

# What's New in MS Search Program

The basic software of the NIST MS Search Program for Windows Version 2.2 is unchanged.

**The following features were added to the current version 2.2, distributed with NIST 14 as compared to ver. 2.0g, distributed with NIST 11 or with NIST 12 MS/MS library.**

## *Features Relevant to Electron Ionization Mass Spectra (Low Resolution)*

Using [Retention Index in library searching](#): RI deviation of a hit from the search spectrum is used to penalize the Match Factor.

[Indexing older user libraries to use saved in them RI values](#) for library searching

Displaying relevant RI values in the hit list

[Import](#), export and [editing](#) compound RI

New definitions of derivatives are used for finding replicate spectra

[Displaying derivative precursor structures](#) in the Lib. Search hit list

## *Features Relevant to All Mass Spectra*

All small molecule mass spectra in NIST libraries have InChIKeys linked to PubChem. These InChIKeys may also be used for [searching compound information on the Web or for exact structure search in libraries](#).

Importing and displaying Glycan structures in KCF format; adding them to mass spectra in a user library

## *Features Relevant to Tandem and High Resolution Mass Spectra*

Displaying additional MS/MS information in hit lists and on spectrum plot

Editing accurate peak m/z, intensities, and peak annotations in the built-in [mass spectrum editor](#)

In-source/EI with accurate ion m/z spectra - High resolution mass spectra

- Import of [‘in-source’ tandem spectra](#), or [EI with accurate ion m/z spectra](#), which do not include a defined m/z of the precursor ion. To import these spectra, select "in-source / EI (accurate m/z)" in [Spectrum Import Options](#).
- In-source/EI with accurate ion m/z spectra have accurate ion peak m/z and intensities. The accuracy is set in [Spectrum Import Options](#).
- [In-source/EI with accurate ion m/z](#) spectra may be searched in In-source/EI with accurate ion m/z library and/or MS/MS library with the new [In-source HiRes search](#) (ion peak m/z tolerance down to 0.015 ppm or  $6 \times 10^{-5}$  m/z units), Similarity Simple, Identity Normal or MS/MS Presearch OFF [search options](#). For this purpose, adding [Reverse Search](#) option may be useful.
- In-source/EI with accurate ion m/z spectra may be [added to a user library](#). Currently, in-source/EI with accurate ion m/z spectra added with NIST MS Search to a user library may be searched with In-source HiRes search only with Presearch OFF option. Rebuild the library with Lib2NIST to get in-source HiRes presearch benefits.

**A number of new features have been added to the version distributed with the 2012 release of NIST MS/MS Libraries.**

A new MS/MS search for small molecules

- Improved ability to identify compounds with a few dominant peaks – a common situation for Tandem MS.
- More robust scoring for spectra acquired on imperfectly-tuned instruments

New Mass Spectrum Import Options

- Control the number of decimal places in the input m/z
- Set peak intensity threshold
- Set Prepend/Overwrite/Ask automatic mass spectrum import option
- Import [in-source or EI spectrum with accurate ion peak m/z and intensities](#).
- Import MS/MS and in-source spectra [from JCAMP files](#) (October 2012, nistms32.dll version 2.1.5.3 or later)

**Note that the following features were added to the previous version (distributed with NIST 11, version 2.0g)**

A new [Exact Mass Search](#) and similar to it [Exact Mass Constraint](#) allow searching for an exact mass of a precursor using its mass or formula or m/z of the precursor ion. This

search and constraint take into account the presence or absence of an electron, adducts or losses; accepts uncertainty in ppm or millimass units; and allows searching for isotopic or monoisotopic peaks.

Ability to use accurate mass to [search exact fragment ion mass](#) values in NIST libraries mainlib, replib, and nist\_msms using Exact Mass in the Any Peaks Search.

Display of the exact mass with any spectrum that has an associated chemical formula.

Manage and search as many as 127 separate MS libraries instead of 16, which was the limit in previous versions.

Optional display of columns in Hit Lists containing the number of synonyms and the number of other databases in which the hit appears. The hit list can be sorted according to the values in these columns.

Maximum number of spectra in a database has been increased from 786,420 to 1,048,560.

Full compatibility with most recent NIST Peptide MS/MS libraries.

Copy selected lines from Hit List or Spec List into Windows Clipboard as tab-delimited Unicode text; this may be conveniently pasted into Excel.

Spectra in the MSP text format can be imported from the Windows clipboard into the Spec List via a Right Mouse button (RMB) menu option.

Searching for words in mass spectrum text information window (use context menu).

Optional exclusion of Homologues from the results of Structure Similarity search.

Import spectra from mzXML and mzData files.

Multiple CAS Registry numbers for a spectrum. A small fraction of spectra in NIST 14 has more than one CAS r.n. Each CAS r.n. may be found in the CAS Number search.

[Tags for searching](#) in [Spectrum Text Info](#) have been added to Tags in Comment constraint

New [MS/MS Instrument Type Constraint](#)

An alternative peak matching method to improve the reliability of the score when searching noisy MS/MS spectra.

When a Tag from the "Comment" field is displayed separately, its display in the "Comment" field is suppressed.

# Introduction

Functionality is provided in five tab-activated Windows, each window having its own specific functions. The general features of all windows are described in detail.-

## [Program Windows](#)

There are a number of controls for the windows, some are specific to each window type and some are general for all the windows - [Controlling Window Display](#).

The right mouse button has many functions. It is context specific, but there are some general features that are common to all of the windows. Try using the right mouse button in each of the windows to see the options it brings up - [Right Mouse Button Controls](#) .

Finally there are some general points about the program that should be kept in mind when using the program - [General Points](#).

To view **Table of Contents** of this help file, press Show button,  [Show](#), on the top left of this window.

To view [What's New in MS Search](#), click this link.

# NIST MS Search Program Windows

## [Library Search Window](#)

Used for finding and comparing user provided (unknown) spectra to the best matching spectra in the NIST or user libraries. Searches may add a variety of [constraints](#) (name, elements present, common compound, etc.)

## [Other Search Window](#)

Used to look up compounds in the library by formula, identification number within a library, [molecular weight](#), "any" peaks, Chemical Abstracts Service registry number, sequential search with [constraints](#), and NIST spectral number. With the exception of the "any peaks" search, these searches only use chemical identity information, not spectral information.

## [Names Search Window](#)

An incremental name search - as you type in the chemical name, names beginning with the letters entered are shown.

## [Compare Window](#)

Use this window for a more flexible comparison of spectra. This is especially useful for comparing more than two spectra or for comparing spectra not arising from a library search.

## [Librarian Window](#)

Provides a set of tools for building and maintaining spectra and chemical structures in a user library.

## [MS/MS Window](#)

Provides convenient viewing of the nist\_msms library contents. Is not displayed if nist\_msms library is missing.

# Controlling Window Display

## Zooming

In a window showing a mass spectrum, the mouse can be used to zoom a rectangular section of the window. Placing the mouse on the screen at any point and holding down the left mouse button will form an anchor for a box. The box is drawn by diagonally dragging the mouse. Dragging horizontally or vertically will not zoom in perpendicular direction. See also [m/z range](#) item under Options menu. Zooming may be reversed by using the right mouse button and then selecting "zoom out" or by double-clicking the zoomed window.

## Splitter Bars and Display Control Tabs

The relative amount of the space taken by any window can be adjusted by using the mouse to move the splitter bars - the boundary lines between the windows. As you move the mouse between windows, you will notice that the mouse cursor changes from a normal arrow into a small resize symbol (similar to a horizontal or vertical equals sign "="). With this cursor symbol showing, you can hold the left mouse button down and drag the splitter bar to resize the window. This resizing can be done in either the horizontal and vertical directions depending upon which border you drag. The windows can be reduced to zero size, in this case the splitter bar alone will be barely visible. Moving the mouse above the upper inner edge of a window allows to discover the hidden splitter bar.

At the bottom of some windows are a series of tabs. These change the way information is shown in that window. The different types of windows have different controls on the tabs.

## List Display Tabs

### Names

All lists of chemical compounds can be displayed as names or as names with structures. The selection of the Names tab forces the names only to be shown.

### Structures

Selection of the Structures tab allows the structures to be shown. Normally both the structure and the name will be shown. You may use the properties (right mouse button)

of a window to show only structures.

## **Compare Display Tabs**

### **Difference**

For each compare window (in either the Library Search window or the Compare Window) there are 4 ways to show the spectra being compared, each with a separate tab.

The Difference tab will show peaks from the first (often upper) spectrum and second (often lower) spectrum along with the difference. The first spectrum will point upward and the second spectrum will point downward from a common axis.

If the peak is only in the first spectrum it will be shown in red (default), if it is only in the second it will be blue (default), if it is in both the difference will be shown in green (default) with the direction (up or down) indicating which of the two displayed spectra has the largest value for that peak.

### **Head to Tail**

The Head to Tail tab will show only the two spectra being compared. The first spectrum will point upward and the second spectrum will point downward from a common axis.

The peaks in the first spectrum it will be in red (default). The peaks in the second spectrum will be in blue (default).

In this mode the structure may be displayed on the difference spectrum - set by properties (right mouse button). The structures are color coded to the mass spectral peaks.

### **Side by Side**

The Side by Side tab will show only the two spectra being compared. Both spectra point upwards from a common axis. Two peaks with the same  $m/z$  are plotted side-by-side, with spaces separating successive  $m/z$  values.

By default, peaks in the first spectrum are in red and peaks in the second spectrum are blue.

In this mode, the structure may be displayed on the difference spectrum - set by properties (right mouse button). Structures are in the same color as corresponding

spectra.

## Subtraction

The Subtraction tab creates a spectrum derived by subtracting the second spectrum from the first. No peak can go below zero. The resulting spectrum is plotted in green (default). Use this mode to subtract the spectrum of a component from the spectrum of mixture.

## Spectral Display Tabs

### Plot/Text

The Plot/Text tab generates a window that is separated into two parts by a splitter bar. One part is a graphical display of the spectrum, the other is a textual display. The graphical portion contains a depiction of the spectrum and chemical structure. The Text portion of the window displays the name, formula, [molecular weight](#), the CAS number, the NIST number, the ID number within the library from which the spectrum came, other databases which list the chemical, the contributor or comment, and any synonyms. Here is [a full list of items](#), which may be displayed in the text windows.

Like most windows, the actual items displayed may be set using the window's properties control.

**Displaying tagged comment fields** - If comment/contributor field contains specially formatted expressions, Tag=value or Tag="value" where Tag is any combination of letters and value should be quoted if it contains spaces then these expressions may be displayed on the plot and in separate lines in textual display but not in the comment/contributor field. To display these tagged fields, select **Comment field options** from the View menu and enter the tags. Select **Show tag=value on plot** is also available. See also [Tags in Comment Constraints](#).

### Plot

The Plot tab simply suppresses the Text window, showing only the graphical spectrum and structure.

# Menu/Tool Bars

## Menu Bar

**File** - Opens a file for import of spectra. Spectra imported will be placed in the Spectrum List (Spec List)

**Search** - Brings a list of all relevant search modes otherwise accessible in Other Search window. If a mode is chosen here, the proper window will immediately open. Searching for a spectrum or a structure in a library is not in this menu.

**View** - There are always three choices on this menu - Toolbar, and Status Bar, which can be turned on or off, and [Comment field options](#). The Status Bar at the bottom is used to display messages about the current actions of the program. In addition, in the library search window, the current mode of search and the mode of the search of the active hit list are displayed in the Title Bar (at the top of the window)

For most windows, there are other options.

**Options** - [m/z range](#) displayed in the Compare and Library Search Window, [Library Search Options](#), and [Replicates](#).displayng options

**Tools** - Five associated programs can always be brought up from this menu.

- MS Interpreter which examines the structural consistency of a mass spectrum using the structural data stored in the database and
- AMDIS for deconvolution of GC/MS data files.
- [Structure Search Database Builder](#) updates user library structure indexation. Depending upon the window and the search done, other tools may be available. For example it is possible to see the substructures derived from a hit list in the Library Search Window ([Substructure Identification](#)) and to see the [substructures used in structure search](#).
- (Re)Index Exact Mass creates index files for [Exact Mass Search](#)
- (Re)Index RI creates index files for displaying Experimental RI, contents of RI column in hit lists, and [using RI in EI searches for Match Factor correction](#).

In addition, the search option parameters including automation can be set from this menu.

**Window** - Set the arrangement of windows and shift to various windows.

**Help** - Bring up the Help system, open NIST Web site in the browser, or display About

box.

**Minimize/Maximize/Exit** - Standard Windows control buttons for resizing the window.

## Main Toolbar

**Cut/Copy/Insert** - Standard Windows commands to act on the highlighted data (Cut and Copy) or to act on data in clipboard (Insert = Paste).



**Print Report** - Prints a report of the current search. Note, in general any window, except the incremental name search window, can be printed by use of the right mouse button.



**MS Interpreter** - Examines the structural consistency of a mass spectrum using the structural data stored in the database



**AMDIS** - Deconvolution of GC/MS data files.



**Set Default Layout** - Sets the current layout of a window as the default layout of that window.



**Change Layout** - Alternates between a series of layouts - current, horizontal, vertical.



**m/z range** - Change spectrum and spectra difference display



**Switch to Caller** - If the program was called from another program, such as AMDIS or an instrument company data system, return the focus to the calling program.



**About** - Displays NIST MS Search version, build date, copyright information, etc.

## Short Cut Bar

The short cut bar is a series of buttons and dialog boxes that allow certain common operations to be performed with a single mouse click. Hovering mouse pointer above a button displays the button name.



**Go** - Search using the currently highlighted spectrum in the Spectrum List (Spec List). Uses the current libraries and [constraints](#).



**Structure Search** - Search using the structure of the currently highlighted data from the Spectrum List (Spec List).



**Import** - Short cut to File/Open for importing data from a file.



**Library Search Options** - Select options, constraints, and libraries for a target spectrum search.



**History Lists** - Go back to previous hit lists. The last 100 searches are retained. You may override this and automatically clear the history when you close the program. To clear the history on exit, right click the mouse in any window, select Properties, and select the "Hits List" tab. By checking "**Clear History on Exit**" you will delete the history list and the associated files.



**Clear History** - Clears the history list and deletes all associated files.



**Replicates On/Off** - If you are searching with the NIST main library, you have the option of automatically showing all replicate spectra (alternate spectra of a compound). These are always indicated in the hit list, but are not shown by default. You can change that behavior with this short cut.



**Best Matching Only** - Display only one best matching spectrum of each compound. In all searches except MS/MS search, compounds are identified by their CAS Registry Numbers. In MS/MS search, compounds are identified, in addition to CAS number, by their names (case-sensitive).



**View Hit List Search Options** - Show the options (libraries, [constraints](#)) used in the active hit list. In addition it will show the type of search performed.



**View Substructure/MW/Cl-Br Information** - Brings up a list box containing data on the chemical information about the compound derived from an analysis of the hit list.

# **Right Mouse Button Controls**

## **Context Sensitive**

Pressing the right mouse button displays a set of selections that depend on the context (window). In all windows it allows setting window properties, and in most windows it allows the spectra or the structure to be cut, copied, and pasted. In addition, for any window with a single spectrum, it allows various spectrum search/transfer options.

The exact action of the right mouse button will be discussed in each section.

## **Search**

Whenever there is a mass spectrum in a window, the right mouse button gives the user the option to immediately perform a library search with that spectrum. If the spectrum has a structure associated with it, the structure can also be searched to find similar structures.

## **Control Window Properties**

Each window has a number of adjustable display properties. In general these include the font typeface and size, as well as the color of the text or spectra. For each window there are specific controls for changing the amount of information shown.

## General Points

The NIST Mass Spectral Search Program (MSSP) conforms to the general Microsoft Windows conventions. Thus multiple selections can be made of spectra either with the Control or Shift keys in coordination with the mouse or the arrow keys. Highlighting the first name (or structure) in a list with the mouse and then going to the fifth name (or structure) and selecting it with the mouse while holding down the shift key will select the first through the fifth name. If the Control key, rather than the shift key were pressed, only the first and the fifth will be selected. The resulting set of selections can then be copied or cut to the Windows clipboard or sent using the right mouse button to various other windows in the program.

Text may be highlighted for copying to the clipboard, and text in the clipboard may be pasted into appropriate windows (for example the formula search window or the editor name window).

In general, the following **standard and non-standard Windows short cut keys** are active

Short cut key	Action	Short cut key	Action
Ctrl-C, Ctrl-Insert	Copy	F1	Brings help
Ctrl-V, Shift-Insert	Paste	Ctrl-A	Selects all spectra in a list
Ctrl-K	Copy selected lines from a list of spectra	Del	Deletes selected spectra from Spec List
Ctrl-X	Cut	Ctrl-Left Mouse Button	Select/deselect a single spectrum
Ctrl-F6, Shift-Ctrl-F6 Ctrl-Tab, Shift-Ctrl-Tab	Move through all open windows	Shift-Left Mouse Button	Select multiple spectra
Ctrl-F4	Close the current window	Alt-F4	Close the program

Some of these actions may be done with context menu opened with the right mouse

button.

Leaving the mouse cursor on a tool bar control will in general bring up a brief description of the control. The splitter bar is used extensively and where multiple column lists appear, the column width may be adjusted with the mouse from the top of the column.

Information on the status bar located at the bottom of the main window may be helpful.

**Molecular weight** always refers to the sum of the most common isotopes (nominal MW), not the average molecular weight, thus  $C_2H_2Cl_4$  will have a molecular weight of 166, the highest intensity parent ion is 168, and the average molecular weight is 167.85. Throughout this program the MW will be 166.

**Exact mass** or **Monoisotopic mass** is the sum of the masses of the atoms in a molecule using the principal (most abundant) isotope for each element.

**Accurate mass** - Experimentally determined mass of an ion that may be used to determine an elemental formula.

Note: accurate mass and exact mass are not synonymous. The former refers to a measured mass and the latter to a calculated mass.

**Dragging and dropping of spectra with the mouse** is generally allowed. It is not possible to insert something into a hit list since this is the result of the program search. However, you may drag a spectrum from the hit list into the compare, or drop a spectrum from any window into the compare window. This action does not affect the hit list resulting from the search.

# Library Search Window

The principal function of the program is to compare unknown spectra against the NIST library. This is performed within this window. From within this window you may

[Import spectra to be searched against the library](#)

[Search spectra for matches in the library](#)

[Search structures for matches in the library](#)

The search can be constrained by formula, compound name, and other parameters. In addition the search can use only the NIST library or only user supplied libraries or both. These and a number of other parameters can be set using the

[Library search options dialog](#)

Once you have a set of results, you may alter the way that they are displayed by

[Controlling window display](#)

Most of these options can be accessed easily from the

[Menu short cut bar](#).

The search results are displayed in the [Hit List](#).

Spectra for searching are in the [Spec. List](#). The same Spec. List is in the [Librarian](#) window.

# Importing Spectra to be Searched

## File

Spectral data can be imported from any of a number of text file formats by opening a file or right-clicking Spectrum List (Spec. List) and selecting Import. In each case the spectra are placed in the Spectrum List (Spec. List) before any other action. The standard NIST format (\*.msp) or JCAMP (\*.jdx) are recommended. In addition, a number of text file formats will be correctly read by the program.

Once a file has been selected, a [dialog box](#) will appear showing all of the mass spectra found by the program in the file.

See also [Retention Index Import with a MSP or JCAMP file](#)

## AMDIS

Using AMDIS it is possible to directly export extracted GC/MS component spectra either one at a time or in batches to the program.

## Direct Import from Data System

For some instrument data systems direct export of the spectra from the instrument is possible. Generally it is possible to use the macro language that comes with the instrument to bring spectra directly from the instrument data system.

## Import File

The title displays Number of Spectra found.

On the left is a window showing the spectra to be imported. As they are selected the spectrum will appear at the bottom of the dialog. The spectra can be imported one at a time or in groups using normal windows keyboard and mouse controls. The data can be imported by pressing Import All button or, once the selection is made, by pressing Import Selected button.

Various [Spectrum Information](#) may be imported from a text file in addition to the peak list.

The [Library Search Options](#) button will bring up the tabbed dialog box for setting the type of search, the constraints, and the list of libraries.

In addition, [Import Options](#) will bring up a dialog that will allow for corrections needed for certain instruments for mass errors in the instrument or to convert accurate masses into nominal masses. This dialog will allow you to multiply the mass by a constant (1.0000 is the default) or add a constant (0.000 is the default) to each mass. In all cases the resulting number will be rounded to the nearest integer (100.4 becomes 100 as does 99.6); these mass values are used in EI searches.

There are also controls to set Intensity threshold, max. number of decimal places in accurate m/z values, etc.

If the amount of memory required to import a MS/MS or in-source/EI with accurate ion m/z spectrum exceeds a threshold, smaller peaks are automatically removed from the spectrum. In this case, a note is added to the Comment field, for example, IntensityThresholdUsed=0.185%, which means that peaks with intensities below 1.85 have been removed. Typically, 8,000-9,000 peaks are left in the spectrum.

The Import Options dialog also has a check box for including synonyms. If the data files have synonyms contained in them, these will be imported if the Include Synonyms box is checked.

Active intensity threshold(s) are displayed in the middle of the dialog box. The selected spectrum preview is in the lower half of the dialog box.

Other controls make MS Search stop asking whether the imported spectra should prepend or overwrite contents of Spec. List.

See also [Importing Retention Index with a Spectrum](#)

To import a file, select from the menu File / Open, or click Import button , or right-click Spec List and select Import. Select a file and select Open.

# Using Tags in Spectral Import

(Applies to nistms32.dll version 2.1.5.4 build November 2012 or later)

## .MSP Files

All tags are listed in "Tag(s) used in MSP file" column of [this table](#). The first line of a spectrum must start with the Name tag. The spectrum must end with a peak list. Below is a minimal (3 lines) example of a spectrum; this spectrum has a single ion peak with m/z=149 and intensity 150:

```
Name: Scan 25
Num peaks: 1
149    150
```

Any additional information in format

Tag: value

must be between the Name and Num Peaks lines. All tags are recognized ignoring letter case. Any tag should not have embedded spaces, except Num Peaks tag.

There are 2 types of tags: (i) tags, which correspond to **internal tags**, \$:nn, and (ii) tags, which do not. For example, tag PrecursorMZ belongs to type (i) because it corresponds to the internal tag \$:04. Tag Formula belongs to type (ii).

**Type (i) tags** are recognized ignoring letter case, underscores, `_`, dashes, `-`, and slashes, `/` and `\`. These tags may be represented in two different ways. For example, internal tag "\$:04" corresponds to the "PrecursorMZ" tag. Precursor m/z=285.142 may be represented in the input MSP file with the line

```
PrecursorMZ: 285.142
```

or

```
Synon: $:04285.142
```

Representation with a synonym is compatible with NIST 08 software.

Tags differing only in CAS, CAS#, or CASNO at the end of the tag are equivalent. For example, tags located on the same line below are equivalent:

```
Related_CAS#, Related_CAS, RELATEDCASNO, RELATEDCAS
Salt/mix_CAS#, Salt/mix_CAS, SALTMIXCASNO, SALTMIXCAS
```

**Type (ii) tags.** The underlined part of the tag shown in the "Tag(s) used in MSP file" column of [this table](#) is required for importing corresponding data from a MSP file into NIST MS Search. The rest of the tag is ignored. . For example, lines (without quotes)

"form:C2H5" and "Formula: C2H5" in the input MSP file are equivalent. For the same reason, tags CAS, CAS#, CAS\_Reg\_No and CASNO are equivalent. However, line "CAS reg no:500-00-0" may not be recognized because its tag, "CAS reg no", has spaces.

## JCAMP-DX (version 4.24 or higher) files

For version 4.24 description, see R.S. McDonald, P.A. Wilks; Appl. Spectrosc., 1988, **42**(1), 151-162. Version 5.0 for mass spectrometry is described in P. Lampen e.a.; Appl. Spectrosc., 1994, **48**(12), 1545-1552.

For an example, see file MSMS\_JCAMP\_Example.JDX usually provided with NIST MS Search.

The first line in a spectrum must have label ##TITLE=. Line with ##END= may be omitted.

JCAMP-DX labels are recognized ignoring letter case, spaces, underscores, dashes and slashes.

In addition to the following labels (characters ## and = are omitted):

TITLE, JCAMPDX, DATA TYPE, END, XYDATA, PEAKTABLE, XYPOINTS, NPOINTS, ORIGIN, OWNER, SAMPLE DESCRIPTION, CAS NAME, NAMES, MOLFORM, CAS REGISTRY NO, MW, .NOMINAL MASS, .MONOISOTOPIC MASS

NIST MS Search also recognizes "User-defined labels" corresponding to **internal tags**, \$:mn (see [this table](#)). To create such a JCAMP-DX label, add "##\$" before and "=" after the corresponding MSP tag or SDfile data header found in [this table](#). For example, the internal tag "\$:04" corresponds to the MSP tag "PrecursorMZ" and SDfile data header "PRECURSOR M/Z". Therefore, JCAMP-DX labels

##\$PRECURSOR M/Z=, ##\$Precursor\_m/z=, and ##\$PRECURSORMZ= are accepted by NIST MS Search as equivalent.

In addition, each of the following lines below contain labels, which are treated as equivalent:

```
##INLET=,           ##.INLET=,           ##$SAMPLE_INLET=
##IONIZATION MODE=, ##.IONIZATION MODE=, ##$IONIZATION=
##SCAN RANGE=,     ##.ACQUISITION RANGE=, ##$MASS RANGE=
##.SPECTROMETER TYPE=, ##$INSTRUMENT TYPE=
##SPECTROMETER/DATA SYSTEM=, ##SPECTROMETER SYSTEM=, ##$INSTRUMENT=
```

Labels corresponding to an **internal tag** and differing only in CAS, CAS#, or CASNO at the

end of the label are treated as equivalent:

##\$Related CAS#=  
##\$Related CAS= , ##\$RELATEDCASNO=  
##\$SALT MIX CAS#=  
##\$SALT MIX CAS=, ##\$SALTMIXCASNO=  
##\$SALTMIXCAS=

In case of **JCAMP-DX Revision 4.10**, only labels typical for HP ChemStation and Agilent MSD ChemStation are recognized.

## Import Options

Import Options are applied to spectra read from a file, sent to MS Search by another program (for example, AMDIS), or inserted from Windows clipboard. The spectra are imported into Spec. List.

**EI Spectrum m/z Rounding** allows for corrections needed for certain instruments for mass errors in the instrument and/or to convert accurate masses into nominal masses needed for searching in EI libraries. This will allow you to multiply the mass by a constant (1.0000 is the default) and/or add a constant (0.000 is the default) to each mass. In all cases the resulting number will be rounded to the nearest integer (100.4 becomes 100 as does 99.6).

Note that these corrections affect the results of searching MS/MS or In-source/EI with accurate peak m/z spectra in EI library.

**Tandem Spectrum Accuracy** settings are for reducing the number of decimal places in precursor ion m/z and product ion m/z values in [MS/MS spectra](#). For example, selecting 4 decimal places would round 123.456789 to 123.4568 and leave 123.45 unchanged.

**In-source/EI accurate ion m/z** settings are for reducing the number of decimal places in peak ion m/z in [in-source/EI with accurate ion m/z spectra](#).

**Intensity threshold** allows setting relative (0-100%) and absolute ( $0-10^{10}$ ) thresholds for mass spectral peak intensity.

Note that too high absolute threshold may eliminate all peaks in a mass spectrum. Such a spectrum is not displayed at the bottom part of the [Import File](#) dialog and is not imported even if selected. Active intensity thresholds are displayed in the middle of the [Import File](#) dialog.

To import spectra previously exported from NIST RI library, deselect both Intensity threshold check boxes.

**Adding spectra to Spec List** controls provide 3 possibilities:

- Prepend - add the spectra to the top of Spec List;
- Overwrite - replace Spec List spectra with the imported ones;
- Ask - the program will ask if the user wants to prepend or overwrite the Spec List. In this mode, a request to prepend or overwrite made by an application sending mass spectra to MS Search will be accepted by MS Search.

**Spectra without precursor ion m/z value** are box contains two radio buttons. They allow or disallow import of accurate ion m/z from spectra that are not [MS/MS](#) and do not have "*Spectrum\_type: in-source*" line. Select "[In-source / EI \(accurate m/z\)](#)" to import a spectrum with accurate m/z; select **EI (nominal m/z)** to import rounded to integer values m/z.

[MS/MS spectra](#) are always imported with accurate ion m/z.

**RI type if unspecified** drop down list box. Choose Column Type for RI values of unspecified column type in imported spectra. The column type will be saved in the imported spectrum. Unspecified column type may be left unchanged. RI values must be between 1 and 16,383.

**Include Synonyms** check box. If the data file has chemical name synonyms contained in it, these will be imported if the Include Synonyms box is checked.

**Set default values** button clears intensity thresholds and sets other controls to default.

To open Import Options dialog, select from MS Search menu Options, Spectral Import options or select Import Options from the [Import File](#) dialog.

## Searching Spectra and Structures

Once a spectrum has been placed in the [Spectrum List \(Spec List\)](#), there are a number of ways to [search](#) the database with the spectrum:

### Double Click on the Spectrum Name or Structure

Once the spectrum name appears in the Spectrum List, you may search with the current constraints by double clicking with the left mouse button on the spectrum name in the list.

### Use the Right Mouse Button

The right mouse button will bring up a menu which, if the mouse is over a spectrum will allow you to search the library using the spectrum pointed to by the mouse. This mode is active in all windows and functions with names, structures or spectra.

### Use the Go Button

The small Go button  at the top of the window will cause the search to start with the current conditions.

### Automatic Searching

It is possible to have the program automatically search spectra that are placed in the Spectrum List (Spec List). This is done by setting Automation to On in the [Library Search Options](#) Box - Access from the [Short Cut Bar](#) , [Library Search Options](#) button



. Automation can be selected in the Search Dialog Tab or in the Automation Dialog Tab. This dialog box is also available from the menu by selecting Tools / Library Search Options.

### Structure Similarity Search

In addition to searching the database using the spectrum, it is also possible to search based on the structure if known. To do this the right mouse button is used and the Structure Similarity Search is chosen. In addition the small button with a representation of cyclopropane (triangle) on the [Short Cut Bar](#)  can be used.

Glycan structures cannot be searched.

## Structure Search Using InChIKey

In all small molecule spectra available in NIST 14 Libraries, MS Search v.2.2 displays InChIKey line in the Text Information window, for example:

[InChIKey: FQZQXPXKJFOAGE-SNXKPFKBSA-N](#) [Non-stereo](#)

Clicking the first link (it represents 27-character InChIKey) would open Web browser and search PubChem for exact structure match (ignoring differences between certain types of tautomeric forms). Clicking the second link, [Non-stereo](#), would search PubChem for the first 14 characters of InChIKey thus also ignoring differences in stereo chemistry, isotopic substitution, and protonation.

Searching for all 27 or only the first 14 characters of InChIKey using [Name Fragment](#) constraint in Sequential Search (available in Other Search window) performs the same structure search in NIST 14 libraries.

Alternatively,

`$inchikey=FQZQXPXKJFOAGE-SNXKPFKBSA-N`

may be entered in Tags in Comment constraint

## Setting Search Options and Constraints

The tabbed dialog box that appears when the Short Cut Bar icon for Library Search

Options  or menu Options / Library Search options is selected provides a number of options. These options are retained and used in succeeding library searches.

The tabs of the dialog are:

**Search** - Set the types of searching

**MS/MS** - Set **MS/MS spectrum** or **In-source/EI with accurate ion m/z spectrum searching options** for searching in libraries of MS/MS or in-source/EI with accurate ion m/z spectra

**Libraries** -Set which libraries are searched and the order of the search

**Automation** - Control the automatic searching from other programs or on import of data

**Limits** - Control the range of mass and abundance to be used in searching

**Constraints** - Apply additional filters on the search.

**RI (GC)** - Set Retention Index Options

# Library Search Options

To access these options, select *Library Search Options* from the *Options* menu or click *Library Search Options* button  in the Lib. Search window and select Search tab.

## Spectrum Search Type

### Identity

**Quick (EI)** - Uses a single prescreen (or presearch) algorithm. The prescreen identifies spectra in the library with features in common with the submitted spectrum. This subset is then compared in detail to product spectrum match factors, which are used to arrange library spectra in a hit list.

**Normal (EI)** - Uses a more extensive set of prescreens - a total of four different prescreen criteria. Typically this will find more spectra for comparing to the submitted spectrum than the Quick search. Note that if your spectrum is not passed through prescreen it cannot be found in the hit list. For spectra that have no closely matching spectra in the library, it may be possible to improve results by searching without a prescreen. (See Presearch)

For both Quick and Normal the final match is calculated using a dot product of the two spectra, with higher m/z peaks having more weight than lower m/z peaks.

**MS/MS** - Searches for a [MS/MS spectrum](#) in a library of MS/MS spectra. Set [MS/MS and In-source/EI with Accurate Ion m/z Spectrum Search Options](#) in MS/MS tab. This may be used for high resolution spectrum searching.

**In-source HiRes** - Searches for an [in-source/EI with accurate ion m/z](#) or [MS/MS spectrum](#) in a library containing in-source / EI accurate ion m/z or MS/MS spectra. The library must be built with Lib2NIST ver. 1.0.4.28 (07/05/2013) or later and have files peak\_em0.inu and peak\_em0.dbu. Set search options in [MS/MS and In-source/EI with accurate ion m/z Spectrum Search Options](#) located in MS/MS tab. Unlike MS/MS search, this search does not compare precursor m/z values. Currently, in-source/EI with accurate ion m/z spectra added with NIST MS Search to a user library may be searched with In-source HiRes search only with Presearch OFF option. To search such a library with Default or Fast Presearch, (re)build the library with Lib2NIST v.1.0.4.28 or later.

### Similarity

**Simple (EI)** - In the Similarity search, a similar set of four prescreen techniques are used but the final match factor is calculated without using m/z weighting. In general this is more likely to produce spectra that are from molecules structurally similar to the compound that produced the submitted spectrum.

**Hybrid (EI)** - Hybrid searching uses both the logic of normal searching plus the logic of neutral loss searching.

**Neutral Loss (EI)** - The neutral losses from the molecular ion are used to produce a neutral loss spectrum. In many cases, this type of matching will provide structurally similar matches.

**MS/MS in EI (EI)** - Search for a [MS/MS spectrum](#) in a library of EI spectra.

All searches marked (EI) use integer peak m/z and integer intensities in the range 1-999. With these searches, any mass spectrum may be searched in any library, which contains mass spectra.

## Spectrum Search Options

### Reverse

Use "reverse" logic which ignores "impurity" (non-matching) peaks in the submitted spectrum. The match factor *is not penalized* for peaks in the target spectrum that are not present in the library spectrum.

### Penalize rare compounds

In an identity search, it is possible to penalize a match by up to 50 points (out of 1000) depending on how common the compound is. The compound is judged to be common if it is contained in a number of other chemical databases. As more databases contain the compound it is assumed to be more common. The same [databases](#) that are used in the constraints are used.

### Presearch

**Default** - Normal operation, use presearch screening before spectrum-by-spectrum match factor calculation.

**Fast** - Use almost the same presearch screening to select on average 3 times less spectra than in Default.

**Off** - Search the entire database using only the detailed match algorithm. This option can take much longer and may be helpful when no close matches are obtained using the default search.

**MW** - Restrict search to only those molecules with a user supplied [molecular weight](#). Use this when the molecular weight of a substance is known.

See also [MS/MS and In-source/EI with Accurate Ion m/z Spectrum Search Options](#) for specifics of MS/MS searching and [Retention Index Options](#) for using RI in EI searches.

## Structure Similarity Search Options

**Match Number of Rings** - Forces the number of rings to be the same in the matched compounds as in the target structure.

## Other Options

**Automation** - If checked do an automatic search on any spectrum or structure brought into the Spectrum List (Spec List). Note, if you use AMDIS and wish to automatically search spectra found by AMDIS with the NIST MS Search Program you must check this option. Under the Tab [Automation](#) the parameters that control the printing some aspects of the searching are set.

**Auto Report** - If checked will automatically print a report. The parameters as to how much to print are set under the Tab [Automation](#) in this dialog.

**Apply Limits** - If checked, activates the limits set under the [Limits](#) Tab in this dialog.

**Use Constraints** - If checked, the constraints that are set under the [Constraints](#) Tab in this dialog are active. The constraints set limits on what is accepted in the final hit list.

**NOTE:** Each of these Other Options can be set and activated under the appropriate Tab dialog. This allows a quick confirmation of which are active and allows them to be activated or deactivated without going to the different tabs of this dialogs. To change the conditions you **must** go to the individual tabbed dialogs.

## Retention Index Options

To access these options, select *Library Search Options* from the *Options* menu or click *Library Search Options* button  in the Lib. Search window and select RI tab.

The value of a retention index for a compound is dependent upon the type of column on which the retention data was taken. Since the database has RI values for multiple column types, the column type to be used for searching must be specified. The column type may be specified in the spectrum submitted to the database, or the program determines it from these controls. Generally, only RI of the same column type should be compared.

**Use Retention Indices (RI) comparison for Lib. Search Match Factor correction** - turn on penalties for RI mismatch between search and library spectrum when using EI-type search (Identity Normal, Quick, Similarity Simple, Hybrid, Neutral Loss, MS/MS in EI.)

Choosing one of the three radio buttons,

- **Standard Non-Polar (e.g. DB-1)**
- **Semi-Standard Non-Polar (e.g. DB-5)**
- **Standard Polar (e.g. DB-WAX)**

will determine the type of column to use:

- if no column type is specified with the RI in the search spectrum
- if the search spectrum has RI for multiple column types
- for displaying RI in a hit list in case no RI comparison is done in the search, including Other Search window and Structure Similarity search results

For comparison or displaying RI values, the program will only use the column type that is specified. If the RI value of the Selected Column Type is not available for a library spectrum, the library spectrum is considered as having no RI value. There may be two exceptions to this rule:

- 1) option "Use other non-polar column type data if selected non-polar type data is absent" allows the program to use an alternate non-polar column type **ONLY** if data for the selected type is absent. This exception applies to the search spectrum and to the found spectrum.
- 2) option "Column type specified in search spectrum will override selection", see below.

**The following three options can be used to modify the selection made**

**above**

**Column type specified in search spectrum will override selection** - If the RI value of the Selected Column Type is not available for a search spectrum but another Column Type RI is, this other Column Type is considered the Selected Column Type for searching this spectrum. This option does not affect RI of Unspecified or Any (n-alkane standard) Column Types. See also [Special Case \(a\)](#) below.

**Use other non-polar column type data if selected non-polar type data is absent** - This option is only active if one of the two Non-Polar column types is selected. If the Selected Column type is not available but other non-polar column type RI is available for a spectrum, then that RI value is used.

**User library spectra with unspecified column type assumed to have selected column type** - This treats RI of Unspecified column type in user library spectra as RI of Selected Column Type. If this option is not selected, RI values of unspecified column type are not displayed in the RI column or used in library searching.

**RI penalty** is subtracted from the Match Factor.

The RI penalty is calculated from the absolute value of the difference between the search spectrum RI and the library spectrum RI

$$\text{DeltaRI} = | \text{RI}_{\text{search spectrum}} - \text{RI}_{\text{library spectrum}} |$$

along with the RI parameters: **RI tolerance** and **Penalty rate**.

**penalty** = 0 if  $\text{DeltaRI} < \text{RI tolerance}$ ;  
otherwise,

$$\text{penalty} = (\text{Penalty rate}) \times (\text{DeltaRI} - \text{RI tolerance}) / \text{RI tolerance}$$

Penalty rates are predefined in terms of NIST match factor (exact match = 999): Very weak - 10, Weak - 20, Average - 50, Strong - 100, Very Strong - 200, Infinite - 100,000.

Note that no penalty is applied to the Match Factor of a library spectrum when that library spectrum has no RI or if the search spectrum does not have an RI value.

## **How to find out if RI penalties were applied to the search results**

If RI penalties are or could be applied to the Match factor, MS Search displays *Match no RI* column. Library spectra to which penalties have been applied, have the value of RI used for the calculation of penalties in the *RI* column.

## Special Cases

(a) If the **Column type specified in search spectrum will override selection** option has been selected, in cases where the search spectrum has multiple RI values (i.e. values for multiple column types) none of which match the Selected Column Type, the RI of the same column type(s) in the search and library spectra are compared, and the smallest penalty is applied. This situation will normally only occur if you are searching the NIST library spectrum or if you have included RI data for multiple columns in a spectrum. It does not apply to data from an experiment.

(b) **Case when no RI comparison is possible or required** occurs when

- Option **Use Retention Indices (RI) comparison for Lib. Search Match Factor correction** has not been selected, or
- The search spectrum has no applicable RI, or
- Structure Similarity Search has been performed
- One of Other Searches has been performed

In other words, this case occurs whenever *Match no RI* column is not displayed in the Hit list or some Other Search has been performed.

In this case, the value for displaying in RI column is selected according to the same rules as if the search spectrum has Selected Column Type.

**Column type specified in search spectrum will override selection** option is applied to the found spectrum. If this option were selected, RI value of the found spectrum would be displayed even if there is no match between Selected Column Type and RI Column type(s) available in the spectrum. If this creates an ambiguity, then the first available RI of the column type in the following sequence is displayed:

- Semi-Standard Non-Polar
- Standard Non-Polar
- Standard Polar

For example, if Selected Column type is Polar, the found spectrum has RIs of Semi-Standard Non-Polar and Standard Non-Polar column types, and Other Search has been performed, then the contents of RI column are:

<b>Column type specified in search spectrum will override selection</b> option	<b>Contents of <i>RI</i> column in the hit list</b>
Selected	Semi-Standard Non-Polar RI

Deselected	-
------------	---

## MS/MS and In-source/EI with Accurate Ion m/z Spectrum Search Options

Searching for a MS/MS spectrum involves comparison of peaks in the search spectrum with library spectra whose precursor m/z may be required to fall within a selected tolerance setting.

Searching for an in-source/EI with accurate ion m/z spectrum involves comparison of accurate peaks in the search spectrum with library spectra. If MS/MS spectra are involved in this search, Ignore precursor option applies; however, precursor m/z values are not compared.

### Presearch Options

The meaning of the **presearch types** for MS/MS or in-source/EI with accurate ion m/z mass spectral library searching is different from those used for searching in EI libraries.

Presearch	Search Type	
	MS/MS	In-source HiRes
Default	compare only spectra that have precursor m/z within a tolerance range	presearch close to that used in <a href="#">Default Similarity Simple search</a> . <sup>(1)(3)</sup>
Fast	combines presearch close to that used in <a href="#">Default Similarity Simple search</a> <sup>(4)</sup> with comparing only spectra that have precursor m/z within a tolerance range. <sup>(2)</sup>	presearch close to that used in <a href="#">Default Similarity Simple search</a> <sup>(3)</sup> with reduced number of candidates for mass spectra comparison. <sup>(1)</sup>
Off	compare to all spectra in a library ignoring precursor m/z values	compare to all spectra in a library

<sup>(1)</sup> The library must be built with Lib2NIST, otherwise it is ignored by the search.

<sup>(2)</sup> It is strongly recommended to build the library with Lib2NIST, otherwise presearch ion peak m/z tolerance is 1 m/z.

- (3) Presearch ion peak m/z tolerance is **Product ions ±** mass spectral peak tolerance.
- (4) Presearch ion peak m/z tolerance is the greater of 0.004 m/z and **Product ions ±** mass spectral peak tolerance

MS/MS Search ignores library spectra, which are not [MS/MS spectra](#).

In-source HiRes Search ignores library spectra, which are not [MS/MS](#) or [In-source/EI with accurate ion m/z](#) spectra.

These options are located in the Library Search Options. To access them, select from the menu Options / Library Search options or select Library Search Options button  on the Lib. Search window tool bar, then select Search tab

## MS/MS and In-source/EI with Accurate Ion m/z Spectrum Search Options

**Precursor ±** - precursor m/z tolerance range, 0.015-100,000 ppm or  $6 \times 10^{-5}$  - 500 m/z units (for MS/MS spectra only).

**Product ions ±** - mass spectral peak tolerance range, 0.015-100,000 ppm or  $6 \times 10^{-5}$  - 500 m/z units.

**Ignore precursor** - ignore mass spectral peaks around precursor m/z within the larger out of precursor and product ions tolerance ranges.

- In case of MS/MS search, this range is determined by the search spectrum precursor m/z. If it is missing and presearch is OFF (that is, precursor m/z values are not compared), then the library spectrum precursor m/z, if present, is used.
- In case of In-source HiRes search, the ranges for the search and library spectra may be different: they are determined by each spectrum precursor m/z, if both are present. If only one spectrum has precursor m/z, it is used for both spectra.

**Use alt. peak matching** - use alternative peak matching for calculating dot-product (preferred in most cases).

## Scoring

- Select **Peptide** check box to use Peptide-specific MS/MS search method.
  - Deselect **Peptide** check box to use Small Molecule MS/MS or In-source/EI with accurate ion m/z searches.
- Note that only options Q-tof, Score Threshold, and Reverse/Impure work for a

non-peptide-specific search (select [More Options](#) to access them).

- Turning ON **Ignore precursor** option may make scores obtained with MS/MS and In-source HiRes searches different if the search spectrum and the library spectrum have different precursor m/z.

[More Options](#) - other options specific for searching in MS/MS, In-source/EI with accurate ion m/z, and NIST peptide MS/MS libraries.

These options are located in the Library Search Options. To access them, select from the menu Options / Library Search options or select Library Search Options button  on the Lib. Search window tool bar, then select MS/MS tab.

## MS/MS Spectrum

An MS/MS spectrum always contains a precursor m/z value. In addition, mass spectral peaks in MS/MS spectra may be saved with higher accuracy both in peak m/z and abundance, which are used in [MS/MS spectrum search](#). Since there is no uniform standards, MS Search extracts precursor m/z value from the following fields choosing value with the highest priority:

Priority	MSP file (NIST MS format)	Spectrum record in a library
1 (high)	Mz_exact=m/z (comment)	floating point field (0.01 ppm accuracy)
2	Parent=m/z (comment)	Mz_exact=m/z (comment)
3	Precursormz: m/z (file line)	Parent=m/z (comment)
4	\$04:m/z (synonym)	\$04:m/z (synonym)
5	Xcalibur Filter (comment)	fixed point field (0.01 u accuracy)
6 (low)		Xcalibur Filter (comment)

Internally, a MS/MS spectrum, just as an [in-source/EI with accurate ion m/z spectrum](#), has a secondary set of peaks with integer m/z, which are used to perform EI-specific searches (this set of peaks may not be observed or edited in [Spectrum Information dialog](#).) Therefore, before importing MS/MS spectra, setting proper EI Spectrum m/z Rounding parameters in [Spectrum Import Options dialog](#) may be important.

## Searching MS/MS Library

NIST MS/MS Library can be searched with the [Identity MS/MS Search](#), [In-source HiRes Search](#), any of the [Other Searches](#) or the [Name Search](#). All that is necessary is to include the nist\_msms and/or nist\_msms2 Library in the Include Libs column of the Search dialog box for the Other Searches or as the single library to be searched in the Incremental Name Search. The Library [Spectrum Search](#) and [Structure Similarity Search](#) are also applicable. However, the user must understand that MS/MS spectra in general are different from EI spectra as well as the fact that structures in the MS/MS Library refer to the neutral form of the precursor ions.

## In-source/EI with accurate ion m/z Spectrum

In-source or EI with accurate ion m/z spectra are intended for saving in-source tandem mass spectra (spectra of electrospray ions that contain fragments of ions as well as precursors) or high resolution EI or chemical ionization spectra. An in-source spectrum allows storing high resolution mass spectral peaks and intensities as small as 0.01 (maximum of 999, as before); it cannot have a precursor ion m/z value. Ordinary EI spectra may also be imported and saved as in-source spectra to keep accurate ion m/z and a wider range of intensities.

This version of NIST MS Search can perform the following operations on in-source/EI with accurate ion m/z spectra:

- Import from .MSP, JCAMP, MGF, mzData, and mzXML files
- Export to MSP and JCAMP (Revision 4.10) formats
- Saving into a user library
- Searching in EI, in-source, and MS/MS libraries

In-source/EI with accurate ion m/z spectra may be searched using [In-source HiRes Search](#), which takes advantage of accurate m/z. If the library does not have *peak\_em0* files, only Presearch OFF option may be used, which may take longer time. Also EI-specific searches, including Normal, may be used.

Currently, in-source/EI with accurate ion m/z spectra added with NIST MS Search to a user library may be searched with In-source HiRes search only with Presearch OFF option.

To import a spectrum that does not have precursor m/z value as an in-source/EI with accurate m/z spectrum, one or both of the following two conditions must be met:

- "In-source / EI (accurate m/z)" radio button is selected in [Spectrum Import Options dialog](#)
- The spectrum is in MSP format and has line "Spectrum\_type: in-source" or "SYNON: \$:00in-source" (without quotes)
- The spectrum is in JCAMP format and has line "##\$SPECTRUM TYPE= in-source" (without quotes)

If "In-source / EI (accurate m/z)" radio button is selected and *Spectrum\_type* in the input file does not begin with "in-source" or "accurate", it will be replaced with "in-source". See also [Using Tags in Spectral Import](#).

Internally, an in-source/EI with accurate ion  $m/z$  spectrum, just as a [MS/MS spectrum](#), has a secondary set of peaks with integer  $m/z$  and intensities, which are used to perform EI-specific searches (starting from MS Search v.2.2, this set of peaks may not be observed or edited in [Spectrum Information dialog](#).) Therefore, before importing in-source/EI with accurate ion  $m/z$  spectra, setting proper EI Spectrum  $m/z$  Rounding parameters in [Spectrum Import Options dialog](#) may be important.

The intensities in the secondary set of peaks with integer  $m/z$  may differ from a spectrum imported as EI spectrum. The difference occurs when two or more peaks have accurate  $m/z$  rounded to the same integer  $m/z$  value. In case of in-source/EI with accurate ion  $m/z$  spectrum import, the intensity of such peaks are summed. In the case of EI spectrum import, only the largest intensity is saved. Therefore, a match factor of two identical spectra imported as EI and in-source/EI with accurate ion  $m/z$  may be smaller than 999.

# More MS/MS and In-source/EI with accurate ion m/z Search Options.

## *Notes*

1. Options **Q-tof**, **Reverse/Impure**, and **Score Threshold** work for all MS/MS and In-source/EI with accurate ion m/z searches.
2. All other options affect only peptide searches.

## Score Threshold

**High** - Display a small number of best hits

**Medium** - Display medium number of hits

**Low** - Display all found hits

## Scoring

**Q-tof** - The searched spectrum was measured on a Q-tof instrument.

**Omsa** - Use OMSSA scoring in assigning probabilities of correct identification (for searching in peptide libraries only)

**# Replicates (Prob. only)** - Use known number of replicate spectra in a consensus library to assign probabilities of correct identification.

## Display

**Reverse/Impure** - Display dot-product for the reverse spectra compare in **Rev-Dot** column (the "reverse" logic ignores "impurity" (non-matching) peaks in the searched spectrum; this match factor *is not penalized* for peaks in the searched spectrum that are not present in the library spectrum.)

**T/F -qry (log)** - Display the logarithm base 10 of True/False probabilities ratio that the search spectrum was correctly identified

**E-Omsa (log)** - Display the logarithm base 10 of the OMSSA score (for searching in peptide libraries only)

**T/F - lib (log)** - Display the logarithm base 10 of True/False probabilities ratio that the

library spectrum was correctly identified

**Cysteine modification** - Which modifications to take into account when automatically assigning y/b labels to a spectrum with known from the name sequence of residues and charge and no peak annotations.

See also [Peptide constraints](#)

To access these options, open Library Search Options dialog by selecting from the menu Options / Library Search options or by selecting Library Search Options button  on the Lib. Search window tool bar; select MS/MS tab; select More Options.

# Libraries

## Available Libraries/Included Libraries

In some instances this control is in the Options tab. The window on the left includes all the libraries that are available to the program. At the bottom the total number of spectra and the number of libraries is indicated. The window on the right shows the libraries that are currently in use.

To transfer an available library to a library that will be included in the search, highlight (select with mouse) the library on the left hand side and then select the Add button.

To exclude a library, select that library on the right hand side and press the delete key above. Multiple libraries may be deleted at once. NOTE, this action does not delete the library from the disk, only from the search.

The up and down arrow allow a specific library to be placed in any position. The libraries will be searched in the order that they are shown in the list. Only a single library can be moved at a time with the arrows.

At the bottom of the list of included libraries is a running tally of the number of spectra and libraries to be used in the search.

## Spectrum Search/Structure Search

It is possible to keep different lists of active libraries for each of the search modes. The choices are maintained separately for the spectrum and structure search as well as for the searches under [other search](#) (formula, MW, CAS number, etc).

## Automation

**Number of Hits to Print** - In automation it is possible to set the program to automatically print the results of the search. The number of hits (from the best match factor down) to be printed is set here.

**Include Spectrum Plot in Report** - Show the spectrum in the report.

**Draw Structure in Plots** - Show the structure on the spectrum in the report.

**Apply Maximum Spectrum Length** - Limit the maximum m/z shown.

**Return Focus to Caller upon Completion** - If the program is called from another program, automatically return to that program and bring it to the foreground.

**Automatic Search On** - One of two points to specify that the program should automatically search with any new data added to the Spectrum List (Spec List).

Automation Tab is in Library Search Options dialog, accessible through menu Options / Library Search Options or Lib. Search window tool bar button .

## Limits

**Apply Limits** - Select to make specifications set in this window active. If not checked no other information in this window is used.

**Minimum Abundance** - Set smallest peak that will be used in comparison. Used to force small peaks in the target spectrum to be ignored. Set to ON and specify minimum abundance on the basis of 999; accepted values are 2-999, where 2 means 0.2% of base peak intensity, 50 means 5%, etc. ([Base peak](#) = 999)

**Minimum m/z** - Used to determine the lowest m/z used for match factor calculation. Without this specification, the matching algorithm starts comparison at the HIGHER of the minimum m/z values in the target (search) or the library spectra. This is done in order to attempt to ignore peaks in one spectrum that do not match peaks in the other simply because the starting m/z was higher. For instance, if the library spectrum began at m/z 20 and the target began at m/z 40, library peaks between 20 - 40 m/z could cause the match factor to be seriously penalized. This approach, however, can produce artificially high match factors when either the library or search spectrum are missing low m/z peaks (partial spectra). To avoid this problem a minimum mass may be specified - this should generally be the minimum m/z from the instrument that produced the spectrum for library searching. Another option is to specify the minimum mass is "never greater than" a given value. In this case if the target spectrum and the library spectrum have peaks below the specified value, then the largest of the two minimum m/z values will be used. Otherwise, all peaks below the specified minimum m/z will be ignored.

**Maximum m/z** - When selected (ON) peaks above the specified mass are ignored. Use to exclude spurious high mass peaks in the search spectrum.

# Constraints

**Constraints** are used to restrict the list of found spectra to only those, which satisfy all selected constraints.

**Use Constraints** - To use any constraint, this must be selected. All constraints will be turned off if this is off. Specific constraints must also be individually turned on for them to be used. As you move the highlight bar down the list of constraint types, you will notice that the bottom of the dialog changes. When highlighted, the parameters of the active constraint can be set.

**MW (Nominal Mass)** - The minimum and maximum molecular weight can be set. As elsewhere in this program the [molecular weight](#) refers to the sum of the atomic weights of the most common isotope. This constraint may not be selected together with Exact Mass constraint.

[Exact Mass](#) - Allows you to specify the exact (monoisotopic) or isotopic mass of a compound within the uncertainty interval. The search parameters are the same as in [Exact Mass Search](#). However, no library indexing is required. This constraint may not be selected together with MW(Nominal Mass) constraint.

[Name Fragment](#) - Allows you to limit your search to only compounds containing the specified letters. For instance, to limit the search to alcohols you may wish to use the fragment "ol".

[Elements Value](#) - Allows you to specify the number of atoms for specified elements that must be present in a compound. For example, to limit results to compounds containing a single phosphorus atom, specify P=0.

[Elements Present](#) - Allows you to specify that certain elements must be present (in any number), or that only selected elements be present (H,C will limit the search to hydrocarbons).

**Peaks** - Allows specification of peaks, [Normal, Neutral Loss, Rank and Maximum Mass](#). In each case a range may be specified.

**Other Databases** - A list of [other databases](#) that the compound must be in. The list is processed as a logical OR, that is if 5 of the databases are checked, the compound is found if it is either of the 5 databases

[Tags in Comment](#) - Restrict search results to spectra that contain specified tagged expressions in the contributor/comment field or in text information. See also [Displaying](#)

### tagged comment fields

**Peptide constraints** - provide restrictions on Peptide sequence, Peptide mobile protons, Peptide number of residues. These constraints may be used in searches displayed in Other search window; they are ignored for all Library searches except MS/MS search with Presearch Off or In-source HiRes search. Only Peptide [MS/MS](#) and [in-source/EI with accurate ion m/z](#) spectra are affected.

**MS/MS Charge** may be used in searches displayed in Other search window; they are ignored for all Library searches except MS/MS search with Presearch Off or In-source HiRes search. Only [MS/MS](#) and [in-source/EI with accurate ion m/z](#) spectra are affected. A spectrum that has no charge satisfies charge=0 constraint.

**MS/MS Instrument Type** - Restrict MS/MS spectra in search results to specific instruments. Does not affect spectra in EI libraries. As other Peptide and MS/MS constraints, it rejects ms1 spectra located in MS/MS libraries. Only [MS/MS](#) and [in-source/EI with accurate ion m/z](#) spectra are affected.

## Exact Mass Search and Constraint

Search for the [exact](#) or [accurate mass](#) of a compound using a mass value, accurate m/z value and charge, or chemical formula and optional uncertainty in parts per million (ppm) or millimass units (mmu).

Make sure that correct **types** of the entered values have been selected in drop down list boxes located on the right from each entry field. The **Find** box provides the selection of what you intend to find.

The **Search value** may be formula, mass, or m/z. In case of **m/z**, an additional entry box, **Charge**, is displayed, together with **No electron mass correction in calibration** check box. The latter, if selected, prevents adding or subtracting the mass of the electron (see next paragraph). This is useful if you do not take into account the mass of the electron when calibrating the mass analyzer. In case of formula, **Gain or loss** entry box is not displayed.

Entered **m/z** value is increased or decreased by the mass of the electron in case of positive or negative charge, respectively, and multiplied by the charge to obtain the mass to be searched.

Optional **Gain** or **loss** should be entered as a chemical formula; it is subtracted or added, respectively, to the mass before searching. Gain or loss entry box is not available for the formula.

The **Uncertainty** is applied to the entered value or to the [exact \(monoisotopic\) mass](#) calculated from the formula.

**Blank uncertainty** entry box is interpreted as 0.5 in the next to the last digit of the entered mass or m/z. For example, the uncertainty for 28.00 would be 0.005. In case of a chemical formula, blank uncertainty field means the search calculates the same exact mass as was used for indexing; collisions of these values were found to be extremely rare.

Exact mass or m/z may also be searched among several most abundant isotopes, which may or may not include the monoisotopic mass.

Normally, the selection in the **Find** box is **Monoisotopic precursor mass**. If **Among XX most abundant isotopes** has been selected, then masses of up to 16 most abundant isotopes of the compound are included in the search. For some compounds, the exact (monoisotopic) mass may not be among masses of the isotopes. If the **Search value** is a chemical formula, no matter what was selected in the Find box, the exact

(monoisotopic) mass corresponding to the formula will be searched. The values of the found isotopic masses are not displayed. You may calculate them using the MS Interpreter.

**Exact Mass Search** requires running **(Re)Index Exact Mass** tool (located in the Tools menu) against an older user library to create the necessary index file, exactmw.inu. NIST 14 libraries have already been indexed. User libraries created with NIST MS Search v.2.0g or Lib2NIST v.1.0.3.8 or later create this index. Since adding a spectrum to an older user library with the NIST MS Search updates the exact mass index only for the added spectrum, exact mass re-indexing of the older libraries is necessary. Only spectra that have chemical formulas may be indexed.

**Exact Mass Constraint** does not require indexing. Exact and, if needed, isotopic masses are calculated on the fly.

## Name Fragment Constraint

Only those compounds with the string that is entered into to dialog box will be displayed. Note, the name fragment may in the primary name or in any of the synonyms. It is often possible to limit the type of compound found by using unique name fragments. e.g., "ol" for alcohol, "one" for ketone, etc.

The strings are compared in a case independent manner, spaces, dashes and some commas are also not used in the comparison. Also the following multi-fragment expressions are supported:

COCA && INE - at least one name contains both "COCA" and "INE"

COCA && !! INE - at least one name contains "COCA" and doesn't contain "INE"

COCA !! INE - at least one name contains "COCA"; no name contains "INE"

Spaces around && and !! are optional; up to 8 name fragments are accepted; the total length of the entered string cannot exceed 39 characters.

Note that InChIKey may be searched as a name fragment. However, it is not displayed in Names window.

## Elements Constraints

Elemental composition can be constrained in two ways, either by specific limits on the number of the element present or by constraints on what elements can be present.

### Elements Value

This option allows specification of the elemental composition of the retrieved compounds; for example:

Br > 0 Must have Br in molecule

Cl < 3 Must have less than 3 Cl's (0,1,2)

F = 0 Must not have F.

All compounds that satisfy these criteria will be retrieved regardless of what other elements are present.

### Elements Present

Specify what elements **must** be in the compound. Compounds with elements not in the list will not be retrieved. There are two options

**Each hit contains all of these elements and no others - [All]**

**All elements in each hit are in this list of elements - [Some]**

The following table will illustrate the usage of the constraint.

Entry	Some result	All result
C, H, O	Each compound will have at least one of the elements C,H, or O and no other elements. The compound has <b>some</b> of the elements.	Each compound will have C and H and O and no other elements. The compound has <b>all</b> of the elements.

## Peaks Constraints

All retrieved spectra must follow the peak types and abundance you enter.

The allowed **peak types** are:

### **Normal:**

This is a conventional mass spectral peak; i.e.,  $m/z$  43 in hexane.

### **Loss:**

This is a neutral loss from a molecular ion. The maximum value is 64. If 0 is entered, then the spectrum must have a molecular ion peak; i.e.  $m/z$  15 for loss of methyl.

### **Rank:**

Order of peak in relative abundance from the largest (Rank 1) to the 16th (Rank 16). The base peak is the largest peak in the spectrum. The rank constraint can be used like the classic 8 peak index data.

### **Maxmass:**

The mass of the most significant high mass peak in your spectrum. Retrieved spectra will contain a peak at this mass within the specified abundance range and no significantly larger peaks at higher masses.

The abundance can be:

**Absolute.** Absolute will evaluate all entries as a percentage of the base (largest) peak - values from 0 to 100% are accepted. A 0% entry accepts spectra with no peak at the specified  $m/z$  value.

**Relative.** The intensity entry for the first peak for Relative abundance is exactly the same as the Absolute - 1 to 100% of the [base peak](#). Intensities of all other peaks are relative to that of this first peak. If a range from 50 to 999 is entered, the intensity of the specified peak in the retrieved spectrum must be between one-half and ten times that of the observed intensity of the first peak in the list. When Relative is selected, the allowed intensity range for peaks after the first is 1 to 999.

## Compound in Other Databases

With this constraint you may limit the search to only those compounds that are in other specific databases. The available databases are:

**Fine** - Commercially Available Fine Chemical Index

**TSCA** - Toxic Substances Control Act Inventory

**RTECS** - Registry of Toxic Effects of Chemical Substances

**EPA** - EPA Environmental Monitoring Methods Index

**USP** - U. S. Pharmacopeia/U.S.A.N.

**HODOC** - CRC Handbook of Data of Organic Compounds

**NIH** - NIH-NCI Inventory File

**EINECS** - European Index of Commercial Chemical Substances

**IR** - NIST/EPA Gas Phase IR Database

If the box next to the database is checked then the compound **must** be in the database for the search to report it. If more than one database is selected, a compound will be selected if it is in any one of them.

The box marked **Use constraints** must be checked or no constraints will be applied to the hit list.

## Tags in Comment and Text Info Constraint

Applies to: all searches  
Max. length: 1023 characters

### Tags in Comment

All information in NIST Peptide libraries is represented by the Tag=value convention. Typically, Peptide spectrum comment field may contain following items separated by spaces:

*Tag=string*

*Tag="string"*

*Tag=number*

Spaces may be inside a *string* only if the string is in double quotes. An unquoted *string* may also have space(s) inside a pair of parentheses, ( ). Spaces are not permitted inside any *Tag* or next to the equal sign.

If such a *Tag* is not found then the spectrum is rejected. If *Tag* is immediately followed by a question mark, for example, *Tag?=value*, then the spectrum is rejected only if the *Tag* has been found and its value does not match the constraint.

Comment fields may also contain *Tag* not accompanied by the equal sign and *string*.

All *Tag*, *string*, *substring*, and *word* comparisons described below **ignore letter case**.

Tags in Comment Constraint allows searching these tagged expressions.

See also [Displaying tagged comment fields](#)

### Tags in Spectrum Text Information

Text information displayed for NIST MS/MS Database mass spectra has special tagged text fields, for example:

Instrument type: QqQ

Spectrum type: ms2

Compound type: M

Precursor type: [M+H]<sup>+</sup>

Precursor m/z: 223.0745  
 Collision energy: 6  
 Instrument: Micromass Quattro Micro  
 Sample inlet: direct flow injection  
 Ionization: ESI  
 Collision gas: Ar  
 Pressure: 1.6mTorr  
 In-source voltage: 25

In MS Search versions starting from v.2.0g, Tags in Comment Constraint allows searching these tagged fields. See [Entering Constraint for Tags in Text Info](#).

## Entering Tags in Comment Constraint

Items in the *Tags in Comment* constraint field should be separated by spaces and/or be on different lines. Unlike spectra comments, spaces next to equal sign (as well as >, <, >=, <=, and :) are allowed inside the items.

Spectra returned must satisfy all entered constraints

### Examples

	Item	Comments of the output spectra must have
1	" <i>consensus</i> "	A Tag or a space-delimited string <i>consensus</i>
2	<i>Parent</i>	6 letter sequence <i>parent</i> anywhere in the Comment
3	<i>fullname="r.lecvk.c/2"</i>	Exactly this Tag=string <i>Fullname=R.LECVK.C/2</i>
4	<i>fullname=lec</i>	String following <i>Fullname=</i> contains sequence <i>LEC</i>
5	<i>Mods : ICAT_cl_hi</i>	String following <i>Mods=</i> contains <b>word</b> <i>ICAT_cl_hi</i>
	The following rules apply to strings that are numbers	
6	<i>Diffmod=6.8</i>	String following <i>Diffmod=</i> begins with number 6.8, for example, <i>Diffmod=6.8/8</i> or <i>Diffmod=6.8</i>
7	<i>Diffmod=5.1:6.8</i>	String following <i>Diffmod=</i> begins with

		number between 5.1 and 6.8, for example, <i>Diffmod=6.8/8</i> or <i>Diffmod=6.8</i>
8	<i>Diffmod</i> > 6.8 <i>Diffmod</i> >= 6.8 <i>Diffmod</i> < 6.8 <i>Diffmod</i> <= 6.8	<i>Diffmod</i> = followed by a number greater than 6.8 <i>Diffmod</i> = followed by a number not less than 6.8 <i>Diffmod</i> = followed by a number less than 6.8 <i>Diffmod</i> = followed by a number not greater than 6.8
Note: <b>word</b> is a string delimited by any character but a letter, a digit, or an underscore "_";		
If <b>word</b> contains spaces or equal sign it should be entered in quotes, "".		

To reverse the meaning of constraints 1-5 (that is, change to "spectra must not have"), enter caret ^ as the first letter:

1	^"consensus"	Tag <i>consensus</i> is not present
2	^parent	Substring <i>parent</i> is not present
3	fullname=^"r.lecvk.c/2"	String tagged <i>fullname</i> is not exactly <i>r.lecvk.c/2</i>
4	fullname=^lec	String tagged <i>fullname</i> does not have substring <i>lec</i>
5	Mods : ^ICAT_cl_hi	String tagged <i>Mods</i> has not word <i>ICAT_cl_hi</i>

## Entering Constraint for Tags in Text Info

Special tags are to be used for searching in the spectrum text information. To enter such a special tag, type dollar sign (\$) immediately (no space) followed by the tag displayed in the spectrum text information, replacing spaces, if present, with underscores. In the special tag, dashes, slashes, and underscores may be omitted, as well as # after letters "CAS". The search string and the special tag are case-insensitive. The syntax is almost the same as the one described above for Tags in Comment Constraint.

Below are several examples.



To search for	Enter into the Tags in Comment Constraint
<i>Precursor m/z</i> in range* 102.9–103.2	\$Precursor_m/z=102.9:103.2
<i>Precursor m/z</i> in range* 102.9–103.2	\$precursormz=102.9:103.2
<i>Instrument type</i> containing word QqQ	\$Instrument_type:QQQ
the exact <i>Instrument</i> name	\$Instrument="Micromass quattro"
spectra that have any <i>Related CAS#</i> entry	\$Related_CAS#
spectra that have any <i>Related CAS#</i> entry	\$relatedcas
<i>spectra that have no Precursor m/z</i> entry	^\$Precursor_m/z

\* Note: Spectrum with text information "Precursor m/z: 103,77" also will be found because the first number, 103, is in the range 102.9–103.2. Note that the actual precursor m/z for such a  $ms^n$  spectrum ( $n > 2$ ) is 77.

### List of acceptable special tags

\$aux	\$ion_mode	\$precursor_m/z
\$charge	\$ion_mw	\$precursor_type
\$collision_energy	\$ion_name	\$pressure
\$collision_gas	\$ionization	\$related_cas#
\$compound_type	\$known_impurity	\$retention_index
\$cone_voltage	\$mass_range	\$ri
\$inchikey	\$maximum_intensity	\$salt
\$in-source_voltage	\$msn_pathway	\$salt/mix_cas#
\$instrument	\$notes	\$sample_inlet
\$instrument_type	\$peptide_mods	\$special_fragmentation
\$ion_formula	\$peptide_sequence	\$spectrum_type

Special tags inside parentheses are considered identical:

(\$aux \$notes) (\$in-source\_voltage \$cone\_voltage) (\$retention\_index \$ri)

Letter case, slashes, dashes, and underscores in special tags are ignored.

Tags differing only in CAS, CAS#, or CASNO at the end of the tag are equivalent.

Some of these tags may be found in other NIST libraries, for example, \$related\_cas#, \$inchikey

See also [Spectrum Text Information](#).

For more details see the [formal description of tags in comment constraints](#) and [formal description of reversed tags in comment constraints](#)



# Peptide Constraints

**Important:** These constraints may be used in searches displayed in the Other Search tab; they are ignored for all Library searches except MS/MS search with Presearch Off and In-source HiRes search. The information necessary for applying the constraints must be present in the comments section of library spectra, in spectrum text information, in the spectrum name, which in case of peptide spectra should be the residue sequence and the charge, for example YYYIPQYK/2. NIST Peptide Libraries is an example of such a library. Alternatively, the residue sequence may be in the *Peptide sequence* field; in this case the charge usually is in the *Precursor type* field. MS/MS library **nist\_msms2** containing spectra of commercially available peptides is an example of such a library.

If a spectrum is not a MS/MS or in-source/EI with accurate ion m/z spectrum or it does not contain the necessary information and is located in a MS/MS or in-source/EI with accurate ion m/z library then the spectrum is rejected. If a spectrum is in EI library then the corresponding constraint is ignored.

**Peptide Sequence** - Restrict search results to spectra of peptides that either contain or do not contain specified residue sequences.

A sequence that starts with period must fit N-terminus (e.g., .AGG).

A sequence that ends with period must fit C-terminus (e.g., EHR.).

A sequence that does not include period must fit anywhere in the peptide chain.

Several residues in brackets mean "any of" (e.g., [ET]); a caret (^) following the opening bracket means "none of" (e.g., [^ET]).

A question mark means "any". Each item must be either separated by a space from the previous one or be in a separate line. The total length should not exceed 127 characters including 2 per each newline. For example,

.AGG [ET]HER

means: there should be AGG sequence on N-terminus; the peptide should contain sequence HER preceded by either E or T. For more details see [Formal description of peptide sequence constraint](#).

**Peptide Mobile Protons** - The number of mobile protons is the peptide charge less the number of sequestered protons. The number of sequestered protons is calculated as the number of Arginine, Lysine, and Histidine residues (R, K, and H, respectively) in a peptide sequence. The minimum value is zero.

**Peptide Number of Residues** may also be constrained.

## MS/MS Instrument Type Constraint

This feature will only work with NIST 11 or later MS/MS libraries.

It affects only [MS/MS spectra](#) and [in-source/EI with accurate ion m/z spectra](#); therefore, it ignores all EI spectra, which means applying this constraint to a search in the EI Database does not affect the Hit List.

This constraint has 7 options (check boxes) that may be combined to produce a Hit List containing such spectra that each spectrum fits one of the selected Instrument Types.

The Instrument type field in an MS/MS mass spectrum text information is interpreted in the following way:

Instrument type check box	Instrument type field in MS/MS mass spectrum
Ion Trap	"IT/ion trap" or "QQIT" or "QqLIT"
Ion Trap + FT	IT-FT/ion trap with FTMS
Q-TOF	Q-TOF
HCD	HCD
Triple quadrupole	QqQ
Other	any other than the above three Instrument types
Unspecified	No instrument type specified

Selecting all five would select all mass spectra (the same as if the constraint is disabled).

The version of Lib2NIST included in NIST 14 or a later version is needed to produce such a database because the instrument type is saved in the NIST 14 MS/MS mass spectrum record as a new numeric field; therefore, this constraint would not work with the NIST 08 MS/MS library or even the NIST Peptide libraries created with the previous version of Lib2NIST.

Available in NIST 11 MS Search Q-TOF/HCD check box was replaced with two, Q-TOF and HCD. Unless NIST 11 MS/MS Library has been rebuilt with NIST 14 version of Lib2NIST, selecting Q-TOF for searching this library is equivalent to selecting both Q-TOF and HCD in searching NIST 14 MS/MS Library.

## Substructure Analysis

The substructure identification relies on the analysis of the results of the user search. From an analysis of the substructures of the compounds in the hit list and the degree of match to the unknown, a set of parameters is derived to predict the probability of a set of substructures. The set is predetermined although you have the option of forcing the system to only examine a [subset of the structures](#). Not all substructures have been provided. For each substructure, there is a slightly longer explanation of the nature of the substructure that appears when the substructure is highlighted.

In addition to the probability of substructures being present, substructures that are absent are reported. You should not ignore this information since it can be an important aid in analyzing for the unknown.

Data on the molecular weight and the Cl/Br content are also provided.

## Substructure Identification

Mass spectra of compounds with similar chemical structural features often have similar features. This is the origin of the common finding that hit lists resulting from library searches for spectra that match a submitted spectrum, commonly contain a number of the retrievals with similar structural features. The underlying reason for this finding is that if different compounds fragment through a common intermediate ion, the signature (mass spectral pattern) of this intermediate ion should be found in their spectra. For instance, since the benzyl cation ( $C_6H_5CH_2^+$ ) is often generated in the fragmentation of ionized compounds containing this group, the benzyl signature will often be found within the mass spectra of these compounds. The ability to make use of this signature to deduce that an unknown compound contains a benzyl group depends on a number of factors, including: 1) the fraction of compounds containing benzyl that show this signature; 2) the uniqueness of the benzyl signature and; 3) the ability of the identification procedure to clearly perceive the benzyl signature.

The most obvious way of deducing the presence or absence of a substructure in an unknown spectrum is to look for its characteristic signature. This is the approach most analysts would take and is the basis of a number of proposed computer methods. One drawback with this approach is that the abundance of peaks that constitute these substructural signatures is often highly variable. Even more importantly, it is commonly found that the most informative characteristic signatures are for relatively complex chemical structures that are relatively rare and not likely to have been documented. However, if a large reference library is used that contains several compounds containing this unusual structural fragment, and they are placed high in a hit list, the presence of

this structural feature in the unknown compound can be clearly established.

The present software has been designed to produce and analyze hit lists to perceive these signatures. This is done by the so-called nearest neighbor technique as described in (S.E. Stein, J. Am. Soc. Mass Spectrom., 1995, 6, 644-655). This method computes match factors for each of approximately 500 pre-defined substructures from the relative match factors of compounds in the hit list that contain these substructures. These match factors are then converted to probabilities that the substructure is present or absent in the unknown compound on the basis of a large number of test runs.

To obtain Substructure Identification information from a library search hit list, select from the menu Tools / Substructure Identification.

## Other Search Window

### Molecular Weight (Nominal Mass)

Search for a single [molecular weight](#). For this search, any constraint other than Molecular Weight and Exact Mass may be selected in the [constraints](#) window.

The available and selected libraries for the search are shown in two list boxes located in the Options tab. Only the libraries in the right hand (Included Libs) list will be searched.

### Exact Mass

[Search for the compound exact or isotopic mass](#) using an exact mass value, exact m/z value, or chemical formula. Enter the search values in the Accurate Precursor Mass Search tab. For this search, any constraint other than Molecular Weight and Exact Mass may be selected in the [constraints](#) window.

The [available and selected libraries](#) for the search are shown in the Libraries tab.

### Formula

A chemical formula is entered in the list box, which also maintains a list of previous entries. Parentheses are completely resolved - thus  $\text{CH}_3(\text{CH}_2)_3\text{C}(\text{C}(\text{CH}_3)_3)_3$  is converted to  $\text{C}_{17}\text{H}_{36}$ . Additional [constraints](#) may be added. Note that not all of the normal constraints are active. MW(Nominal Mass), Exact Mass, Elements Value, and Elements Present are disabled. The specification of a formula defines the molecular mass and the elements so they cannot be additionally specified.

The [available and selected libraries](#) for the search are shown in two list boxes in the Options tab. Only the libraries in the right hand (Included Libs) list will be searched.

### Any Peaks

This allows searching using a [small number of peaks](#). Spectra can be searched using the largest peaks, or peaks in certain ranges of total intensity. In addition search by neutral loss can be done.

The [available and selected libraries](#) for the search are shown in the Libraries tab

### Sequential Search

Search through an entire database for matches that satisfy the options set in the [Constraints](#) tab.

The [available and selected libraries](#) for the search are shown in the Options tab

## CAS Number

Search for compounds with specified CAS number (dashes may be omitted). The program will report if the CAS number is not legitimate. A small fraction of spectra in NIST 14 MS Library have additional CAS numbers marked *Related CAS#* and *Salt/Mix CAS#*. These CAS numbers are also included in the search.

The [available and selected libraries](#) for the search are shown in two list boxes below the CAS number entry box. Only the libraries in the right hand (Included Libs) list will be searched. No constraints are allowed.

## NIST Number

Each spectrum in the NIST library is assigned a unique number, the NIST Number. A set of NIST numbers may be entered as a range, for example, 200001-200010.

The [available and selected libraries](#) for the search are shown in two list boxes below the NIST number entry box. Only the libraries in the right hand (Included Libs) list will be searched. NOTE: Typically, only NIST main, NIST replicate, nist\_msms, and NIST RI libraries can be searched with this mode. Only searchable by NIST Number libraries are displayed in the Available Libs list in the NIST Number search mode. No constraints are allowed.

## ID Number

Each spectrum in a library is given a unique identification number. This is a useful way to examine all or portions of user libraries. When you specify a library in the dialog that appears when you select ID Number as the search mode, you will see the library statistics. For example 919 spectra 1-1308 ID range. This specifies that there are 919 spectra in the data base and they have ID numbers from 1-1308. Each time you edit a spectrum, the old ID number is deleted and a new one specified. If you were to search this library with ID number 1-1500 you would get all 919 spectra. No constraints are allowed.

## Controlling Window Display

**Hit List** - On the left side of the Other Search Window is the Hit List. The Hit List contains the results of the search represented in the drop down list in the Short Cut Bar. Unless otherwise specified, this is the most recent Hit List. Using a tab at the bottom of this window, the Hit List can be displayed as names or structures. Hit list may be sorted by clicking a header of any column except # and Lib.

**Spectral Display** - On the right hand side of the Other Search Window is a Spectral Display window - the orientation can be changed with the right mouse button, the relative sizes by moving the splitter bar. The display can be switched between Plot/Text and Plot.

## Short Cut Bar



**Start Search** - Button with binoculars on it (search).

**Select Search mode** - Select type of search

**History list** - Select hit list to display from previous searches.



**Clear History** - Clear hit list history. Delete associated data files.



**Replicates On/Off** - Show replicate spectra for compounds in hit list. Only operates on NIST Main and Replicate library data.



**View hitlist search options** - Shows the constraints used when generating the currently shown hit list and libraries involved.

## Any Peaks Search

All retrieved spectra must follow the peak types and abundance you enter.

The allowed **peak types** are:

### **Normal:**

This is a conventional mass spectral peak; i.e., m/z 43 in hexane.

### **Loss:**

This is a neutral loss from a molecular ion. The maximum value is 64. If 0 is entered, then the spectrum must have a molecular ion peak; i.e. m/z 15 for loss of methyl.

### **Rank:**

Order of peak in relative abundance from the largest (Rank 1) to the 16th (Rank 16). The base peak is the largest peak in the spectrum. The rank constraint can be used like the classic 8 peak index data.

### **Maxmass:**

The mass of the most significant high mass peak in your spectrum. Retrieved spectra will contain a peak at this mass within the specified abundance range and no significantly larger peaks at higher masses.

### **Exact Mass:**

Note: this peak type applies only to **mainlib, replib, and nist\_msms** NIST MS libraries

This peak type is somewhat similar to Normal, but the mass is accurate. An accurate mass (Mass), an m/z value (adjusted for the electron mass in assumption of a singly charged positive ion), or an elemental composition of the ion (Formula) can be entered. The abundance range is from 0 to 100% of the base peak. 0% means abundance < 5 for base peak intensity = 999. Only monoisotopic peaks may be searched.

The uncertainty can be entered in millimass units (mmu) or parts-per-million (ppm). If this field is left blank, an uncertainty of 0.5 in the next to the last digit entered for the mass or m/z is assumed. In this case the uncertainty always is a 0.5 value; i.e., the

uncertainty for 120.12 is 0.005 and 120.120 is 0.0005.

It is important to select the type of what you enter (mass, m/z, or formula) and the units of the uncertainty if it is not left blank.

To enter a chemical formula, select *formula* in the drop-down list box located above the first entry field (this is possible only if “Exact Mass” type has been selected.) If the “Uncertainty” field is left blank, the accuracy of searched monoisotopic mass calculated from the formula is exactly the same as the accuracy of the mass saved in the database.

It is very rare the exact masses of two different elemental compositions are the same; therefore, there is a very high probability that only the spectra where the peak was interpreted with the same elemental composition will be found.

The exact mass values for the peaks in the NIST 14 MS libraries, mainlib, replib, and nist\_msms, have been calculated from the elemental compositions of ions predicted by MS Interpreter only for peaks that are present in a mass spectrum. **Isotopic** interpretations of the peaks are not included in the library. Some of the peaks do not have interpretation; therefore, these peaks cannot be found with the Exact Mass peak type. More than one exact mass may be saved in a library for a peak in a mass spectrum of a compound. For instance, both CO and C<sub>2</sub>H<sub>4</sub> may be assigned to the same peak with m/z 28. Searching for these two exact mass peaks entered as formulas would produce a number of hits containing both peaks. These hits correspond to mass spectra where the peak at m/z 28 was ambiguously interpreted. In an Any Peaks Search of the NIST 14 mainlib, using the “Exact Mass” Peak type with the above two formulas and intensity range of 10–50%, 1 peak is found in 6,189 spectra and 2 peaks are found in 1,396.

## Notes

You must enter the first three fields in the order from left to right. <Enter>, <Tab>, or the mouse may be used to move from field to field. The Accept button or Enter key will put a line of data in the **peaks** area (see figure below). The Delete button will clear the selected line in the **peaks** area. Double-clicking the line will move data from the line into the entry fields.

Changing the "Type" field will result in the clearing the three entry fields on the right from the selected type. Up to 11 peaks can be entered.

Type	m/z	From	To	Hits	Peaks	Hits
N	28	50	100	1078	1	5111
N	44	50	100	4229	2	98

**p e a k s**

**cumulative  
information**

The "Hits" column in the **peaks** area displays the number of spectra in all the libraries searched that contain this peak in the specified intensity range. The "Hits" field in the **cumulative information** is the number of spectra which contain the corresponding number of peaks. This value will change each time another peak is placed in the **peaks** area. The number of peaks in the **Peaks Required** field at the bottom of the window is, by default, the total number of peaks in the **peaks** area. Each spectrum in the hit list will contain at least this number of entered peaks.

When **Use Constraints** is selected, the number in **Max Num of Hits** limits the size of the hit list. Otherwise, max. number of hits is 6000. If the number of hits exceeds 6,000 no output is produced and an error is displayed.

The **abundance** in all **peak types** can be only Absolute. It will evaluate all entries as a percentage of the base (largest) peak - values from 1 to 100% of the [base peak](#) are accepted, except for Exact Mass, where abundance = 0% is allowed.

## Librarian Window

Initially the Librarian Window appears to be the same as the Other Search Window. However, the Librarian Window does not have any search functions. The spectra displayed are exactly the same as in the Spectrum List (Spec List) of the Library Search Window. The difference is that in this window the spectra present can be edited, added to a user library or removed from a user library.

### Controlling Librarian Window Display

**List** - On the left side of the Librarian Window is the Spectrum List (Spec List). The List can be displayed as names or structures.

**Spectral Display** - On the right hand side of the Librarian Window is a Spectral Display window - the orientation can be changed with the right mouse button, the relative sizes by moving the splitter bar. The display can be switched between Plot/Text and Plot.

### Editing/Adding Spectra

To create a new spectrum, select the new document icon . To edit the currently highlighted spectrum select the "ed" icon . In both cases the same window will appear. If you are editing an existing spectrum all of the associated data will appear except accurate m/z and intensity list. If you are creating a new spectrum all of the fields will be blank.

### Editing/Adding Structures

Within the editing window you will find two options for adding a structure. The first "Attach Struct" allows the import of a structure from either an external file (only \*.mol and \*.sdf files are supported), or from another library. The other option "Clipboard Struct" allows importing a structure from the standard Windows clipboard. Note, only some formats are supported within the clipboard. It is possible to use unsupported software tools, but only if the program can support direct input and output of \*.mol files.

### Adding Spectra to a User Library

To add one or more spectra to a user library (and possibly create the library) select spectra in the Spec. List and click **Add to Library** button on the **Short Cut Bar**. Set the destination library in **Choose Library** dialog and select OK to add the spectra. Spectra

cannot be added to NIST libraries.

## Deleting Spectra from a User Library

To delete one or more spectra from a user library, place these spectra into Spec. List, select them, and click **Delete from Library** button on the **Short Cut Bar**. Before the deletion occurs, MS Search will display a message: "*You are about to delete highlighted spectra from <library>*" If you select "Yes" to dismiss the message, the spectra will be irreversibly deleted. If you select "No" the spectra will not be deleted.

## Short Cut Bar



**File** - Opens a file for import of spectra. Spectra imported will be placed in the Spectrum List (Spec List)



**Export from Libraries** - Directly export data by ID number from a library to the Spectrum List.



**Export** - Export selected data from Spectrum List (Spec List) to standard output text file. Two output formats are available: MSP and JCAMP. The former is the native NIST MS Search format for all mass spectra. The latter is mostly for EI and [in-source / EI accurate ion m/z spectra](#); not recommended for [MS/MS spectra](#). To export chemical structures only, choose MOL or SDF format.



**New Spectrum** - Open [Spectrum Edit Window](#) for entry of new data.



**Edit Spectrum** - Open [Spectrum Edit Window](#) with data from currently highlighted entry. Note - data in the NIST main and replicate libraries, as well NIST MS/MS and RI libraries may not be edited. They may, however, be copied into the Editor, edited and saved into a new library or Spec List.



**Delete from Library** - Remove the highlighted spectrum from the library. Multiple spectra may be selected and removed. Does not affect NIST libraries.



**Add to Library** - Adds the highlighted spectrum to a library. Multiple spectra may be selected and added. Does not affect NIST libraries.



**Move to Library** - Removes the spectrum from one library and puts it in another.

Multiple spectra may be selected and moved. Does not affect NIST libraries.



**Delete Library** - Deletes the entire library. Does not affect NIST Mainlib and Replib libraries.



**Create Library** - Creates a new user library. The total number of libraries, including Mainlib and Replib, cannot exceed 127. The program would reject a request to create the 128th library.

See also [Keyboard shortcut keys](#). Typical useful keys are Ctrl-A, Del, Ctrl-K.

## Editing a Spectrum

Spectra can be added or edited from this dialog.

Note - data in the NIST main and replicate libraries, as well NIST MS/MS and RI libraries may not be edited. They may, however, be copied into the Editor, edited and saved into a new library or Spec List.

The **name** of the compound that is primarily used is entered in the Name field.

The **formula** can be entered if known, in addition if the structure has been entered, the program can take the formula from the structure by pressing the From Structure button.

**Synonyms** can be entered, one per line, in the Other Names field. Note, these names can also be notations from lab notebooks or original file references. They are searched in the same way as the name. Synonyms, which begin with an [internal tag](#) (e.g., \$:04) contain tagged [text information](#). See also [Tags in Spectrum Text Information](#).

**Comments** is a free text field; may contain only ASCII characters.

The **molecular weight** will be calculated from the formula, but if there is no formula it may be entered. It will be checked for consistency with the formula.

The **CAS number** is checked for validity. It need not be entered.

Individual peaks may be entered, edited or deleted. To delete, highlight the peak (m/z, Abundance pair) and press delete. To edit, highlight the peak and edit in the three upper boxes. Each entered peak must be accepted. Note accurate m/z and abundance peak list is saved for [MS/MS](#) and [in-source/EI with accurate ion m/z](#) spectra automatically. If you create accurate (non-integer) m/z peaks, to save them with accurate m/z, select **HiRes Spectrum** check box

You can attach a structure. If a structure is present in the clipboard, the **Clipboard Structure** button will be active. Otherwise, the **Attach Structure** button only will be active. Pressing the [Attach Structure](#) button will bring up a dialog to allow you to bring a structure from other user libraries or from a mol or sdf file.

To create or edit RI data, click **Edit RI** button. A spectrum may have up to 3 RI values of different column type, or a single Any column type (n-alkane standard) RI, or a single Unspecified column type RI.

Once the editing is complete choose **Add to Library** to add the spectrum and structure

to an existing library (the option of creating a new library if fewer than 15 user libraries exist is also available). Choose **Replace** to replace the spectrum in the library from which it came (the option is not available for the NIST library). In addition, the spectrum can be added to the Spectrum List (Spec List) with **Add to List**.

To open Spectrum Information dialog, click Edit Spectrum button .

## Choose Library

Actions on libraries - select the library you want to add or move a spectrum to, select the library you want to delete, bring up a selection box with the current set of libraries. There is a limit of 127 libraries including Mainlib and Replib at any one time.

If there are fewer than 127 libraries, the program will allow the specification of a new library name for adding or moving data. To create a new library and choose it for adding spectra, type the new library name instead of choosing from the list.

To choose a library for editing, switch to Librarian window, select a spectrum unless you want to delete a library or create a new empty library. Select the button you need:



Add spectrum to a library



Move spectrum to a library



Delete a library



Create a new library

In all cases except the deletion of a library, the new library name may be typed into the Choose Library dialog or the existing user library name may be selected from the list below. The new library will be created and the selected spectrum, if any, copied or moved into the library.

**RI type if unspecified** drop down list box. Choose Column Type for RI values in the added to the library spectra. The selected column type will be saved in the library spectrum if column type for RI in the spectrum is not specified. Unspecified column type may be left unchanged. RI values must be between 1 and 16,383.

## MS/MS Window

This window permits browsing of MS/MS spectra, also known as tandem mass spectra. They are sorted first by their nominal (integer) molecular weight of the precursor species, then by its formula and then by the level of fragmentation. In some cases, precursor unfragmented spectra are not shown (MS1, also arranged by precursor ion, P, and 'cone' voltage, CV). Tandem spectra are also distinguished by their level (MS2, MS3, .), their precursor m/z, P, and their collision energy, CE (ev for collision cells and % for ion trap).

For beam-type instruments (e.g., triple quadrupoles and Q-tof instruments), the ms2 designation refers to spectra generated by dissociation in a collision cell of a selected precursor ion. For ion trap instruments, the ms2 designation refers to dissociation in the trap following the isolation of the precursor ion.

Only nist\_msms library contents are displayed. nist\_msms2 library and other MS/MS libraries are not displayed in the MS/MS window. In absence of nist\_msms library, the MSMS tab is not displayed.

### Controlling MS/MS Window Appearance

**Fragmentation Tree.** The first time the MS/MS window is selected, the three panes of the display are blank with the exception of a single entry in the ion fragmentation pane, which is on the left of the window display. This entry is the MS/MS Library. There is a box on the left that has a plus sign (+). Placing the mouse pointer on this box and clicking expands to the first two levels (Positive Ions and Negative Ions). There are plus boxes next to both Positive Ion and Negative Ion. Clicking on one of these plus boxes results in the display of the next level. This is a list of molecular weights of analytes for which a spectrum was obtained by MS/MS. Clicking on the plus box next to one of these molecular weight listings results in the display of a list of elemental compositions that will also have plus boxes next to them. Click on one of the elemental composition check boxes will result in a display of a list of the compound names that have that elemental composition. In most cases, this is only a single compound. Clicking on the plus box next to the compound name will result in the display of a list of the spectra (ms1, ms2, etc.) acquired under MS/MS conditions for this compound. If more than one spectrum was acquired, there is a separate entry for each spectrum. Clicking on the spectrum will result in displaying the spectrum plot and text information.

**Spectral Display** - On the right hand side of the MS/MS Window is a Spectral Display window - the orientation can be changed with the right mouse button, the relative sizes by moving the splitter bar. The display can be switched between Plot/Text and Plot.



## Compare Window

The compare window is designed for comparison of spectra only - it has no searching capabilities. It allows for any set of spectra to be compared and unlike the Library Search Window, it can show multiple spectra.

### Controlling Compare Window Display

There are three window areas in the Compare Window. The top window is a Plot only version of the Spectral Display Window. The middle is a standard Difference Display Window and the bottom is a window for displaying multiple spectra in Plot mode only. In general the lower window will contain the first 16 results (default) from the Library Search. Using the properties (Right Mouse Button) dialog box you may set the number of spectra displayed in the lower set of windows as well as the number contained in the window. As you scroll down the window the display brings up the later results from the search.

### Difference Display Options

The Compare Window uses the [standard difference display](#) options. These are common for all spectral difference displays in the program.

# Names Search Window

## Incremental Name Search

As the name is entered, one letter at a time, the list of matching names below is updated. The clear button clears the entry box and the a-z button allows only alphabetic characters to be entered. Otherwise both alphabetic and numeric characters can be entered.

## Controlling Name Search Window Display

**Name List** - The name entry box is at the upper left of the screen, matching chemical names are shown below. As each letter of the entered name is entered, the list of names is updated.

In NIST MS/MS libraries **nist\_msms** and **nist\_msms2**, the name is followed by additional items, namely, Precursor type or MS<sup>n</sup> pathway, Instrument type, Collision energy, and rounded Precursor m/z, for example,

11-Ketofluprostenol [M-H]- HCD 81V P=455.2  
11-Ketofluprostenol [M-H]- IT 35% P=455.2  
11-Ketofluprostenol [M-H]-=>233.3=>189.1 IT 35% 81V P=189.1  
11-Ketofluprostenol [M-H]- IT-FT 35% P=455.2

These additional items may not be entered for searching the name. Instrument types include IT, IT-FT, QQQ, QTOF, and HCD.

## Spectral Display

On the right hand side of the Names Search Window is a Spectral Display window - its orientation can be changed with the right mouse button, relative sizes can be changed by moving the splitter bar. Tabs permit switching between Plot/Text and Plot. The Spectral Display may be updated with each letter entered or the update may be delayed until the correct name is found and selected. To change the behavior use the right mouse button to select properties.

**Note.** Retrieving and displaying all available Retention indices may significantly increase the time of the program reaction to entering the next name character. To reduce this time, in the properties dialog, select **First** radio button under **Retention indices** check box and enter a small number, for example, 2, in the field to the right from **First**. Retention indices are not displayed in MS/MS window.

## Short Cut Bar

**Name Entry** - Enter the name, each additional letter causes updating.

**Clear** - Clears the entry box.

 **Alpha only entry** - Allow alphabetic only or alphabetic and numeric entry. If alpha-numeric entry is allowed "1,2-dimethyl" is allowed (the comma and dash will not be accepted), with alphabetic entry only the "dimethyl" will be accepted and numbers in the library names are ignored.

**Library** - Specify the library that will be searched.

**Select Replicate** - appears only if **mainlib** library is selected and there are replicates in replib. Allows to choose the spectrum to display. Replicate spectra display is affected by settings of [Replicate Display Options](#). The set of displayed replicate spectra does not change automatically after changing Replicate Display Options. You to select another name and then re-select the name again.

# Controlling Library Search Window Appearance

The right mouse button brings up a number of options. In every window the last option is Properties. Selecting this option brings up a tabbed dialog box with controls for how the data in a window is displayed. The program will bring the correct dialog corresponding to the window the mouse is in when the right mouse button is selected.

For the Library Search Window the following controls and corresponding windows are available:

## Hit List

## Plot/Text of Spec List

The options for these two window are controlled by using the [common list control dialog](#)

## Plot of Hit

## Unknown Plot (Search Spectrum Plot )

## Spec List Plot

The options on the three plot windows use [common plot control dialog](#).

## Spec List Text Info

## Hit Text Info

## Unknown Text Info (Search Spectrum Text Info)

The options on the three text windows use a [common text control dialog](#).

## Comp. Result

The options on the comparison of results window use a [common comparison dialog](#)

## Histogram

The controls for the histogram are unique to this window. They are:

**Bar Width** - The width of the bin for the histogram in match factor points. A value of 10 will group all data with a match factor of 900 to 909 in a single bin.

**Logarithmic Y Axis** - If checked, the Y axis will be in logarithmic units rather than linear.

**X Axis : Automatic** - If checked then the values entered will keep the X axis of the histogram fixed. This is especially useful if comparing multiple hit lists to see how the distribution of the results changes.

**Show Libraries Searched** - If checked the libraries used in the search will be displayed.

**Color Settings** - The colors of the axis (scale), graph (bars), highlighted - to show where in the histogram the item selected in the hit list falls - and background can be specified.

**Font Settings** - The font either for this window or for all windows in the program can be set. Only a limited set of fonts are available, not the entire installed base of fonts from the system.

## Controlling Other Search Window Appearance

The right mouse button brings up a number of options. In every window the last option is Properties. Selecting this option brings up a tabbed dialog box with controls for how the data in a window is displayed. The program will bring the correct dialog corresponding to the window the mouse is in when the right mouse button is selected.

For the Other Search Window the following controls and corresponding windows are available:

### Hit List

The options for this window are controlled by using the [common list control dialog](#) - note that there are no match factor related choices since the other searches do not have match factors.

### Plot

The options for this window use the [common plot control dialog](#).

### Text Info

The options the text windows use the [common text control dialog](#).

## Controlling Name Search Window Appearance

The right mouse button brings up a number of options. In every window the last option is Properties. Selecting this option brings up a tabbed dialog box with controls for how the data in a window is displayed. The program will bring the correct dialog corresponding to the window the mouse is in when the right mouse button is selected.

For the Name Search Window the following controls and corresponding windows are available:

### Plot

The options for this window use the [common plot control dialog](#).

### Text Info

The options the text windows use the [common text control dialog](#).

### Incremental

**Structure Size in %** - Setting will change the size of the structure display if structures are being displayed. Value is a scale factor not an absolute number of pixels.

**Automatically display selected spectrum** - If checked, the structure and spectrum will be displayed as the name is typed. On newer computers this will not slow down the searching, but on older computers the searching for the name will be very slow if the structure is displayed as you type. If the box is unchecked, it is necessary to double click on the name to bring up the spectrum and the structure.

**Color Settings** - The colors for the bonds and rings and the atomic symbols of the structures, the compound name when shown with a structure, the active background - that is the currently active name, the list background and the text can be set.

**Font Settings** - The font either for this window or for all windows in the program can be set. Only a limited set of fonts are available, not the entire installed base of fonts from the system.

## Controlling Compare Window Appearance

The right mouse button brings up a number of options. In every window the last option is Properties. Selecting this option brings up a tabbed dialog box with controls for how the data in a window is displayed. The program will bring the correct dialog corresponding to the window the mouse is in when the right mouse button is selected.

For the Compare Window the following controls and corresponding windows are available:

### Plot

The options on the three plot windows use [common plot control dialog](#).

### Comp. Result

The options on the comparison of results window use a [common comparison dialog](#)

### Comp. List

The options are similar to a [common plot control dialog](#) except for the addition of the option to

**Insert best library search hits** - If checked, the number of hits specified from the top of the hit list is automatically inserted to the compare list.

**Overwrite contents of the list** - if checked will automatically overwrite the current contents of the list with new information. Otherwise the data will be added to the top of the compare list.

**No structures** - if checked will not put structures on the compare window.

**Spectra per page** - specifies the number of structures from the compare list that are shown below the compare window.

## Controlling MS/MS Window Appearance

MSMS Window reflects the contents of nist\_msms library.

The right mouse button brings up a number of options. In every window the last option is Properties. Selecting this option brings up a tabbed dialog box with controls for how the data in a window is displayed. The program will bring the correct dialog corresponding to the window the mouse is in when the right mouse button is selected.

For the MS/MS Window the following controls and corresponding windows are available:

### **Plot**

The options for this window use the [common plot control dialog](#).

### **Text Info**

The options the text windows use the [common text control dialog](#).

# Controlling the Librarian Window Appearance

The right mouse button brings up a number of options. In every window the last option is Properties. Selecting this option brings up a tabbed dialog box with controls for how the data in a window is displayed. The program will bring the correct dialog corresponding to the window the mouse is in when the right mouse button is selected.

For the Librarian Window the following controls and corresponding windows are available:

## **Plot/Text of Spec List**

The options for this window are controlled by using the [common list control dialog](#) note that there are no match factor related choices since the other searches do not have match factors.

## **Plot**

The options for this window use the [common plot control dialog](#).

## **Text Info**

The options the text windows use the [common text control dialog](#).

## List Control Dialog

This topic covers 3 lists displayed by MS Search: [Lib. Search Hit List](#), [Other Search Hit List](#), and [Spectrum List](#), which is displayed in [Lib. Search](#) and [Librarian](#) windows. [Names Search List](#) is not covered here

### Structure View Options

**Structure Size in %** - Setting will change the size of the structure display if structures are being displayed. Value is a scale factor not an absolute number of pixels.

**Structures Only** - If checked, only the structures will be shown in the Structures hit list. If unchecked, the Name of the compound and other information will be displayed. The information to be displayed is set by:

### Items to Display

**Order Number** - The number of the item in the list. For hit lists derived from matching to either a structure of a spectrum this is the order of the hit list. For other lists it is arbitrary.

**Brief/Full/No Origin Name** - Radio buttons - that is one must be selected. The origin of a spectrum is either a text file, the library editor, the main or replicate NIST libraries or one of the user libraries. Brief uses a one or two letter code for libraries and a short name for text or editor spectra. Full will use the full library name.

**RI** - if selected, Retention Index column is displayed. See [Retention Index Option](#). Note that in Lib. Search window, RI column is not displayed in case of MS/MS or In-source HiRes search. RI values are not displayed for [replicate spectra](#).

**No. Synonyms** - If selected, **Syn** column is displayed. It contains number of synonyms in a spectrum.

**No. Other DBs** - If selected, **DBs** column is displayed. It contains number of other databases that have the compound.

### Library Search Specific Controls

Search results are displayed in [Hit List](#).

**Match Factor** - If selected the match factor of the submitted spectrum or structure to the

library will be shown along side the name in both the text (names) and structures view of the list

**Reverse Match Factor** - If selected the [reverse match factor](#) for the submitted spectrum to the library will be shown. Note, this value only has meaning for matches of spectra to the library, not structures.

**Probability** - If selected the probability of the unknown spectrum arising from the same compound given in the list- assuming the correct match is in the database.

## MS/MS and In-source HiRes Search Specific Controls

To reach these controls, select [MS/MS tab is Library Search Options](#) and click [More MS/MS Options](#) button to display Peptide Options. The controls are in the Display frame. For these two searches, selecting **Match Factor** displays *Score* column, selecting **Reverse Match Factor** displays *Dot Product* column. Selecting Reverse/Impure in [More MS/MS Options](#) displays *Rev-Dot* column.

## Common Controls

**Color Settings** - The colors for the bonds and rings and the atomic symbols of the structures, the name when used with the structure, the background for the structure and text and text background can be set.

**Font Settings** - The font either for this window or for all windows in the program can be set. Select Unicode fonts to properly display Greek characters in chemical names. **Set for all views** button sets the selected font for all views.

To open List Control Dialog, use right mouse button in a list window and select Properties.

## Text Control Dialog

The following controls are available:

### Display

**Compound Information** - The name, formula, molecular weight, comments, etc for the compound.

**Ten Largest Peaks** - The intensity and m/z values for the ten largest peaks.

**m/z Intensity List** - The entire spectrum

**Synonyms** - The synonyms of the compound if known

**Retention Indices** - The values of the retention indices in the database.

**Note.** Retrieving and displaying all available Retention indices may significantly increase the time of the program reaction to entering the next name character in the [Names Search Window](#). To reduce this time, select **First** radio button under **Retention indices** check box and enter a small number, for example, 2, in the field on the right from **First**.

Retention indices are not displayed in MS/MS window.

**Arrange Peaks by** - Rows or Columns - the peaks are given in m/z order either across the column or down the row

**Wrap text** - if checked all text is wrapped and made to fit on the window. Otherwise the window will allow the text to spill out and indicate this with a scroll bar.

**Noise level %** - if a non-zero value is entered, peaks below that level (as a percentage of the base peak) will **not be displayed**.

**Color Settings** - The titles (field names), text, masses (m/z values), abundances (intensity) and background colors can be set.

**Font Settings** - The font either for this window or for all windows in the program can be set. Select Unicode fonts to properly display Greek characters in chemical names. **Set for all views** button sets the selected font for all views.

To open Text Control Dialog, user right mouse button to open context menu in a spectrum text information window and select Properties.



# Plot Control Dialog

The following controls are available:

## Options

**Label peaks greater than (%)** - if checked the program will use the value entered to determine which peaks to label with m/z values. If not checked the program uses an algorithm to label peaks in a manner that does not cause the labels to overlap.

**m/z Label** - if checked m/z labels are displayed above the mass spectral peaks

**Wrap legend** - if selected the program will wrap the legend below the plot. Normally the legend has the name of the spectrum, but additional components can be added.

**Text Label** - if checked the program will display peptide-specific peak labels stored in peptide spectrum peak annotations or certain other annotations.

**Peak Width** - set width of bars in mass spectrum bar graph display in pixels.

## Show

**Molecular weight** - if selected, the [molecular weight](#), if known, will be added to the legend.

**Formula** - if selected, the formula, if known will be added to the legend

**CAS Number** - if selected, the Chemical Abstracts Registry number, if known will be added to the legend

**Structure on Plot** - if selected, the structure will be placed on the plot if the structure is known. The program attempts to find space on the spectrum to place the structure. If there is insufficient space it will print only the formula.

**Color Settings** - The axis (scale), peaks, labels on the plot can be set to any color, in addition the bonds and rings and the atomic symbols of the structures, if drawn, can be set. Finally the background can be set.

**Font Settings** - The font either for this window or for all windows in the program can be set. Select Unicode fonts to properly display Greek characters in chemical names. **Set for all views** button sets the selected font for all views.

To open Plot Control Dialog, use right mouse button to open context menu in a spectrum plot window and select Properties.

## Comparison Control Dialog

The following controls are available:

**Label peaks greater than (%)** - if checked the program will use the value entered to determine which peaks to label with m/z values. If not checked the program uses an algorithm to label peaks in a manner that does not cause the labels to overlap.

**Structure on Plot** - if selected, the structure will be placed on the plot if the structure is known. The program attempts to find space on the spectrum to place the structure. If there is insufficient space it will print only the formula.

**Color Settings** - The axis (scale), Pure Upper peaks - belonging only to the top spectrum, Pure Lower peaks - belonging only to the lower spectrum, Mixed peaks - belonging to both spectra or difference peaks and background are all set here. Depending on the view you use for the difference (Difference, Side by Side, Head to Tail, Subtraction) some of these color choices may be omitted.

**Font Settings** - The font either for this window or for all windows in the program can be set. Select Unicode fonts to properly display Greek characters in chemical names. **Set for all views** button sets the selected font for all views.

To open Compare Tool Properties Dialog, user right mouse button to open context menu in the Compare window and select Properties.

## m/z range

**This option - on the menu bar under Tools - is only active in the Library Search and Compare windows.** It is also available on the main windows tool bar  .

With this dialog you may choose between

**Normal display** - Display data as intensity vs. m/z for the ion.

**Neutral Loss display** - Display data as intensity vs. neutral loss from molecular ion. (Default)

In the Neutral Loss display all of the x-axis values are negative except the molecular ion. It is also possible to [change the anchor point](#) for the neutral loss from the molecular ion to some other ion.

For both displays (normal and neutral loss) it is possible to set how the display axis is controlled.

**Scale as search/spec list spectrum** - use the m/z scale of the search spectrum to control all other displays. All m/z ranges will be set to the search spectrum. In this case as you scroll down a hit list you will see a common axis that full contains the search spectrum axis. If the hit list spectrum has data at m/z values below or above the search spectrum the axis will indicate this with a large arrow at the appropriate end.

**Scale as hit list spectrum** - use the m/z scale of the hit list spectrum to control the display. In this case the m/z ranges will be set to the hit list spectrum. As you scroll down a hit list you will see the axis shift to that of the hit list spectrum. If the search spectrum has data at m/z values below or above the search spectrum the axis will indicate this with a large arrow at the appropriate end. **NOTE** - in this mode the x-axis may shift with each new spectrum.

**Min/max of both spectra** - similar to above, but now there will never be an arrow to indicate data not shown. Program will use the smallest of the masses on the two spectra for the minimum and the largest of the two maxima for the maximum of the axis. All data for both spectra are visible.

**Fixed** - Set the start and end values of m/z manually. These will not change. If the either spectra has data at m/z values below or above the search spectrum the axis will indicate this with a large arrow at the appropriate end.

## **Set Anchors:**

Allows the reference point (anchor) for the neutral loss to be changed. The dialog box that appears allows setting of the anchor for the search spectrum, hit list spectrum and spectrum list (Spec List) spectrum. Normally these default to the [molecular weight](#) (MW) or the largest mass. If the compound formula (or molecular weight) is known, then the molecular weight is used. Otherwise the default mass for neutral loss is the maximum mass in the respective spectrum.

It is possible to set a specific mass as the anchor. In the Library Search window, this setting can be different for the search and the hit list spectrum. In addition, it is possible to set a value to be used for an entire spectrum list (Spec List) - each spectrum taken from the list will use the specified value. For the Compare Window, the anchor can be set for the search spectrum and the highlighted spectrum.

## Displaying tagged comment fields

If comment/contributor field of a spectrum contains specially formatted expressions,

Tag=value  
or  
Tag="value"

where Tag is any combination of letters, digits, and underscore characters; the value should be quoted if it contains spaces then these expressions may be displayed on the plot and in separate lines in textual display.

To display these tagged fields, enter the tags, either each on a different line or separate them by spaces.

Selecting **Show tag=value on plot** would display tags also on the plot.

These settings affect spectra display in all windows.

Note that tag Contributor is set by the installer to display Contributor field in nist\_msms2 library.

To reach this control, select from MS Search menu View / Comment field options.

See also [Tags in Comment Constraints](#).

## Replicates Display Options

Replicate spectra are spectra of the same compound. There are two ways of finding replicate spectra. The traditional approach is based on CAS Registry numbers: replicate spectra have the same CAS r.n.

Another approach is based on application of IUPAC International Chemical Identifier (InChI) to chemical structures and works only for spectra from Main and Replicate Libraries.

To use this structure-based method in addition to CAS r.n., select **Display InChI Replicates** check box.

With all other options cleared, certain tautomers and structures differing only in stereochemistry or isotopic substitutions are considered as replicate spectra.

InChI Options have the following meaning:

**Isotope** - Do not consider spectra of different isotopes of the same structure as replicate spectra.

**Stereo** - Do not consider spectra of different E/Z or R/S isomers of the same structures as replicate spectra.

**Derivatives** - Consider spectra of various derivatives of the same compound as replicate spectra. Types of derivatives are listed in the [Show Derivative Parents](#) dialog.

These options display only replicates located mainlib and replib

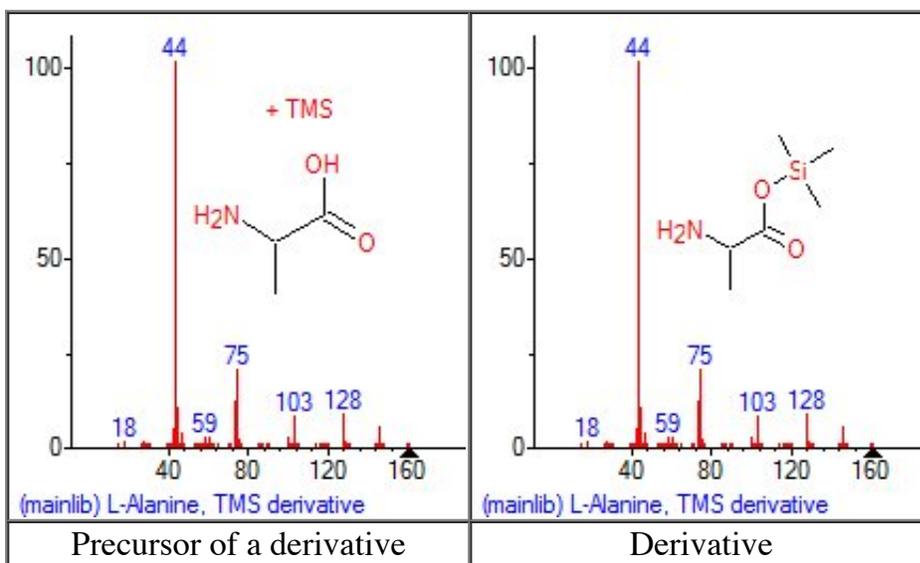
To access Replicates Display Options, select Replicates in the Options menu.

## Show Derivative Parent

To open "Show Derivative Parents for" options dialog, select Lib. Search window, then select "Derivative Parent" from NIST MS Search View menu.

Selecting derivative type(s) affects Lib. Search hit list: structures of derivative precursors are displayed instead of derivatives. Only mainlib and replib hits are affected. Deselecting displayed derivative type(s) immediately changes structure display back to derivative.

The removed attachments and their numbers are listed in the text above the precursor structure. Note that all attachments are displayed removed if at least one of them fits the selected derivative type. For example, instead of TMS derivative of L-Alanine, its precursor would be displayed:



Precursor structure drawing may only be viewed, printed, or copied into the Windows clipboard as pictures. An attempt to save the precursor structure into a molfile, copy it into clipboard, or copy it to Spec.List would save or copy the structure of the derivative.

The same derivative definitions are used to display derivatives, which have smallest precursor in common as replicate spectra in [Replicates Display Options](#).

## Structures in the database

It is possible to [add structures to user databases](#), to [search the database for structures](#) that are similar to the structure of a given compound and to take the structures from the database and put them into other documents or into structure drawing programs.

## Reading structures from files

The File/Open dialog has the option of opening mol, sdf, and KCF (Glycans) file formats. These files contain only the structures. The program will show a preview of the structure for each selected entry. The structures are imported into the Spec List, but only as structures. From the Spec List they may be placed in the clipboard.

If the program is in automatic search mode, the import of the structures will cause the program to perform a structure similarity search on each imported structure.

## Adding structures to user library spectra

A structure may be added to the user library either from a text file (mol or sdf), the clipboard, or a user library.

- To attach a structure to a spectrum located in Spec. List, select this spectrum in Librarian window and click button  (Edit spectrum) on the Librarian Toolbar. [Spectrum Information dialog](#) will open. Select **Attach Struct** button to add a structure from a file or library, or select **Clipboard Struct** button to attach a structure from the clipboard. Select **Add to Library**, or **Replace** (in library), or **Add to List** button to save the spectrum with the newly attached structure. Note that structures from Mainlib are attached according to CAS r.n.
- To attach a structure from the clipboard to the search spectrum, right-click Plot of Search Spectrum in Lib. Search window and select **Attach Clipboard Structure**. A temporary spectrum with the attached structure will be created in the Plot of Search Spectrum window. This spectrum may be saved by right-clicking Plot of Search Spectrum window and selecting Spec List in **Send To** menu. The saved spectrum may be added to a user library.

Structures can be put into the clipboard from the program: right-click a spectrum that has a chemical structure and select from the floating menu Copy Structure to Clipboard.

Structures placed on the clipboard from drawing programs cannot be used unless they are in molfile format.

To use a structure file from a drawing program, the structure must be saved in mol, sdf, or KCF (Glycan) format.

## **Putting a structure on the clipboard**

Using the right mouse button, it is generally possible to copy a structure displayed by the NIST MS Search in the clipboard. From the clipboard the structure can be placed in the Spec List so that it can be searched against the library or it can be added to a user spectrum in the Librarian. The resulting structure cannot in general be pasted into other drawing programs, it can however be pasted into any chemical drawing program which uses the "mol file" format for the clipboard. Pasting into most other applications will result in the text representation of the structure ("mol file" format) being added.

## Sending a structure to a drawing program

Using the right mouse button on a spectrum that has a structure associated with it will bring up a menu choice:

### Send To

Depending upon the location in the program - the choices shown will be Spec List, Compare List, MS Interpreter, Default Structure Editor and any other structure editors that you have defined. The default structure editor is that program that is the default viewer for files with the **.mol** extension. You may add your own editors to this list. To add your editors you need to create a simple plain text file named **autoimp.str** and place it in the same directory as the NIST MS Search program. For each drawing program that you want to use place the following line in this file:

```
"Name" "Fully qualified file name" "%1"
```

For example

```
"DrawIt" "D:\Program Files\DrawAllStuff\DrawIt.exe" "%1"
```

The quotation marks are necessary. The Name is the name that will appear on the Send to menu. The Fully qualified file name is the name of the drawing program with the full path to the drawing program.

The "%1" will allow the program to send the mol file to the drawing program.

**NOTE: Only chemical drawing programs that support command line reading of mol files can be used in this way.**

## Adding a clipboard structure to the Spec List

If a structure is in the clipboard it can be added to the Spec List so that it can be searched. To add a clipboard structure to the Spec List, select the right mouse button and select **Insert Clipboard Structure**.

Once you have selected the Insert Clipboard structure context menu item, a dialog will appear that will allow you to name and add a comment to the structure before it is added to the data.

**NOTE:** The NIST MS Search program can only accept structures put into the clipboard by the NIST MS Search program or chemical drawing programs that use the "mol file" format for the clipboard. If the program cannot accept structures from a chemical drawing program, it is usually possible to use [read mol and sdf](#) files to bring in structures drawn by other programs.

## **Updating User Structure Search Libraries**

Structure similarity search data are not automatically updated when deleting or adding structures and spectra to the user libraries. A separate utility is provided to create the index files needed for structure similarity searches to update the version 2.0 files.

To create or update user structure search databases run "Rebuild Structure Search Databases" from the Tools menu.

# Common Reasonable Losses From Molecular Ions

## ODD LOSSES

M - 1	hydrogen atom	M - H
M - 15	methyl radical	M - CH <sub>3</sub>
M - 19	fluorine atom	M - F
M - 29	ethyl radical	M - CH <sub>2</sub> CH <sub>3</sub>
M - 31	methoxy radical	M - OCH <sub>3</sub>
M - 35	chlorine atom	M - Cl
M - 43	propyl radical	M - C <sub>3</sub> H <sub>7</sub>
M - 45	ethoxy radical	M - OCH <sub>2</sub> CH <sub>3</sub>
M - 57	butyl radical	M - C <sub>4</sub> H <sub>9</sub>
M - 91	benzyl radical	M - CH <sub>2</sub> C <sub>6</sub> H <sub>5</sub>
M - 79	bromine atom	M - Br
M - 129	iodine atom	M - I

## EVEN LOSSES

M - 2	hydrogen molecule	M - H <sub>2</sub>
M - 18	water	M - H <sub>2</sub> O
M - 20	hydrogen fluoride	M - HF
M - 28	CO or ethylene	M - CO or M - C <sub>2</sub> H <sub>4</sub>
M - 30	formaldehyde	M - H <sub>2</sub> CO
M - 32	ethanol	M - CH <sub>3</sub> CO <sub>2</sub> H
M - 36	hydrogen chloride	M - HCl
M - 44	CO <sub>2</sub>	M - CO <sub>2</sub>
M - 46	nitrogen dioxide	M - NO <sub>2</sub>
M - 60	acetic acid	M - CH <sub>3</sub> CO <sub>2</sub> H
M - 90	silanol: HO-Si(CH <sub>3</sub> ) <sub>3</sub>	M - HO-Si-(CH <sub>3</sub> ) <sub>3</sub>

### **Unreasonable suggested loss**

M - 5 to M - 14

M - 21 to M - 25

These losses do not correspond to groups of atoms likely to be lost from most common organic molecules (C H N O S P F Cl Br I); they probably result from noise spikes or ions from an impurity or indicate an incorrect molecular ion or anchor assignment.

The smallest C-containing group that can be lost is CH<sub>3</sub> (15) which is also, by far, the most common neutral loss peak in organic mass spectrometry. Losses >4 or <15 are not generally possible.

## Neutral Fragments Expelled by Simple Bond Breaking

m/z	Fragment	m/z	Fragment
1	.H	47	.CH <sub>2</sub> SH
15	.CH <sub>3</sub>	49	.CH <sub>2</sub> Cl
16	.NH <sub>2</sub>	51	.CHF <sub>2</sub>
17	.OH	54	.CH <sub>2</sub> CH <sub>2</sub> CN
19	.F	55	.C <sub>4</sub> H <sub>7</sub> .C <sub>3</sub> H <sub>3</sub> O
26	.CN	57	.C <sub>4</sub> H <sub>9</sub> .C <sub>3</sub> H <sub>5</sub> O
27	.C <sub>2</sub> H <sub>3</sub>	58	.C <sub>3</sub> H <sub>8</sub> N
29	.C <sub>2</sub> H <sub>5</sub> .CHO	59	.C <sub>3</sub> H <sub>7</sub> O .COOCH <sub>3</sub> .CH <sub>2</sub> COOH
30	.CH <sub>3</sub> NH.CH <sub>2</sub> NH <sub>2</sub>	61	.C <sub>2</sub> H <sub>5</sub> S
31	.CH <sub>3</sub> O.CH <sub>2</sub> OH	65	.C <sub>5</sub> H <sub>5</sub>
33	.SH .CH <sub>2</sub> F .(H <sub>2</sub> O + .CH <sub>3</sub> )	69	.CF <sub>3</sub>
35	.Cl	71	.C <sub>4</sub> H <sub>7</sub> O
40	.CH <sub>2</sub> CN	77	.C <sub>6</sub> H <sub>5</sub>
41	.C <sub>3</sub> H <sub>5</sub>	79	.Br .C <sub>6</sub> H <sub>7</sub>
43	.C <sub>3</sub> H <sub>7</sub> CH <sub>3</sub> CO	81	.C <sub>6</sub> H <sub>9</sub> .C <sub>5</sub> H <sub>5</sub> O
44	.C <sub>2</sub> H <sub>5</sub> NH	91	.C <sub>7</sub> H <sub>7</sub>
45	.C <sub>2</sub> H <sub>5</sub> O .COOH	105	.C <sub>7</sub> H <sub>5</sub> O .C <sub>8</sub> H <sub>9</sub>
46	.NO <sub>2</sub>	121	.C <sub>7</sub> H <sub>5</sub> O <sub>2</sub>
127	.I		

## Neutral fragments expelled by multicentered fragmentation

m/z	Fragment	m/z	Fragment
2	H <sub>2</sub>	45	C <sub>2</sub> H <sub>7</sub> N
17	NH <sub>3</sub>	46	C <sub>2</sub> H <sub>6</sub> O or H <sub>2</sub> O+C <sub>2</sub> H <sub>4</sub>
18	H <sub>2</sub> O	48	CH <sub>4</sub> S
20	HF	54	C <sub>4</sub> H <sub>6</sub>
27	HCN	56	C <sub>4</sub> H <sub>8</sub> C <sub>3</sub> H <sub>4</sub> O
28	CO C <sub>2</sub> H <sub>4</sub>	58	C <sub>3</sub> H <sub>6</sub> O
30	CH <sub>2</sub> O	59	C <sub>3</sub> H <sub>9</sub> N
31	CH <sub>5</sub> N	60	C <sub>3</sub> H <sub>8</sub> O C <sub>2</sub> H <sub>4</sub> O <sub>2</sub>
32	CH <sub>4</sub> O	62	C <sub>2</sub> H <sub>6</sub> S
34	H <sub>2</sub> S	74	C <sub>3</sub> H <sub>6</sub> O <sub>2</sub>
36	HCl	76	C <sub>6</sub> H <sub>4</sub>
42	C <sub>3</sub> H <sub>6</sub> C <sub>2</sub> H <sub>2</sub> O	78	C <sub>6</sub> H <sub>6</sub>
44	CO <sub>2</sub>	80	HBr

## Natural Isotopic Composition and Exact Masses of Common Elements

Element Symbol		Nominal Mass	Exact Mass	Abundance	X+1 Factor	X+2 Factor
Hydrogen	H	1	1.0078	99.985		
		2	2.0141	0.015		
Boron	B	10	10.0129	19.9	4.03nB	
		11	11.0093	80.1		
<a href="#">Carbon</a>	C	12	12.0000	98.91		
		13	13.0034	1.1	1.1nC	0.0060(nC*nC)
Nitrogen	N	14	14.0031	99.64		
		15	15.0001	0.36	0.37nN	
Oxygen	O	16	15.9949	99.76		
		17	16.9991	0.04	0.04nO	
		18	17.9992	0.20		0.20nO
Fluorine	F	19	18.9984	100		
<a href="#">Silicon</a>	Si	28	27.9769	92.2		
		29	28.9765	4.7	5.1nSi	
		30	29.9738	3.1		3.4nSi
Phosphorus	P	31	30.9738	100		
<a href="#">Sulfur</a>	S	32	31.9721	95.0		
		33	32.9715	0.75	0.8nS	
		34	33.9679	4.25		4.4nS

<b>Chlorine</b>	Cl	35	34.9689	75.76		
		37	36.9659	24.24		32.5nCl
<b>Bromine</b>	Br	79	78.9183	50.7		
		81	80.9163	49.3		98.0nBr
Iodine	I	127	126.9045	100		

Assume  $X = 100\%$ ;  $X$  represents the relative intensity of the first peak in a cluster of peaks corresponding to isotopic variants of a given ion. Note that for all isotopic clusters except Boron the lowest mass isotope is the most abundant.

The factor is multiplied by the number ( $n$ ) of atoms of the element in the ion to determine the  $X+1$  or  $X+2$  relative intensity contribution percentage of a given isotope. For example, the contribution at  $m/z=X+1$  due to  $^{15}\text{N}$  for an ion containing three nitrogen atoms would be  $0.37 \times 3 = 1.11$  relative to 100 at  $m/z=X$ .

## Exact Masses and Natural Isotopic Composition of Elements

Element	Symbol	Nominal Mass	Exact Mass	Isotopic Composition
Hydrogen	H	1	1.00782503207(10)	0.999885(70)
	D	2	2.0141017778(4) x	0.000115(70)
Helium	He	4	4.00260325415(6)	1
Lithium	Li	6	6.015122795(16)	0.0759(4)
		7	7.01600455(8)	0.9241(4)
Beryllium	Be	9	9.0121822(4)	1
Boron	B	10	10.0129370(4)	0.199(7)
		11	11.0093054(4)	0.801(7)
Carbon	C	12	12. <b>(REF)</b>	0.9893(8)
		13	13.0033548378(10)	0.0107(8)
Nitrogen	N	14	14.0030740048(6)	0.99636(20)
		15	15.0001088982(7) x	0.00364(20)
Oxygen	O	16	15.99491461956(16)	0.99757(16)
		17	16.99913170(12)	0.00038(1)
		18	17.9991610(7)	0.00205(14)
Fluorine	F	19	18.99840322(7)	1
Neon	Ne	20	19.9924401754(19)	0.9048(3)
		21	20.99384668(4)	0.0027(1)
		22	21.991385114(19)	0.0925(3)
Sodium	Na	23	22.9897692809(29)	1
Magnesium	Mg	24	23.985041700(14)	0.7899(4)
		25	24.98583692(3)	0.1000(1)
		26	25.982592929(30)	0.1101(3)
Aluminum	Al	27	26.98153863(12)	1
Silicon	Si	28	27.9769265325(19)	0.92223(19)

		29	28.976494700(22)	0.04685(8)
		30	29.97377017(3)	0.03092(11)
Phosphorus	P	31	30.97376163(20)	1
Sulfur	S	32	31.97207100(15)	0.9499(26)
		33	32.97145876(15)	0.0075(2)
		34	33.96786690(12)	0.0425(24)
		36	35.96708076(20)	0.0001(1)
Chlorine	Cl	35	34.96885268(4)	0.7576(10)
		37	36.96590259(5)	0.2424(10)
Argon	Ar	36	35.967545106(29)	0.003365(30)
		38	37.9627324(4)	0.000632(5)
		40	39.9623831225(29)	0.996003(30)
Potassium	K	39	38.96370668(20)	0.932581(44)
		40+	39.96399848(21)	0.000117(1)
		41	40.96182576(21)	0.067302(44)
Calcium	Ca	40	39.96259098(22)	0.96941(156)
		42	41.95861801(27)	0.00647(23)
		43	42.9587666(3)	0.00135(10)
		44	43.9554818(4)	0.02086(110)
		46	45.9536926(24)	0.00004(3)
		48	47.952534(4)	0.00187(21)
Scandium	Sc	45	44.9559119(9)	1
Titanium	Ti	46	45.9526316(9)	0.0825(3)
		47	46.9517631(9)	0.0744(2)
		48	47.9479463(9)	0.7372(3)
		49	48.9478700(9)	0.0541(2)
		50	49.9447912(9)	0.0518(2)
Vanadium	V	50+	49.9471585(11)	0.00250(4)
		51	50.9439595(11)	0.99750(4)
Chromium	Cr	50	49.9460442(11)	0.04345(13)

		52	51.9405075(8)	0.83789(18)
		53	52.9406494(8)	0.09501(17)
		54	53.9388804(8)	0.02365(7)
Manganese	Mn	55	54.9380451(7)	1
Iron	Fe	54	53.9396105(7)	0.05845(35)
		56	55.9349375(7)	0.91754(36)
		57	56.9353940(7)	0.02119(10)
		58	57.9332756(8)	0.00282(4)
Cobalt	Co	59	58.9331950(7)	1
Nickel	Ni	58	57.9353429(7)	0.680769(89)
		60	59.9307864(7)	0.262231(77)
		61	60.9310560(7)	0.011399(6)
		62	61.9283451(6)	0.036345(17)
		64	63.9279660(7)	0.009256(9)
Copper	Cu	63	62.9295975(6)	0.6915(15)
		65	64.9277895(7)	0.3085(15)
Zinc	Zn	64	63.9291422(7)	0.48268(321)
		66	65.9260334(10)	0.27975(77)
		67	66.9271273(10)	0.04102(21)
		68	67.9248442(10)	0.19024(123)
		70	69.9253193(21)	0.00631(9)
Gallium	Ga	69	68.9255736(13)	0.60108(9)
		71	70.9247013(11)	0.39892(9)
Germanium	Ge	70	69.9242474(11)	0.2038(18)
		72	71.9220758(18)	0.2731(26)
		73	72.9234589(18)	0.0776(8)
		74	73.9211778(18)	0.3672(15)
		76	75.9214026(18)	0.0783(7)
Arsenic	As	75	74.9215965(20)	1
Selenium	Se	74	73.9224764(18)	0.0089(4)

		76	75.9192136(18)	0.0937(29)
		77	76.9199140(18)	0.0763(16)
		78	77.9173091(18)	0.2377(28)
		80	79.9165213(21)	0.4961(41)
		82	81.9166994(22)	0.0873(22)
Bromine	Br	79	78.9183371(22)	0.5069(7)
		81	80.9162906(21)	0.4931(7)
Krypton	Kr	78	77.9203648(12)	0.00355(3)
		80	79.9163790(16)	0.02286(10)
		82	81.9134836(19)	0.11593(31)
		83	82.914136(3)	0.11500(19)
		84	83.911507(3)	0.56987(15)
		86	85.91061073(11)	0.17279(41)
Rubidium	Rb	85	84.911789738(12)	0.7217(2)
		87 <sub>xx</sub>	86.909180527(13)	0.2783(2)
Strontium	Sr	84	83.913425(3)	0.0056(1)
		86	85.9092602(12)	0.0986(1)
		87	86.9088771(12)	0.0700(1)
		88	87.9056121(12)	0.8258(1)
Yttrium	Y	89	88.9058483(27)	1
Zirconium	Zr	90	89.9047044(25)	0.5145(40)
		91	90.9056458(25)	0.1122(5)
		92	91.9050408(25)	0.1715(8)
		94	93.9063152(26)	0.1738(28)
		96	95.9082734(30)	0.0280(9)
Niobium	Nb	93	92.9063781(26)	1
Molybdenum	Mo	92	91.906811(4)	0.1477(31)
		94	93.9050883(21)	0.0923(10)
		95	94.9058421(21)	0.1590(9)
		96	95.9046795(21)	0.1668(1)

		97	96.9060215(21)	0.0956(5)
		98	97.9054082(21)	0.2419(26)
		100	99.907477(6)	0.0967(20)
Ruthenium	Ru	96	95.907598(8)	0.0554(14)
		98	97.905287(7)	0.0187(3)
		99	98.9059393(22)	0.1276(14)
		100	99.9042195(22)	0.1260(7)
		101	100.9055821(22)	0.1706(2)
		102	101.9043493(22)	0.3155(14)
		104	103.905433(3)	0.1862(27)
Rhodium	Rh	103	102.905504(3)	1
Palladium	Pd	102	101.905609(3)	0.0102(1)
		104	103.904036(4)	0.1114(8)
		105	104.905085(4)	0.2233(8)
		106	105.903486(4)	0.2733(3)
		108	107.903892(4)	0.2646(9)
		110	109.905153(12)	0.1172(9)
Silver	Ag	107	106.905097(5)	0.51839(8)
		109	108.904752(3)	0.48161(8)
Cadmium	Cd	106	105.906459(6)	0.0125(6)
		108	107.904184(6)	0.0089(3)
		110	109.9030021(29)	0.1249(18)
		111	110.9041781(29)	0.1280(12)
		112	111.9027578(29)	0.2413(21)
		113	112.9044017(29)	0.1222(12)
		114	113.9033585(29)	0.2873(42)
		116	115.904756(3)	0.0749(18)
Indium	In	113	112.904058(3)	0.0429(5)
		115+	114.903878(5)	0.9571(5)
Tin	Sn	112	111.904818(5)	0.0097(1)

		114	113.902779(3)	0.0066(1)
		115	114.903342(3)	0.0034(1)
		116	115.901741(3)	0.1454(9)
		117	116.902952(3)	0.0768(7)
		118	117.901603(3)	0.2422(9)
		119	118.903308(3)	0.0859(4)
		120	119.9021947(27)	0.3258(9)
		122	121.9034390(29)	0.0463(3)
		124+	123.9052739(15)	0.0579(5)
Antimony	Sb	121	120.9038157(24)	0.5721(5)
		123	122.9042140(22)	0.4279(5)
Tellurium	Te	120	119.904020(10)	0.0009(1)
		122	121.9030439(16)	0.0255(12)
		123	122.9042700(16)	0.0089(3)
		124	123.9028179(16)	0.0474(14)
		125	124.9044307(16)	0.0707(15)
		126	125.9033117(16)	0.1884(25)
		128	127.9044631(19)	0.3174(8)
		130	129.9062244(21)	0.3408(62)
Iodine	I	127	126.904473(4)	1
Xenon	Xe	124	123.9058930(20)	0.000952(3)
		126	125.904274(7)	0.000890(2)
		128	127.9035313(15)	0.019102(8)
		129	128.9047794(8)	0.264006(82)
		130	129.9035080(8)	0.040710(13)
		131	130.9050824(10)	0.212324(30)
		132	131.9041535(10)	0.269086(33)
		134	133.9053945(9)	0.104357(21)
		136	135.907219(8)	0.088573(44)
Cesium	Cs	133	132.905451933(24)	1

Barium	Ba	130	129.9063208(30)	0.00106(1)
		132	131.9050613(11)	0.00101(1)
		134	133.9045084(4)	0.02417(18)
		135	134.9056886(4)	0.06592(12)
		136	135.9045759(4)	0.07854(24)
		137	136.9058274(5)	0.11232(24)
		138	137.9052472(5)	0.71698(42)
Lanthanum	La	138+	137.907112(4)	0.00090(1)
		139	138.9063533(26)	0.99910(1)
Cerium	Ce	136	135.907172(14)	0.00185(2)
		138	137.905991(11)	0.00251(2)
		140	139.9054387(26)	0.88450(51)
		142	141.909244(3)	0.11114(51)
Praseodymium	Pr	141	140.9076528(26)	1
Neodymium	Nd	142	141.9077233(25)	0.272(5)
		143	142.9098143(25)	0.122(2)
		144+	143.9100873(25)	0.238(3)
		145	144.9125736(25)	0.083(1)
		146	145.9131169(25)	0.172(3)
		148	147.916893(3)	0.057(1)
		150+	149.920891(3)	0.056(2)
Samarium	Sm	144	143.911999(3)	0.0307(7)
		147+	146.9148979(26)	0.1499(18)
		148	147.9148227(26)	0.1124(10)
		149	148.9171847(26)	0.1382(7)
		150	149.9172755(26)	0.0738(1)
		152	151.9197324(27)	0.2675(16)
		154	153.9222093(27)	0.2275(29)
Europium	Eu	151	150.9198502(26)	0.4781(6)
		153	152.9212303(26)	0.5219(6)

Gadolinium	Gd	152+	151.9197910(27)	0.0020(1)
		154	153.9208656(27)	0.0218(3)
		155	154.9226220(27)	0.1480(12)
		156	155.9221227(27)	0.2047(9)
		157	156.9239601(27)	0.1565(2)
		158	157.9241039(27)	0.2484(7)
		160	159.9270541(27)	0.2186(19)
Terbium	Tb	159	158.9253468(27)	1
Dysprosium	Dy	156	155.924283(7)	0.00056(3)
		158	157.924409(4)	0.00095(3)
		160	159.9251975(27)	0.02329(18)
		161	160.9269334(27)	0.18889(42)
		162	161.9267984(27)	0.25475(36)
		163	162.9287312(27)	0.24896(42)
		164	163.9291748(27)	0.28260(54)
Holmium	Ho	165	164.9303221(27)	1
Erbium	Er	162	161.928778(4)	0.00139(5)
		164	163.929200(3)	0.01601(3)
		166	165.9302931(27)	0.33503(36)
		167	166.9320482(27)	0.22869(9)
		168	167.9323702(27)	0.26978(18)
		170	169.9354643(30)	0.14910(36)
Thulium	Tm	169	168.9342133(27)	1
Ytterbium	Yb	168	167.933897(5)	0.0013(1)
		170	169.9347618(26)	0.0304(15)
		171	170.9363258(26)	0.1428(57)
		172	171.9363815(26)	0.2183(67)
		173	172.9382108(26)	0.1613(27)
		174	173.9388621(26)	0.3183(92)
		176	175.9425717(28)	0.1276(41)

Lutetium	Lu	175	174.9407718(23)	0.9741(2)
		176+	175.9426863(23)	0.0259(2)
Hafnium	Hf	174	173.940046(3)	0.0016(1)
		176	175.9414086(24)	0.0526(7)
		177	176.9432207(23)	0.1860(9)
		178	177.9436988(23)	0.2728(7)
		179	178.9458161(23)	0.1362(2)
		180	179.9465500(23)	0.3508(16)
Tantalum	Ta	180	179.9474648(24)	0.00012(2)
		181	180.9479958(19)	0.99988(2)
Tungsten	W	180	179.946704(4)	0.0012(1)
		182	181.9482042(9)	0.2650(16)
		183	182.9502230(9)	0.1431(4)
		184	183.9509312(9)	0.3064(2)
		186	185.9543641(19)	0.2843(19)
Rhenium	Re	185	184.9529550(13)	0.3740(2)
		187	186.9557531(15)	0.6260(2)
Osmium	Os	184	183.9524891(14)	0.0002(1)
		186	185.9538382(15)	0.0159(3)
		187+	186.9557505(15)	0.0196(2)
		188	187.9558382(15)	0.1324(8)
		189	188.9581475(16)	0.1615(5)
		190	189.9584470(16)	0.2626(2)
		192	191.9614807(27)	0.4078(19)
Iridium	Ir	191	190.9605940(18)	0.373(2)
		193	192.9629264(18)	0.627(2)
Platinum	Pt	190	189.959932(6)	0.00014(1)
		192	191.9610380(27)	0.00782(7)
		194	193.9626803(9)	0.32967(99)
		195	194.9647911(9)	0.33832(10)

		196	195.9649515(9)	0.25242(41)
		198	197.967893(3)	0.07163(55)
Gold	Au	197	196.9665687(6)	1
Mercury	Hg	196	195.965833(3)	0.0015(1)
		198	197.9667690(4)	0.0997(20)
		199	198.9682799(4)	0.1687(22)
		200	199.9683260(4)	0.2310(19)
		201	200.9703023(6)	0.1318(9)
		202	201.9706430(6)	0.2986(26)
		204	203.9734939(4)	0.0687(15)
Thallium	Tl	203	202.9723442(14)	0.2952(1)
		205	204.9744275(14)	0.7048(1)
Lead	Pb	204	203.9730436(13)	0.014(1)
		206	205.9744653(13)	0.241(1)
		207	206.9758969(13)	0.221(1)
		208	207.9766521(13)	0.524(1)
Bismuth	Bi	209	208.9803987(16)	1

+ Long lived radioactive isotopes

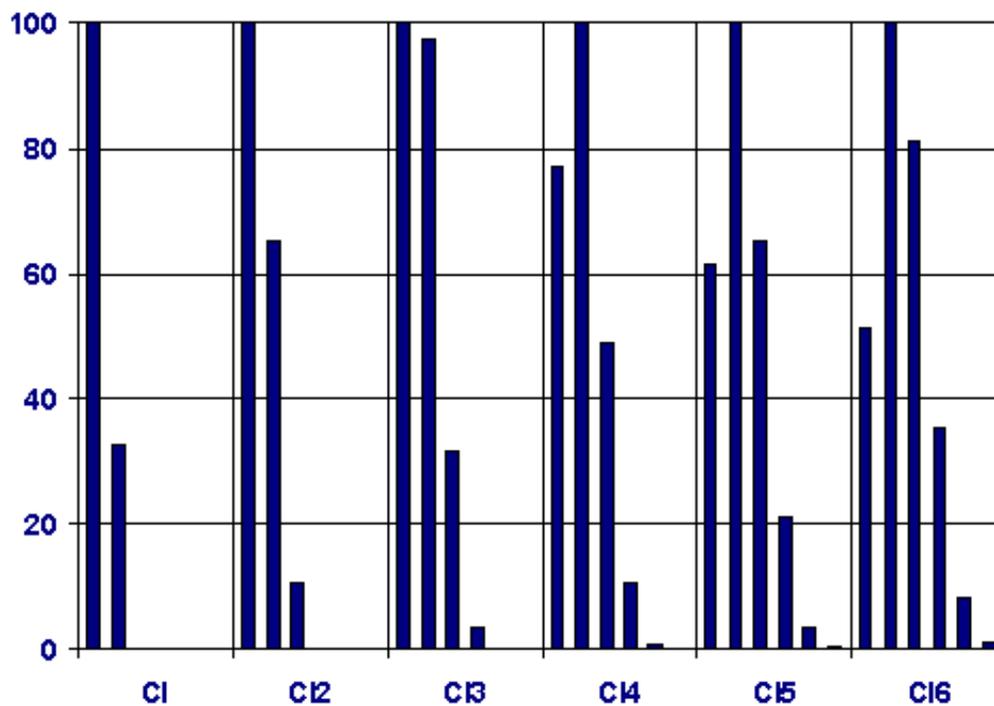
x Frequently used in labeled compounds

## Isotopic Contributions of Carbon, % of monoisotopic

C#	(X+1)+	(X+2)+	C#	(X+1)+	(X+2)+
C1	1.1	0.00	C30	32.5	5.1
C2	2.2	0.01	C40	43.3	9.1
C3	3.2	0.04	C50	54	14
C4	4.3	0.07	C60	65	21
C5	5.4	0.12	C70	76	28
C6	6.5	0.18	C80	87	37
C7	7.6	0.25	C90	97	47
C8	8.7	0.33			
C9	9.7	0.42			
C10	10.8	0.53			
C11	11.9	0.64			
C12	13.0	0.77			
C13	14.1	0.91			
C14	15.1	1.1			
C15	16.2	1.2			
C16	17.3	1.4			
C17	18.4	1.6			
C18	19.5	1.8			
C19	20.6	2.0			
C20	21.6	2.2			

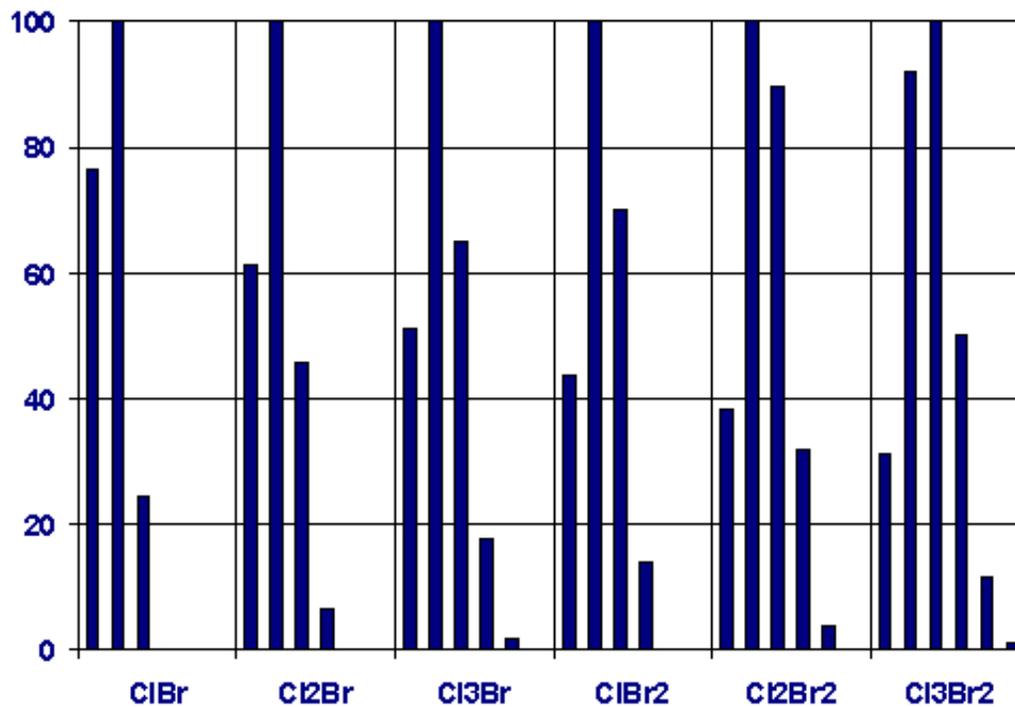
# Cl Isotopic peaks Cl .. Cl<sub>6</sub>

Peaks are at intervals of 2 m/z (X, X+2, X+4,...)



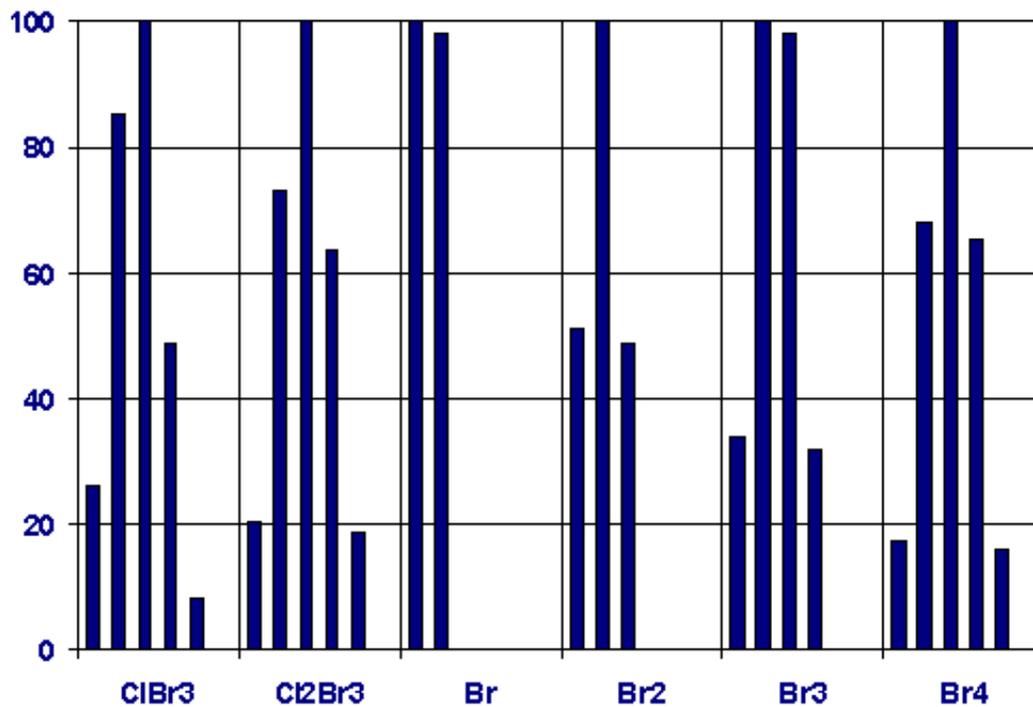
## BrCl combinations ClBr, Cl<sub>2</sub>Br, Cl<sub>3</sub>Br, ClBr<sub>2</sub>, Cl<sub>2</sub>Br<sub>2</sub>, Cl<sub>3</sub>Br<sub>2</sub>

Peaks are at intervals of 2 m/z (X, X+2, X+4,...)



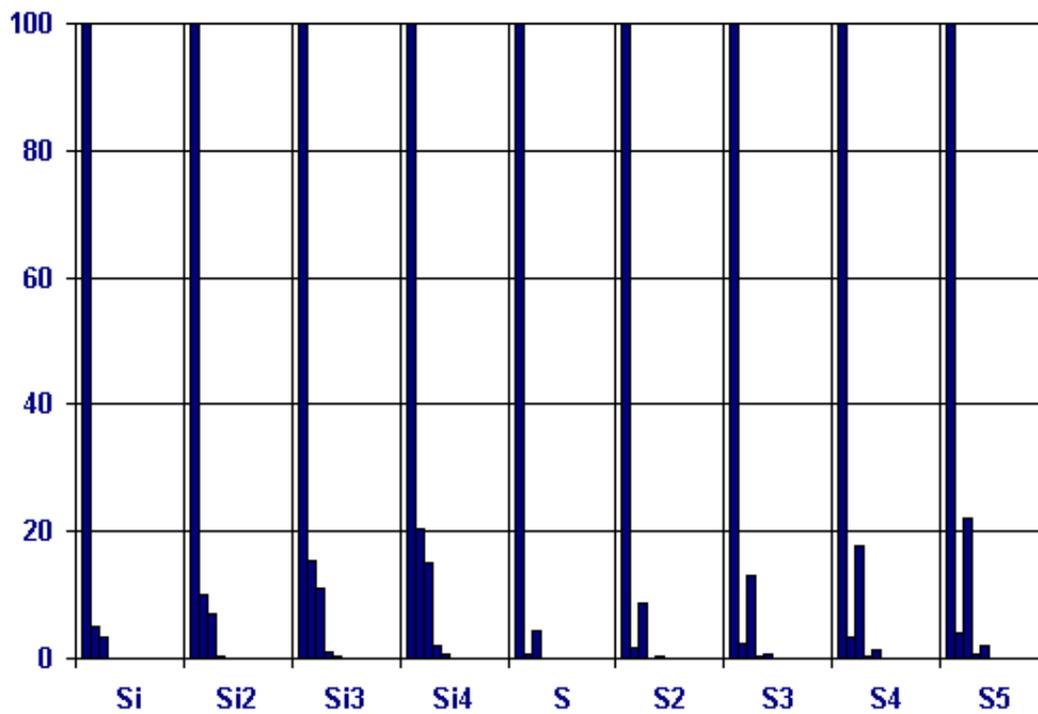
## BrCl combinations - ClBr<sub>3</sub>, Cl<sub>2</sub>Br<sub>3</sub>, Br, Br<sub>2</sub>, Br<sub>3</sub>, Br<sub>4</sub>

Peaks are at intervals of 2 m/z units (X, X+2, X+4,...)



# Silicon and Sulfur Isotopic Peaks

Peaks are at intervals of 1 m/z unit (X,X+1,X+2...)



## Relative Intensity of Peaks Due to Cl/Br Isotopes

Illustration: an ion with 2 Chlorine atoms will have a peaks at 2 and 4 m/z above the nominal mass with intensities of 65% and 10.6% respectively of that of the nominal mass peak.

Cl-Br	X	X+2	X+4	X+6	X+8	X+10
Cl	100	32.5				
Cl <sub>2</sub>	100	65.0	10.6			
Cl <sub>3</sub>	100	97.5	31.7	3.4		
Cl <sub>4</sub>	76.9	100	48.7	10.5	0.9	
Cl <sub>5</sub>	61.5	100	65.0	21.1	3.4	0.2
Cl <sub>6</sub>	51.2	100	81.2	35.2	8.5	1.1
ClBr	76.6	100	24.4			
Cl <sub>2</sub> Br	61.4	100	45.6	6.6		
Cl <sub>3</sub> Br	51.2	100	65.0	17.6	1.7	
ClBr <sub>2</sub>	43.8	100	69.9	13.7		
Cl <sub>2</sub> Br <sub>2</sub>	38.3	100	89.7	31.9	3.9	
Cl <sub>3</sub> Br <sub>2</sub>	31.3	92.0	100	49.9	11.6	1.0
ClBr <sub>3</sub>	26.1	85.1	100	48.9	8.0	
Cl <sub>2</sub> Br <sub>3</sub>	20.4	73.3	100	63.8	18.7	2.0
Br	100	98.0				
Br <sub>2</sub>	51.0	100	49.0			
Br <sub>3</sub>	34.0	100	98.0	32.0		
Br <sub>4</sub>	17.4	68.0	100	65.3	16.0	