

## MassFinder 3

Dr. Detlev H. Hochmuth, Author  
Hochmuth Scientific Consulting  
Hamburg, Germany  
<http://www.massfinder.com>  
Euro 1850, which includes latest version of MassFinder 3, a terpene spectral library with 1982 spectra (includes spectra, retention index, molecular formula, nominal mass based on formula, and stereochemical structure), and free upgrades to your version of MassFinder 3.

Reviewed by O. David Sparkman  
University of the Pacific  
Department of Chemistry  
Mass Spectrometry Facility  
Stockton, CA 95211, USA  
E-mail: [osparkma@pacific.edu](mailto:osparkma@pacific.edu)

During the preparation of the review of Robert P. Adams' fourth edition of *Identification of Essential Oil Components by Gas Chromatography/Mass Spectrometry* (*J. Am. Soc. Mass Spectrom.* **2007**, *18*, 803–806), I saw a screen shot of MassFinder 3.10d. This display was rather intriguing as can be seen in Figure 1 (MassFinder 3.48g). The vertical orientation of the spectrum #/time axis of the chromatogram was reminiscent of the strip-chart recorder outputs used with gas chromatographs and for the output of the total ion current from mass spectrometer ion sources before the days of computerized data systems. The six-position spectral display was also intriguing. From the brief description of the program in the Adams book, it appeared that a further investigation was warranted. The MassFinder software is provided with an approximate 2000 spectra library of compounds that are commonly found in essential oils—that is, flavor and fragrant compositions (natural and synthetic). Each library spectrum has a structure associated with it; and, where appropriate, the structure displays stereochemical features. This library was developed by Wilfried A. König, University of Hamburg, Germany; Daniel Joulain, Robertet S.A., Grasse, France; and Detlev H. Hochmuth, Hochmuth Scientific Consulting, Hamburg, Germany. It is a little surprising, but very refreshing, that Dr. Adams makes reference to the MassFinder 3 software because the Hochmuth Scientific Consulting essential oils library at 1982 compounds and with electronic structures is a direct competitor to the Adams library of 2205 spectra with no electronic structures. It should also be noted that the MassFinder 3 format is among the many formats in which the Adams library is now distributed.

Although the essential oils library is significant to the use of MassFinder 3 in the flavors and fragrance field, the program has far-reaching utility in any field where gas chromatography/mass spectrometry (GC/MS) is used. The program will read most common data-file formats or has utilities to convert them to a JACAMP format. There is a special File Open dialog box for Agilent ChemStation data files. This display provides file selection from the .D file and displays the file acquisition date and information in the sample comments (info) fields of the data file.

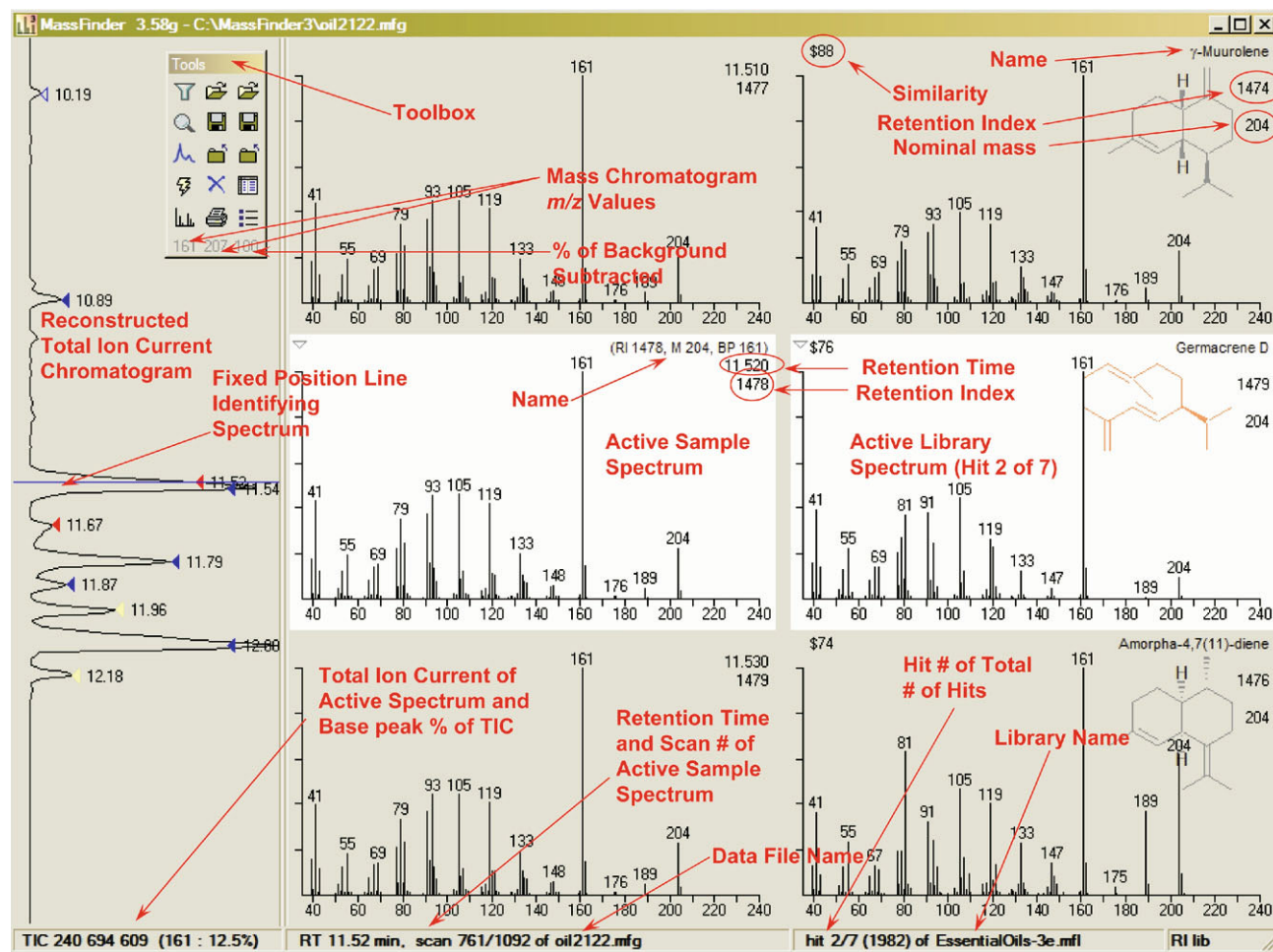
Dr. Hochmuth first wrote MassFinder in 1999 as an academic exercise. The first commercial version (MassFinder 2.0) was introduced in 2001. In 2003, Version 3 was written from a “white page beginning” according to Dr. Hochmuth. In addition to the program's unique display, among many of its one-of-a-kind features is its proprietary search algorithm. When questioned about his understanding of the PBM, IN-COS, and other important search algorithms, Dr. Hochmuth said he knew little about those algorithms and said that his approach was different; moreover, in some cases, it performed better than other algorithms, although there were also some cases where his algorithm did not perform as well. The MassFinder 3 library search uses what it calls a **two-dimensional search algorithm**, which combines a retention index with  $m/z$  value-intensity pairs in matching library and sample spectra. This is why it is important that the retention index values are included with each library spectrum. When a data file is loaded into the program, a retention index is assigned based on the information about the retention time (or scan numbers) for a series of normal alkanes provided with the program. The information in this table can be updated using data files produced on the user's system. A detailed explanation of this table and how it is used and updated with user-generated data is provided as part of the program's documentation.

In the case where two or more compounds in the library have the same or very similar spectra, such as is the case with the regioisomers of xylene, the only differentiating factor is the retention index. The automated assignment of the retention index to the sample spectrum and the use of the retention index associated with the library spectrum is another of the many unique analyte identification features of MassFinder 3.

## Navigation of the Chromatogram

The chromatogram display and the spectral displays are separated by a vertical slider bar. The size and orientation of the six spectra displays are fixed, but their width is a function of the slider bar. Navigation of the chromatogram is a unique feature of the MassFinder 3 program. When the data files are loaded, the reconstructed total ion current (RTIC) chromatogram

Published online April 12, 2007



**Figure 1.** MassFinder 3's desktop. The chromatogram has four different colored  $\nabla$  symbols: the  $\blacktriangledown$  symbol is a component that the user has associated with a library spectrum; the  $\blacktriangledown$  symbol indicates a spectrum that has been evaluated but that is not associated with a library spectrum; the  $\blacktriangledown$  symbol indicates a spectrum associated with a library spectrum by the Auto Report function; and the  $\nabla$  symbol indicates the spectrum at the apex of a chromatographic peak that could not be associated with a library spectrum.

(RTICC) is displayed from spectrum 1 to the last spectrum in the file. The standard click-and-drag command common to all Windows GC/MS data systems is available in MassFinder 3 to display a portion of the chromatogram. However, by holding down the left mouse button with the mouse pointer on the chromatographic display and dragging the mouse, the display will shift in the direction of the drag (lower beginning spectrum number when the drag is down and higher ending spectrum number when the drag is up). This same dragging motion while holding down the right mouse button will result in the displayed range contracting (down) or expanding (up) without changing the current position, which is very convenient. A double click of the left mouse button with the pointer below the baseline restores the full RTICC. The amplitude of the chromatographic peaks can be increased or decreased with the use of the mouse wheel. Optionally, the RTICC can be set to normalize to the spectrum with the highest total ion current when the chromatogram is

zoomed. Mass chromatograms (called ion traces in this program) of up to two different individual  $m/z$  values can be displayed with or without the RTICC. The  $m/z$  values for the mass chromatograms are set by using mouse actions to change the two values on the left and bottom row of the Toolbox (Figure 2). The display of the mass chromatograms is toggled on and off by clicking on the user-defined values. The numeric value on the right of the bottom of the Toolbox represents present intensity subtraction of the peaks that are in the selected background spectrum. Background subtraction occurs only when this number is not grayed out.

There is a horizontal line on the chromatogram display that is positioned to a point that is the center of the middle mass spectrum. This line is located on the spectrum in the chromatogram displayed to the right (middle spectrum position). Any spectrum in the displayed range of the chromatogram can be selected by placing the mouse pointer on the desired position and clicking the left mouse button once. Most GC/MS data

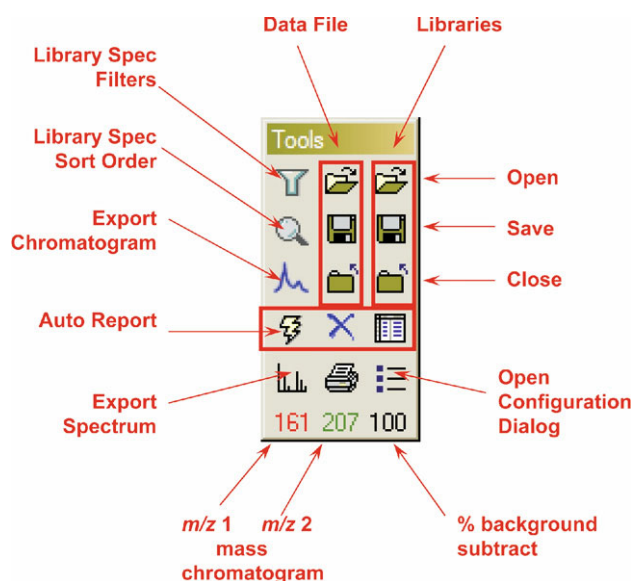


Figure 2. Definition of the Toolbox buttons.

systems selected a desired spectrum with a single or double click of the left (or, in the case of the Agilent ChemStation, the right) mouse button. What is different about MassFinder, however, is that chromatographic display is shifted to place that spectrum under this fixed-position horizontal bar. The number of spectra in the display range remains the same, but the starting and ending spectrum numbers change to acclimate this repositioning.

Another very convenient navigation aspect is the ability to use the mouse wheel to increase or decrease the amplitude of the chromatographic display. MassFinder 3, unlike most GC/MS data analysis programs, makes extensive use of the mouse wheel.

Because the chromatogram has no visible spectrum #/time axis, it can be difficult when you jump from spectrum to spectrum with no numeric reference points, even if the current retention time is always displayed with the spectrum in the status bar. However, as soon as some of the chromatographic peaks are marked, their retention time labels (optionally displayed) provide valuable points of orientation. The author said he has had these comments before and is trying to decide what can be done to improve on the situation without destroying the clean lines that currently exist. There is a keyboard command that allows the display of a zero line so that it is possible to determine the level of background in the data file.

The facile nature of being able to have an almost analogy-type navigation of the chromatograms is one of the things that makes MassFinder 3 so attractive for data manipulation and information extraction.

## Mass Spectral Display

The spectral display next to the chromatogram consists of three consecutive spectra from the data file: the

selected spectrum that is identified by the fixed horizontal blue line on the chromatogram (middle); the spectrum that is one spectrum number greater than the selected spectrum (bottom); and the spectrum that is one spectrum number less than the selected spectrum (top). This makes it very easy to identify a noise spike in a spectrum, spectral skewing, coelution, and so forth. The middle spectrum will appear highlighted compared to the top and bottom spectra. Putting the mouse pointer on the bottom spectrum and clicking the left mouse button once will advance the highlighted spectrum to the next highest spectrum number. Putting the mouse pointer on the top spectrum will move the highlight to the next lower spectrum number. Spectra can be scrolled through by placing the mouse pointer on any of the three spectra and using the mouse wheel. This moving through the chromatogram one spectrum at a time with the previous and next spectra also displayed provides a slow-motion animation of the data file that can be quite revealing. Use of this capability makes spectral skewing and coelution very obvious when present. Clicking on the  $\nabla$  symbol in the upper left corner of the middle spectrum will cause the symbol to turn black ( $\blacktriangledown$ ) and the same symbol will appear on the chromatogram, indicating that this is the spectrum to be used as the background. Now, it is not possible to use an average of several spectra as a background spectrum or to view the average of several spectra in an attempt to compensate for spectral skewing. I consider this to be a disadvantage of the program. MassFinder 3's author said that he has had this complaint from others and plans to add the ability to work with averaged spectra in a future release. If another spectrum is selected to be considered as the background spectrum, the  $\blacktriangledown$  symbol moves from the previous position in the chromatogram to the position of the newly selected spectrum.

Putting the mouse pointer on the spectrum and clicking the left mouse button will result in the spectrum display changing to have a line that includes the retention index of the spectrum, the analyte's possible nominal mass (this function really reports the highest non-isotope  $m/z$  value in the spectrum; it is not very useful except for pure spectra exhibiting a molecular ion peak), and the  $m/z$  value of the base peak (which is usually obvious from the graphic display of the spectrum). The  $\blacktriangledown$  symbol will appear on the chromatogram along with the spectrum's retention time if the appropriate option has been selected in the **Configuration** tab of the program's **Settings and Configuration** dialog box. A second click of the left mouse button with the pointer on the spectrum will cause the name line on the spectrum and the symbol marking the retention time on the chromatogram to disappear. The text in this name field can quickly be edited by using a keyboard command to call the label editor. This label editor stores names that were previously used so that they are available for other chromatographic peaks that may be given the same name, such as "contaminant."

The number of  $m/z$  units displayed in the spectral display is fixed by settings in the **Configuration** tab of the **Settings and Configuration** dialog box. By placing the mouse pointer below the  $x$ -axis of the sample spectrum in the middle display or the library spectrum in the middle display, holding down the left mouse button, and dragging the mouse from right to left or left to right, the display can be scrolled while maintaining the same  $m/z$  range. Placing the mouse pointer in the same position while holding down the right mouse button will allow for expanding (drag to the left) or contracting (drag to the right) the display. Placing the mouse pointer in the same position below the spectrum's  $x$ -axis and double clicking the left mouse button results in the display specified by the limits set in the Configuration to be restored. This fixed-limit display could result in missing important peaks such as a molecular ion peak above the upper  $m/z$  limit or diagnostic peaks such as  $m/z$  30 and  $m/z$  31, which are characteristic in the mass spectra of aliphatic amines and alcohols and may appear below the lower  $m/z$  limit. However, if there are mass spectral peaks that do not appear below the currently displayed lower limit or the currently displayed upper limit, a + sign will be displayed to the left or right of the  $x$ -axis, respectively.

The status bar under the spectra from the chromatographic display shows the spectrum number of the spectrum in the middle display along with the total number of spectra in the data file and the spectrum's retention time. The name of the data file also appears in the status bar below the spectral display. This is somewhat redundant because the data file name is also in the program's title bar, although redundancy is not always bad. The data file and library file name in the status bar make it easy to differentiate between the two types of spectra, and the data file name in the title bar makes it clear what is being processed. When the background subtraction is turned on, the three spectra will have **diff** printed in the upper left corner. An alternate display allows for the background-subtracted spectrum to occupy the center position, the nonbackground-subtracted spectrum to occupy the top position, and the spectrum being used as the background spectrum appearing as the bottom spectrum. This is an especially useful display when trying to separate two components that may be coeluting in the same RTIC chromatographic peak, or where a spectrum from the library is being used with a spectrum of a mixture of two or more compounds.

MassFinder 3 is designed to perform a library search of every spectrum that appears in the center display of the three spectra next to the chromatographic display. The results of the library search are displayed on the right side of the program's display. The library search is complete as soon as the sample spectrum is displayed. There is no delay in waiting for the results of the library search. The library spectra are displayed sorted according to a sort order that is selectable from the **Search Filter** tab of the **Settings and Configuration** dialog box.

The two most often used are **spectral similarity** and **retention index**. These two options (out of a total of seven) can be toggled using a button on the tool bar. When the matched library spectra are sorted by retention, the center spectrum (one adjacent to the sample spectrum) is the one with the most similar retention index and the top spectrum is one entry before the center spectrum with respect to the retention index. When the spectra are sorted by similarity, the top position will be blank; and the matched library spectrum with the highest similarity will occupy the middle position, next to the chromatogram spectrum that was searched. The number of the matched library spectrum will be in the status bar at the bottom of the display following the word "hit" along with its position in the hit queue. The name of the library being used will also be displayed in this region of the status bar. By placing the mouse pointer on the top or bottom library spectrum and clicking the left mouse button, it is possible to scroll through the list of library hits. The mouse wheel can also be used for this purpose. This is a very convenient way to do a comparison of the sample and library spectrum. The match quality is provided as a numeric value in the upper left corner of the library spectrum. If the mouse pointer is placed on the middle library spectrum and the left mouse button is clicked, the name of the library spectrum will be added to the middle sample spectrum. The ▼ symbol will appear on the chromatogram to mark the position of the identified component. Placing the mouse pointer on the library spectrum and clicking the left mouse button a second time will cause the name to disappear from the sample spectrum and the indicator symbol to be removed from the chromatogram.

The three library spectra are used in cases of parametric retrieval when the spectra meet specific criteria such as  $m/z$  value for the base peak, molecular weight, Chemical Abstracts Service registry number (CASrn), name fragment, or an  $m/z$  value peak with a greater than specified relative intensity. These spectra will all have **search filters** in the upper left corner to indicate that they are not related to the spectra displayed from the chromatogram. The Edit Library Entry function has a validity check for CASrn designations to make sure the entered value matches the construction rules; that is, the official checksum algorithm is applied.

A library spectrum can be selected by use of the parametric retrieval techniques and positioned in the center position. If the mouse pointer is placed on this library spectrum and the right mouse button is clicked, MassFinder 3 will do a true reverse search where this spectrum is searched against every spectrum in the data file to see which spectrum in the data file provides the best match. The position of that spectrum will be indicated by the position of the horizontal blue line on the chromatographic display. To our knowledge, MassFinder 3 is the only system for review of GC/MS data that has this feature.

Another interesting feature of MassFinder 3 is the ability to edit library entries. The **Modify Library Entry** dialog box has a check box that will cause the current library spectrum to be replaced by the sample spectrum that is in the center chromatographic spectrum display. It is important to note that any library field can be edited. When a field is edited or a new spectrum is added to a library and is marked as “modified” or “added”, it is possible to do a parametric retrieval for all modified spectra. Structures are easily added to spectra by copying them from the Windows Clipboard. You can use any structural drawing program to create a structure and then copy it as a MOL file or in any metafile graphic format to the Windows Clipboard. There is a button that allows you take a structure from the Windows Clipboard and associate it with the library spectrum. There is another button that allows a structure associated with a spectrum to be copied to the Windows Clipboard. This structure can then be pasted into a structural drawing program for modification or some other program such as Word for display with an exported spectrum.

MassFinder 3's ability to edit library entries, including the spectrum itself, means that its flexibility is considerable. Most other systems require that a spectrum be called from the library, edited, added to the library, and then the original spectrum must be flagged as deleted.

Building new libraries is as easy as adding spectra to existing libraries. With no library open, just use the add-spectrum-to-library command. This will start a new library with the chromatogram spectrum in the center display. The field for the name of the file containing the spectrum, the **Retention Index** (RI) and **Retention Time** fields, and **Scan Number** field are automatically populated. The **Name Field** will be populated with the word **unknown**. This means that a library can be quickly built and then updated with structures, CASrn designations, formulas, and other information at a later time. Because of MassFinder 3's unique ability to assign retention indexes to each spectrum, having the RI associated with a spectrum is very easy. Structures for all created libraries are stored in a single folder inside the folder containing MassFinder 3. These structures are in a Windows metafile format and can be opened in ChemDraw from CambridgeSoft (<http://www.cambridgesoft.com>) for editing. All of the structures, regardless of the presence of multiple libraries, appear to be stored in this single folder. It is not possible to tell which structures are associated with which libraries unless the unique structure file name is known. There is no utility to export libraries. The author, on request, will convert a MassFinder library to another format, but this conversion does not include structures. Libraries can be imported from other formats but not with structures.

In addition to the utility that allows for adding and editing library entries, there is a utility that allows for the addition, modification, and removal of  $m/z$  value-

intensity pairs from a mass spectrum in the library. This utility can also be used to manually create a library spectrum entry.

One problem that exists with most GC/MS data systems concerns an ability to achieve good graphics for inclusion in reports and presentations. MassFinder 3 has addressed the export of graphics with separate utilities for spectra and chromatograms. The same drag-and-drop expansion, scroll, and expand functions that exist for chromatograms and spectra incorporated into the main program exist in the Export displays. This means any type of chromatographic or spectral display can be generated as shown in Figure 3 for the chromatogram and Figure 4 for the spectrum. The spectrum graphic will optionally allow for labeling of mass spectral peaks, showing  $m/z$  values and tick marks on the abscissa, showing tick marks and relative present values on the ordinate, including the text from the spectrum's name field, and, most important, showing structures on the library spectra and sample spectra that have been given library-spectra names. The highlighted sample spectrum and library spectrum are available for output when the Spectrum Export function is evoked from the program's **Toolbox** by simply clicking on the **GC** or **Lib** button at the top of the display, which toggles back and forth between the two spectra.

The color of the graphic is selectable as is the ability to toggle between thick and thin lines. If a structure is displayed with the spectrum, the structure remained black when an alternate color is selected for the display.

The Export Spectrum function also has a Text Export feature that could be used. This would output the labeled mass spectral peaks in a format of ( $m/z$  value) relative percentage; that is, 69 (15). This makes creating documents such as theses or reports much more convenient because users will not have to type  $m/z$  value-intensity lists. Pairs of  $m/z$  values-intensities could be added or removed from this list by clicking on unlabeled peaks (to add them) or labeled peaks (to remove them). All outputs from both the Export Chromatogram and the Export Spectrum function sent an output to the Windows Clipboard or to a printer. In addition to the Text Export and Graphics Export functions of the Spectrum Export, there was the ability to copy all the  $m/z$  value-intensity pairs in a tab-delineated format to the Windows Clipboard for import into Excel.

The Chromatogram Export function has a unique ability to label chromatographic peaks. The labels are orientated vertically rather than horizontally. You can choose to label either peaks that are associated with a library spectrum (identified spectra), peaks that are investigated but are not associated with a library spectrum (what the program refers to as unknown scans), or both. The labels can be just the retention time, the retention time followed by the name, or a tick mark positioned at the position of the spectrum that is labeled. Any combination of these three is possible. A sample spectrum's name field can be populated with the name of the adjacent library spectrum; or it can be

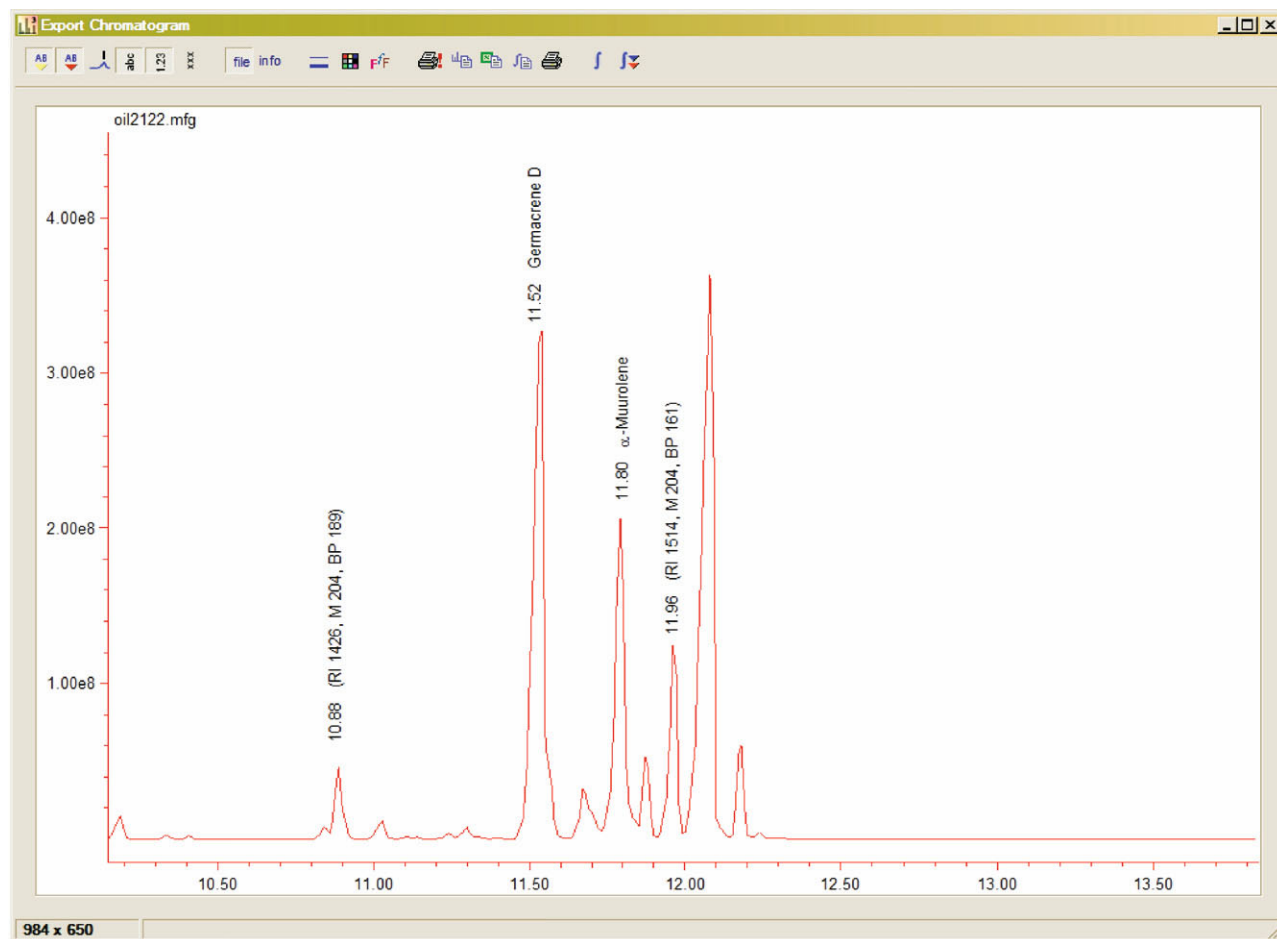


Figure 3. An example of the Export Chromatogram view.

given the default unknown name that contains the retention index value (RI), possible nominal mass value (M), and the  $m/z$  value of the base peak (BP) when the mouse pointer is put on the sample spectrum and the left mouse button is clicked. If the label is added first, its position can be moved relative to the chromatographic peak. If the retention time is then added, this new position will be respected. The tick mark defining the spectrum that is labeled in the chromatogram will always be positioned just above the intensity of that spectrum in the chromatogram.

The Chromatogram Export function allows for peak integration. There are two integration modes: Simple Integration and Advanced Integration. The Simple Integration mode allows the user to select the region of the baseline that is to be integrated. The values reported are a percentage of the summed peak areas. The Advanced Integration mode functions with the Auto Report function. The Auto Report function will examine every spectrum in the chromatogram to determine the best match to the spectra in the library. This is a little like a target compound analysis where the target analytes are in the library. This is an important concept of MassFinder 3 using a library with spectra of specific compounds for a specific type of analysis. The Auto

Report has several settable options that allow the user to customize the ability to identify the analyte. Results of the Auto Report function will be a table that contains the retention time, name (from library spectrum), integrated area, and the percentage that integrated area is of the total area for all the identified chromatographic peaks. In the event that there is no matching spectrum/retention index in the library, but the chromatographic peak is above a set threshold, that peak is identified as an unknown. This table can then be exported to the Windows Clipboard for import into Excel for quantitation purposes. MassFinder 3 sets the baseline for each integration. The Advanced Integration function allows the user to redefine that baseline for each peak to have a better value for the peak area. The MassFinder Auto Report feature does not assign the same library spectrum to more than a single chromatographic peak as can be the case with other GC/MS data systems because of the use of the retention index. Another important point about the Auto Report is the Peak (chromatographic) Finder function. Chromatographic peaks are defined through thresholds and a number of spectra across the peak. These files are determined by the program while viewing a sample chromatographic peak. There is a chromatographic peak evaluation but-

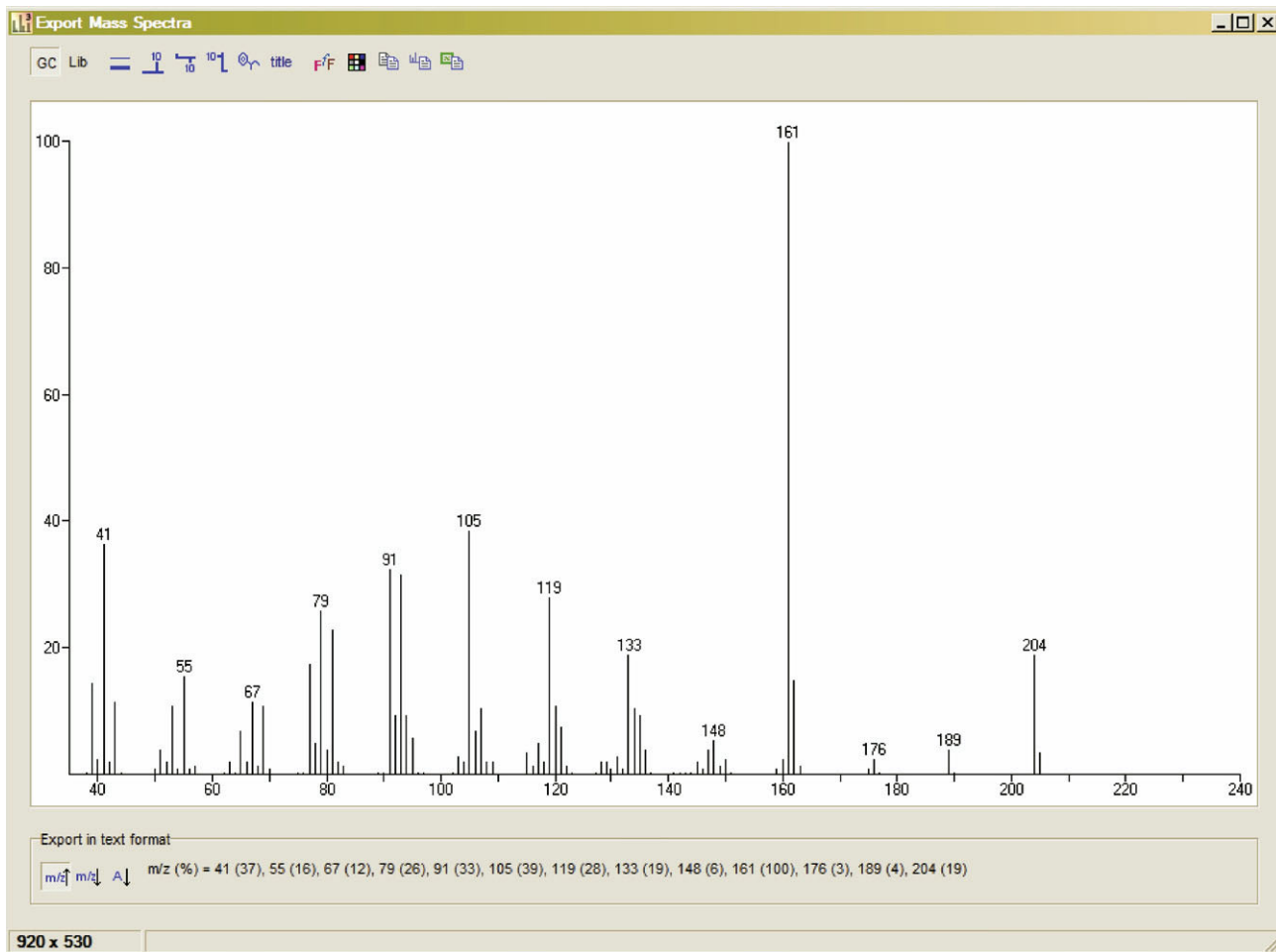


Figure 4. An example of the Export Spectrum view.

ton on the Auto Report tab of the Configuration and Settings dialog box. It should be noted that the Peak Finder will not deconvolute chromatographic peaks representing more than a single analyte. The Auto Report and Peak Finder functions are dependent on chromatography that produces a single peak for each compound (high-resolution gas chromatography).

In the version evaluated, the mass chromatograms, even when displayed without the RTICC, were not set to the Chromatogram Export function; and integration was done only on the RTICC. This possible shortcoming was discussed with the program's author, who stated that he would consider some alternatives that would include mass chromatograms. Another potential shortcoming in the tested version was an easy ability to send a spectrum to the NIST Mass Spectral Search Program to search it against the NIST/EPA/NIH Database or the Wiley Registry. Spectra can be exported in a tab-delimited format to Excel, a NIST-compatible text file created, and then that file opened with the NIST MS Search Program; however, this process could not be considered convenient. The author stated that a future version of MassFinder 3 will include a seamless export to the NIST Program, much like that available from the

Waters, Agilent, Thermo Fisher, Varian, Perkin Elmer, and Agilent GC/MS data systems.

The program uses a USB hardware key that can be installed on an individual computer or on a computer on a network so that other work stations (one at a time) can use it. Installation of the hardware key is accomplished with a standard Install Shield program; however, installation of MassFinder 3 is much more primitive. The user must create a folder and then copy the files for the program from the distribution CD to this folder. There are no program groups or desktop icons created. The author said this was done in deference to the academic community who like to avoid the rigors of a number of formal installations. I found this type of installation to be somewhat amateurish and may be difficult for many. The documentation that accompanies MassFinder 3 is very good and well written, especially considering that the program is from Germany where English is not the first language. The documentation is divided into several sections, and users are advised to open and print all the PDF files and to become familiar with each section before beginning use. As can be seen from Figure 4, this program does not have the standard Windows program appearance.

The tool tips (labels that appear as the mouse pointer moves over the icons of the Toolbox or that are present on various displays) are very well done and provide an abundance of information. The evaluated version had no Help file or Help buttons on any of the displays. There is no automation capability. That is to say, there is no way to automatically apply the Flash Report function to a series of files. This may not be necessary because of the speed of all the various functions of the program.

There is a network version of MassFinder 3 that does offer some convenience and multiple-user support at a reduced cost.

MassFinder 3 looks different, functions differently, and offers different capabilities than those of other GC/MS data analysis programs. It is an interactive program that allows the extraction of a great deal of information for a GC/MS data file. Another very im-

pressive feature of this program is the openness of the author to suggestions for enhancements and improvements. Upgrade notices are sent by e-mail to all users so that they can download upgrades from an FTP site. This download site is open for only a few weeks after the upgrade, but if the download period has expired, the author will provide the current version. This program is well suited to any type of qualitative GC/MS analyses and can prove very useful in any laboratory.

#### *Reviewer's Note*

Some parts of this review may appear to be a little overly didactic. This is because the subtleties of this software can be presented only in this way without being able to view the program in operation. It is these subtleties that set this GC/MS analysis software apart from other programs.