



MZmine 3

Documentation

The MZmine Community

None

Table of contents

1. Welcome to the MZmine 3 wiki!	4
1.1 Libraries in MZmine	4
2. Getting Started	5
3. Main window overview	6
3.1 MS data files and feature lists tab	6
3.2 Main content pane	6
3.3 Main menu	6
3.4 Task overview	6
3.5 Page Contributors	7
4. LC-MS Workflow	8
4.1 Raw data processing	8
4.2 Feature processing	8
4.3 Page Contributors	9
5. LC-IMS-MS Workflow Overview	10
5.1 Supported formats	10
5.2 Feature detection workflows	10
5.3 Graphical comparison of LC-MS and LC-IMS-MS data	11
5.4 Page Contributors	11
6. Module documentations	12
6.1 Data import	12
6.2 Mass detection	14
6.3 Mobility scan merging	16
6.4 ADAP chromatogram builder	17
6.5 IMS Expander	18
6.6 Ion mobility trace builder	19
6.7 Smoothing	21
6.8 Local Minimum Resolver	23
6.9 CCS Calibration and calculation	26
7. Visualisation	29
7.1 Raw data visualisation	29
8. Additional IMS resources	30
8.1 Ion mobility spectrometry terminology	30
8.2 Graphical comparison of LC-MS and LC-IMS-MS data	32
9. Performance options	33
9.1 Preferences	33

9.2 Logs	34
9.3 Command-line arguments	34
10. How to contribute	36
10.1 Contribute to the MZmine documentation	36
10.2 Creating a new page	38
10.3 Page Contributors	39
11. Acknowledgements	40

1. Welcome to the MZmine 3 wiki!

Here you can find general processing guides, module documentations and video tutorials.

Want to get started with MZmine 3? Check our [Getting started](#) page!

[Main window overview](#)

[Performance options](#)

[LC-MS workflow](#)

[LC-IMS-MS workflow](#)

[Raw data visualisation](#)

[Acknowledgements](#)

1.1 Libraries in MZmine

- [Apache XML Graphics](#) - EPS image export
 - [Chemistry Development Kit](#) - Isotope pattern and molecular calculations
 - [Freehep](#) - EMF image export
 - [Google Guava](#) - Utility classes
 - [JDK Documentation](#)
 - [JChemPaint](#) - 2D molecule visualization
 - [JFreeChart](#) - TIC, Spectra and 2D visualizers
 - [Jmol](#) - 3D molecule visualization
 - [jmzml](#) - mzML file import
 - [jmzTab](#) - mzTab file import and export
 - [NetCDF-Java](#) - NetCDF file import
 - [VisAD](#) - 3D visualizer
 - [WEKA](#) - Clustering and other machine learning algorithms
 - [Bruker TDF SDK](#) - Native tdf/tdf file import (requires VC++ 2017 redist.)
 - [Thermo raw file parser](#) - Native Thermo raw import
-

Last update: February 18, 2022 17:51:32

2. Getting Started

Before creating your first project, we recommend setting up some things.

1. Set a temporary file directory. Go to *Project* → *Set preferences* → *Temporary file directory*. This requires a restart to take effect.
 - a. We recommend setting the directory to an SSD with enough space for fast processing and visualizations.
 - b. On Windows, old temporary files are deleted when a new session is started.
2. MZmine 2 projects cannot be imported due to changes in the data structure.
3. MZmine 2 batch files cannot be imported due to parameter optimizations.

You can familiarize with the new GUI here: [Main window overview](#)

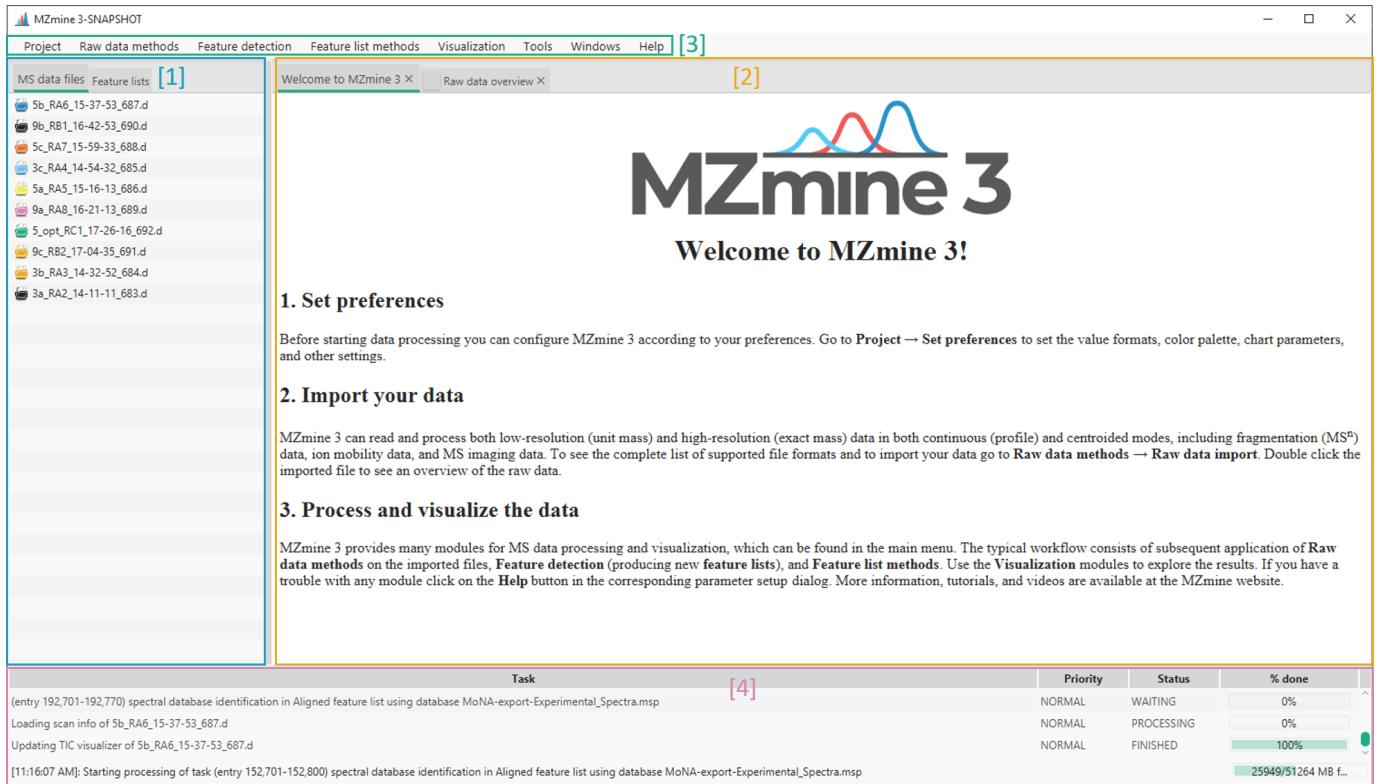
A quick insight to data processing workflows can be found here: [LC-MS workflow](#) or [LC-IMS-IMS workflow](#)

You can also check out the new processing wizard under *Processing wizard* in the main menu.

Last update: February 18, 2022 17:07:50

3. Main window overview

The MZmine 3 main window is made up of mainly four important building blocks.



3.1 MS data files and feature lists tab

[1]: The (raw) ms data and feature list tabs. Here you can find your imported data files and processed feature lists. *Hint: you can also import files by dragging & dropping them to the ms data tab.*

3.2 Main content pane

[2]: The main content pane. Visualisations such as a raw data overview or a feature list can be viewed here. This pane can also contain multiple tabs. Every tab can also be opened in a new separate window by right-clicking on the header.

3.3 Main menu

[3]: The main menu. Here you can find methods to import and process your data files and feature lists and visualise the results. Furthermore, projects can be saved and preferences can be set.

3.4 Task overview

[4]: The task overview. Current tasks are displayed and their status and progress are indicated. Tasks can also be canceled by right clicking on a task.

3.5 Page Contributors

SteffenHeu

Last update: January 18, 2022 21:51:26

4. LC-MS Workflow

The workflow proposed herein is intended as a general pipeline for untargeted LC-MS (or LC-MS/MS) data preprocessing. The main goal is essentially to turn the highly-complex LC-MS raw data into a list of features, and corresponding signal intensity, detected across the analysed samples. Such feature lists can then be exported for further downstream analysis (e.g., identification, search against spectral libraries, statistical analysis, etc.). A schematic representation of the workflow is shown below:



References: - 10.1039/9781788019880-00232 - 10.1007/978-1-0716-0239-3_3 - 10.1016/bs.coac.2018.08.003

4.1 Raw data processing

The raw data processing consists of essentially two steps: [Data import](#) and [Mass detection](#)

4.1.1 Raw data import

Either open (e.g. mzML) and native vendor (e.g. Thermo, Bruker) data formats can be imported in MZmine 3. All the supported formats are listed here ([LINK to Doc](#)). For more details see the [Data import](#) module.

4.1.2 Mass detection

This step produces a list (referred to as "mass list") of the m/z values found in each MS scan across the LC run that exceed a user-defined threshold (i.e. noise level). For more details see the [Mass detection](#) module.

4.2 Feature processing

The goal of the "Feature processing" is to obtain a list of all the detected features (characterized by a RT and m/z value) from the raw LC-MS data.

4.2.1 Chromatogram building

The first step in the "Feature processing" is to build the so-called extracted ion chromatograms (EICs) for each detected mass (see "Mass detection"). There are two modules in MZmine 3 that can fulfil this task: [ADAP chromatogram builder](#) (widely used) and [Grid mass](#) (create docs).

The "detected" features in each file are listed in the so-called "feature lists", which are then further processed.

(e.g. to) and aligned to connect corresponding features across all samples.

4.2.2 Smoothing in retention time dimension (optional)

- Optional, depends on the LC peak shape
- For more details see the [Mass detection](#) module.

[Smoothing](#)

4.2.3 Feature resolving

[Local minimum resolver](#)

4.2.4 De-isotoping

- Remove features produced by natural isotopes
- For more details see the De-isotoping module (add link).

4.2.5 Gap-filling

Gap-filling can be performed on the aligned feature lists to cope with missing features that might be artifacts of the feature-detection process. - For more details see the Gap-filling module.

4.3 Page Contributors

[SteffenHeu](#) (35.85%), [tdamiani](#) (64.15%)

Last update: February 18, 2022 15:06:54

5. LC-IMS-MS Workflow Overview

Compared to regular LC-MS, LC-IM-MS data is more complex due to the additional separation dimension. Since some terms might not be straightforward for new users, a basic explanation of IM separation principles and the terminology used within this documentation is provided [here](#).

5.1 Supported formats

- Vendor formats:
 - .tdf (Native Bruker LC-IMS-MS and MALDI-IMS-MSI format)
 - .tsf (Native Bruker MALDI-IMS-MS (single shot) format)
 - .mzML
 - Created via [MSConvert](#) from native Bruker data
 - Created via [MSConvert](#) from native Waters/Agilent data
-

5.2 Feature detection workflows

Ion mobility data can be processed in MZmine 3 in two ways. The first few steps are different for the two workflows (see below).

1. [LC-IMS-MS workflow via ADAP Chromatogram builder and IMS expander \(recommended\)](#)
2. [LC-IMS-MS workflow via Ion mobility trace builder / Recursive IMS builder](#)

While these lists might seem fairly similar, there are some differences in the processing approach. The LC-IMS-MS workflow builds ion mobility traces from the data in the mobility scans, whilst the LC-MS workflow builds EICs from the summed frames. Imagine setting the noise level to 200 and occurs in 15 mobility scans with an intensity of 180 and in 20 with > 200, but the ion mobility trace builder required at least 25 data points. At the same time, the ADAP chromatogram builder was set to a minimum highest intensity of 1000 in a frame, which might have been reached due to summed intensities.

Since intensities are summed to build frames, the LC-MS workflow can be more sensitive to less abundant compounds whilst the LC-IMS-MS workflow will directly detect ion mobility traces and be less susceptible to noise. For .mzML data, this might not make a significant impact, because the frame is built from the thresholded mass list. This means that the main impact on sensitivity for low intense compounds is made by the user's noise level selection. However, for native Bruker raw data, the summed frame is automatically built via the functionality of the vendor library without the user's influence. Therefore, the more low abundant compounds might be detected, if the LC-MS workflow is chosen.

5.2.1 LC-MS workflow (recommended)

LC-IMS-MS data can also be processed via the regular LC-MS modules. If necessary, detected features can be expanded into the mobility dimension.

For this workflow, generation of summed frame spectra via the [Mobility scan merging](#) module is a mandatory step, if the data was imported from an .mzML file (automatically generated via native Bruker import).

- [Data import](#)
- [Mass detection](#)
- [Mobility scan merging \(mzML data\)](#)
- [ADAP Chromatogram builder](#)
- [Smoothing in retention time dimension \(optional\)](#)
- [Resolving in retention time dimension](#)
- [Expanding EICs in mobility dimension](#)
- [Smoothing in mobility dimension \(optional\)](#)
- [Resolving in mobility dimension](#)
- [Smoothing in rt and mobility dimension \(optional\)](#)

- Some recognised features might have rather noisy signals (in rt and mobility dimension) after the mobility resolving step. If smoother shapes are required, the smoothing can be reapplied afterwards. In that case, smoothing can be applied to both dimensions at once.

5.2.2 LC-IMS-MS workflow

The LC-IMS-MS workflow will directly build [ion mobility traces](#) from the raw data in the mobility scans. This workflow does not necessarily require summed frame spectra. However, if extracted ion chromatograms shall be visualized via the [Chromatogram visualizer](#), the frame intensities are used. In case these are not present, the chromatograms will be blank. Note that feature intensities from the LC-IMS-MS workflow might not exactly match the frame chromatograms due to summing being executed prior to thresholding (for native Bruker data). Furthermore, multiple isomers might hide behind a single chromatographic peak.

- [Data import](#)
- [Mass detection](#)
- [Ion mobility trace builder](#)
- [Smoothing in retention time dimension \(optional\)](#)
- [Resolving in retention time dimension](#)
- [Smoothing in mobility dimension \(optional\)](#)
- [Resolving in mobility dimension](#)
- [Smoothing in rt and mobility dimension \(optional\)](#)

- Some recognised features might have rather noisy signals (in rt and mobility dimension) after the mobility resolving step. If smoother shapes are required, the smoothing can be reapplied afterwards. In that case, smoothing can be applied to both dimensions at once.

5.3 Graphical comparison of LC-MS and LC-IMS-MS data

[Data comparison](#)

5.4 Page Contributors

[SteffenHeu](#) (96.88%), [tdamiani](#) (3.12%)

Last update: February 18, 2022 13:31:51

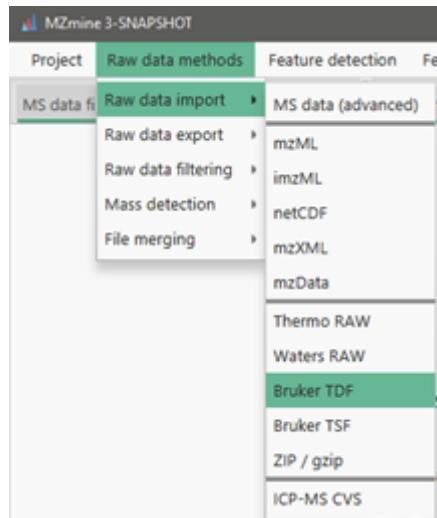
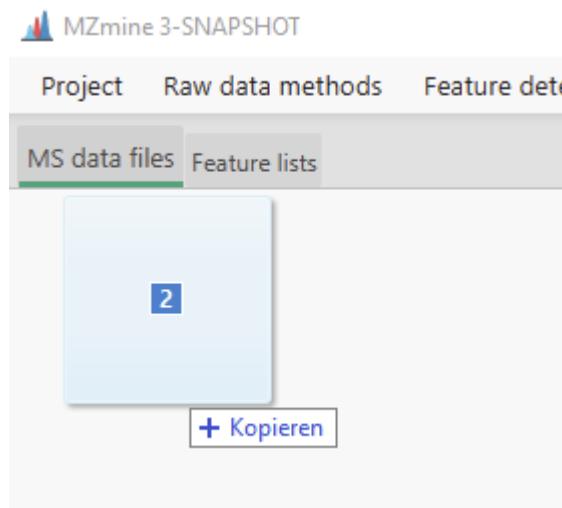
6. Module documentations

6.1 Data import

6.1.1 LC-MS data

Raw data can be imported via the main menu **Raw data methods → Raw data import**. Note that multiple data files/folders can be dropped into the **MS data (advanced)** dialog. If individual modules are used, folder based formats can only be imported as one folder at a time. When using the **MS data (advanced)** dialog, inexperienced users should deactivate the direct mass detection steps, since they alter the raw data on the import. Mass detection is then performed, when the scans are loaded and only peaks above the noise level are imported.

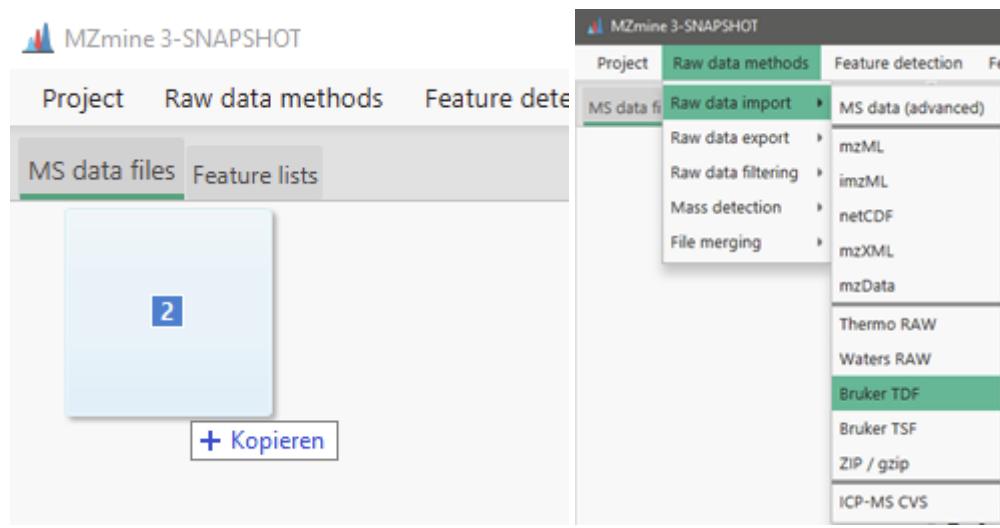
Alternatively, you can simply drag & drop the raw data into the raw data list of the main window.



6.1.2 LC-IMS-MS data

As any other data format, ion mobility data can be imported via the main menu **Raw data methods → Raw data import**. Note that multiple .tdf data folders can be dropped into the **MS data (advanced)** dialog. The Bruker TDF import can only select a single folder. When using the **MS data (advanced)** dialog, inexperienced users should deactivate the direct mass detection steps, since they alter the raw data on the import. Mass detection is then performed, when the scans are loaded and only peaks above the noise level are imported.

Alternatively, you can simply drag & drop the raw data into the raw data list of the main window.



Last update: January 18, 2022 21:51:26

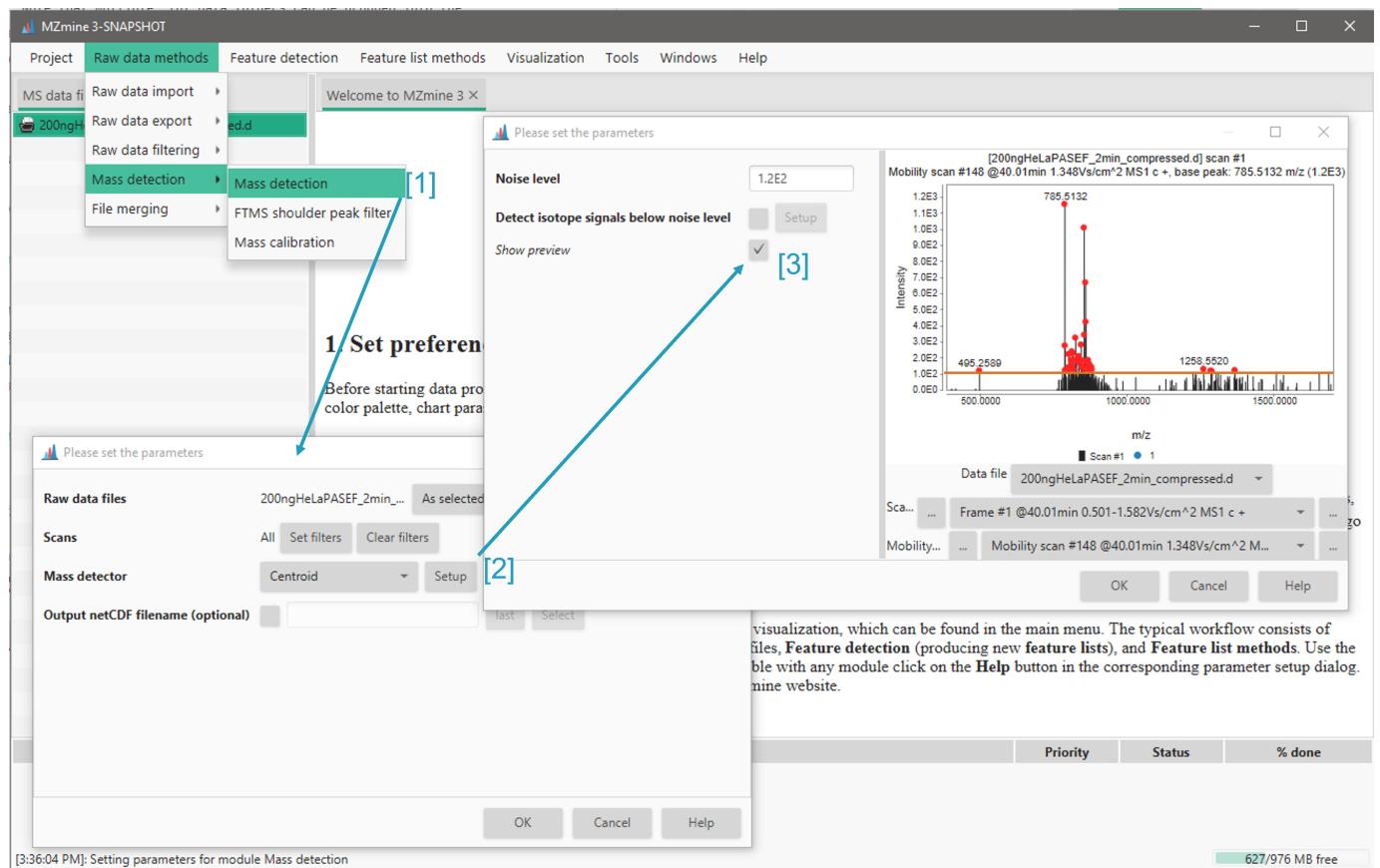
6.2 Mass detection

6.2.1 LC-MS data

The mass detection steps perform noise filtering (by a threshold) and centroiding of profile raw data. The raw data format can either be centroided or in profile mode. If the data is centroided, the **centroid** mass detector can be used. Profile data requires a different mass detector such as **exact mass**.

The mass detection is launched via **Raw data methods** → **Mass detection** → **Mass detection** ([1] in the figure). In the dropdown menu [2], an applicable mass detector should be chosen and configured via the **Setup** button [2]. By selecting the **Show preview** checkbox, a scan can be selected to adjust the noise level.

The output of the mass detection step, can be referred to as **mass list**, since it will only contain a list of selected m/z values.



SETTING THE NOISE LEVEL

Choose the noise level to detect (= red dot) actual peaks but filter out detector noise. The detector noise is usually determined by a lot of signals of the same intensity.

Detect isotope signals below noise level

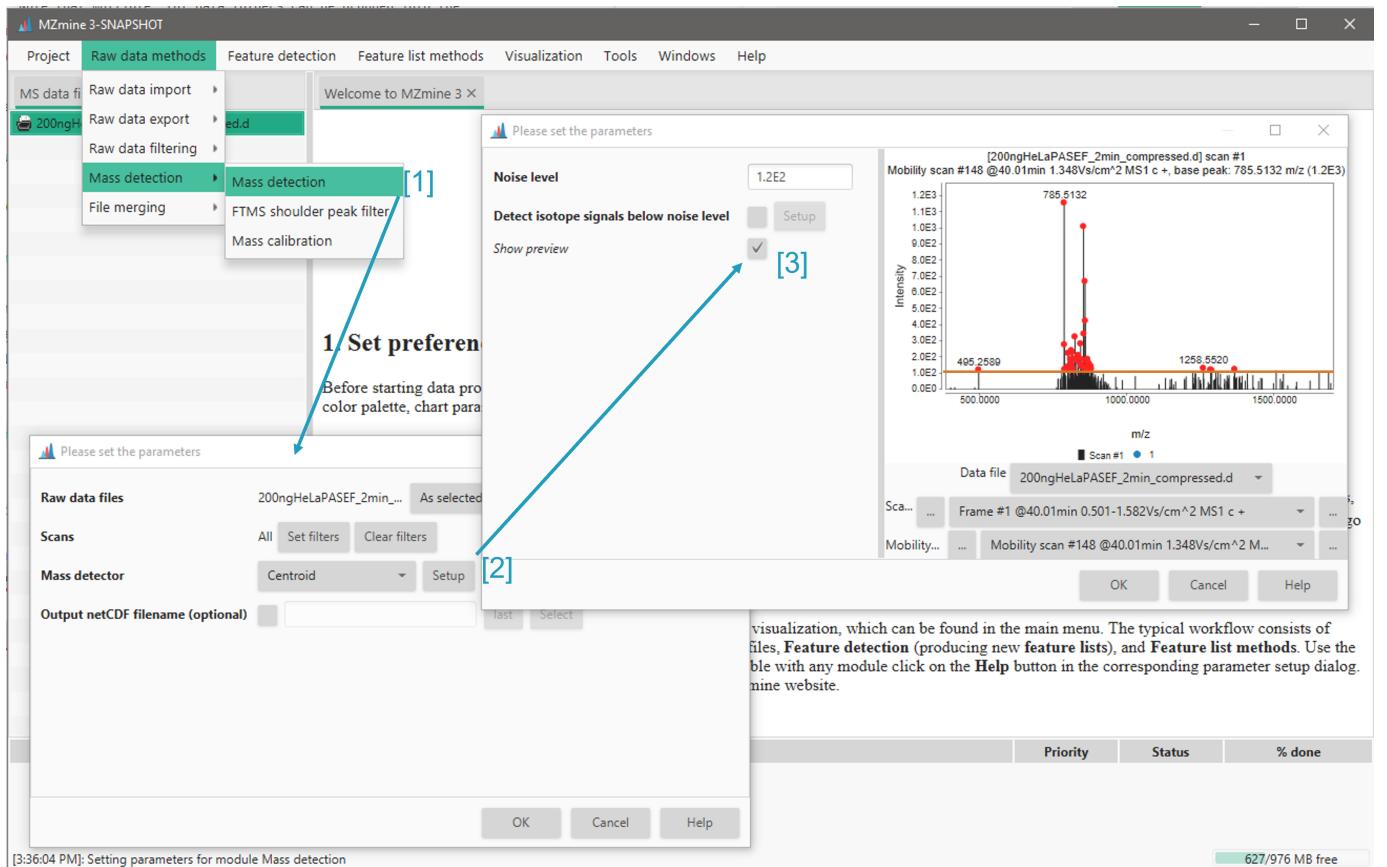
Centroid and **exact mass** detectors provide an additional parameter called **detect isotope signals below noise level**. This option allows to include additional peaks, corresponding to isotope masses, in the resulting mass list. By clicking the **Setup** button, you can specify the following parameters: * **Chemical elements** - chemical elements which isotopes will be considered (select from the periodic table). * **M/z tolerance** - the maximum allowed difference between theoretical isotope m/z and measured m/z present in the scan to be considered the same. * **Maximum charge of isotope m/z** - should be a positive integer (further denoted as K). All peaks having m/z equal to (up to m/z tolerance) theoretical isotope mass (any isotope of any specified chemical element) divided by any of 1, ..., K will be included in the mass list. The default value is 1.

6.2.2 LC-IMS-MS data

The mass detection steps perform noise filtering (by a threshold) and centroiding of profile raw data. Native Bruker raw data is already centroided, therefore the centroid mass detector should be used. Waters .mzML raw data might come as profile data, which requires a different mass detector such as **exact mass**.

The mass detection is launched via **Raw data methods → Mass detection → Mass detection** ([1] in the figure). In the dropdown menu [2], an applicable mass detector should be chosen and configured via the **Setup** button [2]. By selecting the **Show preview** checkbox, a scan can be selected to adjust the noise level. Note that a mobility scan should be selected to determine the noise level. However, the same noise level will be applied to frames, too.

The output of the mass detection step, can be referred to as **mass list**, since it will only contain a list of selected m/z values.



SETTING THE NOISE LEVEL

Choose the noise level to detect (= red dot) actual peaks but filter out detector noise. The detector noise is usually determined by a lot of signals of the same intensity.

Last update: February 18, 2022 13:40:55

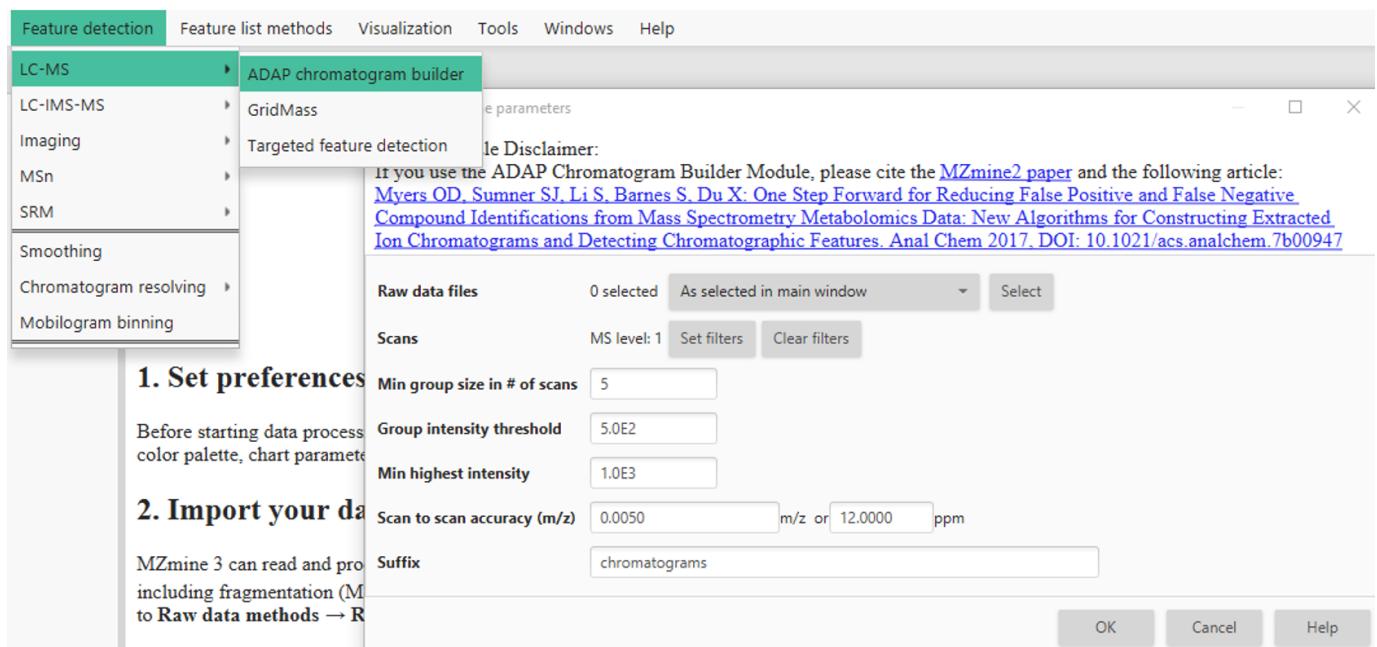
6.3 Mobility scan merging

If a .mzML file is imported, the merged frame spectrum must be created via the **File merging → Mobility scan merging** module. This is required to gain access to MZmine's regular LC-MS functionality. This step uses the centroided and thresholded data produced by the [mass detection](#) step.

This step is not required when importing native Bruker .tdf or .tsf data from .d folders. When importing native Bruker data, a merged spectrum for the frame is created automatically by the vendor library.

Last update: January 18, 2022 21:51:26

6.4 ADAP chromatogram builder



Scan selection

The scan selection parameter specifies the scans that shall be processed for feature detection. Usually, setting the ms level to 1 is sufficient. If a calibration segment is present, it can be cut out via the retention time filter in the scan selection.

Min group size in number of scans

This parameter specifies the number of consecutive detections of the same m/z value in a chromatographic peak (rt dimension). This means that a single m/z has to be detected in, e.g, 5 scans with an intensity higher than zero. This parameter helps to filter noise. Usually no less than 5 should be set here if the MS1 acquisition rate is sufficient.

Group intensity threshold

Specifies a minimum intensity that the number specified by **min group size** have to exceed. In this example, the intensity in at least 5 scans must be above 5E2.

Minimum highest intensity

The highest point of a potential EIC must exceed this value.

Scan to scan accuracy (m/z)

The **m/z tolerance** specifies the scan-to-scan tolerance for EICs. This tolerance depends on the mass accuracy and resolution of the instrument. Usually, a good starting point for optimisations are 0.005 and 5-10 ppm for Orbitrap instruments, while 0.005 and 10-15 ppm can be used for TOF instruments.

Last update: January 18, 2022 21:51:26

6.5 IMS Expander

The IMS expander will search for data points in mobility scans for existing features. This requires prior chromatogram building (see [ADAP Chromatogram builder](#) and resolving in retention time dimension (see [Resolving](#)).

Parameter settings

M/Z TOLERANCE

If selected, a tolerance will be applied to the feature's detected m/z while searching for data points in mobility dimension. Otherwise, the accepted m/z range is determined by the feature's m/z distribution in accumulated frame spectra.

Recommended setting: selected, 0.003 m/z and 15 ppm

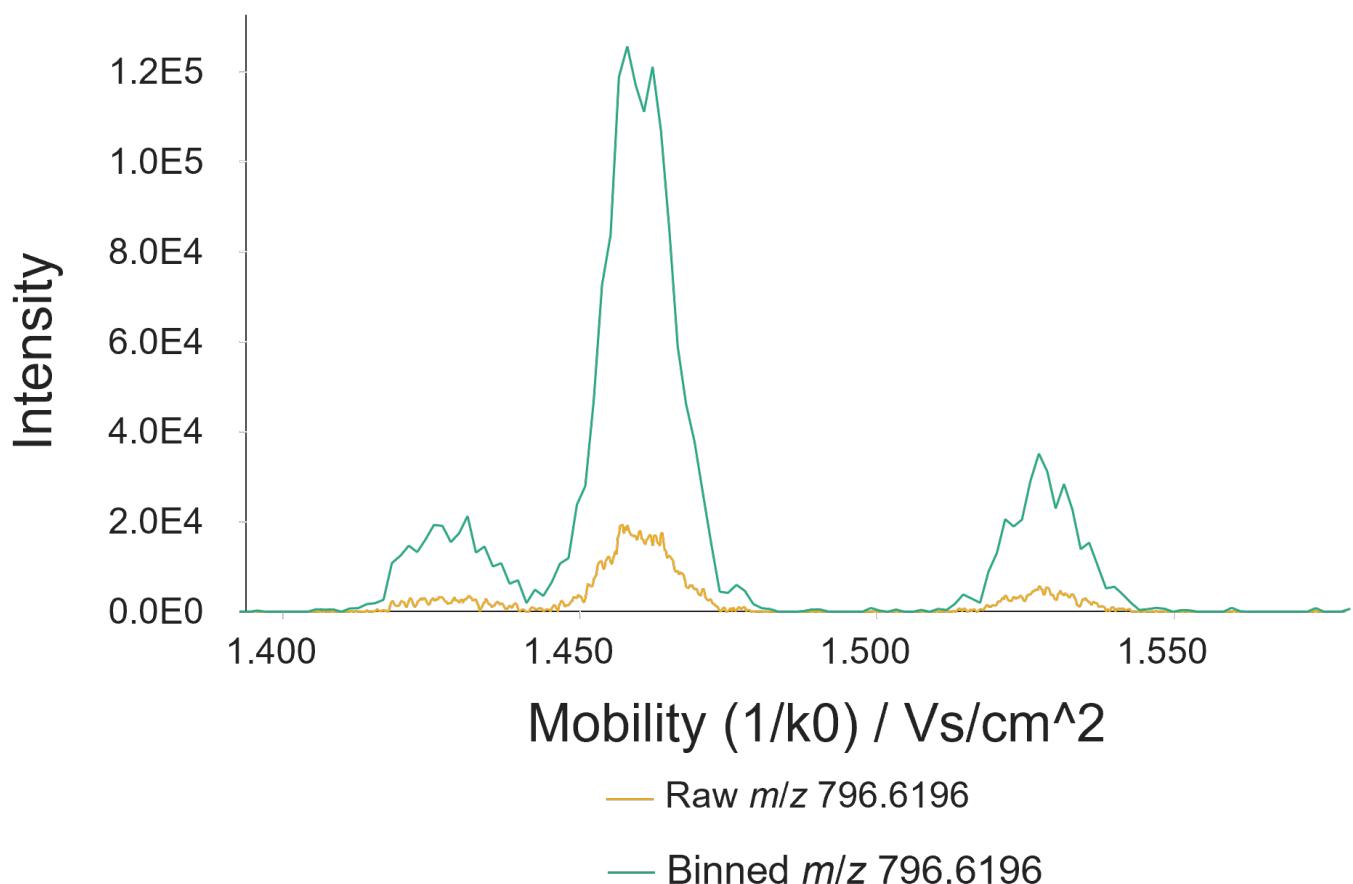
RAW DATA INSTEAD OF THRESHOLDED

Enables searching in mobility scan raw data instead of the thresholded (=mass detected) data. Only possible for centroid raw data files.

OVERRIDE DEFAULT MOBILITY BIN WIDTH (SCANS)

If selected, the default number of binned mobility scans can be overridden. Useful for data with high mobility resolution.

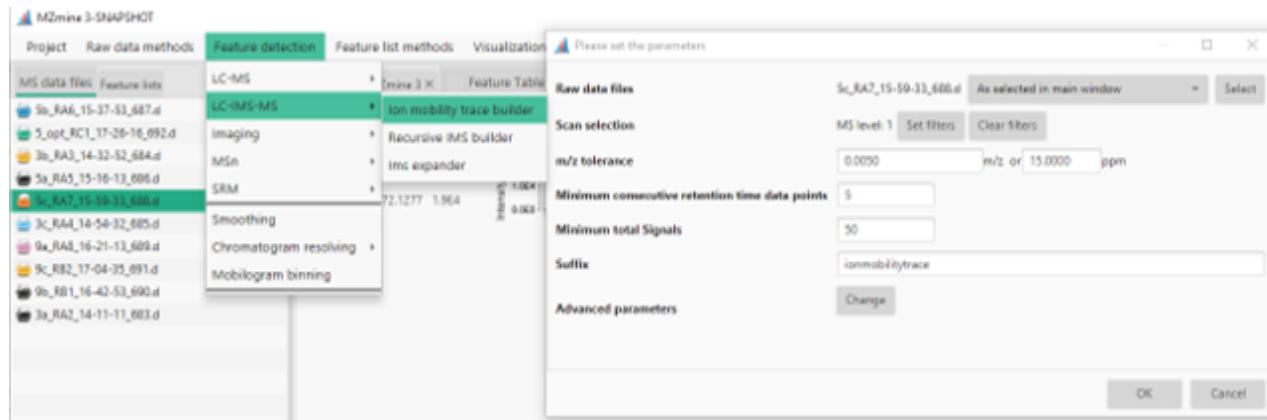
Binned mobilogram example



6.6 Ion mobility trace builder

The **Ion mobility trace builder** will build ion mobility traces from the raw data. Alternatively, the **Recursive IMS builder** can be used, which requires less ram but takes longer.

The **Ion mobility trace builder** is found under **Feature detection → LC-IMS-MS → Ion mobility trace builder**.



Scan selection

The scan selection parameter specifies the scans that shall be processed for feature detection. Usually, setting the ms level to 1 is sufficient. If a calibration segment is present, it can be cut out via the retention time filter in the scan selection.

m/z tolerance

The **m/z tolerance** specifies the scan-to-scan tolerance for ion mobility traces. This tolerance window may need to be set higher than for classic LC-MS feature detection (e.g. to 0.005 m/z and 15-20 ppm instead of 10 ppm) due to lower intensities therefore less accuracy in individual mobility scans compared to LC-MS scans. Note that the overall accuracy is achieved via LC-IMS-MS is the same due to the higher number of scans.

Minimum consecutive retention time data points

This parameter specifies the number of consecutive detections of the same m/z value in a chromatographic peak (rt dimension only). This means that a single m/z has to be detected in, e.g. 5 frames with an intensity higher than zero. This parameter helps to filter noise. Consecutive detections in the mobility dimension do not affect this parameter. Usually no less than 5 should be set here if the MS1 acquisition rate is sufficient.

Minimum total signals

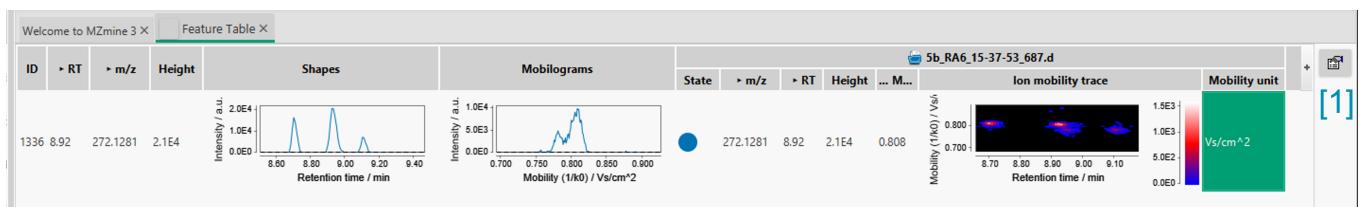
Specifies the total number of peaks in the mobility dimension in all mobility scans. Every "dot" in an ion mobility trace represents a single datapoint. (see [ion mobility traces](#))

Advanced parameters

For most applications, these parameters do not need to be set/changed. For high mobility resolved data the mobilograms might become noisy due to less ions reaching the detector at the same time. By default, the number of binned scans is set to cover about 0.0008 Vs/cm² per bin. The effect of binning can be seen [here](#). If you are unsure about the nature of your data, you can perform trace building with the standard parameters and apply/preview the binning afterwards via the **Feature detection → Mobilogram binning** module.

Processing result

After performing ion mobility trace detection, a feature list is created in the feature list tab (see [feature lists tab](#)). In the feature table, multiple columns are created. The displayed columns can be set via the button on the right of the feature table ([1]).



The **shapes** displays an EIC of the ion mobility trace (intensities summed in rt dimension). The **mobilograms** column shows a mobilogram for the ion mobility trace (intensities summed in mobility dimension). The shapes and projections can be smoothed and resolved. However, the ion mobility trace is always represented by the raw data and remains unaltered. After resolving, the shapes and mobilograms have to be recalculated from the raw data, which is why the smoothing is lost after resolving.

Last update: February 18, 2022 13:44:49

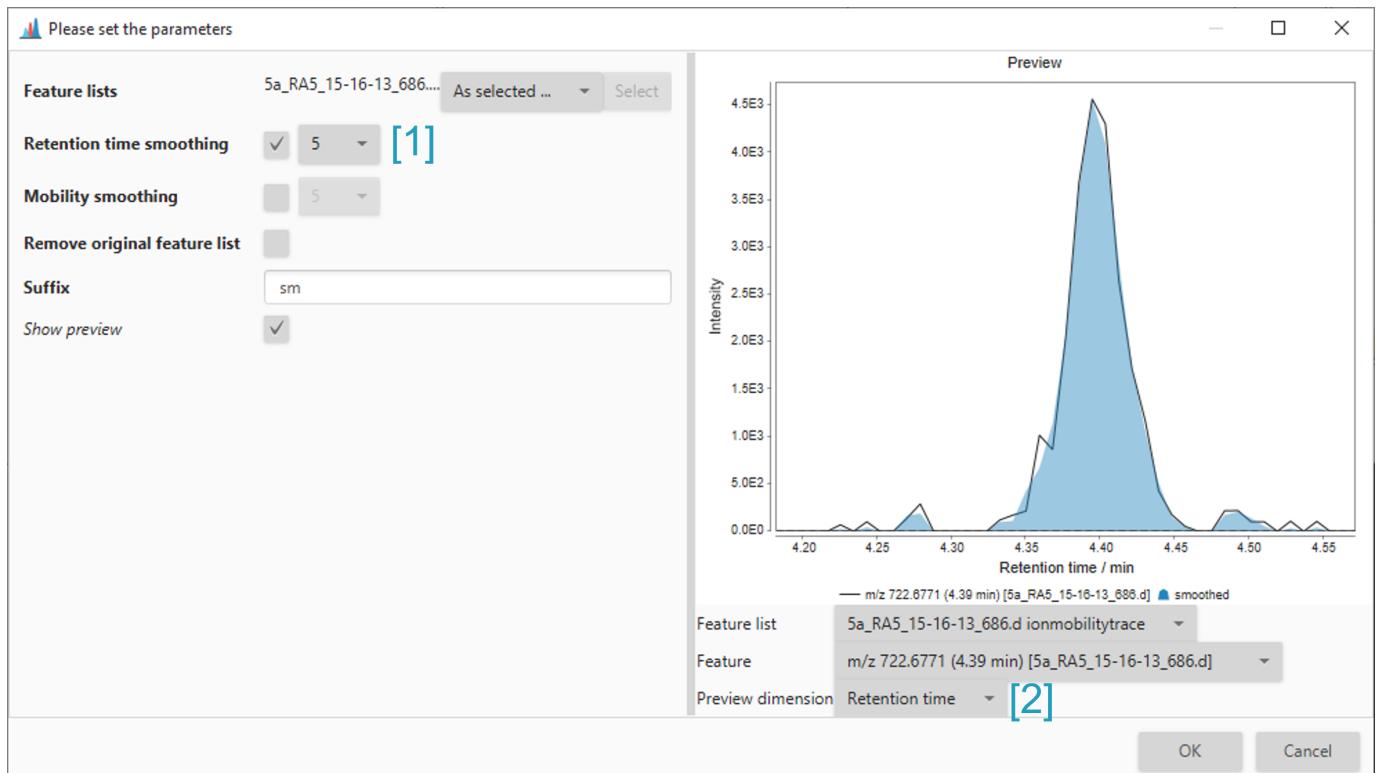
6.7 Smoothing

6.7.1 Retention time dimension

Smoothing chromatograms is optional. The necessity of smoothing in RT dimension is determined by the noisiness of chromatographic peaks. These can be influenced by the overall spray stability, instrument accumulation times, transfer efficiency and many more.

The number of data points to be smoothed in rt dimension can be set at [1]. Note that the correct preview dimension is selected at [2].

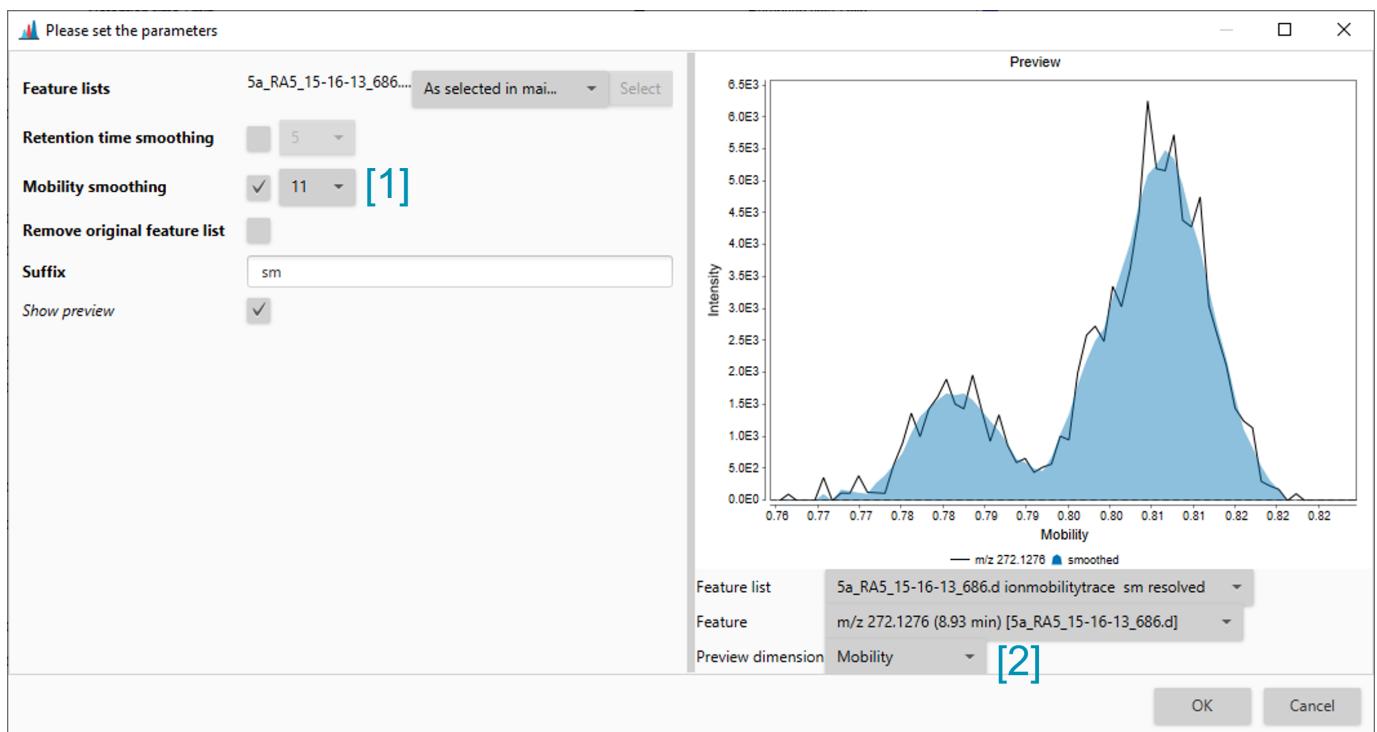
For large batch modes, the **Remove original feature list** parameter should be selected. While parameters are being optimised, this is not recommended, because removing a feature list cannot be undone.



6.7.2 Mobility dimension

After resolving a feature in RT dimension, the mobilograms will be recalculated from the raw data (the resolved ion mobility trace). Therefore, a smoothing step is necessary if the data requires it. The smoothing dialog is opened via **Feature detection → Smoothing**

Select to smooth the mobility dimension [1] and select it as preview dimension [2]. The filter with depends on the number of spectra acquired in the observed mobility range. Usually, a value between 5 and 15 should be appropriate.

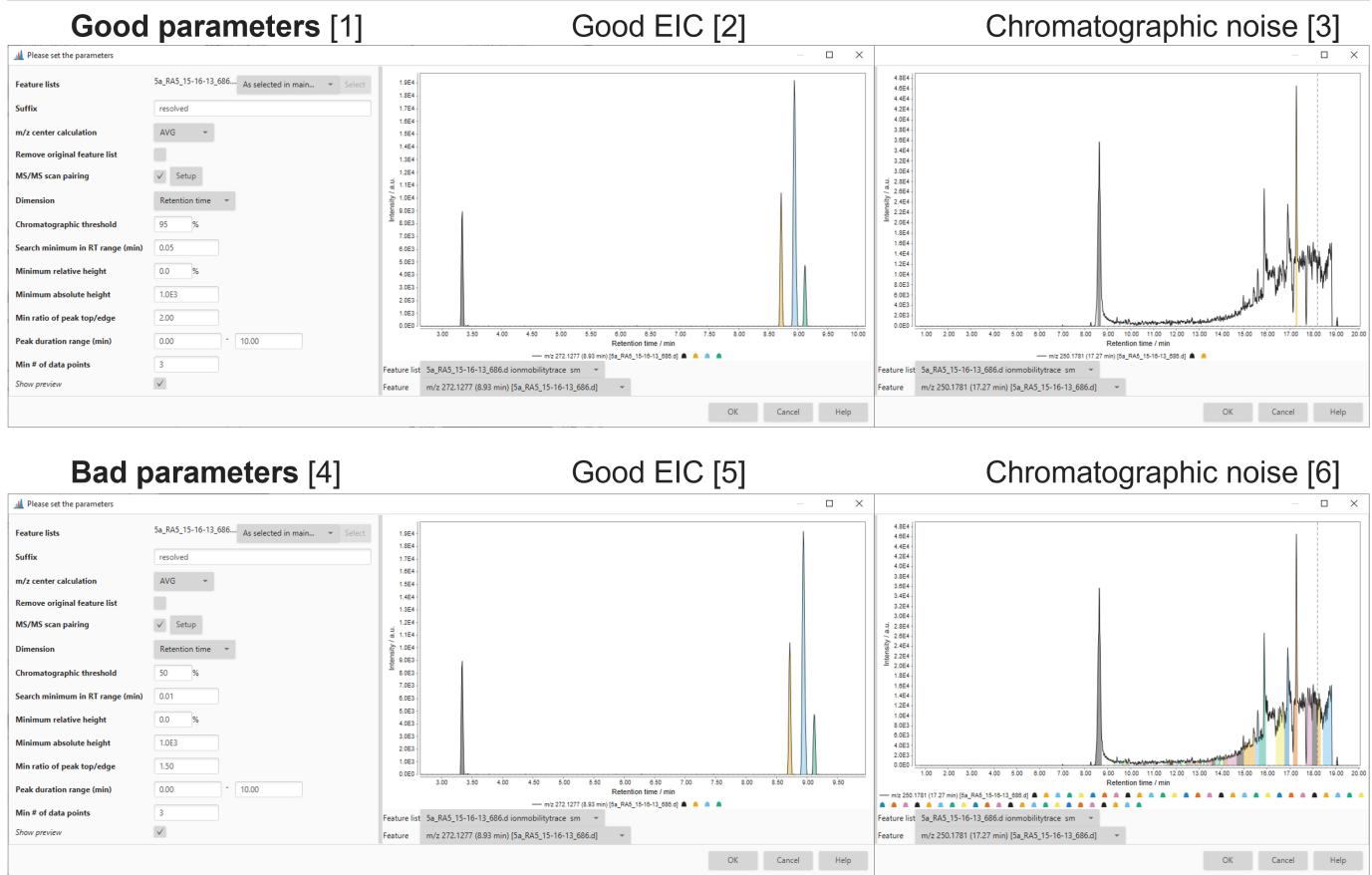


Last update: January 18, 2022 21:51:26

6.8 Local Minimum Resolver

Resolving traces/chromatograms into individual features, is one of the most **crucial** steps of data processing. Well optimised parameters [1] can lead to recognition of all good features in a "good" EICs [2] and to few noise recognised as feature in EIC that also contains chromatographic noise [3]. On the other hand, poorly optimised parameters [4] can still lead to recognition of all features in a good EIC [5], but recognise a lot of noise as feature in a noisy EIC [6].

Therefore, we recommend optimising the parameters on good EICs and checking the results of these parameters with a noisy EIC. Most of the time, a noisy EIC can be found by sorting the feature table with decreasing area.



Parameter settings

MS/MS scan paring

Selecting this parameter will pair DDA MS/MS spectra to the resolved features. This is optional at this stage, because it will be executed again during resolving in the mobility dimension.

Dimension

The dimension to be resolved can be selected here. Select *Retention time*.

Chromatographic threshold

This parameter is crucial for removing noise from chromatograms. If this parameter is set to, e.g., 50, the lowest 50 % of intensities will be removed. Since the *all* retention time in the data file are used for this determination, this value should be rather high (e.g., 95 %) to begin with and only lowered if necessary.

Search minimum rt range (min)

Determines the step size that will be scanned for individual peaks. Setting this value too low, can cause peak edges to be cut off, setting it too high might lead to incomplete separation of narrowly eluting compounds.

Minimum relative height

Determines the minimum relative intensity of a individual feature in relation to the highest intensity in the chromatogram. May lead to discrimination of low intensity features.

Minimum absolute height

Determines the minimum absolute intensity of a feature to be recognised by the algorithm. This parameter depends on what you want to detect, the instrument and detector type. Usually, Orbitrap instruments report higher intensities than TOF instruments. However, the noise level is also higher for Orbitrap than for TOF instruments. For TOF instruments 1E3 or even 5E2 can be appropriate, whilst Orbitraps can require 1E5 or 5E4.

Min ratio of peak top/edge

Describes the minimum ratio of the highest point of a peak to the lowest point of a peak. This mostly affects detection of low intensity and not-baseline-resolved signals and should be optimised using such a signal as an example.

Peak duration range

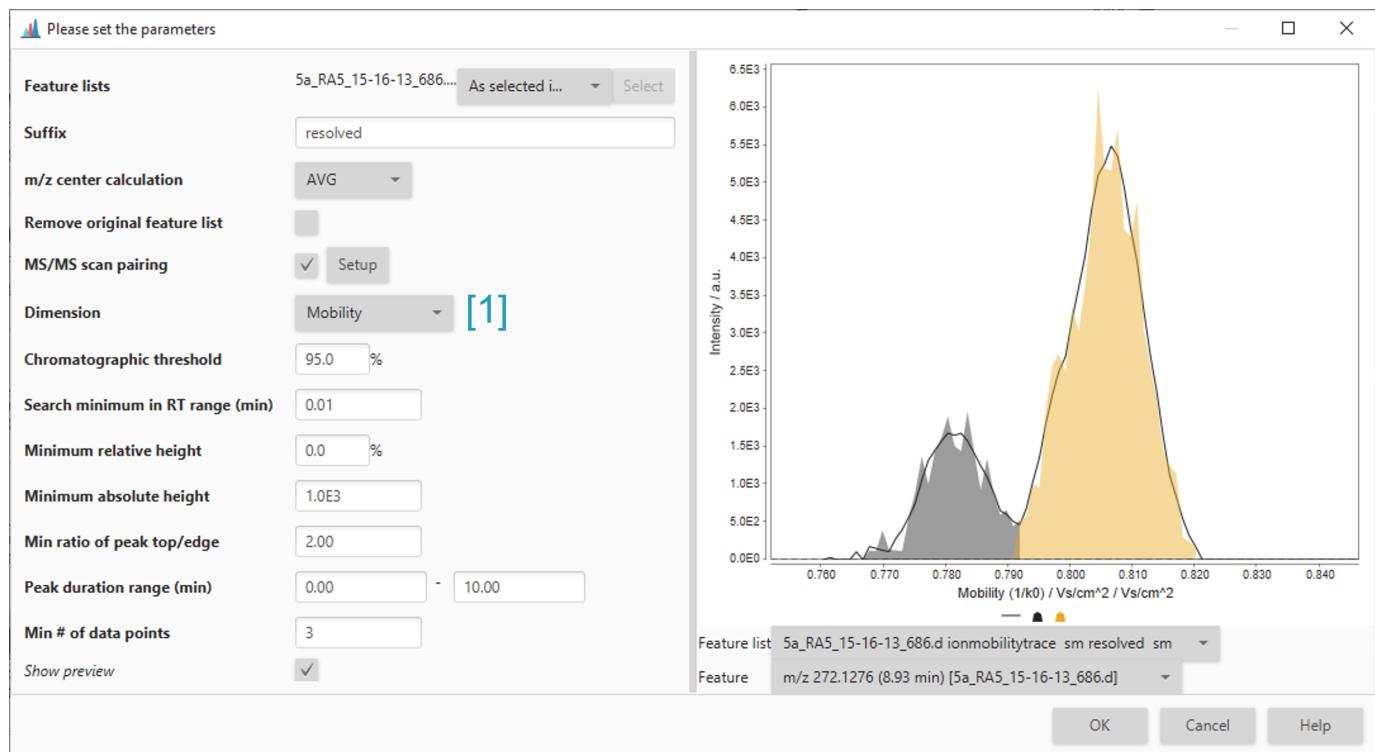
Describes valid peak lengths. Can be used to filter out very short or long noise signals.

Minimum number of data points

Can be used in addition to **Peak duration range** to filter out noise. Should be set no lower than 5 in most cases.

6.8.1 Ion mobility data

In general, the same principles apply as in the retention time resolving step. However, a few differences shall be noted. In the screenshot you can also see, that the resolved mobilograms are recalculated from the raw data and previously apply smoothing steps are therefore lost and must be reapplied if necessary.

**Dimension**

Mobility has to be selected as a dimension to resolve mobilograms [1].

Chromatographic threshold

Since there are less scans in mobility dimension (e.g., 400 - >1000 per frame, depending on the instrument type and setting) than in rt dimension (e.g. 5000 for LC-MS depending on acquisition rate), the threshold should be lowered to 80 or less.

Search minimum range

This value determines the search range in mobility dimension. Therefore, this value has to be set lower when resolving a TIMS (Bruker data) mobilogram, because the absolute numerical values are smaller (e.g., 0.01). When resolving mobilograms from Waters or Agilent data (mobility as drift time in ms), the values are higher and must therefore be increased.

Other parameters such as minimum intensities or minimum number of data points should be adjusted depending on what you want to detect.

Last update: January 18, 2022 21:18:57

6.9 CCS Calibration and calculation

Accurate determination of CCS values requires a valid CCS calibration and molecule charge states to be detected.

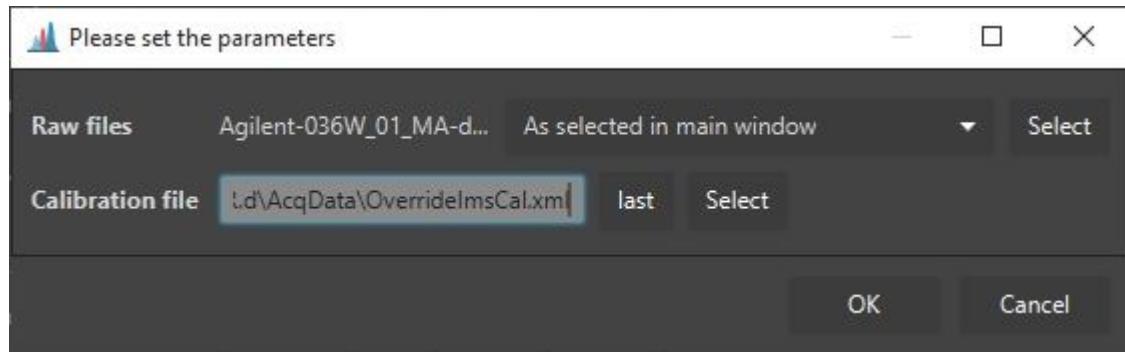
- **timSTOF** raw data can be recalibrated using data analysis and imported in MZmine. The recalibrated data will be used by default. (see [Calculating CCS values](#))
- **mzML** raw data requires the determination of a calibration function from the raw data (e.g. as detected features) or as import from an external file. (see [Creating or importing a CCS calibration](#)

6.9.1 Creating or importing a CCS calibration

Importing a CCS calibration

Agilent calibration data can be imported from the "OverrideImsCal.xml" file in the Agilent raw data folder.

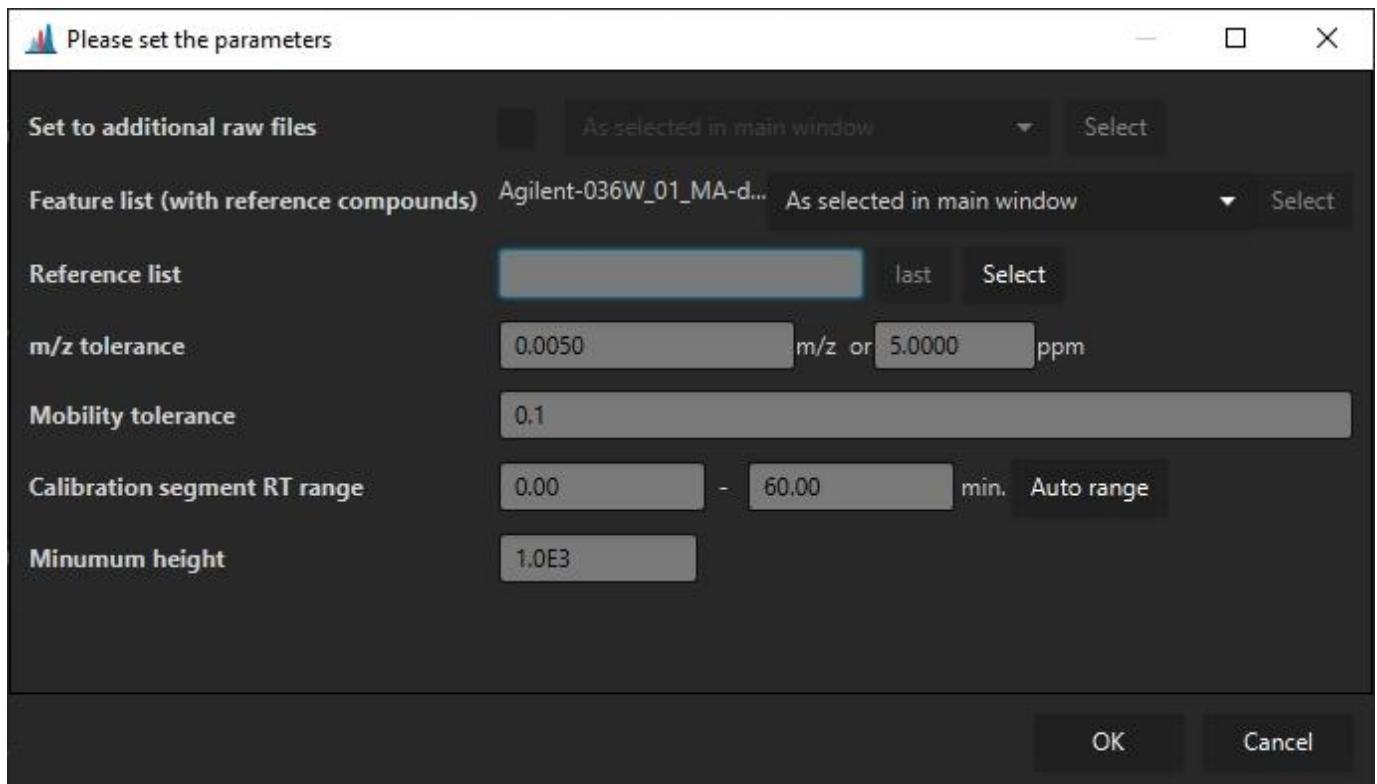
The calibration import is accessed via **Feature list methods -> Processing -> External CCS Calibration**. Then select the calibration "OverrideImsCal.xml" from the raw data folder, and select the raw data files the calibration should be applied to.



Reference CSS calibration

If a mobility calibrant is infused during an HPLC run of every sample, a CCS calibration can be calculated on a per-run basis. Otherwise, a single run can be used to calibrate multiple files.

The calibration module can be accessed via **Feature list methods -> Processing -> Internal reference calibration**.



Set to additional raw files If a calibration calculated from a single feature list shall be applied to multiple other raw files, the raw files can be selected here. This requires only a single raw file to be selected.

Feature list (with reference compounds) Specifies (a) feature list(s) that contains the reference compounds. If multiple feature lists are selected, every feature list will be searched for reference compounds, and the calibration will be used for the raw data files in the particular feature list. This means that no raw data file may be selected. (Cannot set multiple calibrations to a single raw file.)

If a single feature list is selected, the calibration may be applied to additional raw data files via the **Set to additional raw files** parameter.

Reference list Specifies a ".csv" reference list of for CCS calibrant ions. Must contain the columns "mz", "mobility", "ccs", "charge". Columns must be separated by ";". The ion mode may be specified via the charge of the ion, e.g., as 1 or -1. Only the correct polarity will be used to calculate the calibration.

m/z tolerance The m/z tolerance for the reference compounds.

Mobility tolerance the mobility tolerance to detect the reference compounds.

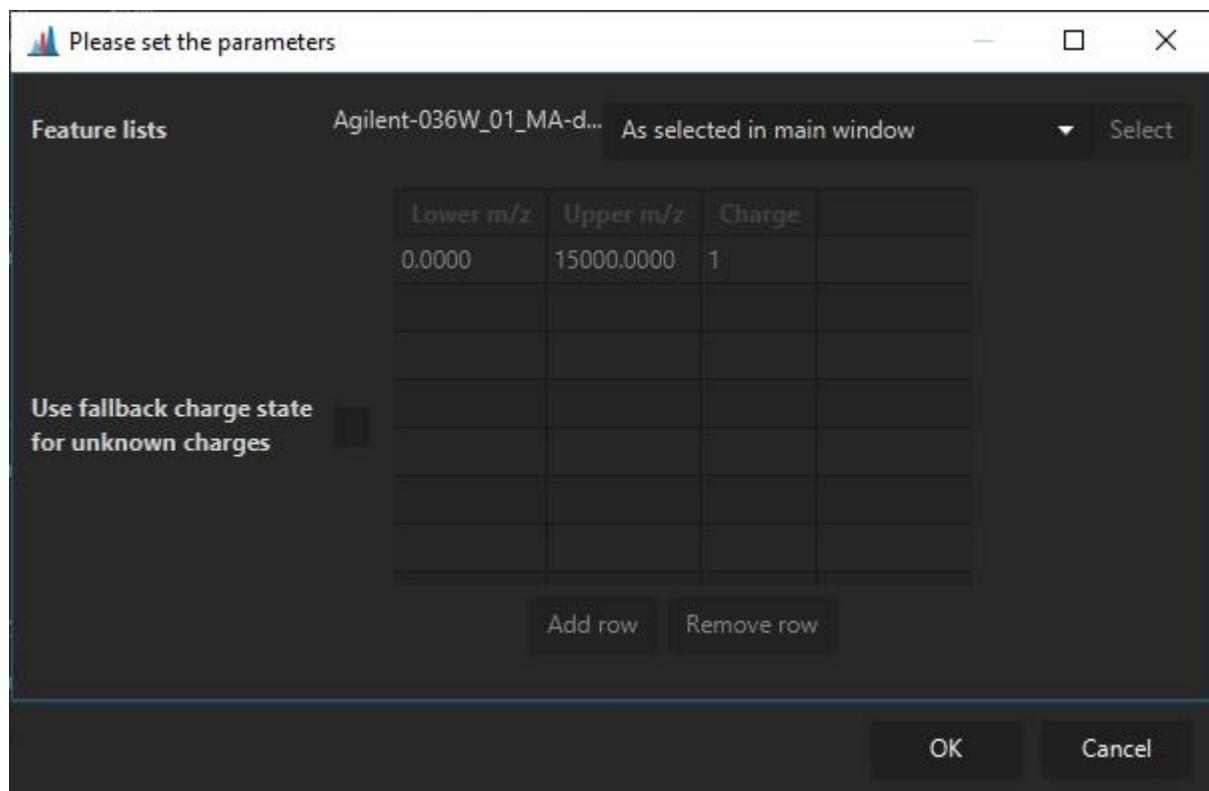
Calibration segment RT Range Specifies the rt range that shall be searched for calibrant ions. Usually either the beginning or end of a HPLC run.

Minimum height A minimum intensity for reference compounds to be used as calibrant signals for determination of the calibration.

6.9.2 Calculating CCS values

After a calibration as been set (Agilent/Waters/Bruker mzML) (Bruker tdf works out-of-the-box) CCS values can be calculated via **Feature list methods -> Processing -> Calculate CCS values**.

Here, a default charge state may be set, in case it could not be determined. Otherwise, the charge state determined via the isotope pattern will be used.



Last update: January 18, 2022 21:51:26

7. Visualisation

7.1 Raw data visualisation

todo

7.1.1 Raw data overview (LC-MS)

todo

7.1.2 Ion mobility raw data overview (LC-IMS-MS)

todo

7.1.3 Chromatogram plot

todo

7.1.4 Spectrum plot

todo

7.1.5 Page Contributors

[SteffenHeu](#)

Last update: January 18, 2022 21:51:26

8. Additional IMS resources

8.1 Ion mobility spectrometry terminology

8.1.1 Background

Ion-mobility mass-spectrometry, here simply referred to as ion-mobility (IM), is an analytical technique where ions are separated through a gas-filled mobility cell prior to the MS acquisition. Ions drift through the IM cell with different velocity based on their interaction with the buffer gas, which allows for the separation of different shaped molecules. Modern devices are able to perform IM separation on a millisecond timescale, typically within 10 to 100 ms. Thus, IM nicely fits in-between LC separation (~seconds timescale) and MS detection of TOF instruments (~microseconds timescale). This allows LC-IM-MS instruments to acquire several MS spectra during each [accumulation](#), without incurring sensitivity loss. For example, assuming a typical 100 μ s MS-acquisition time of TOF analyzers, around 1000 spectra can be recorded within 100 ms of IM separation. Therefore, as opposed to LC-MS, multiple MS (or MS₂) spectra are associated to each RT in LC-IM-MS data. A more detailed explanation of LC-MS and LC-IMS-MS raw data structure is provided [here](#).

Trapped ion mobility spectrometry (TIMS) - Todo

Trapped ion mobility spectrometry (TIMS) reverses the concept of traditional drift tube IM. Rather than moving ions through a stationary gas, TIMS holds ions stationary against a moving gas and then releases them according to their mobility. Video: <https://www.youtube.com/watch?v=cWjz32wky2A>

Time-dispersive ion mobility spectrometry (DTIMS and TWIMS) - Todo

Time-dispersive IM devices include "traditional" drift tube (DTIMS) and travelling-wave (TWIMS) devices. In classic DTIM, ions migrate through an inert buffer gas under the influence of a weak electric field, whereas collisions with buffer gas molecules retard the progress of the ions. As larger ions have more collisions with the gas, they are more strongly retarded than their smaller counterparts. Thus, smaller ions, having a smaller cross section, arrive earlier at the detector than ions with a larger collisional cross section (CCS). The ion mobility K is then defined as the ratio of the analyte's steady-state net drift velocity to the applied electric field, and it is convention to calculate the reduced ion mobility K_0 at standard pressure and standard temperature, often reported as the inverse reduced ion mobility $1/K_0$.

8.1.2 Terminology

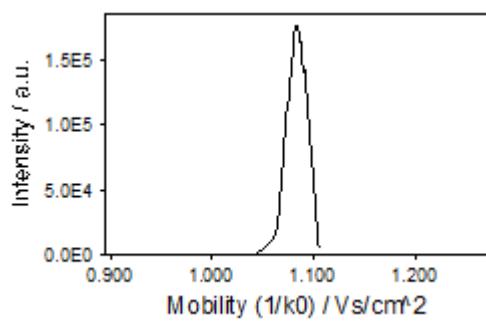
Accumulations, Mobility Scans and Frames

Although mainly used for TIMS, the term "*accumulation*" refers to the pack of ions gathered at the head of the IM device prior to the release and separation in the IM cell. As explained [above](#), since the accumulation-separation cycle typically last ~100 ms, multiple MS spectra (referred to as "*mobility scans*" in MZmine) are acquired during each cycle. The set of *mobility scans* collected during each IM separation constitutes a "*frame*". A *frame* can be seen as the IM separation of a single *accumulation*, along which multiple MS spectra are collected. Several *frames* are contained within one LC peak. Thus, the frame number are a natural unit to measure chromatographic RT. See [here](#) for more details.

--FIGURE HERE--

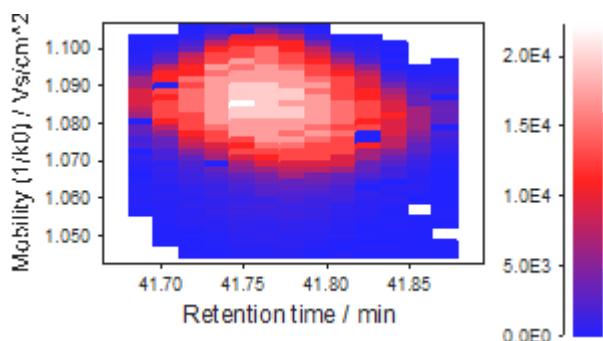
Mobilograms

A "*mobilogram*" represents the intensity of an m/z or m/z range along the mobility axis. A *mobilogram* may be build from multiple frames and summed or built from a single frame.



Ion mobility trace

An "ion mobility trace" basically represents a mobility resolved extracted ion chromatogram (EIC).



Cross Collisional Section

IM-derived CCS values can be used as an additional molecular descriptor to support the compound unknown identification process.(Paglia et al. 2014) However, the number of acquired spectra per run increases from several thousand to several million, requiring memory-efficient software and new processing algorithms.

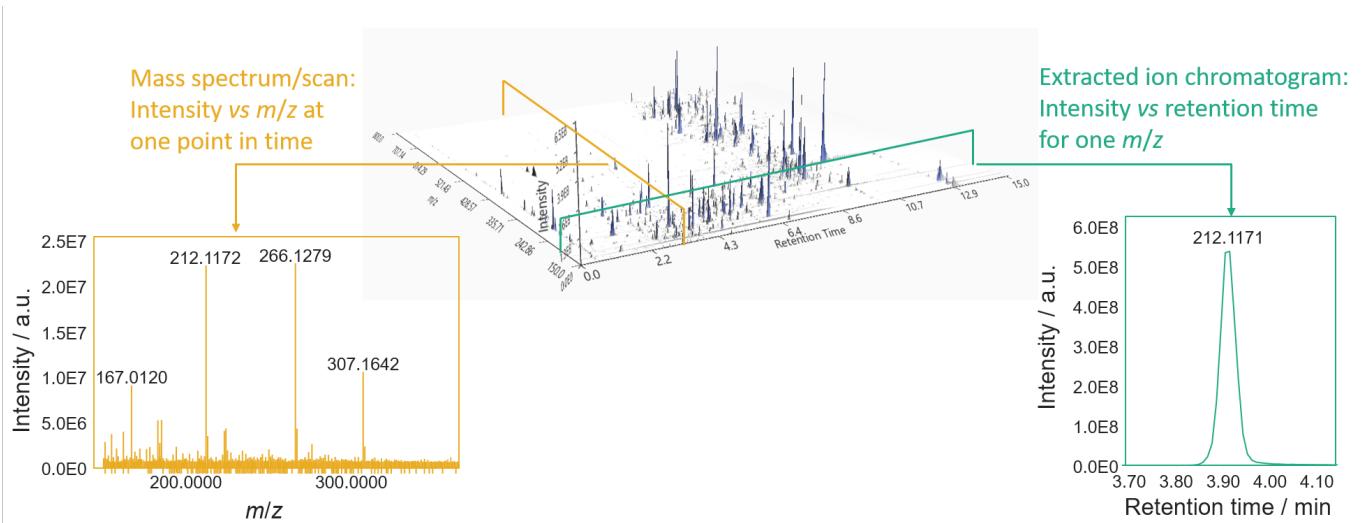
8.1.3 References

<https://doi.org/10.1074/mcp.TIR118.000900>

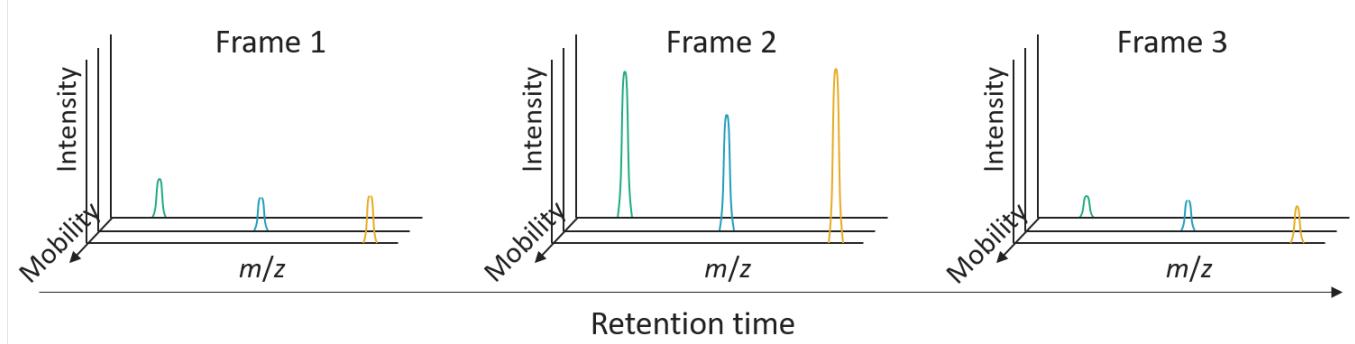
Last update: February 14, 2022 22:31:14

8.2 Graphical comparison of LC-MS and LC-IMS-MS data

Classic LC-MS data consists of three dimensions: m/z, intensity, and retention time. At every retention time, a whole mass spectrum is acquired (yellow). Putting all scans together creates a three-dimensional plane. By slicing the three-dimensional data at a single m/z (+- a tolerance), EICs can be created (green).



On the other hand, ion mobility resolved data consists of a three-dimensional data plane at each retention time. The three dimensions being m/z, intensity, and mobility (as drift time (ms) or inverse reduced mobility $1/k_0$ [$\text{Vs}/(\text{cm}^2)$]). The 3D projection of regular LC-MS data can be created by summing all mobility scans of a frame to create a frame spectrum. (see [Mobility scan merging](#))



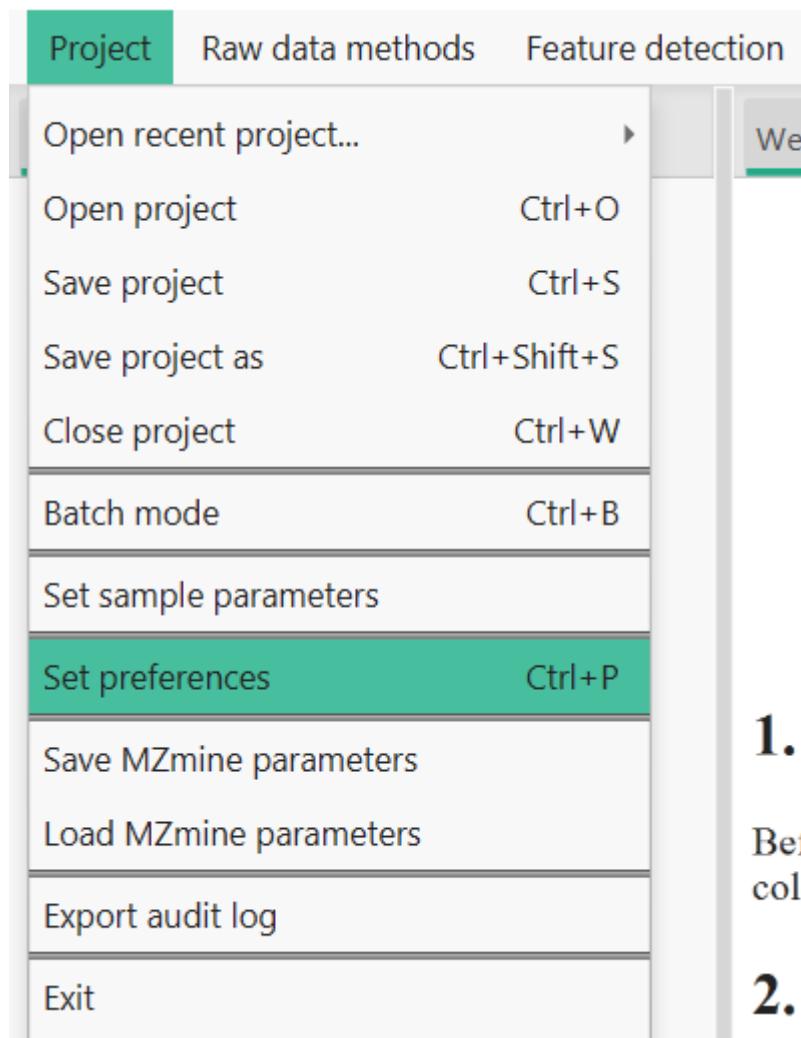
Last update: January 19, 2022 09:34:46

9. Performance options

This section contains information on how to tune MZmine 3 for different systems.

9.1 Preferences

The preferences can be changed in MZmine's graphical user interface by accessing *File/Set preferences* from the menu. The choices will be stored in a (hidden) *.mzmine3.conf* file in the user's home directory (Windows: *C:\Users\USERNAME*) once MZmine is closed.



Important preferences

1.

Be
col

2.

The screenshot shows the 'General' section of the MZmine preferences window. It includes the following settings:

- Number of concurrently running tasks:** Set to the number of CPU cores (8) (dropdown menu with value 8)
- Keep in memory:** ALL (dropdown menu)
- Temporary file directory:** D:\tmpmzmine (text input field)
- Use proxy:** Setup (button)
- R executable path:** (empty text input field) with 'last' and 'Select' buttons.
- Send anonymous statistics:** checked (checkbox)

Below this section are links to 'Other' and 'Visuals'.

Memory options

The screenshot shows the 'General' section of the MZmine preferences window, specifically focusing on memory options. The 'Temporary file directory' dropdown menu is open, showing the following options:

- NONE
- ALL** (highlighted in green)
- FEATURES
- MASS_LISTS
- RAW_SCANS
- MASSES_AND_FEATURES

9.2 Logs

Currently, the logs are written to an *mzmine_0_0.log* file in the user's home directory. Please submit your log files together with any issues on [GitHub](#).

9.3 Command-line arguments

Command-line arguments offer a variety of options that generally override the corresponding parameters in the preferences.

Windows

An easy way to start MZMine with arguments is to create a shortcut to the MZmine.exe, right-click, and define the target with additional arguments. This example runs MZMine in batch mode (headless), imports the specified batch file, overrides the memory management to **none** (which is the default), effectively using memory mapping to store and access spectral, centroid,

and feature data from temporary files stored in the defined temp directory. By leaving out the *memory* or *temp* arguments, the values stored in the current *preferences* file will be used, or the default values if no *preferences* file was found.

Start MZmine batch with memory mapping (DEFAULT)

```
"C:\Program Files\MZmine\MZmine.exe" -batch "D:\batch\my_batch_file.xml" -memory none -temp "D:\tmpmzmine"
```

Start MZmine batch on machines with enough memory (RAM) with -memory all

```
"C:\Program Files\MZmine\MZmine.exe" -batch "D:\batch\my_batch_file.xml" -memory all -temp "D:\tmpmzmine"
```

9.3.1 Argument table

Argument	Options (default)	Description
-batch	a path, e.g. "D:\batch.xml"	Path to batch file
-memory	none , all, features, centroids, raw, masses_features	Defines what data is kept in memory (RAM) or otherwise memory mapped to the temp directory. Generally this setting should be <i>none</i> . If memory is no issue this option might be set to <i>all</i> process all spectral and feature data in memory. The option <i>masses_features</i> keeps centroid mass lists and features in memory while memory mapping raw spectral data.
-temp	a path, e.g., "-temp "D:\tmpmzmine\"	The defined directory should be on a fast drive (usually SSD > HDD > network drive) with enough free space. Local drives are usually preferred. MZmine uses memory mapping to efficiently store and access spectral and feature data. This can lead to a considerable temporary consumption of disk space. Make sure that the selected drive has enough space (maybe 20 GB + 1 GB/10 files; generously over estimated).

Last update: December 9, 2021 21:42:48

10. How to contribute

10.1 Contribute to the MZmine documentation

1. Make a GitHub Account

You'll need to make a [GitHub Account](#).

2. Click Edit Button on Page You Want to Edit

MZmine 3 Documentation

[Home page](#)
[Main window overview](#)
[LC-MS workflow](#)
[LC-IMS-MS workflow](#)
[Raw data visualisation](#)

LC-IMS-MS Workflow

Supported formats

- Vendor formats: *
 - .tdf (Native Bruker LC-IMS-MS and MALDI-IMS-MSI format) *
 - .tsf (Native Bruker MALDI-IMS-MS (single shot) format)
- .mzML *
 - Created via [MSConvert](#) from native Bruker data *
 - Created via [MSConvert](#) from native Waters data

Table
Edit this page
Backg
termir
Mot
forn
Fran
Mot
Ion
Raw d
Raw
Mas
Se

3. Fork the Repository When Prompted (only the first time)



You need to fork this repository to propose changes.

Sorry, you're not able to edit this repository directly—you need to fork it and propose your changes from there instead.

[Fork this repository](#)

[Learn more about forks](#)

4. Make the Edits in MarkDown

mzmine_documentation / docs / Ion-mobility-data-proc Cancel changes

Spaces 3 Soft wrap

```

1 # LC-IMS-MS Workflow
2 ## Supported formats
3
4 * Vendor formats:
5 *
6     * .tdf (Native Bruker LC-IMS-MS and MALDI-IMS-MSI format)
7 *
8     * .tsf (Native Bruker MALDI-IMS-MS (single shot) format)
9 * .mzML
10 *
11     * Created via [MSConvert](https://proteowizard.sourceforge.io/download.html) from native Bruker
12     data
13 *
14     * Created via [MSConvert](https://proteowizard.sourceforge.io/download.html) from native Waters
15     data
16
17 **Note**: mzML via MSConvert from Agilent raw data can be imported, but certain restrictions might
18 hinder processing workflows due to the nature of the raw data format.
19
20
21 ***
22
23 ## Background information and terminology
24
25 Since ion mobility spectrometry (IMS) resolved data is more complex due to the additional dimension
26 when compared to regular LC-MS data, some terms shall be clarified before going into details of the
27 processing steps.
28
29 ### Mobility separation and data format
30
31 Ion mobility separation usually occurs on the millisecond timescale, fitting nicely in-between
32 liquid chromatography (LC) (few seconds per chromatographic peak) and mass spectra acquisition of
33 TOF instruments (several micro seconds). Therefore, the mobility dimension can be resolved by
34 acquiring multiple spectra during a mobility separation (e.g. 1000 spectra per 100 ms). This leads
35 to multiple mass spectra acquired at one IMS accumulation. Thus, at one retention time, multiple
36 spectra are acquired. A detailed comparison of LC-MS and LC-IMS-MS raw data can be

```

Attach files by dragging & dropping, selecting or pasting them.

5. Propose Changes

Please describe the change you are making.

Commit changes

update mobility resolving step

add msms pairing description in mobility resolving step

steffen.heuckeroth@gmx.de

Choose which email address to associate with this commit

- o Commit directly to the `master` branch.
- Create a new branch for this commit and start a pull request. Learn more about pull requests.

6. Create Pull Request

The screenshot shows a GitHub repository page for 'mzmine / mzmine_documentation'. The 'Pull requests' tab is selected. A yellow banner at the top indicates that 'SteffenHeu-patch-1 had recent pushes 1 minute ago'. A green button labeled 'Compare & pull request' is visible.

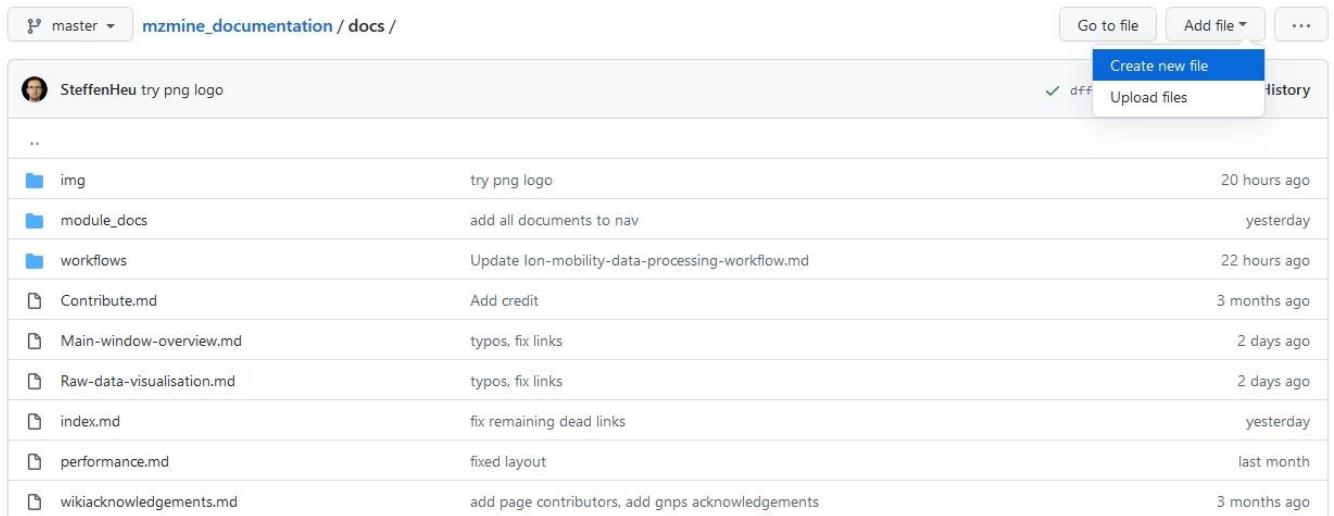
7. Finalize Pull Request with Description

The screenshot shows the GitHub pull request creation interface. The title 'update mobility resolving step' is entered. The description 'add msms pairing description in mobility resolving step' is written. A note says 'Able to merge. These branches can be automatically merged'. A 'Create pull request' button is at the bottom right.

10.2 Creating a new page

Follow steps 1 - 3.

Navigate to mzmine_documentation/docs in your fork and create a new file



mzmine_documentation / docs /		
SteffenHeu try png logo		...
 img	try png logo	20 hours ago
 module_docs	add all documents to nav	yesterday
 workflows	Update Ion-mobility-data-processing-workflow.md	22 hours ago
 Contribute.md	Add credit	3 months ago
 Main-window-overview.md	typos, fix links	2 days ago
 Raw-data-visualisation.md	typos, fix links	2 days ago
 index.md	fix remaining dead links	yesterday
 performance.md	fixed layout	last month
 wikiacknowledgements.md	add page contributors, add gnps acknowledgements	3 months ago

Follow steps 4 - 7.

10.3 Page Contributors

[SteffenHeu](#)

This page was adapted from the [GNPS documentation](#).

Last update: January 20, 2022 14:49:05

11. Acknowledgements

We would like to point out that this wiki was set up in tight collaboration with the [GNPS](#) staff. We highly appreciate your help!

Last update: October 12, 2021 11:51:46