

3050 Spruce Street, St. Louis, MO 63103 USA
Tel: (800) 521-8956 (314) 771-5765 Fax: (800) 325-5052 (314) 771-5757
email: techservice@sial.com sigma-aldrich.com

Product Information

MLSDiscovery™ Software Manual

Description

MLSDiscovery[™] is a software tool designed to streamline and simplify the tedious process of building authentic standard libraries for mass spectrometry. The program is tailored to work with the Mass Spectrometry Metabolite Library (MSMLS[™]), Large Scale Metabolite Library (LSMLS[™]), and other metabolite libraries distributed by IROA Technologies LLC.

The layout of the software user interface is intuitive and the first section of this manual will concentrate on the recommended workflow. This is followed by detailed descriptions of user interface to help clarify the possible questions and explain options not previously mentioned. The final section discusses the use of the library and software for sample analysis.

MLSDiscovery user guide video can be used for additional instructions for operating the program. The video can be accessed using this link:

MLSDiscovery User Guide or by entering this address: http://www.anymeeting.com/wyhfnkefitp/E151DC878346

Procedures

Definitions:

<u>Note</u>: The definitions below are intended to avoid misunderstanding and ambiguity within the scope of this manual only.

Experimental sample – solution of one or more compounds intended to be analyzed within a single LC-MS/LC-MSMS run; each sample may be analyzed by more than one chromatographic/MS method combination. Currently MLSDiscovery does not support mixed polarity files.

Multiplex plate – individual plate from the Mass Spectrometry Metabolite Library. The library is distributed in 96 well plate format together with a plate map specifying the location (plate, number, row, and column) of each single authentic standard. Compound identifiers and descriptors such as CAS numbers are provided. Please refer to the individual product information sheets for further information.

Note: The plate map contains descriptors and represents information gathered from multiple databases and therefore may contain errors. We suggest that the information provided is carefully reviewed. To help build a better database, please report any discrepancies.

Experiment design – library construction involves analysis of multiple mixtures of authentic standards, possibly by more than one LC-MS/LC-MSMS method. Experiment design is a definition of relationship between experimental samples, raw data files, and analytical methods. It determines which samples contain which authentic standards and which raw data files correspond to the analysis of specific samples by a specific method.

Starting the Program for the First Time

The requirements of the program are:

- The most recent version of Java 8 must be installed and callable.
- 2. The computer should have at least 8 GB of RAM.
- 3. The computer is running Windows 7 or higher.
- 4. You will need administrative privileges on the computer you wish to install it on.

<u>Notes</u>: Installation of the software requires administrative privileges on the computer you wish to install it on. The installer will by default install and register the software, and leave a shortcut on the desktop.

IMPORTANT! Before you install and run the program read the following!

Due to the security policy in Windows, when you install the program in the default location (e.g., C:\Program files) there are restrictions on what you can do inside this program directory, which includes the modification of database. There are two ways around it -1) install the program in a different location (e.g., in the user's personal directory) or 2) change the permissions on the folder and all the files inside it when you install in the default location (you must have administrator's permissions on the computer to do it). To follow are instructions to change permissions on the folder.

To execute this software package, there are three links/folders that you need to be aware of:

- 1. The first is the shortcut to the MLSDiscovery software itself that was placed on the desktop during the installation of the software.
- The second is a folder, which you will create and where the project is stored so you may return to it at any time. As you will learn, a complete copy of the raw data files is saved within the project, as are the preferences and libraries you have extracted.
- The third is a folder where the data are stored. All raw data files should be exported as **Centroided** data files. These are smaller and will process faster and more accurately.

After the program is downloaded, an Installer Wizard will pop up on your computer screen. Installation and use of the program require acceptance of the license agreement. Upon reading and accepting the terms of the license agreement, the wizard will bring you through a series of screens for installation, including the selection of a destination folder. Once installation is complete by clicking the "Finish" button, the user has the option to start the MLSDiscovery program.

Open the program. If this is the first time the program is opened, you will be asked to set the default working directory. This is a folder where the project is stored. A complete copy of the project, including raw data files, project preferences, and libraries that are generated will be saved in this directory. Please consider the free disk space on your computer carefully when making this decision since projects may take considerable space. Projects may be created/copied/moved to any directory of your choice and the default directory location may be changed at any time by editing program preferences.

You will also require a separate folder where the LC-MS and MSMS data files for the project are stored. MLSDiscovery software supports mzXML and mzML open formats for raw data files. Raw data should be **centroided**, using profile MS data may produce incorrect results and will dramatically increase memory and processor requirements.

Standard Workflow

The standard software workflow for creating a new compound library includes the following steps:

- Creating new project
- · Creating experimental samples
- Managing compound libraries
- Defining composition of each experimental sample (which authentic standards it contains)
- Importing raw data for one or more LC-MS analysis methods
- Associating samples with corresponding data files
- Performing compound search
- Curating search results
- Exporting library data for use in subsequent experiments
- Merging libraries (if more than one library has been created)

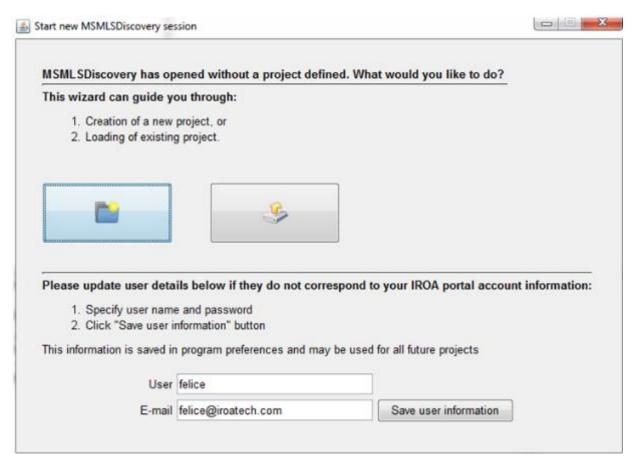
User may interrupt the work at any stage and return to analysis at a later date.

<u>Note</u>: Complete copies of the raw data files are saved within the project, as are the preferences and libraries you have extracted.

Creating a New Library Generation Project

At program startup the user is presented a window, which allows the creation of a new project or to open an existing one (see Figure 1). Specify the name and email address. Once loaded the program will store this information and you will not be required to enter it again.

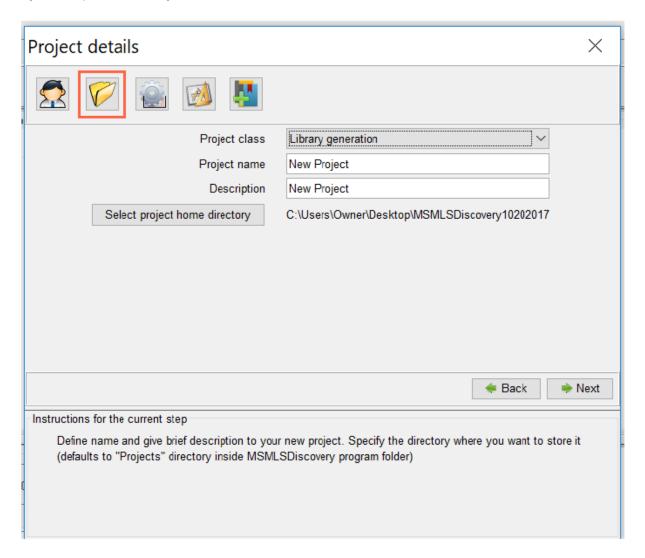
Figure 1. Startup Task Chooser



To initiate a new project, click "Create new project" button (), which will bring up the first page of a New Project Wizard, project definition (see Figure 2).

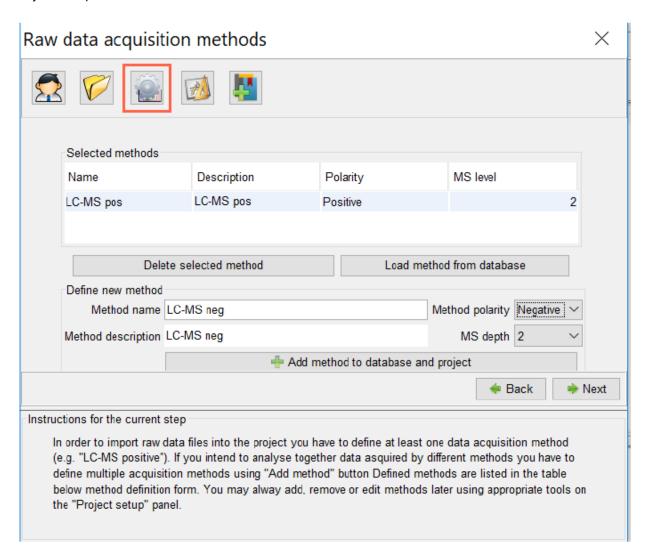
Note: Clicking "Open existing project" button () will bring up file chooser dialogue. By default, it will open the default working directory where the projects are saved. Navigate to the folder of the project you want to open, select the project file (one with ".cfproj" extension) and click "Open" button to load the project.

Figure 2. Project Setup Wizard - Project Definition



Under Project class, select "Library generation" from the dropdown menu. Specify the name and description of the experiment. As an example, if for this project you are building a library for Plate 1, you might name the project "Plate 1". Choose a project home directory for where all the project information will be stored and click the "Next" button to move to the method definition screen (see Figure 3).

Figure 3. Project Setup - Method Definition

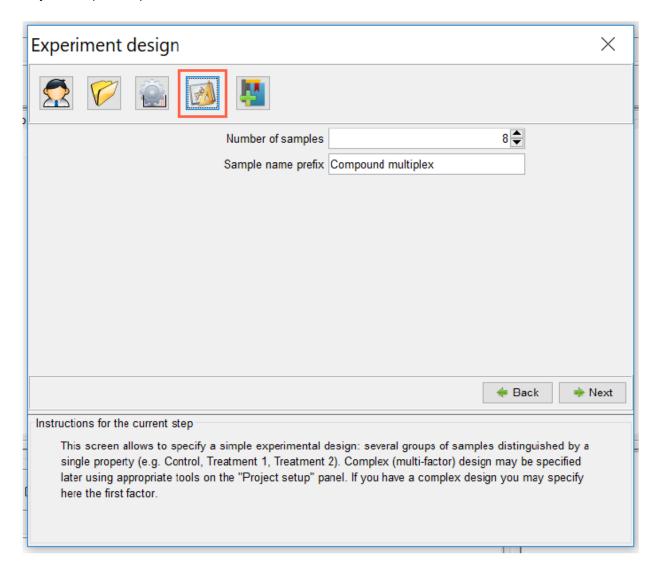


Specify name, description, polarity, and MS depth for each analysis method you have used to collect raw data. (In this project example, you would add all methods used to collect raw data for Plate 1; for example, you might name the method LC-MS pos for data collected using positive polarity for Plate 1). There are two MS depth levels: select level 1 for MS and level 2 for MSMS data. Use the "Add method" button to add the specified method to the project (it will show up in the "Defined methods" table). You may remove a method from the list using "Delete selected method button". This may be useful if you made a mistake in method definition. In addition to manually defining methods you may use the "Load method from database" button to add previously defined and stored methods.

You can edit method information, add new, and delete existing methods later in the program as well. When you have finished adding all methods, click "Next" to move to sample definition screen (see Figure 4).

Creating Experimental Samples

Figure 4. Project Setup - Sample Definition



For the sake of simplicity, sample definition in wizard is limited to specifying the sample name prefix and number of multiplex samples. In this example, let's say you have pooled each row in Plate 1, here you would specify 8 samples representing the 8 rows (A-H) in Plate 1 you will examine. Once the number of samples are added, click the "Next" button.

Select Standard Compound Library Version for the Project

After the definition of the experimental design, the "New Project" wizard (\times) will ask the User to specify the

Metabolite Library of Standards (MLS) that was used to create the data files () i.e. MSMLS, FAMLS™, OAMLS™, BACSMLS™ etc. From the "Select standard library" pull-down menu, select the correct library. If you are not sure which library version you have, either contact your supplier or use your plate map (provided to you at time of purchase) to verify compounds in the "Selected sample composition" window (see **Error! Reference source not found.**B).

After clicking "Finish and close wizard" button the new project with specified parameters will be created and loaded into the program.

Note: There is also a button () on the Multiplex designer panel (see below) that can be used to modify/switch to the Metabolite Library of Standards used to create the data files.

Multiplex Designer Panel

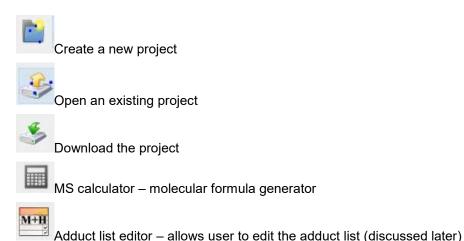
The Multiplex designer panel will now pop up (see **Error! Reference source not found.**5). This is where you will finish defining the experimental design. The samples you have created will be listed in the "Multiplex samples" table (see **Error! Reference source not found.**B).

Multiplex designer overview toolbar

Actions associated with buttons on the toolbar of the Multiplex designer panel (see **Error! Reference source not found.**5) are described below.



Multiplex designer toolbar



Library database editor – once the libraries are generated this is used to manage the compound libraries. This tool allows the user to edit the library data, including name, molecular formula, descriptors, and retention time.



Show acquisition methods manager



Open "Program preferences" tool used to set search parameters



Unassigned to date (space holder)



Save project and exit button – the project can be left at any time and will be saved

Second row of icons:



Edit name, project description, and user information



Switch MLS library – specify which MLS library was used to create the data files



Add multiple samples



Add a single sample



Delete selected samples



Find library compounds in the samples - this is the button to run the program after set-up is completed

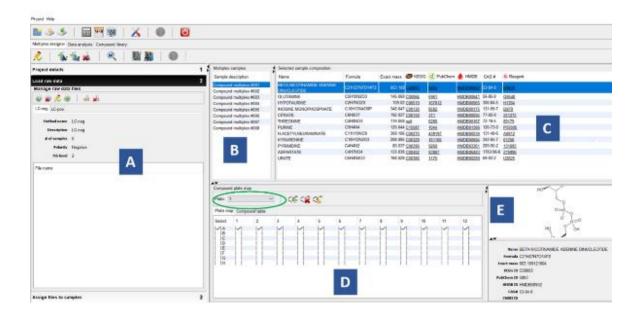
<u>Note</u>: After making changes to any preferences, make sure you click on the modified the program will save the new settings.



icon to apply the changes. Once

You are now in the "Multiplex designer panel" (see **Error! Reference source not found.**5). Now that the program knows which library was used to generate the data files, assign the correct compounds to each of the multiplex samples. To associate a sample with the authentic standards it contains, first select the correct Plate number by using the "Plate selector" dropdown menu (circled in green, see **Error! Reference source not found.**D). Then highlight a sample in the "Multiplex sample" table (see **Error! Reference source not found.**B).

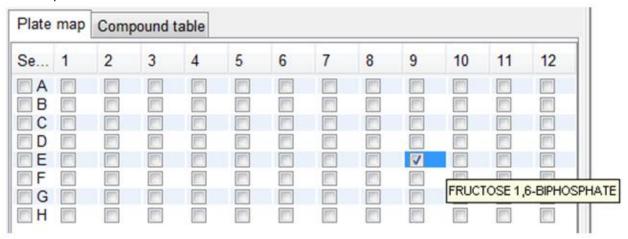
Figure 5.Multiplex Designer Panel Overview



There are two ways to select compounds; the "Plate map table" (see Error! Reference source not found.) and/or the "Compound table" (see Error! Reference source not found.). Both table tabs are located in the lower middle portion of the "Multiplex designer window" beneath the green circle (see Error! Reference source not found.D), and the user may switch between them by clicking on the appropriate tab.

The "Plate map table" (see **Error! Reference source not found.**D) mimics the layout of the 96 well plate used to supply MLS authentic compounds (MSMLS, LSMLS, or BACSMLS etc.). Individual compounds may be selected by checking corresponding boxes, or an entire row may be selected by checking the individual (A-H) boxes in the first column. The tooltip will display the compound name when you point the mouse to a specific table cell; clicking and selecting the cell will bring up the compound structure and other information in the lower right portion of the window (**Error! Reference source not found.**E).

Figure 6.
Plate Map Table



The "Compound table" (see **Error! Reference source not found.**) allows only single compound selection, clicking on each compound row also brings up the selected compound structure and details in the window of **Error! Reference source not found.**E.

Figure 7. Plate Compound Table

Row	Column	Name	Formula	Exact mass
A		1 NAD	C21H27N7O14P2	663.109
A		2 L-GLUTAMINE	C5H10N2O3	146.069
A		3 HYPOTAURINE	C2H7NO2S	109.02
A		4 INOSINE 5'-PHOSPHATE	C10H13N4O8P	348.047
A		5 CITRATE	C6H8O7	192.027
A		6 L-THREONINE	C4H9NO3	119.058
A		7 PURINE	C5H4N4	120.044
A		8 N-ACETYLNEURAMINATE	C11H19NO9	309.106
A		9 L-KYNURENINE	C10H12N2O3	208.085
Α		10 PYRIMIDINE	C4H4N2	80.037
A		11 D-ASPARTATE	C4H7NO4	133.038
A		12 URATE	C5H4N4O3	168.028
В		1 CYTIDINE	C9H13N3O5	243.086
В		2 L-SERINE	C3H7NO3	105.043
0		3 LOYOTENIE	0211781020	404.00

After selecting the desired compound(s), click the "Add selected compounds to sample" button () on the toolbar at the top of the window and the selected compound(s) will be linked to the "Compound multiplex" sample (see Error! Reference source not found.B) and the "Selected sample composition" table (see Error! Reference source not found.C) will be populated with the selected compound(s). To switch between different multiplex plates, use the "Plate selector" dropdown menu (circled in green, see Error! Reference source not found.D).

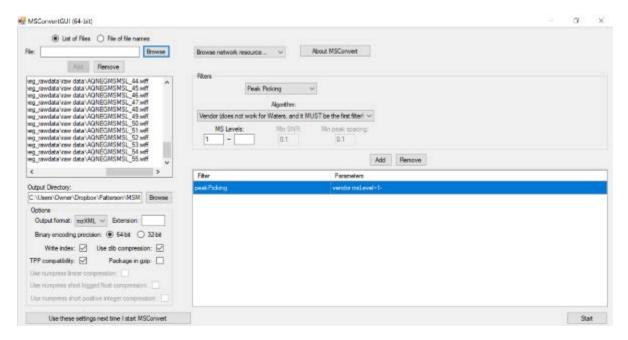
Now is a good time to verify that you are using the correct library. Compare the Excel plate map (plate, row, and column information) provided with your library, to the compounds listed in window "C" to make sure they are the same and the correct library has been selected in program preferences. If you need to change the compounds that were assigned to the samples this can be easily done. Compounds can be individually removed by selecting the compound

in window "C" and clicking the "Delete selected sample" button (), or the entire list of compounds for a highlighted sample may be removed by clicking the "Remove all compounds from sample" button ().

Importing Raw Data

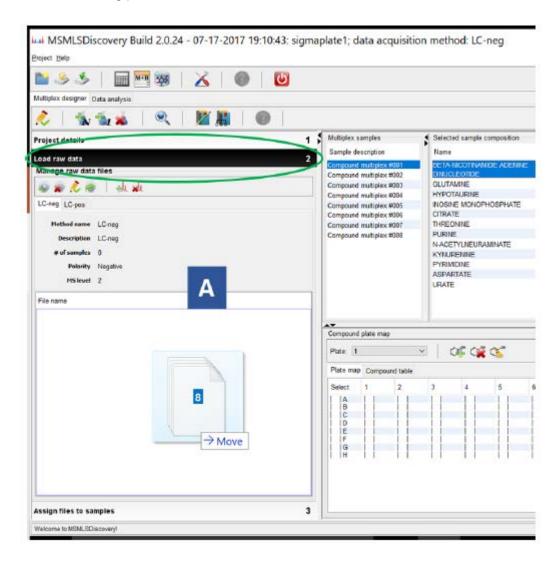
Prior to loading raw data files, it is very important to convert the data files to centroided mzXML. Most vendor files can be processed using ProteoWizard msconvert (open-sourced tool, Windows only, since it relies on instrument vendor libraries). A screenshot of the ProteoWizard MSConvert user interface with recommended settings for centroiding MS data is shown (see **Error! Reference source not found.**). Click the "Add" button so the vendor peak picker gets added to the filters.

Figure 8.MSConvert user interface



To import raw data, click the "Load raw data" bar of the accordion panel (see **Error! Reference source not found.**A) to open the data import panel (see **Error! Reference source not found.**). If you have more than one analytical method defined for the current project, select the corresponding tab of the panel and load the files by either "dragging-and-dropping" into the "File name" portion of the table (as shown) or by clicking "Add raw data files" button () on the panel toolbar and navigating to the directory containing the project raw data and selecting the required data files.

Figure 9. Raw data loading panel

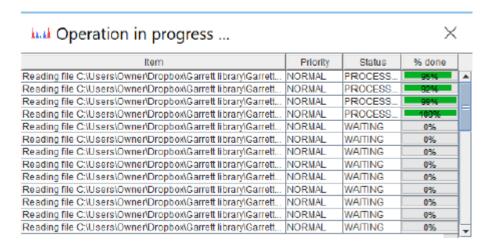


Individual raw data files may be removed from the project. To do this, highlight the file(s) you want to remove in the file table of the selected method tab and click button on the panel toolbar.

To activate the data set for analysis, click on the tab of the chosen data set and then click button on the panel toolbar. Green checkbox icon in the tab title will indicate that method was activated.

During file import the "Task monitor" screen is activated to show operation progress and blocks the main program window until the process is completed. Every operation of the software that requires noticeable time, such as loading raw data files, is run as an independent thread. (see **Error! Reference source not found.**).

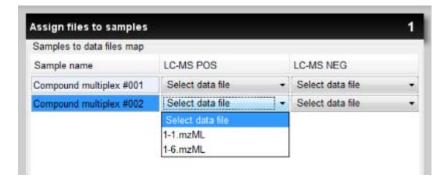
Figure 10.
Task Monitor Panel



Associating Samples with Corresponding Data Files

After importing the raw data files it is necessary to assign them to the corresponding samples to let the software know which compounds to look up in which data file. This is done by choosing the files from the dropdown menus in a "Sample to data files map" (see Figure 11). Click the "Assign files to samples" bar in the accordion plate (see Figure 5A) to display the mapping table.

Figure 11.Samples to Raw Data Files Mapping Table

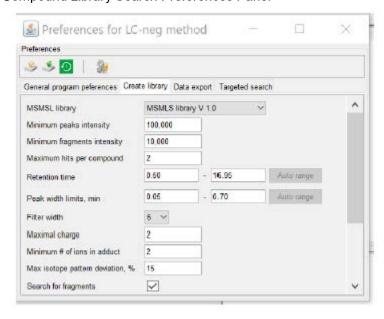


Specify Search Parameters

After file assignment is completed you are ready to specify search parameters and run the search. To examine and modify the search parameters open the "Preferences" window by clicking "Edit program preferences" button



Figure 12.
Compound Library Search Preferences Panel



Each parameter is described below, and recommended settings are provided.

- Minimum peak intensity this is the minimal intensity setting for the base peak.
 - This depends on the data, once the raw data are loaded it makes sense to browse through the scans and create an average MS of an empty area to estimate noise. One may estimate this setting from the total ion current (TIC) chromatogram or Base Peak Chromatogram (BPC). See "Generating chromatograms".
- Minimum fragments intensity this is the minimal intensity for fragments; default is set at 10%, that is 10% of the minimal intensity for the base peak.
 - o May be 10% of the minimal intensity for the base peak, rise higher to avoid junk.
- Minimum hits per compound the recommend setting is at least 2 or 3 when running library generation, as the same m/z may occur in the same sample more than once.
- Retention time setting is dependent on chromatography; we recommend setting the minimal value so that it will
 exclude the void volume.
- Peak width limits, min this is the minimal and maximal allowed width of chromatographic peak.
 - o minimal should allow the fit of 3-4 MS scans for the peak to be reliable.
- Filter width width of smoothing filter (Weighted Moving Average Filter) used during chromatographic peak extraction.
 - Specified as number of scans to take moving average across.
 - Smoothing does not influence integration results; it is used for a better definition of peak start and end in noisy data
- Minimum number of ions in adduct we recommend setting this to at least 2 to exclude junk peaks.
- Maximum isotope pattern deviation, % we suggest 10% for library generation.
- Maximal charge this is the largest charge to consider when looking for mass spectral features.
 - For a targeted search, this may be increased since with low intensity peaks, ratios may be distorted.
- Search for fragments fragments will be identified when this parameter is checked.

- Fragment lookup cutoff this is defined as the % intensity of the major adduct; we suggest setting at 10-25% and then adjust empirically. (Do not set too low as you could end up with a lot of small peaks).
- Fragments RT tolerance we suggest using the same value as set for minimal peak width (above); as fragments are expected to co-elute with major adducts.
 - Differences in RT are usually due to poor peak shape.
- Minimal Signal/Noise for fragments –we recommend setting to 2 or 3 initially and adjusting empirically as necessary.

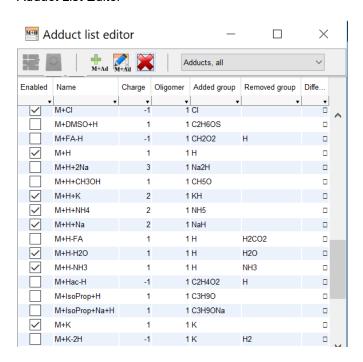
Specify Adducts to Search

The software maintains a database of adducts and mass modifications to help with the interpretation of the mass spectra. Lists may be edited and extended by the user. In the Multiplex designer panel (see **Error! Reference source**

not found.), click on the "Show adducts/modifications editor" button () to open the "Adduct list editor" (see Error! Reference source not found.). This editor allows the user to control which adducts to include in the search by

checking individual boxes in left column of the window. The "Create new adduct" (M+Ad) button allows the user to define a new adduct not listed in the table; the "Edit adduct" (M) button allows the user to edit a selected adduct; and the "Delete selected adduct" (M) button allows the removal of an Adduct.

Figure 13. Adduct List Editor

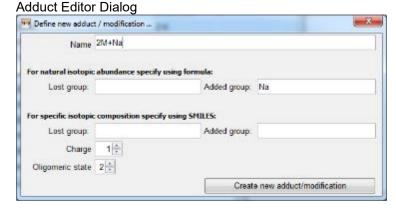


"Add" and "Edit" commands call the same dialog (see **Error! Reference source not found.**), but in case of "edit" it is pre-populated by the values for selected adduct or modification.

Lost and added groups should be specified as valid molecular formulas. If you need to use specific isotopic composition, lost and added groups should be specified using SMILES strings.

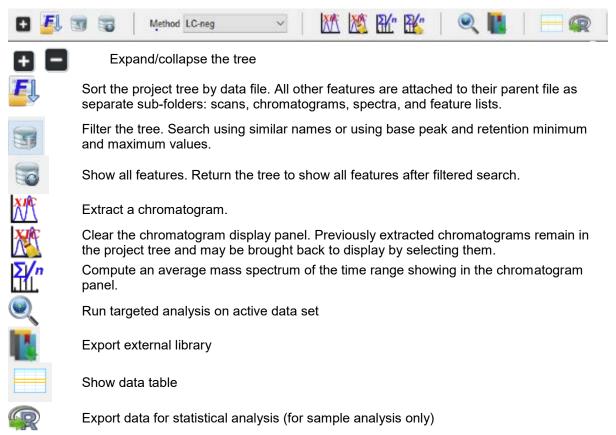
The dropdown menu on the adduct list editor toolbar allows user to display all adducts, adducts for specific polarity or mass modifications. To export complete list of adducts to text file, choose any option except "Mass differences". The latter option is used to view, edit, and export only mass modifications. Both adducts and mass modifications may also be imported into the program. Only entries with unique names (not already in the list) will be imported. Formats of the files for adducts and modifications are slightly different and may be examined by exporting the default lists supplied with the software.

Figure 14.



Raw Data Review

Once the data files are loaded they can be examined in detail.



Instructions for the use of the chromatogram tools displayed above are provided in the following section.

Generating Chromatograms

When examining LC-MS data it is often useful to visualize chromatographic profiles of specific masses to see if they form a chromatographic peak, examine the shape and magnitude of the peak and determine which masses co-elute and have similar peak shapes. Discovery provides the tools to create several types of chromatograms from raw MS data:

- Total Ion Chromatogram (TIC) is constructed by plotting the sum of areas of all the peaks in each scan against time.
- Base Peak Chromatogram (BPC) is constructed by plotting the area of the highest (base) peak in each scan against time.
- Extracted Ion Chromatogram (XIC) is constructed by plotting the area of the peaks having m/z within a narrow window against time.

To extract the chromatogram:

- 1. Call chromatogram dialog by clicking the the button on the raw data overview toolbar.
- 2. Use the "Plot type" selector to choose the type of chromatogram (TIC, BPC, or XIC).
- 3. If chosen, plot type is XIC define mass to extract (Extract M/Z), precision measurement unit (ppm or mDa), and extraction window width (M/Z window).
- 4. The Retention time window applies to all types of chromatograms. Use the "Auto range" button to set the time range from start to finish of the run.
- 5. Click "Extract chromatograms" to extract the specified type of chromatogram for raw files selected in the list on the left side of the dialog.

New chromatograms are displayed on the Chromatogram panel. Depending on the sorting options, chromatograms appear in the project tree in the "Chromatogram" folder in the root of the project or in the "Chromatogram" folder under each separate data file. ClusterFinder™ keeps track of extracted chromatograms and prevents duplicates from accumulating, so if chromatograms are extracted but the display does not update, then the data is already in the project and may be displayed by selecting the appropriate items in the project tree.

To clear the chromatogram panel, click the 🍱 icon.

To load the chromatogram in the viewer, click on its name in the project tree. If any chromatograms are already loaded the newly selected one will be overlaid on top of them.

To delete chromatogram(s) from the project:

- 1. Highlight chromatograms by clicking each one while pressing "Ctrl" key.
- 2. Right-click to activate the context-sensitive menu and choose "Remove chromatogram".

Generating average mass spectra

It is sometimes desirable to know the most prominent mass-spectral features within a specific retention time range. For this purpose, the creation of an average mass spectrum for a specific range is useful. Essentially the whole mass range is divided into small intervals and the intensities of the peaks from every scan that fall within each of those mass intervals are averaged.

To create average mass spectrum:

- 1. Create chromatogram (any type, depending on the specific question in mind).
- Zoom to the desired region of the chromatogram (RT range) to average the spectrum. Averaged spectra will be created only for the files that have their chromatograms displayed.
- 3. Click the icon to extract averaged spectra. Depending on the number of files, data density and retention interval width, this may take a few seconds.

Depending on the sorting options, the average spectra will appear either in the project tree in the "User spectra" folder in the root of the project, or in the "User spectra" folder under each separate data file. The program keeps track of extracted average spectra and prevents duplicates from accumulating, so if spectra are exacted but the display does not update, the data may be already in the project and may be displayed by selecting the appropriate items in the project tree.

To clear spectra panel click the **k** icon.

To load the spectrum in the viewer, click on its name in the project tree. Spectra may not be overlaid and are shown one at a time.

To delete spectra from the project:

- 1. Highlight spectra by clicking each one while pressing the "Ctrl" key.
- 2. Right-click to activate the context-sensitive menu and choose "Remove spectrum".

Performing Compound Search

After search parameters are specified you may start the search by clicking "Find library compounds" button () on a window toolbar. The button is available both on "Multiplex designer" and "Data analysis" panels. When the "Find library compounds" button is selected you will be asked "Are you sure you want to discard existing search results?" Select "yes". The view will switch to "Data analysis" panel and task monitor window will appear to show the progress of the search. The program will search for compounds, fragments and adducts as directed and available.

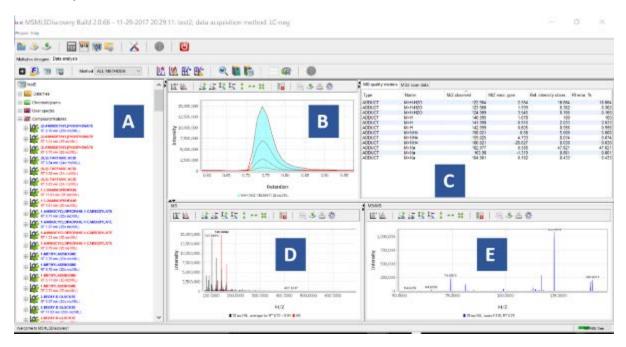
The masses of adducts are calculated as follows:

- 1. Using the molecular formula of the free compound in neutral form (when no innate charge is present, like in choline), the number of atoms is calculated for each element; taking into account whether the adduct is a monomer, dimer, etc.
- 2. Calculate the numbers of atoms added when an adduct is formed. For example, if the adduct is with formate (HCOO-), then 1 hydrogen, 1 carbon and 2 oxygens are added to the element counts for the free compound (as calculated in step 1).
- 3. Calculate the numbers of atoms lost. For example, if there is a water loss, subtract 2 hydrogens and 1 oxygen from the total counts for the respective elements obtained in step 2. At this stage, the software checks if the loss is possible. The total count for the element cannot be negative after the loss. If it is, the adduct is not created. If there is no loss this step is omitted.
- 4. Calculate the mass of the adduct as a neutral molecule based on the element count from step 3.
- 5. Correct the mass from step 4 for the charge by adding/subtracting the mass of electron multiplied by the charge, and then dividing by the charge. This is the expected *m/z*.

Search Results and Curation

Once the search is completed, results are displayed in a tree format in the "Data Analysis" panel and are ready for curation (see Figure 15A).

Figure 15.
Data Analysis Panel



By default, all compounds found in data files for all analytical modes are listed in the alphabetical order. Compound names are color coded (red for positive mode, blue – for negative). Clicking on the compound in the tree will bring up the following information in the corresponding panels:

- Extracted ion chromatograms for base peaks of the detected adduct(s) see Error! Reference source not found.B;
- Mass spectrum in a table format with mass and relative intensity errors comparative to theoretical values see
 Error! Reference source not found.C;
- MSMS (M/Z and intensity values table) if available for any of the MS1 adduct base peaks see Error!
 Reference source not found.C, "MS2 scan data" tab;
- Average mass spectrum plot (MS1) across the compound peak and detected adduct(s) mass spectrum plot see
 Error! Reference source not found.D;
- MSMS plot, if data available MSMS see Error! Reference source not found.
- The "Method" dropdown menu in the window toolbar allows one to switch feature list for individual analytical methods (e.g., show only positive mode or only negative) or display all the features in a single list.

Curation is intended to produce a complete and reliable library of compounds for the purpose of compound identification in data sets. It may include any of the following tasks:

- · Correcting mistakes where the automatic algorithm matching has failed
- Evaluating peak shapes and intensities
- Comparison of retentions times across positive and negative polarities
- Evaluating the quality of spectrum bins
- Editing primary identification for the bins
- Editing bin parameters (retention window, etc.)

The user may delete individual peaks and spectrum bins from the tree. To do that, click on the peak or bin item to highlight it, right-click to activate the context-sensitive menu, and then choose either "Delete bin" or "Remove peak" depending on the type of object selected. The user will be prompted to confirm the operation. To delete multiple bins,

highlight them by clicking each one while pressing the "Ctrl" key and then click the ** icon on the main toolbar.

Note: Deleting bins and peaks is NOT reversible.

When curation is complete, click the "Save library to database" button" ().

Exporting Library for Use in Subsequent Research Experiments

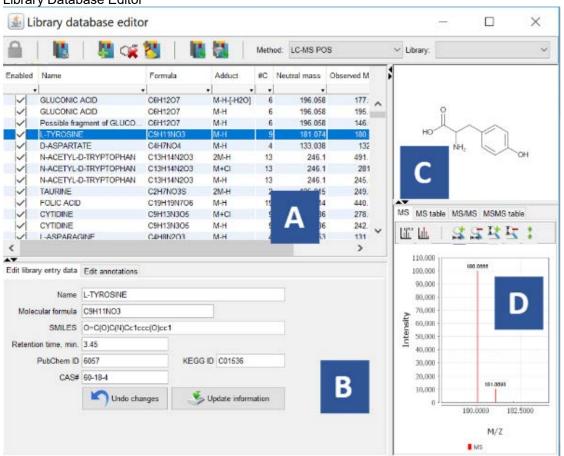
After the data curation, the library may be exported in tab delimited (TSV, plain text) format or Chemical Exchange Format (CEF, developed by Agilent Technologies, XML-based). Export is possible for all analytical modes simultaneously and data can be saved separately for each method (recommended) or combined for all methods. Please see next section for details.

Library Database Editor

Library database editor (see Error! Reference source not found.) is intended to manage compound libraries stored

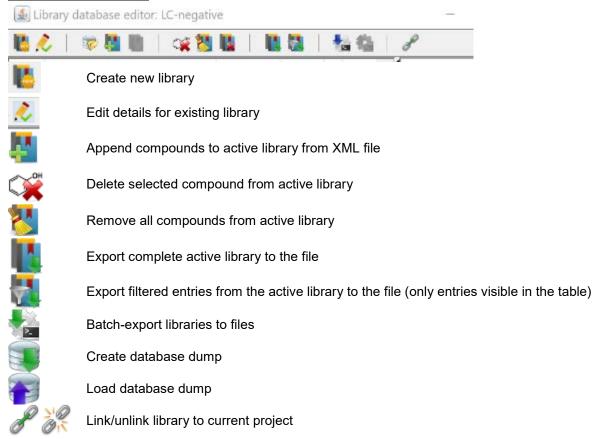
in the internal database of the software. To activate the editor, click button on the main software toolbar. Make sure following curation, libraries were saved to the database using the "Save library to database" button" () in the data analysis panel.

Figure 16. Library Database Editor



The editor window includes the library listing table (A) which may be filtered and sorted using header rows, structure display (B), feature editor panel (C), and mass spectrum panel (D).

Library Editor Toolbar



Several parameters for the library entry may be edited (see Figure 18) by the user. The form is automatically populated by selecting the feature of interest in the table as highlighted below.

Figure 17.Library Database Editor

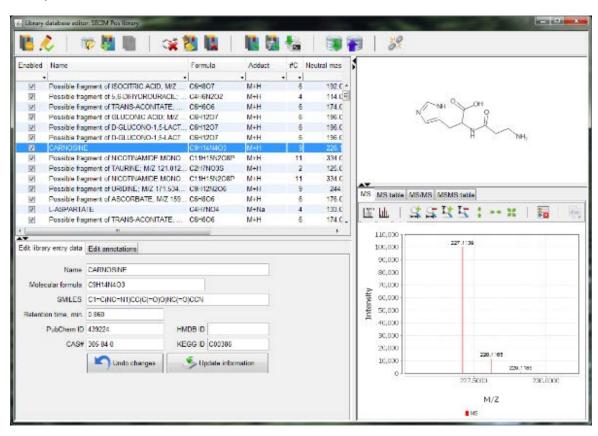
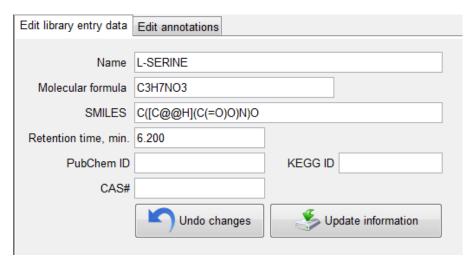


Figure 18. Library Entry Data Editor



All the fields are verified before saving the edited data. Undo button will roll back the changes only if "Update information" was not clicked yet.

Every library feature may have multiple text annotations. To add, edit, or delete text annotations use the "Edit annotations" tab (see Figure 19).

Figure 19. Library Annotations Editor



Annotation editor toolbar:



Specifying Libraries Using Database Manager

It is possible to attach or change the library for the method in the active project at any time using library database manager. The method must be set as active first. Load the desired library using "Load library from database" button

(), then click on the "Chain" icon () to link library to the method. The icon will change to indicating that library is linked to the method. The library name will also appear in the corresponding method tab on the project setup panel.

Importing File-Based Libraries or Adding New Libraries

To use previously created file-based libraries they must be first imported into the database.

To import the library from file, create a new empty library in the database by clicking "New library" button (specify the name, description (optional), and associated separation method in the "Create new Library dialogue (see Figure 21).

Then use "Append library from disk ..." button () to select and load library file. Depending on the size of the library import may take from a few seconds to a few minutes.

Figure 20. Library Manager Window

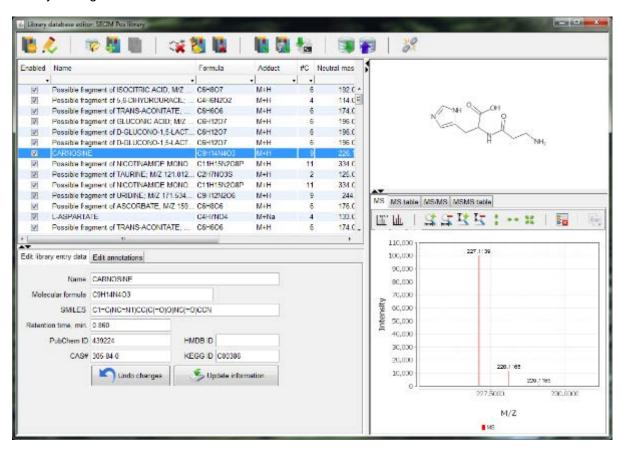
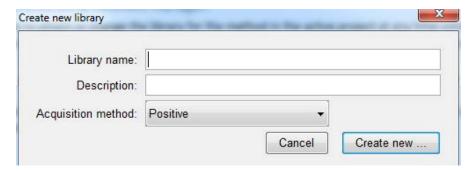
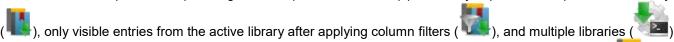


Figure 21.New Library Dialog



Exporting Libraries to File

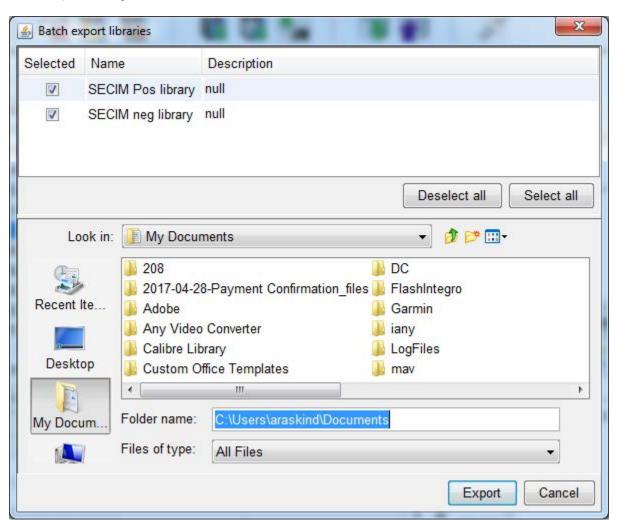
There are several options to export single or multiple libraries to file(s). You may export the complete active library



To export individual libraries, in the "Data analysis" screen click "Export compound library to file" button (Library database editor which will bring up a new window with a toolbar allowing to choose export format — tab delimited, MSP format (called NIST format in the drop-down menu), or Agilent Compound Exchange Format (CEF) (see Figure 22). Select location to store the library. Then, again in the "Data analysis" screen, save the library

to database by clicking on "Save library to database" icon ().

Figure 22.
Batch Export Dialog



To combine and export combined libraries, again use the Library database editor (see Figure 23). Here the user can append libraries (), export library (), and export filtered libraries

To apply library filter, you may type in the top row of any column (just below the table header) to select a subset of records. For example, in the top row of the "name" column if you type "DEOXY", only the compounds that start from "DEOXY" will appear (see Figure 24).

Figure 23. Library Export Dialog

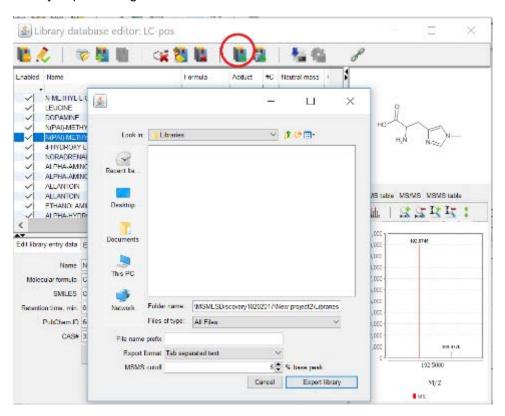
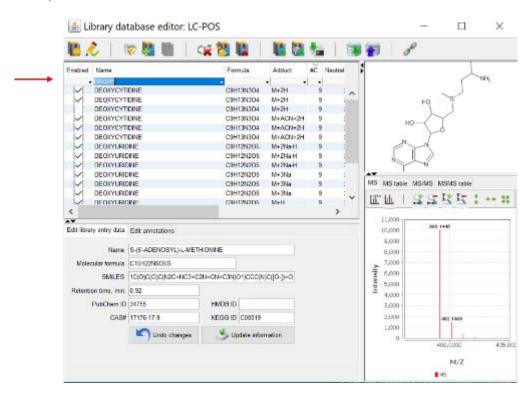


Figure 24. Library Database Editor



Libraries can be imported into different software programs or used in other independent library search products, for example Agilent's "Personal Compound Database Library" (PCD) or used directly in the MLSDiscovery software program to analyze samples, see Section 4. Libraries can be also directly imported into the IROA ClusterFinder software program for metabolic profiling. To use your new library in the sample search mode select it at the last page of the project setup wizard as shown above for the library generation project.

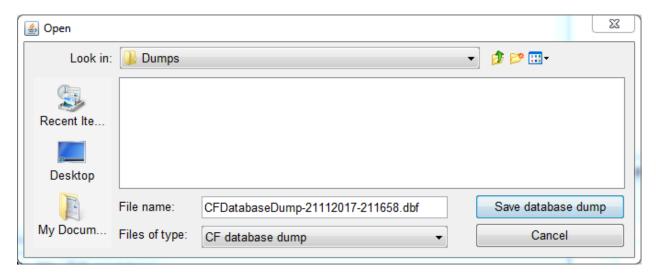
For batch library export, select the "Show export multiple libraries dialog" icon () to reveal the Batch export libraries" screen (see Figure 22) and select the libraries you wish to batch export and then click the Export button.

Exporting and Importing Complete Database Dump

This procedure is mainly intended to safely and easily transfer the user-generated data when upgrading to the new version of the program, but it also may be used to create data backups. When database dump is invoked all libraries, acquisition methods, adducts, and mass differences are extracted from the database and saved as a single compressed file. File name is generated automatically and contains the timestamp. The dump files have extension "dbf".

To initiate the dump, click icon on the library manager toolbar, change the destination for the dump and file name if necessary, and click "Save database dump" button to create the dump (see Figure 25).

Figure 25. Create Database Dump



To load an existing dump, click the icon on the library manager toolbar (the warning will pop up, read through and click "OK" to continue), select the dump file and click "Load database dump" button to load the dump into the database.

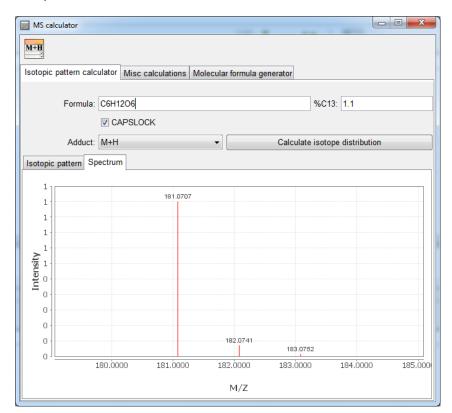
<u>Note</u>: **IMPORTANT**: If you are loading the dump into the copy of software that has already been used, note that all the user-generated information in the database will be erased and replaced by the data from the dump!

MS Calculator

MS calculator is a general-purpose mass spectrometry toolbox. It may be called by clicking the button on the toolbar. MS calculator has tree panels: isotopic pattern calculator (see Error! Reference source not found.), miscellaneous calculations (see Error! Reference source not found.), and molecular formula generator (see Error! Reference source not found.).

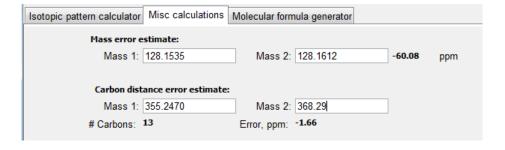
Isotopic pattern calculator allows user to simulate isotope distribution for a given molecular formula, taking into account specific level of ¹³C incorporation and type of adduct. Data are presented in both graphical and table formats. CAPSLOCK checkbox simplifies typing in of the formulas not containing 2-letter element symbols.

Figure 26. Isotopic Pattern Calculator



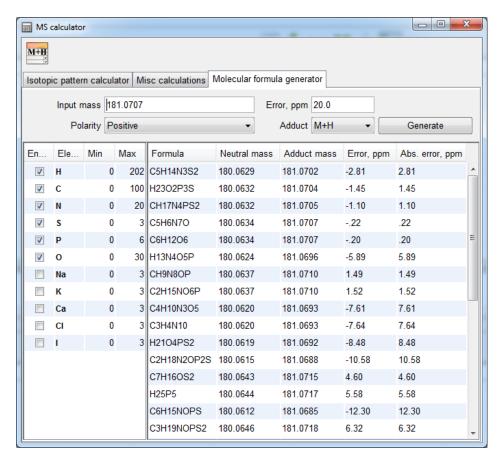
Miscellaneous calculations panel allows user to quickly estimate mass errors to help during the manual curation of the data.

Figure 27.MS Calculator – Miscellaneous Calculations



Molecular formula generator allows the user to calculate possible formulas given the m/z value, adduct type, mass accuracy, and defined ranges of element atom numbers. Mass and adduct type may be populated automatically when the generator is called through the context menu on the spectrum bin (see **Error! Reference source not found.**).

Figure 26.
Molecular Formula Calculator



Close the Discovery program and you will find that you have the entire project saved which can be re-opened at any time.

Sample Analysis

Now that you have built the library using the MSMLS, LSMLS, or other MLS authentic compounds and Discovery software, they may now be used for sample analysis. Locate the saved libraries generated in previous section. Libraries will be in either Chemical Exchange Format (CEF, developed by Agilent Technologies, XML-based) or tab delimited (TSV, plain text) format.

Initiating a Sample Analysis Project

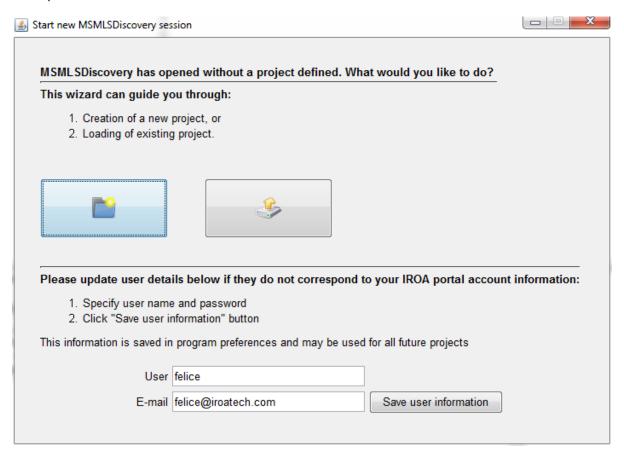
To begin sample analysis, either open the program again or to initiate a new project by clicking "Create new project"



button (

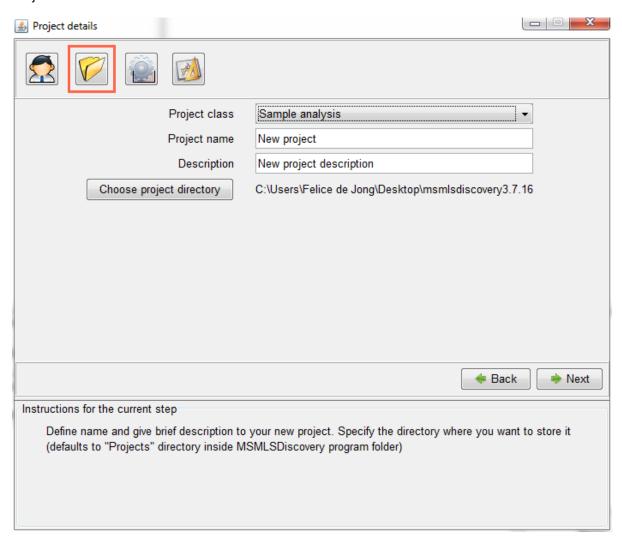
) which will bring up the first page of a New Project Wizard (see Error! Reference source not found.).

Figure 27. Startup Task Chooser



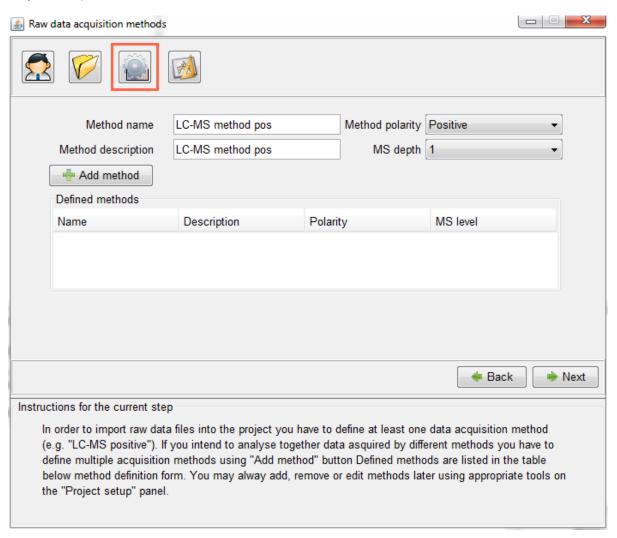
In the "Project selection screen" (see **Error! Reference source not found.**), specify "Project class" by using the arrow to select "Sample analysis". Specify the name and description of the experiment, change project home directory if required, and click "Next" button to move to the method definition screen (see **Error! Reference source not found.**).

Figure 28. Project Selection



Specify name, description, polarity, and MS depth for each analysis method you have used to collect the raw data of the samples you wish to analyze. Use "Add method" button to add specified method to the project (it will show up in the "Defined methods" table below). You may edit method information, add new and delete existing methods later as well. When finished click "Next" to move to sample definition screen (see **Error! Reference source not found.**).

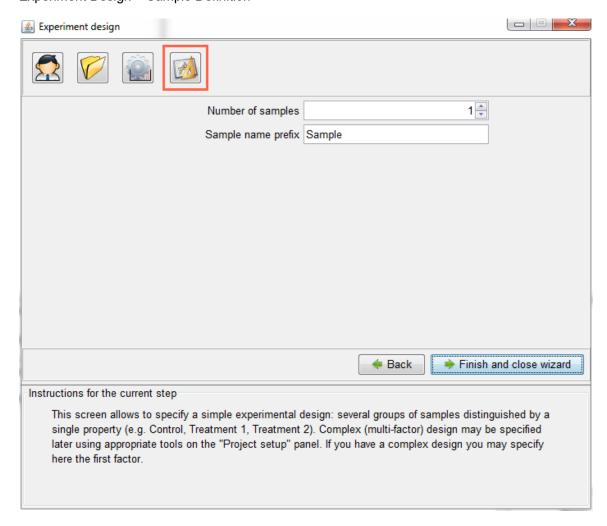
Figure 29. Project Setup – Method Definition



Defining and Loading Experimental Samples

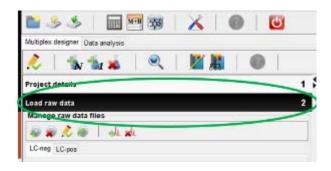
Specify the number of samples and modify sample name as desired by clicking on name sample name prefix, then select "Finish and close wizard" (see Figure 30).

Figure 30. Experiment Design – Sample Definition



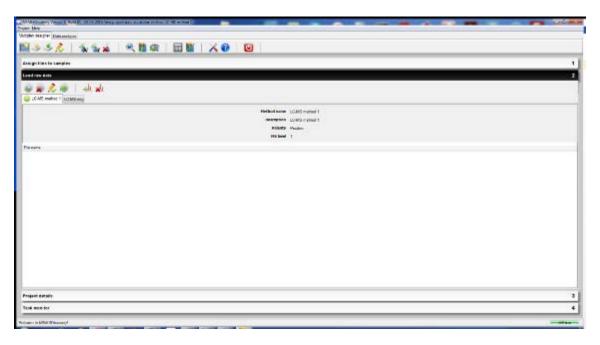
Select the "Load raw data tab" (see Error! Reference source not found.).

Figure 31. Load Data Tab



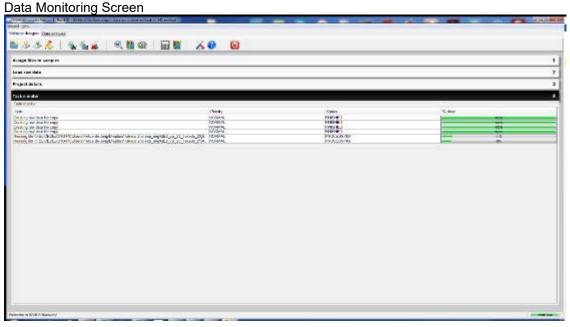
Load data by clicking "Add raw data files" button () on the panel toolbar (see Figure 32) and navigating to the directory containing the raw data. Be sure to load <u>centroided mzXML</u> files. Click on each method and repeat to add all data files for each acquisition method.

Figure 32. Load Data Tab



During file import, the "Task monitor" section is activated to show operation progress (see **Error! Reference source not found.**).

Figure 33.



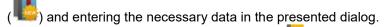
After importing the raw data files, assign the files to their corresponding samples by selecting the "Assign files to samples tab" and choosing the files from the dropdown menus for each of the defined methods (see **Error! Reference source not found.**).

Figure 34.Data Monitoring Screen

	mod 1	
odec; Felb		
Authories designer Data seraignite		
🖺 🍮 🏂 🌡 🐁 🖦 📗 🍭 🗎 📾		
Assign files to samples		
Samples to data files map	- HARRISTON OWN	1.10000
Europie rums	LC-M8 method 1	LC-M
cardrol #001	Select calls de	w Sales
cartrel #001 cartrel #002	Select colords Select colords	* See
carried #002	Select code nie	* Sales
corrol 6002 corrol 6003	Select colo ida Select colo ida	* Selection

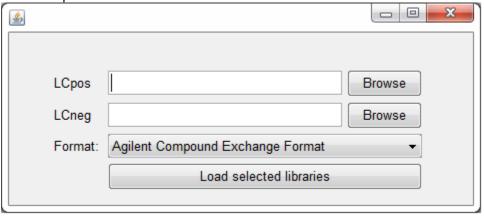
Load Compound Libraries Created During Library Generation

In order to use previously created file-based libraries they have to be imported into the database first. Instructions were previously presented under "Library Database Editor" in the "importing file-based libraries or adding libraries section". To import the library from file, first create a new empty library in the database by clicking "New library" button



Then use "Append library from disk ..." button () to select and load library file (see Figure 35). Depending on the size of the library import may take from a few seconds to a few minutes.

Figure 35.
Load Compound Libraries



If the library is already in the data base, then load the desired library using "Load library from database" button



When the correct library is loaded, be sure to click on the "Chain" icon () to link library to the method. The icon

will change to findicating that library is linked to the method. The library name will also appear in the corresponding method tab on the project setup panel.

Specify Search Parameters and Find Library Compounds in the Samples

Specify search parameters. Most of the parameter names are self-explanatory, except smoothing width, which is specified as a fraction of the total run time (duration of chromatographic separation in LC-MS). It also should be noted that minimum peak intensity filter is applied to the highest (usually monoisotopic) peak of the adduct, not to any mass peak in general. See the previous section on select search parameters.

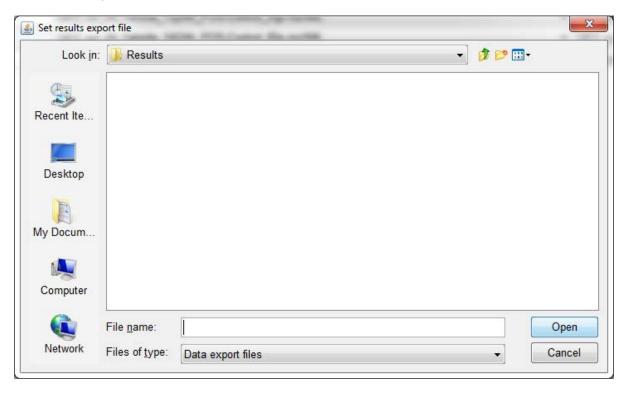
After search parameters are specified you may start the search by clicking "Find library compounds" button () on a window toolbar. The button is available both on "Multiplex designer" and "Data analysis" panels. When the "Find library compounds" button is selected you will be asked "Are you sure you want to discard existing search results?" Select "yes". The view will switch to "Data analysis" panel and task monitor window will appear to show the progress of the search.

Export Results

Once the search is completed results are displayed in a tree format (see Error! Reference source not found.).

Results can be exported by clicking on the "Export Results" button (") on the window toolbar will produce a "Save file" dialog prompting for the destination of the file with quantitative data for every detected library compound in every sample analyzed (see Figure 36). Provide a file name for the project and click "open".

Figure 35. Set Results Export File



We hope the MLSDiscovery program is useful and welcome any feedback. Please send your comments and suggestions to info@irotach.com.

MSMLS, LSMLS, FAMLS, OAMLS, BACSMLS, MLSDiscovery and ClusterFinder are trademarks of IROA Technologies, LLC.

SS,AA,KVG,PD,MAM 04/19-1