University of Delaware Mass Spectrometry Facility

Polymer MALDI-TOF Data Analysis Guide

Author: Katherine Martin Date: 01/18/2023, updated 03/29/2023

Molecular Formula Calculation

Molecular Formula identification can be done on the MALDI-TOF computer in the Bruker **flexAnalysis** software by clicking *Tools > SmartFormula Manually*. There are also web-apps available for this such as **ChemCalc** (https://www.chemcalc.org/, see https://pubs.acs.org/doi/10.1021/ci300563h) or the **Advanced Molecular Formula Generator** also from researchers at the EPFL MS Facility (https://ms.epfl.ch/applications/theoretical-calculations/).

Exact Mass and Isotopic Pattern Modeling

On the MALDI-TOF computer in **PolyTools** through *View > Mass Calc* is a basic mass calculator. The separate Bruker **IsotopePattern** app is for exact mass and isotopic pattern modeling of molecules by resolution, but can run into memory issues with more complex modeling. Xcalibur Qual Browser can also be used for isotopic pattern modeling by resolution and is available on computers at the computer table in the lab.

For web-apps, isotopic pattern modeling is available through **ChemCalc**, mentioned above. Additionally, **EnviPat** and its corresponding R package (https://www.envipat.eawag.ch/, see http://pubs.acs.org/doi/abs/10.1021/acs.analchem.5b00941) is a useful tool for more complicated modeling. It is good to become familiar with larger polymer isotopic complexity and how this appears under different resolving powers.

Note: *If non-commercial software is used in data analysis, it should be cited. Guidelines for this are given on the respective sites.*

PolyTools

Please see the final section of this guide for more information and a walkthrough example.

The MALDI-TOF computer has a copy of Bruker **PolyTools (ver. 1.31)** which can be used for end group, repeating unit, and cation adduct suggestions for homopolymers and homopolymer mixtures and to automatically calculate number average molecular weight (M_n), weight average molecular weight (M_w), polydispersity index (pdi), degree of polymerization (DP), percentage of intensity in the series out of the complete peak list (%I), and count of number of *m/z* in the series (cnt).

Please note that PolyTools is a suggestion software and performs automated calculations. It is best used with an understanding of reasonable ID and the process behind parameter calculations. It can only account for α and ω end groups. Additional *m/z* will end up in the residual or be incorrectly combined into default suggestions for end groups or cations. This version does not make mass defect plots, but information from it can be used to manually

create these. The help guide for PolyTools can be accessed through the ? > menu options in the software.

Polymerix

The software **Polymerix (ver. 3.01)** by Sierra Analytics is also available on the MALDI-TOF computer. This software is much more powerful than PolyTools and can handle spectra processing (including changing peak width), ultrahigh mass resolution data, and copolymer analysis. It can account for resolution changes across a spectrum; it additionally calculates Z-average molecular weight (M_z), and DP_n, DP_w, with DP; allows double bond equivalent (DBE) and element ratio restriction; and can account for negatively charged, as well as multiply charged data and mixed adducts, among other features.

It can accept files in ASCII mass/intensity peak list format. See the Polymerix Guide for guidance in use. This is also a useful resource for better understanding polymer MS analysis in general and the "manual" process of polymer characterization.

PolyTools Analysis Example

File Input and Table Editing

PolyTools helps best for polymers with common, known, or suspected end groups and polydispersity <1.5. It can accept peak lists created with flexAnalysis or from other sources and mass spectra from flexAnalysis. The easiest way to use it is to bring processed files, which need to be well-calibrated and have a properly made peak list, directly into PolyTools by selecting *Tools* > *PolyTools* in the menu options of flexAnalysis. Alternatively, outside data can be brought into PolyTools through *File* > *Open* as ASCII text files in mass/intensity or mass/height/area peak lists that are space, tab, or comma separated. For the submitted peak list, it is important that all relevant peaks are identified and have a reasonable identified intensity for polymer characterization parameters to be calculated correctly.

PolyTools will automatically attempt to make repeating unit and end group suggestions based on the previously used parameters in the software. **Please be aware that depending on the last used settings and polymer sample, these default suggestions are not often correct, consider all, and only, the repeating units and end groups available in the loaded table, and are based on the allowed mass tolerance and minimizing the residual (e.g. will suggest adducts like Ag even if Ag wasn't used as an additive and would not make sense to form without addition).**

If the expected repeating unit or adduct is not already included in the drop-down options (the loaded table), then a table will need to be created and loaded. This is done in the separate "Polymer Table Editor" app on the computer. Open the default table file std.mtb in the folder path D:\Methods\PolyTool. Then "Save Table" and save a duplicate of this renamed, such as including user initials. This renamed file can be edited to add alternate monomer, end group, and cation options for PolyTools to draw from. Guidelines for this are given in *? >*

PolyTools > PolyTools Manual > Section 5. Table Editor. The new table must be reloaded in PolyTools through *File > Load Table*.

PEG700 Example

This example is with the mass spectrum of a polyethylene glycol (PEG700) standard. This is a relatively simple spectrum where end groups are already known. Beginning with the spectrum open in flexAnalysis, the first step is processing the data. First confirm the spectrum is not processed by selecting the *Process* tab. If “Undo All Processing” is greyed out, then it is not processed. Either keep or delete, as desired, already applied processing settings. This is starting with an unprocessed spectrum.

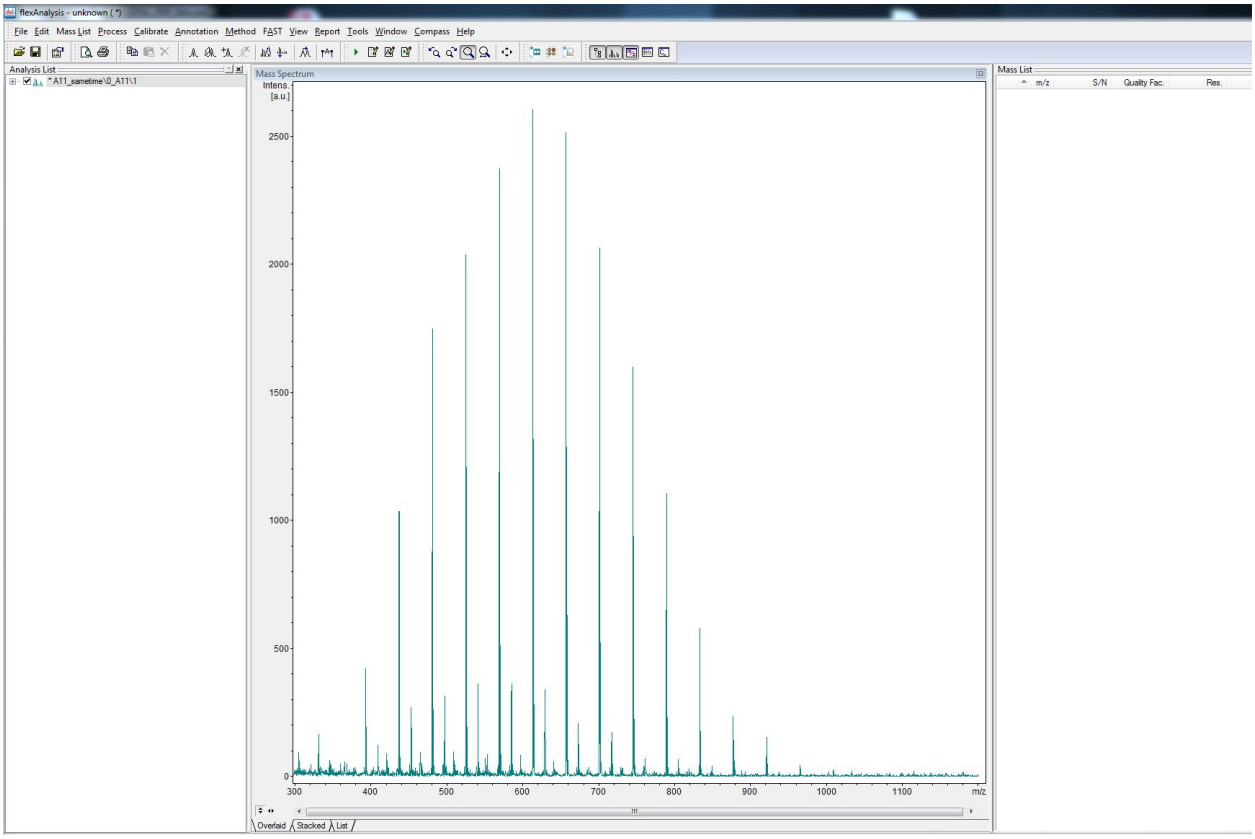


Figure 1. Unprocessed PEG700 spectrum in flexAnalysis

This was done in reflectron positive ion mode, so the spectrum has isotopic resolution. Zoomed out, there is an obvious main repeating unit series and what appears to be 2 other repeating unit series of lower signal intensity. These are the 2 signals seen between the larger repeating peak signals, where the first is ~+16 *m/z*, and the minor second is ~+28 *m/z* from the preceding “main series” peak. There are also other minor peaks that do not clearly belong to a repeating unit series.

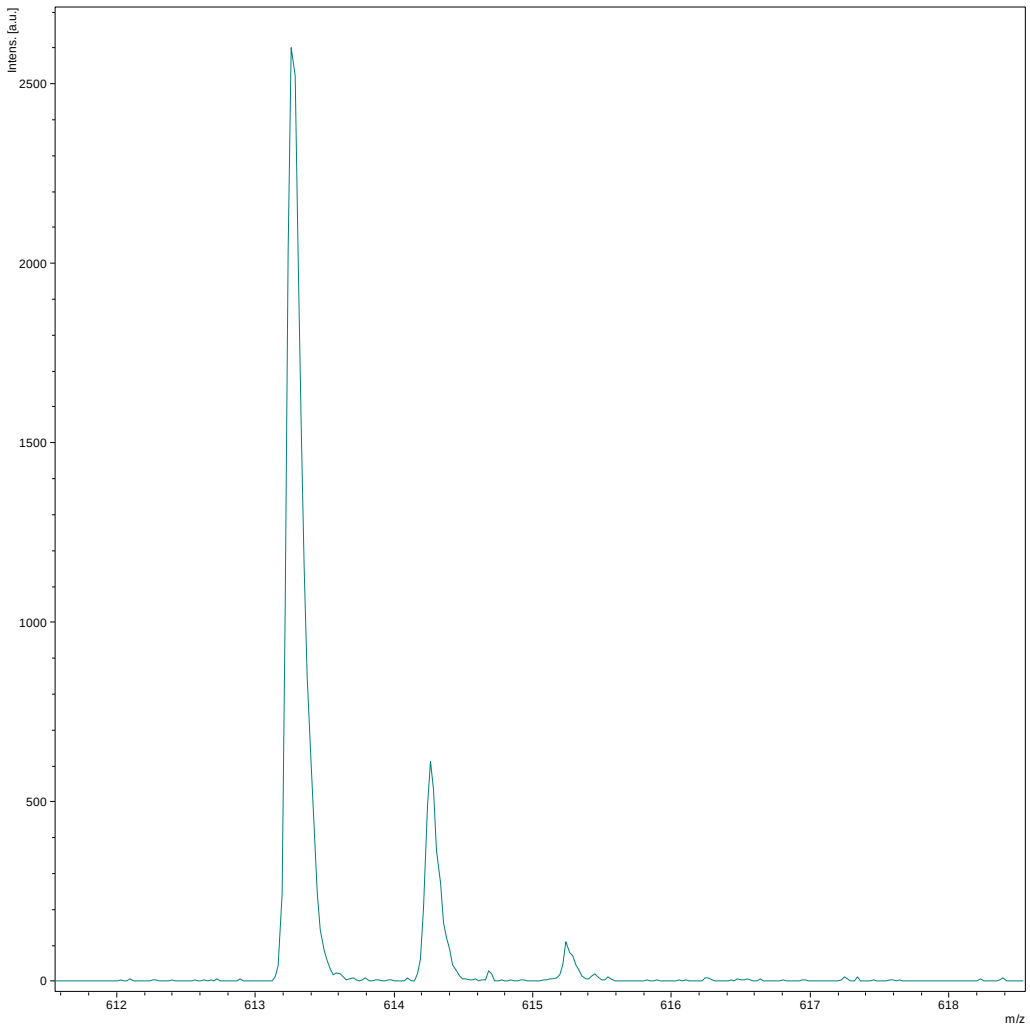


Figure 2. Zoomed in view of the highest intensity peak cluster in the main repeating unit

Zooming into the highest intensity peak at around the 613 *m/z* part of the spectrum, shows that what appears as a single peak zoomed out, is actually multiple peaks for all larger signals. Here, this is the isotopic pattern. The sample is a PEG, so the +1 and + 2 *m/z* peaks from the first peak can be assumed to reflect combined contributed signal from primarily a single ^{13}C replacing a ^{12}C for the first +1 *m/z* trailing peak, and primarily combined contributed signal from both two ^{13}C for ^{12}C replacements as well as an ^{18}O for ^{16}O, though interpretation of this changes depending on end groups, adducts, and chain length. While this is important to spectral interpretation, this is not something that is understood or needed by PolyTools. De-isotoping can be done in PolyTools, but otherwise, PolyTools cannot interpret isotopic peaks unless the loaded table included isotopic entries.

This zoom is important for determining processing parameters. To view settings, select *Method > Edit Processing Parameters > Processing*. The spectrum here was lightly smoothed (*Process > Smooth Mass Spectrum*) for better peak identification and integration. Smoothing depends on the data needs, but is usually done as lightly as needed while still preserving data

features of interest. The observed peak width is important for selecting appropriate smoothing width. Most data needs would use the SavitzkyGolay algorithm option in flexAnalysis.

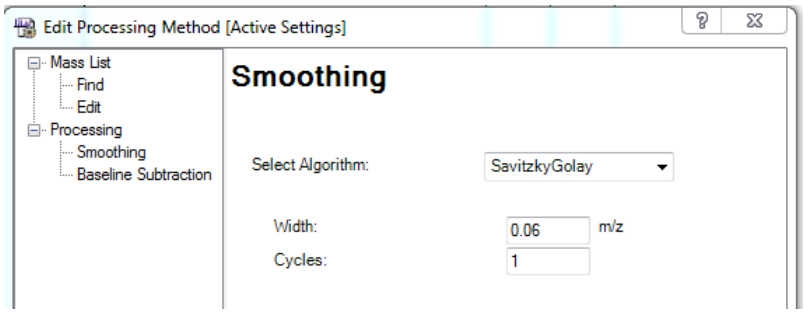


Figure 3. Smoothing settings used here

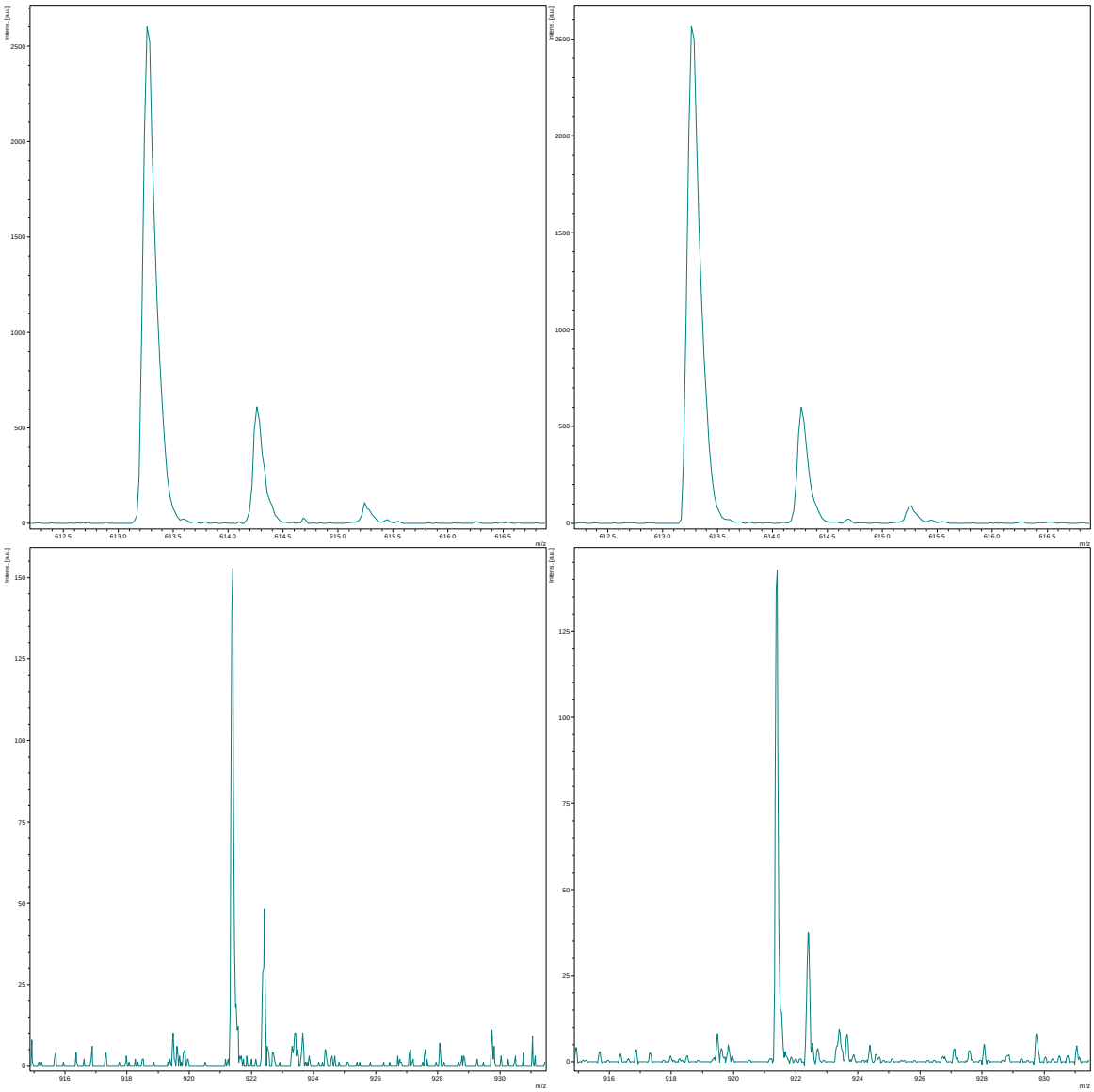


Figure 4. The effects of the selected smoothing (left=unprocessed, right=smoothed) on larger (top) and smaller (bottom) peaks in the spectrum

Smoothing, particularly over-smoothing, can introduce spectral biases and errors through smoothing artifacts. It is important to zoom in and check that smoothing is having the desired effects across the spectrum range.

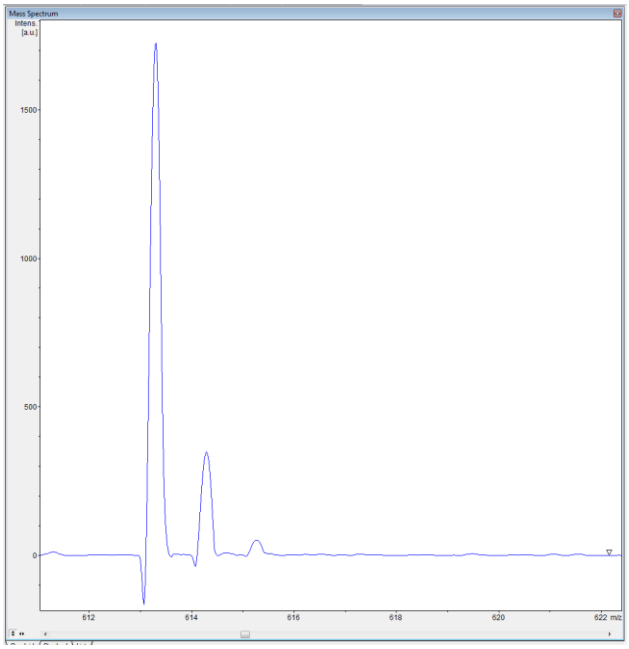


Figure 5. An example of SavitzkyGolay, 0.2 *m/z* width, 1 cycle smoothing introducing smoothing artifacts

The baseline was subtracted (*Process > Subtract Mass Spectrum Baseline*) using the TopHat algorithm, though this was not necessary for this data, as the baseline was level and not elevated. Processing and Mass List settings can be saved and reloaded to save time when processing similar spectra.

After smoothing and baseline subtraction, as applicable, the next step is creating a peak list. Peak list settings are also within “Edit Processing Method”. The Mass List Find settings refer to what automatically happens when the “Find Mass List” button is pressed to identify peaks and create a peak list. For detailed guidelines see the flexAnalysis Manual (flexAnalysis 3.4 User Manual.pdf on the computer). Briefly, Snap is meant to only identify the *m/z* of the first peak (usually the monoisotopic) in a closely spaced cluster, assuming it is an isotopic cluster, but with the area calculation from the cluster. This is more commonly used in peptide analysis. For all peaks to be considered separately, even in an isotopic cluster, and for data not at isotopic resolution, Centroid should be selected as the peak detection algorithm. Other parameters are explained in names and the manual, but key considerations are Maximal Number of Peaks and Peak Width in the Centroid algorithm.

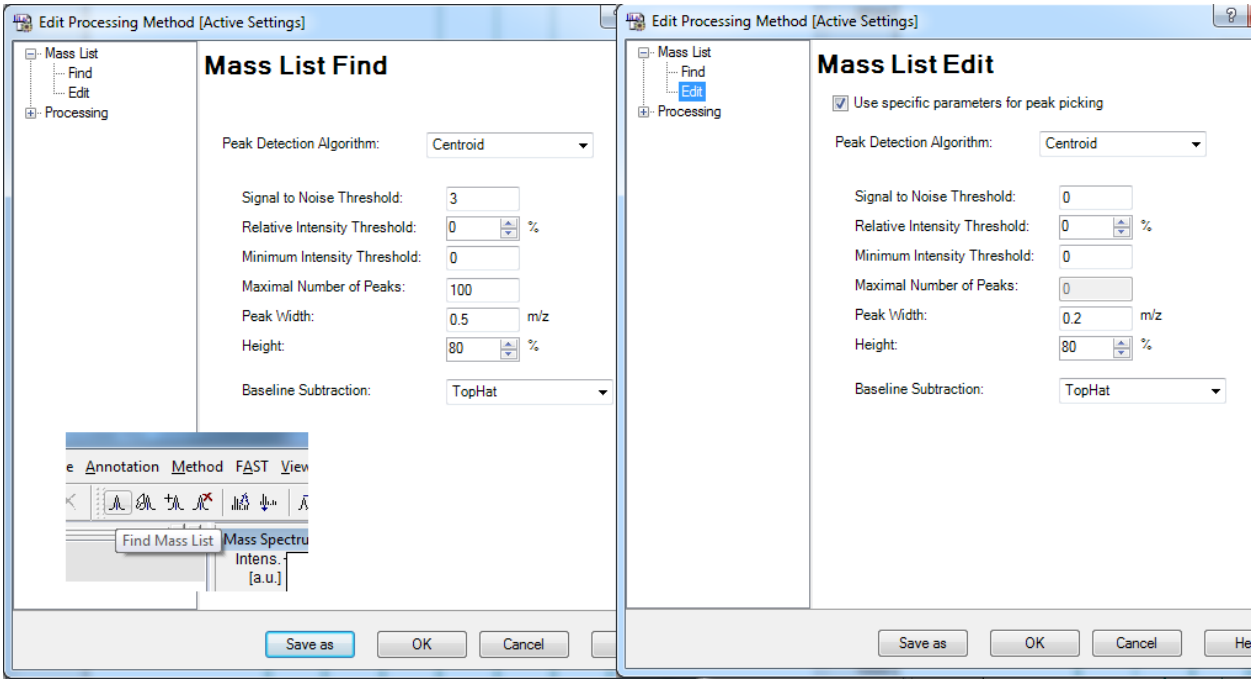


Figure 6. Peak list find settings (left) used here for automatic peak picking and peak list edit settings (right) used for manual peak picking. The inset shows the “Find Mass List” button.

Maximal Number of Peaks must be edited to fully cover the spectrum of interest. For example, in a well-resolved spectrum of a larger average molecular weight polymer, number of relevant peaks can reach into the hundreds or higher, and a max number of identified peaks of 100 would not be appropriate for characterization. Peak Width is an expectancy value and has strong control on recognition of peaks and accuracy of integration and must be set appropriately for the spectrum. Peak Height refers to the upper part of the peak used to identify the *m/z*, 80% is a typical setting, but may need to be adjusted depending on the extent of peak asymmetry.

Centroid was used with the above settings (Fig. 6), as Snap would be subject to baseline bias on isotopic peak visibility, though PolyTools results would be similar. Signal-to-Noise and peak recognition restrictions were kept low because the data had good resolution, minimal complexity, and small peaks that were clearly real data. No restrictions were wanted for manual peak list editing, and manual peak list additions were all very small peaks in repeating unit series that were missed by the initial automatic assignment.

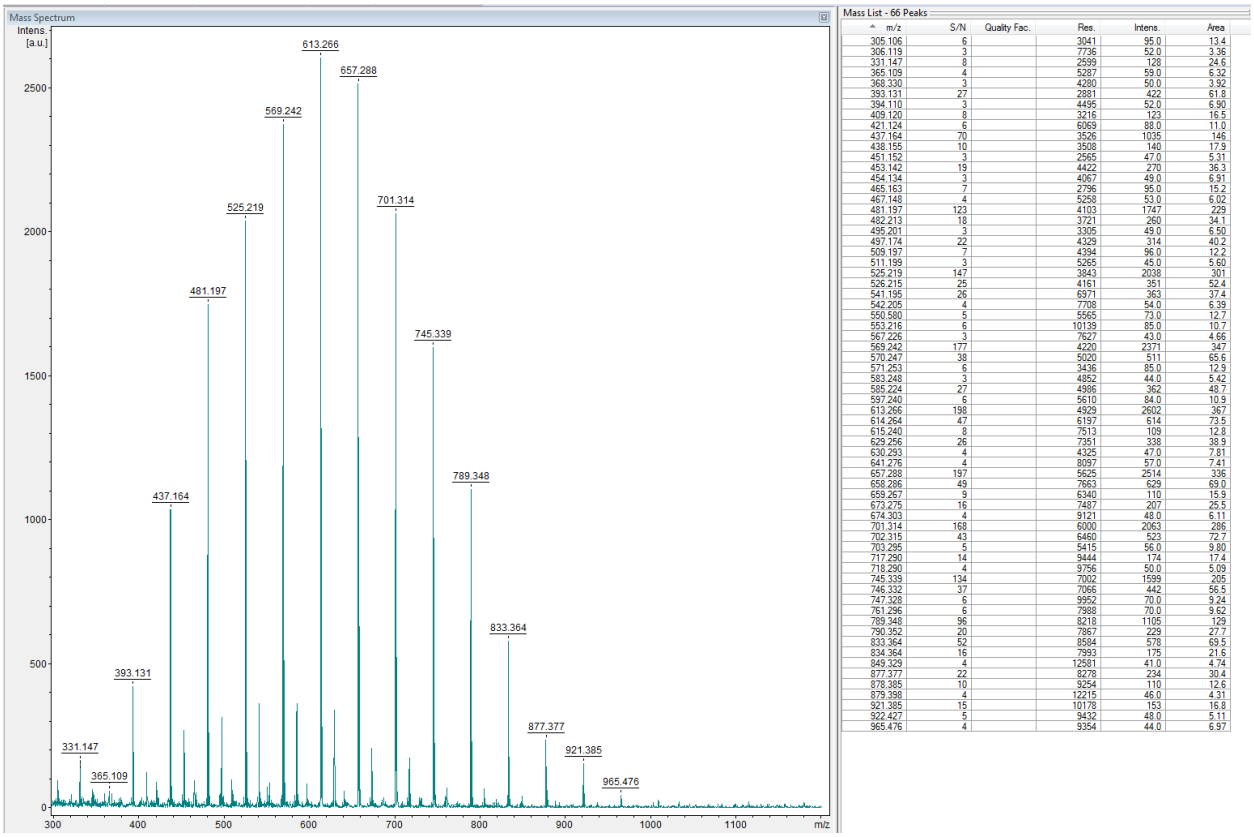


Figure 7. Automatic peak list assignment

Pressing the “Find Mass List” button identifies peaks using the Mass List Find settings (Fig. 6). While this is a good initial peak list to start from, there is some assignment of what appears to be background noise or ions, and under-assignment of the smaller peaks of interest that were part of repeating unit series. To address this, “noise” or undesired peaks were deleted from the peak list.

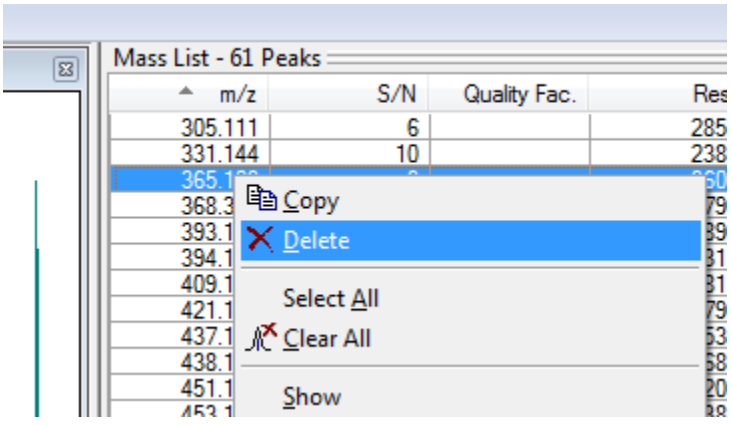


Figure 8. Deleting undesired peak assignments

Then peaks of interest that were not initially identified by the first “Find Mass List” pass, were identified through the “Mass List Edit” button (to the right), clicking them to manually add them to the peak list. This uses the Mass List Edit settings.

Some “real data” peaks may still not be recognized through Mass List Edit peak picking settings, even with low restrictions. These can be force-added to the peak list through “Edit Direct”, the button to the right of Mass List Edit. However, if used, this only adds *m/z*, intensity, and relative intensity directly from coordinates because it does not use a peak picking algorithm, and so intensity may not correctly match that of the surrounding peaks. This could bias polymer parameter calculations. The right-most button in this row clears the peak list.

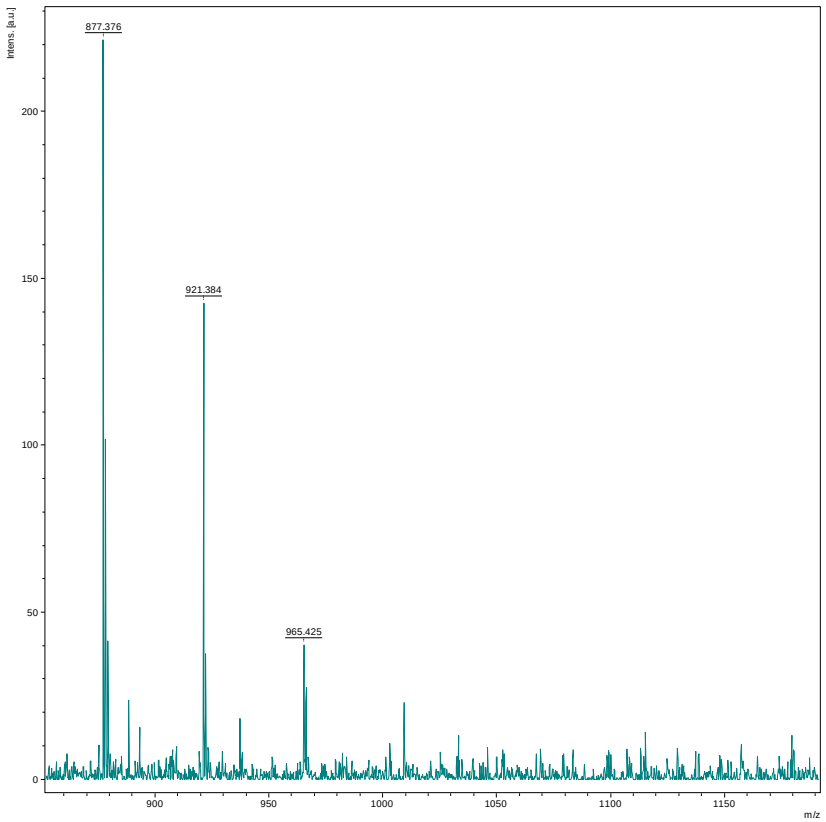


Figure 9. Examples of small peaks (unlabeled) that were initially not identified for the peak list before manual addition

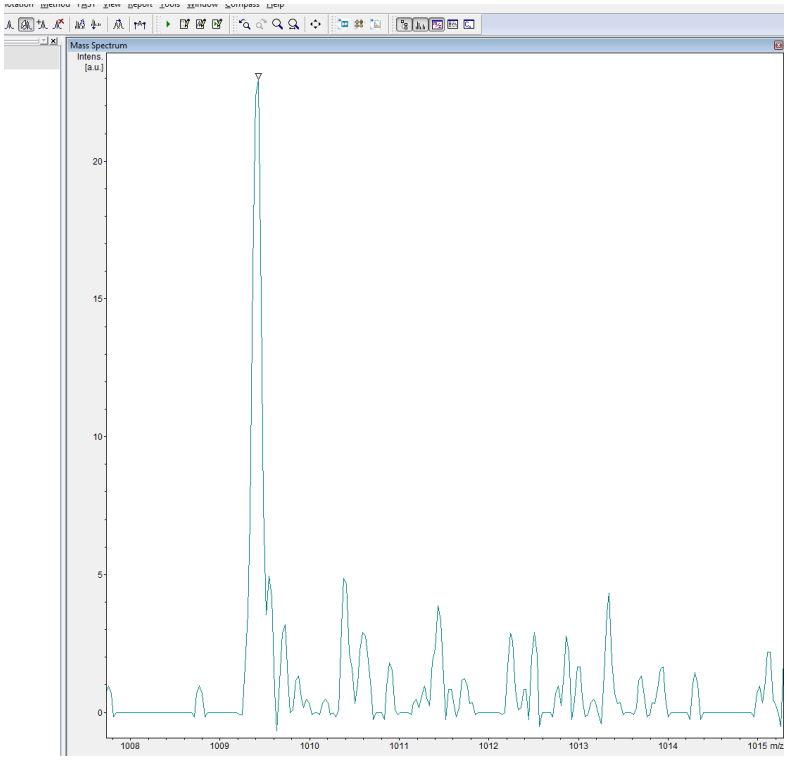


Figure 10. Zooming in for manual peak picking through Mass List Edit (triangle pointer for selection)

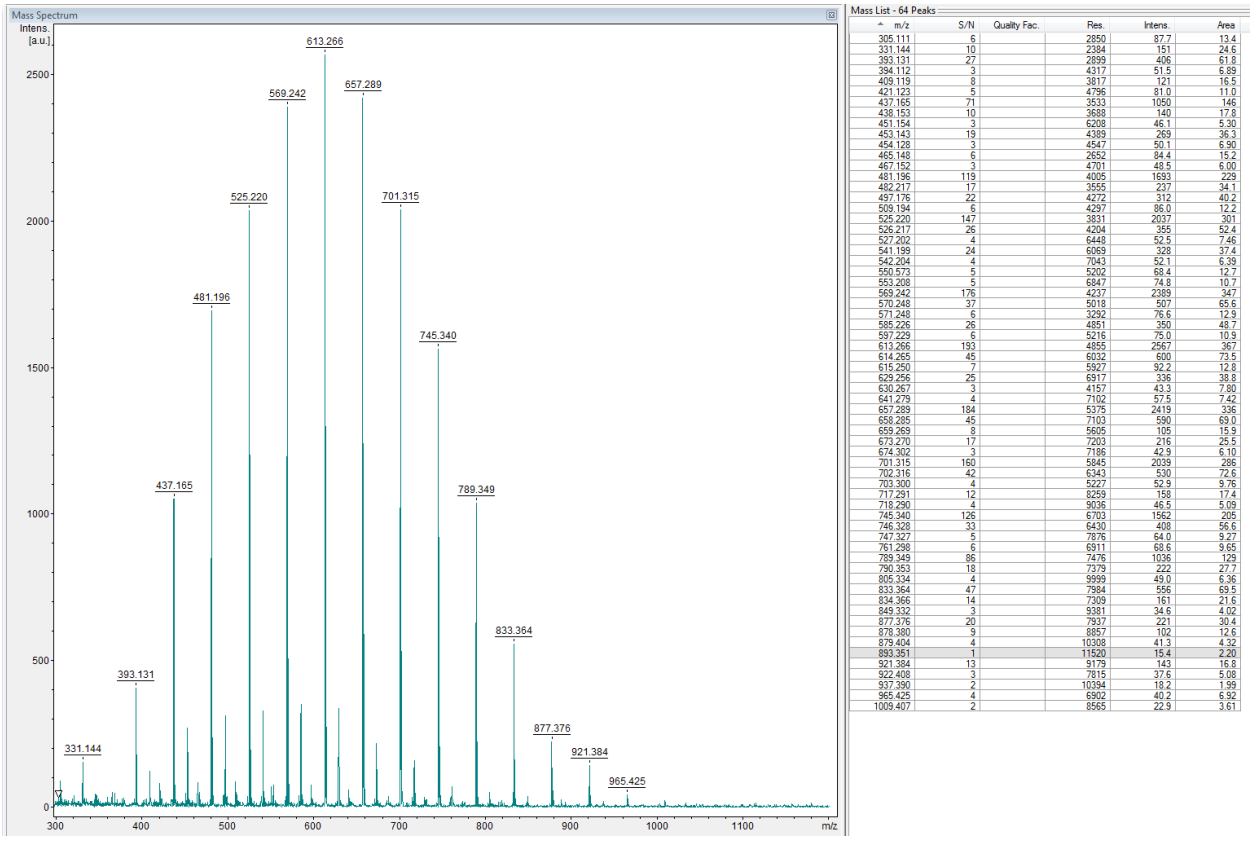


Figure 11. Final processed spectrum

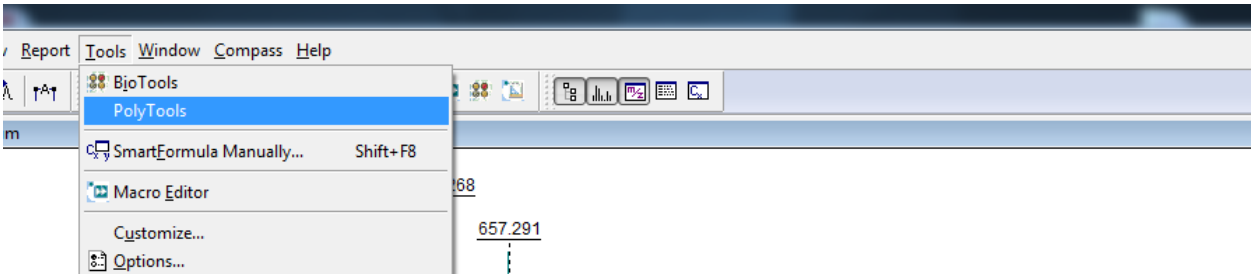


Figure 12. Bring processed spectrum from flexAnalysis into PolyTools

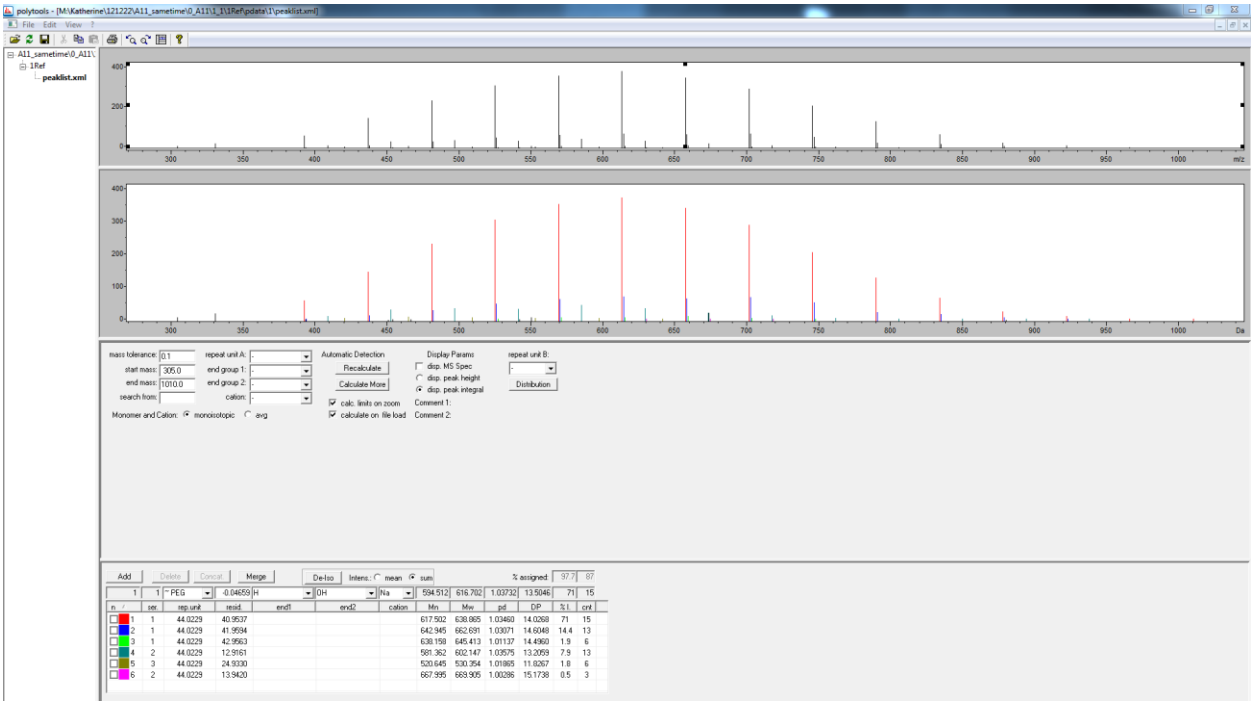


Figure 13. PolyTools default suggestions

The spectrum is ready to be analyzed in PolyTools. PolyTools tries to automatically identify repeating units, end groups, and adducts using the loaded table. In the second plot, *m/z* not identified as part of repeating series are in black. These may be matrix peaks, background contamination, or other ions in the analyzed sample. The first identified repeating unit series (n1, red) is suggested to have a PEG repeating unit, H and OH end groups, with sodium adduct ionization $[M+Na]^+$. In this table, the residual currently reflects *m/z* unaccounted for by the repeating unit, so the sum of end group 1, end group 2, and any adducting ion.

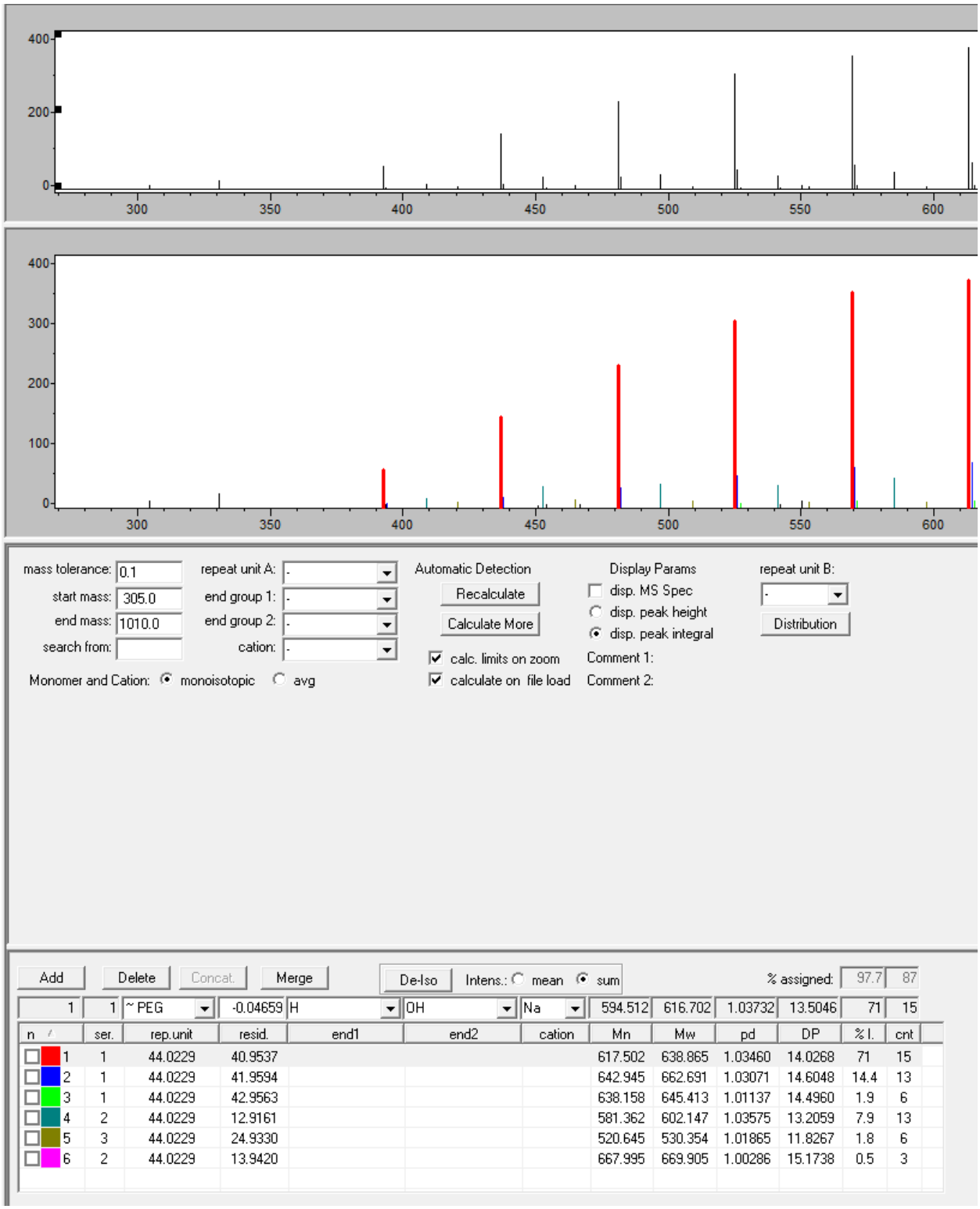


Figure 14. Selecting first identified series

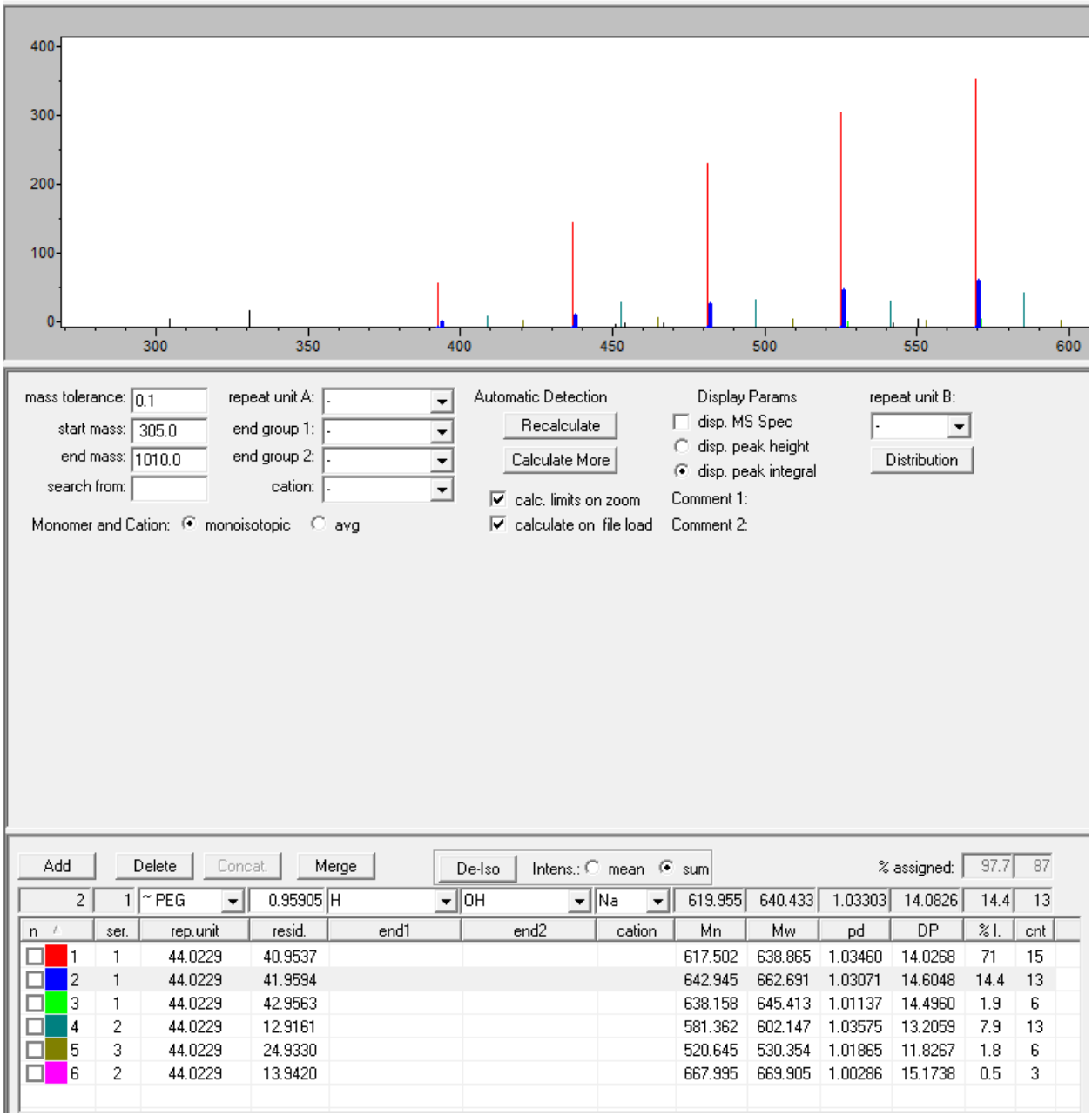


Figure 15. Selecting second identified series

Here this first series is correct, though other default suggestions are incorrect. The second and third series are isotopic series of n1. Next is to put in known parameters. It was already known from the standard that the repeating unit was PEG (C_2H_4O) and H and OH should be end groups.

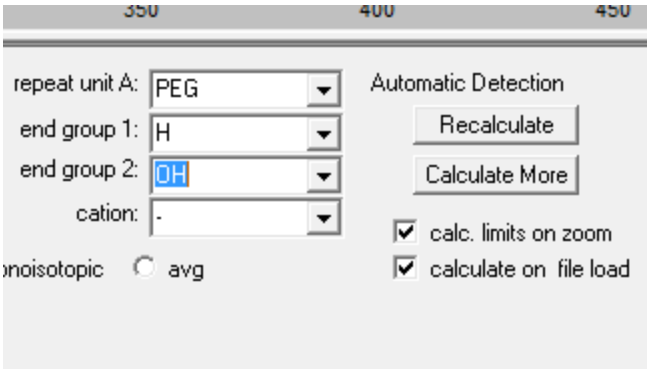
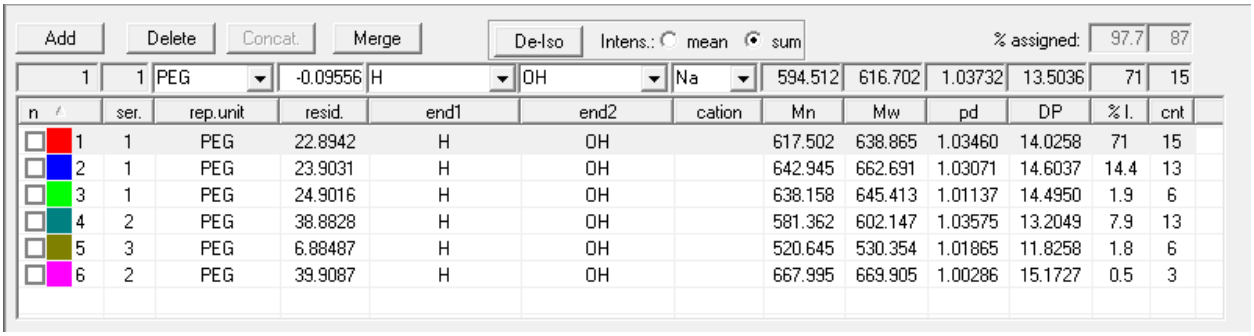


Figure 17. Changing input end groups and recalculating

These are input and "Recalculate" is pressed. For this spectrum mass tolerance was kept at 0.1, and can be seen to catch all expected repeating unit peaks, but this needs to be selected based on the spectrum being analyzed. Suggestions can also always be manually edited and added to view calculations.



n	ser.	rep.unit	resid.	end1	end2	cation	Mn	Mw	pd	DP	% I.	cnt
1	1	PEG	22.8942	H	OH		617.502	638.865	1.03460	14.0258	71	15
2	1	PEG	23.9031	H	OH		642.945	662.691	1.03071	14.6037	14.4	13
3	1	PEG	24.9016	H	OH		638.158	645.413	1.01137	14.4950	1.9	6
4	2	PEG	38.8828	H	OH		581.362	602.147	1.03575	13.2049	7.9	13
5	3	PEG	6.88487	H	OH		520.645	530.354	1.01865	11.8258	1.8	6
6	2	PEG	39.9087	H	OH		667.995	669.905	1.00286	15.1727	0.5	3

Figure 18. Suggested repeating unit series after repeating unit and end groups were entered

Now the residual only reflects *m/z* unaccounted for after consideration of the repeating unit and end groups, so the ~23 is Na, ~24 and 25 are the +1 and +2 isotopic peaks of this series, ~39 is ^{39}K, ~40 is the +1 isotopic peak of this series, and ~+7 is the third repeating unit series. Isotopic series were deleted because they were unneeded here. This could also be handled by using the De-Iso button.

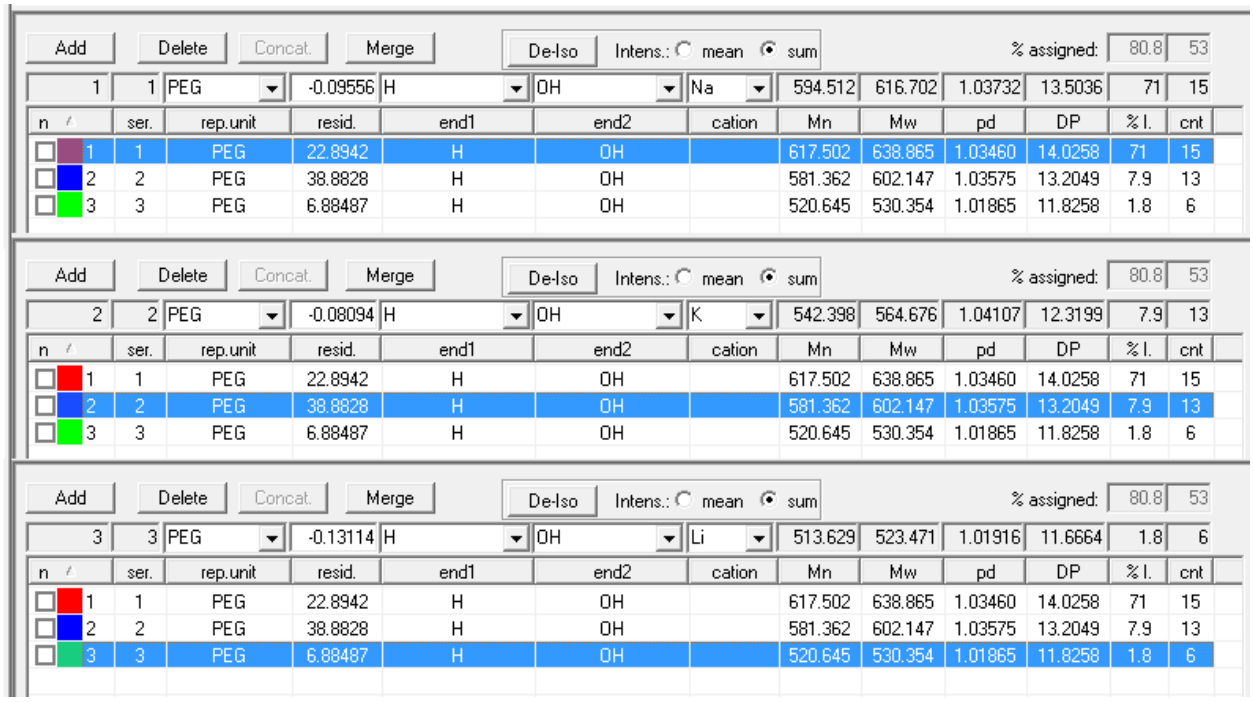


Figure 19. New suggested series

PolyTools suggests these to be $[M+Na]^+$, $[M+K]^+$, and $[M+Li]^+$ ionization forms. These identifications have low residuals, so they are good mass matches. The residual now only reflects unaccounted for *m/z* after consideration of repeating unit, end groups, and cationization. The first two of these PolyTools suggestions are correct, and make sense as PEGs preferentially ionize though alkali metal ionization pathways and Na^+ and K^+ are commonly present from glass and other materials used during sample storage and preparation. While a good mass match, lithium adducts would not be expected to form without an addition or specific contamination source. This third minor series, which was +28 *m/z* from the main sodium adduct series, might reflect something like a carboxyl end group (COOH vs OH), $[M+Na]^+$, due to degradation or initial standard end group purity. While interesting, this minor series was deleted for parameter calculation.

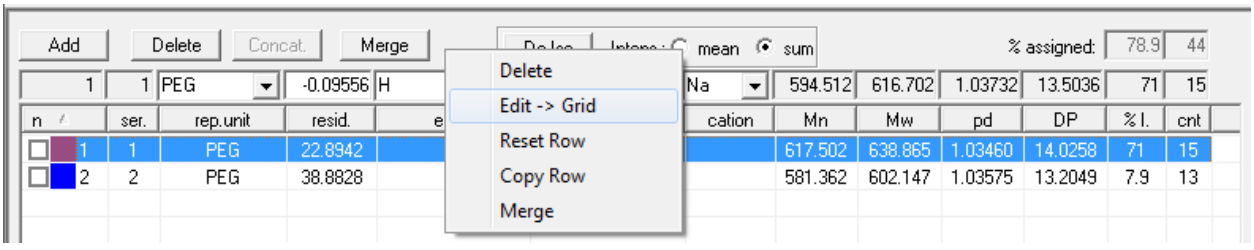


Figure 20. Select Edit -> Grid to calculate considering suggested end groups and ionization

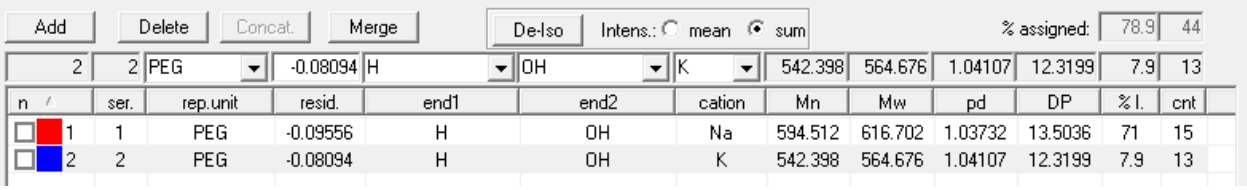


Figure 21. Final calculated results

After selections are added to the grid, parameters are recalculated to reflect the assignments. For example, Mn goes down to account for contributed mass from the adducting cation, so it reflects the polymer. Mass residuals are after summing contributions from all parts and are very low and can be accounted for by mass accuracy and error.

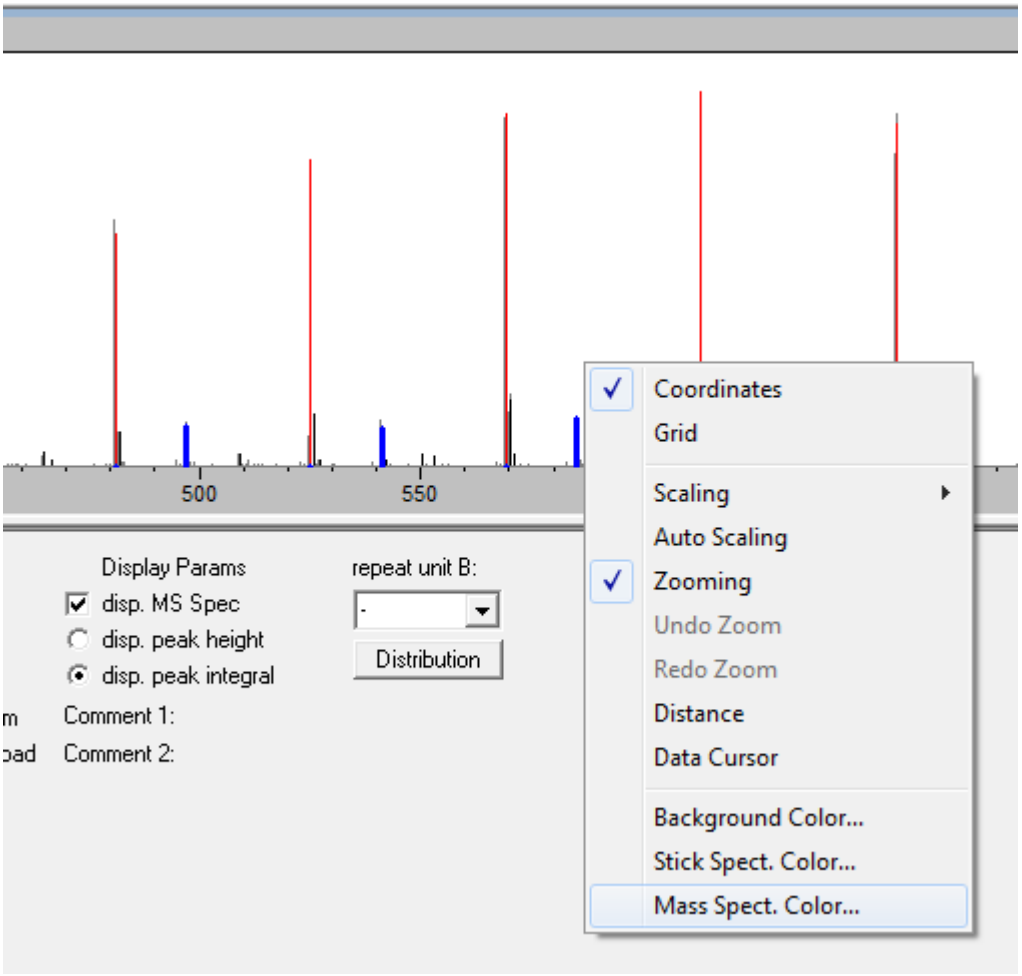


Figure 22. Plotting MS spectrum (grey) and selecting color

Depending on how results were generated, the spectrum can usually also be added to the repeating series intensity plot by checking the “display MS Spectrum” box. The color of this can also be edited. Unfortunately, the plotting color of peaks not assigned to a repeating series is the same as the plotting color of theoretical *m/z* (here black), though this color can be changed with “Stick Spect. Color”, they both change together. Plot background color can also be changed. Series colors are standardized to series number and cannot be edited.

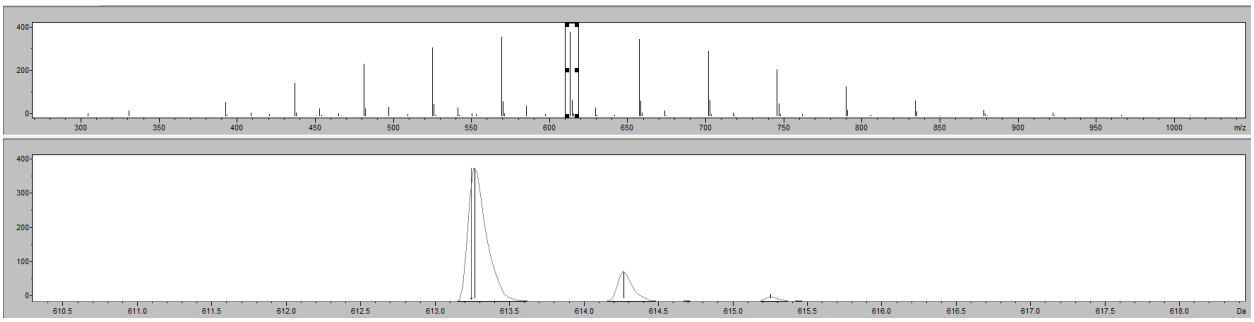


Figure 23. Zoomed in view of the highest intensity peak in the main series, showing the experimental spectrum (grey), experimental *m/z* from the main series in the peak list (red), and the theoretical *m/z* (black)

References

Patiny, L., Borel, A. (2013). ChemCalc: a building block for tomorrow’s chemical infrastructure, *Journal of Chemical Information and Modeling* 53(5), 1223-1228.

Loos, M., Gerber, C., Corona, F., Hollender, J., Singer, H. (2015). Accelerated isotope fine structure calculation using pruned transition trees, *Analytical Chemistry* 87(11), 5738-5744.