MsMetrix

MsXelerator Tutorial:
Impurity Profiling and Peak Purity Analysis: Part III

This document describes some new features and tools for Peak Purity Analysis. Please also check the other documents on impurity profiling and Peak Purity Analysis.

1. Check Impurities using the HeatMap Procedure in the Browser:

From the Browser zoom in around your Peak of interest. To create a HeatMap use: Plots > HeatMap. Depending on the Resolution of your instrument you can also make HeatMaps for high resolution data. Figure 1 shows a heatmap in nominal mass resolution (1 Da) that zooms on the data at a level of 1% compared to the main peak in this region. You can clearly see co-eluting peaks at different masses. Zoom in to see the masses for the peaks, or use the data cursor tool from the Icon-Bar. Click on a peak to see the (x,y) information (time, mass).

When selecting the High Resolution HeatMap option, the mass resolution will be 0.05 Da by default. Please zoom in on the peak of interest and to reduce the amount of data for creating the plot.

The graphical procedure using a HeatMap can be of great value to quickly see impurities, especially if the impurities have masses close to the mass of the Main Peak. Sometimes it will be less easy to see if the peaks are really co-eluting or very close in retention times. To check, Overlay the individual mass traces (EIC) in the Browser. You can enter the masses to plot in the Select m/z window (on the right) or just click on a mass in the Mass Spectrum Plot. To overlay a second EIC, hold the Ctrl key and click on another mass in the Mass Spectrum. If you want to do High Resolution EIC plotting, be sure that the option: Options > Use Direct Accurate EIC Plotting from MS plot is checked. The resolution of the plot can be set from the options menu or entered in the Resolution Field above the MS Plot.

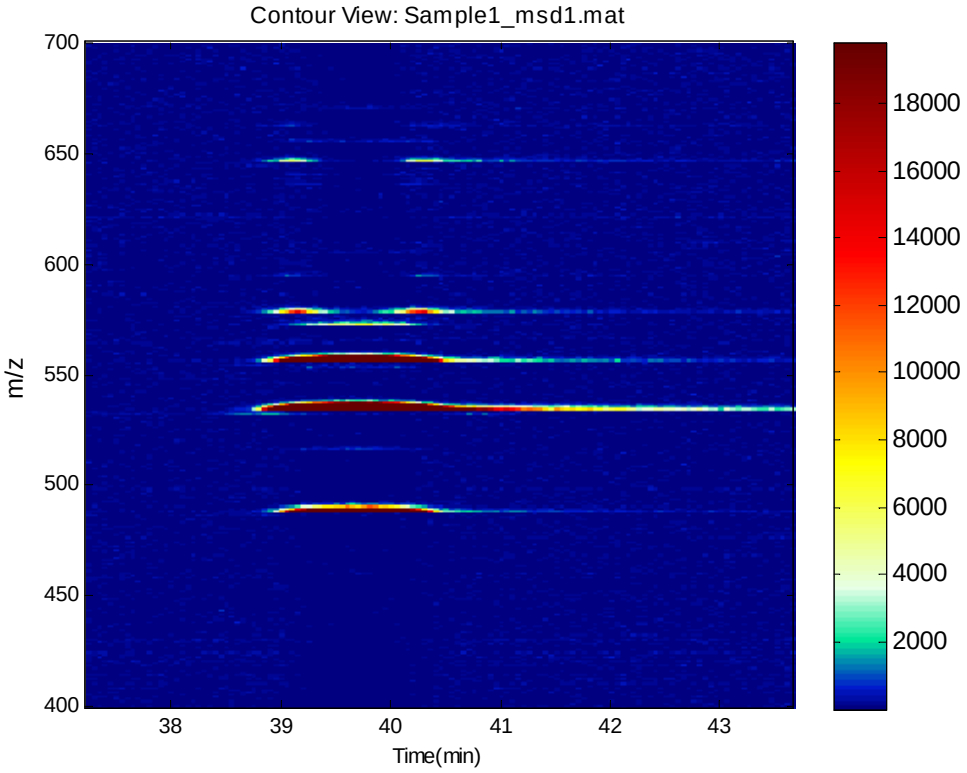


Figure 1: Nominal HeatMap Plot showing the main peak and its impurities.

2. Perform Peak Purity Analysis using Principal Component Analysis:

Principal Component Analysis (PCA) is a multivariate data processing technique that is able to check the purity of your peaks. This technique can be applied to LC/MS and GC/MS data. The procedure is easy and can be very sensitive to detect minor impurities close to the main peak. PCA will not be able to find impurities if these are perfectly co-eluting with the main peak. In such a case these impurities are regarded as fragment peaks or adduct, being part of the main peak.

In the Brower, zoom in on your peak of interest. From the menu select: Task > GCMS: Peak Purity PCA. You will be asked to set two points for baseline correction in the TIC window. Click before and after the peak to select a proper range. Normally you will use both markers to do a baseline correction, see Figure 2.
The next question is to set a threshold on the peak intensities compared to the main peak. Try to use the Relative procedure. Select a threshold of e.g. 2.5% relative to the intensity of the main peak in this window. To perform more sensitive peak purity analysis you can set a value of 1% or even lower. This depends on the noise in your data and the type of application.

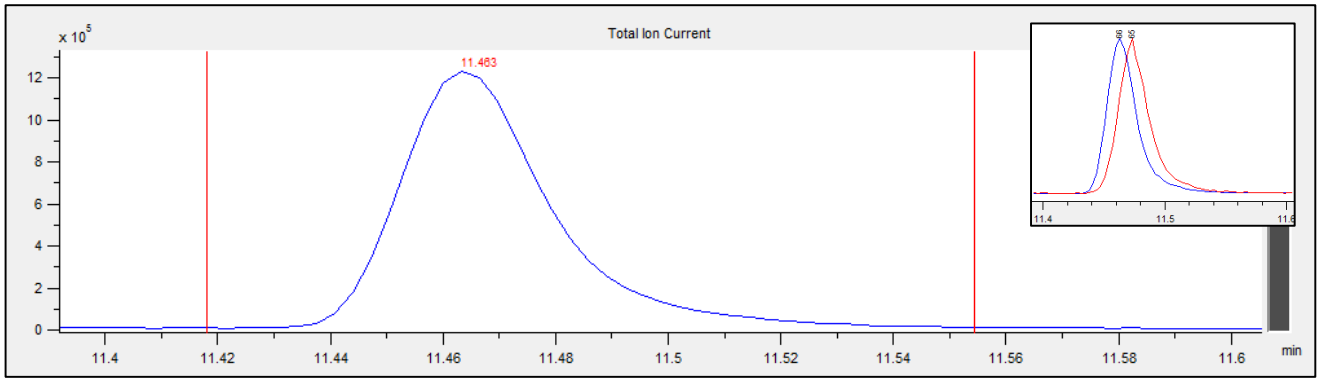


Figure 2: Selected Region for Peak Purity Analysis using PCA procedure (GC/MS Example). The inset show normalized mass chromatograms for the main peak and one of the impurities.

During the next question, select to perform the PCA on the raw data. A graphical overview as shown in Figure 3 will be shown. Figure 3 shows three sub-plots.

The top window is the PCA Score Plot. It shows the distribution of the (nominal) masses in two Principal Component dimension and is useful to see which masses cause the main variation. The first principal component (PC) explains 99.6% of the variance in the data. This relates to the main peak and all co-eluting masses. However, the second PC still explains 0.4% of the variance. Not much, but clearly more than we would expect from random noise.

The largest mass in the first PC is m/z 66. The largest mass in the second PCA dimension is m/z 65. You can overlay the mass chromatogram traces in the Browser, to see the retention time shifts. (see inset of **normalized** traces in Figure 2), a clear impurity, eluting to the right of the main peak.

The second sub-plot of Figure 3 is the so-called residual loadings plot. It displays the chromatograms after the Peak has been corrected for the variation of the first dimension. In the case of no impurities, this plot would display noise only. From figure 3B, you can clearly see that the traces are not zeros or noisy only. There is a very clear structure present. This means that an impurity is present, the main mass for this impurity is m/z 65 (check the second pc in subplot 3A).

Subplot 3C shows the Residual Plot after correcting for the Main Peak **and** the variance of the impurity from figure 3B. You can now see that the plot is dominated by noise. This means that probably no additional impurities are present at a level near the noise level.

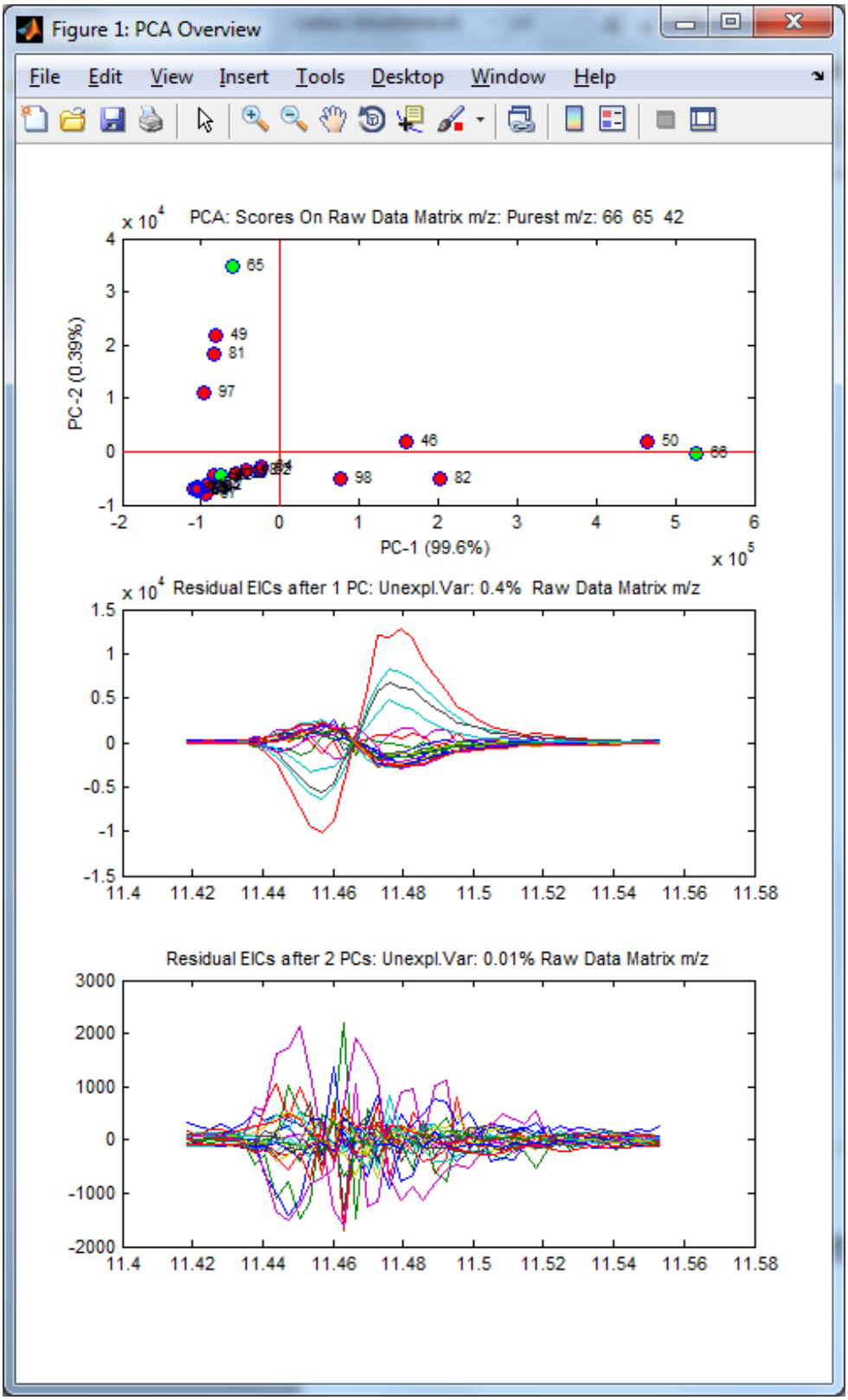


Figure 3: PCA Overview Plot: top – PCA score plot showing all masses in two dimensions. Middle – residual loading plot showing clear structure which means presence of a extra component. It only explains 0.4% of the variance in the data, so this must be a small impurity. The bottom plot shows that after two pc's no more un-explained variance is present.

3. Check Peak Purity using MPeaks Peak Picking and Clustering:

The last procedure for Peak Purity Analysis is to use MPeaks to find all peaks in the data. From the Browser start MPeaks. Try to use default peak picking settings as a first step. See MPeaks tutorials and the Manual for a description of MPeaks and its settings. Take some time to optimize the MPeaks settings.

Attention: If you have high resolution Centroided data, you might want to use High Resolution Peak Picking. To do this, activate the Accurate m/z check box in the upper Peak Picking Parameters Window.

Running Peak Picking:

Zoom in on the peak of interest and select: Current, from the Time Parameters Window. This will use only the zoomed region for peak Picking. Press the Run button from the buttons displayed at the bottom of the window. MPeaks will produce a (nominal) peak picking table, showing all the detected peaks, their masses, retention times, Heights, relative Height compared to largest Peak, Area etc. etc. The largest peaks will have a relative height of 100%.

The table can be sorted on mass, retention time or height. If applicable, nominal mass table can now be converted to accurate mass using the Accurate Mass Features from the Menu. For low resolution data it is probably best to only use the Average Accurate Mass Conversion option.

Exploring Peak Picking results:

You can directly explore the detected peaks in MPeaks by selecting masses using the mouse. The chromatograms will be plotted in overlay. You can normalize the traces and automatically zoom in on the time range of interest. You can also easily switch between mass chromatogram and mass spectrum.

To graphically check all peaks near the mean peak for co-elution, sort the table on retention time and click on the main peak in the table first. Be sure that auto-zooming is active with the appropriate zooming range (set from the options). Be sure to overlay the Normalized TIC (Click on the TIC check box in the top right corner and select to plot in overlay). Now you can go through your table and visually detect co-elution of the detected peaks relative to the TIC. This is a fast and sensitive procedure.

You can also use the MPeaks Viewer, for some special plots. Figure 4 shows the so-called MPeaks Viewer Dot Plot, handy for Peak Purity Analysis. All detected peaks are plotted in a time / mass contour window. The TIC is plotted in overlay in this plot. From the GC/MS examp
le you can clearly see that a peak right of the main peak elutes. MPeaks can perform peak detection at very low levels. Very small peaks were left out in the plot below.

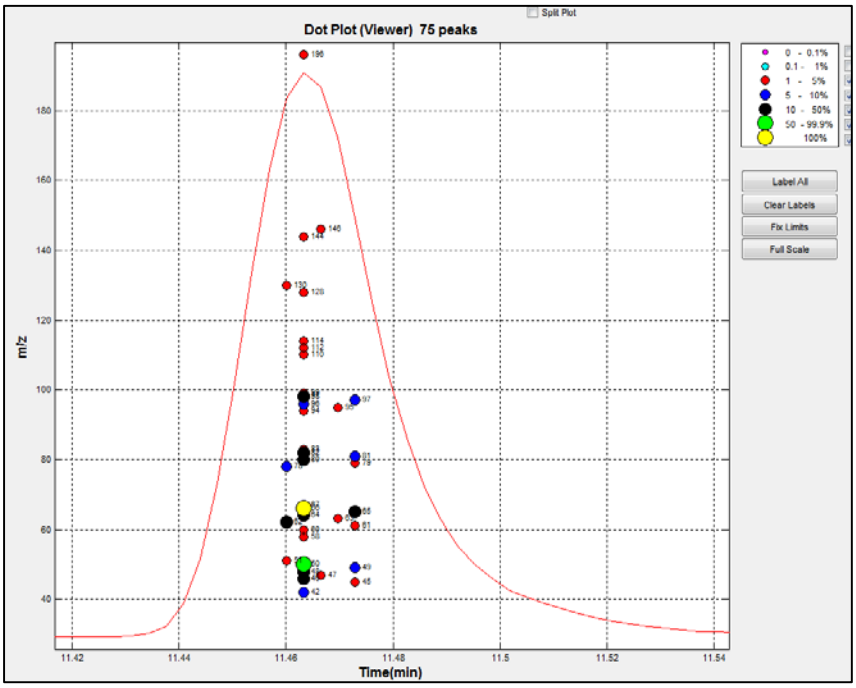


Figure 4: MPeaks Viewer Dot Plot (GC/MS example), showing detected peaks in overlay with the TIC. From the plot it is clear that a smaller peak elutes on the right of the main peak. Different Fragment ions are visible.

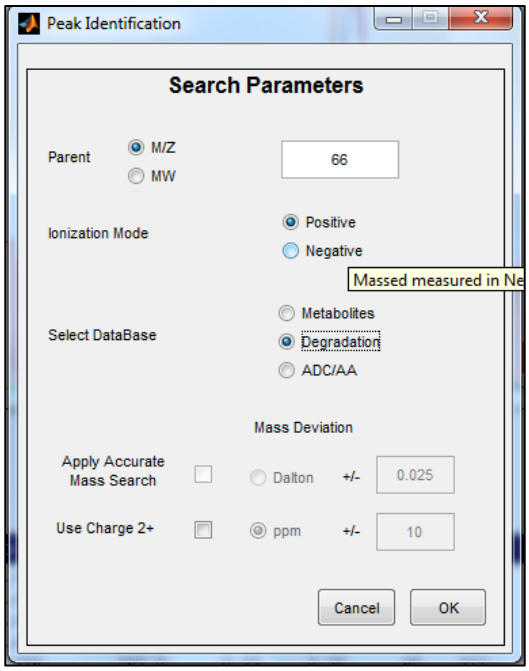
Clustering Peaks:
Clustering is a procedure to group peaks having the same peak shape and retention time. The goal is to group all fragment or adducts ions and to separate the main peak from the impurities. From the MPeaks menu select Clustering. You will get a window to set the clustering parameters. If you have high resolution data, clustering can be performed on high resolution mass chromatograms.

The third column of the table will now display the clustering value. Similar Peaks regarding retention time and peak shape will be in the same cluster. For Peak purity analysis check the cluster before and after the cluster in which your main peak lies. Overlay the peaks to see if they are co-eluting or not.

Identification:
For Degradation profiling, MPeaks can try to identify the detected peaks relative the mass of the main peak. The current data base has some 60 modifications like oxidation, reduction etc. the masses of all peaks in the table are compared to the mass of your product peak.

Click on the entry of your product peak. Next Start the Identification tool by pressing: Menu > Identification > Find Metabolites / Degradation Products. The mass of your "Parent" will be displayed in the m/z edit box. Select the type of ionization and the data base to use (Degradation). Press OK to perform the ID. The possible degradation products (nominal in the example) will be displayed in the table.

The data bases can be adapted by the user to contain list of ions applicable to your type of research.



Peak	m/z	Cluster	tR	PH	%PH	%Area	Degradation ID MW(65)	
1	66	5	11.463	264832.0	100.00	100.00	[+0] Parent Molecule	
2	50	5	11.463	235776.0	89.03	91.23	[-16] Reduction	R-CO-NH(OH) -> R-CO-NH
3	82	5	11.463	131840.0	49.78	49.93	[+16] Hydroxylation *	R-H -> R-OH
4	98	5	11.463	81568.0	30.80	30.39	[+32] Adduct Methanol *	H -> H(MeOH)
5	64	5	11.463	40048.0	15.12	15.82	[-2] Oxidation	Ring aromatisation
6	62	5	11.460	38544.0	14.55	15.50	[-4] Oxidation	Ring aromatisation
7	80	5	11.463	27336.0	10.32	10.87	[+14] Ketone *	CH2 -> C=O
8	96	5	11.463	16305.0	6.16	6.38	[+30] Oxidation	R-CH3 -> R-COOH
9	67	5	11.463	11668.0	4.41	4.44	[+1] Hydrolysis	R-CO-NH2 -> R-COOH + N
10	130	5	11.460	5074.0	1.92	1.86	[+64] Adduct ACN	H -> Na(ACN)
11	84	5	11.466	457.0	0.17	0.20	[+18] Hydrolysis	R + H2O
12	79	6	11.473	3909.0	1.48	1.61	[+13] Deamination	R-CH2-NH2 -> R-CH2-CH(
13	68	7	11.466	442.0	0.17	0.41	[+2] Reduction *	R-CHO -> R-CH2-OH

Figure 6: MPeaks ID, showing main product peak and possible degradation products