ionization techniques are not checked for consistency.

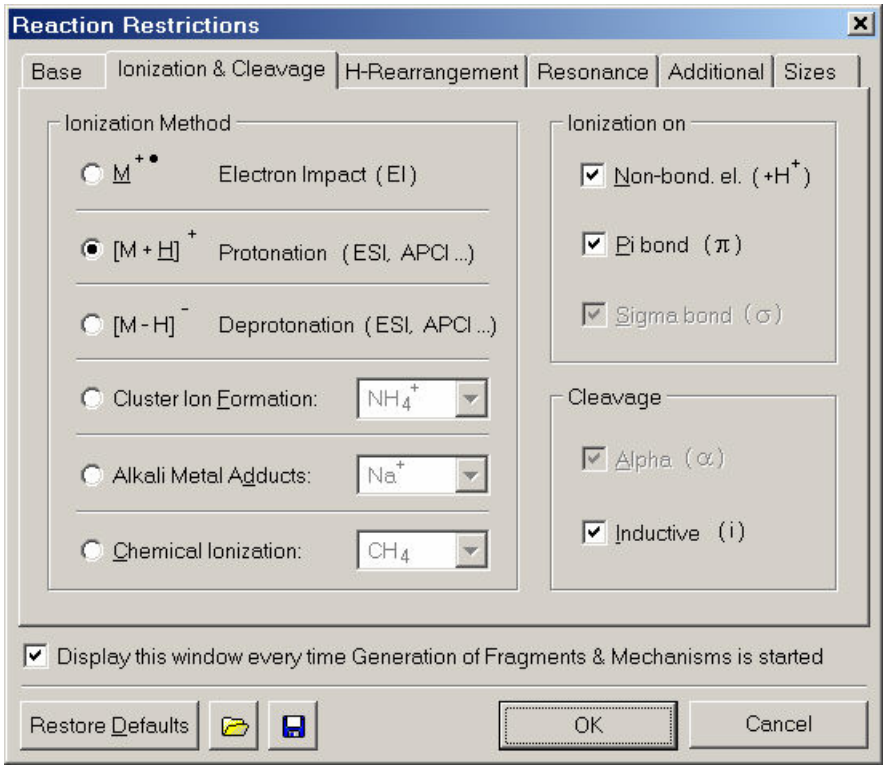


Figure 69. Reaction Restriction dialog strict showing the Ionization & Cleavage page

H-Rearrangement Page

The second tab in the Reaction Restriction window is **H-Rearrangement**. This contains controls setting the four basic hydrogen rearrangements listed in Table 2.

Table 2. Hydrogen rearrangement types

rH_A	Radical site rearrangement
rH_B	Charge site rearrangement (α, β)
rH_C	Charge site rearrangement (γ)
rH_R	Charge-remote rearrangement

The hydrogen rearrangements that involve radical (odd electron ions) rH_A are set by default for hydrogen transfer from a steric optimal atom, usually from an γ-atom (McLafferty rearrangement). Hydrogen shift to adjacent position ($rH_{1,2}$) is activated by default and cannot be deactivated.

There are two possible reasons for changing the default setup of rearrangements. First, if you are missing an important peak and you suspect an unusual rearrangement, you can compel hydrogen transfer from an α, β, γ or δ-atom. Second, you might want to simplify a mechanism by deactivating rearrangements that cause redundant reaction steps.

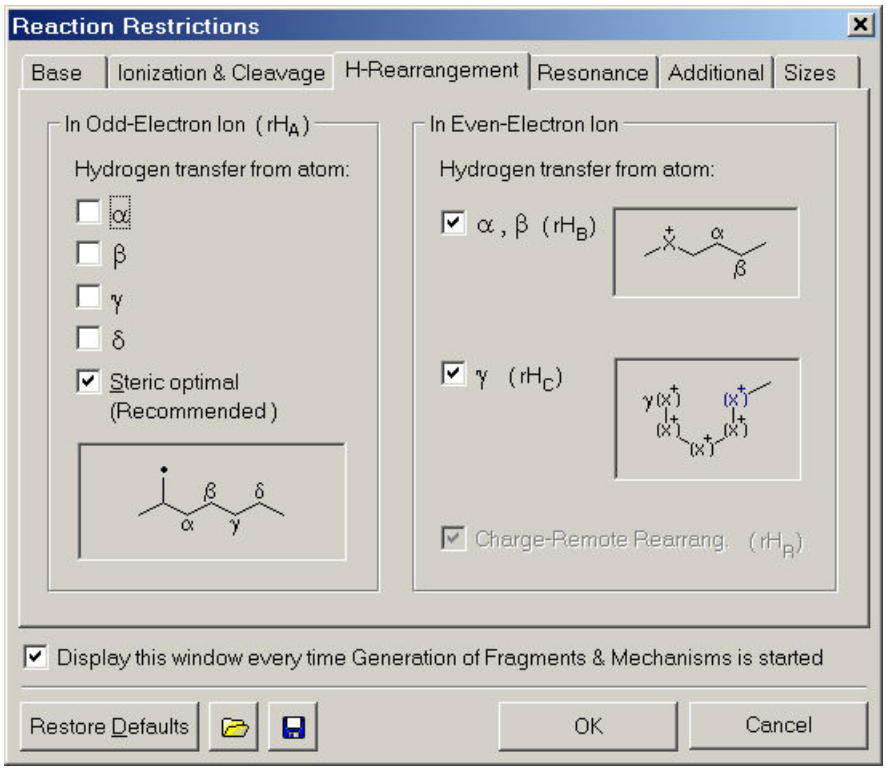


Figure 70. Reaction Restrictions dialog window showing the Hydrogen Rearrangement page

Resonance Page

Mass Frontier generates fragmentation and rearrangement mechanisms along with electron shift reactions (resonance reactions). Because these reactions might, even for small structures, cause a large number of by-products, by default, the resonance reactions are not depicted. To keep the reaction network clear, the program performs a reduction of reaction complexity, by not displaying resonance reactions. Thus, elementary reaction steps that include resonance reactions are merged into a single step. If you encounter unclear elementary reaction steps, instruct the system to display all resonance reactions by clicking the **Yes** button in the Display Resonance Reactions box in the Resonance page in the Reaction Restrictions dialog window. See Figure 71.

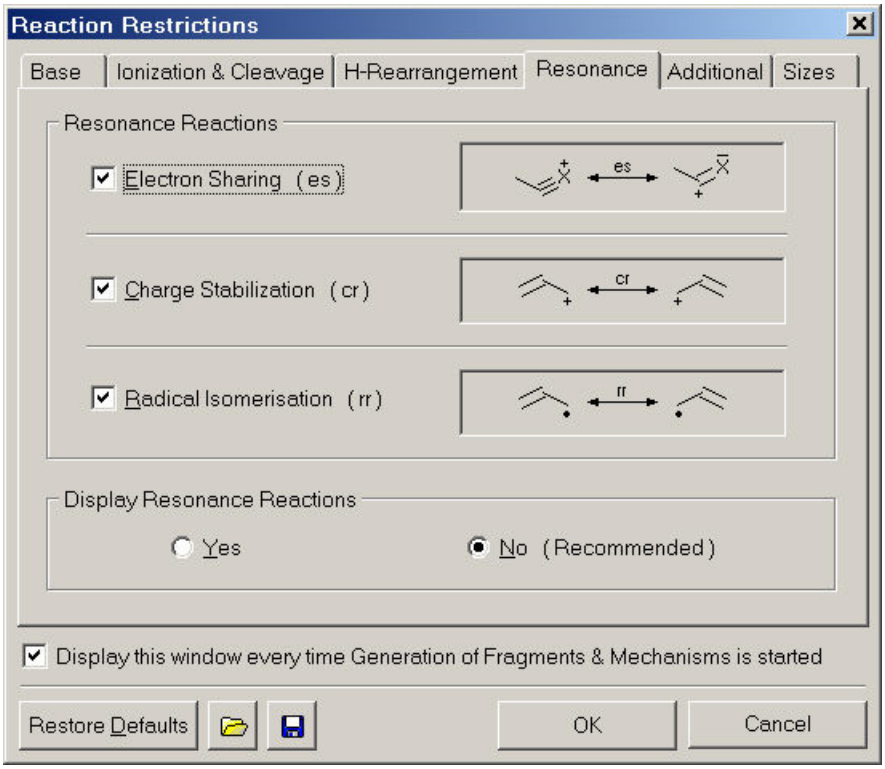


Figure 71. Reactions Restrictions dialog window showing the Resonance page

Additional Page

By default, cleavage on an aromatic ring is not activated. However, when you dealing with small aromatic compounds, activate bond cleavage on aromatic rings by clicking the **Cleavage** box in the **Allowed on Aromatic Systems** box in the Additional page in the Reaction Restrictions dialog window. See Figure 73. For example, when you generate fragments and mechanisms of phenol, the important fragment corresponding to the peak *m/z* 66 can be generated only if cleavage on aromatic systems has been activated.

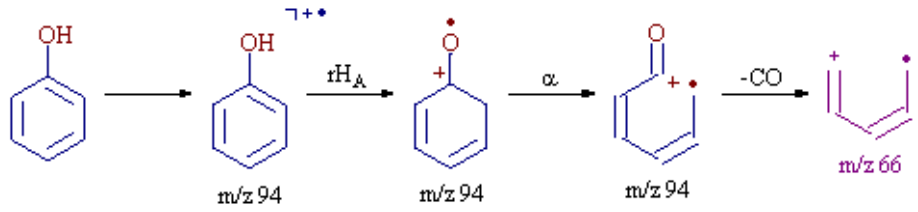


Figure 72. Structures of phenol, showing bond cleavage on an aromatic system

Cleavage on aromatic systems is deactivated by default. This is due to the large numbers of fragments can be generated from large aromatic compounds and the aromatic bond is a very strong bond.

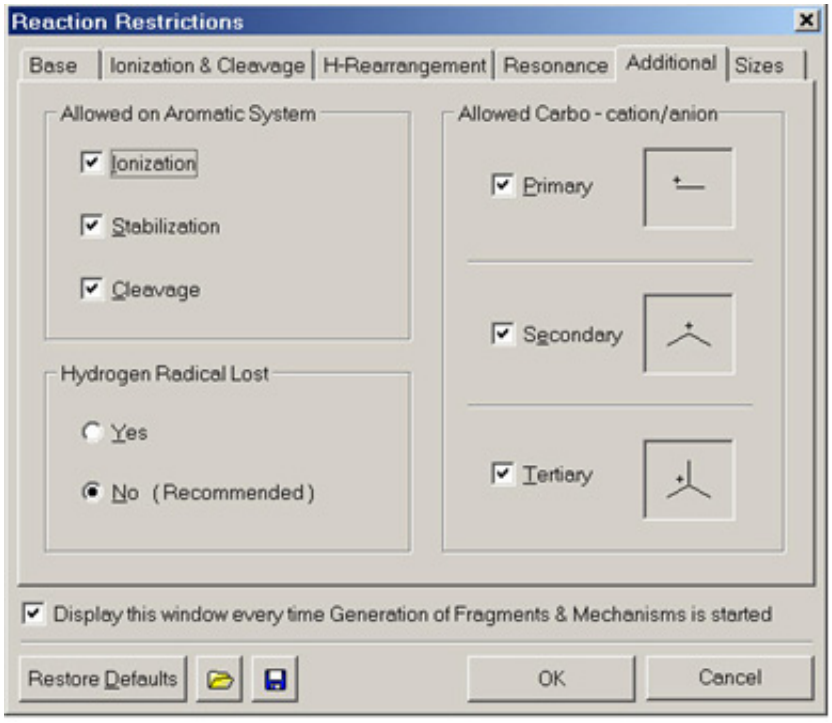


Figure 73. Reaction Restriction dialog window, showing the Additional page

Sizes Page

Use the Sizes page to limit the size and complexity of a reaction pathway generation. See Figure 74. The Reaction Steps Max Number box gives the number of cascaded fragment reactions. Increasing this number could exponentially increase the number of fragments produced for a given reaction path. Generally, keep this number small, and if additional fragments need to be generated, select individual fragments to use as starting points for additional reactions.

Large and structurally complicated molecules can produce a large number of reactions. Because generated reactions take up a large amount of memory, the number of temporarily generated reactions is limited. The system generates the fragments and mechanisms so that the fragments that are generated first are the most important. If the reactions limit is reached, the generation stops, and a message box appears as a reminder. Even if the generation is stopped, the fragments and mechanisms generated up to that point are displayed. Therefore, when a generation is stopped, the most important fragments have probably already been generated. However, if you are missing an important fragment, and you believe this is due to the generation being interrupted, increase the number of internal reactions in the Reactions Limit box.

The *m/z* precision of calculated fragments can be set in the Mass Settings dialog window. Choose **Options > Mass Settings**.

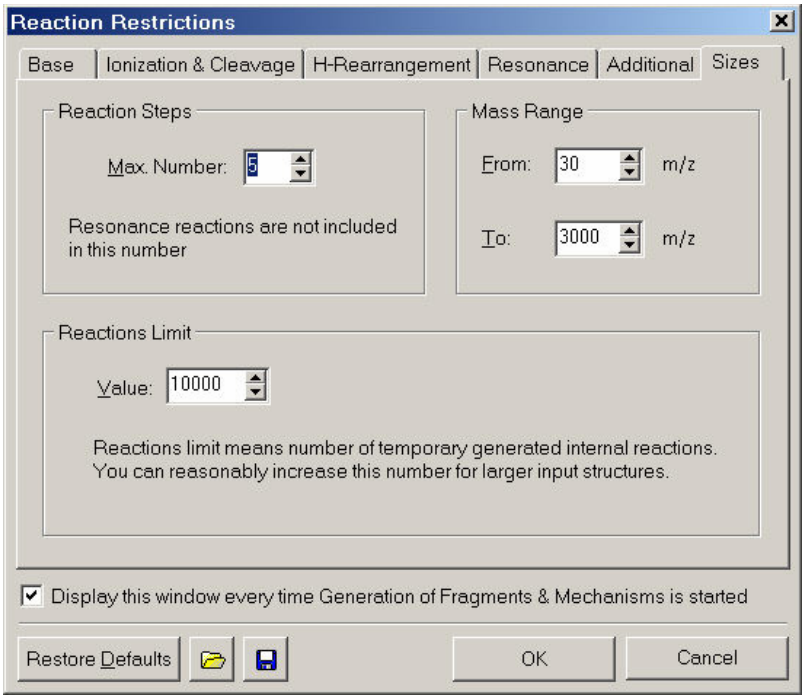


Figure 74. Reaction Restrictions dialog window showing the Sizes page

Generated Fragments Linked with Spectrum

Mass Frontier has the ability to link generated fragments with a mass spectrum. If you start a generation of fragments and mechanisms from the Database Manager, the generated fragments are automatically linked with peaks in a spectrum according to their *m/z* values to help explain peaks in spectrum. After a generation, highlighted (or explained) peaks are displayed in a different color (by default red) in the original mass spectrum. Selecting a highlighted peak reveals all the mechanisms leading to it.

If an unexplained peak is likely to be an isotopic peak of an explained peak, this is depicted in a third color (by default, green). Selecting such a peak reveals all the mechanisms leading to the inferred fragment, which can produce this isotopic profile.

Remember that the inability to predict energies and barriers in ionized molecules prevents the prediction of all the peaks in a mass spectrum. Thus, fragment predictability usually ranges between 50-90 percent. However, a prominent unexplained peak can be valuable for the interpretation or identification of an unknown. An unexplained peak can indicate a compound-specific mechanism that occurs in molecules with similar structural features or with a common substructure. There are a number of mechanisms that have only been observed in a specific group of compounds and cannot be applied generally when proposing fragmentation and rearrangement pathways.

If you suspect a compound-specific mechanism of fragment formation, verify your assumption by conducting a substructure search and then comparing the explained and unexplained peaks in the spectra retrieved by the substructure search.

Eliminating Generated Fragments Not Present in a Spectrum

The Fragments & Mechanisms module shows all *m/z* values that have been generated with given Reaction Restrictions settings. If the generated fragments are linked with a spectrum (that is, the fragments were generated from a still existing Database Manager window which contains the given spectrum), you can eliminate the *m/z* values that do not have corresponding peaks in the spectrum. In some cases, Mass Frontier generates a large number of theoretically possible fragments with a variety of *m/z* values. It is, therefore, useful to display only those fragments that can be linked with a spectrum (explained peaks).

To eliminate *m/z* values that cannot be linked with peaks in a given spectrum

- Click the **Show *m/z* Values for Explained Peaks Only**  button in the Fragments & Mechanisms window.

If you eliminate *m/z* values that do not correspond to a spectrum, these values are not permanently deleted. While the button is in the down position, Mass Frontier removes these values from the tab control or from the box, depending on the selected settings of the *m/z* selector (Mechanism Layout dialog window). Reset the button to its up position to restore the original state and list all generated *m/z* values.

Simulation of MS^n Experiments

With Mass Frontier, you can simulate fragmentation processes in MS^n experiments. See Figure 75. The fragments and mechanisms can be generated not only from neutral compounds, but also from ions. You can also select a product fragment (parent ion) in a Fragments & Mechanisms window and begin the generation there. An unlimited number of consecutive secondary ion decomposition reactions can be simulated.

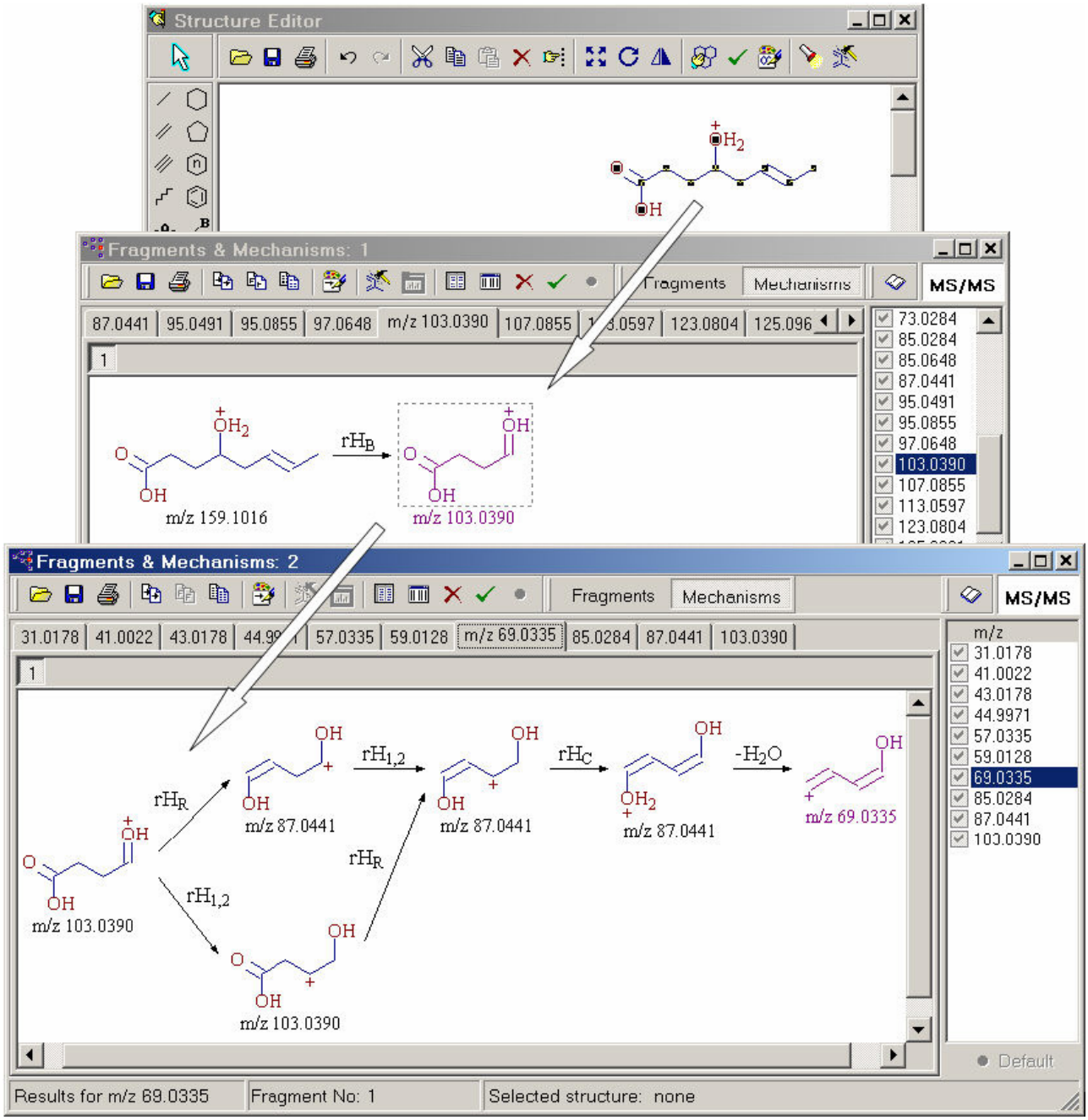


Figure 75. Fragments & Mechanisms window showing the simulation of an MS^n experiment

Unexplained Peaks

The fact that a peak cannot be explained by Mass Frontier because the corresponding fragment was probably formed by a compound specific mechanism, can be helpful in the identification of characteristic structural groupings that give rise to the peak. For example, the phthalates produce a characteristic ion with *m/z* 149, which is formed by a highly specific mechanism. The peak at *m/z* 149 can be easily recognized as a contaminant from elasticized polymers. Mass Frontier is not able to explain this peak because its corresponding fragment is formed by an unusual hydrogen rearrangement and cyclization, which are not supported. To distinguish between a randomly unexplained peak and a compound-specific peak, you need to find some examples in the library. Using a substructure search, you can retrieve compounds that contain a phthalate group as a common substructure. After the generation of fragments and mechanisms of the retrieved examples, the prominent peak corresponding to the phthalate group remains unexplained in the majority of cases. For example, a phthalate with a functional group at position 3, 4, 5, or 6 has its prominent peak shifted to higher masses by the mass of this functional group. Such an unexplained prominent peak, present in the spectra of structurally similar compounds, can be a strong indicator of a compound-specific fragmentation process. This information can help in the identification of a substructure under investigation.

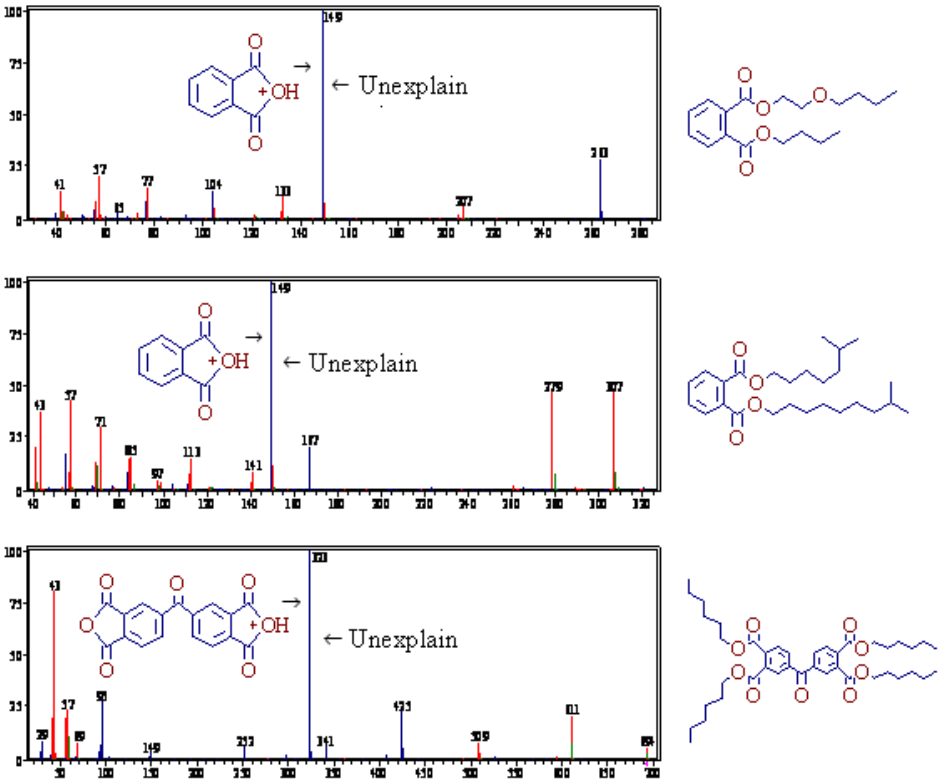


Figure 76. Spectra and structures of phthalate analogs

Too Many Proposed Fragments for a Peak

For a molecule that contains a larger number of locations for possible reaction initiation (heteroatoms and π-bonds), Mass Frontier can generate several theoretically possible ions with identical mass-to-charge ratios. Usually, these ions are isomers of the same fragment, but sometimes they are structurally different. If the program generates two or more different fragments with identical mass value, you need to choose the correct one. As before, a substructure search can provide the necessary information.

Figure 77 shows how to select the most likely fragment from several possible candidates. The spectrum of 2,3,3a,4,5,6-Hexahydro-10-methyl-1H-pyrazino[3,2,1-j,k]carbazole exhibits a base peak at *m/z* 198 (M-28). See Figure 77. For this peak, Mass Frontier selected twenty-two possible ions that are mostly isomers of two fragments.

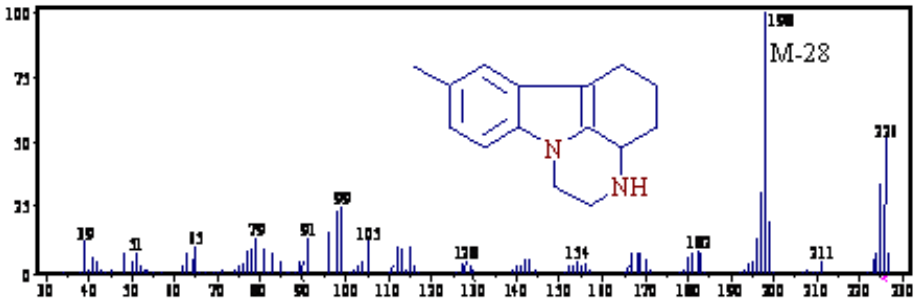


Figure 77. Spectrum of 2,3,3a,4,5,6-Hexahydro-10-methyl-1H-pyrazino[3,2,1-j,k]carbazole

The first fragment is formed by a retro-Diels-Alder reaction involving an α-bond on a carbazole group. The second fragment is the result of ethylene loss from a piperazine group. See Figure 78.

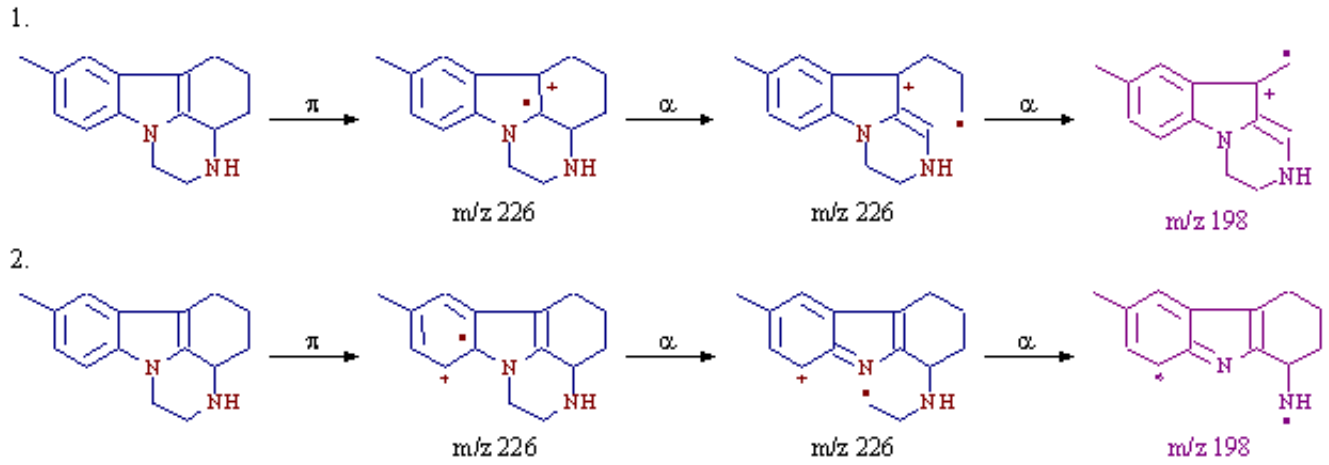


Figure 78. Two possible mechanisms of fragment formation

To determine the most likely fragmentation process for the formation of the fragment *m/z* 198 (M-28), begin a substructure search of the investigated structure, excluding the methyl group, to determine whether the spectra of the retrieved compounds also contain a base peak or a prominent peak at M-28. See Figure 79.

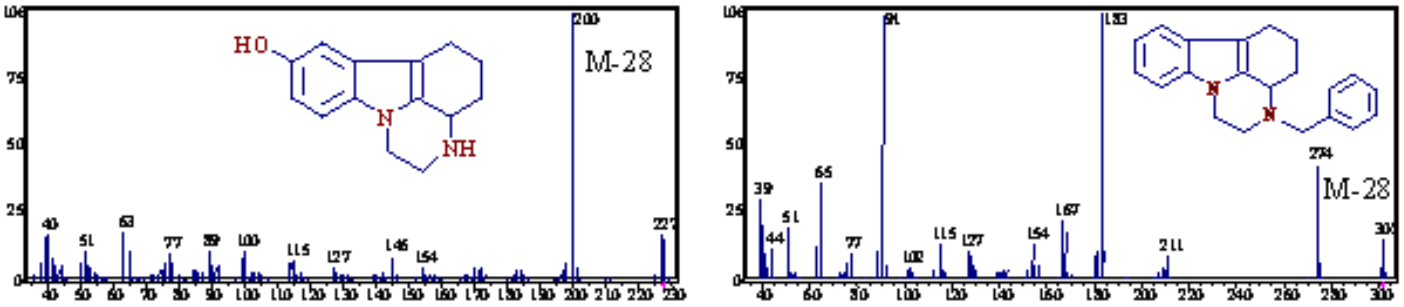


Figure 79. Spectra from a substructure search

When the peak M-28 is confirmed as the characteristic peak of this compound class, retrieve two different groups of spectra. See Figure 79. The first group of retrieved spectra shows compounds with a common substructure that are involved in the first mechanism. See the spectra on the left side of Figure 80. Similarly, the second group of spectra shows compounds with a common substructure that are involved in the second mechanism.

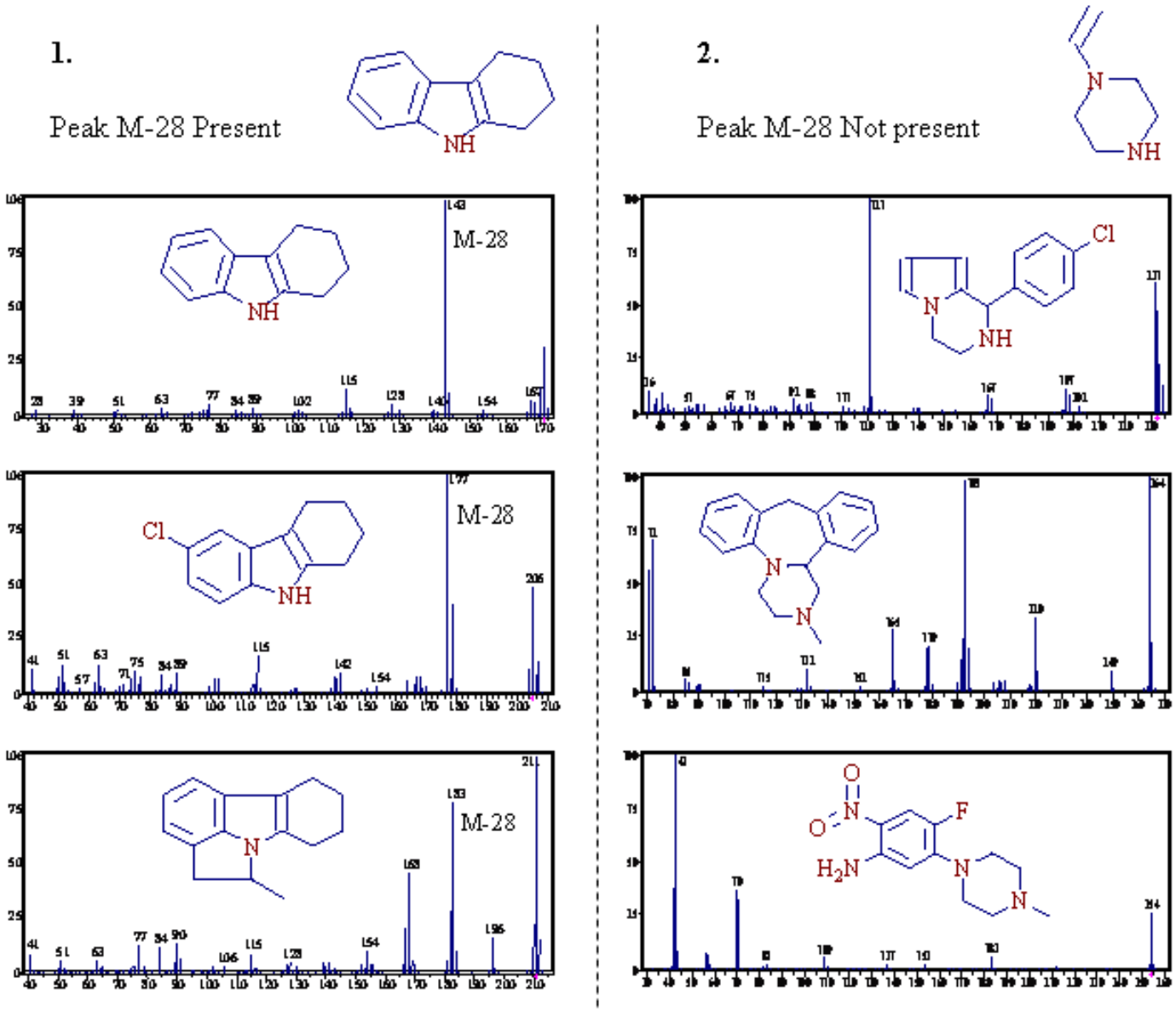


Figure 80. Two groups of spectra showing compounds with common substructures that are involved in two different mechanisms

Two separate substructure searches of 1,2,3,4-Tetrahydrocarbazole and Piperazine result in the retrieval of two groups. You can see which group yields a prominent peak at M-28.

These two groups of spectra prove that the peak *m/z* 198 in the spectra of 2,3,3a,4,5,6-Hexahydro-10-methyl-1H-pyrazino[3,2,1-j,k]carbazole is formed by a Retro-Diels-Alder reaction (Mechanism 1).

Bar Code Spectra

Mass Frontier can automatically predict possible fragments from a chemical structure using general rules and library mechanisms, along with primary determination of the structural plausibility of generated ions. The prediction of mass spectra is hindered by the difficulty of predicting thermochemical data, the thermodynamic stability of product ions, and reaction rates. However, generated fragments and their mass-to-charge ratio values can be used for creating bar code spectra. A bar code spectrum contains peaks at predicted mass-to-charge ratio values with identical (maximal) intensity. See Figure 81.

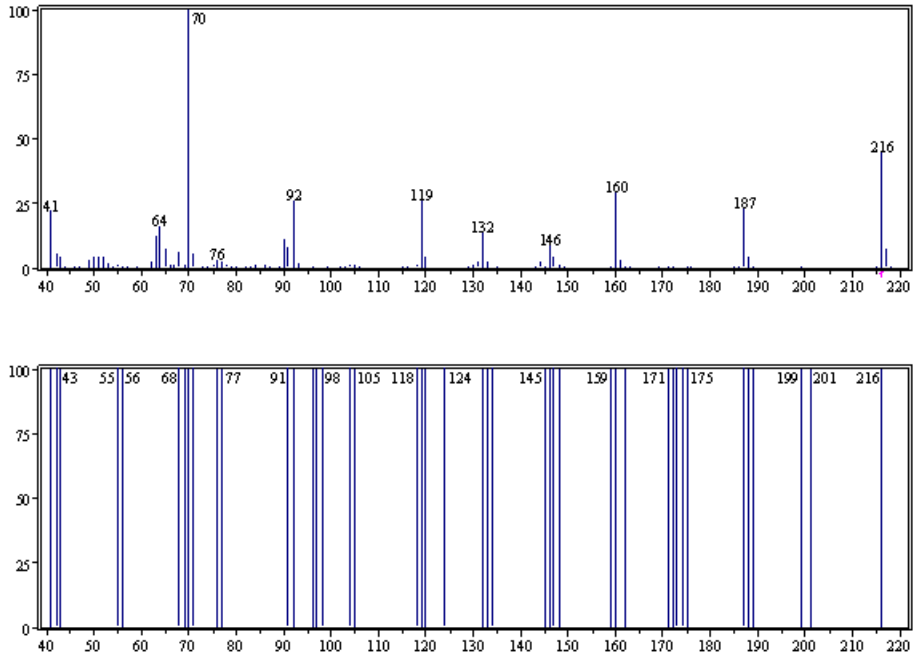


Figure 81. Experimental and Bar Code spectrum of Pyrrolidine[2,1-c]-2H,5H-1,4-benzodiazepin-2,5-dione,1,3-dihydro- (Electron Ionization mode)

With Mass Frontier, you can create bar code spectra from predicted fragments.

To create a bar code spectrum

Click the **Bar Code Spectrum** button in any Fragments & Mechanisms window.

The created bar code spectrum is placed in a Spectra Manager window. Bar code spectra are automatically linked with their original Fragments & Mechanisms windows. If you click any bar code peak in Spectra Manager,

the program displays corresponding fragments, along with its (their) formation mechanisms. This link remains in place as long as the corresponding Fragments & Mechanisms window exists.

Use bar code spectra in several strategies for investigating spectra-structure relationships. The primary purpose of generating bar code spectra is that they allow the possibility to identify spectral differences in structurally similar compound classes, for which mass spectra are not available. To study fragmentation dissimilarities between structurally related compounds, it can be easier and quicker to compare two or more bar code in Spectra Manager than to manually compare fragments and their mass-to-charge ratios between Fragments & Mechanisms windows.

For example, if you are interpreting an unknown spectrum and you have established two structural proposals, do the following:

1. Draw both structures separately in the Structure Editor.

2. Generate fragments for both structure to find out which structure belongs to the unknown spectrum.

3. Create bar code spectra, and place them in the same Spectra Manager window.

4. Compare the bar code spectra in the Compare Spectra page in Spectra Manager. In the Difference Spectrum box, you see the specific peaks that this pair of spectra do not have in common.

You can then compare these specific peaks with the unknown spectrum, and take a closer look at the fragmentation mechanisms of these peaks. If a specific peak is present in the unknown spectrum and the mechanism of formation seems to be plausible, select the most likely structure. The approach using bar codes, is superior to comparing explained peaks, because it can be applied to a large number of structural proposals simultaneously.

The following example demonstrates how to identify the correct structure for an unknown spectrum. Assume two structures are suggested: 1-Butanone, 1-phenyl- and 1-Propanone, 2-methyl,1-phenyl-. A comparison of their bar code spectra shows that the peak at *m/z* 120 is present in the unknown spectrum and only one of the bar code spectra. The formation mechanism for the peak *m/z* 120 reveals a classical sequence of McLafferty rearrangement and α-cleavage that is very common for 1-butanones or larger ketones. See Figure 82.

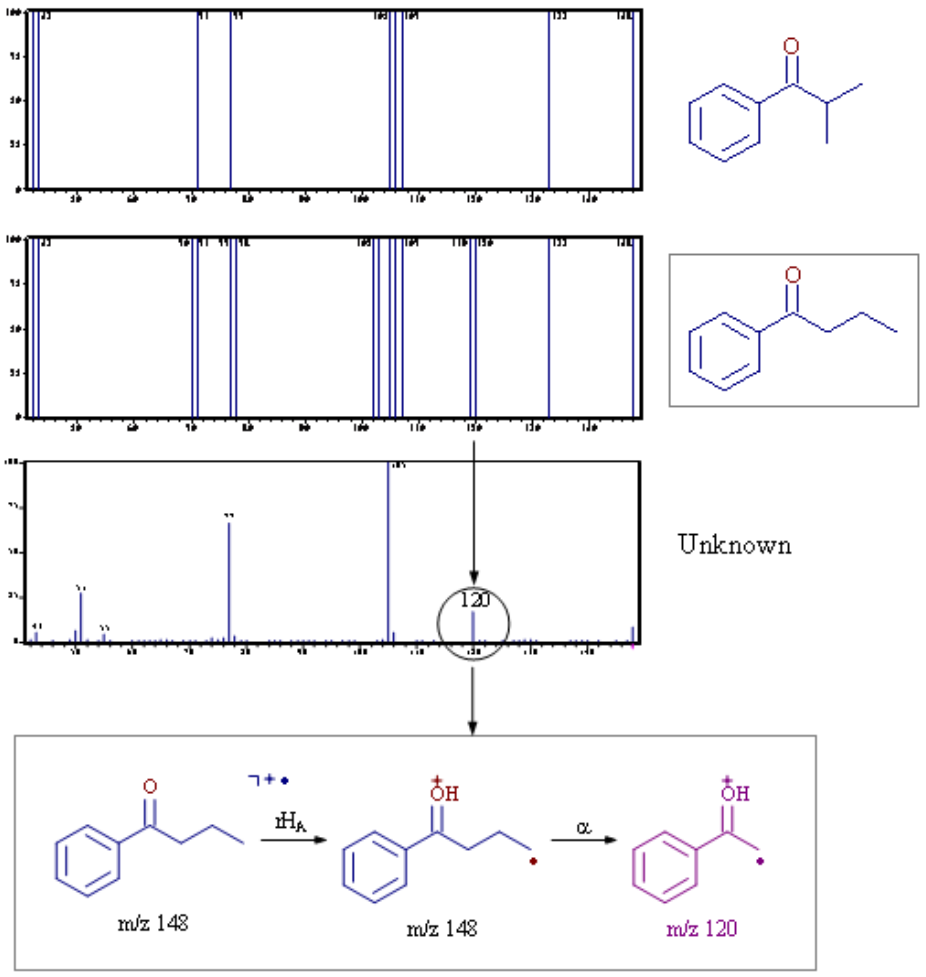


Figure 82. Postulating structures from isobaric possibilities using Bar Code spectra

Chapter 7 Fragmentation Library

Use the Fragmentation Library module to create and manage fragmentation mechanism databases. This module contains a graphical editor for entering fragmentation reactions which can be stored in a database, together with complementary information for the reaction. All fields of the database can be queried, for example, authors, ionization method, or mass analyzer. All library structures from the reactions are also fully searchable.

This module contains an expert system that automatically extracts a decomposition mechanism for each fragmentation reaction in the database and determines the compound class range that the mechanism can be applied to. Mass Frontier uses this expert system to apply database mechanisms to a user provided structure and automatically predicts the fragmentation reactions for a given compound.

The current version of Mass Frontier comes with almost 5000 fragmentation schemes that contain around 19000 reactions collected from mass spectrometry literature. The collected mechanisms are stored in the HighChem database that automatically appears when the Fragmentation Library module is opened.

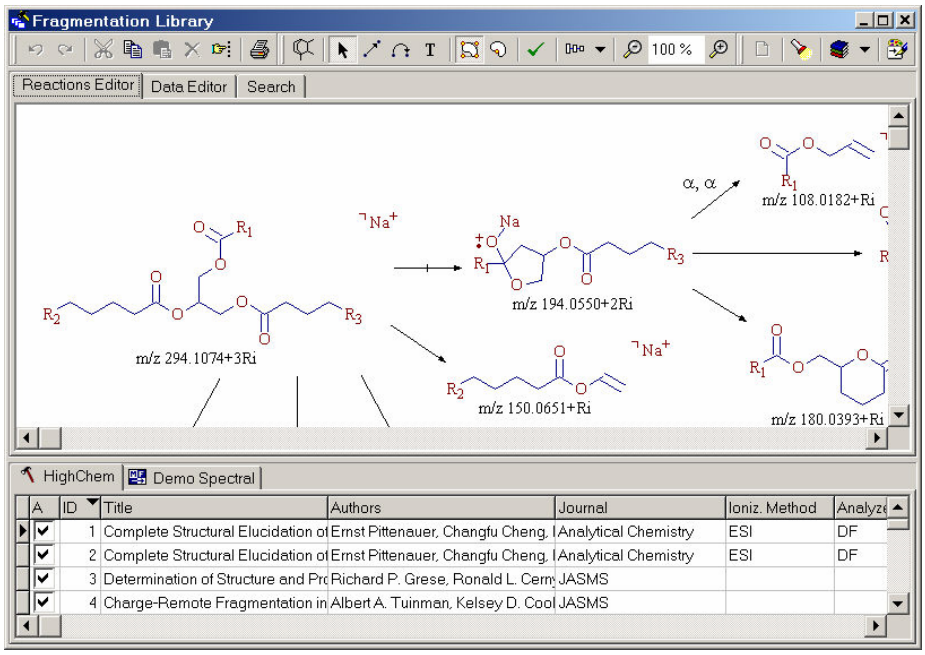


Figure 83. Fragmentation Library window

This chapter contains the following sections:

- Fragmentation Library Toolbar
- Drawing Fragmentation Reactions
- Active Record
- Additional Information
- Saving Records
- Mechanism Extraction
- Reaction Symbols
- Using Library Reactions in Fragmentation Prediction
- Search Utilities
- HighChem Fragmentation Library

Fragmentation Library Toolbar

Figure 84 depicts the Fragmentation Library toolbar.

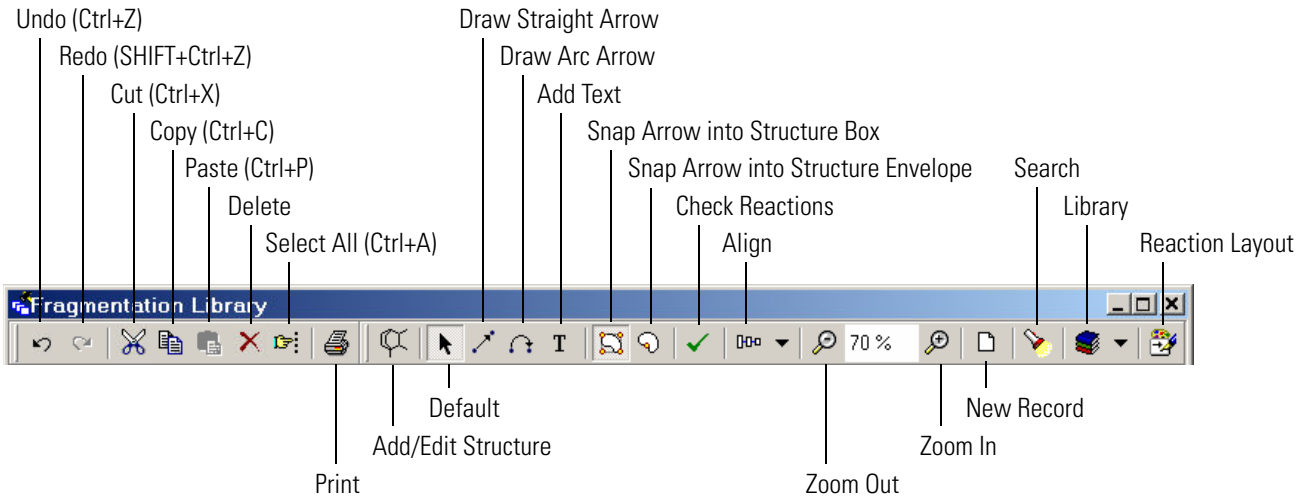


Figure 84. Fragmentation Library toolbar

Drawing Fragmentation Reactions

The Fragmentation Library module includes a graphical editor for entering and editing fragmentation reactions. To open the Reaction Editor, select the Reaction Editor tab in the top-left corner of the Fragmentation Library module.

The **Reaction Editor** buttons on the module buttons bar can be detached and placed anywhere in the program. Right-click to access some drawing actions.

To add a new structure

To add a structure, use one of the following methods:

- Double-click on the drawing canvas.
- Click the **Add/Edit Structure** button on the button bar.
- Right-click on a blank canvas and choose **New Structure.**

The Structure Editor appears, where you can draw your fragment. This Structure Editor is a dialog window, which means you cannot open another window unless it is first closed. To place your fragment on the canvas, click **OK**. The fragment can be moved anywhere on the canvas by dragging the structure with the mouse.

To edit an existing structure

To add an existing structure, use one of the following methods:

- Double-click a fragment.
- Select the structure and click the **Add/Edit Structur**e button.
- Point the mouse cursor at a structure, right-click and choose **Edit Structure**. The Structure Editor appears.
- To confirm your changes, click **OK**. To reject the changes, click **Cancel**.

To add a straight reaction arrow

To add a straight reaction arrow, use one of the following methods:

- Click the **Draw Straight Arrow** button and then click the cursor where you want to place the arrow.
- Point the cursor at a structure, right-click, and choose **New Arrow**.

In order for the expert system to correctly extract mechanisms, the structures in a reaction scheme must be properly connected by arrows. The system considers stand-alone structures and disconnected arrows as errors and ignores them. If a structure is connected with at least one arrow, the selection squares are displayed in green. If a structure is not connected, the selection squares are displayed in yellow. The same color-coding applies to arrows. If an arrow is connected with a structure, this arrow end is shown in green. If the arrow end is not connected, it appears in yellow.

To help you properly connect the drawing objects, small red selection squares appear around the structure or arrow when you move them, if the object is close enough to connect.

Active Record

Mass Frontier applies the mechanisms of fragmentation reactions to a structure you provide and automatically predicts fragmentation reactions for the given compound. However, some mechanisms might be erroneous or too general and the expert system might apply them even though they are not desired. To exclude a particular record from the generation of fragments, click to clear **Active Record**. See Figure 85.

This option is especially useful for stopping the use of library reactions consisting of small fragments whose mechanisms can coincidently fit to virtually any compound. This can cause the generation of an extremely large number of fragments, which dramatically slows the generation process and makes reviewing the predicted fragments difficult.

Record activity can be changed in the Data Editor's Active Record box or in column A of the record grid. These two boxes are linked and always work together.

Note Changes to an Active Record box are immediately updated in the database and do not need to be saved.

To activate or deactivate all the selected records, right-click the records and choose **Activate All Selected Records** or **Deactivate All Selected Records**.

To sort active and inactive records in a library grid, click in the column header. To restore the ID number ordering, click the ID header tab in the library grid. It is possible to sort any column by clicking in its grid column header. To display the reverse order, click again in the header.

By default, the Fragmentation Library window shows all the library records in a corresponding library grid. If you need to list a subset of library records, you can hide selected records. Right-click the selected records and select **Hide Selected Records**. To restore the complete list of library records, click the **Reset** button in the Search tab.

Note Hidden records are not deleted and are fully searchable.

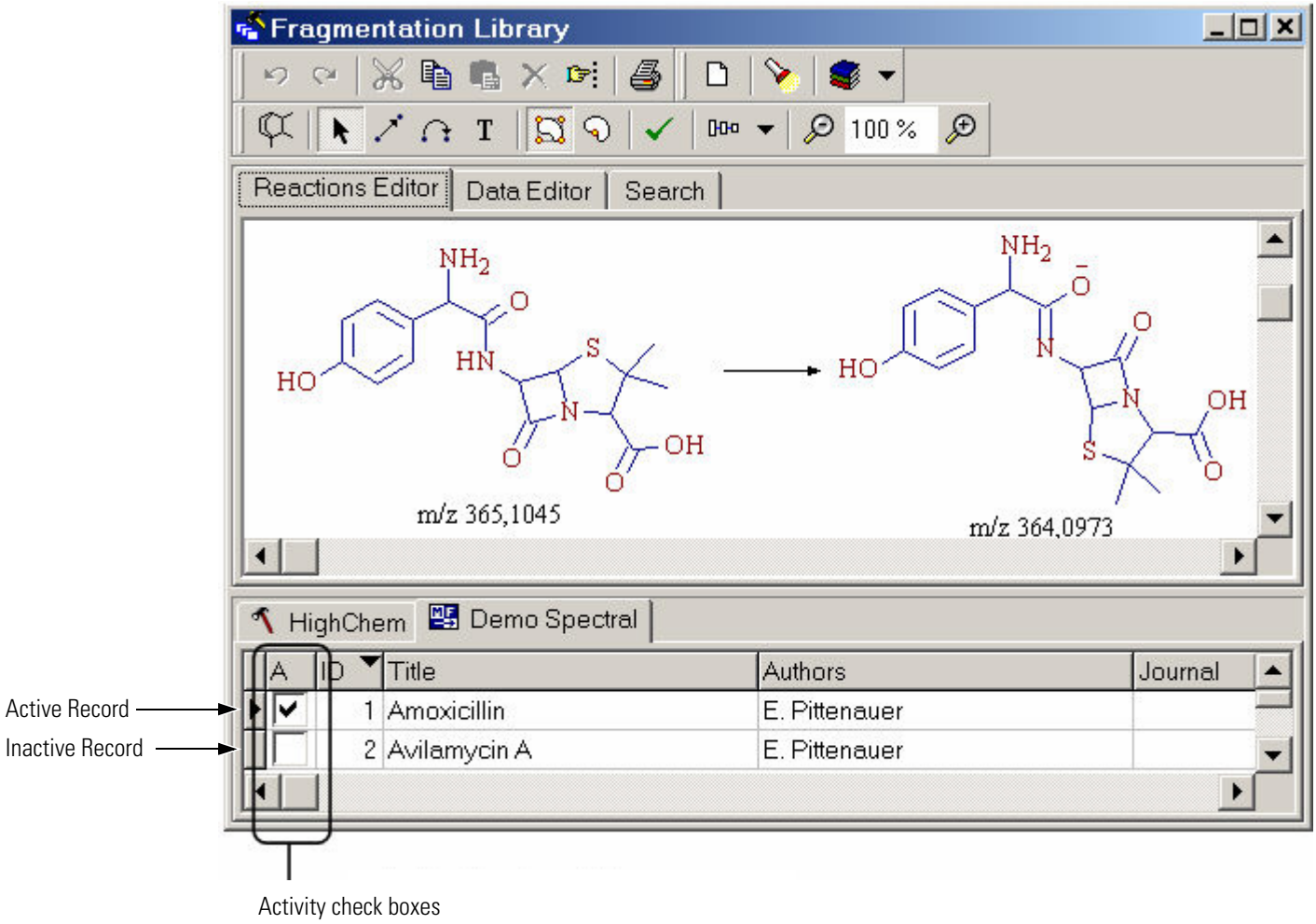


Figure 85. Fragmentation Library window showing active and inactive records

Additional Information

In addition to a fragmentation scheme, text data can also be maintained in a particular database record. To enter or edit text data, select the **Data Editor** tab in the top-left corner of the Fragmentation Library module. There are four available fields:

- Record activity
- Source of fragmentation scheme (journal, author, and so on)
- Experimental technique and instrument info
- Comment

Data Editor fields can be edited at any time, even for an existing record. To keep the changes, the record must be saved. Changes to the Active Record box are immediately updated in the database and do not need to be saved.

Note The Active Record field (labeled A) in the Data Editor is linked with the box in column A in the record grid in the lower half of the window. See Figure 86. Both of these controls always work together.

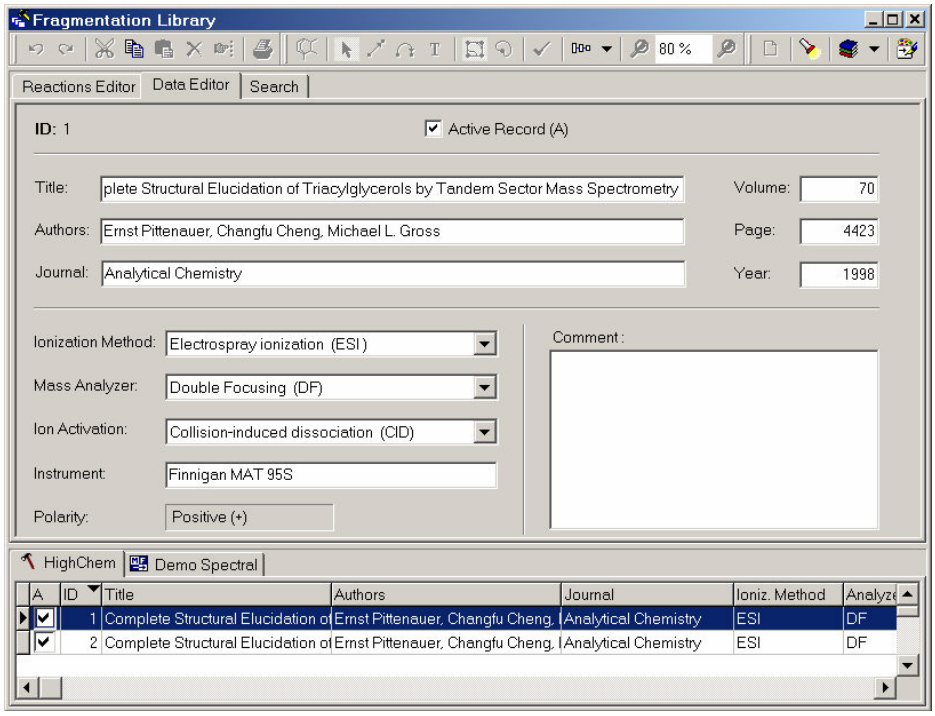


Figure 86. Data Editor page of the Fragmentation Library module

Saving Records

To make a change permanent in Fragmentation Library, it must be saved.

To save changes in Reaction Editor or Data Editor

1. Click any in the record grid.
2. Click the line up or line down key.
3. Click the **Library** button and choose **Save Record To Library**.

When saving a record, the software performs the following sequential actions:

1. The fragmentation scheme in Reaction Editor is checked for formal correctness. If the software detects an error, you are prompted to either save the scheme as it is, or return to the record to correct the problem.
2. A fragmentation mechanism for every single unimolecular reaction in the fragmentation scheme is extracted using advanced algorithms. This process can take several seconds for complex reactions.
3. The reaction scheme, additional text, numerical data, and the extracted mechanisms are stored in the database.

If the check for formal correctness encounters an error, you are prompted to either save the record as it is, or to go back and make appropriate changes. If you save a scheme with just a single error, the reaction mechanisms are not extracted and the entire record is ignored in the prediction of fragments and mechanism feature. You can return to an erroneous record at any time to fix such a problem and enable the scheme for fragmentation prediction.

Reaction symbols appear only if a scheme is correctly saved.

Mechanism Extraction

At the core of the automated prediction of fragmentation and rearrangement pathways is a computer-based system for the extraction of unimolecular decomposition mechanisms from reactions you provide. This complex software system decodes the underlying principle of fragmentation mechanisms from reaction drawings and builds a knowledge base of fragmentation events. The system works automatically and replaces the need for manual input of atom-atom correspondence in precursor and product ion pairs. A computer learning procedure has been developed to allow the processing of specific fragmentation details. Similarly, in experimental mechanistic studies, labeled or generic structures can be used to direct the desired dissociation route. By means of deuteria or substituents (R) participation, the decoding algorithm unambiguously extracts the underlying mechanism.

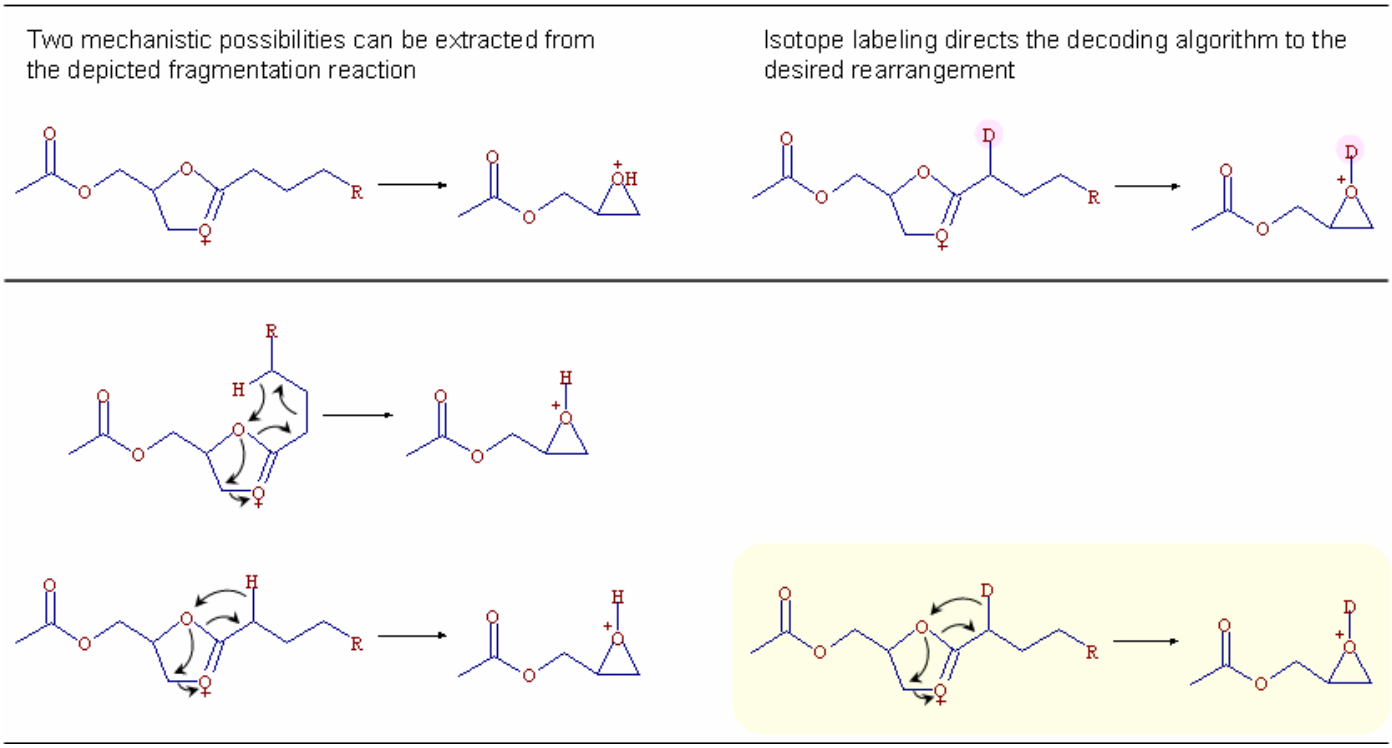


Figure 87. Virtual isotope labeling can be used to specify the rearrangement course

Because many reactions stored in a fragmentation library follow general fragmentation rules that can be predicted by the use of preprogrammed unimolecular reactions, the library distinguishes between class specific mechanisms and general fragmentation reactions. After a reaction has been entered into the database, the system attempts to identify a general

fragmentation rule and assigns the relevant reaction symbol above the arrow. See "Fragmentation, Rearrangement and Resonance Reactions" on page 110.

In terms of fragmentation prediction, the determination of the preferred ionization site is as important as detailed knowledge of the fragmentation mechanism. The system supports virtual generation of charged molecules based on library ionization reactions using the exact location of the positive or negative charge, or the unspecified charge location symbol that can be assigned to a structure drawing.

Reaction Symbols

Mass Frontier's expert system automatically extracts fragmentation mechanisms from a reaction drawing after a scheme has been saved. If a reaction follows one of the preprogrammed general fragmentation rules, the arrow is captioned with the particular rule abbreviation. See "Fragmentation, Rearrangement and Resonance Reactions" on page 110. The abbreviations for general fragmentation rules are found in the **Options > Reaction Mechanism Overview** main menu.

Even if a reaction is formally correct, it might not be possible to derive a reaction mechanism from a drawing you provide. This can occur if you enter an unfeasible fragmentation mechanism or because the unimolecular reaction is incomprehensible to the expert system. If a mechanism cannot be extracted, the reaction arrow has a cross through it. See Figure 88. In this case, the mechanism is reduced to the exact precursor and product structures and only the identical neutral or ionic precursor is matched with the structure you provided in the fragmentation prediction process.

The software sometimes decodes a mechanism from a drawn reaction but the atom-matching procedure is unable to find the corresponding atom counterparts on both sides of the reaction leading to partially recognized mechanisms. When this occurs, the reaction arrow appears with a small line through it. This kind of reaction can be used only for fragment prediction for some input structures according to the fragmentation algorithms decision. A partially recognized reaction mechanism might not be selected for fragments prediction, even when the input fragment looks similar to the precursor in the library reaction.

To overcome such a problem with unrecognized or partially recognized mechanisms, try to decompose complex one-step mechanisms into a number of reaction steps and then save these in a fragmentation library.

Note Reaction symbols for an unrecognized or partially recognized reaction do not appear immediately the reaction is drawn, but only after the record has been correctly saved in a library.

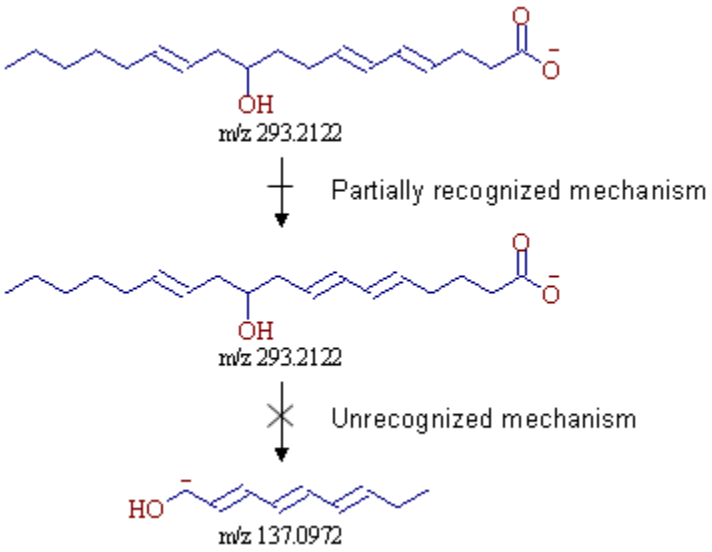


Figure 88. Reaction symbols used for partially recognized and unrecognized mechanisms

Using Library Reactions in Fragmentation Prediction

The primary purpose of library reactions is to create a database of fragmentation and rearrangement mechanisms that can be applied to structures you supply to predict the decomposition pathways occurring in a mass spectrometer. The library mechanisms are a significant extension to the general fragmentation rules which might not cover all the complex processes for a broad spectrum of ionization and ion activation techniques. The advantage of library mechanisms is the flexibility they give to alter predicted fragmentation pathways and entering highly specific mechanisms that apply for a limited class of compounds. Because the current library contains around 19000 reactions, the fragmentation predictability is much higher than if only general rules apply. This is especially true for low-energetic experiments such as ESI or APCI that often yield complicated skeletal rearrangements and unusual ring closures.

Any recognized and active mechanism that has been saved into a library serves as a knowledge base for the prediction of fragmentation pathways. After you have drawn and saved a reaction into any fragmentation library, use the mechanism template for prediction fragmentation pathways for any structure that the derived mechanism can be applied to. See Figure 89. The range of applicability depends on many factors, but structurally similar compounds with a common ring scaffold usually exhibit identical fragmentation mechanisms.

Note Fragment stability and general ion energetics depend on many thermodynamic parameters and even a slight structural dissimilarity between two molecules can result in large differences in the course of fragmentation pathways. For example, two identical structures with a simple hydroxy group difference can occasionally exhibit completely different spectra. Thus, the derived fragmentation analogy based on library reactions might not always reflect the real fragmentation events for structural analogues.

If a fragmentation reaction was predicted using a library reaction, double-click the Lib arrow caption to see the corresponding mechanism. Both template and generated fragmentation mechanisms are displayed in red in the Fragmentation Library window and the Fragments & Mechanisms window.

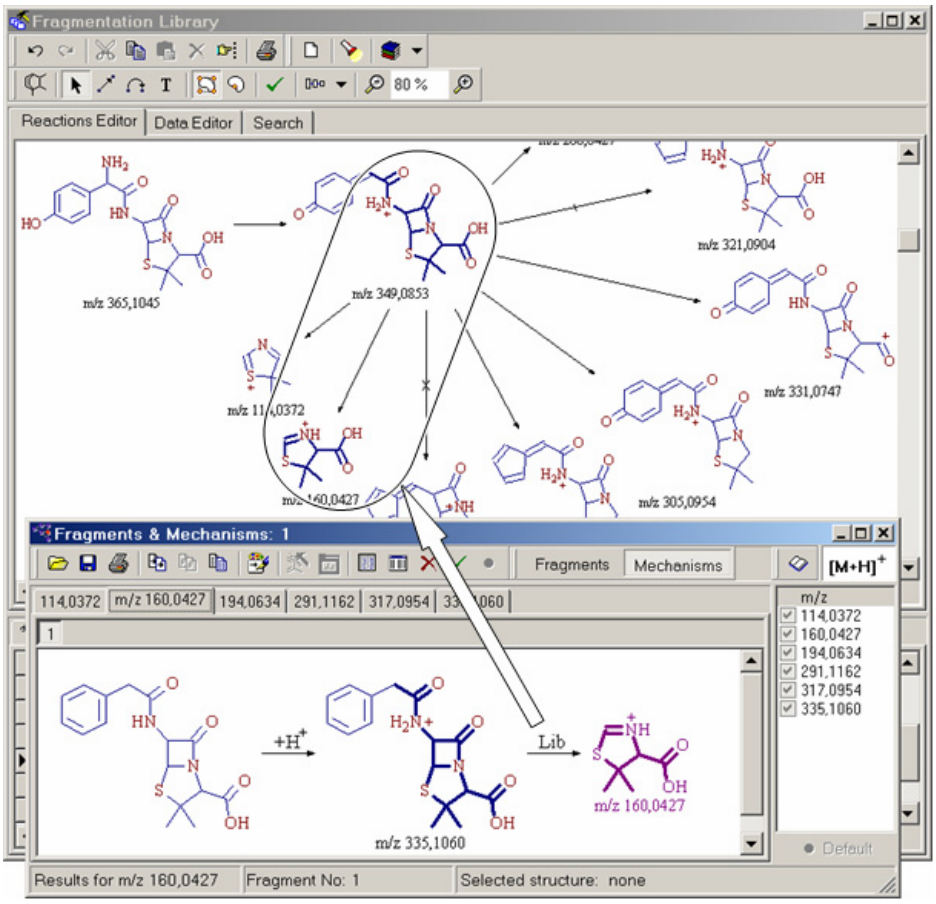


Figure 89. Fragments & Mechanisms window showing predicted rearrangement reaction based on mechanism template stored in Fragmentation Library window

Search Utilities

Fragmentation libraries are fully searchable by various criteria. Search criteria can be combined to narrow your search. All searchable items are found in the Fragmentation Library window in the Search page. To conduct a search, type the desired query in relevant search box or select one of the predefined options and click the **Search** button. The search results are listed in the grid box in the lower part of the window.

Note To restore the complete list of library entries in the grid box after a search has been performed, click the **Reset** button in the Search page.

Use a substructure search to select all the reactions in a library that match all the reaction precursor or product structures with your query substructure. The search results display the query substructure in red in the library structure, in the same manner as for the Substructure Search procedure in the Database Manager.

Because fragmentation libraries largely contain ionic structures that can undergo resonance reactions, Mass Frontier offers a resonance substructure search. This feature ensures the correct retrieval of all resonance structures, even if the query structure is in a different resonance form than the library structure. This feature works fully automatically, so you do not need to be concerned about the particular resonance state of the ionic structures. However, be aware of this functionality when reviewing search results, as the query and library structures in positive search results might appear to be different if the resonance reaction are possible. In this case, do not consider this difference as an error, but rather be aware that there might be complex resonance variations.

It is possible to ignore charges and radicals in substructure searches if their location is ambiguous. When searching structures with an unspecified charge location or substituents, be sure to review the search rules that apply. See "Substructure Search Rules" on page 85.

Note It is possible to search in one library at a time.

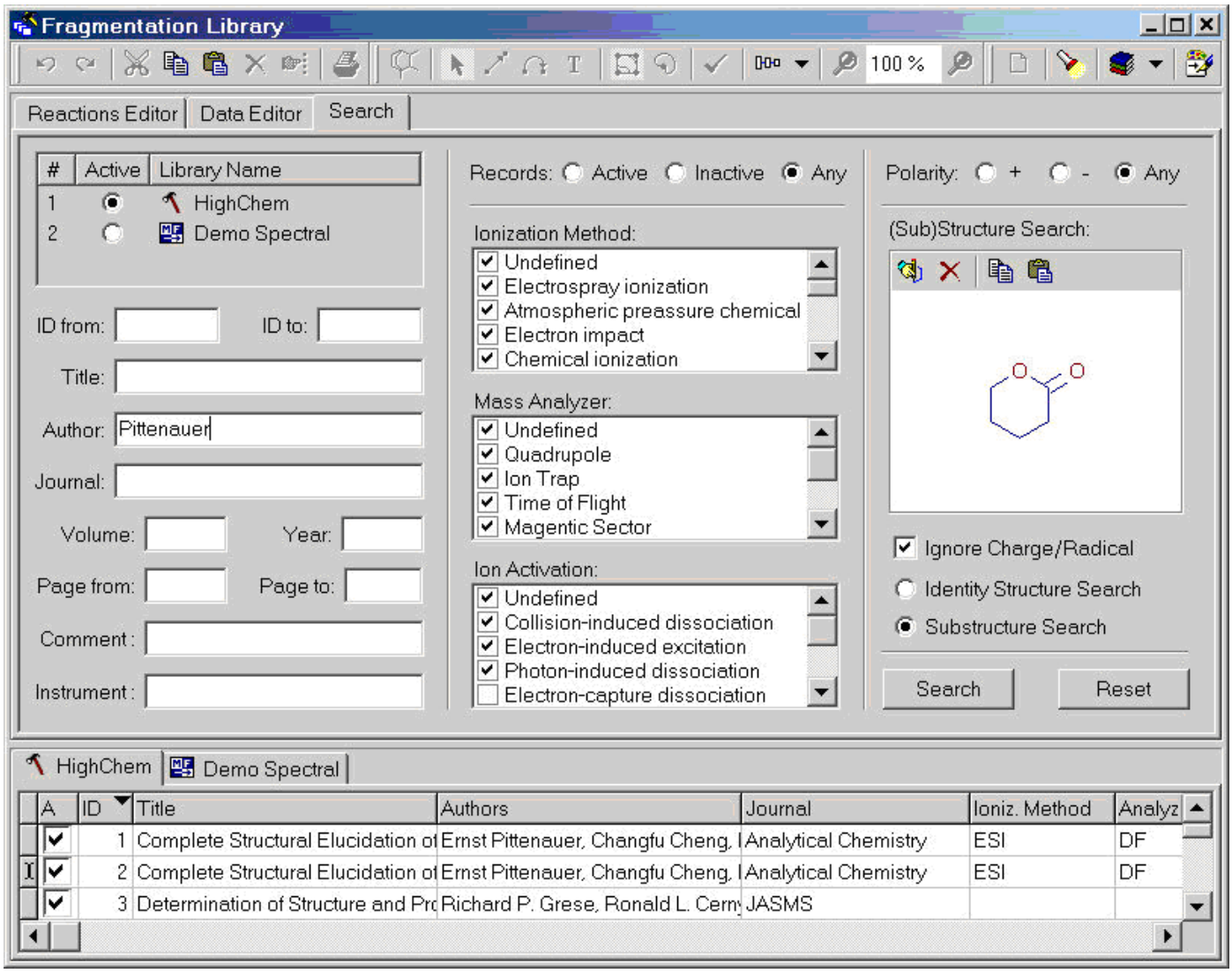


Figure 90. Search Page of the Fragmentation Library window

HighChem Fragmentation Library

HighChem has been collating fragmentation mechanisms published in all the available printed media dedicated to mass spectrometry for a number of years and these have been entered into the computer system. Each reaction, along with the chemical structures, has been manually drawn in Reaction Editor and saved in the HighChem Fragmentation Library, which currently contains around 100000 individual reactions. Fragmentation pathways are accompanied with complementary information such as the title, authors and source of the information. This library collection serves as a knowledge base for the prediction of fragmentation pathways from user provided structures.

To ensure high-quality data, fragmentation mechanisms have been evaluated in two stages: manual and automatic. The manual evaluation included accuracy and plausibility assessments of reaction mechanisms and consistency checking between fragment masses and peak *m/z* values, if the spectrum was available. The automatic evaluation includes simple element, charge, and radical consistency checks on both sides of the reaction, in addition to newly developed algorithms for complex electron mapping that has revealed formally erroneous mechanisms. Numerous problems and errors regarding mechanisms were uncovered by both stages and either appropriate corrections were made, or these mechanisms were excluded from the library.

Chapter 8 Fragments Comparator

Use the Fragments Comparator module to process and compare fragments generated in Fragments & Mechanisms windows. This module supports fragments generated using any ionization method or that originate from different structures. Using fragment marks in the Fragments & Mechanisms window, you can export only a selected set of fragments into the Fragments Comparator module. The fragments are organized in columns where each column represents a set of fragments provided by an associated Fragments & Mechanisms window. The Fragments Comparator module can hold an almost unlimited number of fragment sets, limited only by system resources.

The Fragments Comparator is an integral part of the Fragments & Mechanisms module. If you double-click any fragment in the Fragments Comparator module, the associated mechanism appears in the Fragments & Mechanisms window.

Note The Fragments Comparator is able to recall only mechanisms that are present in an open Fragments & Mechanisms window. If you close the associated Fragments & Mechanisms window, the link is lost.

The comparison feature is especially useful when analyzing the fragmentation products of structurally related compounds. Common fragments point toward a common substructure in terms of fragmentation. The most interesting are however the fragment differences. They can indicate fragments along with corresponding peaks in a spectrum that are characteristic for the distinction of structural details. Predicted *m/z* values of fragments that are different for structurally related compounds will attract your attention when examining spectral differences in the Database Manager in the Compare Tab.

The Fragments Comparator window consists of Table and Structures views. The Table view lists the numerical *m/z* values of fragments and Structures view displays structural drawings of possible fragments. Because the Fragments & Mechanisms module can generate several isobaric isomers for a single *m/z* value, the Structures view imports only the first fragment for each generated *m/z* value. Fragments are imported into these two views simultaneously but the information is managed independently. If a column in one view is moved or deleted, this action does not affect the other view.

In the Table view, you can select a column or a part of a column, and copy the data. The *m/z* values can then be pasted into Excel or any spreadsheet program.

In the Structure view, the cell can be resized by using the Cell Size track bar.

Columns can be moved in both views. The columns for imported fragments (left part) can be deleted.

Note The Fragments Comparator can display structures of fragments only if the associated Fragments & Mechanisms window is still open. If you close the associated Fragments & Mechanisms window, the corresponding column of fragments in the Structures tab are removed.

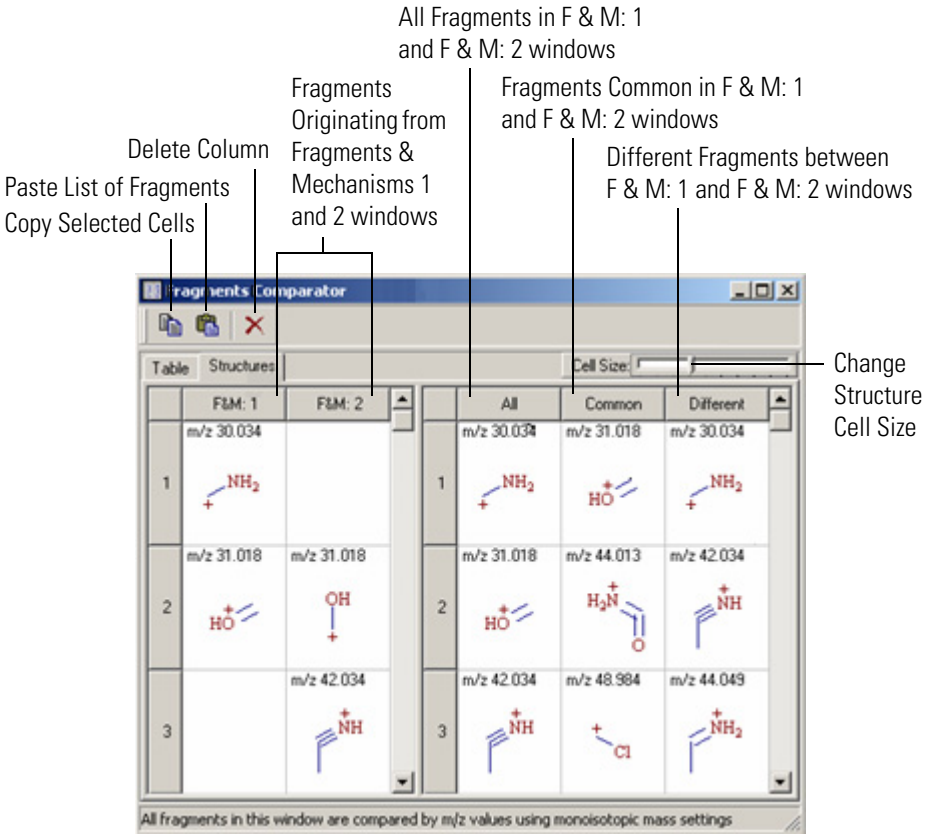


Figure 91. The Fragments Comparator window showing Structure tab

Both the Table and Structures views in the Fragments Comparator window are divided into two parts. The left part contains columns of imported fragments for each Fragments & Mechanisms window. The right part contains three columns that show three types of comparison results. The

first result column shows all available fragments (logical OR), the second column shows all common fragments (AND), and third column shows different fragments (NAND).

Note All fragments are compared by *m/z* values using resolution mass settings. Choose **Options > Mass Settings**. The fragments are usually predicted in several isomeric forms, therefore a structural comparison would not be reasonable. Because the fragments are compared by *m/z* values the calculated precision, defined in the mass settings, has a significant influence on the comparison results.

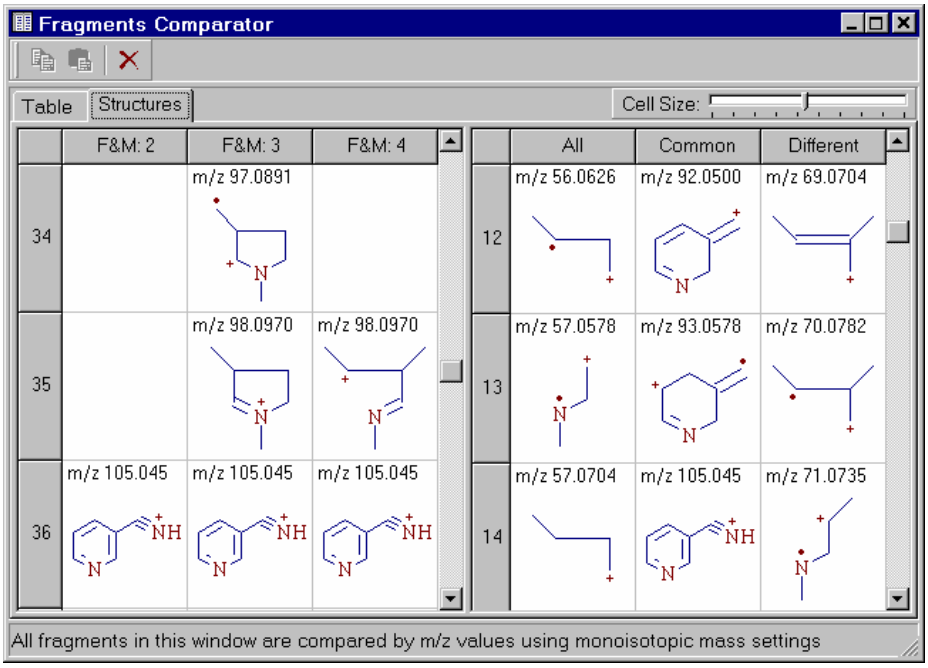


Figure 92. The Fragments Comparator window

Chapter 9 Mass Spectra Classification

Computer methods of analyzing mass spectral data center on three fundamental methodologies: library search techniques, expert system procedures, and classification methods. Classification is a powerful enhancement of library search and fragmentation prediction methods. The computer-oriented methods available in Mass Frontier complement each other, but are based on different principles. This provides possibilities for creating alternative strategies and for responsible data interpretation.

In contrast to statistical software packages, with Mass Frontier you can directly apply classification methods to mass spectral data. All preprocessing and data transformation procedures are carried out automatically which ensures the classification methods are used correctly even by those with no knowledge of multivariate statistics.

This chapter contains the following sections:

Spectra Classification

The primary goal of spectra classification is to find correlations between the properties of compounds and their mass spectra. Because physical and chemical properties and biological activities of chemical compounds are, to a large extent, functions of molecular structure, the results of classification analysis reflect structural features that are determined by fragmentation ions appearing in a mass spectrum. The advantage of classification methods is that detailed knowledge of the complex spectra-structure relationship is not required to get satisfactory results. Classification strategy in Mass Frontier is based on a user-friendly, graphic presentation of the results, which you can view on the screen.

Mass Frontier contains three classification methods: Principal Component Analysis (PCA), Fuzzy Clustering (FC), and Self-Organizing Maps (SOM), which are a special class of Neural Networks. These methods are based on different principles and enable you to explore complex data from various perspectives. PCA uses multivariate statistics, fuzzy clustering assigns data to clusters, and SOM is based on competitive learning.

In the multivariate statistic, each spectrum can be considered as a single point in an *n*-dimensional space, with the intensities being the coordinates of this point. A dimension (axis) of that space represents a mass-to-charge ratio (*m/z)* of the considered peak. Therefore, the dimensionality is determined by the *m/z* value of the last peak in the spectrum. For example, the EI spectrum of hydrogen exhibits two peaks at *m/z* =1 (intensity 2%) and *m/z* = 2 (100%). This spectrum can be viewed as a point in a two-dimensional space with the coordinates 2, 100. In reality, spectra have a higher dimensionality than two. If the dimensionality is too high, or several coordinates are equal to zero (usually a mass spectrum does not have peaks at every *m/z* value), the classification methods might not provide the required results. Therefore, a reduction of dimensionality is carried out either before a spectrum is placed in *n*-dimensional space, or during the classification process.

The basic hypothesis of multivariate statistical methods is the assumption that the distance between points (spectra) in an *n*-dimensional space is related to a relevant property of the compounds which represent these points. If the points are close enough to form a cluster or a separated region, assume that the compounds that correspond to these points exhibit common or similar properties. To ensure the results of the classification methods have statistical significance, place a large amount of spectra (usually one or more groups, each with 10 - 1000 spectra) in the same *n*-dimensional space. Then apply multivariate statistical methods, with various parameters, to evaluate these points (spectra). The objective of a classification process is to separate these points (spectra) into two or more classes according to the desired structural or other properties.

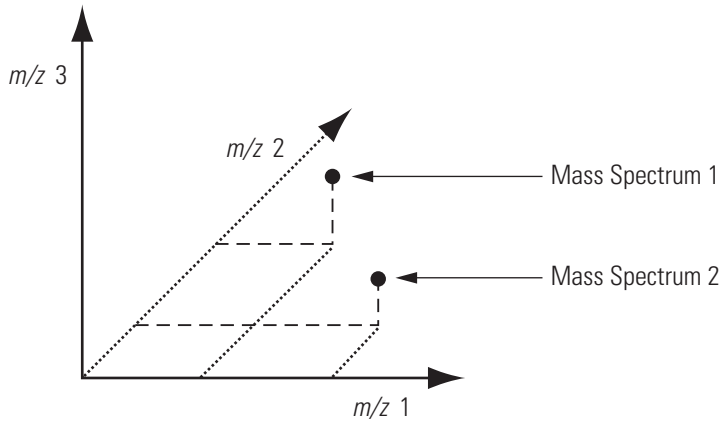


Figure 93. Representation of two spectra as points in a 3-dimensional space

Principal Component Analysis (PCA)

Mass Frontier offers a classification method called principal component analysis (PCA). The central idea of principal component analysis is to reduce the dimensionality of a data set in which there are a large number of interrelated (that is, correlated) variables, while retaining as much as possible of the variation present in the data set. In the case of mass spectrometry, the data set consists of the mass spectra of different compounds. The mass spectra are expressed as the intensities of individual *m/z* ratios (variables).

Use PCA to find a new coordinate system that can be expressed as the linear combination of the original variables (mass-to-charge ratios *m/z*) to describe major trends in the data. Mathematically, PCA relies upon eigenvalue/eigenvector decomposition of the covariance or the correlation matrix of the original variables. PCA decomposes the data matrix X as the multiplication of two matrices P (the matrix of new coordinates of data points) and T' (transposition of the coefficients matrix of the linear combination of the original variables):

$$\mathbf{X = P \cdot T'}$$

Generally, data can be adequately described using far fewer coordinates, also called principal components, than original variables. PCA also serves as a data reduction method and a visualization tool. When the data points are plotted in the new coordinate system, the relationships and clusters are often more apparent than when the data points are plotted with the original coordinates.

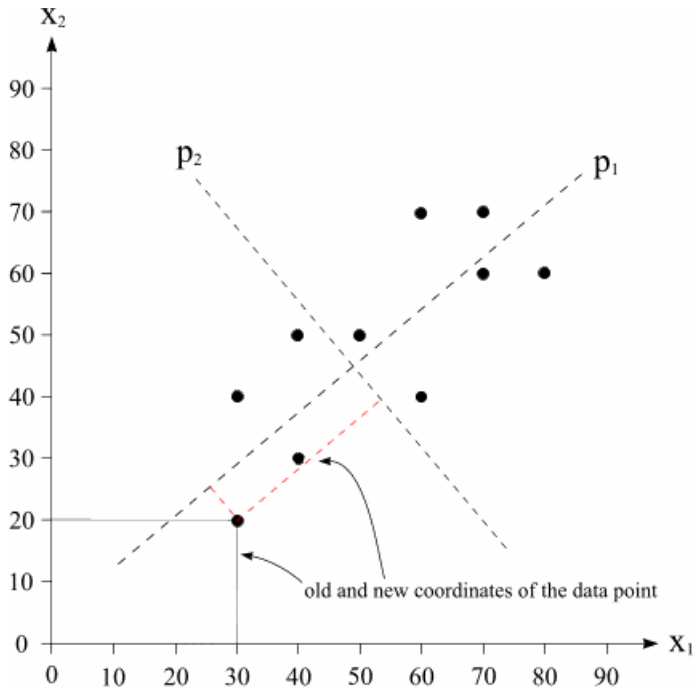


Figure 94. Geometrical interpretation of PCA

With geometrical interpretation of PCA, the axes of the new coordinate system – principal components p1 and p2 – are created as the linear combinations of the original axes. New coordinates (principal components) are orthogonal (perpendicular) to each other. There is greater variation in the direction of p1 than in either of the original variables, but very little variation in the direction of p2. The first principal component describes the direction of the greatest variation in the data set while the second principal component describes the direction of the second greatest variation (and so on, for data sets with more than two variables).

Neural Networks (Self-Organizing Maps)

Self-organizing maps (SOM), sometimes called Kohonen networks, are a special class of neural networks. A self-organizing map consists of neurons placed at the nodes of a two-dimensional lattice. The neurons become selectively activated to various input mass spectra or classes of spectra in the course of a competitive learning process. The neurons compete among themselves to be activated or excluded. SOM can be considered as a nonlinear generalization of PCA.

Use self-organizing maps to transform a set of *n*-dimensional input spectra into a discrete two-dimensional map, and to display this transformation. Each input spectrum presented to the network activates a neuron according to a complex set of interrelationships between spectra. In SOM, each mass spectrum must always activate a neuron and this spectrum is shown on the particular neuron. Spectra that activate the same neuron belong, in terms of classification, to the same pattern. To ensure that the self-organizing process has a chance to develop properly, expose the networks to a certain number of different spectra. Therefore, with Mass Frontier, use a minimum of 10 spectra in a self-organizing process.

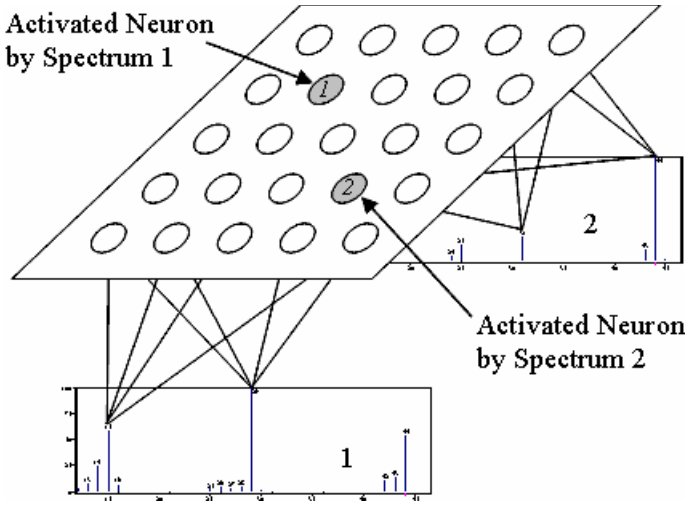


Figure 95. SOM Architecture used in Mass Frontier. Each spectrum activates a neuron in the map

Note Different results from the SOM classification method are produced for an identical data set if the input data is processed in a different order. This order sensibility is an inherent feature of neural networks and is not a result of faulty algorithms.

Fuzzy Clustering

Cluster analysis is a technique for grouping data into clusters to find common structural features in spectral data. Membership degrees between zero and one are used in fuzzy clustering instead of crisp assignments of the data to clusters. Fuzzy clustering is based on the dot product distance between the center of clusters and experimental spectral points. The dot product is calculated from mass spectral *n*-dimensional space with the intensities being the coordinates and *m/z* values dimensions. The dimensionality of spectral space is determined by the number of peaks whose intensities are above the predefined threshold. The number of clusters is usually defined *a priori*.

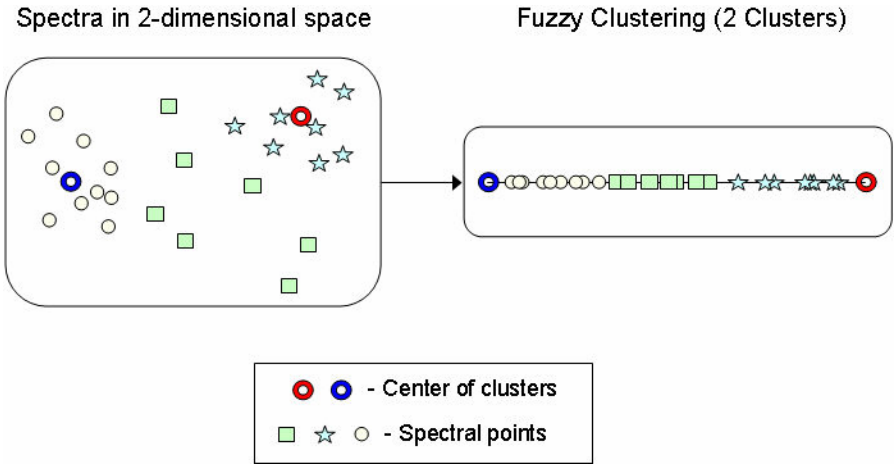


Figure 96. Transformation of two-dimensional spectral space into one-dimensional Fuzzy Clustering model

Spectra Transformation

Various mathematical transformations of mass spectra can increase classification efficiency. Better separation of classes can be achieved in some cases if transformed, instead of original spectra, are submitted to classification. In addition, some transformation procedures reduce the number of variables and lower the dimensionality of the spectral space, which shortens the computing time.

Because the most common neutral loss is 14 (loss of CH_2), the logical spectra transformation is into modulo-14 spectra, which thereafter can be used as input data for PCA. Modulo-14 spectra are calculated as the sum of peak heights at mass-to-charge ratio values shifted by 14. Each modulo-14 spectrum has 14 dimensions (transformed mass-to-charge ratio values) that are significantly lower than regular spectra. Mass Frontier offers modulo-14 transformation with (A1) or without (A2) normalization of such spectra.

Classification of mass spectra assists the interpretation of structurally related compounds. Because the characteristic peaks in spectra of structurally related compounds can be shifted due to various substituents, it can be difficult for classification methods to recognize structural similarity. Overcome this difficulty by transforming spectral data into auto-correlation spectra.

The auto-correlation function:

$$A(\Delta x) = \int f(x).\, f(x + \Delta x) dx$$

is suitable for detecting periodicity in a series of spectra. In Mass Frontier, you can choose auto-correlation transformation with (B1) or without (B2) normalization of mass spectra. Because auto-correlation does not reduce the space dimensionality and requires computing time to be calculated, a classification that uses this transformation is the most time-consuming procedure among the transformation methods.

With Mass Frontier, you can submit original (not transformed) spectra (C1) to classification as well.

Mass Frontier offers the following spectra transformations:

- Modulo-14 spectra normalized A1
- Modulo-14 spectra not normalized A2
- Auto-correlation spectra normalized B1
- Auto-correlation spectra not normalized B2
- Original spectra (not transformed) C1 (default)

No general rule exists concerning the selection of an appropriate mass spectrum transformation. Classification methods can be used for a broad range of problems, and each of them might need a different spectrum transformation. To reliably find the transformation that provides the best separation of classes for particular groups of spectra, experiment with all of them. Subsequently, use the transformation which provides you with the most information when dealing with comparable data.

Chapter 10 Spectra Classifier

You can use Mass Frontier to classify spectra according to physical or chemical properties such as point of origin, toxicity, aromaticity, and so on. For spectra that cannot be found in a library, classification can involve identification of substructure types or compound classes (structure elucidation) in order to establish and confirm structural proposals. Use classification in cases when only structurally related compounds need to be retrieved from a complex chromatogram (metabolite research).

Use the Spectra Classifier module to retrieve and organize spectra, which can then be submitted for Principal Component Analysis (PCA), Self-Organizing Maps (SOM) and Fuzzy Clustering classification. Spectra are organized into groups with their own names and graphic representations. When spectra are organized into groups, it is easier to distinguish between the classes of spectra that develop from the submitted groups of spectra after the classification process.

In classification analysis, it is important to distinguish between groups and classes. It does not make sense to speak about classes prior to the classification process. Classes can be assigned only after the selected classification method clearly shows the occurrence of clusters or regions that consist of spectra with the desired properties. Prior to classification, spectra are allocated only to groups according to user-defined criteria; that is, *a priori* information. Often the objective of the classification process is to obtain classes that closely resemble groups of spectra submitted to classification

This chapter contains the following sections:

Spectra Classifier Window

The Spectra Classifier window is the gateway to PCA, SOM and Fuzzy Clustering. Any spectra that you want to classify must first be sent or pasted to the Spectra Classifier window. A spectrum or spectra is automatically assigned to a group. Mass Frontier automatically assigns a name to a new group, which you can rename. The program also assigns a graphic representation to each group that is prepared for classification. Up to 255 groups of spectra can be added to the Spectra Classifier and each group can consist of an unlimited number of spectra. Conversely, it is also possible for a group to contain only a single spectrum. The points are assigned symbol types and colors according to settings in Classification Layout. These settings only show group membership and have no influence on the results of the analysis.

To open Spectra Classifier window

Click the **Spectra Classifier** button on the tool bar in the main window or choose **Tools > Spectra Classifier**. An empty Spectra Classifier window opens. See Figure 97.

Only one Spectra Classifier window can be opened at a time. If you click the **Spectra Classifier** button or choose **Tools > Spectra Classifier**, and the Spectra Classifier is already open, this window becomes active.

To choose spectra for classification

Use one of the following methods to choose spectra for classification:

- Select one or more records in a Database Manager window that you want to classify and click the **Add Selected Records to Spectra Classifier** button.
- Select one or more scans or components in Chromatogram Processor window that you want to classify and click the **Add Selected Scans or Components to Spectra Classifier** button.
- Use the Copy (Database Manager, Chromatogram Processor) and Paste commands to add spectra to the Spectra Classifier module.

There are two panes in the upper half of the Spectra Classifier window. The left one is a container of groups of spectra that have been sent from a Database Manager window. You fill this pane successively with the spectra you have chosen for classification. Each group of spectra is visually represented as a single line in the left or right pane. The right pane contains the groups of spectra that are actually classified if you click the **Classify Now** button. Move a group from one pane to another by clicking **Add** or **Remove**. With these two panes, you can have spectra available in the left pane for inclusion in future classification runs.

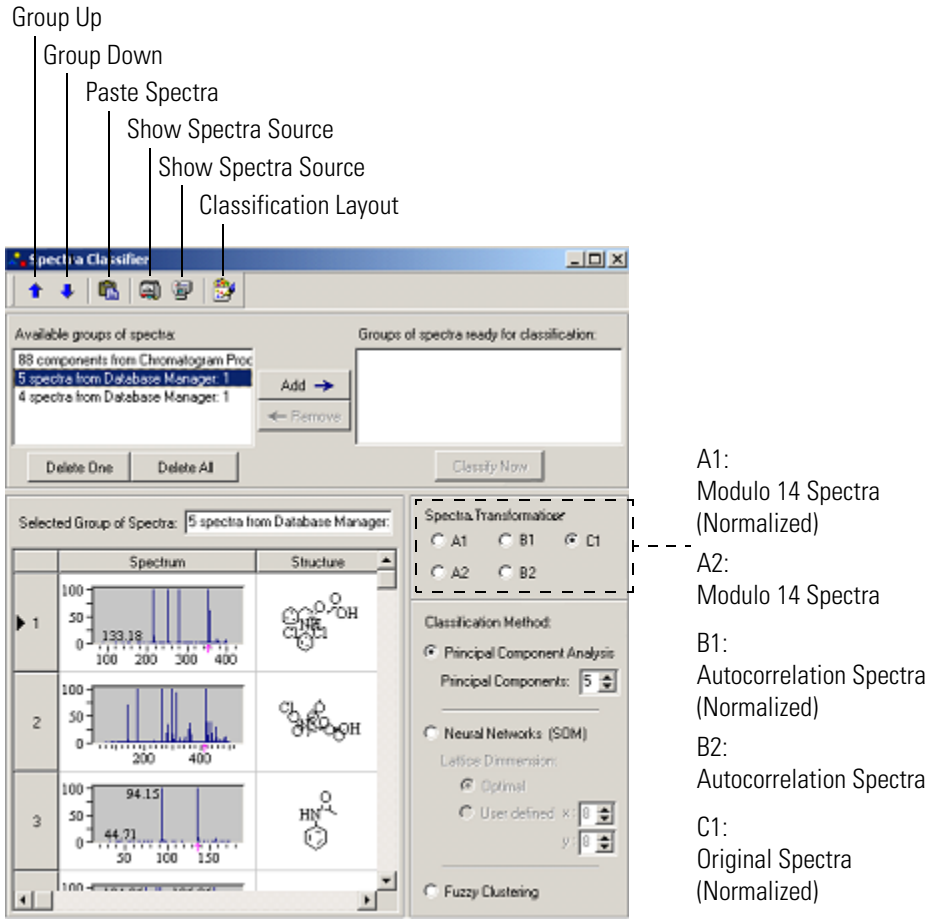


Figure 97. Spectra Classifier window

If you select a line (group) in either of the two panes, all corresponding records are previewed in the grid that displays spectra, structures, and additional information. If you double-click a line (group) in either of the two panes, the Database Manager window where the corresponding spectra originate, opens automatically and these records are selected.

Note Spectra Classifier can hold only spectra that are still present in the original Database Manager or Chromatogram Processor windows. In addition, once you have created a group of spectra in Spectra Classifier, do not move the spectra from the original Database Manager window or clear the component spectra from a chromatogram in Chromatogram Processor. If you violate either of these conditions, this group of spectra is removed from Spectra Classifier.

Spectra Classifier contains five buttons for selecting the transformation to be applied to the spectra. See Figure 97. Due to the long names of transformation techniques, only short code names are displayed. However, if you point the cursor at a code name, the full name appears.

You can choose from three classification methods: Principal Component Analysis, Neural Networks, and Fuzzy Clustering. If you choose Principal Component Analysis, you can set the number of principal components that are calculated. If you choose Neural Networks, the lattice size can be set in two ways; either you assign the x and y size, or the program sets the optimal size automatically (recommended method). The Fuzzy Clustering method has no options.

Once the data has been prepared, start the classification by clicking the **Classify Now** button. All groups of spectra listed in the right pane are classified according to the chosen options (transformation and classification method). After the classification is completed, the results are displayed either in a Spectra Projector window or in a Neural Networks window. While a classification is being processed, you can continue working with the program.

Classifying Mass Spectra

The following example clarifies how the classification of mass spectra can be carried out. Suppose you want to classify the derivatives of nicotine and caffeine. You have an EI spectrum of an unknown compound that is either a metabolite of nicotine or caffeine, and you need to find out to which of these two compound classes it belongs. To complete this task, retrieve sufficient amount of spectra for each group (nicotines and caffeines) which are then submitted to a Principal Component Analysis, to form clusters of each group. Do this by performing a substructure search, where the search query is the structure of nicotine and caffeine. If sufficient amounts of spectra (more than 10 spectra) for each group are found, select these records and add them, separately for each group, to the Spectra Classifier. Then assign an appropriate name to each group by using the Selected Group of Spectra box. You can use the copy and paste commands. Then, select a line in the left pane of Spectra Classifier and click **Add**. Repeat this for the next group.

Mass Frontier automatically assigns a graphic representation to each group. This is shown in the right pane for each line. Change the graphic representations of groups by choosing **Menu > Projection Layout**. Note that the colors and type of graphic representation are only used to distinguish the groups in the plots and have no influence on the results of the analysis. See Figure 98.

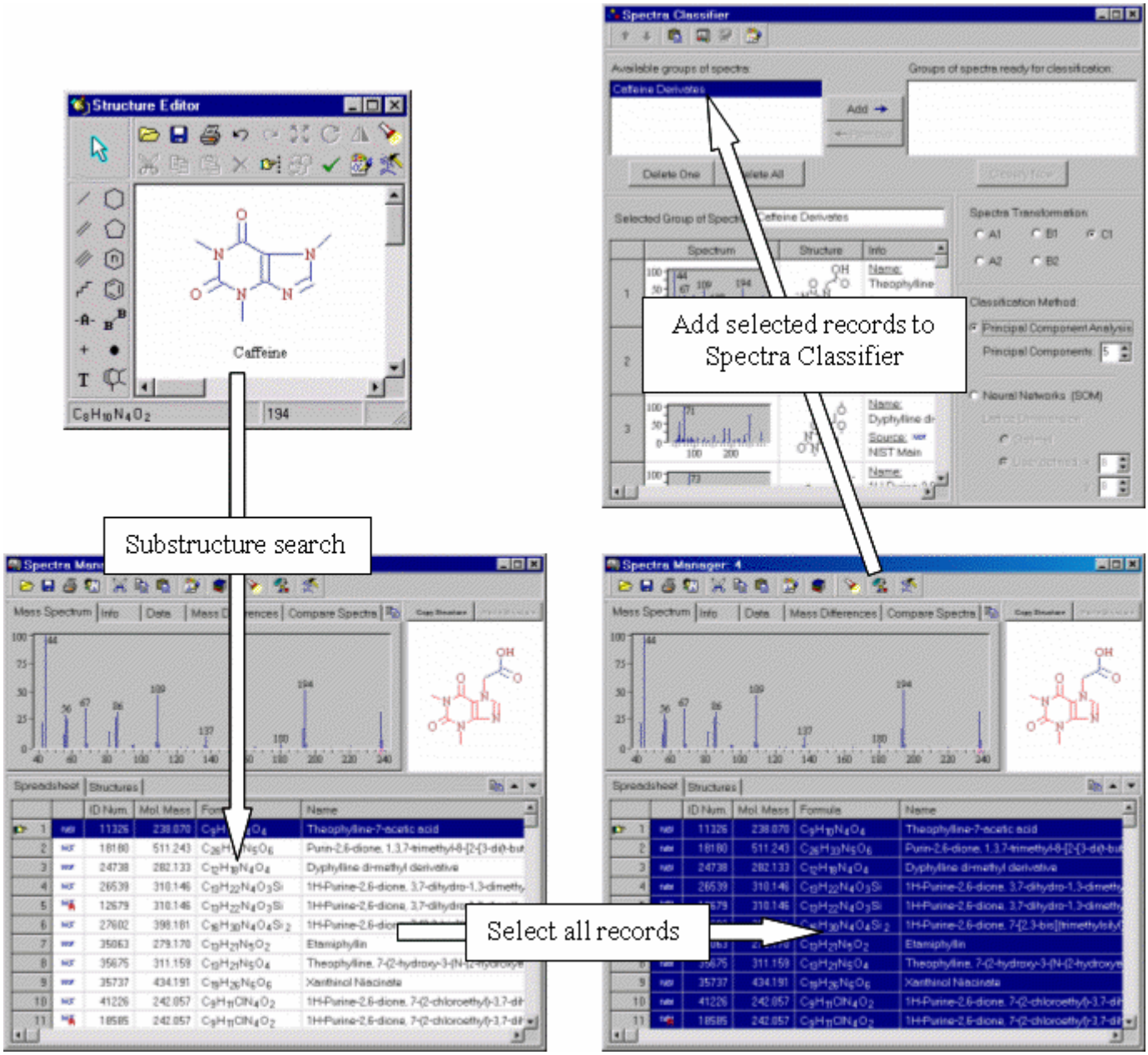


Figure 98. Retrieving data for a classification task

When all the groups of spectra you want to classify are in the right window, click the **Classify Now** button to launch the classification process. After the generation is finished, a new Spectra Projector window appears.

Maintaining Groups of Spectra

With Mass Frontier, you can save and open references to records with the extension .ref that are present in the Database Manager spreadsheet. Use this feature to maintain groups of spectra that are sent from the Database Manager to the Spectra Classifier window. If you have retrieved spectra for classification, save these records as references, rather then save the spectra as ASCII files (.jcamp or .msp). In contrast to .jcamp or .msp files, a reference file stores the locations, where all relevant information about each spectrum is saved (spectrum, structure, experimental conditions, and so on). In addition, reference files are easy to manipulate which means you can add or remove one or more records to them. Deconvoluted spectra from the Chromatogram Processor cannot be saved as references.

Chapter 11 Spectra Projector

Use the Spectra Projector module to display the results of Principal Component Analysis (PCA) and Fuzzy Clustering (FC). In Spectra Projector, mass spectra are projected as points on a two-dimensional plane or a three-dimensional twistable space, according to principal components or cluster combinations you define. Use classification analysis to find classes of spectra on the projection that exhibit common or similar properties. With the Spectra Projector, you can place an external spectrum onto a projection in order to examine its class membership.

Classification can also be applied to spectra that cannot be distributed *a priori* into groups. In this case, classify only one group of spectra, and attempt to find points on the projection that represent compounds with the desired properties. Use the Spectra Projector to select such points in order to examine their spectra and structures.

This chapter contains the following sections:

Generating Spectra Projector Window

A Spectra Projector window cannot be directly opened from the program desktop. It must be generated from the Spectra Classifier module, by selecting the Principal Component Analysis or Fuzzy Clustering options and clicking the **Classify Now** button. See You cannot alter classification results once they have been generated. If you want to remove a spectrum (point) from a projection plane, you must remove this spectrum from the input data, and then launch a new classification process. An unlimited number of Spectra Projector windows can be opened at any given time in the program.

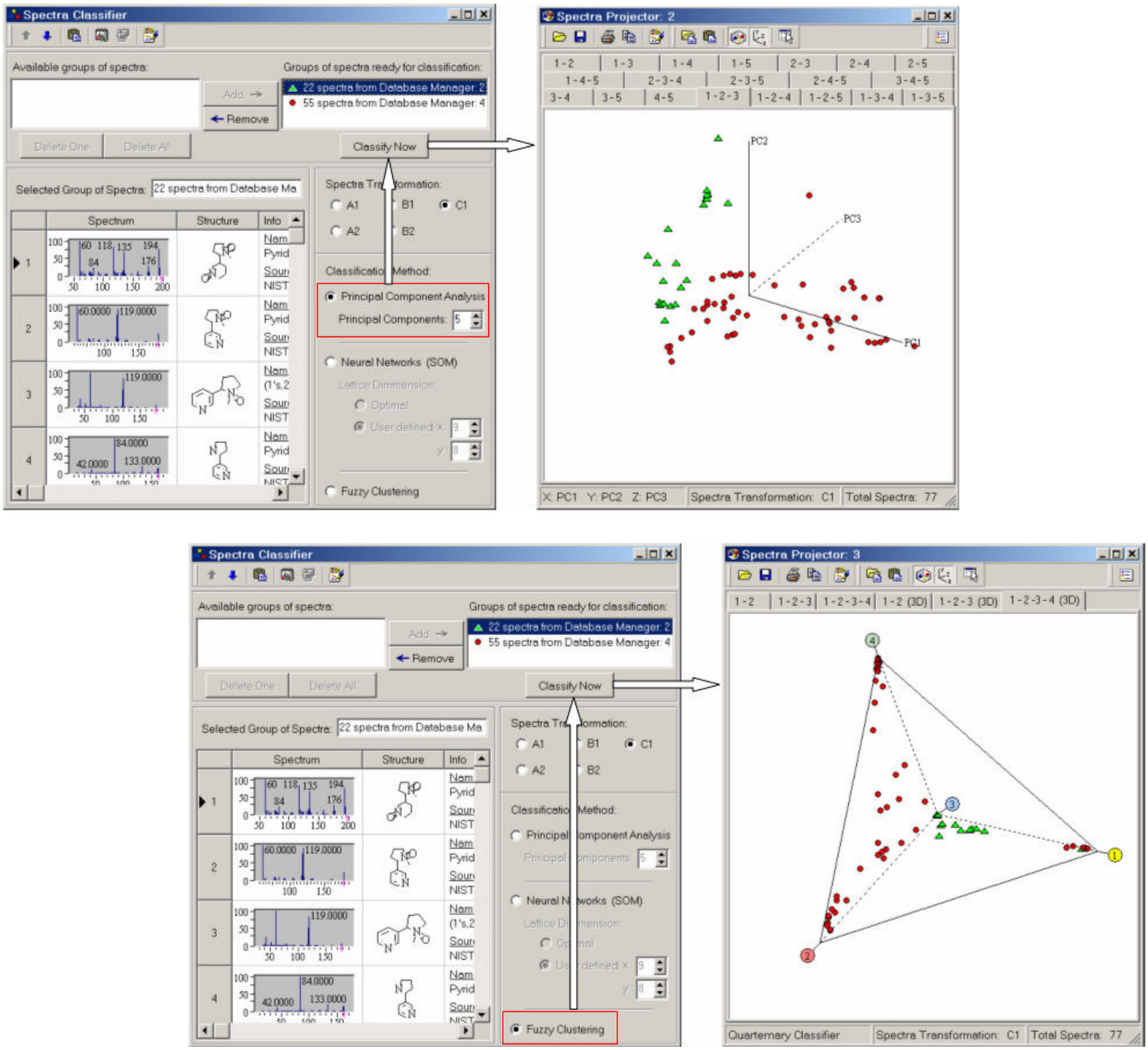


Figure 99. Generating Spectra Projector windows from Spectra Classifier using PCA and FC

Spectra Projector Window

The Spectra Projector displays spectra as points on a two-dimensional plane or in a three-dimensional space. A combination of two principal components, which make up the 2-D projection plane, can be selected in the tab control. If you are viewing PCA results using the 3-D projection mode, you can select a combination of the three associated principal components.

For example, if you click the 2-D projection tab captioned 2-5, the spectra are projected onto the plane of the 2nd and 5th principal components. The number of tabs is determined by the number of principal components that have been selected in the Spectra Classifier module. The Spectra Projector displays a tab for every possible combination of the selected principal components. For example, if you have generated Spectra Projector with 3 principal components, 3 tabs are available (1-2, 2-3, and 1-3). For the 5 principal components (which is the default), 10 tabs are available (1-2, 1-3, 1-4, 1-5, 2-3, 2-4, 2-5, 3-4, 3-5, and 4-5). The same principle applies to the 3-D projection mode. Note that in the 3-D mode, the number of combinations of principal components for more than five components is significantly higher in comparison to 2-D. Therefore, to analyze your spectra in the 3-D mode, do not use more than five principal components. Set the number of principal components in the Spectra Classifier window prior to the PCA calculation.

The Status bar in a Spectra Projector window is divided into three parts. See Figure 100. The left part informs you which principal components have been selected. The middle part displays the type of spectra transformation that has been used for classification. The right part shows how many spectra have been classified (the total number of spectra from all groups).

Use the Spectra Projector to enlarge any region of the projection plane independently, for each combination of principal components, by using the left mouse button. To get the original scale, click the **Zoom Out** button, which appears in the top-left corner of the projection plane after any change of scale. In the 3-D mode, you cannot enlarge the projection.

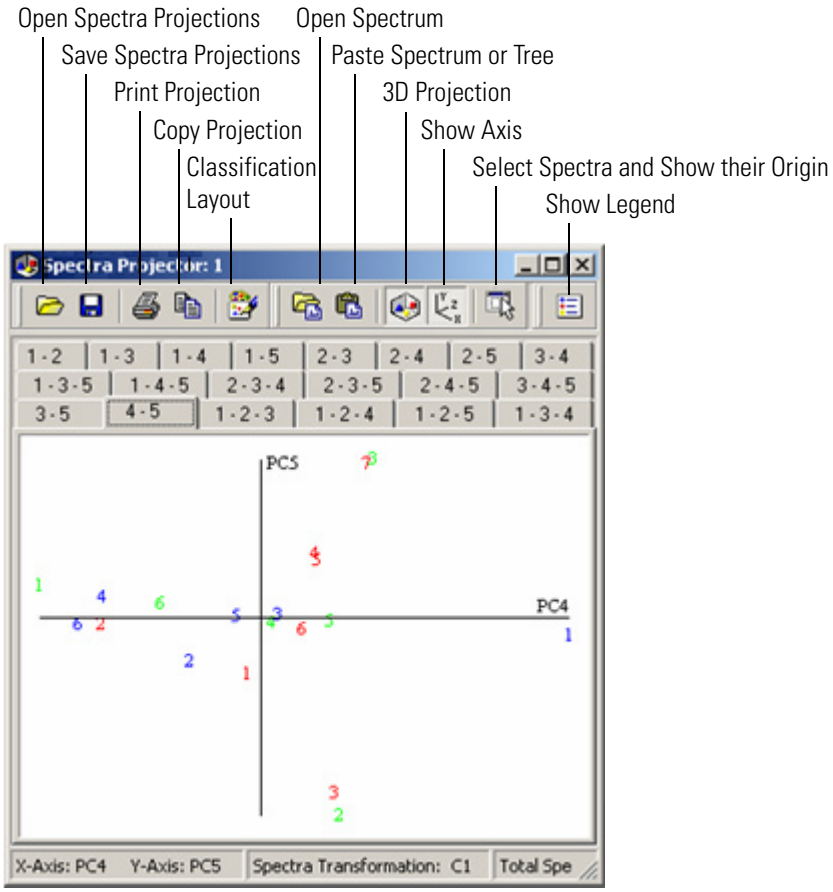


Figure 100. Spectra Projector window

3-D Projection Mode

Use Principal Component Analysis and Fuzzy Clustering to reduce the dimensionality of a spectral space to a level comprehensible to the human eye. Because a paper or screen plot is inherently 2-D, the method of choice is usually 2-D Principal Component Analysis or Fuzzy Clustering. However, if a 3-D projection on a 2-D computer screen is accompanied with motion, you might perceive this simulation as 3-D space. To increase the space feeling, the 3-D plots are interactive; that is, the plot can be rotated with the mouse. Spectra Projector enables interactive 3-D viewing of your data.

The 3-D projection mode used in Mass Frontier stems from the idea that spectra classification of complex data set in a 3-D space can be more effective and reliable than a 2-D plot. This idea is based on the fact that the human brain recognizes visual patterns in 3-D space efficiently.

When analyzing Principal Component Analysis or Fuzzy Clustering results, do not rely entirely on 2-D plots. Two or more clusters which overlap in a 2-D plot might be separated in 3-D space, or two seemingly separate clusters discernible in a 2-D plot might appear to be too close together in 3-D mode to be considered as two different objects.

To view PCA or FC results in 3-D mode

- Click the **3-D Projection**  button.
- Left-click anywhere in the classification plot and move the mouse to rotate the plot interactively. Note that the arrow cursor is replaced by a circle during the dragging step.

To better appreciate the three-dimensionality of the space, display the x-, y-, and z- axes by clicking the **Show Axes**  button, which is next to the **3-D Projection** button.

Note It is not possible to enlarge a part of the plot in 3-D projection mode. However, as in the 2-D mode, the selection of particular spectra and the display of their origins is possible.

Opening and Saving of Classification Results

With Mass Frontier, you can save and open projection planes that have been generated with particular initial settings (transformation method and number of principal components).

To open spectra projection planes

Do one of the following:

- Click the **Open Spectra Projections** button in any Spectra Projector window.
- Choose **File > Open > Spectra Projections**.

To save spectra projection planes

Do one of the following:

- Click the **Save Spectra Projections** button in any Spectra Projector window.
- Choose **File > Save > Spectra Projections**.

Projection planes are saved as graphics, without the possibility of recalling the original spectra, which are displayed as points. If you need to access the spectra with structures from projection planes, you must save the references to records in the Database Manager, as described in Chapter 10. In this case, the classification must be regenerated from the original data set that was saved as references. However, it is possible to add an external spectrum.

Accessing Spectra from Spectra Projector

It is an important part of classification analysis to know which spectrum is represented by a point on a projection plane. Because Mass Frontier links corresponding modules, spectra with structures or chromatographic components can be recalled from any projection plane.

To access a single spectrum, scan, or deconvoluted component

Double-click a point on the projection. The linked module appears on top of all other windows, the corresponding record or component is selected, and the spectrum is shown.

To access several spectra in a region

Do one of the following:

- Select a region with the mouse cursor while holding down the right mouse button.
- Click the **Select Spectra And Show Their Origin**  button and then select a region with the cursor while holding down the left mouse button.

If the selected spectra originate from the Database Manager, you are prompted about whether to copy the records with spectra and structures to the last active Database Manager window or to a new one. If your spectra originate from a chromatogram (Chromatogram Processor), the corresponding scans or components appears in the same color as they appear in the Principal Component Analysis or Fuzzy Clustering plot.

Note Spectra Projector is able to recall only those records that are present in a Database Manager window. If you close the Database Manager window that is the source of the input data for classification, the link between a point and its spectrum is interrupted. The program automatically warns you if you try to close a Database Manager window that is linked with one or more Spectra Projector windows. In addition, to keep the links between points and spectra intact, you must not move records in, or between, Database Manager windows by using the cut and paste commands.

If the data originated from Chromatogram Processor, the window must still be open.

Adding An External Spectrum

Use principal component projections of mass spectra to find classes of compounds with common or similar properties. Sufficiently separated classes can then be used to investigate the class membership of an unknown spectrum. Use the Spectra Projector to add an external spectrum, which was not used for classification, to the projection plane. If the added spectrum is clearly projected into a particular class region, assume that this compound has similar, usually structural, properties.

Figure 101 shows that the two groups are separated into clusters. You can add the unknown spectrum to the projection plane by pasting or opening its spectrum into the Spectra Projector. The projection of this external spectrum clearly shows that it belongs to the nicotine class (squares are PCA and triangles are FC). This result can be confirmed or rejected by using the fragmentation pattern of nicotine.

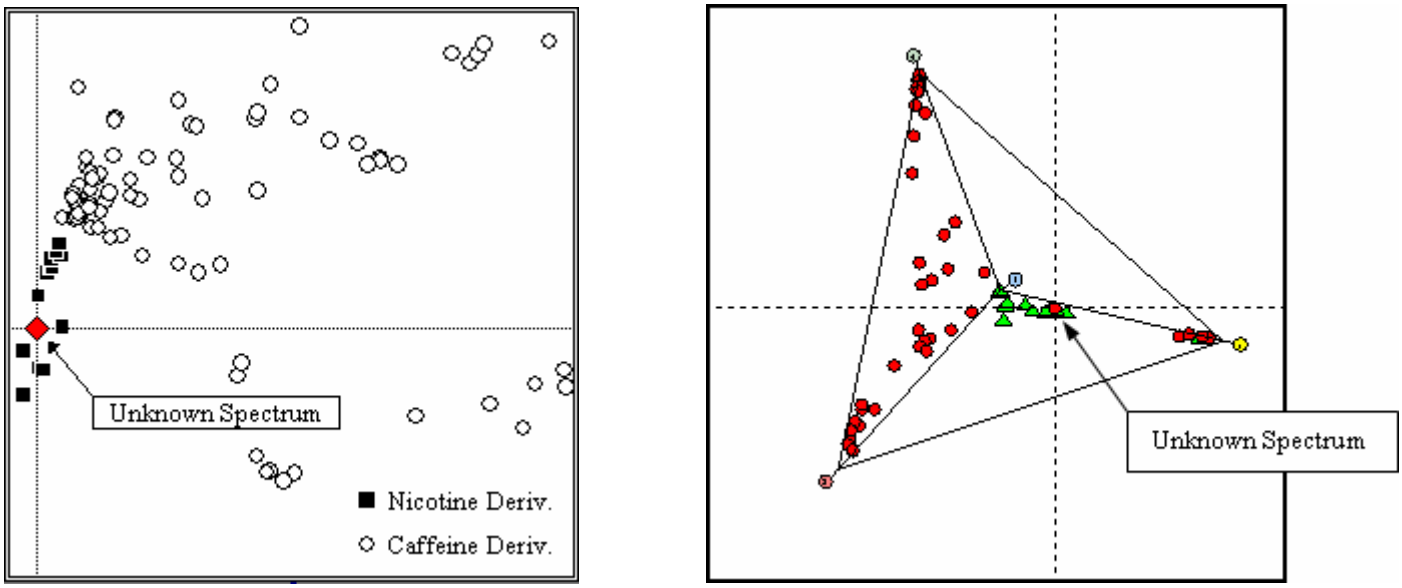


Figure 101. Projection of unknown metabolite into the nicotine class using PCA and Fuzzy Clustering

Chapter 12 Neural Networks

Use the Neural Networks module to display classification results from Self-Organizing Maps (SOM), which are a class of neural network. A self-organizing map is a network of neurons, arranged in the form of a two-dimensional lattice. The size of a lattice can either be calculated automatically or user-defined. During classification, neurons become selectively activated to various input spectra as a result of a competitive learning process.

Use classification analysis using SOM to find classes of spectra on the map that exhibit common or similar properties. If one or more spectra activate the same neuron, assume the spectra belong to a common class. In this case, the spectra exhibit certain similarities. In addition, spectra that activate neighboring neurons, and those neurons that have low Euclidian distance between each other (shown by border line thickness), can also be considered as related.

In neural networks, each mass spectrum must always activate a neuron and this spectrum is shown on the particular neuron. Neurons are displayed as rectangles on the screen. Spectra are represented as symbols or numbers and are placed onto neurons. Because the spectra are located in discrete objects the interpretation of SOM is relatively easy, in contrast to PCA and FC where you do not have to deal with diffuse clusters. However, it might happen that larger numbers of neurons are activated by a single spectra and this advantage is lost.

Spectra that activate the same neuron belong, in terms of classification, to the same pattern. Therefore, all spectra drawn inside a neuron box are equal for classification purposes and their graphical positions within a neuron are irrelevant.

Note Different results from the SOM classification method are produced for an identical data set if the input data is processed in a different order. This order sensibility is an inherent feature of neural networks and not a result of faulty algorithms.

This chapter contains the following sections:

Generating Neural Networks Window

Neural Networks are generated from the Spectra Classifier module. As in the Spectra Projector, the Neural Network module cannot be opened directly from the program desktop. Once the network has been generated, classification results cannot be altered. If you want to remove a spectrum (symbol) from a network, you must remove this spectrum from the input data, and then launch a new classification process. See Figure 102. An unlimited number of Neural Networks windows can be opened at any given time in the program.

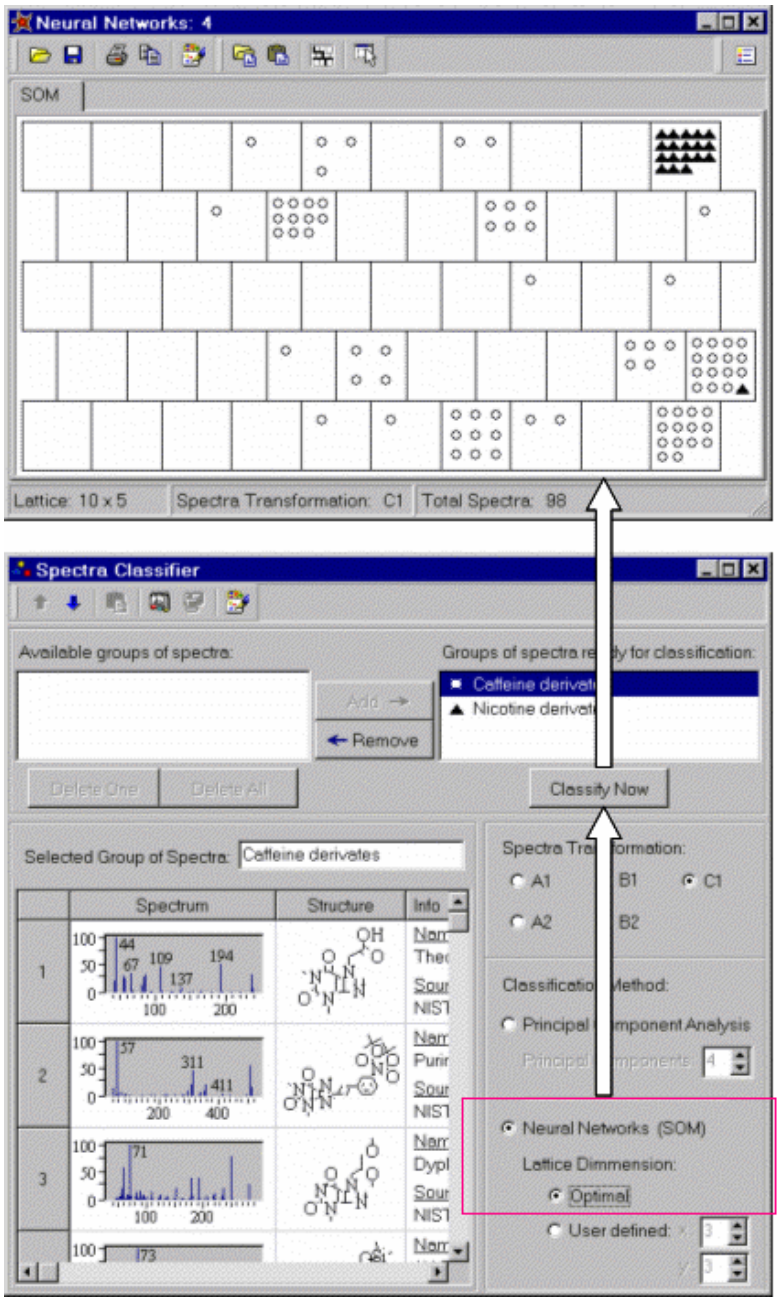


Figure 102. Generating a Neural Networks window from Spectra Classifier

Neural Networks Window

A Neural Networks window displays a topological map of neurons that are drawn as rectangle. Every neuron has a minimum of two and maximum of six neighboring neurons. Spectra are displayed as symbols or numbers within neurons.

The Status bar in a Neural Network window is divided into three parts. The left part informs you about the lattice dimension. The middle part displays the type of spectra transformation that has been utilized for classification. The right part shows how many spectra have been classified (the total number of spectra from all groups).

Use the Neural Networks module to enlarge any region of the lattice, by using the left mouse button. To return to the original scale, click the **Zoom Out** button, which appears in the top-left corner of the projection plane after any change of scale.

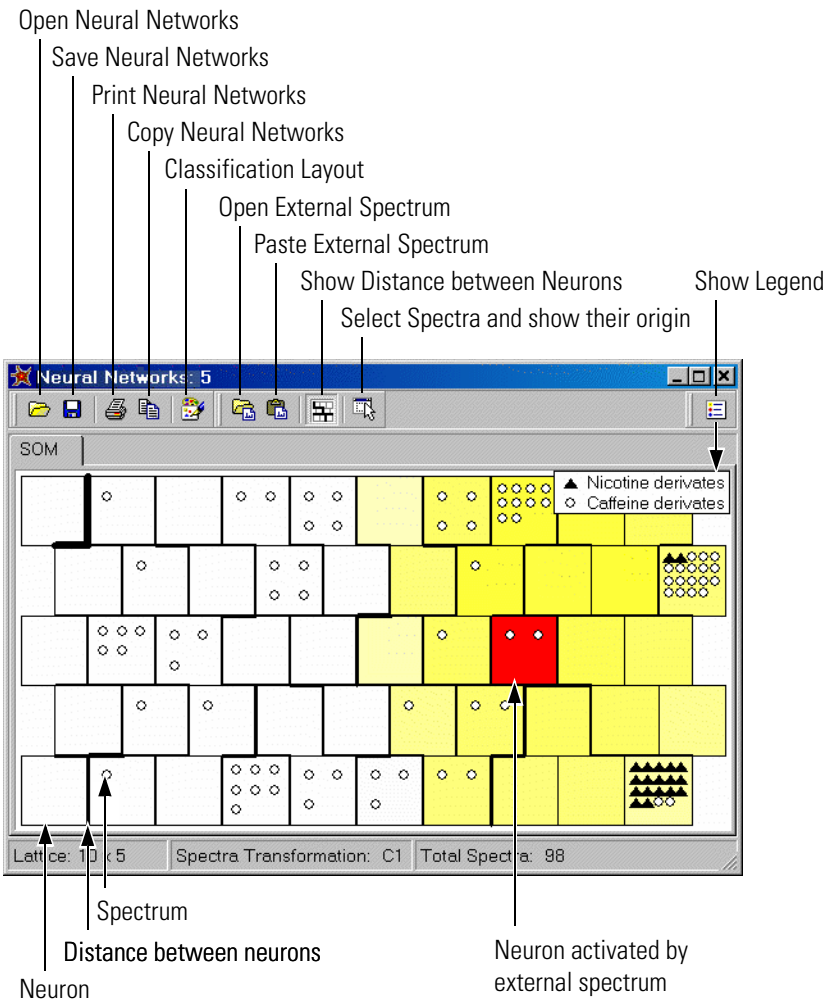


Figure 103. Neural Networks window

Although neurons are displayed in a regular 2-D lattice, the actual Euclidian distances between neurons vary. To show the approximate distances between neighboring neurons, the borders of the neuron are shown using different thickness. The line thickness is proportional to the distance between immediate neighbors. The thinner the line, the closer together are the neurons.

To show neuron distances

Click the **Show Distances between Neurons** button. See Figure 103.

Use information about neuron distances when classifying an unknown spectrum. The distances show how a neuron that has been activated by an unknown correlates with neighboring neurons. The distances can also suggest that spectra from a common might have been activated by multiple close neurons.

Note Spectra that activate the same neuron belong, in terms of classification, to the same pattern. Therefore, all spectra drawn inside a neuron box are equal for classification purposes and their graphical positions within a neuron are irrelevant.

Working with Neural Networks

With Mass Frontier, you can save and open neural networks that have been generated with particular initial settings such as transformation method and lattice dimension.

To open or save neural networks

Do one of the following:

- Click the **Open Neural Networks**  in any Neural Networks window.
- Choose **File > Open > Neural Networks**.

To save neural networks

Do one of the following:

- Click the **Save Neural Networks**  button in any Neural Networks window.
- Choose **File > Save > Neural Networks**.

Neural Networks are saved as graphics, without the possibility of recalling the original spectra, which are displayed as symbols. If you need to access the spectra from lattice, you must save the references to records in the Database Manager, as described in Chapter 10, "Spectra Classifier." In this case, the classification must be regenerated from the original data set that was saved as a reference.

It is an important part of classification analysis to know which spectrum is represented by a symbol on a lattice. As Mass Frontier links corresponding modules, spectra with structures or chromatographic components can be recalled from any neural network.

To access a single spectrum, scan, or deconvoluted component

Double-click a symbol on the lattice. The linked module appears on top of all other windows, the corresponding record or component is selected, and the spectrum is shown.

To access several spectra in a region

Do one of the following:

- Select a region with the mouse cursor while holding down the right mouse button.
- Click the **Select Spectra And Show Their Origin**  button and then select a region with the mouse cursor while holding down the left mouse button.

If the selected spectra originate from the Database Manager, you are prompted whether to copy the records with spectra and structures to the last active Database Manager window or to a new one. If your spectra originate from a chromatogram, the corresponding scans or components appear in the same color as they appear in the Neural Networks window.

Note The Neural networks module is only able to recall records that are present in a Database Manager window. If you close the Database Manager window that is the source of the input data for classification, the link between a symbol and its spectrum is interrupted. The program automatically warns you if you try to close a Database Manager window that is linked with one or more Neural Networks windows. In addition, to keep the links between symbols and spectra intact, you must not move records in, or between windows. The Database Manager must still be open.

If the data originated from the Chromatogram Processor, the window must still be open.

Chapter 13 Chromatogram Processor

Use the Chromatogram Processor module for extraction and processing of mass spectral scans from GC/MS or LC/MS files and from MS data. Component detection and spectra deconvolution techniques are available. This module also offers spectral averaging and background subtraction features. Three types of chromatograms can be displayed: Total Ion Chromatogram (TIC) of MS and MS/MS, and Selected Ion Chromatogram (SIC). Use the SIC feature to display individual mass chromatograms of ions that are characteristic for a specific compound of interest.

Mass spectral scans, deconvoluted components, and various types of MS data can be classified using Principal Component Analysis (PCA), Self-Organizing Maps (SOM), or Fuzzy Clustering (FC) methods. Spectra can be searched in a library for positive compound identification, or the spectra can be copied to the Database Manager module for further processing and archiving. Several chromatograms can be opened simultaneously. Fully customizable chromatogram and mass spectrum layouts are available. Use the Chromatogram Processor to copy chromatograms with extracted spectra import into reports, spreadsheets, or other Windows programs. MS data from Xcalibur is displayed in tree structures allowing the user to clearly view the dependencies. Using text boxes, you can annotate chromatographic scans or components. This module does not include target analysis, and automatic quantitation of ions is not available.

This chapter contains the following sections:

- Selected Ion Chromatogram
- Thresholding, Baseline Correction and Smoothing
- Automated Component Detection and Spectra Deconvolution
- Processing Xcalibur MS^n Data

Chromatogram Processor Window

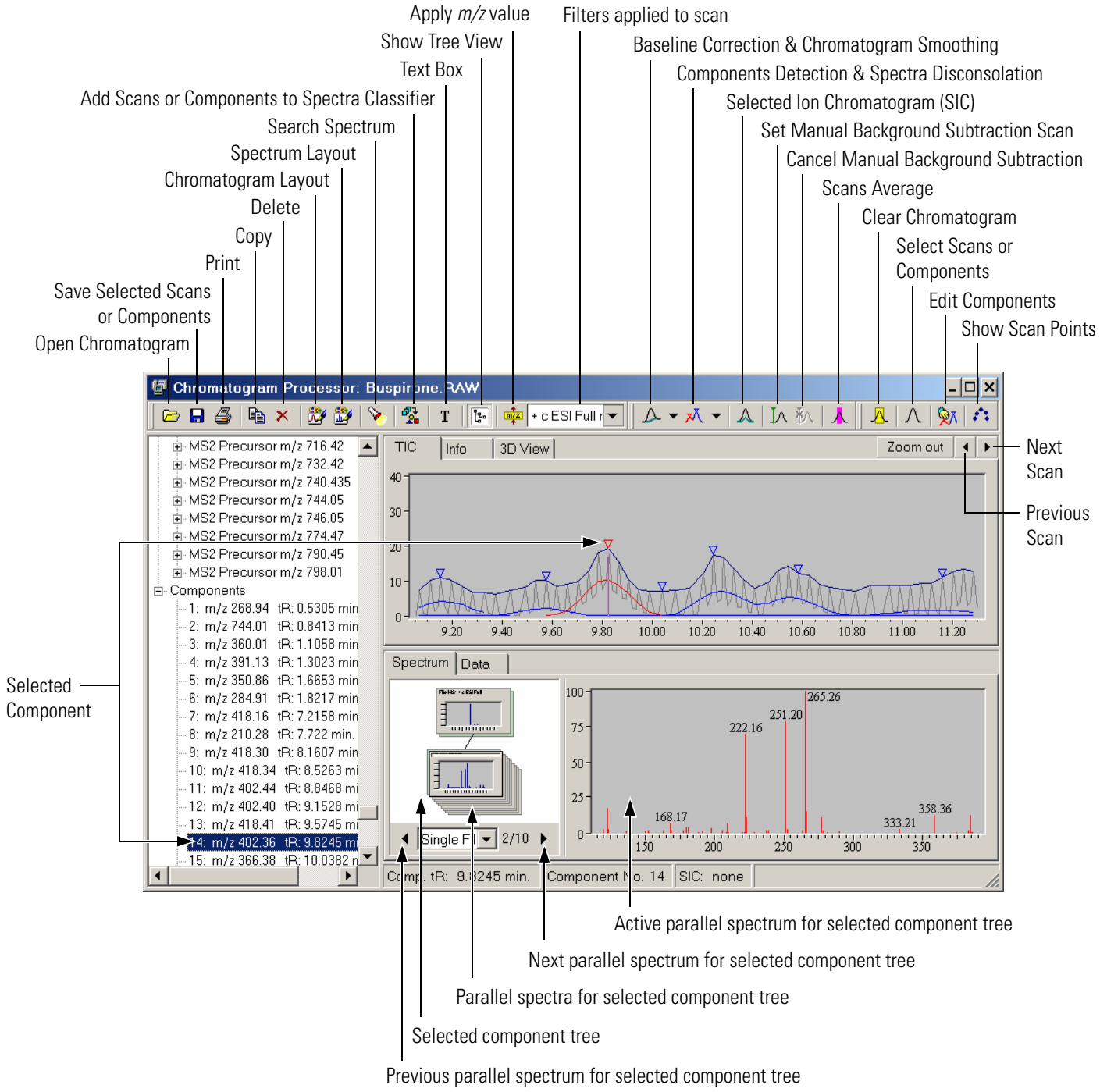


Figure 104. Chromatogram Processor window

Data File Formats

Mass Frontier supports the following data file formats for importing GC/MS and LC/MS files:

- Xcalibur .raw files (MS and MS^n)
- Finnigan LCQ, GCQ, ITS40, and Magnum
- Varian Saturn
- ChemStation .hp
- netCDF
- DOS, Windows, and UNIX .jcamp

These files can be imported to the Chromatogram Processor but cannot be exported. Single scans can be saved in .jcamp or .mps format.

Mass Frontier supports centroid-type data for mass spectra. Centroid mass spectra are displayed as a bar graph. Profile-type data is not supported.

Due to various netCDF standards implementation, some .cdf files might not be readable in Mass Frontier.

Note Features connected with chromatographic spectral trees are supported for Xcalibur .raw files only (Data Dependent files).

Opening Chromatograms

To load a GC/MS or LC/MS file to a Chromatogram Processor window

1. Click the **Chromatogram Processor** button on the speed bar or choose **File > Open > GC/LC/MS**.

2. After an Open GC/LC/MS File window appears, select the appropriate file type and the extension in the Files of type box.

3. Choose the file to open and click the **Open** button.

With Mass Frontier, you can open chromatograms with up to 30000 scans. If you open a file that contains more than 30000 scans, the chromatogram is displayed up to the retention time corresponding to scan number 30000. The rest are ignored and are displayed. The program does not allow the opening of GC/MS or LC/MS files from a CD-ROM. If you have files on a CD-ROM, copy these files to the hard drive to make them accessible for importing. In addition, disable the Read Only file attribute.

To change the file attribute

1. Select all the files whose attributes to change in Windows Explorer.

2. Right-click the selected files and a pop-up menu appears.

3. Select **Properties** in the pop-up menu.

4. Check the Read Only attribute box that is located at the bottom of the Properties dialog window.

Mass Frontier comes with GC/MS and Xcalibur MS demonstration files which are located in the /Chromatograms directory.

TIC Page

Total Ion Current (TIC) chromatograms are displayed in the TIC tab in the Chromatogram Processor. Both the retention time and the intensity axis can be zoomed by holding down the left mouse button and dragging a rectangle around the region you want to rescale. To unzoom the TIC, click the **Zoom Out** button in the top-right corner.

To select a scan from the TIC, click anywhere in the chromatogram pane and the scan corresponding to the selected retention time appears in the Scan tab. The active scan is indicated by a vertical line, which is purple by default. To move the scan point to the next or previous scan, click the arrow button of the direction you want to move the scan point. Retention times (tR) and scan numbers of active scans are displayed in the status bar.

Info Page

The Info tab in the Chromatogram Processor window shows additional information saved in the data file. Each GC/MS or LC/MS file contains a different list of items describing sample, instrument, experimental conditions and other information. See Figure 105. Mass Frontier extracts all additional information from the file and puts them into the Info tab.

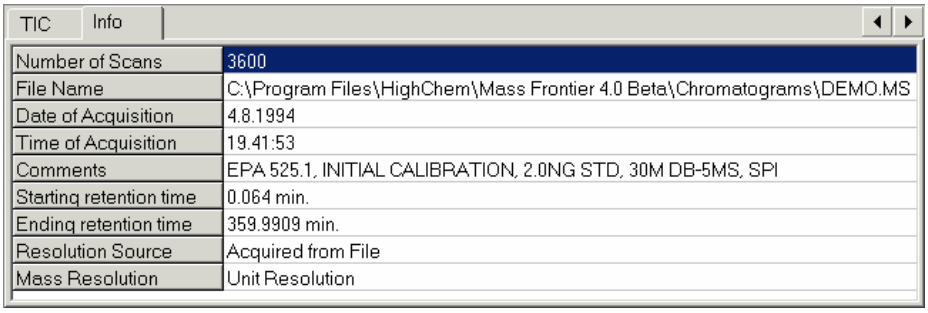


Figure 105. Info page of the Chromatogram Processor window

Spectra Averaging

With Mass Frontier, you can extract the average mass spectrum from several scans. Click the **Spectra Average** button and mark off the region of the scans to be averaged by dragging a rectangle around the region. The marked region is displayed in the same color as the active scan line (the default is purple). The average spectrum is displayed in the Scan tab. See Figure 106. To cancel the spectra average mode, click anywhere in the chromatogram pane and the single scan mode is restored.

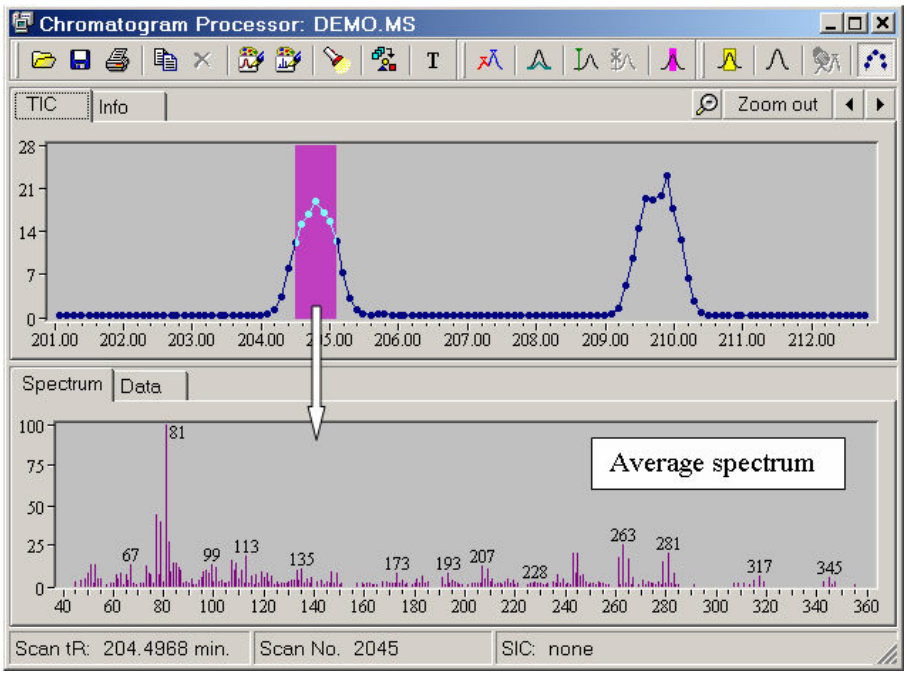


Figure 106. TIC page showing a selected region of several scans to be averaged

Background Subtraction

To eliminate the background signal from an active scan or from the average of spectral scans, perform a manual background subtraction. Set the location of two representative background scans by clicking the **Manual Background Subtraction** button and then left-clicking the scan point in the chromatogram pane. The background scans are indicated by a vertical line, which is depicted in a different color (the default is green) from the scan line. The resulting spectrum is displayed in the Scan tab. To cancel a background subtraction, click the **Cancel Background Subtraction** button. Use the Selected Ion Chromatogram feature to choose the two most representative background scans. Such background scans can be used in the Manual background subtraction option in the Automated Components Detection and Spectra Deconvolution procedures.

To see the effects of the background subtraction on a scan, copy both the extracted and original spectra to the Database Manager and compare them by using the spectra comparison routine in the Compare Spectra tab.

Background subtraction can be used in conjunction with the Selected Ion Chromatogram feature. In this case, the Selected Ion Chromatogram profiles are extracted from background subtracted scans.

Processing Extracted Spectra

A spectrum obtained from the average of scans, from scans with a subtracted background or from component deconvolution can be further processed in the Database Manager, the Spectra Classifier, or can be directly searched for in a library from the Chromatogram Processor. To transfer the extracted spectrum to the Database Manager or Spectra Classifier, use the **Copy** button. After clicking the **Copy** button in the Chromatogram Processor window, paste the mass spectrum to a Database Manager or Spectra Classifier window. In addition, when you click the **Copy** button in the Chromatogram Processor, the chromatogram graphic and the spectrum from the Scan tab are copied to the Windows Clipboard and can be used for creating reports in any Windows application.

The Chromatogram Processor window contains a **Search** button for the direct searching of spectrum from the Scan tab in libraries. To search the extracted spectrum in libraries, click the **Search** button in the Chromatogram Processor window and the Spectra Search dialog window appears.

Mass Spectra obtained in Chromatogram Processor can be classified by using Principal Component Analysis and Neural Networks.

To choose scans or deconvoluted spectra for classification

Select scans or components that you want to classify and click the **Add Selected Scans or Components to Spectra Classifier**  button.

Selected Ion Chromatogram

With Mass Frontier, you can display a chromatogram for a selected ion in a different color. In Xcalibur, the selected ion chromatogram (SIC) is sometimes called an individual or single ion chromatogram, ion profile, or selected ion monitoring (SIM). The program can display up to three selected ion chromatograms per window.

To display an SIC for a particular mass-to-charge ratio

Do one of the following:

- Click on the spectral peak of interest on the Scan page.
- Select the Data tab and click the *m/z* value in the mass-to-charge ratio table.
- Click the **Selected Ion Chromatogram** button.

In either case, a dialog window appears where you can change or add the mass-to-charge ratio and confirm your choice. Use this dialog window to select a color for a particular *m/z* value.

Mass Frontier extracts the SIC from the original file. This process can be time-consuming for chromatograms with a large number of scans. However, because Mass Frontier is a multithreading application, you can still use other windows during the SIC extraction.

Use a selected ion chromatogram to verify automated component detection and spectra deconvolution results. An SIC helps you to quickly recognize mixture components in a peak region. Examine an SIC of model peaks for any component you want to use in further analysis. Remember that some structural (alkanes) or optical isomers produce almost identical mass spectra and even if you can clearly see two or more maxima in a peak region, an SIC might not reveal a multicomponent profile.

Use an SIC to determine whether the composition of the background changes over the course of a run. To view the background profile, extract the SIC of a base peak, or prominent peak, from a scan which is clearly in a non-peak region. If the SIC of a background peak has a variable profile around the peak you are focusing on, choose two scans with different SIC profiles for background subtraction.

Background subtraction can be used in conjunction with the SIC option. In this case, the SIC profiles are extracted from background subtracted scans.

Note SIC peaks can appear larger on the screen than the corresponding TIC peaks do. In reality, SIC signals are disproportionally lower than TIC signals, especially if the data was not acquired using the selective ion monitoring mode. See Figure 107. It would be impractical to examine minute SIC peaks displayed together with TIC on the screen. Therefore, SIC data are normalized to the maximum of the TIC signal and visually enlarged for display purposes.

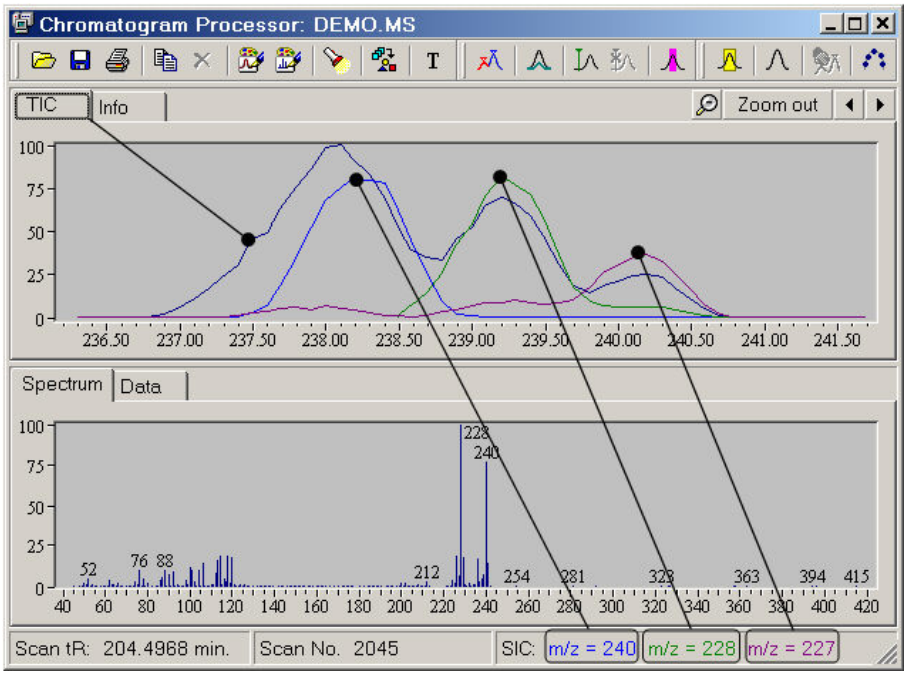


Figure 107. Chromatogram Processor, showing selected ion chromatograms of ions with *m/z* 240, 228, and 227

Thresholding, Baseline Correction and Smoothing

The interfacing of gas and liquid chromatography with mass spectrometry often produces a signal that is not associated with information of interest. Mass Frontier provides tools for GC/MS and LC/MS data processing to improve the useful signal using a variety of advanced algorithms. The following types of data processing techniques are used in Mass Frontier:

- Thresholding
- Baseline Correction and Noise Elimination
- Smoothing

The main purpose of chromatographic data processing is to prepare data to be valid for the automated detection and deconvolution of chromatographic components. There are two different ways of combining data processing and component detection. You can preprocess data in advance by using a variety of methods for every processing type separately with number of customizable options and then begin component detection. You can also perform data processing at the same time as component detection by using predefined methods and options optimized for common chromatograms in the same window used for component detection.

Thresholding

Thresholding is a data processing method that analyzes every scan to reduce ion intensities or delete spectral peaks if algorithmic criteria have been met. The main purpose of thresholding is to eliminate noise or impurity (column bleeding) peaks or peaks originating from minor components that are not of interest.

To apply thresholding to your data, click the **Chromatogram Processing** button and choose **Thresholding**. The dialog window appears with the following parameters:

- Apply Threshold to Top Stage Only

  If checked, only MS^1 (full scan) scans are processed.

- Minimal Remaining Peak Count

  If checked, the number of the most intense peaks specified in the Minimal box are affected by thresholding.

- Algorithmic

  If checked, one of the following algorithms is used: Linear Fit, Histogram or Median.

- Apply Threshold to Spectra with Peak Count Higher than

  If checked, algorithmic thresholding is applied to spectra with a number of peaks higher than the specified value.

- Maximal Allowed Threshold

  If checked, the algorithmic threshold is applied to spectral peaks with an intensity lower than the specified value.

- Manual

  If checked, all peaks that exhibit an intensity lower than the value given in the Threshold box are deleted.

Baseline Correction and Noise Elimination

Baseline correction and noise elimination algorithms analyze ion profiles (selected ion chromatograms) of all ions appearing in spectra over the entire region of a chromatogram. In contrast to thresholding, where the individual scans and their spectral peaks are independently analyzed and modified, baseline correction and noise elimination analyze and modify spectral peaks in conjunction with identical peaks in a given retention time range.

To apply baseline correction and noise elimination to your data, click the **Chromatogram Processing** button and choose **Baseline**. The dialog window appears with the following parameters:

- Process Top Stages Only

  If checked, spectra in MS^1 stage (full scan) are processed.

- Segmentation

  If checked, the chromatogram is divided into discrete parts to which the algorithms are applied separately. This option can provide better results if a chromatogram exhibits diverse shape, peak density and baseline characteristics over the retention time scale. Avoid segmentation if a Loess derivative filter is used.

- Use Baseline Correction

  If checked, one of the following methods is applied to your chromatogram: Top-Hat filter, Savitzky-Golay derivative filter, or Loess derivative filter.

- Use Noise Elimination

  If checked, apply the following methods individually or simultaneously: Counter filter to reduce chemical noise, and Quantile filter to reduce electronic noise.

Smoothing

Smoothing is a process by which ion profiles (not total ion chromatogram) for every ion (*m/z* value) found in the data file are averaged with their neighbors in a time series. Smoothing can increase correct component detection and can eliminate spikes that cause false positive results.

To apply smoothing to your data, click the **Chromatogram Processing** button and choose Smoothing. The dialog window appears with the following parameters:

- Smooth Top Stages only

  If checked, smoothing is applied only to MS^1 stage (full scan).

- Segmentation

  If checked, the chromatogram is divided into discrete parts to which the algorithms are applied separately. This option can provide better results if the chromatogram exhibits diverse shape, peak density and baseline characteristic retention time scale.

- Method

  This box includes the Savitzky-Golay, Median and Loess smoothing methods. In contrast to Median and Loess methods, the Savitzky-Golay method requires time-equidistant scans.

Automated Component Detection and Spectra Deconvolution

Mass Frontier incorporates an advanced automated system for detecting chromatographic components in complex GC/MS or LC/MS runs and extracting mass spectral signals from closely coeluting components (deconvolution). Individual mass spectra or spectral trees obtained after deconvolution can be searched in libraries or classified using Principal Component Analysis and Neural Networks.

Mass Frontier identifies components using the following algorithms:

- Rapid Component Detection (RCD) Algorithm
- Joint Component Detection (JCD) Algorithm
- Total Extraction Component Detection (TECD) Algorithm
- Direct Infusion Algorithm

These algorithms involve the combined use of the following procedures:

1. Noise examination and signal filtering.
2. Baseline definition and demarcation of chromatographic peaks.
3. Background scan determination and background subtraction.
4. Component candidate detection and model ion selection (*m/z*).
5. Component candidate confirmation or rejection.
6. Spike elimination
7. Calculation of exact component retention time.
8. Spectra deconvolution using linear algebra.

The Components Detection and Spectra Deconvolution system works automatically. Use the system for broad types of chromatographic runs, for both GC/MS and LC/MC analyses, for clean and noisy signals, and for simple and more complex chromatograms. However, some parameter changes might be needed to optimize the system for specific applications. Use this automated procedure for small and medium size organic compounds and not for the processing of proteins, peptides, oligonucleotides, or other biomolecules.

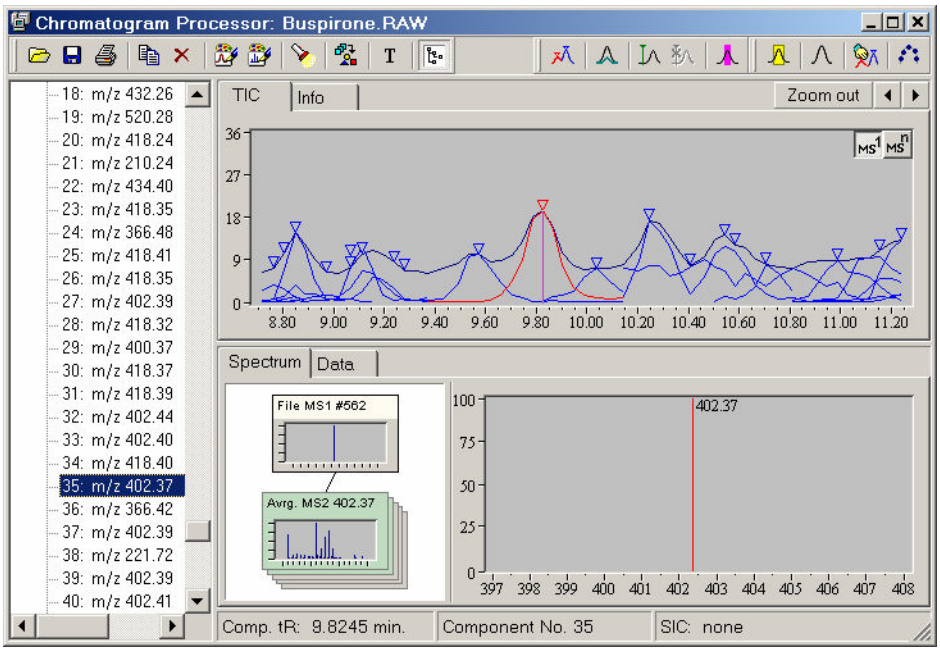


Figure 108. Deconvoluted data dependent chromatogram, showing spectral

Detected components are marked with a triangle in the original chromatogram. As the program calculates the precise retention time (tR) of each component, the component triangle might appear between scan points.

Note To display a deconvoluted spectrum of a component rather than of a scan, click the triangle that marks the detected component in the chromatogram pane. To display an original scan at the position of a detected component, click above the triangle.

If you select a component by clicking its triangle, a deconvoluted spectrum appears in the spectrum pane. This spectrum can be processed in the same way as any spectra in the program, for example, searched in a library or copied to the Database Manager.

If you move the mouse cursor over a component triangle, a tool tip informs you about the component number, precise retention time and the model ion *m/z* value that were used in the automated detection and deconvolution processes.

To copy or process more than one detected component in the Database Manager or in Spectra Classifier modules, you must first select them.

To select one or more components

Click the **Select Scans or Components** button and select the desired components with the mouse cursor while holding down the left mouse button.

To select all detected components in a run

1. Right-click anywhere in the chromatogram pane. A pop-up menu appears.

2. Click **Select All**.

The program makes a strict distinction between original scans and detected components and does not mix them. When selecting a chromatographic region, components are preferred to scans. If you select a chromatographic region that contains components, only these components are selected. If you select a region that does not contain any components, all scans in that region are selected.

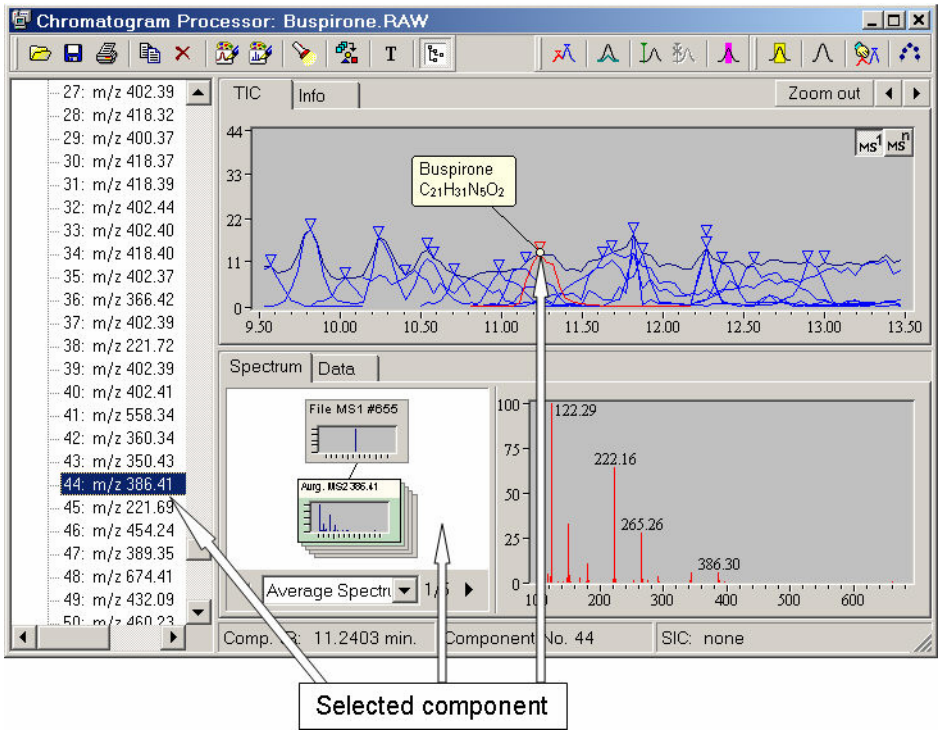


Figure 109. Chromatogram Processor window, showing items connected with selected component generated from LC/MS/MS data

RCD Algorithm

The Rapid Component Detection (RCD) algorithm is based on a "model ion" that represents a characteristic ion (*m/z* value in MS^1 spectra) for every detected component. The algorithm starts with component candidate detection and model ion selection and continues with the correlation of model ion profiles to confirm or reject a candidate.

To start automated component detection and the spectra deconvolution procedure using RCD algorithm

1. Click the **Components Detection and Spectra Deconvolution** button and choose RCD.

2. When the parameters setup dialog window appears, change the settings as needed, and then click **Calculate**.

The RCD algorithm dialog window contains various parameters that can be optimized for specific types of analysis. Note that these parameters are interdependent, so a change of one parameter can also affect algorithms linked with other parameters.

- Threshold of Total Signal

  The program uses a different threshold level from the one given in the data file. In especially noisy chromatograms, setting the threshold higher can reduce the number of false positive results. However, if the algorithm is too restrictive and is missing some chromatographic peaks, lower the default value.

- Minimum Model Ion Abundance

  The algorithms search for spectral peaks that have the most rapid rise and fall of signal in a peak region. This peak is called a model ion. To eliminate random fluctuations, a model ion must exhibit a minimum abundance value.

- Smoothing

  Use this option when noisy data needs to be analyzed. Mass Frontier automatically determines the smoothing factor according to the signal-to-noise ratio. You can change this value or switch off smoothing. The program uses an exponential filter similar to the analogue RC filter.

- Spectra Difference Factor

  To eliminate false positive component detection, the adjacent components must show some degree of spectral dissimilarity that is represented as the match factor used in library searching. The spectra Difference Factor value is the minimal match factor between spectra that detected components might exhibit.

- Background Subtraction

  If you choose the Automatic option, the program attempts to find a region of relatively constant signal intensity before and after every peak and sets two background scans there. If you want to use the Manual option, set the background scans by using the **Set Background Subtraction Scan** button before starting the detection and deconvolution process. Two manual background scans can be set anywhere in the chromatogram. If you choose **None**, background subtraction is not performed.

- Precursor Ion Subtraction

  If a product ion chromatogram is being deconvoluted, the selected precursor ion can be subtracted from all scans to improve component detection. However, if a component does not fragment and only a precursor ion can be observed, do not apply subtraction because this component might be overlooked. Note that Mass Frontier does not support scan events that are often used in connection with product ion scanning.

- Retention Time Range

  Detection and deconvolution calculations are time-consuming processes. The computing time needed depends on a number of factors, the most significant of which are the number of scans and the number of mixture peaks. To speed up your work, select only a part of a chromatogram to be analyzed. Other regions are ignored.

- Spectra Deconvolution

  Two extraction algorithms can be applied depending on the intended use of the component spectra. If components are intended for a library search, use Sharp spectra deconvolution. If the purpose of component detection is target analysis, use Soft deconvolution. Generally, Sharp deconvolution subtracts peaks from coeluting components with a higher multiplication factor and thus produces spectra with fewer peaks and lower intensities of isobaric peaks than Soft deconvolution.

JCD Algorithm

The Joint Component Detection (JCD) algorithm is based on the statistical analysis of all ion profile maxima. Ion profiles (ion chromatograms) with comparable shapes and maxima belonging to a limited time range are considered as a single component. The algorithm extracts individual mass spectral peak abundance profiles to produce a purified spectrum or spectral trees and generates the peak shape of a representative component. The JCD algorithm is recommended, but this requires significant computer processing resources.

To start automated component detection and the spectra deconvolution procedure using the JCD algorithm

1. Click the **Components Detection and Spectra Deconvolution** button and choose JCD.

2. When the parameters setup dialog window appears, change the settings as needed, and then click **Calculate**.

The RCD algorithm dialog window contains the following parameters that can be optimized for specific types of analysis. Note that these parameters are interdependent, so a change of one parameter might also affect algorithms linked with other parameters.

- Mass Merge Power

  Specifies the mass difference within which the algorithm merges spectral peaks into one *m/z* value. A low value might result in more components (oscillating ions) and a high value can result in fewer components (merging ions).

- Average Peak Width

  Specifies the chromatographic peak width in scans that the algorithm uses to identify a potential component. If a value is too high, this can result in the loss of narrow peaks. If a value is too low, it might split a real component into two different components.

- Baseline Correction

  If checked, an automated baseline correction for each ion profile is applied using the Top-Hat algorithm. Use this option only for chromatograms with an elevated baseline.

- Smoothing

  If checked, the smoothing of each ion profile is performed using the Average Peak Width value and the Loess algorithm

- Noise Modification

  Determines how to adjust the intensity values of spectral peaks where abundances are comparable to noise level. Choose one of the following methods:

  – None

    Spectral peak intensities are not altered

  – Elimination

    Spectral peaks with intensities lower than the specified value are eliminated. Use this method if low abundant peaks are not of interest.

  – Transition (default)

    Artificial noise is added to replace random spikes with constant noise for better detection of low abundant components.

- Analyze MS Stages

  Determines which MS stages are considered in the analysis of ion profiles and for detecting components. The remaining stages are used only for spectral tree reconstruction. The choices are as follows:

  – Top Stages (default)

    Analyzes only the top-stage ions present in the data, where top means MS^1, or MS^2 if MS^1 is not present. The algorithm builds the component tree by joining the corresponding lower stage spectra that meet the following criteria: they occur within the component envelope and the software detected the precursor *m/z* in the top stage as a component ion. (If the deconvoluted MS^1 spectrum contains peaks that have been further isolated, the corresponding MS^2 spectra are assigned to the spectral tree)

  – Lower Stages

    Analyzes all ions except those from the top MS stages. The algorithm calculates the spectrum of the top stage from the total ion abundance of the top stage. The resulting spectral trees do not have as much depth as do those from the Top Stages option.

– All Stages

  Analyzes all ions regardless of the MS stage. The resulting spectral trees do not have as much depth as do those from the Top Stages option.

Use the Top Stages level in most cases. Use the Lower Stages level if a Data Dependent experiment is set in such a way that ions are isolated according to a predefined list of *m/z* values and top-stage spectra are noisy. Use the All Stages level for the preliminary analysis of complex data.

- Eliminate non-model ion stages

  If checked, the product ion spectrum along with subsequent stages is deleted if its precursor ion is not a model ion (usually the most intense) in the top stage.

- Retention Time Range

  Enables you to specify only a part of a chromatogram to be analyzed (time range), which can speed up your work. In this case, the algorithm ignores regions outside the specified range.

- *m/z* Range

  Enables you to specify only a part of a chromatogram to be analyzed (*m/z* range), which can speed up your work. In this case, the algorithm ignores regions outside of the specified range.

- Baseline Threshold

  Specifies the peak baseline as a percentage of base peak height.

- Merging Factor

  Specifies a limit for the time difference of the ion profiles maxima. Too high a value can cause the merging of randomly coeluting components. Too low a value can split a component into more false-positive components.

- Sharpness Tolerance

  Specifies the degree (percent) of similarity of ion profile shapes. If two ion shapes meet the specified percentage for the sharpness tolerance, as well as other parameters, the algorithm merges the ions into a single component.

- Wide Component Merge Mode

  Specifies whether a limit for the time difference of the ion profile maxima is used in a combination of the following parameters: Average Peak Width and Merging Factor. Select this box (default) to avoid splitting a component into ion peaks that are detected as redundant components in chromatograms with wide peaks. Clear this box if an incorrect component merge occurs.

TECD Algorithm

The Total Extraction Component Detection (TECD) algorithm creates spectral trees for every section of a chromatogram that starts with an MS^1 scan. Each generated tree is then divided into subtrees based on the value of the Minimal Tree Depth parameter. Generated subtrees are merged into components based on the precursor ion *m/z* value identity and the spectral tree match factor similarity.

To start automated component detection and the spectra deconvolution procedure using the TECD algorithm

1. Click the **Components Detection and Spectra Deconvolution** button and choose TECD.

2. When the parameters setup dialog window appears, change the settings as needed, and then click **Calculate**.

The TECD algorithm dialog window contains the following parameters that can be optimized for specific types of analysis. Note that these parameters are interdependent, so a change of one parameter can also affect algorithms linked with other parameters.

- Minimal Tree Depth

  Specifies the minimum number of tree sections the algorithm creates from the initial spectral tree. The value determines the MS stage where a division takes place.

- Tree Match Factor

  Specifies the minimum percentage that two spectral trees within adjacent tree sections must match before the algorithm considers the two spectral trees as the same component. Matching spectral trees are defined as having identical precursors up to the level specified by the Minimal Tree Depth value and a Tree Match Factor value that exceeds the specified value.

- Wide Component Merge Mode

  Enables a comparison of the spectral trees for potential matching and merging, not only in adjacent sections, but also in sections up to the distance specified by the Allowed Gap value.

- Allowed Gap

  Specifies the maximum distance between nonadjacent tree sections over which to compare the spectral trees for potential merging.

- Retention Time Range

  Enables you to specify only a part of a chromatogram to be analyzed (time range), which can speed up your work. In this case, the algorithm ignores regions outside of the specified range.

- *m/z* Range

  Enables you to specify only a part of a chromatogram to be analyzed (*m/z* range), which can speed up your work. In this case, the algorithm ignores regions outside of the specified range.

Direct Infusion Algorithm

The Direct Infusion algorithm is a spectral tree construction utility rather than a component detection method. It creates one or more spectral trees from a single raw file by reading of various MS^n scans acquired in one run and constructing a tree according to their MS stage and precursor ion *m/z* values.

To start the spectral tree construction utility using the Direct Infusion algorithm

1. Click the **Components Detection and Spectra Deconvolution** button and choose Direct Infusion.

2. When the parameters setup dialog window appears, change the settings as needed, and then click **Calculate.**

The Direct Infusion algorithm dialog window contains various parameters that can be optimized for specific types of analysis. Note that these parameters are interdependent, so a change of one parameter can also affect algorithms linked with other parameters.

- Minimal Tree Depth

  Specifies the MS stage that has only a single precursor. The value determines the MS stage where a tree branch division takes place. If the default value 1 is used, the algorithm creates a single tree from all the scans in a run (file). If the value 2 is used and the analyzed run contains MS^2 spectra with different precursor *m/z* values, the algorithm creates separate trees for each MS^2 scan with a unique precursor ion *m/z* value.

- Include Upper Spectra

  If checked, the resulting spectral trees contain actual scans in the stage above the level set in Minimal Tree Depth. If unchecked, the spectra above the stage set in Minimal Tree Depth. contain spectra with a single peak equal to the precursor ion of the product spectra.

- Average Scan

  If checked, every tree node contains only average spectra.

- Calculate Envelope

  If checked, the ion profile (envelope) for each tree is calculated. Use to preview scans that belongs to a particular tree.

Processing Xcalibur MS^n Data

A spectrum obtained from the average of scans, from scans with a subtracted background or from component deconvolution can be further processed in the Database Manager, Spectra Classifier, or can be directly searched for in a library from Chromatogram Processor. To transfer the extracted spectrum or tree to the Database Manager or Spectra Classifier, use the **Copy** button. After clicking the **Copy** button in the Chromatogram Processor window, paste the mass spectrum to a Database Manager or Spectra Classifier window. In addition, when you click the **Copy** button in the Chromatogram Processor, the chromatogram graphic and the spectrum from the Scan tab are copied to the Windows Clipboard and can be used for creating reports in any Windows application.

The Chromatogram Processor window contains a **Search** button to directly search spectrum from the Scan tab in libraries. To search the extracted spectrum in libraries, click the **Search** button in the Chromatogram Processor window and the Spectra Search dialog window appears.

Classify mass spectra or spectral trees obtained in the Chromatogram Processor by using the Principal Component Analysis, Neural Networks, and Fuzzy Clustering methods.

To choose scans or deconvoluted spectra for classification

Select scans or components that you want to classify and click the **Add Selected Scans or Components to Spectra Classifier**  button.

Use the Chromatogram Processor to view and process Xcalibur Data Dependent experiments and product ion scanning .raw files and extract spectral MS^n trees from chromatograms. If Xcalibur data is opened in Chromatogram Processor, a tree view control appears on the right side of the window. Use this tree view control to select one or more product scans at any MS^n stage. If deconvoluted components are present, they are listed at the bottom of the tree control.

If component detection and spectra deconvolution procedures are applied to data-dependent chromatograms, the program generates components as spectral trees. The tree components can be processed in the same way as regular spectra. They can be edited in Components Editor, searched in spectral libraries or classified by using the Spectra Classifier. Spectral trees in the Chromatogram Processor can only be viewed. To edit them, copy and paste them into the Database Manager, where all editing utilities are available.

Note that the spectrum in the bottom right of the Chromatogram Processor window is displayed for the selected tree node spectrum.

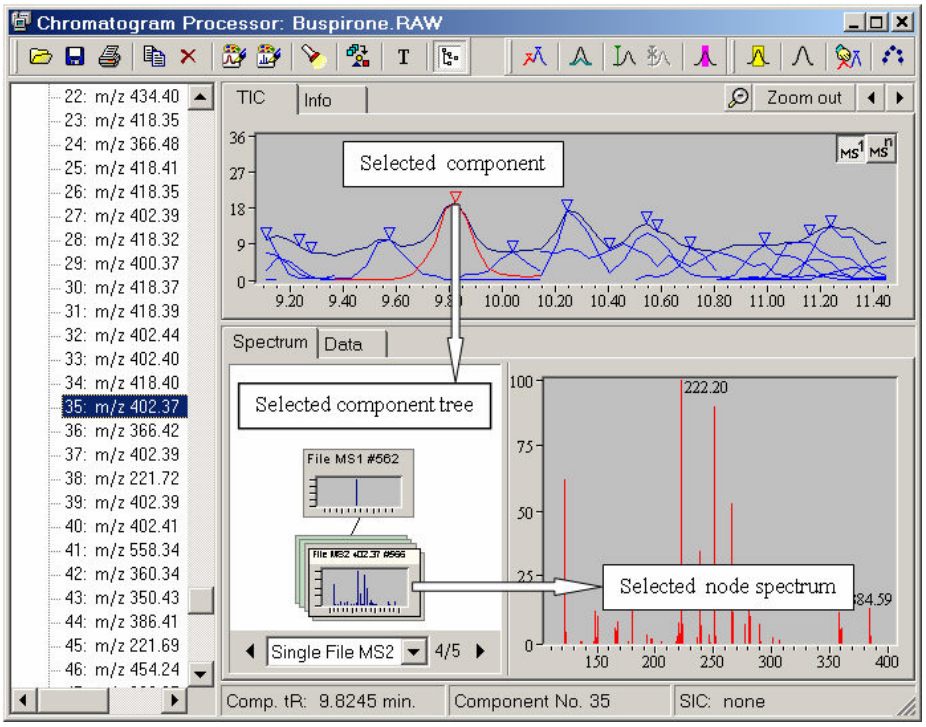


Figure 110. Chromatogram Processor window showing selection cascade for deconvoluted spectral tree

Chapter 14 Components Editor

Use the Components Editor to edit, search and organize chromatographic components stored in a library or generated by the Components Detection and Spectra Deconvolution features. This module has a complete set of management tools to delete unrelated components, add chemical structures, edit extensive data fields, process spectral trees, annotate spectral peaks, and sort search match lists for every component in a processed chromatogram. The Components Editor modal window is accessible from the Database Manager or Chromatogram Processor window. The modal window is a program module that pops up over an applications frame window. When the modal window is present, no other application window can be used. To close this module, click either the **OK** or **Cancel** button.

The Components Editor closely resembles visually the Database Manager module; however, the processing item is a chromatographic component rather than a database record. To make the modules easy to distinguish, the Components Editor has a light blue bar on the left site of the window. Both modules handle almost identically. See Chapter 4, "Database Manager," for an explanation of module functions.

Each chromatographic component is represented by a single row. The columns contain supplementary component information. One of the columns lists Model Ion values that have been used for component detection. These values help you to orient yourself and find components of interest. In most cases, the model ion is the base peak in the full scan spectrum; however, if closely coeluting components have isobaric base peaks, the algorithms select different model ions to distinguish the components.

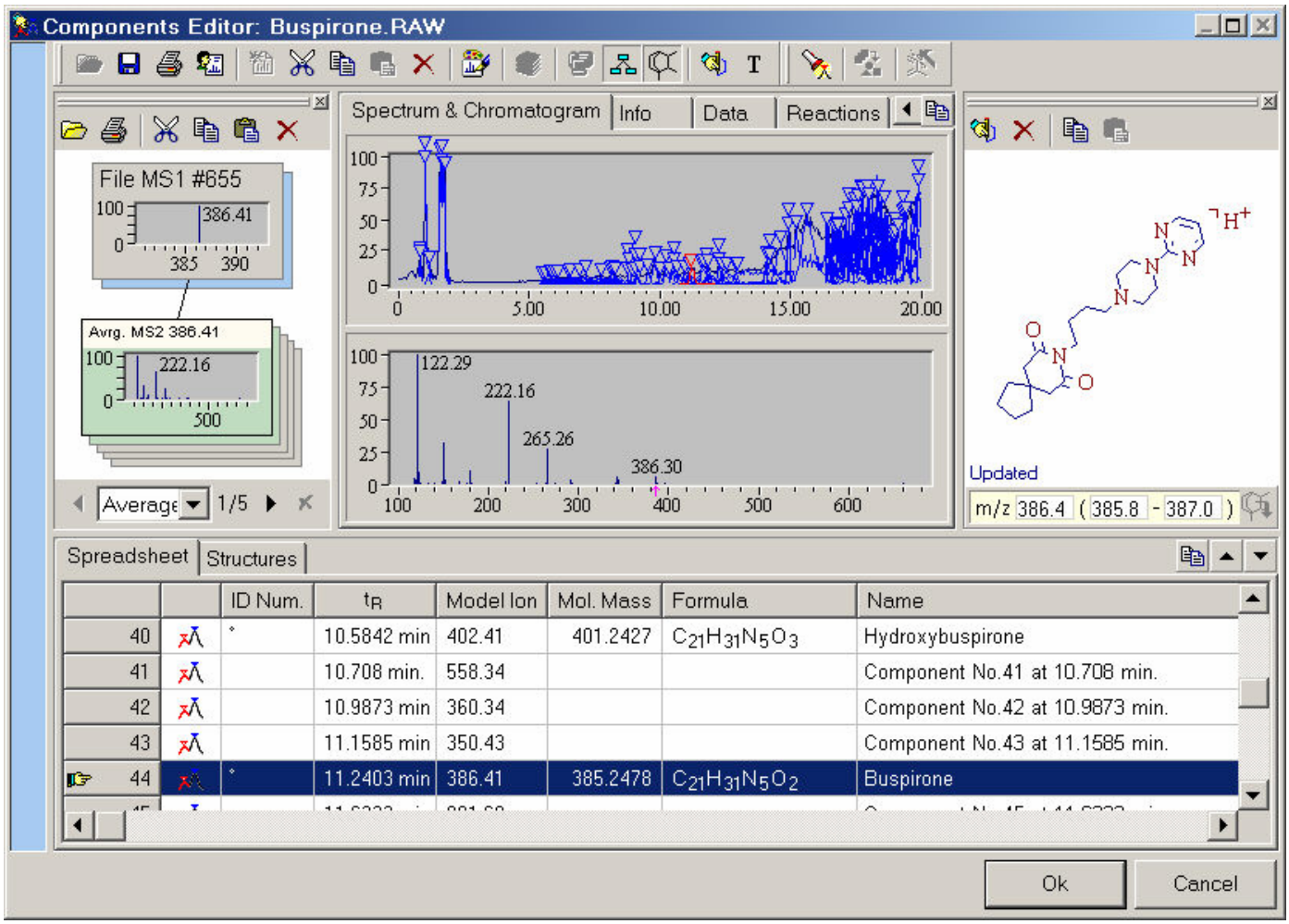


Figure 111. Components Editor window

Searching Components

Use the Components Editor to search one selected subset or all the chromatographic components in the libraries at once. To search selected components, select components in the spreadsheet and click the **Search** button in the tool bar and then choose the **Search Selected Components** item from the pop-up menu. To search all the components from the processed chromatogram, click the **Search** button and choose **Search All Components**.

Note Searching a large number of components might take a considerable amount of time depending on the library size.

After the search is completed, the highest match factor for each component search is listed in the Match column in Spreadsheet. If the search procedure did not lead to any plausible matches, the match factor is not displayed. If at least one match was found, the text in the name filed of the component row appears in red.

To process the match list of a component, select the component row and open the Hit Selector window by clicking the **Hit Selector** button. In the Hit Selector, you can review the match list and accept a library record that correspond to the component by selecting the match and clicking **OK**. If you accept a library record for a component, all relevant information (structure, name, molecular mass, ion types and so on) is adopted and entered in the component fields. If you decide to reject the match list, click **Cancel**. You can undo an accepted match by clicking the **Reject Library Hit** button that appears next to the **Hit Selector** button.

Note The Hit Selector window lists the best matches found during the library search. The match (hit) factor describes the exactness of the match to your component.

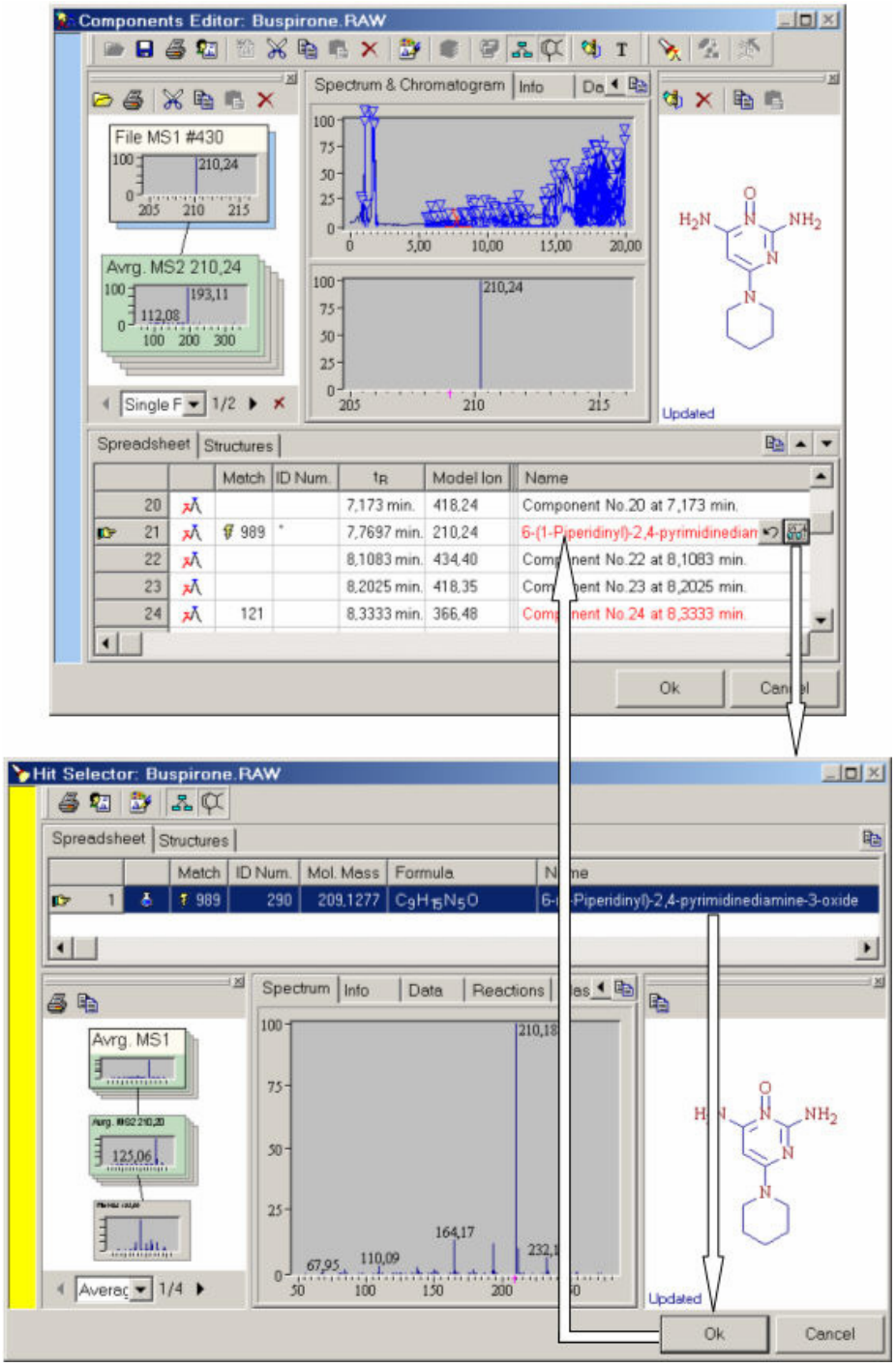


Figure 112. Components Editor and Hit Selector windows showing assignment of a library search match to a chromatographic component

Chapter 15 Mass Settings

All calculated masses displayed in Mass Frontier are monoisotopic masses. The monoisotopic mass of an ion is the mass of the isotopic peak which is composed of the most abundant isotopes of its elements. Because the program's calculations support only single-charged ions (z = 1), the calculated—not measured—exact masses in structure-based modules are equal to their *m/z* value.

This chapter contains the following sections:

Resolution

Resolution settings are used for the differentiation of adjacent peaks (spectra) and *m/z* values (calculations). Because Mass Frontier processes only centroid spectra, the differentiation is defined here as the spacing between resolved peaks (spectra) or as the smallest difference in *m/z* values that is distinguished (calculations). No additional parameters, such as peak width, are taken into account.

Mass frontier supports the following resolution types:

- Unit Resolution
- Resolving power M/ΔM
- Mass Resolution ΔM

where,

M = *m/z* (mass-to-charge ratio)

$\Delta M = M_2 - M_1$ (M_1, M_2 are two adjacent peaks)

Peaks (*m/z* values) that fall into the ΔM band are merged into a single peak (*m/z* values).

To change resolution settings

Choose **Option > Mass Settings** from the main menu.

To set the Resolution setting for all calculations in Mass Frontier, you must define Resolution type in the **User Defined Type** box and set the appropriate value and unit in the **Unit** and **Value** boxes.

For imported spectra or GC/LC/MS chromatograms from a file, there are two options for the setting of the resolution (**Resolution & Precision of Experimental Spectra** box). You can either adopt the resolution saved in a file (**Acquired from Source** box), or set your own resolution settings to be used for calculations (**User Defined Type** box). Note that it is not possible to improve the resolution of experimental data by using User Defined Type. Do not set resolution values that are better than the resolution of the mass spectrometer actually used for data acquisition. Use this option to artificially reduce the resolution of experimental data when working with low- and high-resolution spectra in the same data set (spectra search, target analysis, or classification).

Note Changing the Resolution in Mass Settings dialog window does not affect spectra that are opened (with the Chromatogram Processor) and fragments that have already been calculated with the (Fragments & Mechanisms and Fragments Comparator modules). If you require spectra or fragments with a new setting, you must reload these spectra or regenerate these fragments.

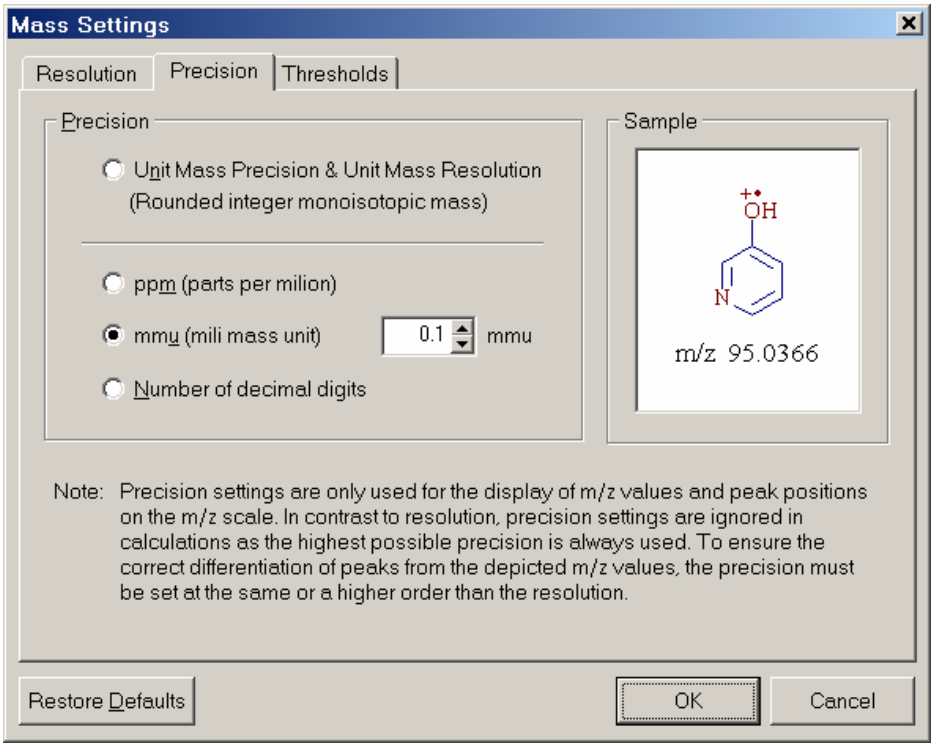


Figure 113. Resolution page of the Mass Settings dialog window

Precision

Precision refers to *m/z* values and is defined as the position of the last digit relative to the decimal point that is displayed. Precision settings are used only for the display of *m/z* values and peak positions on the *m/z* scale. In contrast to resolution, precision settings are ignored in all calculations as the highest possible precision is always used. To assure the correct display of *m/z* values, set the precision to the same or a higher order than the resolution.

Note When working with experimental spectra, some of the digits of the displayed *m/z* values might not be significant.

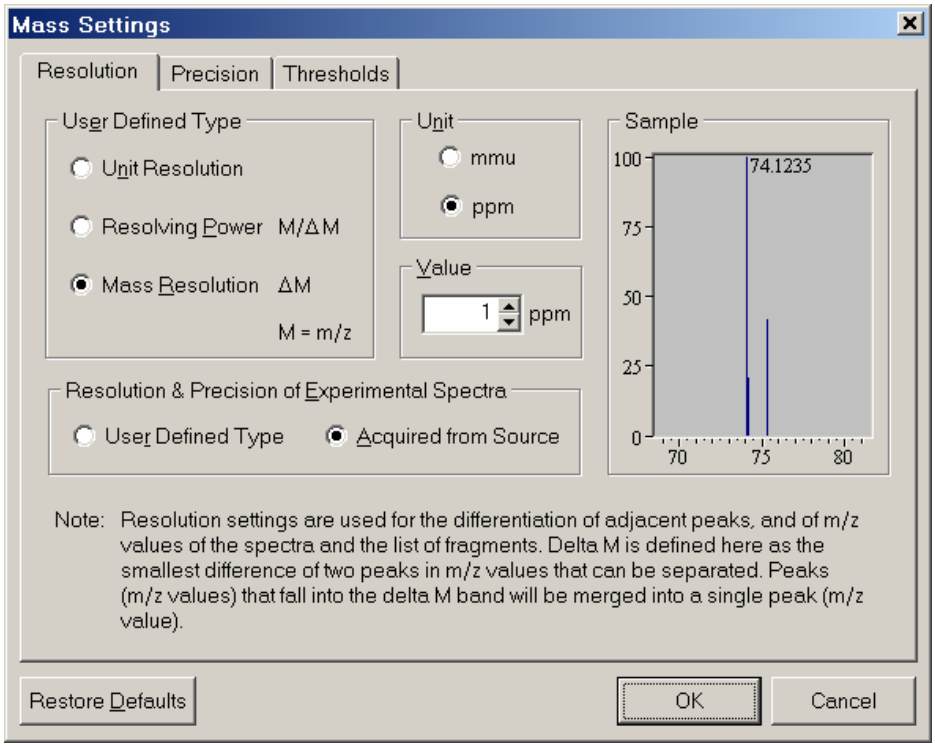


Figure 114. Precision page of the Mass Settings dialog window

Chapter 16 Microsoft Office in Mass Frontier

Mass Frontier supports the Object Linking and Embedding (OLE) features of Microsoft Excel and Word documents; that is, these programs and their native data files can be opened inside the Mass Frontier desktop. You do not need to leave Mass Frontier to work with Excel and Word documents because they become part of Mass Frontier. Use this feature with embedded Excel documents for data exchange between Mass Frontier modules and Excel worksheets.

Note The Mass Frontier installation package does not include the Microsoft Office software. In order to use the Object Linking and Embedding features of Excel and Word, these programs must be purchased separately. Install Microsoft Office prior to installing Mass Frontier.

This chapter contains the following sections:

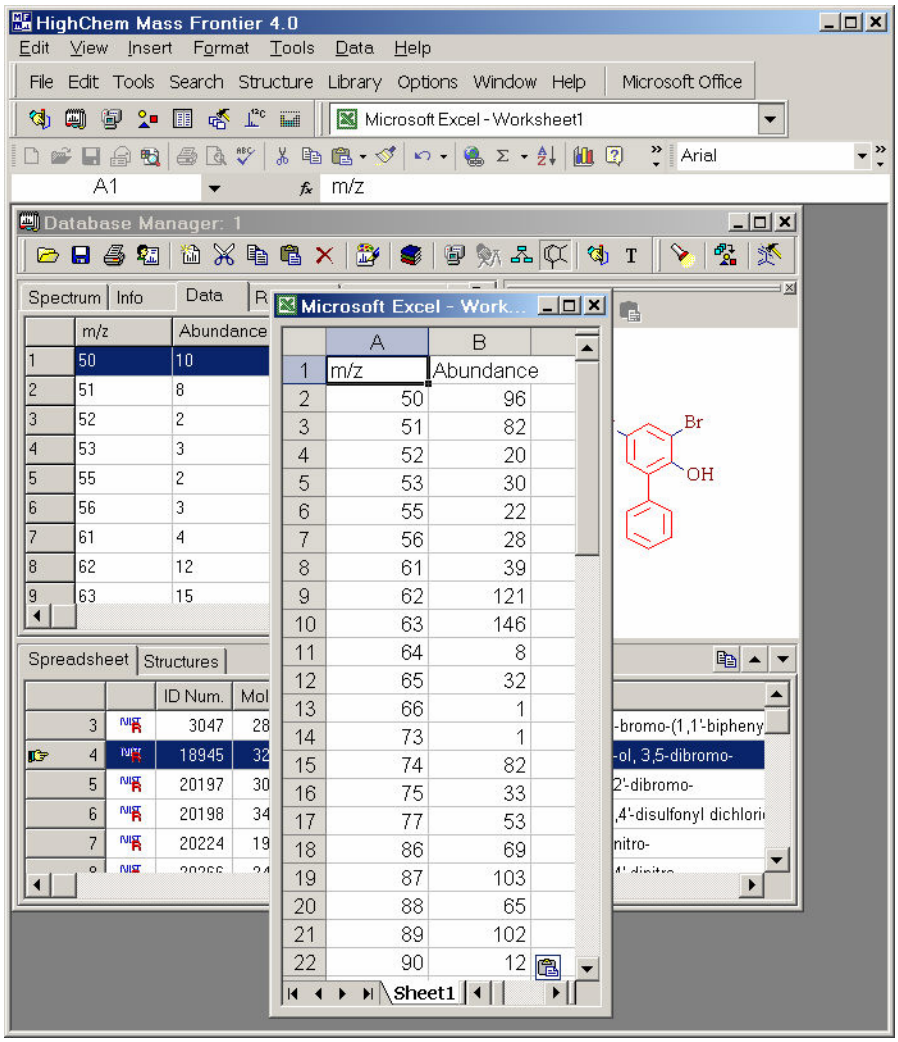


Figure 115. An Excel worksheet embedded into Mass Frontier

Data Exchange between Excel and Mass Frontier

There are three basic types of data you can exchange between Excel and Mass Frontier: text tables, graphics, and mass spectra in table format. The data can be exchanged by using the copy and paste commands. Mass Frontier does not support the import or export of native Excel .xls files directly into the Database Manager. However, Excel documents can be embedded into Mass Frontier and data can be exchanged by using the Clipboard. Choose **Microsoft Office > Open Microsoft Excel Document.** In this case, a new window opens and the Excel menu and toolbar appear.

Note Embedded Excel or Word windows have their own main menus (displayed above the Mass Frontier menus) and buttons (displayed below Mass Frontier buttons). Microsoft Office controls (menus and buttons) are visible only if an Excel or Word window is active. These controls do not contain Open and Save commands. If you want to open or save an Office document, you must choose **Microsoft Office** on the Mass Frontier main menu.

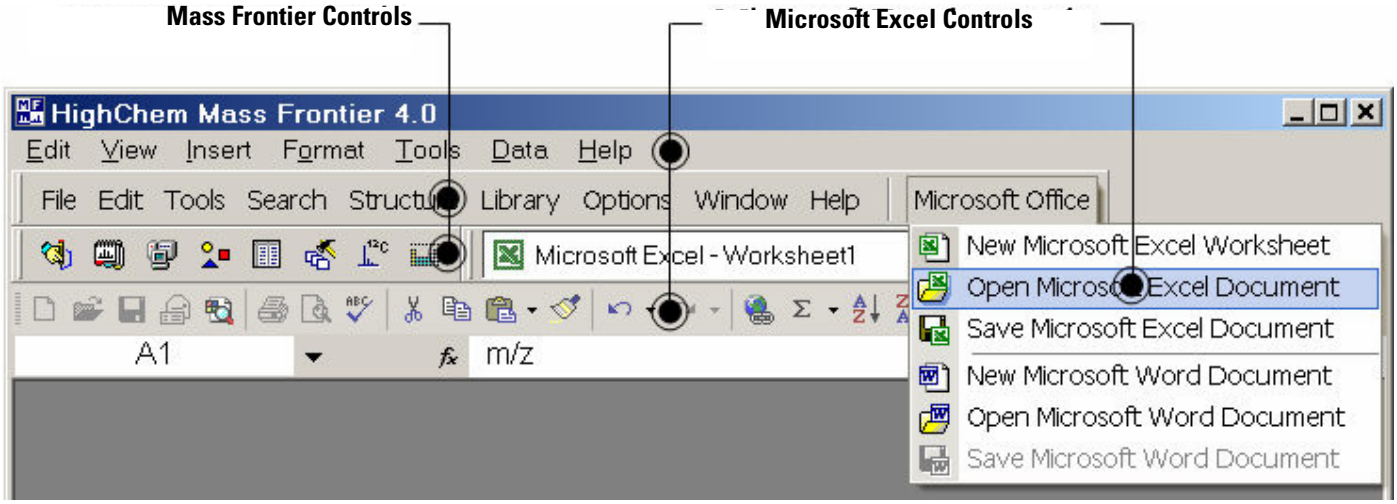


Figure 116. Locations of Excel and Mass Frontier controls

Exporting Data to Excel

You can export four types of data from the Database Manager to Excel: text tables, graphics, mass spectra in numerical table format, and spectral trees. Because all these types can be accessed at the same time, you must specify the type of data you want to export to Excel.

To export additional information associated with a record

1. Choose the Info tab (Info tab data must be visible).
2. Click the small **Copy Selected Tab** button in the top-right corner of the Info tab control.
3. Paste the data by clicking the **Paste** button in Excel.

To export mass spectrum or mass differences graphics

1. Choose the Mass Spectrum or Mass Differences tab (a Mass Spectrum or Mass Differences spectrum must be visible).
2. Click the small **Copy Selected Tab** button in the top-right corner of the Mass Spectrum tab control.
3. Paste the graphic into Excel by clicking the **Paste** button.

To export a mass spectrum in numerical table format

1. Choose the Data tab (an *m/z* and abundance table must be visible).
2. Click the small **Copy Selected Tab** button in the top-right corner of the Data tab control.
3. Paste the table into Excel by clicking the **Paste** button.

To export records data in the spreadsheet

1. Select one or more records in the spreadsheet.
2. Click the **Copy Selected Rows** button to the right of the Spreadsheet tab control.
3. Paste the data into Excel by clicking the **Paste** button.

Importing Spectra from Excel

Mass spectra in table format or spectral trees stored in Excel can be imported to the Database Manager by using the Clipboard. Spectral tables can be organized horizontally or vertically. In order to correctly interpret *m/z* values and abundance, follow one of these conventions:

1. If the spectral table is vertical, the first column must be the *m/z* value and the second must be abundance.

2. If the spectra are oriented vertically and the first column is abundance and the second column is the *m/z* value, caption the first row of the first column Abundance and the first row of the second column *m/z*.

3. If the spectral table is oriented horizontally, the first row must be the *m/z* value and the second row must be Abundance.

4. If your spectra are oriented horizontally and your first row is abundance and the second row is the *m/z* value, caption the first column of the first row Abundance and the first column of the second row *m/z*.

5. More than one spectrum can be imported at a time. In this case, your first column or row, depending on the orientation, must be the *m/z* values and all other columns (or rows) must be Abundance.

To import spectra from Excel to Mass Frontier

1. Verify that your table actually contains mass spectra.

2. Select the table you want to export to Mass Frontier.

3. Click the **Copy** button in Excel or, if your document is embedded, click the **Copy** button on the Excel toolbar.

4. Click the **Paste** button in the Spectra Manager window.

Mass Frontier supports standard tables with separated numbers, so spectra can also be imported from other programs.

Using Excel as Spectrum Editor

Excel can be used as a spectrum editor for Mass Frontier by using the export and import features. Use this tool if, for example, you have an experimental spectrum with a number of noise peaks at high *m/z* values that you want to delete, or you want to extract part of a spectrum that is important to your report or presentation. Do not add, delete, or alter prominent peaks.

Chapter 17 Formula Generator

Use the Formula Generator to calculate a list of theoretical molecular formulas that best fit an *m/z* value. Because the number of possible molecular formulas for a given *m/z* value is closely related to the mass tolerance, elements used, and the maximum allowed number of atoms for each isotope, always carefully evaluate these parameters before a generation starts. The *m/z* value can either be manually entered or automatically taken from any spectral peak in the Database Manager or the Chromatogram Processor. You can also use the Formula Generator to calculate the isotopic pattern for any generated formula.

This chapter contains the following sections:

To open a Formula Generator window

- Click the **Formula Generator** button on the main tool bar.
- Or choose **Tools > Formula Generator** from the main menu.

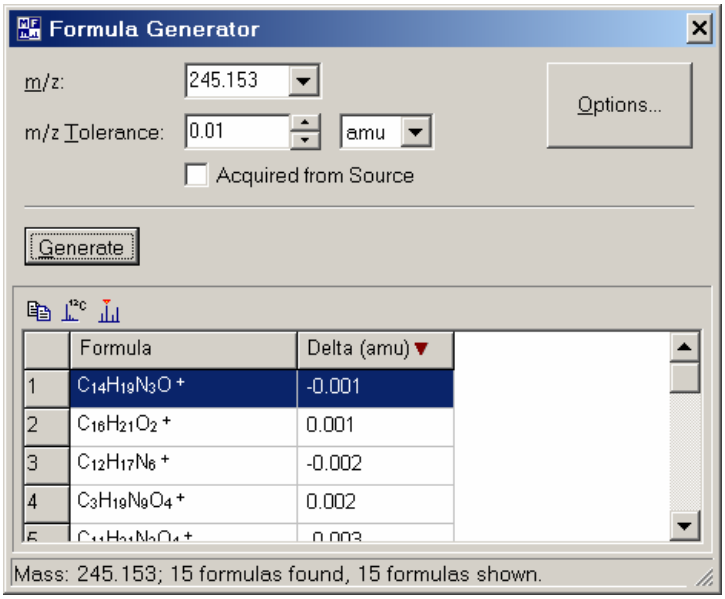


Figure 117. Formula Generator window

Formula Generation from a Peak

With Mass Frontier, you can calculate the possible elemental composition for any spectral peak displayed in the Database Manager and Chromatogram Processor windows by simple peak picking. The advantage of this method compared to manual entering of the *m/z* value is that all the decimal places are accurately transferred to Formula Generator and the *m/z* tolerance value is automatically calculated from the source spectra if this option is enabled in the Acquired from Source box.

To generate possible formulas for a peak

1. Click the **Option** button and review the predefined settings. Give special attention to the Charge setting (Limits tab) according to the polarity of the MS mode used for spectra acquisition. Be sure to check the maximum number of elements used for calculation in the Elements in Use tab.

2. Click the **Pick a Peak**  button. The Formula Generator window minimizes to a bar.

3. Click on a spectral peak in the Database Manager or the Chromatogram Processor window.

   A list of possible formulas appears.

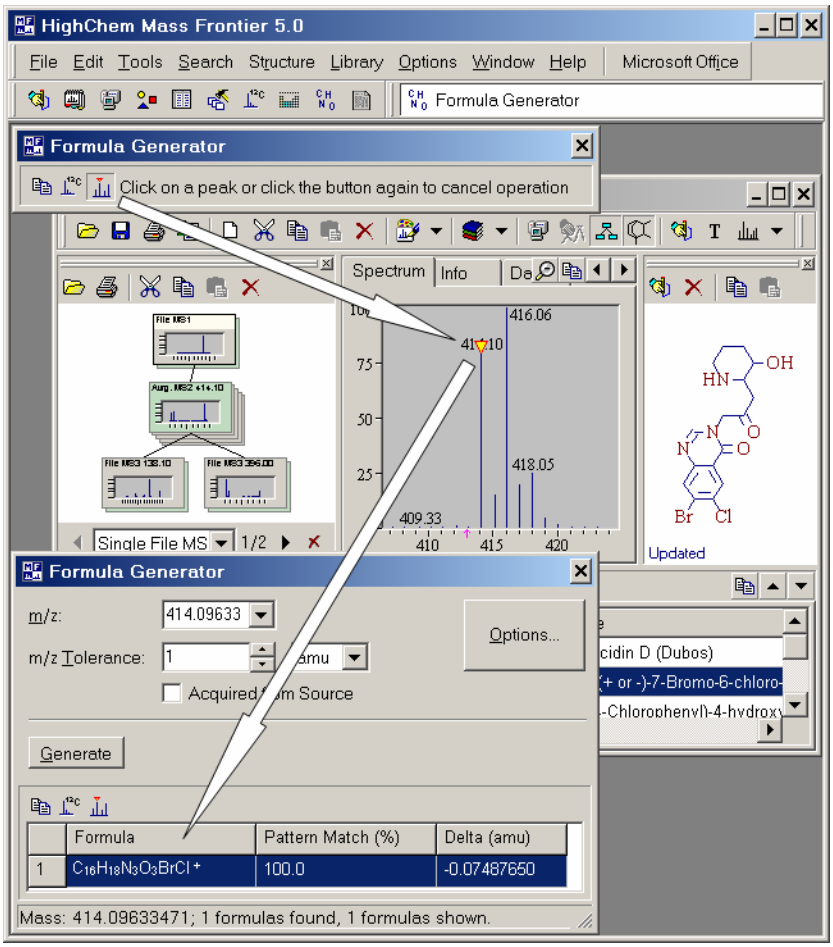


Figure 118. Formula Generator window

Formula Generator Options

Use Formula Generator options to speed up calculations, restrict the number of possible formulas, and set the charge state and polarity.

Available parameter settings include the following:

- Charge

  Charge state and polarity.

- Nitrogen rule

  Decide whether to use the nitrogen rule in the elemental composition calculation. If yes, the choices include even-electron (radical cation) or odd-electron (protonated) ion.

- Hydrogen Count

  Use to specify whether to use an algorithm for the exclusion of implausible formulas with improbable high numbers of hydrogen.

- Valence Test

  Use to specify whether to exclude formulas if the atoms cannot be connected in any way using valences common in organic chemistry.

- Ring plus Double Bond Equivalent (RDBE)

  Only formulas are displayed for which the ring plus a double bond equivalent is within the range *From* – *To*. The RDBE limits the calculated formulas to those which make sense chemically.

- Elements in Use

  Use this tab to specify isotopes of particular elements with a maximum number that is considered for the calculation of formulas.

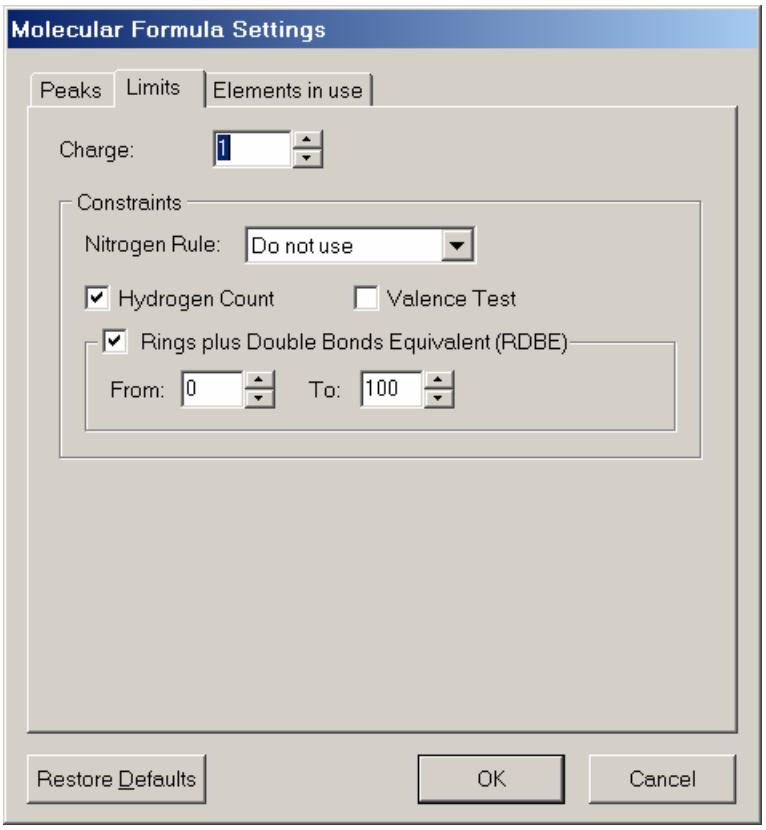


Figure 119. Molecular Formula Settings window

Chapter 18 Report Creator

Use the Report Creator to create customizable reports. You can create reports by extracting any information from the modules on the screen. Both text and graphics can be extracted and you can specify the order in which objects from the module appear in the report. Reports can be created from the following modules: Structure Editor, Database Manager, Fragmentation Library, Chromatogram Processor, Spectra Projector, Neural Networks, and Fragments & Mechanisms. Reports can be printed or exported to PDF files and the report layout saved.

This chapter contains the following sections:

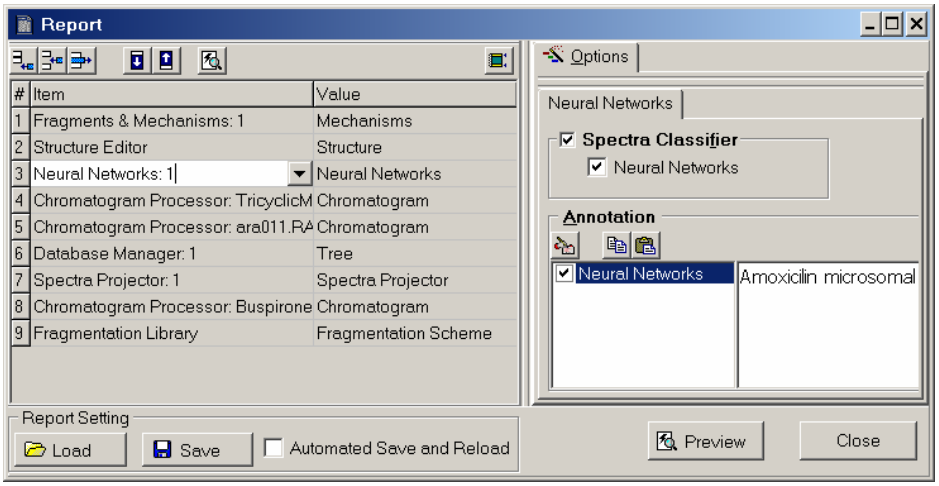


Figure 120. Report Creator window

Report Creator Window

The Report Creator window is divided into two parts. See Figure 121. Use the left part to manage the modules from which you want to create a report. Select a module by highlighting its name and its corresponding customizable items are displayed on the right side. Use the right part to select and annotate a number of text or graphical objects for the selected module.

Use the left part of the Report Creator window to add or remove a module from reporting or to change the order in which they appear in the report.

Note A module can be listed only once on the left side of the Report Creator.

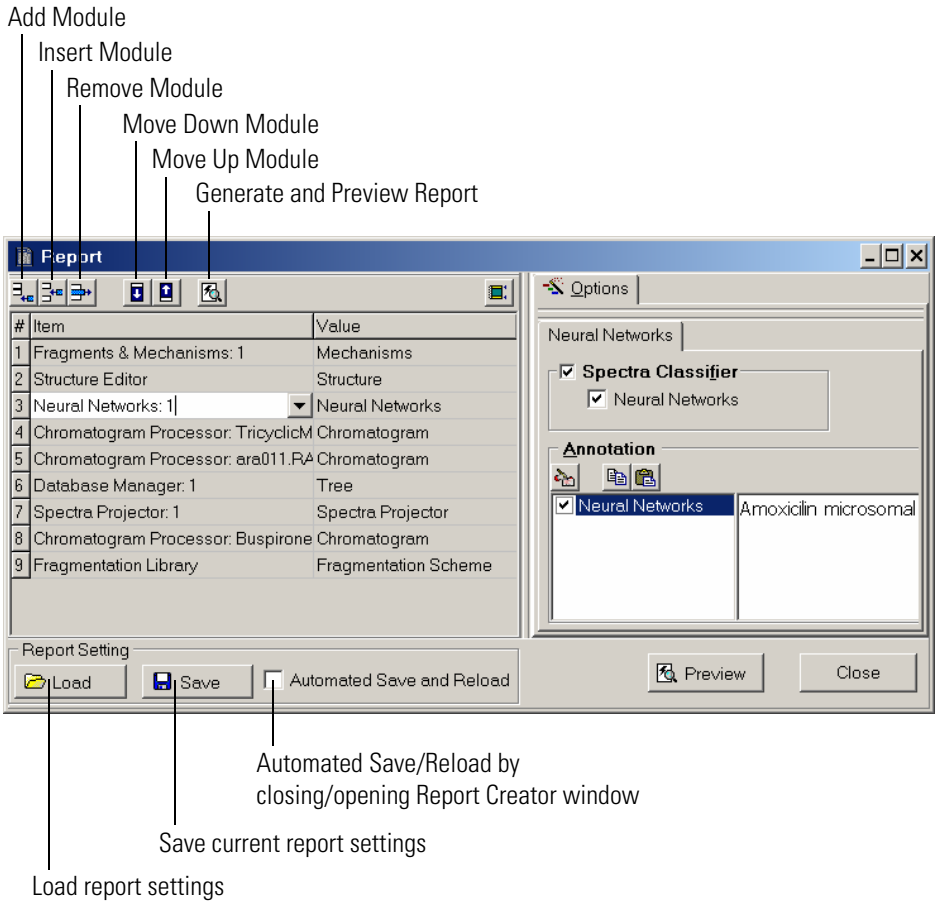


Figure 121. Report Creator window

Creating Reports

Create Reports by using the Report Creator only from windows that are open. You cannot report data that is stored in Mass Frontier but does not appear on the screen when Report Creator is launched.

To create a report

1. Open one or more modules with the corresponding data you want to report.

2. Click the **Report Creator** button on the main tool bar, or choose **Tools > Report Creator** from the main menu

When the Report Creator opens, all modules available for reporting are listed on the left side. If you click on a module name, options and annotation fields specific to the selected module appear on the right side.

To change the general report settings such as header, footer, separation lines, page breaks, or orientation, click the **Report Layout** button in the Report Creator window.

Once the modules and their objects have been selected and annotated, generate a report preview by clicking the **Preview** button. See Figure 122.

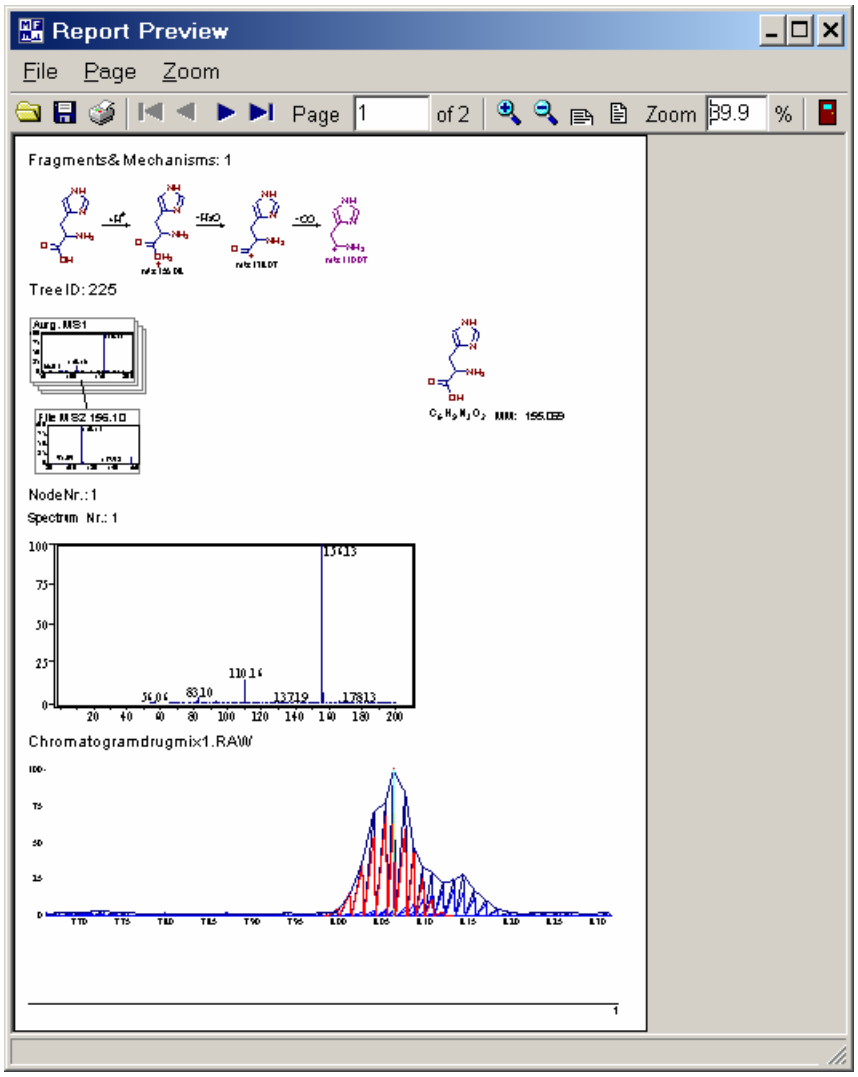


Figure 122. Report Preview window

Index

I

J

K

L

M

N

O

P

R

S

T

X