

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

Purpose

This document provides basic guidance for user interaction with QTOF/TOF data in the Qualitative Analysis ("Qual.") software.

Applicability & Authorization

Any BioCORE staff or research user who wishes to interact with their QTOF/TOF data for inspection, background spectra subtraction, etc. at the level of a single data file. Prior training from BioCORE staff is required before using the LC/MS QTOF to tune the instrument and collect data for analysis.

Instrumentation & Software

The Agilent 1290 UHPLC multi-sampler liquid chromatography system linked to the Agilent 6550 Quadrupole Time-Of-Flight (QTOF) mass spectrometer is used to collect MS or MS/MS data. Qualitative Analysis version B.07 or higher will be used for this SOP and this software is installed on the "QTOF" data analysis computer.

Background

All data inspection, evaluation, background subtraction, and manual peak picking is performed on single data files in this program. This program is not to be used for any kind of Quantitative analysis.

Certain functions in Qual. can be applied to multiple data files at the same time. These functions include chromatogram extraction, chromatogram integration, spectrum extraction, signal-to-noise calculation, and more.

Although Qual. can be used for basic data evaluation, it has powerful utility in untargeted metabolomics. After basic inspection of data files, a user will choose a "Find Compounds" option for finding compounds using different methods. After finding compounds, a user can "Identify Compounds" by searching a library (MS/MS data) or database (MS1 only spectra). These options will be described in more detail below.

Opening data files in Qual.

- File> Open data file> navigate to the folder containing your data (.d files)> can select 1 or multiple files.
- If files have been previously opened in Qual. and results have been saved with the .d file, then under Options in the Open data file window select the box for 'Load result data'. You can also

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

select 'Load results method' if you would like to load the methods used to generate the results for this data file when it was last saved.

Method Explorer> Chromatogram

Define Chromatograms

Here you can define and generate multiple chromatograms at the same time from one or more data files. There is an option to 'Integrate when extracted' for each chromatogram, and this will use the current settings within the Chromatogram > Integrate (MS) method explorer. A TIC is the default spectra when QTOF/TOF data is imported into Qual. but the total ion chromatogram is useless so extracting chromatograms is important.

-BPC means base peak chromatogram, and extracts the largest ion from the TIC.

-EIC (most common) means extracted ion chromatogram, and extracts selected ions (1 or more) from the TIC. Within the EIC options under Advanced, it is best to set the Single m/z expansion to Symmetric +/- 20 ppm. You can also exclude masses (1 or more) if desired. You may want to exclude the reference ions @ 121 and 922 m/z.

-pNLC extracts neutral losses, such as mass of water.

-Instrument Curve and Other Chromatograms can be used to extract instrument details such as DAD (UV trace), pressure, or temperature.

Extract Chromatograms... (from Chromatogram Results)

An alternative to method editor is to right click on any chromatogram in the Chromatogram Results window and select Extract Chromatograms... This functions identically to Define Chromatograms but the user can only extract one chromatogram at a time and only from a single data file.

Extract EIC> At Maxima in Ranges (from Spectrum Preview)

Another alternative for extracting an EIC is from the Spectrum Preview window. When walking along any chromatogram in the Chromatogram Results window, the Spectrum Preview will show the current spectra for the selected scan. Using left click-and-drag to select ions of interest in the spectrum preview (1 or more ranges can be selected by holding in Control), right click inside the range and select Extract EIC> At Maxima in Ranges. This will create an EIC for the maximum abundance ion within the selected ranges over the full RT of the MS scan for this data file only.

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

Extract EIC from MS1 level or MS/MS level data

-When collecting QTOF data in Auto-MS/MS or Targeted-MS/MS modes, the TICS generated have a jagged appearance. This jagged appearance exists in both the TIC as well as in the EIC derived from MS/MS level data. The reason for this is due to how the QTOF instrument cycles through MS and MS/MS scans when doing these scan types.

-[Extract EIC from MS data](#): this will extract MS1 level only data from the TIC. For example, if targeted-MS/MS was performed and known precursor ions were included in the QTOF acquisition method, then these precursor ions will be highly represented in the peaks in the extracted EIC. No MS/MS level data (product ions of the known precursors) is included in this EIC.

-[MS Chromatogram tab](#): common settings to select are MS Level=MS, Scans=Scan, m/z values= full range of data collected, a smaller range that includes precursors, or multiple m/z values separated by commas.

-[Advanced tab](#): Single m/z expansion window should be Symmetric (ppm) +/- 20. Default of 100 is much too large.

-[Excluded Masses tab](#): option to exclude known masses. Users may wish to regularly exclude the QTOF reference calibration masses @ 121 and 922 m/z, however these can be useful for visual referencing.

-[Integrate when extracted](#): when checked, this option will automatically perform integration (peak picking) on the extracted EIC based on the setting in the Chromatogram> Integrate (MS) tool.

-[Extract EIC from MS/MS level data](#): this will extract MS/MS level data only from the TIC. If targeted-MS/MS was performed on known precursors, then all product ions resulting from collision energy applied to these precursors will be represented in the extracted EIC. No MS1 level data is included in this EIC. Note that often precursor ions will show up in the MS/MS scan due to incomplete fragmentation in the collision cell.

-[MS Chromatogram tab](#): common settings to select are MS Level=MS, Scans=Scan, m/z values= full range of data collected, a smaller range that includes precursors, or multiple m/z values separated by commas.

-[Advanced tab](#): Single m/z expansion window should be Symmetric (ppm) +/- 20. Default of 100 is much too large.

-[Excluded Masses tab](#): option to exclude known masses. Users may wish to regularly exclude the QTOF reference calibration masses @ 121 and 922 m/z, however these can be useful for visual referencing.

-[Integrate when extracted](#): when checked, this option will automatically perform integration (peak picking) on the extracted EIC based on the setting in the Chromatogram> Integrate (MS/MS) tool.

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

Chromatogram> Integrate (MS)

- This tool is used to choose settings for automated peak picking from EICs containing MS1 level data.
- Integrator tab: choose peak-picking algorithm. Agile 2 is the newest algorithm and is suggested by Agilent. Agile 2 is better than Agile 1 for small peak detection and MS/MS. Agile 2 is parameter-less.
- Peak Filters tab: used to specify how to select peaks and includes options to filter by height or area, specifying absolute counts or % of largest peak, and specifying the maximum number of peaks to pick.
- By right clicking inside an EIC (single EIC): right click inside an EIC with MS1 data and select Integrate Chromatogram. This will use the settings in the Integrate (MS) tabs.
- From EICs in Data Navigator (multiple EICs): user can select multiple EICs in the Data Navigator window and then click Integrate Chromatogram in the Method Editor. This will integrate according to the settings selected.

Chromatogram> Integrate (MS/MS)

- This tool allows for the automated peak picking within EICs containing MS/MS data.
- Integrator tab: choose peak-picking algorithm. Agile 2 is the newest algorithm and is suggested by Agilent. Agile 2 is better than Agile 1 for small peak detection and MS/MS. Agile 2 is parameter-less.
- Peak Filters tab: used to specify how to select peaks and includes options to filter by height or area, specifying absolute counts or % of largest peak, and specifying the maximum number of peaks to pick.
- By right clicking inside an EIC (single EIC): right click inside an EIC with MS/MS data and select Integrate Chromatogram. This will use the settings in the Integrate (MS/MS) tabs.
- From EICs in Data Navigator (multiple EICs): user can select multiple EICs (with MS/MS data) in the Data Navigator window and then click Integrate Chromatogram in the Method Editor. This will integrate according to the settings selected.

-Note that peak integration can also be performed as part of an Unknowns Analysis Workflow, to be described later.

Chromatogram> Calculate Signal-to-Noise

This tool is used to calculate the SNR (signal-to-noise ratio). Signal can be defined by Height or Area, Noise can be defined in different ways. Common definitions used are peak-to-peak and root-mean-square (RMS). Noise region can then be determined automatically within set parameters (desired length and minimum length of noise region) or specific noise regions can be set. To do this, input time ranges using dashes (-)

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

and commas (for example: 1.5-3.0, 4.5-5.0). Finally, you can select a check box that will automatically compute SNR when chromatograms are integrated. Again, this can be applied to all highlighted EICs in Data Navigator by hitting Calculate Signal-to-Noise. Note that noise regions used in the SNR calculation will automatically show as bolded lines on the EIC.

Chromatogram> Adjust Delay Time

Most users will not use this regularly. This tool can be used to adjust the delay time for the MS1 level data (EIC from MS1 Scan) relative to the DAD1 signal (UV trace) so that the UV trace and MS1 TIC are overlaid properly in time. This adjustment is required because the LC stream flows through the UV detector before entering the MS instrument. Adding a time delay of 0.05 min to the MS1 level data will shift the EIC retention time by 0.05 min to the left to approximately line up with the UV trace if the flow rate was 0.5 mL/min during acquisition.

Method Explorer> Spectrum

Anchoring

When spectra are created they will be automatically anchored (linked) to the EIC from which they originated. The anchoring exists by default, but can be inactivated by deselecting the Linked Navigation icon. If a user wished to delete specific spectra but not the EIC from which it originated, then it is critical to turn off the anchoring to the EIC.

Setting a background spectrum (for normalizing User Spectra of interest)

-A user may wish to subtract a background spectrum from user spectra of interest to remove contaminating ions, background ions, or ions that are not of interest. Only one spectra can be assigned as background for manual spectra extraction (see below). Also, this background spectra only applies to a single data file. There are three ways to set a background spectrum.

-[By walking the chromatogram \(single point spectrum\)](#): using the Walk Chromatogram tool will open the Spectrum Preview window where you can right click and select Move to Background Spectrum.

-[By selecting ranges in the chromatogram](#): using the Range Select tool in Chromatogram Results, a user can select a range (or multiple ranges by holding Control) and then double click inside the range to bring up a Spectrum Preview representing the range(s) selected. Right clicking in the preview window you can select Move to Background Spectrum.

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

-[From MS Spectrum Results/User Spectra in Navigator](#): if for some reason you want to set the background spectrum from an existing User Spectra, clicking the spectra of interest in Data Navigator will open the MS Spectrum Results window. Right clicking in this window, or right clicking the User Spectra in Navigator, you can select Move to Background Spectrum.

Creating User Spectra

-User Spectra can be extracted manually or automatically from integrated EIC peaks. Additionally, User Spectra can be background spectra subtracted manually or automatically.

-[By walking the chromatogram \(single point spectrum\)](#): using the Walk Chromatogram tool will open the Spectrum Preview window where you can right click and select Copy to User Spectra.

-[By selecting ranges in the chromatogram](#): using the Range Select tool in Chromatogram Results, a user can select a range (or multiple ranges by holding Control) and then double click inside the range to bring up a Spectrum Preview representing the range(s) selected. Right clicking in the preview window you can select Move to Background Spectrum.

-Note that the manual spectra extraction methods above are influenced by the active selection for Manual spectrum background in the Manual Extraction tab of the Extract (MS). If the background is set to None, then the spectrum will be exactly as in the Spectrum Preview window, but if the background is set to Current background spectrum (if set in Data Navigator) then the extracted spectrum will automatically have this background subtracted. This is denoted by the word Sub in the name of each User Spectra in Navigator.

-[By selecting an integrated peak in the EIC](#): if an EIC has been integrated and peaks picked, use the Peak Select tool and double click on any peak to automatically extract the spectrum from the peak. This will automatically add the spectra to the User Spectra and will be represented by the retention time (rt) range for that peak. *Note that this semi-automatic spectra extraction is not subject to influence by the Manual Extraction tab. Rather, it is influenced by the Peak spectrum background selection and other settings in the Peak Spectrum Extraction (MS) tab within the Extract (MS) tool.* If a background subtraction method is selected here, then the peak spectrum will have a background subtracted when the peak is double-clicked in the EIC.

-Any spectra that is background subtracted with have the word Sub in the name.

-[By using the Extract \(MS\) or Extract \(MS/MS\) tools for automated extraction \(see next section\).](#)

Spectrum> Extract (MS)

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

-This tool is used to extract mass spectra from integrated chromatogram peaks, and therefore should be used after the user has integrated their desired EICs containing MS data (but not MS/MS). The Manual extraction tab only influences manual spectrum extractions as outlined above, and can be ignored for automated extraction.

-Peak Spectrum Extraction tab: these parameters affect what data is extracted from peaks. For Spectra to include, standard is to Average the scans that represent >10% of the peak height (ignoring the lowest 10% close to background). This improves mass accuracy vs. At apex of peak, which results in higher abundance values. For TOF spectra, it is standard to Exclude if above 10% of saturation (this means that if a peak shows saturation [a flat peak], only then the scans within 10% of saturation will be ignored). User can select m/z ranges from which to extract spectra if desired, but it is common to use the m/z range from the EIC. User can select to never return an empty spectrum (which wouldn't be useful for anything). Finally, user can choose a peak spectrum background subtraction method or time range if desired. It is standard to select None unless a user has a specific reason for subtracting a background.

-Peak Location tab: this tab determines how peaks are centroided. Maximum spike width is the number of data points a peak must have in order to be considered a peak. This prevents extraction of noise. The Required valley is the relative amount of the peak height that a peak must "reach" on its way back down towards baseline in order for it to be considered a "real" peak. This applies to peaks that are not baseline resolved. Best to use the default values of width 2 and valley 0.7. Poorly resolved isobars might benefit from reducing the required valley.

-Charge State tab: the isotope model can be set here. It is standard to select Common organic molecules (no halogens) when doing metabolomics work, unless you have a good reason to suspect halogens may be present. Limit charge states based on the compound class of interest; for small molecule applications, it is common to limit charge state to a maximum of 1-2. The peak spacing tolerance and ppm window for isotope groupings affects the error tolerance for isotope clusters. This should almost never be changed from default values (0.0025 and 7.0 ppm).

-Peak Filters tab: this tab is used to filter which peaks are selected for spectrum extraction. Setting peak filters will reduce noise in the spectra. Options include absolute peak height, relative peak height, or both. If these settings are too stringent, you will eliminate isotope peaks, which is undesirable. These settings depend on what you are trying to extract. General settings include filtering by Absolute height >100 counts. Optional is to include Relative height >1% of largest peak, but this may filter out some isotope peaks and smaller peaks if the largest peak is of very high abundance.

-From EIC using Peak Selection Tool (only selected peaks): using the peak selection tool you can click and

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	ACH B0-103	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

drag across multiple peaks which will appear highlighted. User can then click Extract Peak Spectrum button in the Method Editor> Extract (MS) window.

-[From EIC using right click> Integrate and Extract Spectrum \(all peaks\)](#): on any EIC chromatogram, the user can right click and select Integrate and Extract Spectrum. This will first integrate the EIC based on the current settings in Chromatogram> Integrate to generate peaks and then automatically apply these peaks to the Spectrum> Extract settings and generate User Spectra accordingly. *Note that if the EIC contains MS1 data, then this process will use the Integrate (MS) and Extract (MS) settings, and an EIC containing MS/MS data will be processed using the Integrate (MS/MS) and Extract (MS/MS) settings.*

Spectrum> Extract (MS/MS)

-This tool is used to extract mass spectra from an integrated EIC containing MS/MS data.

-[Peak Spectrum Extraction \(MS/MS\) tab](#): these parameters affect what data is extracted from peaks. For Spectra to include, standard is to Average the scans that represent >10% of the peak height (ignoring the lowest 10% close to background). This improves mass accuracy vs. At apex of peak, which results in higher abundance values. For TOF spectra, it is standard to Exclude if above 10% of saturation (this means that if a peak shows saturation [a flat peak], only then the scans within 10% of saturation will be ignored). User can select m/z ranges from which to extract spectra if desired, but it is common to use the m/z range from the EIC. User can select to never return an empty spectrum. Finally, user can choose a peak spectrum background subtraction method or time range if desired. It is standard to select None unless a user has a specific reason for subtracting a background.

-[Peak Location tab](#): this tab determines how peaks are centroided. Maximum spike width is the number of data points a peak must have in order to be considered a peak. This prevents extraction of noise. The Required valley is the relative amount of the peak height that a peak must "reach" on its way back down towards baseline in order for it to be considered a "real" peak. This applies to peaks that are not baseline resolved. Best to use the default values of width 2 and valley 0.7. Poorly resolved isobars might benefit from reducing the required valley.

-[Charge State tab](#): the isotope model can be set here. It is standard to select Common organic molecules (no halogens) when doing metabolomics work, unless you have a good reason to suspect halogens may be present. Limit charge states based on the compound class of interest; for small molecule applications, it is common to limit charge state to a maximum of 1-2. The peak spacing tolerance and ppm window for isotope groupings affects the error tolerance for isotope clusters. This should almost never be changed from default values (0.0025 and 7.0 ppm).

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

-**Peak Filters tab:** this tab is used to filter which peaks are selected for spectrum extraction. Setting peak filters will reduce noise in the spectra. Options include absolute peak height, relative peak height, or both. If these settings are too stringent, you will eliminate isotope peaks, which is undesirable. These settings depend on what you are trying to extract. General settings include filtering by Absolute height >100 counts. Optional is to include Relative height >1% of largest peak, but this may filter out some isotope peaks and smaller peaks if the largest peak is of very high abundance.

-**From EIC using Peak Selection Tool (only selected peaks):** using the peak selection tool you can click and drag across multiple peaks which will appear highlighted. User can then click Extract Peak Spectrum button in the Method Editor> Extract (MS/MS) window.

-**From EIC using right click> Integrate and Extract Spectrum (all peaks):** on any EIC chromatogram, the user can right click and select Integrate and Extract Spectrum. This will first integrate the EIC based on the current settings in Chromatogram> Integrate to generate peaks and then automatically apply these peaks to the Spectrum> Extract settings and generate User Spectra accordingly. *Note that the EIC must contain MS/MS data and the process will use the Integrate (MS/MS) and Extract (MS/MS) settings, not the MS1 only settings.*

-Note that spectrum extraction can also be performed as part of an Unknowns Analysis Workflow, to be described later.

Spectrum> Extraction Data Format

This tool can be used to set preferences for the format of Chromatogram data and Mass spectral data when it is extracted. Best is to use "Profile when available, otherwise Centroid" for both settings because Profile data contains more signal and information.

MS Spectrum Results Window> Delta Mass Caliper tool

This tool can be used to calculate the loss of mass between two ions in a mass spectra, such as the loss of water (18 m/z) or salt (23 m/z). This is also useful to look for common contaminants like adducts which show up as multiple peaks with 22 m/z or 44 m/z separation.

Find compounds> Find by Molecular Feature (molecular features = potential compounds)

-This is typically used in an Unknown Compound Discovery Workflow

-The Molecular Feature Extraction (MFE) tool is an extraction algorithm. The extraction algorithm starts by

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

organizing MS1 data (MS1 level only!) into a 3D array with RT, m/z, and abundance. It will look for and remove noise, before looking for mass signals associated with a common RT and other relationships. It will group m/z ions that are present at a common RT. Then m/z are grouped into isotope clusters and other mass relationships such as adducts, dimers, and charge states. Finally, isotope clusters are grouped into molecular features (aka compounds). Note that when a data file is subject to MFE it will bring along the MS/MS data for the ride, but this data is not used in any of the algorithms.

-MFE will generate extracted compound chromatograms (ECC) which is an EIC that sums all of the ion species associated with a compound (molecular feature). Ions in ECC are expressed as volumes, not abundance, and represent the total area of all ions associated with the compound.

-The list of compounds extracted by MFE are not identified by default, and this extraction should be followed by a Compound Identification process (next section).

-Extraction tab (data extraction): these settings affect how data is extracted from a data file. Target data type should be Small molecules (chromatographic). Input data ranges can be restricted by RT or m/z range, depending on user. Peak Filters should be set relative to the background present when looking at the integrated EIC, and users should set the minimum peak height count a bit above the level of the constant noise background. Alternately users can set a minimum signal-to-noise ratio (only applicable if profile data was collected).

-Ion Species tab (data extraction): these options impact how the MFE algorithm creates relationships between ions during extraction. Users should include all expected and common ions and neutral losses commonly found when doing metabolomics. If users are collecting ions in positive polarity then more positive ions will be expected and users should select all default positive ions. Certain ions (such as NH₄⁺) may be expected in higher abundance if the chromatographic solvents contain that ion (NH₄⁺ from ammonium formate). H⁺ and Na⁺ are very common positive ions. If samples were derivatized during sample preparation, then the derivatized ions can be added manually in this window so that the MFE algorithm will look for these ions.

-Charge State tab (data extraction): these options impact how the MFE algorithm creates relationships between ions during extraction. Users should use the default options including peak spacing tolerance of 0.0025 m/z plus 7.0 ppm, Common organic molecules as the Isotope model for Unknown Metabolomics Analysis workflows, and a maximum charge state of 2 because small molecules usually have only 1-2 charges.

-Peak Filters (MS/MS) tab (data along for the ride): Ignore this tab, it plays no role in MFE. MS/MS level data simply comes along for the ride in MFE, and the settings on this tab have no impact on data

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

extraction or display.

-Compound Filters tab (display only): these options affect what compounds are displayed; they do not affect the extraction (compounds can be extracted that are not displayed). Users can include or exclude compounds displayed based on relative height and/or absolute height and/or limit to the largest # of compounds. Compound Quality score is a measure of how “good” a feature is and it is based on the SNR (signal-to-noise ratio), the RT peak/width, the co-elution score, and the change in calculated m/z. The default Quality score is 80. Users can also restrict the RT and charge states of compounds to display.

-Mass Filters tab (display only): this tab allows the user to exclude masses or include masses with a window around the mass, and only impacts what compounds are displayed. By default this tab is not commonly used.

-Mass Defect tab (display only): this can be used to filter out impurities, if known, and only impacts that compounds are displayed. By default this tab is not commonly used.

-Results tab (display only): this tab affects how compounds are displayed. The check box for Delete previous compounds is checked by default so that the compound list is replaced any time the Find by MFE is run. User can choose to highlight first or all compounds. User should at a minimum Extract MFE spectrum and Extract ECC, and often Extract EIC for MS1 level data. If MS/MS data is available (from auto-MS/MS or targeted MS/MS data files), the user can Extract MS/MS Spectrum and choose per collision energy (if >1 CE was used during data collection) which is recommended, or Extract average spectrum for all CE. Precursor tolerance window of 20 ppm is standard. Users should not click Deisotope MS/MS spectrum as isotopes are important for qualitative assessment of spectra. Users can also limit the display to the largest # of compounds, if only top hits are of interest.

-Advanced tab (display only): this tab affects how compounds are displayed. User can choose to Include All compounds regardless of the number of ions associated with the feature, and if this is chosen there is a further choice of filtering by including, excluding, and including only compounds with indeterminate neutral mass. User default is to Include all compounds regardless of ion number and to exclude compounds with indeterminate neutral mass.

-Find Compounds by Molecular Feature: when all selections have been made, clicking this button will result in a pop up window that allows the user to select the data files you wish to have the MFE algorithm applied to. Click Find and the algorithm will begin the process of extracting the features from each data set and then filtering the features for display. Finally, running MFE will give a table of extracted molecular features called compounds that fit the display filters. Popular columns in this table are Score (MFE), Ions, Base Peak, Mass, m/z, RT, Algorithm, Volume, and others depending on the user preferences.

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	ACH B0-103	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

-QC checking the compounds extracted by MFE is important. The user should look for compounds with low volume by sorting the table by Vol, and looking at the EIC for each low volume compound. The user should 1) make sure that the peak in the EICs are well above background noise, and 2) check for good isotope ratios in the mass spectra.

-Note that Find Compounds by MFE is the default Compound Finding algorithm performed as part of an Unknowns Analysis Workflow, to be described later.

Find Compounds> Find by Auto MS/MS

-This compound finding tool requires a data file that was acquired by Auto MS/MS. This algorithm mines data from the TIC and any ion identified at the MS1 level as a precursor having MS/MS product ions will be considered in the compound-finding algorithm. This tool will return compounds with MS1 and MS/MS level data.

-This tool does not extract anything, it only mines the data. This tool simply filters the data for display.

-Processing tab (filtering only): can set thresholds for compound identification, including a retention time window around MS1 precursor ions, MS/MS TIC thresholds, mass match tolerance, and limiting to largest # compounds. There are also Compound fragment filters and Persistent background compounds.

-Excluded Masses tab: option to exclude masses and include a window around excluded masses.

-Results tab: option to delete previous compounds, highlight compounds, Extract EIC, Extract ECC, Extract MS (with option to subtract current background spectra), Extract MS/MS, and extract separate MS/MS spectrum per collision energy (recommended).

-Remember that all 'extraction' settings were set as part of the Data Acquisition (DA) method, and the options above only apply to filtering for display.

Find Compounds> Find by Targeted MS/MS

-This compound finding tool requires a data file that was acquired by Targeted MS/MS. This algorithm mines data/filters data for display from the data collected in the DA method. This tool will return compounds with MS1 and MS/MS level data.

-Integrator tab: user can select integrator algorithm. Agile 2 is recommended.

-Processing tab: set chromatogram peak width and option to limit to the largest compounds.

-Cpd TIC Peak Filters: user can set filter based on Peak height or Peak area absolute or relative to largest peak; option to limit to largest # peaks.

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

-**Peak Spectrum tab:** users can select how to create spectra from compound peaks included based on filtering parameters above. It is common to average scans >10% of peak height and to exclude if above 10% of saturation. User can set m/z ranges or use the default (m/z range of the chromatogram) and to never return empty spectrum. Option to subtract peak spectrum background and to define background in various ways. Default is to select None.

-**Results tab:** Option to delete previous compounds, highlight results, and extract MS/MS chromatogram, Extract MS spectrum, and Extract separate MS/MS spectrum per collision energy (these options are recommended).

-Remember that all 'extraction' settings were set as part of the Data Acquisition (DA) method, and the options above only apply to filtering for display.

Find Compounds> Find by Integration

-This tool will look at and integrate a TIC. It will find the predominant ion in each peak and extract the ion information from the peak. This tool includes many tabs and settings for data extraction and filtering.

-This method is not very useful unless you have pure compounds with unknown peaks... ie. very rare!

When to change "Find By" settings?

-Anything that changes ion abundances (especially isotope abundances)

-Changes in Sample Preparation

-Changes in Acquisition Rate or Number of Transients

-Changes in Source conditions

-Changes in Retention Time (LC solvents, gradient, etc.)

-Changes in Column, even if the new column has the same lot #

Find Compounds by Formula> Find by Formula – Options

-This tool uses MS1 level data ONLY, and therefore can be used with any data type because all QTOF modes collect MS1 level data.

-The User has to tell the Find by Formula what to look for by including formula information including Name, Molecular Formula, and Mass (based on formula).

-**Formula Source tab:** user must choose a source of formulas by entering formulas manually, attaching a .CEF file, by selecting an available Database or Library with compound information, or selecting Worklist (this option only applies if formula information was included in the DA worklist). User can select the

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	ACH B0-103	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

maximum number of formula matches and by default the automatically increase for isomeric compounds option is checked.

-Formula Matching tab: Match tolerance options allow user to select ppm window for matching to formula. Chromatogram extraction options allow user to select window around the mass to extract the EIC. Note that a formula/mass needs to be extracted in order to get a match score. The closer to the edge of the chosen window the lower the match score will be.

-Positive and Negative Ions tabs: user should select what is likely to be detected. Select all positive ions if DA was performed in positive polarity and charge state 1-2 for small molecules.

-Scoring tab: this tab affects what weighting is given to formula matching scoring parameters. Best not to change the default scoring or Expected data variation numbers. If a user doesn't know what the default values are, please load the 'default.m' method. The Mass score is based on the proximity of mass to the theoretical mass. Scores for isotope abundance and isotope spacing are based on the composition of the formula. RT matching may not be beneficial unless a library obtained using the exact same LC parameters as the current run is being used for Find by Formula. Can zero this weight if you don't wish to include RT in score. The Overall Score is calculated as follows: $SUM(\text{weight} \times \text{score}) / SUM(\text{raw scores})$.

-Results tab: user can select what to do when algorithm is finished, including Delete previous compounds, Highlight first compound, Extract EIC, Extract cleaned spectrum, and Include structure (if available). User can Extract MS/MS spectrum and extract per CE (recommended) with additional precursor tolerance settings (if MS/MS data was acquired). Note this is just bringing MS/MS data along for the ride, it is not used in extraction or Find By algorithm. User can finally choose to Extract raw spectrum and select Prefer profile data if available, and to Clip extracted raw spectra outside of a given mass spectral window (ppm). This is encouraged to see nice, clean mass spectra for Formulas/Compounds.

-Result Filters tab: use this to set display options for Found Formulas, including Only generate cpds for matched formulas, setting Warn if score and Do not match if score for formula matches, and Single ion matching Warn If and Do not match if second expected ions should be above a certain number of counts.

-Fragment Confirmation tab: this tab is only applicable to an "All Ions" workflow.

-After performing Find by Formula tool a Compound Table will be created. User should make sure important columns are showing in the table. In the Target Screening category, select Flags, Mass, Score. In the Compound Identification category, select Name, Formula, m/z. Note that the 'Flags' column will give information about why the flag is there when the pointer hovers over the flag.

-Note that Find Compounds Formula will use MS1 ions only and match them to theoretical ions from formulas found in a library, database, .CEF file, or an input list of other formulas. This is much less useful

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	ACH B0-103	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

than running Find by Molecular Feature algorithm to create compounds from actual data collected, and then processing the compound list by Searching a Database/Library, which is part of the Unknowns Analysis Workflow (see below).

Identify Compounds> Search Database

-This tool will take all the information in the compound table and compare it to available Database(s). Note that these should be Agilent Databases, as online databases were not acquired on an Agilent QTOF.

-Remember that Databases contain only MS1 level data, no MS/MS data, and may or may not contain RT.

-Matches and scores based on accurate mass match, isotope abundances, and isotope ratio matching.

-Search Criteria tab: User should set Values to match as 'Mass' (this is the strength of a QTOF), and set a match tolerance of low ppm (5 ppm is standard). The option to match by RT is almost never useful and presents a large barrier to correct matching in most cases.

-Note that RT matching here might have utility if running QC samples as part of a method SOP that would use a home-made Database or Library with accurate RT information for that method.

-Note that RT match could also be useful for isomer identification if isomers are present in data file.

-Database tab: User selects a Database path.

-Peak Limits tab: This is related to spectral peaks and the default is 5. This means that up to 5 isotope peaks will be searched for each molecular species [M+H]⁺, [M+Na]⁺, etc.

-Ions tabs: set expected ions and charge states. User should select all positive ions if running in positive polarity and should select charge range of 1-2 if doing small molecule applications.

-Scoring tab: do not change default weighting for scoring, unless you wish to remove RT matching from scoring (common if RT will not match, user should set it to '0'). Do not change expected variation values.

-Search Mode tab: user can select how to search database for neutral or ionic entries when searching masses from simple ions. Ionic is default.

-Search Results tab: user can limit search to top # of hits, based on overall score.

-User will need to select specific compounds or the full compound list before pressing Search Database for Compounds in Method Editor.

-User will look at Compound List table to see if the compound Name was populated along with Score (DB), ID Source, Hits (DB).

-If MS1 level data is sufficient for identification and/or was the only data collected, then a Database Search is the best way to identify compounds, and may be part of an Unknowns Analysis Workflow.

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

Identify Compounds> Search Library

-Libraries contain both MS1 level data and MS/MS level data, often at multiple standard CE values (10, 20, 40V).

-It is best practice for users to collect MS/MS level data at these three CE values. This must be setup at the DA method stage before data collection (see DA methods SOP).

-Scoring uses the mass accuracy and relative intensity of ions. The more accurate the m/z and the more reproducible the intensity (isotope ratios), the higher the match score will be.

-Library search can be performed as both a Forward Search and a Reverse Search. A Forward Search has a lower threshold because it forces the sample spectrum onto the library spectrum. A Reverse Search has a higher threshold because it forces the library spectrum onto the sample spectrum, and the library spectrum is always much cleaner than the sample spectrum.

-Libraries tab: user can select one or more libraries to search from and if multiple are chosen, can select to search all or stop at first library match. User sets max number of hits per compound (if compounds match multiple library entries).

-Search Criteria tab: user should restrict spectral comparison based on Instrument type and if collision energies in data file match library, user can also select Restrict based on Collision energy. User can exclude precursor ion from Reverse Score when ratio of precursor intensity to total library MS/MS intensity exceeds 90%. This is a way to remove a dominant ion that was resistant to fragmentation from over-influencing the Reverse Score by removing it if it is very high abundance. The 90% default is reasonable for most users.

-RT Scoring tab: user can set RT matching here, if desired. This is often not used, but can be useful if the library was created using the same LC gradient as the data file, or for QC analysis.

-Peak Filters tab: user can limit spectral ions to those with a minimum abundance and set the max # largest peaks.

-Tolerances tab: used to set precursor ion and fragment ion mass expansion windows for mass accuracy. Note that fragment ions have lower mass accuracy than precursor ions in the QTOF, and this is why the expansion window is larger.

-Remember to show desired columns in the Compound Results table, including ID source, Library Score, Library Name, etc.

-Note: When doing small molecule work, changing the collision energy often does not affect MS/MS product ion abundance more than 20-30%. Running an internal standard should always match the MS/MS fragmentation pattern and abundances precisely.

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	ACH B0-103	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

-Note that user can always run Database Search, Library Search, and Generate Formulas in a single Unknowns Analysis Workflow (see below), and even combine all three with weighting.

Identify Compounds> Generate Formulas

- This tool is completely theoretical and uses *in silico* predictions to predict molecular formulas (compounds) based on the mass, peak height, peak spacing, user-defined type, charge, elements, charge carriers, mass error, and probability score.
- MS/MS ions are used to improve the specificity of formula generation (if MS/MS data were acquired)
- Formulas can be generated for all compounds highlighted in the Data Navigator following a Compound Finding algorithm.
- Users are encouraged to acquire MS/MS data to improve the predictions and generate more discriminate formulas.

Unknowns Analysis Workflow (aka MS Target Compound Screening Workflow)

Almost all of the settings applied in the MS Target Compound Screening Workflow in the Method Editor rely on the settings in the corresponding sections elsewhere in the Method Editor, such as the 'Find By' and 'Identify Compounds' algorithms. Therefore, the user must understand these options as outlined above.

-If the Method Editor does not show a drop down item called MS Target Compound Screening Workflow, then the user needs to go to Configuration> Configure for Workflow> MS Target Compound Screening. A pop up box will give options to Save current method, Reload current method, Load default method, Use current layout, or Load workflow's default layout. Users will often Save current method, Reload current method, and Use current layout in order to simply add this drop down item to the users current window and method. User can save this section by saving the method.

-Extraction Data Format: user should always select Profile when available, otherwise Centroid, because Profile always contains more information.

-Chromatograms: optional. User can select integration algorithm (Agile 2) and peak filters if peak integration is performed as part of a File Open Action or Worklist Action.

-Mass Spectra: optional. User can select spectral peak extraction settings if spectral peak extraction is performed as part of a File Open Action or Worklist Action.

-Find by Molecular Feature: this will be the most useful algorithm for finding compounds in an unknowns analysis. User will choose many settings that impact how molecular features (compounds) are extracted

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

from MS1 data. MS/MS level data is not used in the algorithm and is only along for the ride. Users should refer to the detailed description of this algorithm found earlier in this document.

-Find by Formula: this algorithm is not as useful as Find by Molecular Feature for compound extraction, which uses much more detailed information. User can choose many setting that impact how this algorithm extracts compounds from a known list of formulas (excel, database, or library). Users should refer to the detailed description of this algorithm found earlier in this document.

-Identify by Database Search (MS1 data only): user selects one (1) database path to be used for database searching during compound identification. User should select to match by Mass as the default. User can include RT matching if using a database generated under the same LC conditions as the sample.

-Identify by Library Search (MS/MS data): user can select multiple libraries to be used for compound identification, and set forward and reverse matching scores, search criteria, peak filtering, and tolerances for precursor and fragment ions. See detailed descriptions found earlier in this document.

-Identify by Formula Generation: this algorithm is not very useful because it produces completely theoretical formulas. However, this option is often included in the Unknowns Analysis Workflow because it can generate possible formulas for compounds that were not identified by Database or Library searches in the workflow.

-Match Scoring: this section includes the score weighting and data variation used in the Identify by Formula Generation in silico prediction algorithm. User should not change the defaults.

-Compound Report: this section is used to specify Reporting options. User can specify many different options depending on what data records are desired. Microsoft Excel and PDF-based reporting are available, and many report template options can be selected.

-Automation: this tab facilitates execution of various automations, including File Open Actions, Worklist Actions, Print Compound Report, and Run Compound Automation Steps.

>File Open Actions and Worklist Actions tabs allow the user to specify various combinations of actions that will be run in sequence from top to bottom when executed. For example, a user can include Extract Defined Chromatograms, and Integrate and Extract Peak Spectra, and save this in a Qual. method (ex. "Unknowns_Analysis_Workflow.m" method file). When opening a new file the user can load this workflow method (.m) and then execute File Open Actions, and the settings saved in the method editor for the individual actions of extracting chromatograms, integrating chromatograms, and extracting peak spectra will be automatically applied to the data file.

>Similarly, the user can setup and execute Worklist Actions as desired. Worklist actions are similar to File Open actions except they can be restricted to certain RT time or ranges using the Worklist RT Ranges tab.

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

>Users can execute the Print Compound Report automation after all other Unknowns Analysis automations have been performed.

>Users will execute the Run Compound Automation Steps automation to perform the real 'unknowns analysis workflow'. The automation settings are specified in the Analysis Options tab. Compound mining drop down list allows the user to specific one (1) Compound Finding algorithm. It is not possible to perform multiple compound finding algorithms in a single automation. The default for unknowns analysis is Find by Molecular Feature. User can then select multiple Compound Identification tools. User should select Search a Database, Search a Library, and Generate Formulas – Only compounds without database hits. This way, any unidentified compounds can have a theoretical formula predicted for possible user follow up. User can select to Show only identified compounds, but by default this is unchecked.

Unknowns Analysis Workflow Protocol

1. Import Data File(s) of interest.
2. Open Method (.m) file and ensure MS Target Compound Screening drop-down is open in Method Editor.
3. Check settings in the appropriate sections involved in the automation steps.
 - a. Find by Molecular Feature (most common Compound mining algorithm, uses MS1 data)
 - b. Identify by Database Search
 - c. Identify by Library Search
 - d. Identify by Formula Generation
 - e. Optional – File Open Actions & Worklist Actions require additional settings to be modified in the method editor
 - f. Optional – Compound Report setting can be modified if a report in Excel or PDF is needed after Finding and Identifying Compounds
4. Optional – Run File Open Actions on data files
5. Optional – Run Worklist Actions on data files (can restrict by RT)
6. In Automation> Analysis Options tab:
 - a. Select Find by Molecular Feature
 - b. Check the box for Search a database for each compound
 - c. Check the box for Search a library for each compound
 - d. Check the box for Generate formulas for each compound
 - i. Check the radio button for Only compounds without database hits
7. Click 'Run Worklist Actions Now' to execute the Unknowns Analysis Workflow.

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____

Faculty/ Department or Institute/Centre	Cumming School of Medicine, Department of Pediatrics, Neurocritical Care Program, ACH BioCORE	Date Created:	January 4 2019
Location	<i>ACH B0-103</i>	Created By:	Matthew Rosin
Supervisor	Dr. Michael Esser & Dr. David Sinasac	Revision #:	1.0
		Revision Date:	January 4 2019
		Revised By:	Matthew Rosin

- a. This will find compounds by molecular feature extraction, then attempt to identify compounds by comparing them to a database (MS1 data only), a library (MS/MS level data), and if not found in a database, it will try to generate predictions of theoretical formulas matching the compound.
- b. Note that in Qual. B.07, there is an error in the automation that will stop the algorithm from generating predicted formulas at the first encounter of a compound that was already matched to a library. This may result in most if not all of the compounds not being subjected to formula generation, and an error message will pop up.
 - i. A time-consuming work-around for this is to look at the Compound List and select all compounds that did not have a library hit, and run the Identify by Formula Generation.
 - ii. We do not yet know why library-matched compounds cause errors in the formula generation algorithm.
8. The quality of the compound identification results can be evaluated in various ways:
 - a. Look at the overall Score, which represents a weighting of each identification algorithm.
 - i. Note that a low overall score might simply reflect that different compounds were identified in the Library vs. Database searches, and that the individual score for matching the library might be very high (>97%) and likely to be the correct match.
 - b. Look at the individual score for the Library Match and Database Match. A high score for library or database match is desirable and gives confidence in the compound identification.
 - c. If MFG was used, look at both the MFG score and the mass difference 'Diff (ppm)'. A high score and low mass difference is desirable, but the formula is still only a prediction.
 - d. For compounds that were matched to a Database, the user should use the Show Predicted Isotope Distribution tool in the MS Spectrum Results window when looking at the MFE spectrum and check to see how closely the ions match the predicted isotopes.
9. User should inspect compound results and if necessary modify the compound finding and identification settings such that the user has a high confidence in the compound list produced.
10. Compounds can be individually removed from the Compound List by right clicking in Data Navigator and selecting Delete.
11. Save the data file, which will now include all results and the Compound List.
12. Export the data file as a .CEF file in the same path destination as the .d data file.
13. The .CEF file can be imported into Mass Profiler or Mass Profiler Professional for basic or advanced analysis, respectively.
14. User can also create compound reports in excel or PDF format.

Printed copies are UNCONTROLLED unless signed by an authorized BioCORE lab personnel below:

Personnel Name: _____

Date Printed: _____