KnowItAll Software Training

LC Expert

Automatic LC-MS Processing and Analysis

How to use KnowItAll LC Expert to Perform Automatic LC-MS Searching and Analysis

Purpose

These exercises demonstrate how to use KnowItAll LC Expert to automatically analyze LC-MS chromatograms.

Objectives

These exercises will teach you how to:

- Use KnowItAll LC Expert to deconvolute chromatograms into peaks for further analysis
- Perform an untargeted databases search
- Perform a targeted accurate mass search
- Apply MSforID search algorithm

Background

LC-MS chromatograms are rich in information. Analysis is challenging and curated libraries are time consuming to search through. LC Expert application allows for the automatic deconvolution of the chromatogram into peaks, which can be further analyzed and then searched for known and unknown targets. MSforID search algorithm is included for high accuracy LC-MS searching. Users of LC Expert are encouraged to create user libraries with their in-house compounds to streamline their workflows using KnowItAll.

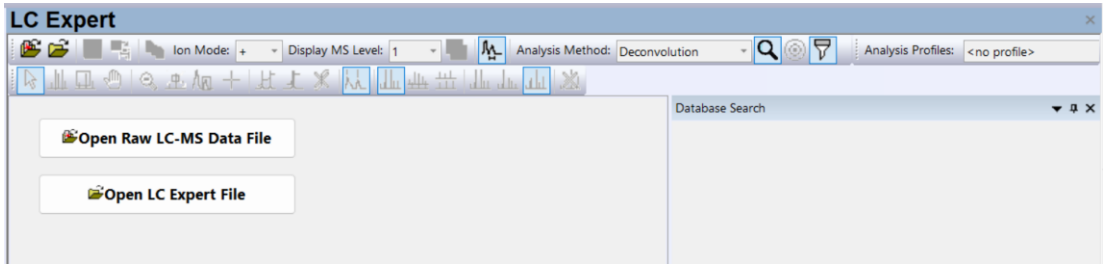
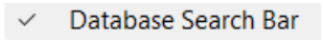
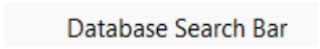
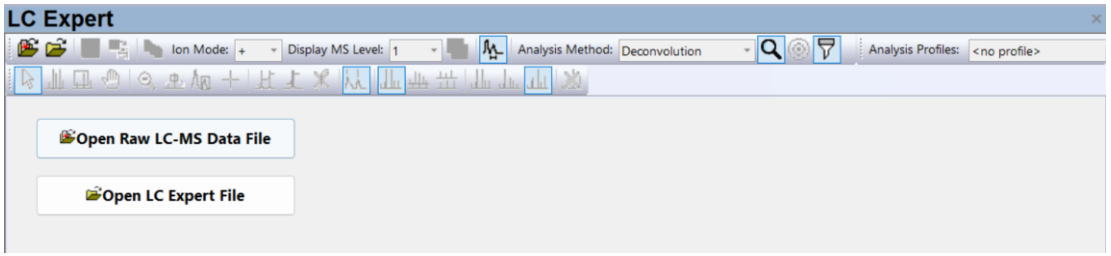
Training Files Used in This Lesson

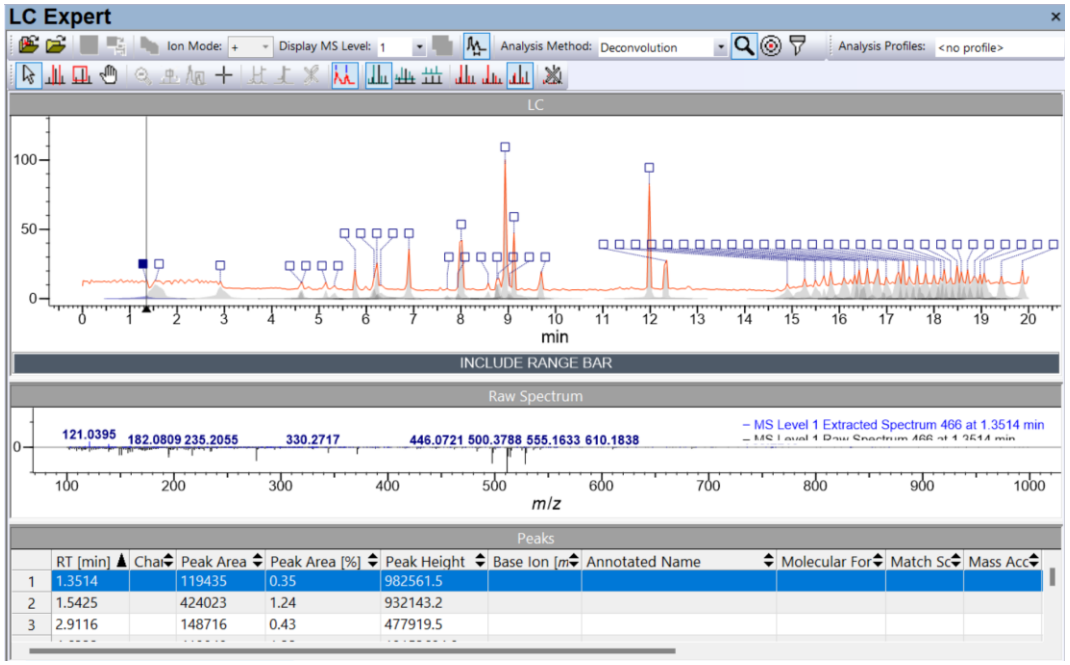
- Folder files in C:\Users\Public\Documents\Wiley\KnowItAll\Samples\LC-MS

KnowItAll Applications Used

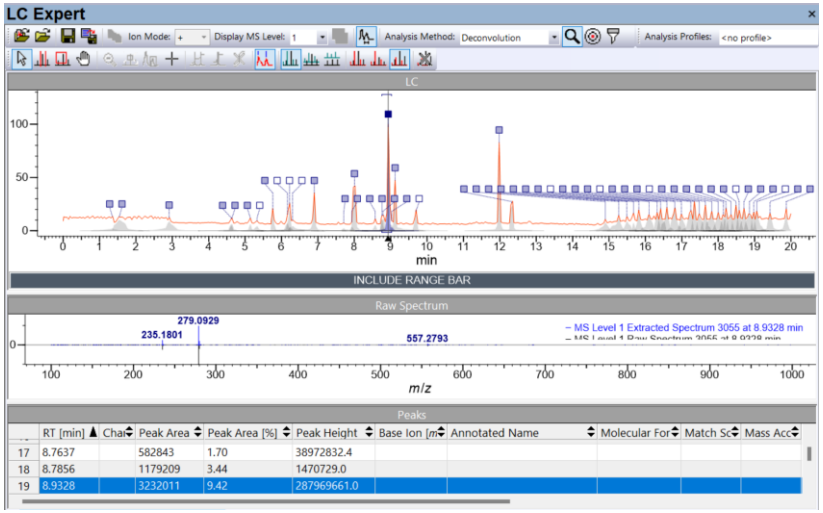
- KnowItAll LC Expert
- KnowItAll SearchIt

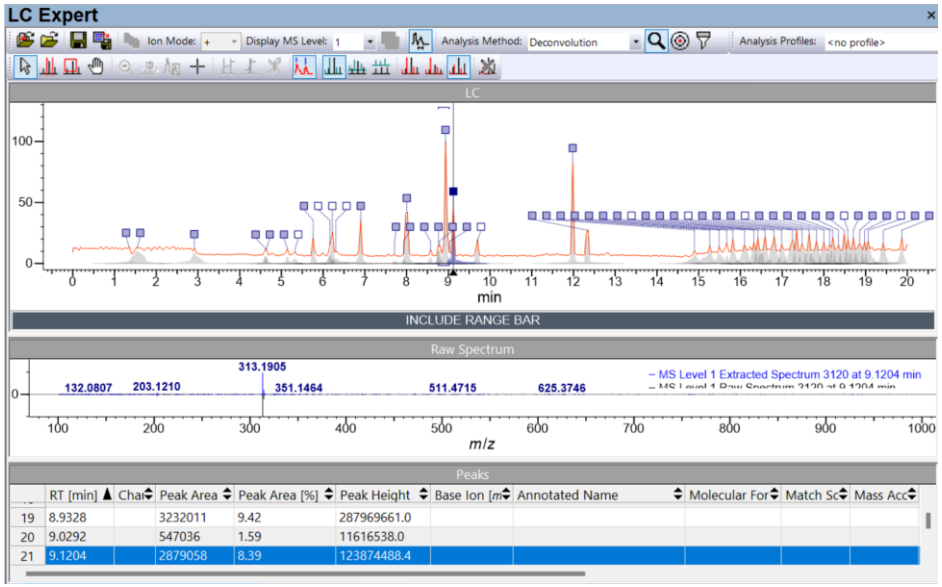
Example: Open a Chromatogram in LC Expert

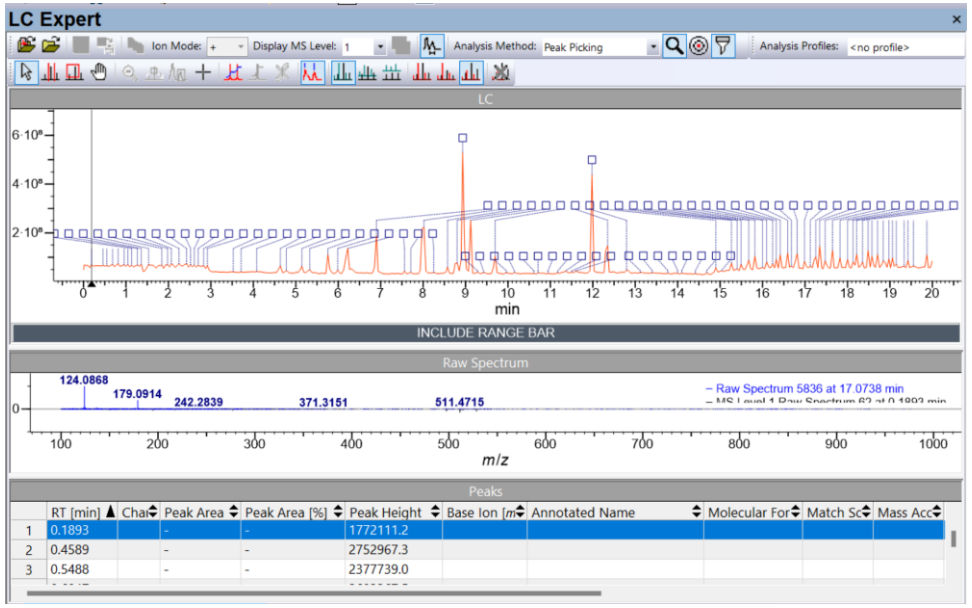
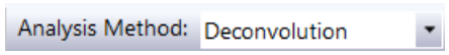
	Action	Result
1	Open **LC Expert** application by clicking on the icon (), typically found in the **Spectral Processing** toolbox.	**LC Expert** application is displayed:  *Note:* **LC Expert** must be in your current license to have access to the application.
2	To simplify the next portion of the training, hide the **Database Search** panel by clicking **View > Database Search Bar** to remove the checkmark (*e.g.,* as shown in figure below). *Before Deselection:*  *After Deselection:*  	Upon deselection of the **Database Search** bar, the panel is hidden.  *Note:* The **Database Search** bar will be discussed in the subsequent section.

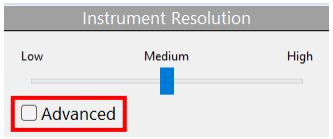
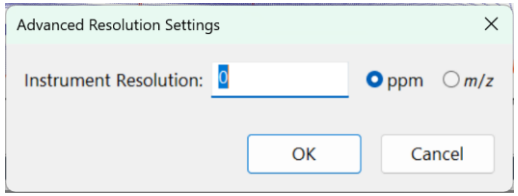
	Action	Result
3	Click **Open Raw LC-MS Data File** button. Navigate to **"C:\Users\Public\Public Documents\Wiley\KnowItAll\Samples\LC-MS\"**. Select "**TESTMIX2_180504_MAS011_06 .mzXML**" and open in the application.	The file opens in **LC Expert** application:  • The top panel displays the **Chromatogram** with **Peak Boxes** representing each peak. • The middle panel is the **Raw Spectrum** panel in which the extracted spectrum and raw spectrum are displayed. • The bottom panel is the **Peaks Table** which shows the deconvoluted peaks generated from the **Chromatogram**. Each row in the table includes information about each peak. • The detected ion polarity is displayed in the **Ion Mode** panel, located on the **Standard Toolbar**.

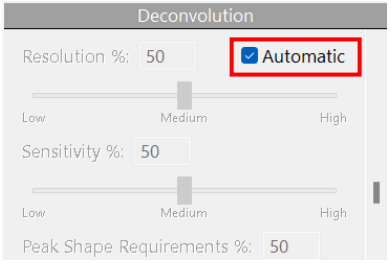
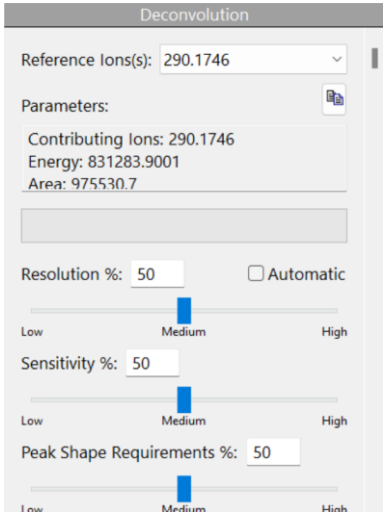
	Action	Result
4	Click the **Include Range Bar** with left mouse button and drag left/right to select a region for analysis. *Note:* This can also be achieved by clicking on the **Include Range Bar** with the right mouse button. On the proceeding pop-up **Include Ranges** dialog, click **Add**. A space below **Low Range** will appear to manually input a value. Do the same below **High Range** to manually input a value. Click **OK**.	The **Include Range Bar** can be used to isolate analysis regions in the **Chromatogram.** Outside of these regions (shown in gray coloration on the **Chromatogram**), there is no deconvolution or additional analysis occurring: 
5	To remove the isolated regions from the **Include Range Bar**, click the trash can icon () on the right side of the **Include Range Bar**.	The isolated regions are removed from the **Chromatogram**, and the full chromatographic region is deconvoluted: 

	Action	Result
6	Click on the **peak box** (☐) for the tallest peak (located at 8.93 min).	Upon peak selection for the peak at 8.93 min: • The **Peak Box** is shaded to confirm selection. • The **Peak Area** is shaded with blue coloration, displaying the deconvoluted peak area. • There is a **Bracket** above and below the peak, to visualize the retention time region for the peak. • The related row in the **Peaks Table** is highlighted. 

	Action	Result
7	Click on another row in the **Peaks Table** (*e.g.*, row 21).	The related peak in the **Chromatogram** becomes selected: • The **Peak Box** is shaded with a darker coloration. • The **Peak Area** is shaded. • The retention time region is indicated by the **bracket**.  *Note:* If a peak's component can be identified by the user, its name can be manually entered into the **Peaks Table** by double clicking on the associated cell in the **Annotated Name** column.

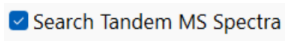
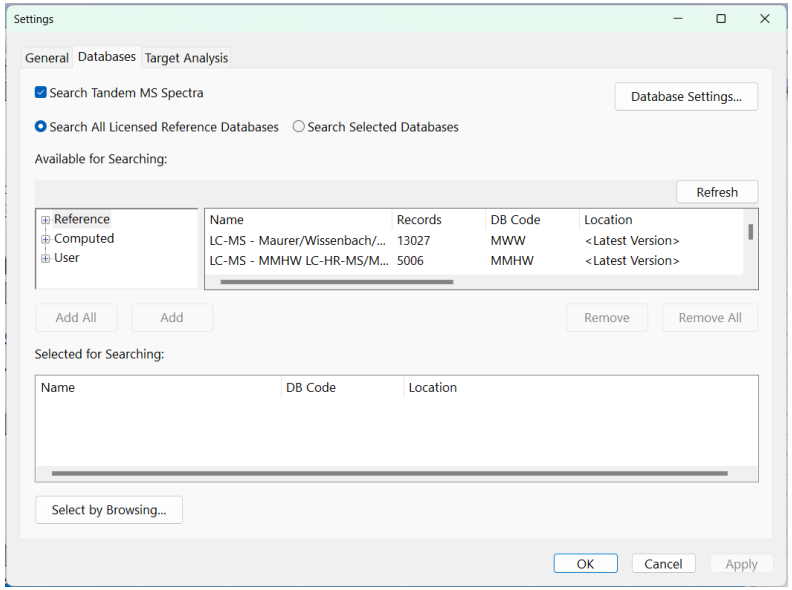
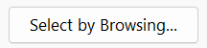
	Action	Result
8	*Note:* By default, the **Analysis Method** dropdown menu on the **Standard Toolbar** will be set to **Deconvolution** mode. In **Deconvolution** mode, correlated MS scans will be *grouped* together (into peaks) and the averaged MS1 scan for the peak is provided as the **MS Level 1 Extracted Spectrum** in the **Raw Spectrum** pane. The peak information (*e.g.,* **Peak Area**, **Peak Area Percent [%]**) populates in the **Peaks Table**. If preferred, the **Raw MS Spectrum** information for a *single* time point can be analyzed in the **Peaks Table** by changing the **Analysis Method** to **Peak Picking** menu option.	When the **Analysis Method** is set to **Peak Picking**, a single time point (from the **raw MS spectrum**) is used to characterize the peak and thus the **Peak Area** and **Peak Area [%]** columns are replaced with ' – ' symbols. Conversely, when **Deconvolution** mode is used as the **Analysis Method**, related MS scans are grouped together to for the peak, and thus **Peak area** and **Peak Area [%]** information is available in the **Peaks Table**. 
9	Before proceeding to Step 10, check that the **Analysis Method** is set to **Deconvolution** menu option.	The **Deconvolution** option for **Analysis Method** is selected: 

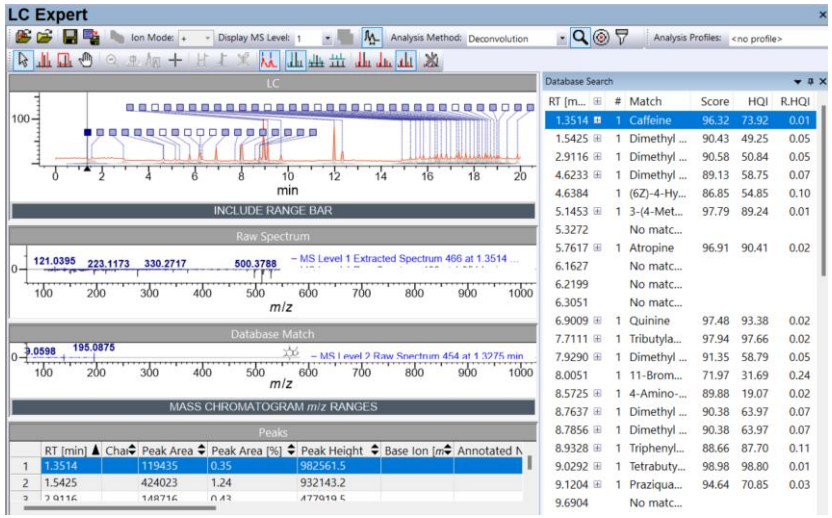
	Action	Result
10	To adjust the peak deconvolution settings, select **View > Deconvolution Settings** bar.	The **Deconvolution Settings** bar is opened. The **Deconvolution Settings** bar can be used to tune the deconvolution and instrument resolution: 
11	To change the **Instrument Resolution**, go to the **Deconvolution Settings** bar and click the checkbox in the **Instrument Resolution** panel next to **Advanced**. 	The **Advanced Resolution Settings** popup appears. 
12	Click **Cancel** on the popup to close the dialog.	The popup closes.

	Action	Result
13	Go to the **Deconvolution Settings** bar and deselect the checkbox next to **Automatic** in the **Deconvolution** panel. 	The **Resolution**, **Sensitivity** and **Peak Shape Requirements** become available for adjusting. The slider bars or numeric values can be used to control the deconvolution settings. As the deconvolution settings are changed, the number of peaks in the **Peaks Table** and **Peak Boxes** on the **Chromatogram** are changed. 
14	Click to reselect the checkbox next to **Automatic**. Select the exit icon (✕) on the **Deconvolution Settings** panel to hide the panel.	The application resumes automatic deconvolution for the peaks. The **Deconvolution Settings** panel becomes hidden.
15	To save the file, select **File > Save LC Expert File**.	An LC Expert Analysis file is saved to the location of your choosing. TESTMIX2_180504_MAS011_06.lca This file can be reopened to re-analyze datasets, or continue processing in the future.

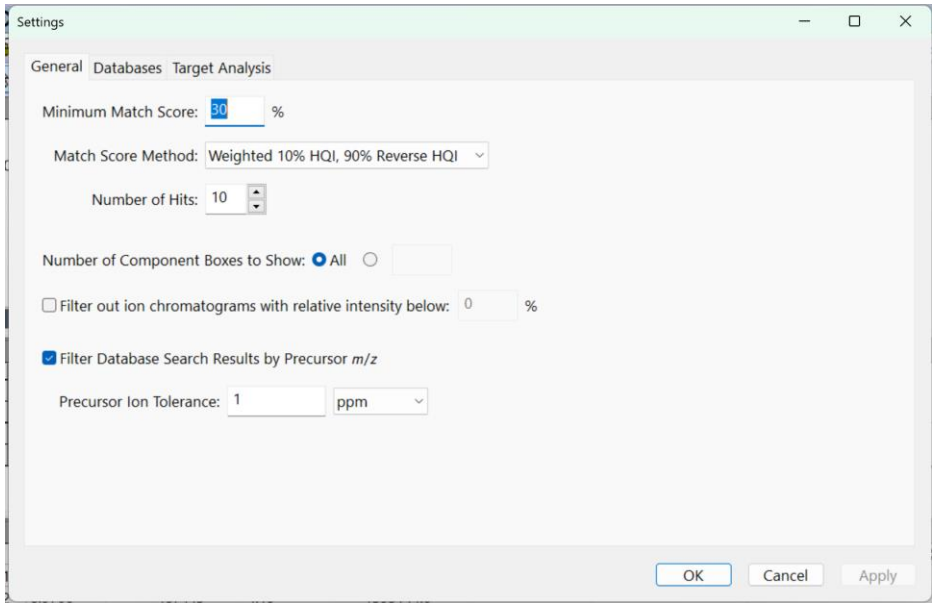
Example: Perform an Untargeted MS2 Search

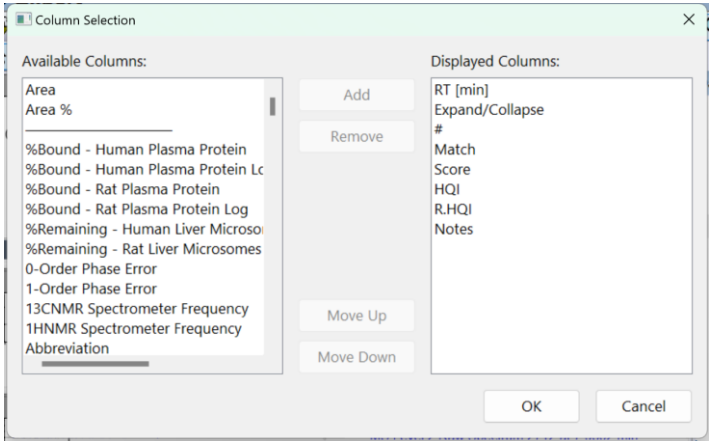
This section explains how to perform an untargeted library search for MS^n data.

	Action	Result
16	Continue using the chromatogram file from Step 2. In the **Settings** popup that appears, select **File > Settings**. Choose the **Databases** tab. Make sure the **Search Tandem MS Spectra** checkbox is selected. 	The **Settings** dialog opens.The **Databases** menu displays search settings. The specific databases displayed in the panel below “Available for Searching” depend on the user license. 
17	To choose the databases available for searching, select the radio button for **Search All Licensed Reference Databases**. Alternatively, select database(s) can be searched by using the radio button for **Search Selected Databases** and individually adding the desired libraries.	When the radio button for **Search All Licensed Reference Databases** is selected, then the **Selected for Searching** window is unavailable. When the radio button for **Search Selected Databases** is selected, then the **Selected for Searching** window is available. *Note:* Specific available databases depend on the user’s license. User databases can be added for searching by selecting the **Select by Browsing** button () and navigating to the desired database file (.sdbx).

	Action	Result
18	Click **Apply** then **OK** to save any changes made in the **Settings** dialog. Select **View > Database Search Bar**.	The **Settings** dialog is closed. The **Database Search** panel is visible on the riht side of the display window: • The deconvoluted peaks are searched using the selected libraries. • The peak retention times (**RT [min]**) in the **Database Search** panel align to the peaks in the **Peaks Table**. • Clicking on a row in the **Database Search** panel highlights the related row in the **Peaks Table**, and the peak in the **Chromatogram**. • The best search match for the MS2 spectrum is displayed as the top hit for each peak retention time.  The software performs a cosine similarity search to match the MS^n spectra agaisnt the applied databases. Using the default settings, the database search results are filtered by: • Matching ion polarity using the ion polarity applied in the **chromatogram**. • Precursor ion m/z. *Note:* Specific database matches in the **Database Search** panel will depend on the used licensed databases for searching, and the settings to be configured at Step 19.

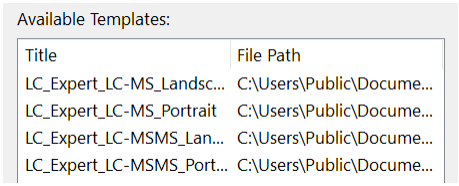
	Action	Result

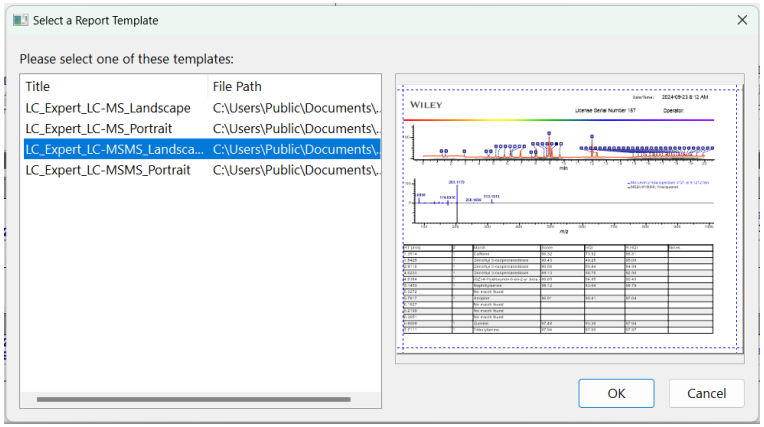
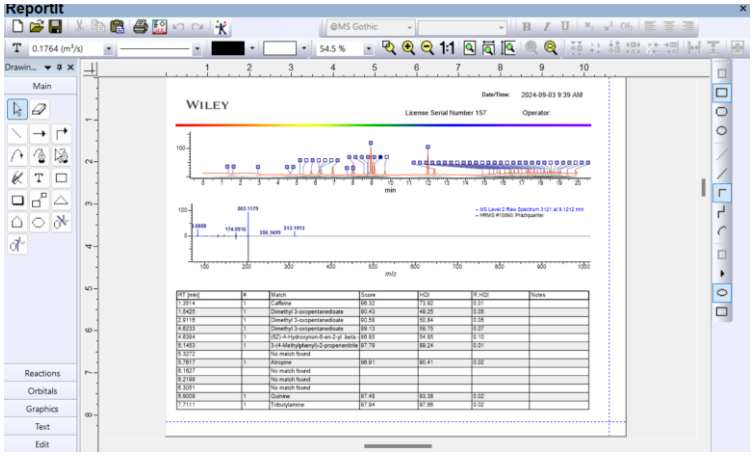
19	Go to **File > Settings**. Remain on the **General** tab, and change the Precursor Ion Tolerance to 1 ppm.	The **Settings** dialog is launched. The last applied settings will be retained in the application. There are settings for the execution of the cosine similarity search:
		• **Match Score Method** which defines whether HQI or R.HQI (Reverse HQI) should be prioritized.
		  ○ By default, R.HQI is more heavily weighted in **LC Expert**.
		  ○ The scoring method can be changed using the dropdown menu.
		• When selected, the checkbox for **Filter Database Search Results by Precursor m/z** forces the database search results to be filtered by the precursor m/z of the query spectrum.
		  ○ If deselected, query results will not be filtered by precursor m/z and all m/z values will be accepted.
		• **Precursor Ion Tolerance** provides the match tolerance for Precursor ion m/z.
		 

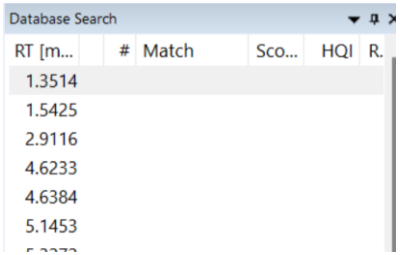
	Action	Result
20	Click **OK** in the dialog window, then click on the **expand** icon (⊞) next to a search match in the **Database Search** panel.	The top 10 best matches are displayed (or less than 10 matches if less than 10 were identified). Specific matches depend on the applied licensed databases, and the settings configured in Step 19:  *Note:* Double click on the cell next to the desired row in the **Notes** column to add a note.
21	Right click on the **Database Search Table** and select **Edit Component Table Columns**.	The **Column Selection** dialog window is launched. 

Database Search

RT [m...	#	Match	Score	HQI	R.HQI	Not
9.1204	1	Praziqua...	94.64	70.85	97.28	
	2	1-(4-Pyri...	94.58	58.61	98.57	
	3	Praziqua...	94.39	70.74	97.01	
	4	Praziqua...	93.86	70.27	96.48	
	5	Praziqua...	90.14	67.80	92.62	
	6	1-(4-Pyri...	89.31	55.35	93.08	
	7	2-Piperidi...	88.84	53.10	92.81	
	8	N,N-Di(p...	87.95	52.55	91.88	
	9	N-(2,6-Di...	87.91	56.77	91.37	
	10	(S)-1-(Pyr...	84.62	60.97	87.25	

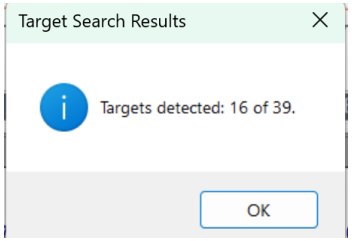
	Action	Result
22	Select a column in the **Available Columns** list, *e.g.,* "CAS Registry Number". Click on the **Available Column** and select **Add**. Click **OK** to proceed.	The column is added to the **Displayed Columns** cell at the bottom of the list. Upon clicking **OK**, the dialog window is closed. The new column is added as a column to the **Database Search Table**.  Additional columns can be added to the **Database Search Table**, to add helpful metadata when scanning search results.
23	Go to **File > Edit Report Templates** to import report templates.	The **Report Templates** dialog window is launched. If there are templates already in the dialog, skip the next step. 
24	Click **Add** then navigate to: **"C:\Users\Public\Documents\Wiley\ KnowItAll\Report Templates\LC Expert"**. Select all 4 report templates in the folder. Click **Open**.	The report templates are added to the **Available Templates** box. 

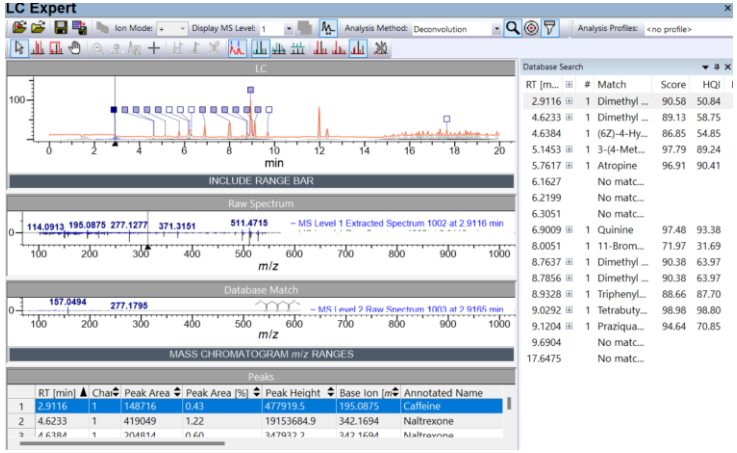
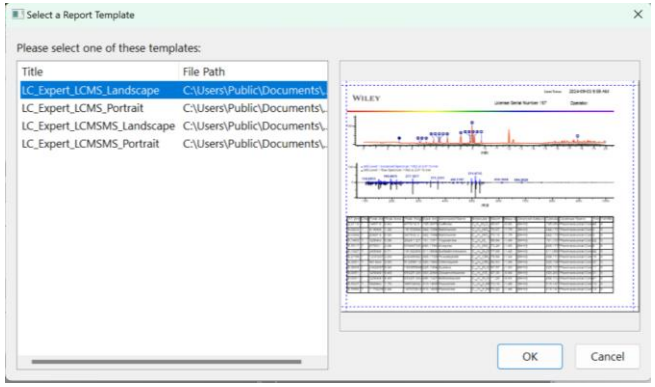
	Action	Result
25	Click **Close** on the dialog window. Using the **Transfer To** (Transfer to:) bar, click **ReportIt** (ReportIt).	The **Select a Report Template** dialog window is launched with the active chromatogram previewed in the selected report template. 
26	Click on **"LC_Expert_LC-MSMS_Landscape"** then **OK** on the dialog window.	The **Report** is generated, displaying the MS^n search information from the **Database Search Table**: 

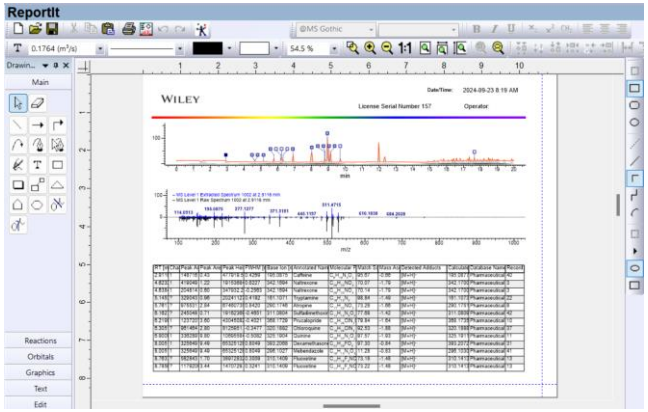
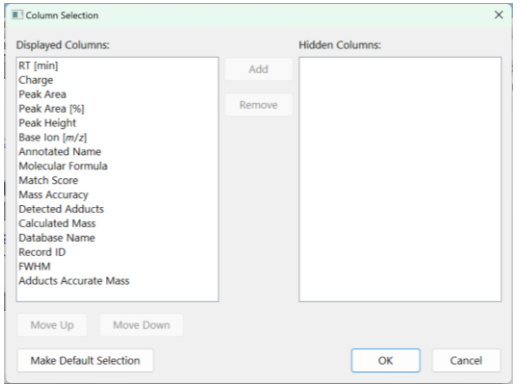
	Action	Result
27	Ue the **Back Arrow** icon ( ) to return to **LC Expert**.	**LC Expert** is opened.
28	*Note:* To stop the database search from taking place in the software background, deselect the related icon ( ) located on the **Standard Toolbar**, or turn off the database search setting given in Step 16. For the next section, the **Database Search** remains selected.	Upon deselecting the setting, the **Database Search Table**'s search results depopulate: 

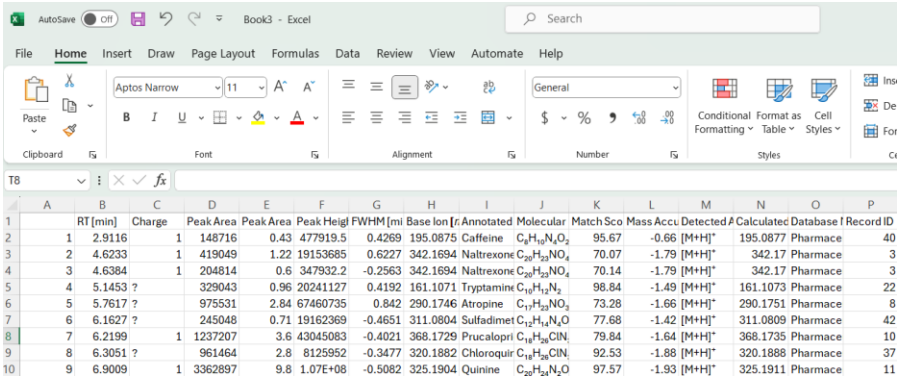
Example: Accurate Mass Searching

This section describes how to perform an accurate mass search within the chromatogram file. LC Expert's Targeted Analysis workflow searches the chromatogram for a list of compounds in a target list using the exact mass of the targets.

	Action	Result
29	Continue with the **Chromatogram** from the previous section. Click on the **Targets** icon () or choose **Analysis > Targeted Analysis**. Navigate to **"C:\Users\Public\Public Documents\Wiley\KnowItAll\Samples\LC-MS\"**. Select **"Pharmaceutical Compounds.sdbx"** and click **Open**. After reading the **Target Search Results** popup, click **OK** to close the popup window. *Note*: The sdbx file imports a list of compounds as targets to search for in the chromatogram. Individual targets can also be searched for by transferring a structure from **ChemWindow** into **LC Expert** using the **Transfer To** bar.	Upon selecting the **Targets** button, a **File Explorer** window opens. After opening the sdbx file: • The **Target Search Results** popup provides the number of found compounds in the chromatogram.  • The **Peaks Table** updates with the detected compound information: ○ **Annotated Name** is the compound record name from the sdbx file. ○ **Base Ion [m/z]** is the base ion from the MS1 extracted spectrum. ○ **Molecular Formula** for the identified compound (*i.e.,* target). ○ **Match Score** is the match score calculation using the target's accurate mass and the calculated exact mass. ○ **Mass Accuracy** is the mass accuracy calculation using the target's accurate mass and the calculated exact mass. ○ **Detected Adducts** is the adduct which is detected in the extracted spectrum and applied in the exact mass calculation. ○ **Calculated Mass** is the exact mass for the target with the detected adduct. ○ **Database Name** is the name of the imported sdbx file used as the target list. ○ **Record ID** is the specific record ID from the sdbx file to identify the detected target. *Note:* The **Peaks Table** and **Database Search Table** may be filtered at this step, which is discussed in Step 30.

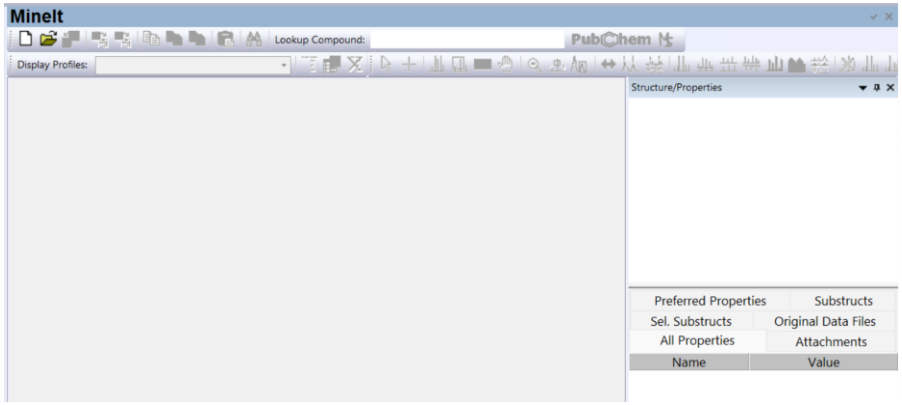
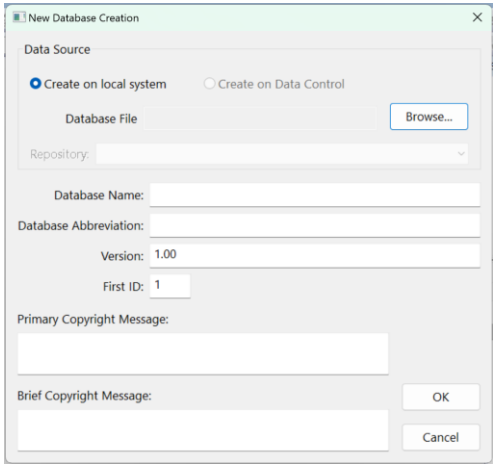
	Action	Result
30	The **Peaks Table** can be filtered/unfiltered by selecting/deselecting the **filter** icon ( ).	When the filter is selected: • The **Peaks Table** hides any row that does not have a detected target. • The **Peak Boxes** on the **Chromatogram** are filtered to match the displayed rows in the **Peaks Table**. • The **Database Search** table will only display matches for peaks visible in the **Peaks Table**. 
31	Select the **Transfer To** (Transfer to:) bar to transfer to **ReportIt** (ReportIt).	The **Select a Report Template** dialog window is launched: 

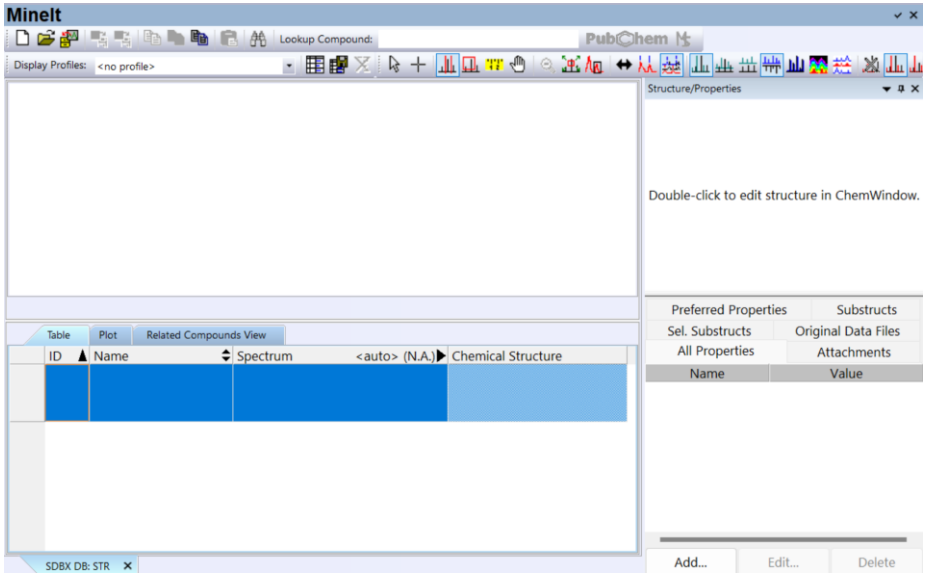
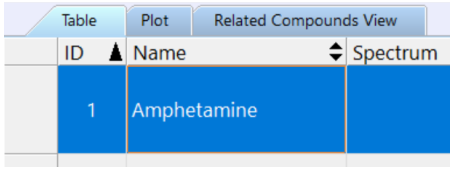
	Action	Result
32	Select "**LC_Expert_LC-MS_Landscape**" on the dialog window.	The **Report** is generated displaying the accurate mass information from the **Peaks Table**:  *Note:* The active display for the **Chromatogram** will be retained in the report (*e.g.*, filtered versus unfiltered **Chromatogram** from Step 30).
33	Use the **Back Arrow** icon () to return to **LC Expert**.	**LC Expert** is opened.
34	The **Peaks Table** can be modified by right clicking directly on the table and selecting **Edit Column Display**. Click **Cancel** to close the dialog window.	The **Column Selection** dialog window is launched, which allows for hiding columns and rearranging their order. 

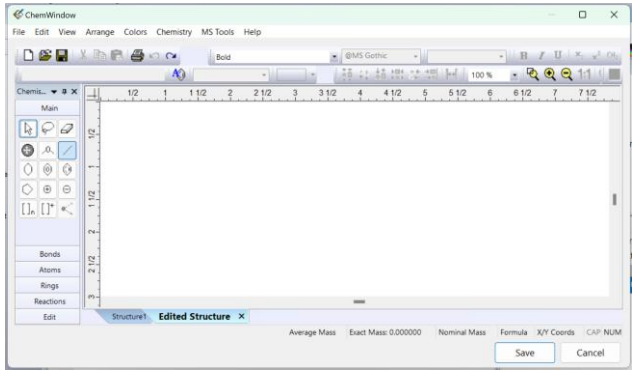
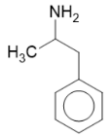
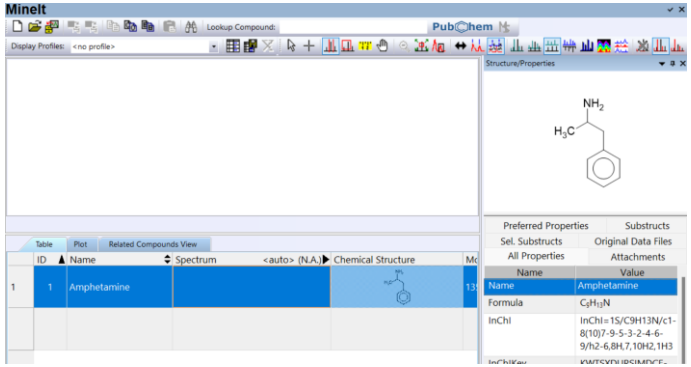
	Action	Result
35	The **Peaks Table** can be copied into a document by right clicking on the table and selecting **Copy Table to Clipboard**.	The table is copied into a document:  The active display for the **Peaks Table** will be retained in the report (*e.g.*, filtered versus unfiltered **chromatogram** from Step 30).

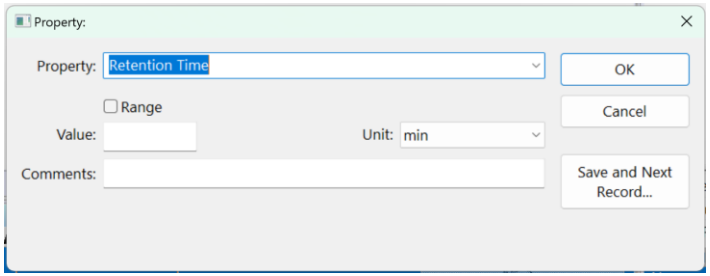
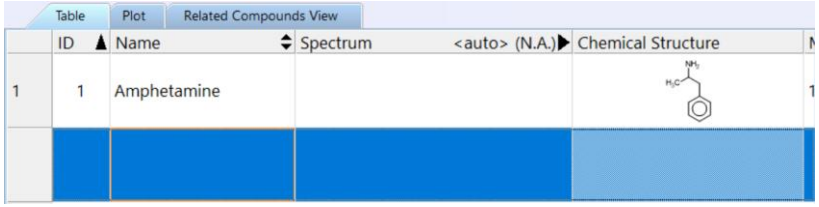
Example: Create a User Database for Accurate Mass Searching

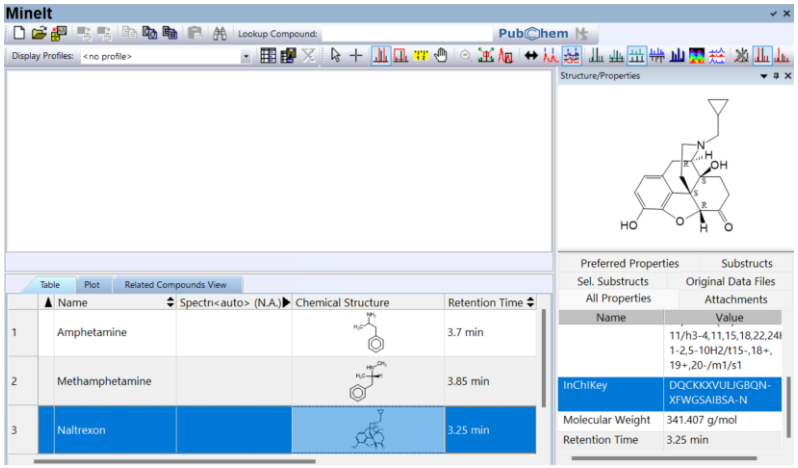
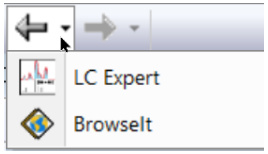
This section describes how to prepare the user database that is used for accurate mass searching, such as the sample file in the previous section.

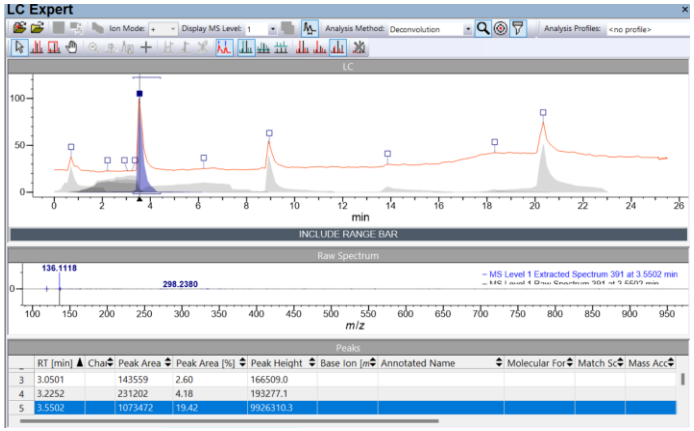
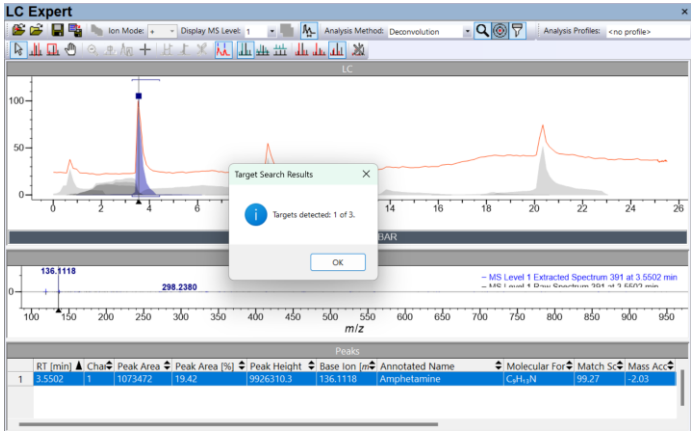
	Action	Result
36	To execute accurate mass searching within a **Chromatogram**, a user database with compounds is required for searching. Begin by opening the **MineIt** application (), typically found in the **Data** toolbox.	The **MineIt** application is displayed: 
37	Create a user database: Select **Database > New**.	The **New Database Creation** dialog window is launched. 

	Action	Result
38	Fill in the required database information: • **Database File** save location. Click **Browse** to select the location. • **Database Name**. • **Database Abbreviation** (3 letters). Click **OK** to proceed.	A blank database is opened in **MineIt**: 
39	Double click on the **Name** cell in the first row of the **Table**.	The **Property** dialog window is launched.
40	Enter “Amphetamine” in the proceeding popup. Press **OK** to save.	Amphetamine is displayed as the **Name** property in the first record (ID 1). 

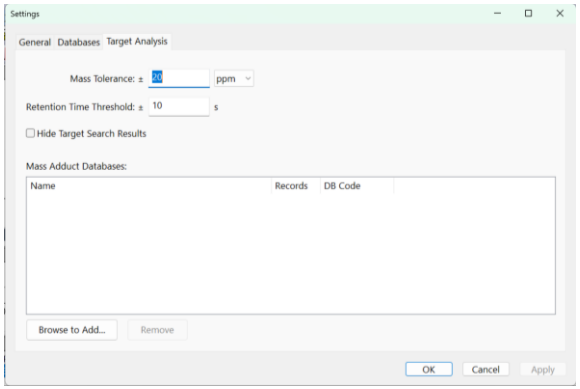
	Action	Result
41	Double click in the **Structure/Properties** window where it reads **"Double-click to edit structure in ChemWindow"**.	**ChemWindow** is launched as a popout window: 
42	Copy the given SMILES string to clipboard: c1(ccccc1)CC(N)C Then select **Edit > Paste Special** and select **SMILES.**	The struture for amphetamine appears in **ChemWindow**. 
43	Click **Save** to proceed.	The structure is displayed in the **MineIt** user database: 

	Action	Result
44	To add the compound's expected **Retention Time**, navigate to the **Structure/Properties Table** and select **Add**. Type **Retention Time** in the **Property** dialog window.	The **Retention Time** property value appears in the dialog window: 
45	Enter the value "3.7" in the **Property dialog window** in the cell next to **Value**. Retain the default units of "min". Select **OK** to proceed.	The **Retention Time** is added to the record.  *Note:* **Retention Time** is not required for the database file. If a value for **Retention Time** is not provided, then the full **Chromatogram** will be scanned for the compound.
46	Add more compounds to the database. Begin by clicking on the next row in the **Table**.	Upon clicking on the next row in the **Table**, it becomes highlighted to depict that it is active: 

	Action	Result
47	Add the following compounds into the user database by repeating steps 39-46. **Name**: Methamphetamine **SMILES**: c1cccc(c1)C[C@H](C)NC **Retention Time**: 3.85 min **Name**: Naltrexone **SMILES:** *Provided in the Result cell* **Retention Time**: 3.25 min	**Naltrexone SMILES:** O[C@]12[C@@]3(N(CC[C@@]11C4=C(C(=CC=C4C3)O)O[C@]1(C(CC2)=O)[H])CC1CC1)[H] Methamphetamine and Naltrexon are added to the database.  This user database of in-house compounds can be used for Exact Mass searching in **LC Expert**.
48	Mouse over to the **Previous Application** icon ( ) and click on the down button ( ). Choose **LC Expert** to return to the selected application.  	**LC Expert** application is opened.

	Action	Result
49	Using the open raw file icon ( ) located in the **Standard Toolbar** of **LC Expert**, navigate to **"C:\Users\Public\Public Documents\Wiley\KnowItAll\Samples\LC-MS\"**. Open *one* of the raw **Chromatogram** files: • **Compound_1_POS.mzML** • **Compound_2_POS.mzML** • **Compound_3_POS.mzML**	The selected **Chromatogram** opens in **LC Expert** (*e.g.*, Compound_1_POS.mzML): 
50	As executed in Step 29, click on the **Targets** icon ( ) or choose **Analysis > Targeted Analysis**. Navigate to and select the structure database file saved in Step 38.	One target from the database is found in the chromatogram:  • Amphetamine is detected in "Compound_1_POS.mzML". • Methamphetamine is detected in "Compound_2_POS.mzML". • Naltrexone is detected in "Compound_3_POS.mzML".

	Action	Result

51	*Note:* User settings for **Targeted Analysis** can be updated by selecting **File > Settings** and choosing the **Target Analysis** tab.	The available **Targeted Analysis** settings are:
		• **Mass Tolerance** for accurate mass deviation tolerance.
		• **Retention Time Threshold** for the retention time tolerance in seconds.
		• **Hide Target Search Results** checkbox prevents the **Targets Search Results** popup from appearing when selected.
		• **Mass Adducts Databases** allows the user to import additional adducts in an sdbx file, to be used in the accurate mass search.
		  ○ By default, only [M+H] or [M-H] adducts are scanned for in the **Chromatogram.**
		  ○ To add additional adducts, click the **Browse to Add** button and navigate to "Additional Adducts.sdbx" found in the LC-MS samples folder (See Step 52 for more information).
		 

52	*Note:* Users can create their own adduct libraries by modifying the sample sdbx file **"Additional Adducts.sdbx"** from the LC-MS samples folder (**"C:\Users\Public\Public Documents\Wiley\KnowItAll\Samples\LC-MS\"**), or by creating their own database file following the standards provided in this sample file.	To prepare an adduct library, the following information is required:
		• **Name:** used for adduct labels.
		• **Formula:** used to calculated the isotopic adduct radio.
		  ○ **KnowItAll** has been designed to recognize adduct losses by incorporation of a subtraction (-) symbol, *e.g.,* adduct [M-H] is be depicted as -H, and adduct [M+Cl-H] is depicted as Cl-H.
		• **Selected Ion Charge:** gives the ion charge and polarity where the adduct should be scanned.
		  ○ *E.g.,* The adducts [M-H] and [M+Cl-H] should be -1 and -2 correspondingly.
		  ○ Positive ions do not need a plus (+) symbol because the adduct is assumed to be positive (unless specified with a – symbol, denoting a negative adduct).

MSforID Searching

Introduction to MSforID Searching

The many challenges for preparing tandem MS search libraries and algorithms for unknown compound identification are well known and documented. Nonetheless, search tools and databases remain a critical part of the tandem MS workflow. The **MSforID** search method was designed to address these challenges, such as demonstrating a robustness against instrumental variability when searching quality databases and a high tolerance to variability in peak fragmentation patterns (*i.e.,* between the correct database match versus its experimental spectrum).[3] **MSforID** was positively evaluated using different instruments (*e.g.,* QqTOF, QqLIT, QqQ, LIT, LIT-FTICR and QTRAP) by different manufacturers and in different laboratories.[2-4]

The approach for **MSforID** searching is to compare the search query against a library of compounds where multiple CID spectra exist for each compound record. The compound records contain multiple spectra measured at different collision energies creating the series of spectra for the compound. The **MSforID** search algorithm then compares the query spectra to the *series* of CID spectra for the compound (Figure A). This is dissimilar to typical databases search algorithms that compare the query to a *single* spectrum per match (Figure B).

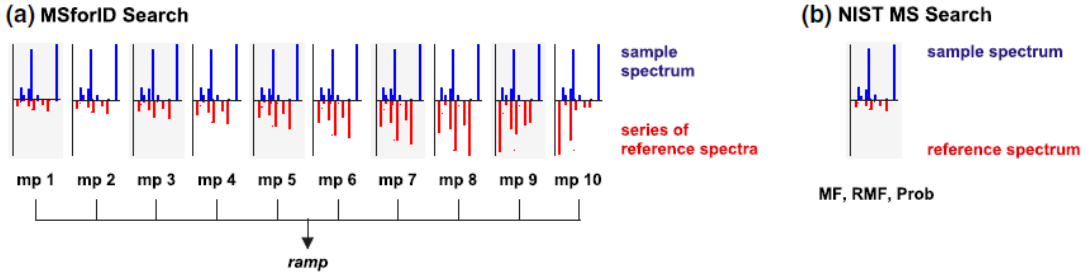


Figure A) MSforID and **B)** NIST MS Search identify search methods. (*Reprinted from Ref. 3*)

The MSforID Algorithm

The **MSforID** algorithm measures the average similarity of a query spectrum to the series of compound reference spectra. It is a probability-base matching algorithm that analyzes:

- The mass deviation for the precursor ion between the query spectrum and the database compound record.
- The number of matching fragments between the query and the database spectrum.
- The mass deviations and intensity differences for matching fragments.

For each search result, the algorithm calculates the **Average Match Probability (AMP)** for the compound's database record that contains the series of spectra. The **Relative Average Match Probability (RAMP)** is subsequently calculated, which is the normalized **AMP** value compared to the search results for the specific query (*i.e.*, from 0-100).[3] Search results are presented in **KnowItAll**'s **MineIt** by descending **RAMP** values, and the highest **RAMP** value is considered the best match. A **RAMP** value of >40.0 is considered a very good match score.[5]

Using the MSforID Search Tool in KnowItAll

Three search methods are available for **MSforID** searching in **KnowItAll**'s **SearchIt** application: (1) **Standard Search** (default), (2) **Composite Search**, and (3) **Direct Search**. The recommended search algorithm is the **Standard Search**, which applies the main published algorithm.[3] The **Standard Search** compares all spectra in the database record to the query spectrum (as in Figure A) to compute the **RAMP**. Differently, the **Composite Search** compares the single averaged spectrum for all spectra in the database record to the query spectrum using an adapted version of the **MSforID** algorithm. The averaged spectrum is calculated in real-time during the search, and the **Composite Search** can be faster when using very large databases. The **Direct Search** is a revised edition of the **MSforID** algorithm that aims to remove false positives from the hitlist.

Preparing In-House MSforID Libraries

The "Wiley Registry of Tandem Mass Spectral Data – MS for ID" database contains highly curated spectra for use with **MSforID** searching in **SearchIt**. To prepare user libraries in-house that are highly curated for accurate **MSforID** searching, the **MSforID** database standards[1] are recommended:

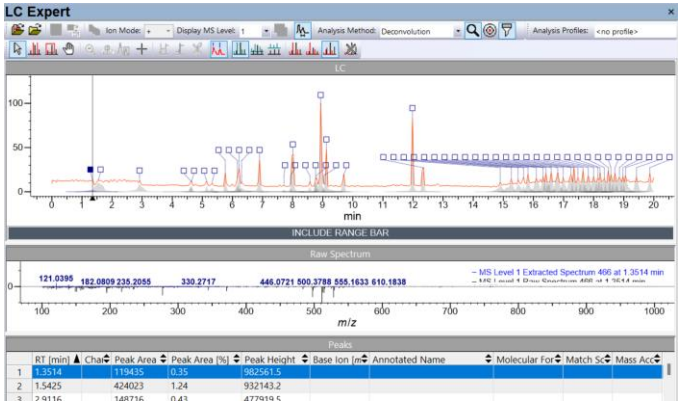
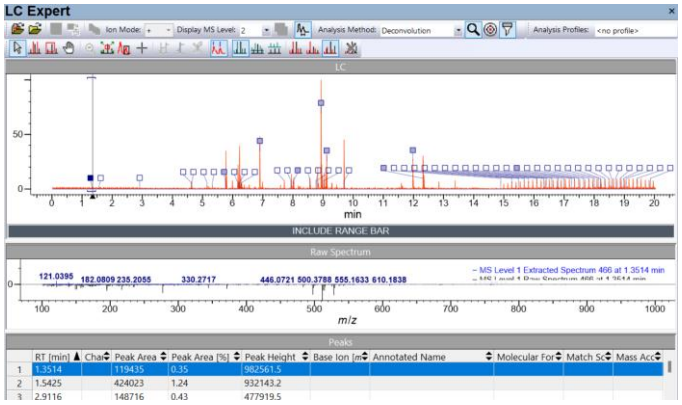
1. Measure mass spectra for the standard compounds at multiple collision energies (*e.g.,* from 5 to 50 eV).
2. Filter low abundant signals in the standard spectra (*e.g.,* less than 0.01%).
3. Prepare database records in MineIt using one precursor ion (*e.g.,* M+H). If compound spectra detected from different adducts are available for your library, separate these out into different records (*e.g.*, M+H spectra in one record and M+Na spectra in a second record).

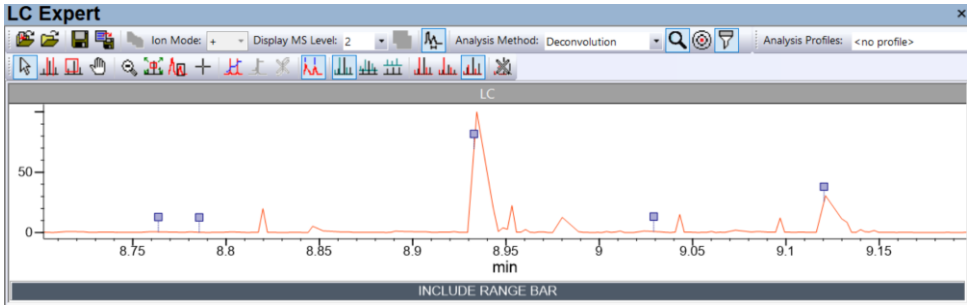
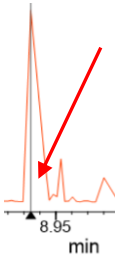
References & Additional Reading on MSforID

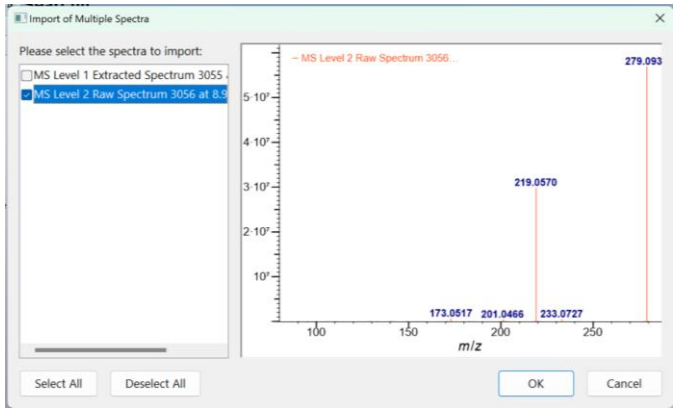
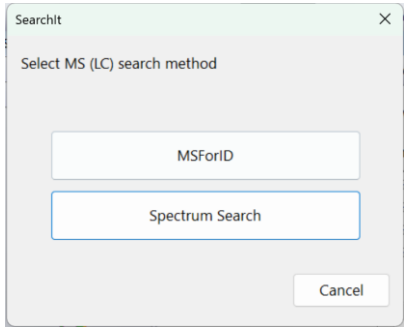
1. M. Pavlic, K. Libiseller, H. Oberacher. Combined use of ESI-QqTOF-MS and ESI-QqTOF-MS/MS with mass-spectral library search for qualitative analysis of drugs. *Anal. Bioanal. Chem.* **2006**, *386*, 62-82. doi: 10.1007/s00216-006-0634-8
2. H. Oberacher, M. Pavlic, K. Libiseller, B. Schubert, M. Sulyok, R. Schuhmacher, E. Csaszar, H. Köfeler. On the inter-instrument and the inter-laboratory transferability of a tandem mass spectral reference library: 1. Results of an Austrian multicenter study. *J. Mass Spectrom.* **2008**, *44*, 485-493. doi: 10.1002/jms.1545
3. H. Oberacher, M. Pavlic, K. Libiseller, B. Schubert, M. Sulyok, R. Schuhmacher, E. Csaszar, H. Köfeler. On the inter-instrument and the inter-laboratory transferability of a tandem mass spectral reference library: 2. Optimization and characterization of the search algorithm. *J. Mass Spectrom.* **2008**, *44*, 494-502. doi: 10.1002/jms.1525
4. H. Oberacher, W. Weinmann, S. Dresen. Quality evaluation of tandem mass spectral libraries. *Anal. Bioanal. Chem.* **2011**, *400,* 2641-2648. doi: 10.1007/s00216-010-4598-3
5. H. Oberacher, G. Whitley, B. Berger, W. Weinmann. Testing an alternative search algorithm for compound identification with the 'Wiley Registry of Tandem Spectral Data, MSforID'. *J. Mass Spectrom.* **2013**, *48*, 497-504. doi: 10.1002/jms.3185

Example: MSforID Search

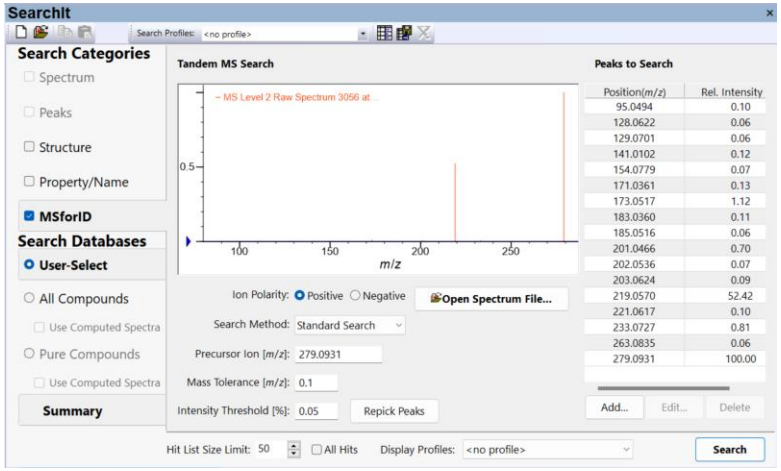
This section describes how to execute an MS^n library search using the MSforID search algorithm using the SearchIt application.

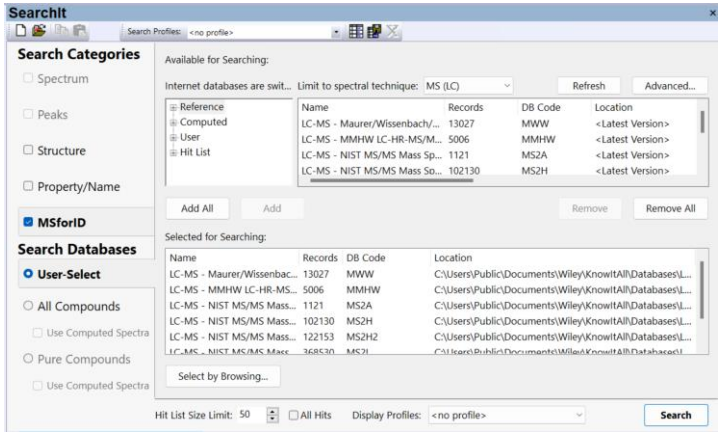
	Action	Result
1	Using the open raw file icon ( ) located on the **Standard Toolbar** in **LC Expert**, navigate to **"C:\Users\Public\Public Documents\Wiley\KnowItAll\Samples\LC-MS\"**. Open **"TESTMIX2_180504_MAS011_06 .mzXML"** (**C:\Users\Public\Public Documents\Wiley\KnowItAll\Samples\LC-MS**).	The **Chromatogram** opens in **LC Expert**: 
2	Use the **Display MS Level** dropdown menu located on the **Standard Toolbar** to select "2".	By changing the value to "2" on the **Display MS Level** dropdown menu, the **Chromatogram** MS2 scans become visible. Use the MS Level panel to interchange the spectra display between MS1 and MSn levels: 

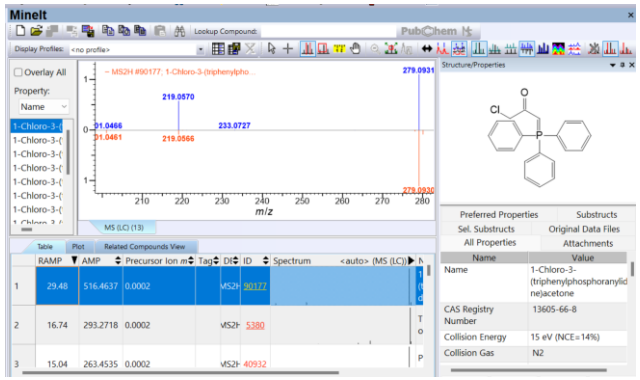
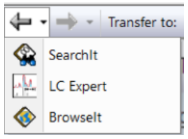
	Action	Result
3	Right click on the **Chromatogram** and select **Horizontal Zoom Mode**. Click and drag the left mouse button in the region around 9 min to zoom in on the **Peak** at 8.93 min. Left click on the **Chromatogram** and choose **Selection Mode** to deactivate the zoom cursor.	The **Chromatogram** is zoomed in on the **peak** at 8.93 min. 
4	Left click directly on the **chromatogram** at 8.93 min where the tall peak is.  *Note:* **Peak boxes** are reserved for componentized MS1 **peaks**, therefore do not select the **peak box** and instead click directly on the chromatogram.	The raw MS2 scan measured at 8.93 min is selected, and a black vertical line displays the location of the active MS2 scan (as in previous image). To verify that a MS2 spectrum is selected in the **Raw Spectrum** (black font) pane, verify the header information in the **Raw Spectrum**, which should read “MS Level 2 Raw Spectrum”. The **Extracted Spectrum** (in blue font) remains as an MS1 componentized spectrum. If the **Raw Spectrum** (black font) reads “MS Level 1 Raw Spectrum”, click again on the **chromatogram** to update the MS scan to an MS2 spectrum. 

	Action	Result
5	On the **Transfer to** bar (Transfer to:), select **SearchIt** (SearchIt).	The **Import of Multiple Spectra** dialog window appears. The “MS Level 2 Raw Spectrum” option is selected by default:  *Note:* This dialog is used to specify which MS spectral scan will be transferred using the **Transfer to** bar , *i.e.*, either the MS1 **Extracted Spectrum** or the MSn **Raw Spectrum**.
6	On the **Import of Multiple Spectra** dialog window, retain the default selection “MS Level 2 Raw Spectrum”. Select **OK** to continue.	The **SearchIt** import dialog window appears with 2 different search options: • “MSForID”, to open a new **MSforID** search tab in **SearchIt**. • “Spectrum Search”, to use an alternative search algorithm (*e.g.*, cosine, adaptive, *etc*.). 

	Action	Result

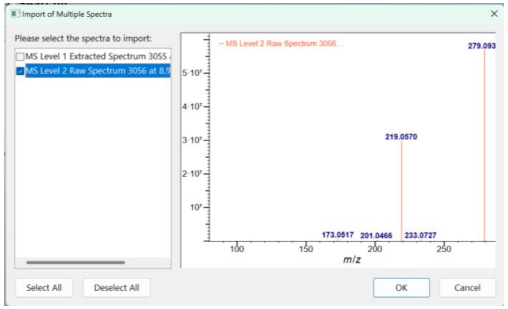
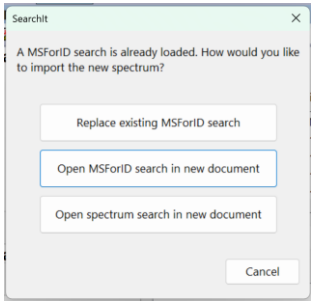
7	Select **MSforID** on the **SearchIt** dialog window.
	
	Note: The **Intensity Threshold (%)** can be increased or decreased. Select **Repick Peaks** button (Repick Peaks) to update the **Peaks to Search** list using the updated threshold. The triangle symbol ( ) reveals the minimum peak height.
		The MS2 raw spectrum opens in **SearchIt**'s **MSforID Search** window. The **MSforID Search** window prepopulates the information:
		• **Ion Polarity** which is the ion polarity information in the raw chromatogram file, if included in the raw file.
		  ○ If this information is not included in the raw file, then **positive** will be selected by default, and this could be updated to **negative** by selecting the opposite radio button.
		• **Search Method** is the specific **MSforID Algorithm** that will be applied in the search. The last use search will be selected as the menu option.
		  ○ **Standard Search** (default)
		  ○ **Composite Search**
		  ○ **Direct Search**
		• **Precursor Ion (m/z)** is the MS2 scan's precursor ion information, if included in the raw file.
		• **Mass Tolerance (m/z)** parameter sets the tolerance for MS spectrum peak m/z deviations.
		• **Intensity Threshold (%)** parameter sets the minimum peak height for the MS spectrum peaks.
		 
		Note: A popup warning will display on the **Warning** dialog window if the raw spectrum is not detected to be an MS2 spectrum. This may be because the raw file does not contain MS Level information (*e.g.*, such as imported .jdx files), or it is the wrong MS Level (*e.g.,* MS Level = 1). Click **Confirm** to bypass the warning and import the MS spectrum into the window or **Cancel** to stop the process.

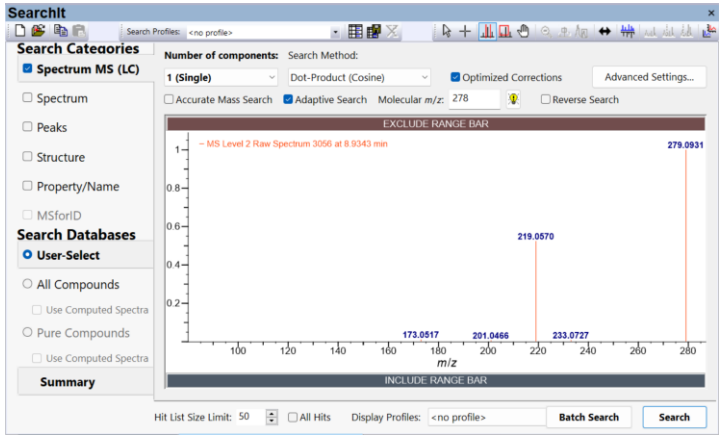
	Action	Result
8	In the **MSforID Search** window, change the **Mass Tolerance** value to “0.01" m/z.	The **Mass Tolerance** is decreased: Mass Tolerance [*m/z*]: 0.01 *Note:* The **Mass Tolerance** plays an impactful role in the calculation and final search results.
9	In the **Search Databases** tab, click on **User-Select** option. Define databases for searching using the **User-Select** databases tab by clicking **Add** for desired LC-MS databases.	The databases selection dialog window is displayed. Available LC-MS databases depend on the specific user license: 

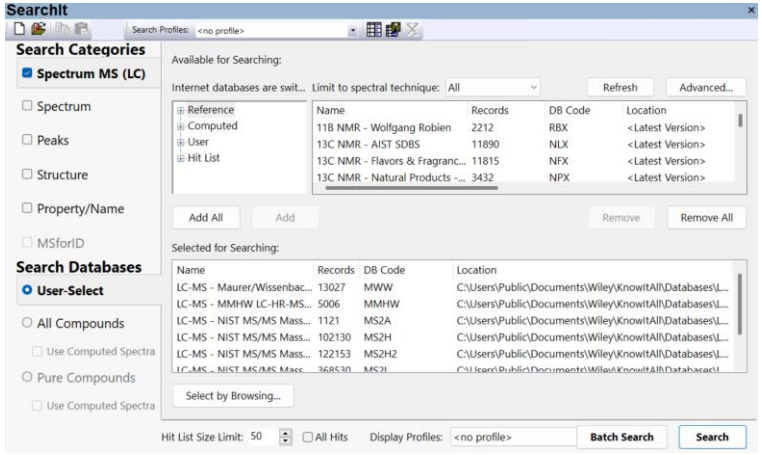
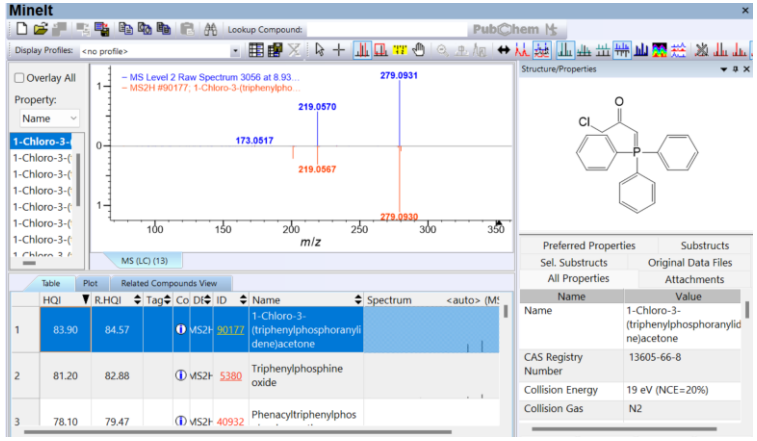
	Action	Result
10	Select **Search** button (Search) to execute the search.	The best match for the spectrum query is displayed in **MineIt**. The **Table** display the columns: • **AMP (Average Match Probability),** which is the average probability that the reference compound record could be the query spectrum. • **RAMP (Relative Average Match Probability)**, which is the **AMP** value relative to the total search results. ○ By default, the search results will be organized by decreasing **RAMP** value. • **Precursor Ion m/z Difference** which is the m/z difference between the query's precursor ion and the database record.  *Note:* Each row in the search results **Table** represents a match for the compound record's *series* of spectra, and each compound record can only be returned once as a match result. In the **Table**, each compound record's MS spectra will retain the order in which they exist in the database, *i.e.,* the first spectrum in the series will always be displayed first in the search results.
11	Navigate to the **Previous Application** icon and click on the down button ( ). Choose **LC Expert** to return to the selected application. 	**LC Expert** application is opened.

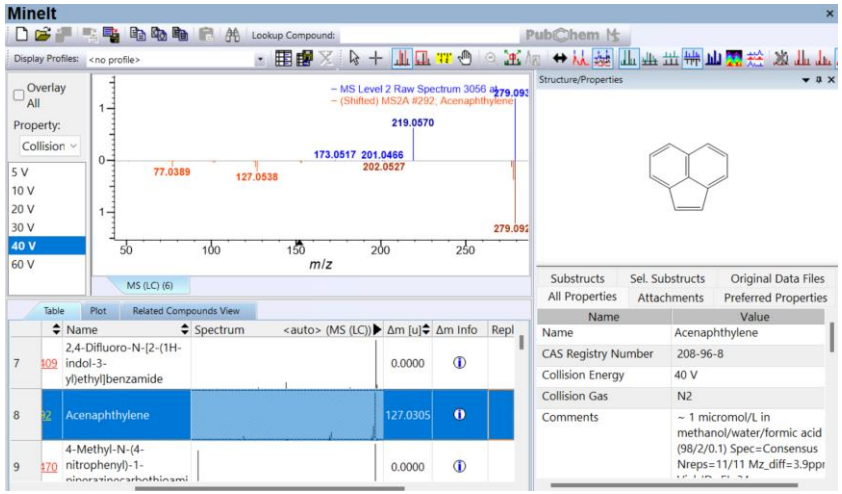
Example: Spectrum MSn Searches using SearchIt

This section describes how to execute an MS^n library search using spectral search algorithms in SearchIt application.

	Action	Result
12	Continue with the **Chromatogram** from the last section. In **LC Expert**, with the scan from Step 4 still selected, use the **Transfer to** bar (Transfer to:) to send the spectrum to **SearchIt** (SearchIt).	The **Import of Multiple Spectra** dialog window appears. The "MS Level 2 Raw Spectrum" option is selected by default.  *Note:* This dialog is used to denote which MS spectral scan should be transferred using the **Transfer to** bar , *i.e.*, either the MS1 **Extracted Spectrum** or the MSn **Raw Spectrum**.
13	On the **Import of Multiple Spectra** dialog window, retain the default selection "MS Level 2 Raw Spectrum". Select **OK** to continue.	The **SearchIt** import dialog window appears with 3 different search options: • "Replace existing MSforID search", to override an existing search in **SearchIt**. • "Open MSforID in new document", to open a new search window in **SearchIt**. • "Open spectrum search in new document", to use an alternative search algorithm (*e.g.*, cosine, adaptive, etc.). 

	Action	Result
14	Select “Open spectrum search in new document”.	The MS2 raw spectrum opens in **SearchIt**’s **Spectrum Search** dialog window. **SearchIt** recognizes that the spectrum search type is LC-MS, denoted by “Spectrum MS (LC)”.  For further control of the spectrum search: • Click **Advanced Settings** to launch a dialog window with more search controls (*e.g.,* **Instrument Resolution** for **Accurate Mass Search**). • **Include Range Bar** and **Exclude Range Bar** allow for including/excluding specified areas on the spectrum.
15	In the **Spectrum MS (LC)** window, select the checkbox next to **Accurate Mass Search**.	**Accurate Mass Search** is selected:  **KnowItAll Accurate Mass Search** retains high resolution information included in the spectrum query and database record when performing the search.

	Action	Result
16	Click on **User-Select** option under **Search Databases** tab. Use the **User-Select** databases tab to choose databases for searching. Click **Add** to select an LC-MS database for searching. *Note:* Available LC-MS databases depend on the specific user license.	The databases selection dialog window is displayed: 
17	Click **Search** to execute.	The best match for the spectrum query is displayed in **MineIt**: 

	Action	Result
18	Use the **Previous Application** arrow () to return to **SearchIt**.	**SearchIt** application is opened with the previous query loaded.
19	Click on **Spectrum MS (LC)** to modify the spectrum search settings. Select the **Adaptive Search** checkbox and enter “279.0931” as the value for **Molecular m/z** if not already detected.	The **Adaptive Search** is selected.  **Adaptive Search** method allows for scanning the spectrum peaks for available functional groups or molecular replacements between the query spectrum and the database spectrum, extending the database library to similar compounds that are not in the available spectral space. *Note:* **Adaptive Search** method is available for low resolution data by deselecting the **Accurate Mass Search** checkbox.
20	Click **Search** to execute.	The best match for the spectral query is displayed in **MineIt**:  For the **Adaptive Search** results, regarding the spectrum query and the database record: • **Δm** column gives the difference in compound mass between the query and the database compound. • **Δm info** column contains a selectable info icon () that informs on the peak shifts that occurred to create the query result. • **Replacement** column gives the group replacement if known.