MALINTO – MALDI Interpretation Tool

Program Manual of *MALINTO* version 1.0 (10/31/2022)

GNU Octave Software

For the MALDI interpretation tool (*MALINTO*) the programming language GNU Octave was used. GNU Octave is similar to Matlab and available as free software which can be downloaded using the following link: https://ftpmirror.gnu.org/octave/windows/ (last accessed on October 31, 2022). Although newer Octave versions are continuously released, the download and installation of "octave-6.4.0-w64-installer.exe" is highly recommended for the correct performance of *MALINTO*.

Start of the Program

The *MALINTO* software package includes following program and example files:

- MALINTO_GUI.m; graphical user interface of *MALINTO*, this script is started to run the MALDI interpretation tool
- cartesian_product.m; calculation of theoretical mass list without filters
- maldi_masslist.m; filtering of theoretical mass list, assignment of peaks from MALDI experiment to theoretical mass list
- maldi_statistics.m; calculation of comonomer and terminating group ratios
- Example_SampleFiles.xlsx; raw data from MALDI measurements, exported from the Bruker FlexAnalysis® software
- Example_MassLists.xlsx; Excel export after calculation of theoretical mass list, assignment of experimental peaks, and running of the statistics program. Includes input data (monomeric repeating units, end groups, adducts). This file can be imported to *MALINTO* for further processing or to extract the input data. Background information on the examples is given in the corresponding article "*MALINTO*: A New MALDI Interpretation Tool for Enhanced Peak Assignment and Semiquantitative Studies of Complex Synthetic Polymers" (https://doi.org/10.1021/jasms.2c00311)
- Example_MassLists_Comments.xlsx; calculations performed by *MALINTO* are repeated in Excel to better explain the principles (functionalities and terminating groups, conditions to exclude chemical preposterous structures, statistical evaluations,…); general formulas are given in the corresponding article

After opening the GNU Octave software the directory of the *MALINTO* files has to be chosen on the left side (1) as shown in the overview of the GNU Octave interface in Figure 1. The Excel mass lists of the experiments have to be saved in the same folder in .xlsx format. The "MALINTO_GUI.m" file is selected by double clicking (2) and the program is started by choosing the highlighted button (3). A separated window opens which is shown in Figure 2. If Octave is newly installed, an error message concerning the encoding might pop up which is solved by choosing "SYSTEM (CP 1252)". This setting can be alternatively checked and changed in the menu available via Edit/Preferences/Editor/File handling.

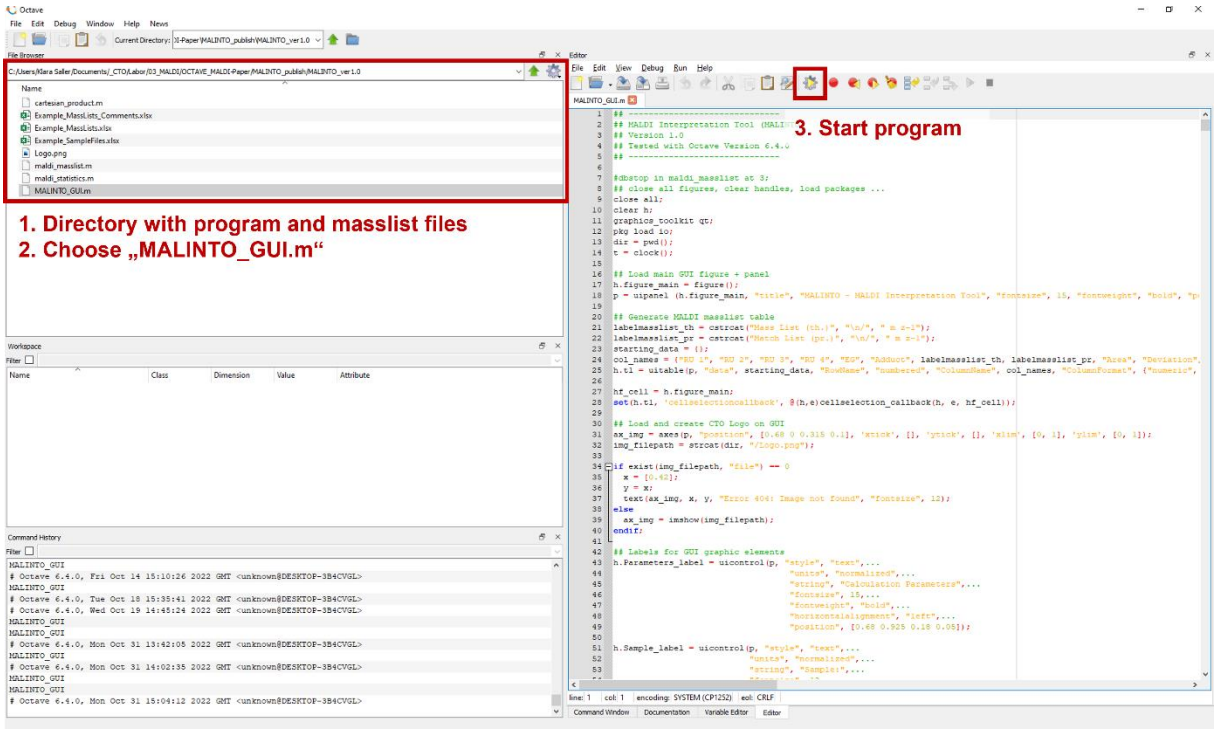

Figure 1: GNU Octave interface to start the graphical user interface of the MALDI interpretation tool (*MALINTO*).

Theoretical Mass List

The main window of the program (Figure 2) is filled with masses and functionalities of the monomers, end groups and adducts either manually or by importing a previously generated export.

1) Monomeric Repeating Units
   For addition polymers (chain growth polymers, polyurethanes, etc.) the mass of the repeating unit equals the mass of the monomer. In case of condensates, the monomer mass has to be subtracted by the mass of the condensation product in order to obtain the mass of the repeating unit in the polymer. As shown in Figure 2, the monoisotopic mass of isophthalic acid (IPA, 166.02661 Da) is reduced by one water molecule (18.01057 Da), the same accounts for neopentyl glycol (NPG, 104.08373 Da) and 1,10-decanediol (DD, 174.16198 Da).
   Additional to their mass, monomers can differ in kind and number of functionalities. Diacids can be reacted with diols to give polyesters, the same principle applies to polyamides (diacid/diamine) and polyurethanes (diisocyanate/diol). Using monomers with a higher number of functionalities gives branched polymer chains with an increased number of terminating groups. By specifying the kind and number of functionalities, chemically unlikely structures are removed during mass list calculation (e.g. product of 2 isocyanates). In case of AB monomers such as lactic acid, both COOH and OH functionalities are present in the monomer, allowing it to react with itself. Double bonds are specified in the same way, both functional group columns are filled with the amount of double bonds present (e.g. 1 for styrene, 2 for divinylbenzene).

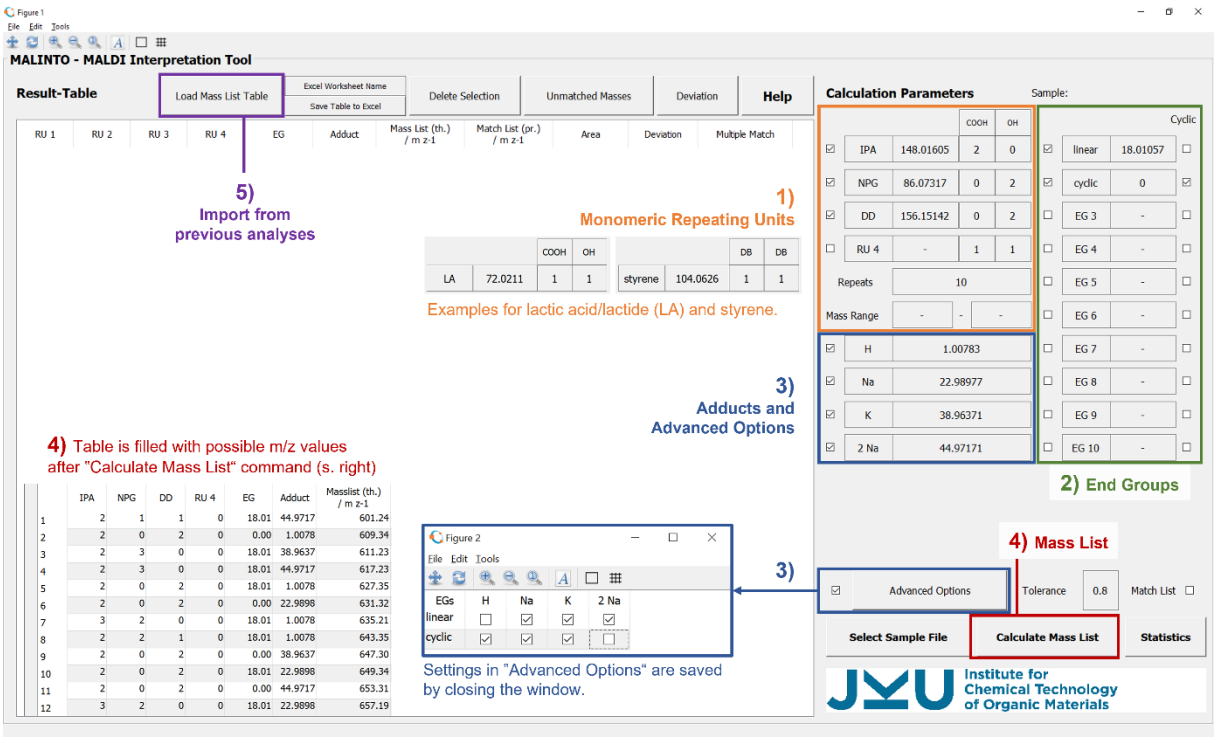

Figure 2: Main window of *MALINTO* including necessary input data 1) monomers/repeating units, 2) end groups, 3) adducts and advanced options for end group - adduct combinations, 4) calculation of mass-list, 5) import from previously generated excel export (e.g. "Example_MassLists.xlsx" which is part of the download package).

The length of the calculated mass list can be limited either by giving the maximum number of repeating units or by defining a mass range. The monomers can be selected and deselected for further sample evaluation. Frequent changes of the input data during the same session, especially selecting and deselecting, might lead to error messages in the command window because *MALINTO* generates different tables in the background whose sizes do not necessarily match the new input data. This problem is easily solved by closing and restarting *MALINTO*. It is recommended to previously generate an Excel export (13, Figure 5) to simply reimport the input data (5) after restart.

2) End Groups

Depending on the kind of polymers, end groups can either be varied by initiators and terminating agents (for example in anionic or ring-opening polymerizations) or distinguish between linear and cyclic species (for example in polycondensation reactions). In the case of branched structures, multicycles may be present (see Example 3 in corresponding article) which are identified by repeated subtraction of the condensation product. If the end group presents a cyclic structure, it has to be marked as “cyclic” (left to the end group’s mass) which triggers further events such as correction of number of free functionalities prior to statistical evaluation.

For the shown example with isophthalic acid, neopentyl glycol, and 1,10-decanediol linear polyesters with acid and/or alcohol terminating groups can be formed. This ratio of free functionalities will be calculated in the statistics tool using the number of the individual monomers and the sum of corresponding functionalities.

3) Adducts and Advanced Options

Up to 4 different ions can be chosen as adducts for the calculation of the mass list. The shown examples are often observed for polyester systems if no special purification is

carried out. Additional to H, Na, and K adducts, a proton/sodium exchange may occur if free carboxylic acid functionalities are present (sodium adduct described as "2 Na"). Since this requires a linear structure, special combinations of end groups and adducts may be excluded from the mass list in the "Advanced Options" settings. These settings are saved by closing the window and setting a tick next to the button.

4) Calculate Mass List
After entering the required input data, a theoretical mass list can be calculated by clicking on the specified button. All selected monomers / repeating units are combined up to the chosen number of repeats, after which end group and adduct masses are added. Since several mathematically possible combinations do not represent real chemical structures, several filters are applied as described in the corresponding article and shown in the "Example_MassLists_Comments.xlsx" file. The theoretical mass list is sorted by m/z values and can be used for manual assignment of peaks, e.g. already during MALDI measurements.

5) Load Mass List Table
If a set of data has already been exported as Excel file, the input data can be reimported by choosing "Load Mass List Table" and selecting respective file and tab (optional). The *MALINTO* download package includes several examples in the file "Example_MassLists". These files (or tabs) should be copied prior to further editing in order to prevent problems during import.

Assignment of Peaks – Match List

After conducting the MALDI measurement and performing calibration, a deconvoluted peak list has to be generated including the areas / integrals of the peaks. In this special case, this peak list is generated by the Bruker FlexAnalysis® software and can be exported as Excel file. Prior to analysis via *MALINTO*, this file has to be placed in the same directory as the program files. After entering (1–3) or importing (5) the input data, following steps are required which are highlighted in Figure 3.

6) Select Sample File
Files need to be in a .xlsx format, eventually they have to be converted first. If the file includes more than 1 tab, the desired sample can be selected in the software. The tab (sample) name is displayed on the upper right corner.

7) Match List and Tolerance
Choose "Match List" to compare the mass list of the selected sample to the calculated theoretical mass list. Peaks are assigned if the m/z value matches one or more theoretical values within an adjustable tolerance range.

8) Calculate Mass List
Performs the calculation as described in 4) including the comparison to the selected sample file. If practical m/z values can be assigned within the specified tolerance, these m/z values including deviation to theory and corresponding peak area are included. If more than one theoretical value lies within the tolerance, peak data are added to all possible assignments and "nan" (not a number) appears in the "Multiple Match" column.

These assignments have to be corrected manually by selecting unfit lines and using the "Delete Selection" button (see Figure 4 with further details).

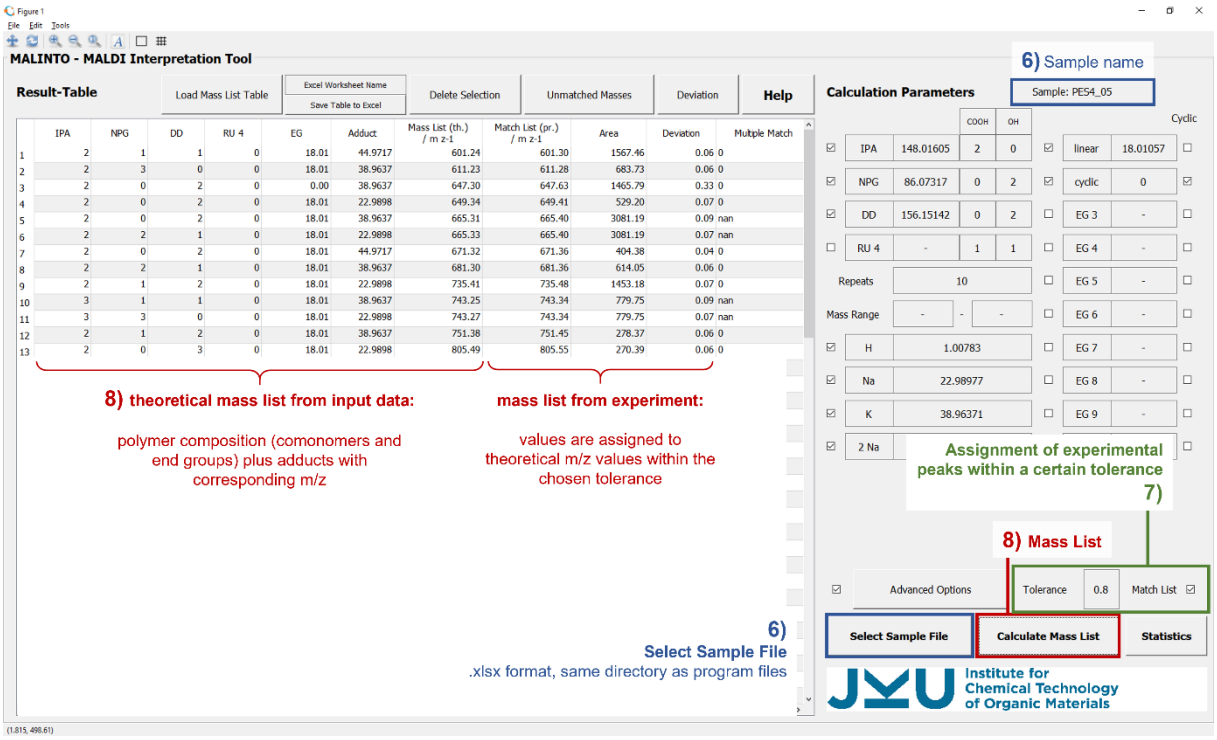

Figure 3: In order to assign peaks from a MALDI measurement, 6) a sample file, 7) a tolerance value and the option "Match List" are selected. After using 8) "Calculate Mass List", the table fills with the experimental peaks and their mathematically possible assignments.

Corrections of the Match List

In some cases, the match list has to be manually corrected due to either wrong or ambiguous assignments. The number of ambiguous assignments drastically increases with the number of comonomers, end groups, and adducts and depends on the actual molar masses (How many different combinations have nearly the same m/z value?). Several tools implemented in the *MALINTO* software help with the identification of the right assignment and highlight possible problems. These are shown in Figure 4.

9) Deviation
The deviation of the experimental to the theoretical m/z values gives hints about the accuracy of the peak identification and the calibration. While a general trend is usually plausible, strongly varying deviations might indicate wrong assignments. Additional to the deviation values in the column, a graph is shown after clicking on the "Deviation" button (10) which helps to identify such entries. The row number is shown on the x-axis to facilitate finding the entries in the list which can further be selected and deleted (12). For the identification of the exact row number using the graph, the cursor can be placed on a certain data point and the coordinates will be given in the lower left corner of the window. The deviation graph in the example shown in Figure 4 also indicates two different series which correspond to the ambiguous assignments. After correcting these, the deviation graph can be checked again by clicking on the corresponding button (10). The graph is not updated in real time.

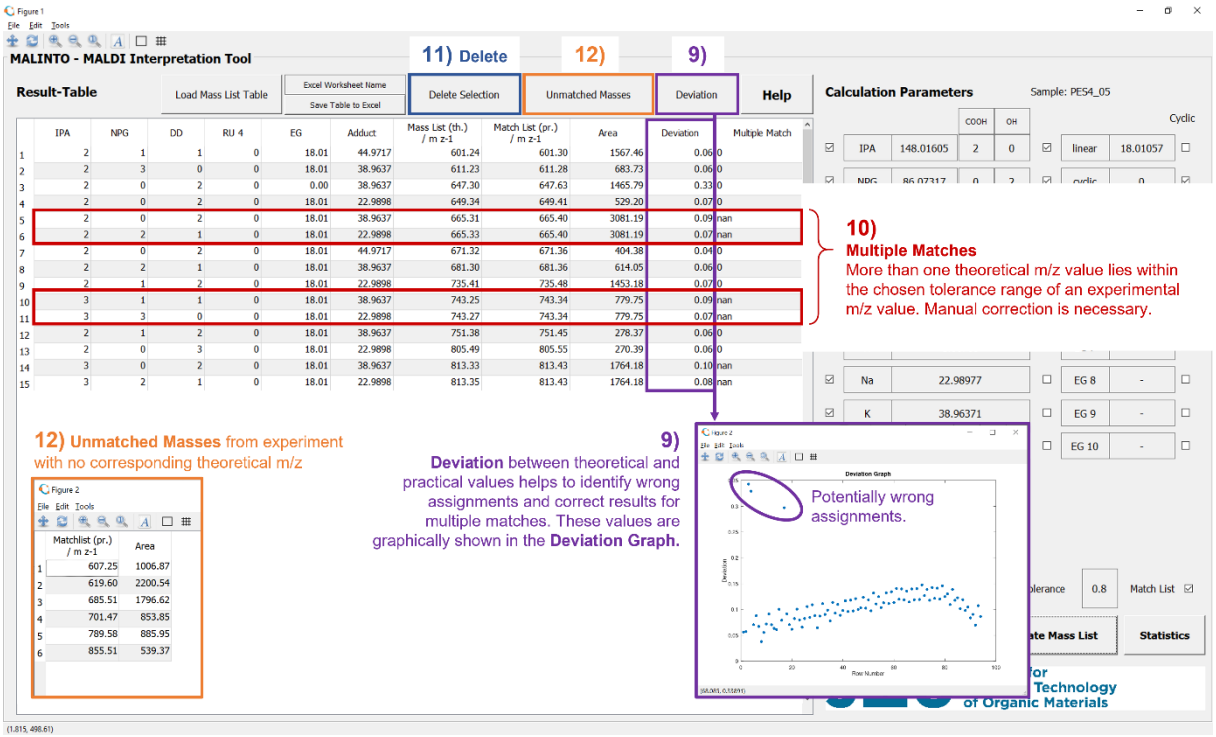

Figure 4: The match list might result in wrong or ambiguous assignments whose identification is facilitated by the 9) Deviation column and graph as well as the 10) Multiple Match column. These assignments should be deleted (11) prior to ongoing investigations. Experimental peaks which cannot be assigned with the generated mass list are shown by choosing "Unmatched Masses" (12).

10) Multiple Matches

The "Multiple Match" column highlights ambiguous assignments by giving the entry "nan" (not a number). "nan" can be optically well distinguished from "0" which indicates a unique assignment. The identification of the right assignment usually requires experience with or further knowledge of the sample. It is usually advisable to thoroughly investigate the series of a newly examined sample, not only using the software but additionally looking at the original MALDI mass spectrum to identify the different series. In the example shown in Figure 4, the mass of sodium adducts of one polyester species is similar to the mass of potassium adducts of polyesters with other compositions. Since experiments were carried out using sodium trifluoroacetate as ionization agents, sodium adducts were dominant for low molecular weight species which could be assigned unambiguously. Potassium adducts were observed as well but only for species which also had intense peaks for the corresponding sodium adduct. Hence, polymer species with the sodium adduct were chosen as the right assignment. After identifying such patterns, it was tested to work exclusively with sodium adducts which gave good results and assigning was accelerated significantly (see corresponding publication).

11) Delete Selection

In order to manually correct the match list, wrong assignments have to be deleted prior to statistical evaluations. This can be realized by selecting rows or single cells and choosing the "Delete Selection" (12) button. In both cases the row is completely deleted. Multiple selection is possible by using the "Shift" or "Ctrl" key.

12) Unmatched Masses

If the experimental mass list contains entries, which cannot be assigned by comparison with the theoretical mass list, these are listed as "Unmatched Masses" when choosing the corresponding button (9). A particularly high number of unmatched masses might indicate mistakes in the input data or significant abundancy of unexpected side products. In case of an Excel export, this list will be included.

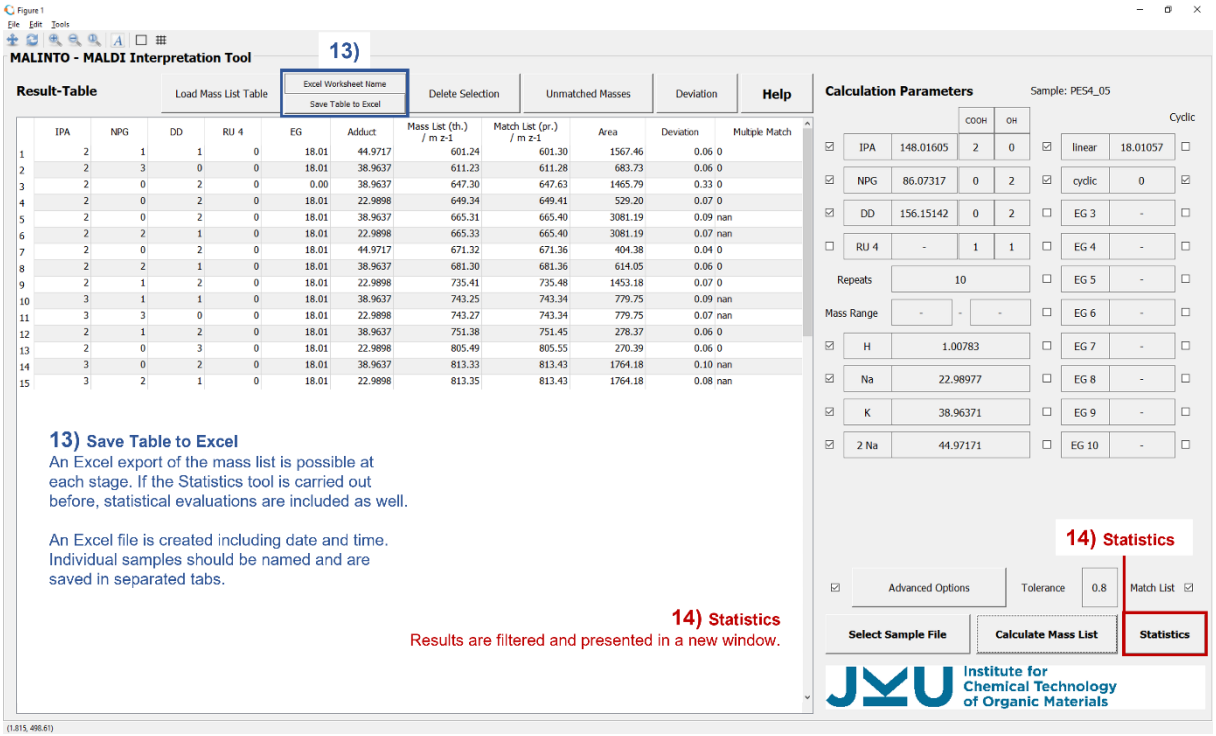

Figure 5: The mass and match list can be exported any time using the "Save Table to Excel" button (13). After clearing of wrong and ambiguous assignments, statistical evaluations can be started (14). Results from the statistics are also exported if the tool has been used prior to saving the Excel file.

13) Save Table to Excel

Both the theoretical mass list as well as the match list can be exported to Excel files at any time during editing (Figure 5). While a new Excel file is created each time the program has been started, different results are stored in individual tabs of this file. The name of these tabs can be specified (13) prior to clicking the button "Save Table to Excel". If the export is performed after any statistical evaluations, these results will be stored in the file as well.

14) Statistics

This button (14) is chosen to run the statistical evaluations of previously checked match lists. The results are shown in a new window.

Statistics – Quantification experiments

While *MALINTO* is already significantly reducing the time needed for the assignment of peaks in MALDI spectra, it can further be used for analyzing ratios of different comonomers, and terminating groups. Quantification via MALDI mass spectrometry is limited due to several factors such as inhomogeneous sample spots, varying ionization efficiencies, and suppression of

higher molecular weights. Nevertheless, previously published results (for example Saller et. al., *Anal. Chem.* 2020, 92, 15221–15228) showed potential for the semi-quantitative use of MALDI MS. Similar studies shall be encouraged by the accelerated MALDI interpretation using *MALINTO*. The statistics window is shown in Figure 6.

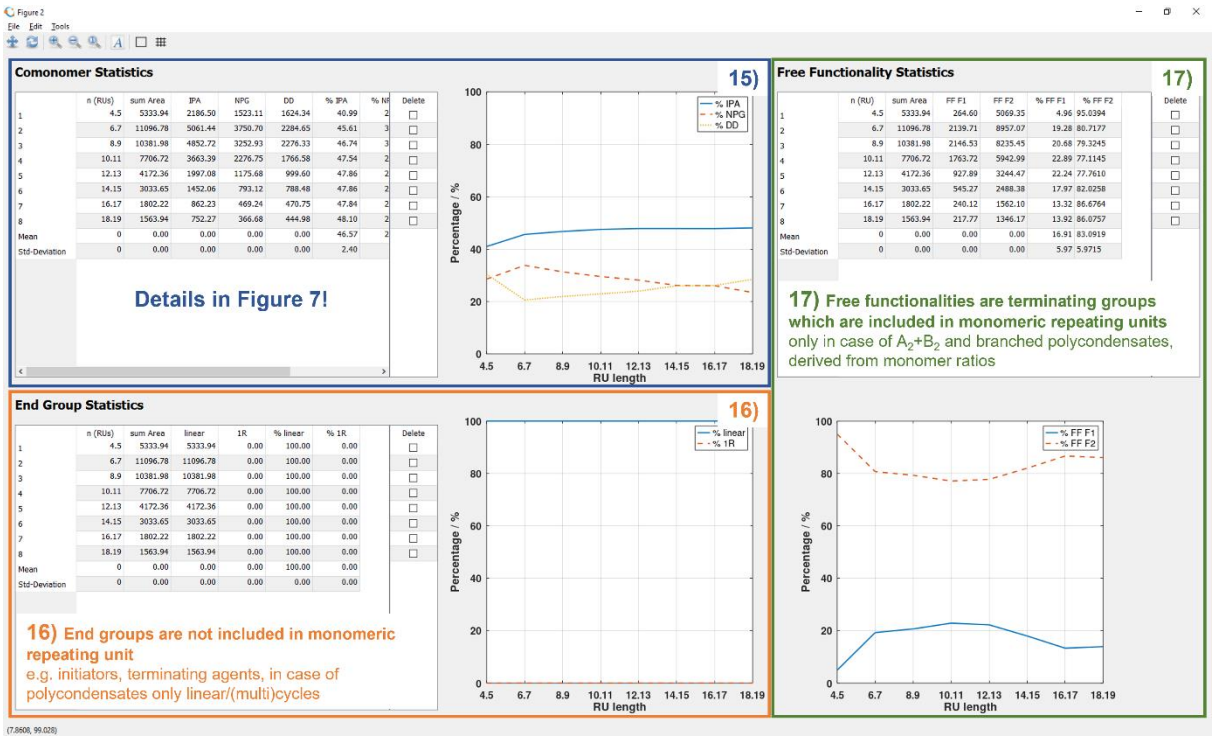

Figure 6: Statistics window including evaluations on 15) comonomer ratios and terminating group ratios. There are two options for terminating groups: 16) End groups are not included in the repeating unit and are part of the required input data for calculating theoretical mass lists. 17) Free functionalities are terminating groups which are derived from monomer ratios if these monomers carry different functionalities (e.g. in A_2+B_2 polycondensates).

15) Comonomer Statistics

To show the structure of the statistical evaluations, details of the Comonomer Statistics are given in Figure 7.

A) In principle, areas of individual peaks are divided according to their polymer composition (in this case comonomers). If peaks are assigned to the same polymer species but with different adducts, areas are summed up.

B) These results are both numerically and graphically summarized for different numbers of repeating units n (RUs) to identify trends or outliers. Two succeeding n (RUs) are combined since in the case of polycondensates even numbers always give mixed terminated chains or cycles while odd numbers only occur if one functionality is in excess.

C) The mean value and standard deviation of results for the individual chain lengths are calculated.

D) Select outliers to be deleted prior to the final calculation which will be performed upon closing the statistics window. This will also affect the results shown in the Excel export. The example shown in Figure 7 reveals a slight trend of the comonomer ratios which is in accordance with further results discussed in the corresponding publication. For individual samples, such a trend should further be investigated instead of taking the resulting mean value as a final result. In such cases, the mean

value is simply depending on the examined m/z range instead of representing the comonomer composition of the whole polymer.

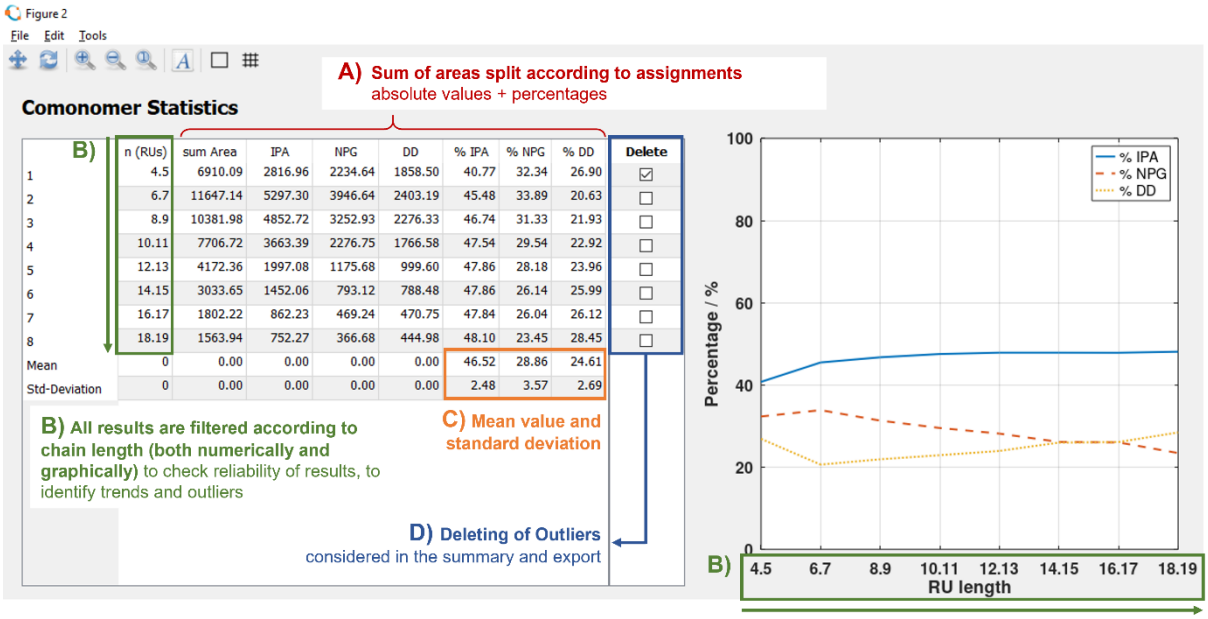

Figure 7: Structure of the individual statistics. A) The sum of areas is split according to the assigned polymer composition. B) Results are filtered according to their chain length in order to identify trends or outliers. As a result, mean value and standard deviation (C) of the relative abundancies are provided. Outliers can be deleted (D) which will be considered in the summary (see Figure 8) as well as the Excel export.

16) End Group Statistics

In order to work on both complex polycondensates and simple polymers the *MALINTO* software evaluates different kinds of terminating groups. End groups are of special importance for simple polymers and investigations on cyclization of (branched) polycondensates. These end groups are not included in the repeating unit and thus, have additional masses which have to be specified in the input data (2). Examples for such end groups are initiators and terminating agents for anionic or ring-opening polymerization products. In case of polycondensates, linear and (multi)cyclic structures are distinguished.

17) Free Functionality Statistics

The actual terminating groups of polycondensates are not caused by certain initiators or terminating agents but the ratio of the individual monomers and their different functionalities. Thus, this kind of terminating groups differ from the classical end groups (16) as defined above and are called "free functionalities". In the given example an excess of the alcohol components mainly leads to polyester chains with two free alcohol functionalities. Additionally, polyester species with two free acid or both functionalities are present which reduces the relative amount of % FF OH.

18) Summary of all statistical evaluations

After optional selection of outliers, mean value and standard deviation of the three statistics are recalculated after closing the statistics window. A summary of the results is given in a new window which is shown in Figure 8.

E) Additional to previously discussed results, the average number of free functionalities Ø FF for the chosen mass range is calculated. This value could be of interest

when investigating branched polycondensates as shown in the corresponding publication. Ø FF can only be compared between similar samples in a defined m/z range. With increasing chain length, more branching monomers can be incorporated which naturally increases the number of free functionalities per polymer chain and thus, the Ø FF value.

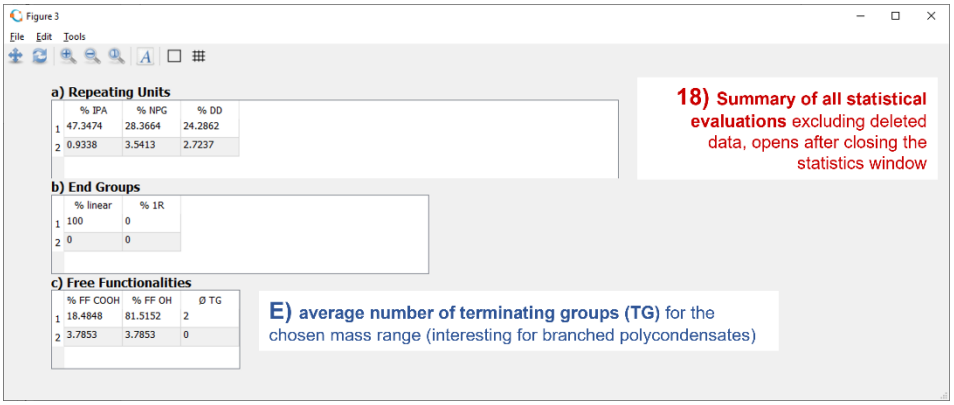

Figure 8: After closing the statistics tool, a summary of the results opens in a separate window. Additionally, the average number of free functionalities within the chosen m/z range is given which might be of interest for branched polycondensates.

Additional Information

The concept of *MALINTO* as well as examples given in the *MALINTO* software package are explained in detail in the article "*MALINTO*: A New MALDI Interpretation Tool for Enhanced Peak Assignment and Semiquantitative Studies of Complex Synthetic Polymers" published in the Journal of the American Society for Mass Spectrometry. The open access article is available using the following link: https://doi.org/10.1021/jasms.2c00311.

MALINTO has been programed in GNU Octave by Daniel C. Pernusch after the concept developed by Klara M. Saller. The project was initiated and supported by Clemens Schwarzinger. The work was performed at the Institute for Chemical Technology of Organic Materials at the Johannes Kepler University Linz in Austria.

Please contact klara.saller@jku.at or clemens.schwarzinger@jku.at for any questions or comments.