Deconvolution in TraceFinder software for unit-resolution GC-MS data

Authors
Amit Gujar[1], John Voss[1],
Jason Cole[1], Xin Zheng[1] and
Daniela Cavagnino[2]
[1]Thermo Fisher Scientific,
Austin, TX, USA
[2]Thermo Fisher Scientific, Milan, Italy

Keywords
Deconvolution, gas chromatography-mass spectrometry, ISQ 7000 GC-MS system, TSQ 9000 GC-MS/MS system, TraceFinder software, unknowns screening, target screening

Goal
To demonstrate the features of the deconvolution plug-in software application for Thermo Scientific™ TraceFinder™ Software, version 4.1.

Introduction
Gas chromatography-mass spectrometry (GC-MS) is a powerful analytical technique for identifying, confirming and quantifying organic compounds in complex matrices.[1,2] Its use covers a wide range of applications in several analytical fields - ranging from pesticides analysis in food, to organic contaminants in environmental samples, and to drugs analysis in biological fluids, as well as to organic impurities in pharmaceutical products. The gas chromatograph separates the analytes of interest from matrix interference and the mass spectrometer is used for identifying the compounds based on the unique fragmentation pattern of the molecule. Extensive libraries, for example NIST/EPA/NIH Mass Spectral Library, are available to compare the generated mass spectra in order to ascertain the identity of the compounds. For a unit mass resolution instrument like the single quadrupole mass spectrometer, the library matching score along with the compound's retention time are used to confirm the identity.

When analyzing compounds in complex matrices, one common occurrence is co-elution of compounds of interest or co-elution of analytes with matrix interferences. Figure 1 shows an example where there are two pesticides co-eluting to form a single chromatographic peak. A NIST library search done at the apex of the peak identifies the compound as Simetryn, but on closer examination of the actual and reference spectrum, one can observe that some of the mass spectral fragments (like *m/z* 184 and 129) are prominently missing in the library spectrum of reference.

Thermo Fisher
SCIENTIFIC
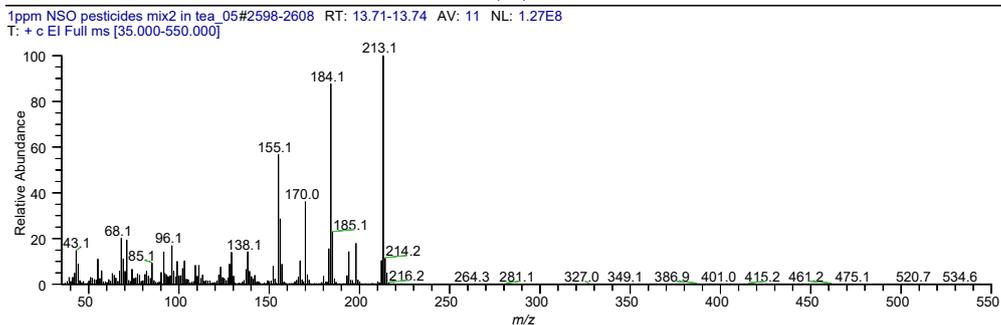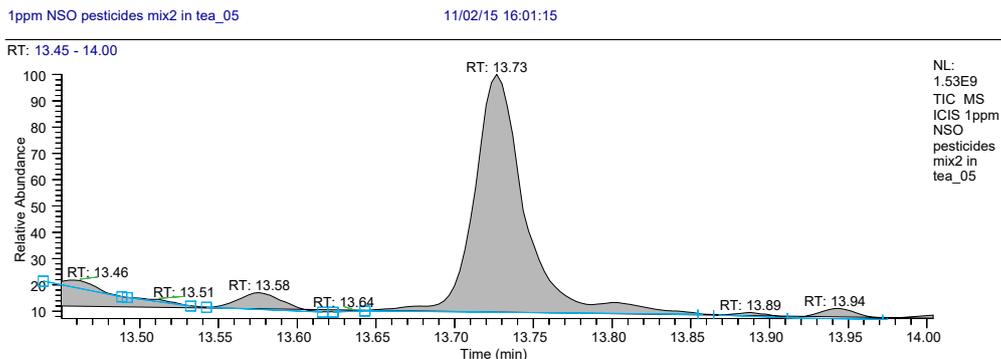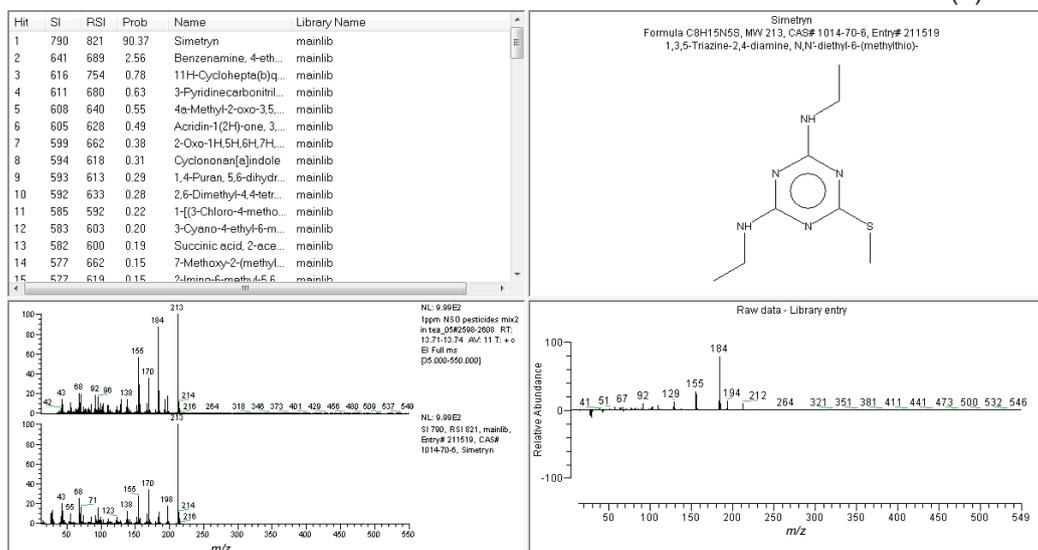
Figure 1. (a) Chromatogram and mass spectrum of a peak found in tea matrix spiked with pesticides; (b) library search of the same mass spectrum showing the NIST identification

By extracting the ion chromatogram (XIC) for masses 184 and 213 in comparison with the total ion chromatogram (TIC), as shown in Figure 2, it is evident there are two different retention times, 13.72 and 13.74 minutes, for two peaks indicating the presence of two different compounds. By performing a library search on the spectrum at 13.74 minutes while subtracting the mass spectrum at 13.72 minutes, the compound can be identified as Fuberidazole.

This manual process of deconvoluting a peak is tedious and requires extreme care and attention, and in a case where the analyte concentration is so low that it is hidden in the noise, the peak cannot be identified.

The deconvolution plug-in application for TraceFinder 4.1 software is designed to automatically deconvolve co-eluted chromatographic peaks into multiple components by aligning mass spectral peaks according to their slightly different retention times. In addition, the software
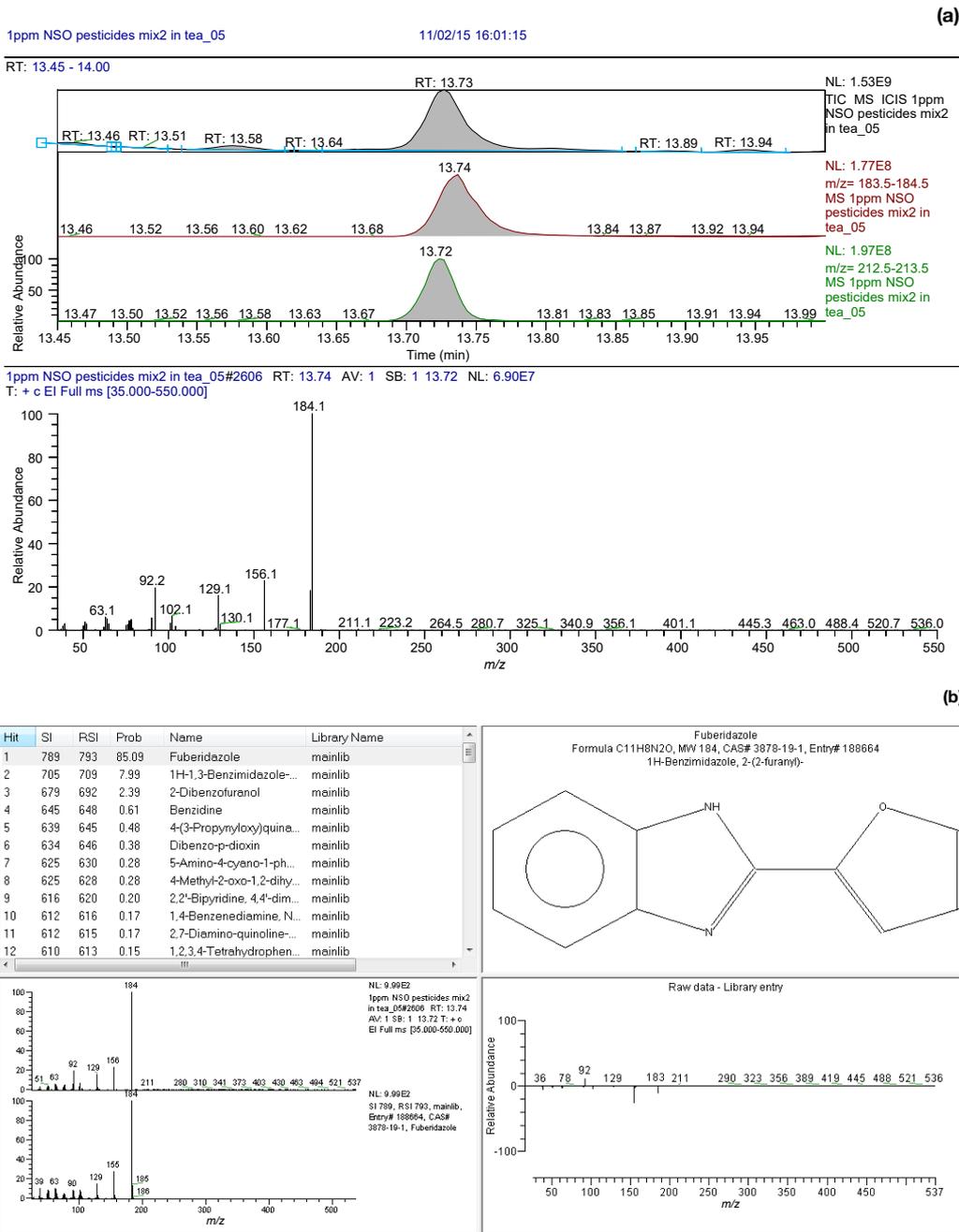
Figure 2. (a) Extracted ion chromatograms (XIC) of *m/z* 184 and 213 and spectra at retention time of 13.74 minutes subtracted by spectra at 13.72 minutes and (b) its library match results

automatically does a library search match and has the ability to export an unknown compound mass spectrum to the NIST format in order to create a user-defined library of compounds. When combined with the unknown screening functionality of TraceFinder software, it can be used to do cross-sample overlay of analytes that are well below the conventional concentrations where a unit-mass-resolution quadrupole instrument can typically detect and identify compounds.

This technical note will give an overview of the various functionalities of the deconvolution plug-in software and how they can be used in different analytical workflows.

Experimental

The analyzed samples were QuEChERS (Quick, Easy, Cheap, Effective, Rugged, and Safe) extracted tea matrix spiked with 1 ppm, 10 ppb, 5 ppb and 1 ppb levels of a custom pesticide mix containing nitrogen, oxygen and sulfur compounds. Another set of samples were prepared

with Restek™ GC Multiresidue Pesticide Standard #7 (GC MR7) (Catalog #32569) spiked in QuEChERS extracted lettuce matrix. The spiked levels ranged from 50 to 5000 ppb. A Thermo Scientific™ TSQ™ 9000 triple quadrupole mass spectrometer equipped with the Thermo Scientific™ ExtractaBrite™ ion source and coupled to a Thermo Scientific™ TRACE™ 1310 gas chromatograph was used to analyze the samples. The mass spectrometer was operated in full-scan mode scanning 35-550 Da with a dwell time of 0.2 s.

Software - Deconvolution plug-in software, version 1.3 was used in conjunction with TraceFinder software, version 4.1.

Results and Discussion

Theory- Determination of unknown compounds by library matching of mass spectra has traditionally followed a 'Chromatogram-first' approach. A 'Chromatogram-first' approach involves peak detection by one of the several available algorithms (like Genesis, ICIS, COBRA, etc.) and matching the mass spectra of detected peaks with the library to determine the closest matching compounds. The manual method to detect and identify unknown compounds also follows this approach but with greater flexibility for which scans (data points across the chromatographic peak) are selected and which background scans are subtracted. One of the flaws of using this approach is that co-eluting compounds can be difficult (or almost impossible at low-levels) to identify. The deconvolution plug-in software takes the 'mass spectrum-first' approach. The workflow of the two approaches is shown in Figure 3. The detailed steps that are taken in the mass-spectrum first approach of the deconvolution plug-in software are shown in Table 1. The most important parameter is the "ion overlay window" % that determines which of the XIC peaks are binned together as one component. An example of a binned component with 90% ion overlay criteria is shown in Figure 4. If the TIC peak of this component is within the threshold specified in the processing parameter options, then it will appear as a deconvoluted component in the peak list table. Next, the software performs a library search for all the binned components if this option is selected during processing. This approach identifies compounds that would be co-eluting chromatographically but whose mass spectral peak apexes are still separated.
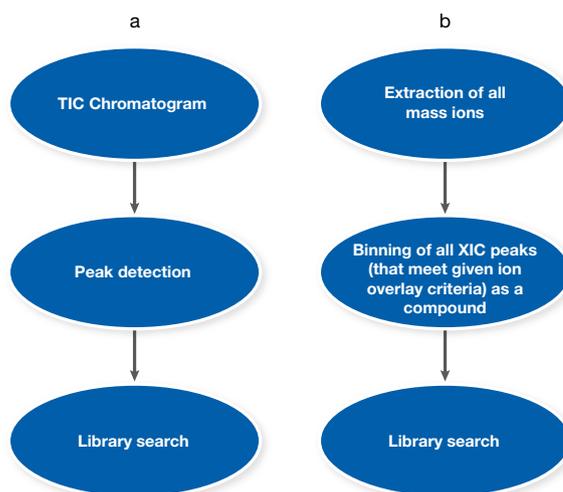
Figure 3. Unknowns analysis workflow - (a) Traditional 'Chromatogram-first' approach and (b) 'Mass spectrum-first' approach of the deconvolution plug-in software

Table 1. Automated steps taken by the deconvolution plug-in software

Step	Overview
1	Extracts all ions (nominal mass) and generates XICs
2	Peak detects each of the XICs
3	Picks up the most intense peak (for any RT and *m/z*)
4	Bins all the peaks that are within the "Ion overlap Window" % parameter (example- 99% of the height) and, if within the TIC threshold- it is considered as one deconvoluted component
5	Library searches this component (if option selected)
6	Repeat Steps 3 to 5 for the next highest peak and continues until no more peaks left
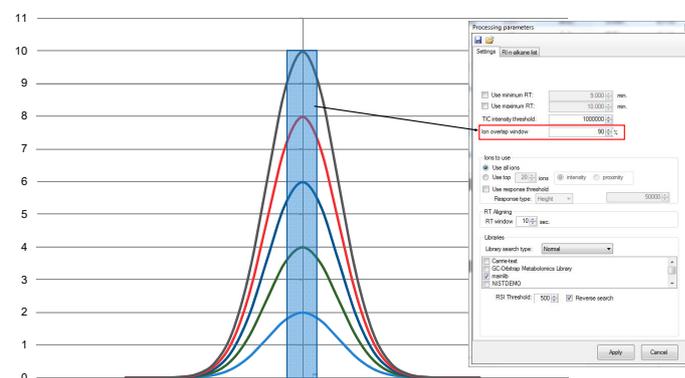
Figure 4. Example of Ion overlay window of 90%

Deconvolution at high levels of analyte- Figure 5 shows the comparison of the QuEChERS extracted tea matrix chromatograms for both control samples and those spiked with the custom pesticide mix at 1 ppm level. One can easily pick out pesticides from the matrix by visual comparison of the two chromatograms and identifying peaks that are absent in the unspiked sample. More challenging are cases where co-eluting pesticides show up as a single chromatographic peak. The 'chromatogram-first' approach mentioned above would give poor match scores for such peaks and thus fails to identify one or the other co-eluting compounds. The deconvolution plug-in software computationally separates these individual compounds from the single chromatographic peak by the procedure described in Table 1. The results are presented in a tabular format in the form of a 'Peak List'. Figure 6 shows the snapshot of the software highlighting the Peak List table. The figure also shows the chromatographic peak of Figure 1 deconvoluted into its components: Simetryn and Fuberidazole. Another example of deconvolution is shown in Figure 7 where the deconvolution plug-in automatically deconvolves and identifies the chromatographic peak at
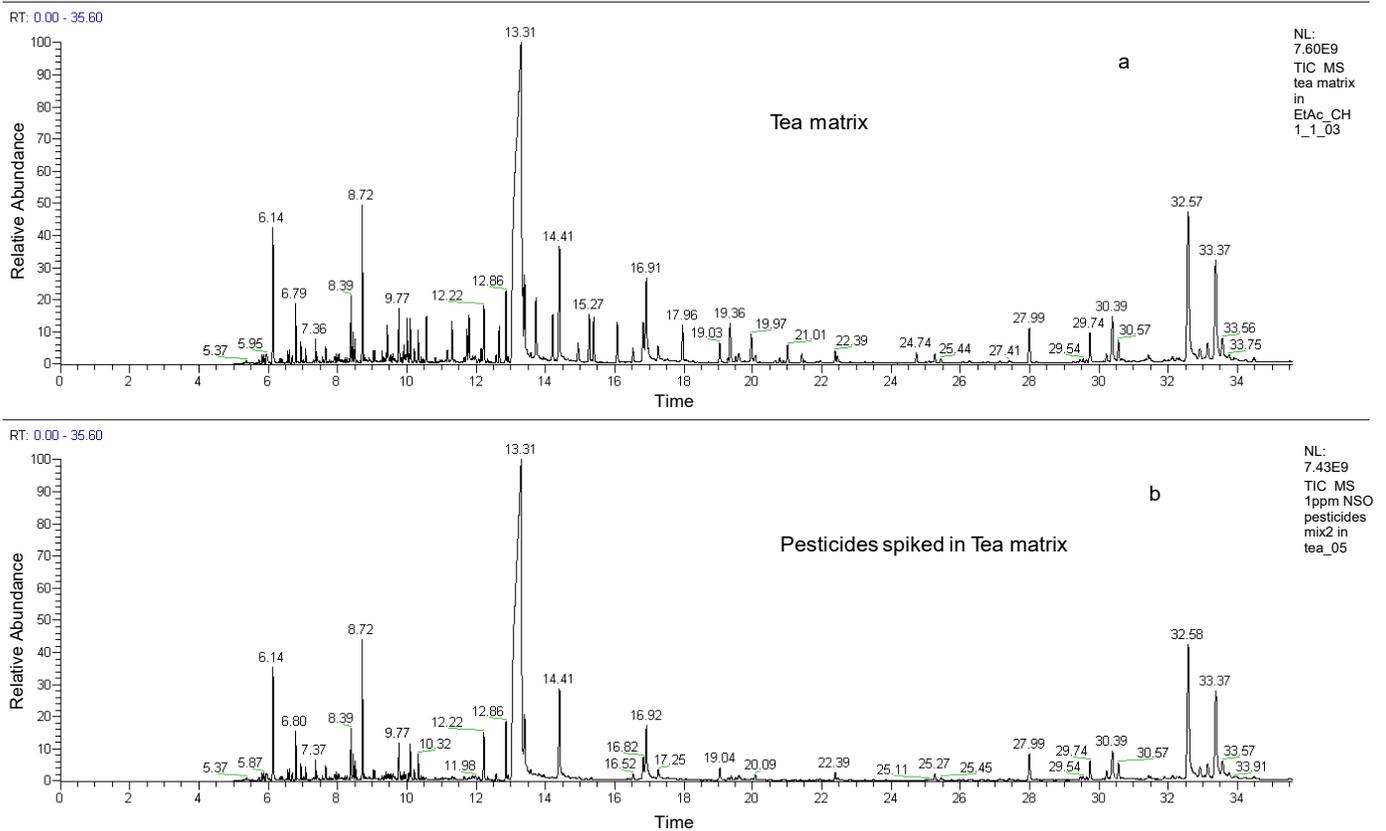
Figure 5. Chromatogram of (a) unspiked, (b) pesticides spiked tea matrix
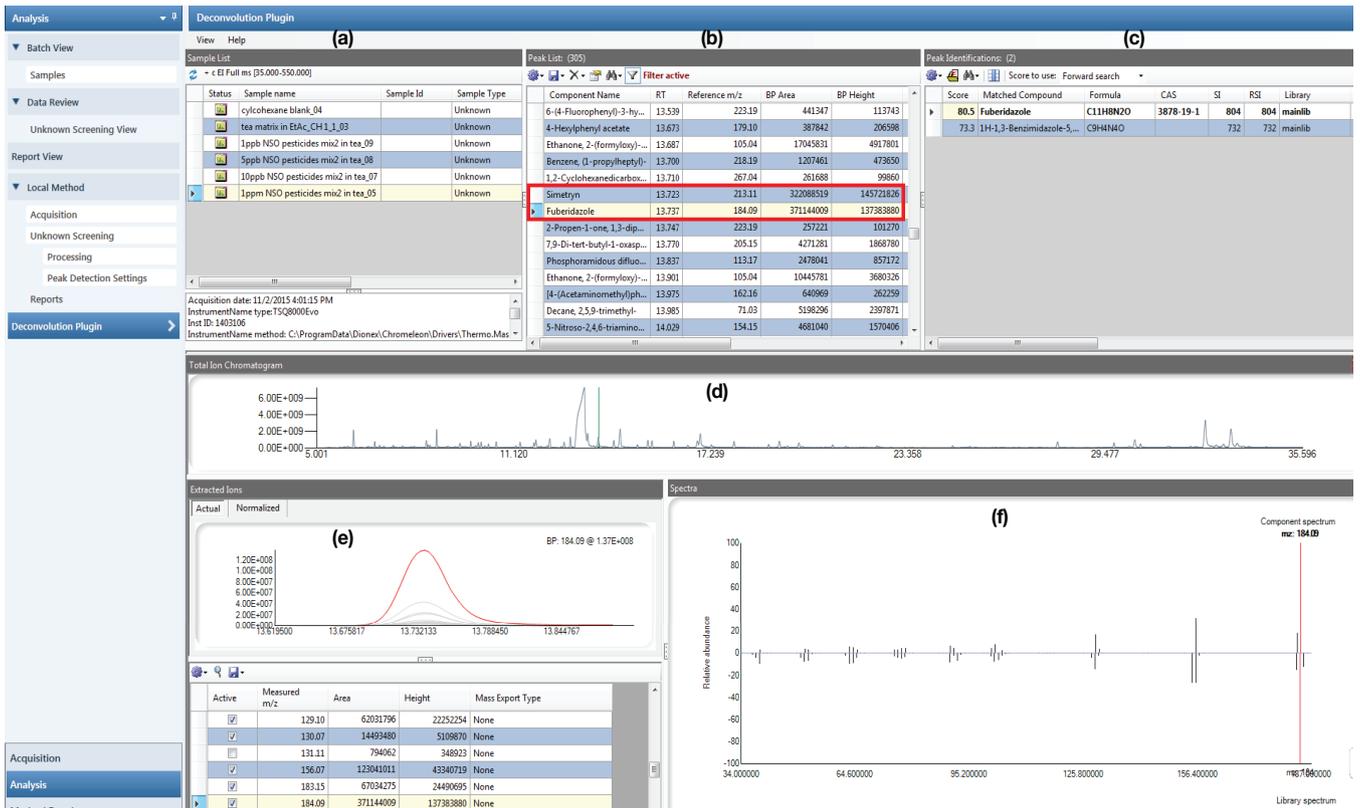
Figure 6. Snapshot of the deconvolution plugin software highlighting separation of simetryn and fuberidazole as individual compounds. Also shown are various section of the deconvolution plugin software- (a) Sample list and information, (b) deconvoluted peak list, (c) Library search results for the selected component in the peak list, (d) Total Ion Chromatogram (TIC), (e) Overlay of extracted ions of the deconvoluted component in the peak list and (f) the component mass spectra and comparison to the library.
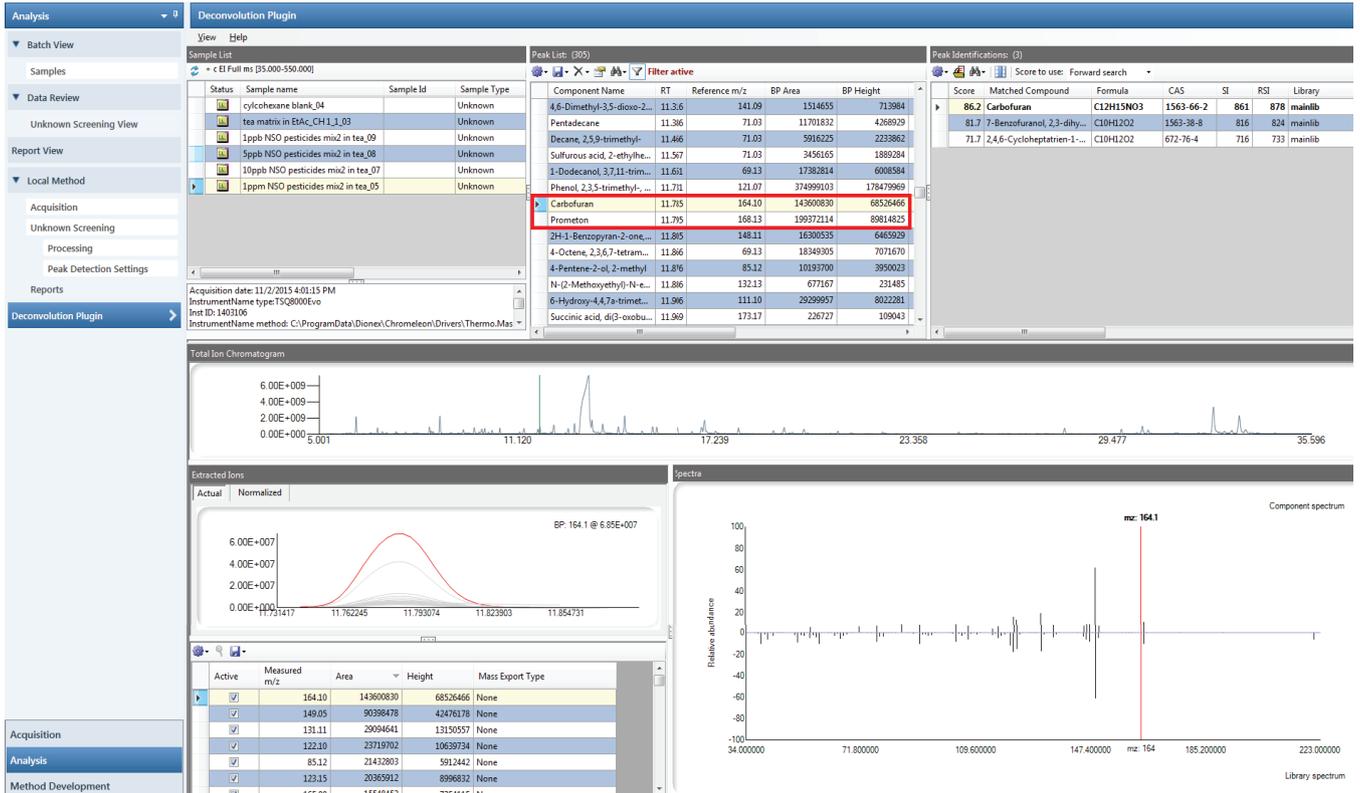
Figure 7. Snapshot of the deconvolution plug-in software highlighting separation of carbofuran and prometon as individual compounds

11.79 minutes as Carbofuran and Prometon. As can be seen from Figure 8- the peak at 11.79 minutes is a single peak but with two apexes for extracted ions at *m/z* 164 and 225. All the aligned peaks associated with carbofuran and prometon are also shown as separate compounds, as determined by the deconvolution plug-in software.

Deconvolution at low levels of analyte- One of the most useful features of the deconvolution plug-in software is the ability to identify compounds at low levels even in the presence of high-level background or interferences. Figure 9 shows that the deconvolution software successfully identified Desmetryn at 10 ppb levels in
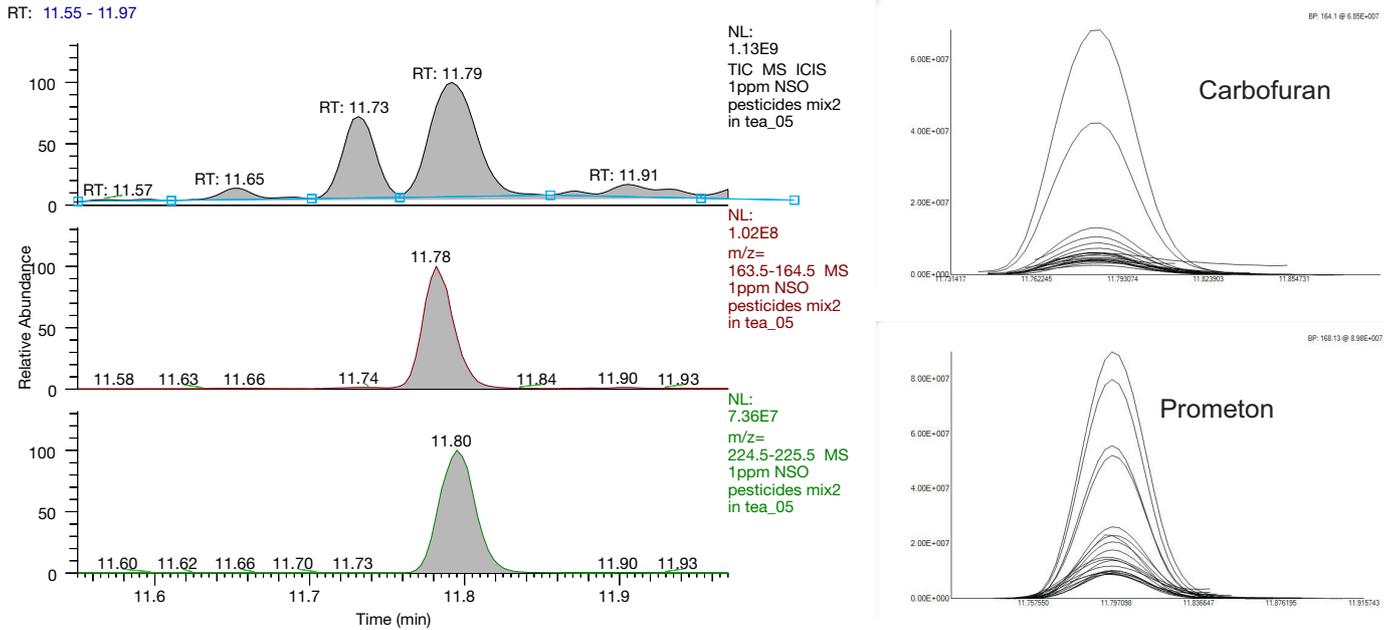
Figure 8. (a) Chromatographic peak associated with co-eluting carbofuran and prometon; (b) deconvolution plug-in software result showing all the binned XIC associated with individual compounds (and eventual library search results)
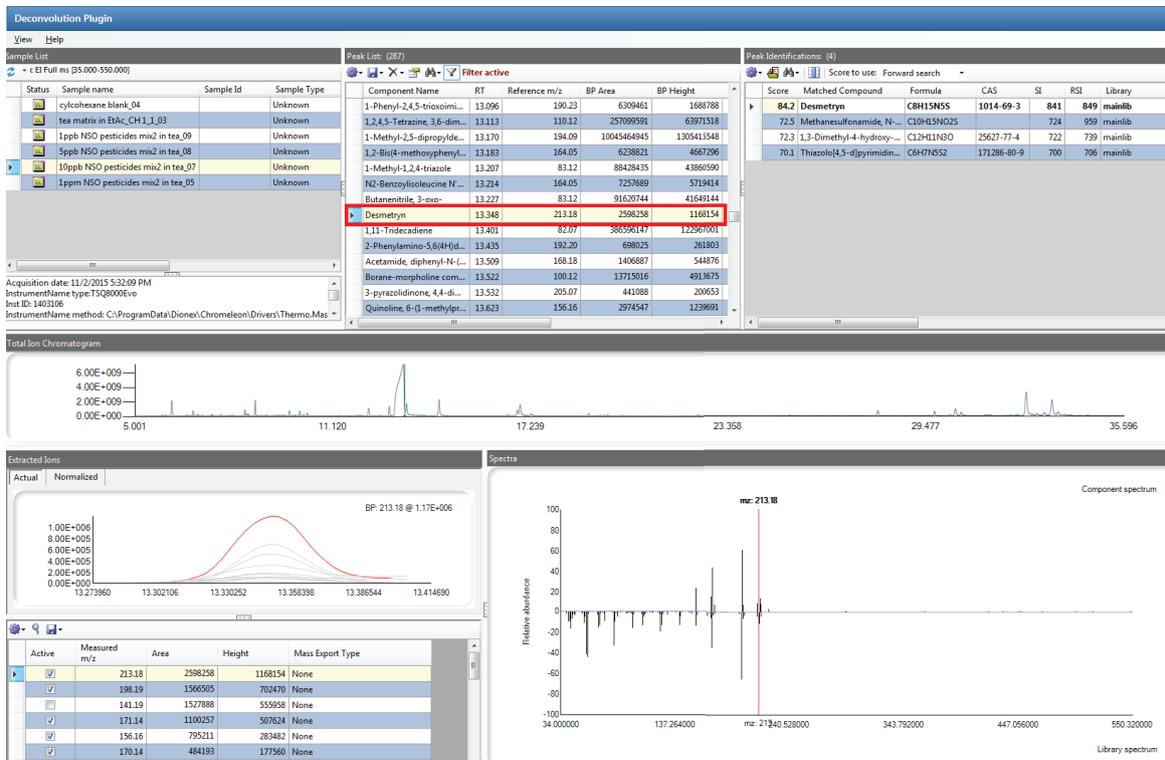
Figure 9. Deconvolution plug-in software correctly identifying Desmetryn at 10 ppb level

the spiked tea extract. The detection and identification of Desmetryn in such a complex matrix is particularly challenging since this pesticide elutes under the tail of an overloaded caffeine peak, as shown in Figure 10, which reports the TIC chromatogram of the sample. It would be almost impossible for a human or a chromatography data system software without a deconvolution algorithm to pick out Desmetryn from this background.

Table 2 shows the NIST library hit results from the GC MR7 pesticides mix spiked in lettuce matrix at various concentration levels. The Deconvolution software is able to pick up all the pesticides even at low-levels. All the pesticides, except Bromopropylate, are number one NIST library even at 50 ppb levels.
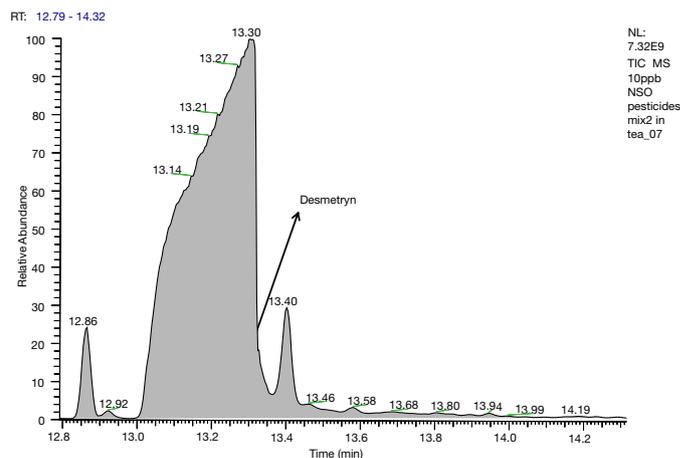
Figure 10. TIC showing co-elution of 10 ppb Desmetryn with the huge caffeine peak

Table 2. Deconvolution plug-in software NIST library hit# for pesticides at various concentrations in lettuce matrix

| Compound | Formula | RT (min) | NIST Library Hit # at Concentration (ppb) | | | | |
			50	200	500	2000	5000
o-Hydroxybiphenyl	$C_{12}H_{10}O$	9.94	1	1	1	1	1
Chlorpropham	$C_{10}H_{12}ClNO_2$	11.01	1	2	2	1	1
Metalazyl	$C_5H_{11}NO_2S$	13.79	1	1	1	1	1
DCPA	$C_9H_9Cl_2NO$	14.80	1	1	1	1	1
Chlozolinate	$C_{13}H_{11}Cl_2NO_5$	15.78	1	1	1	1	1
Fluazifop-butyl	$C_{19}H_{20}F_3NO_4$	18.40	1	1	1	1	1
Chlorobenzilate	$C_{16}H_{14}Cl_2O_3$	18.63	1	1	1	1	1
Carfentrazone ethyl	$C_{15}H_{14}Cl_2F_3N_3O_3$	19.70	1	1	1	1	1
Bromopropylate	$C_{17}H_{16}Br_2O_3$	21.82	3	2	1	1	1

Features of the deconvolution plug-in software-

a) Filter options — Depending on the chromatogram, the deconvolution plug-in software can generate hundreds or sometimes thousands of detected peaks. The filter function is used to narrow down the peak list to a manageable size and show compounds that match a certain criteria. There are three major criteria to filter the deconvolved peaks, as shown in Figure 11. The first filter, (1) is based on minimum ion peak area, height or counts; this threshold eliminates deconvolved peaks that are too small to be confidently identified. It also contains an option to narrow the retention time range in which to search for peaks of interest or on peaks that show certain mass delta because of neutral losses. The second filter option, (2) is based on mass filters; here one can narrow the peaks based on fragment ions of interest or on peaks that show certain mass delta because of neutral losses. For example,

some classes of compounds have common ion fragments like some phenyl methylcarbamate insecticides which have large 121 mass fragments. It is possible to use this option to filter out peaks containing m/z 121. The third option, (3) is based on library search criteria. Here the peak filtration is based on an SI/RSI score. Additionally it is possible to filter the peak list library hits based on name, CAS score, or formula.

b) Exclusion option — This feature, shown in Figure 12, allows exclusion of common peaks from, for example, a matrix blank thus showing only peaks of interest which are not part of the matrix or solvent blanks.

c) Ability to create user defined libraries — The deconvolution plug-in software allows creation of NIST format custom libraries from deconvoluted peaks
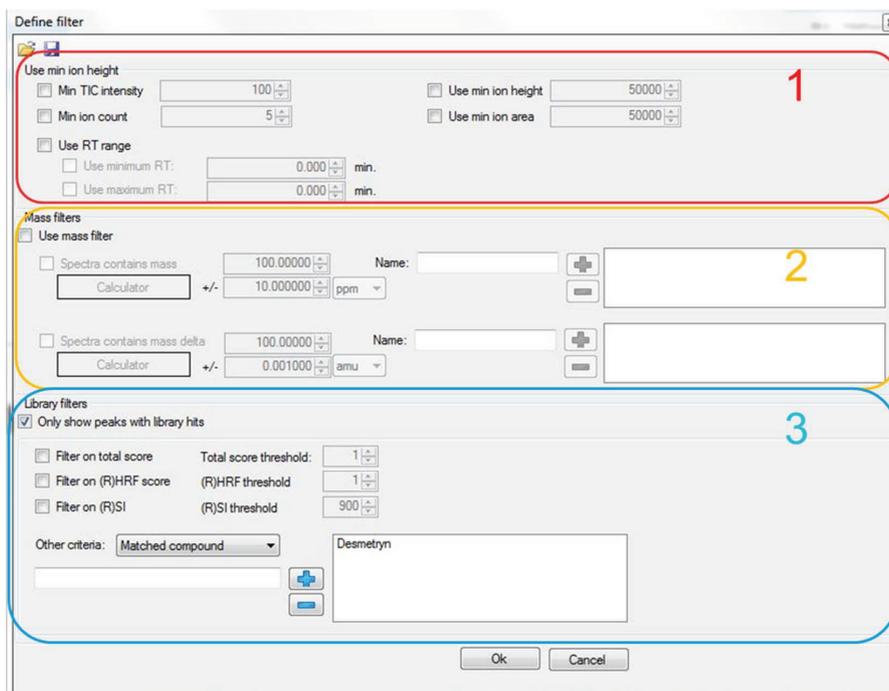
Figure 11. Filter options in the deconvolution plug-in software. Filtering can be based on 1) minimum area, heights or counts, 2) mass filter, i.e. on particular *m/z* or *m/z* delta or 3) library search criteria.

Figure 12. Exclusion feature allows us to choose the injection that will exclude common peaks associated with the matrix

found by the software, or allows to export ions into a compound database (CDB) format file to be used by TraceFinder software's quantitative method.

d) Unknowns screening with heat mapping — Results from the deconvolution plug-in software can be sent to the unknowns screening view of TraceFinder software, which allows the user to do heat mapping across various samples. Figure 13 shows the unknown screening view.

The compound Fuberidazole is highlighted, showing the detectability at levels as low as 1 ppb.

e) Reporting capability- A number of pre-made reports exist in TraceFinder software, for example the "Unknown screening high density report", that reports the results generated by the deconvolution plugin including the deconvoluted peak list along with the SI/RSI match factor scores. In addition, custom reports can be made according specific requirements of the user.
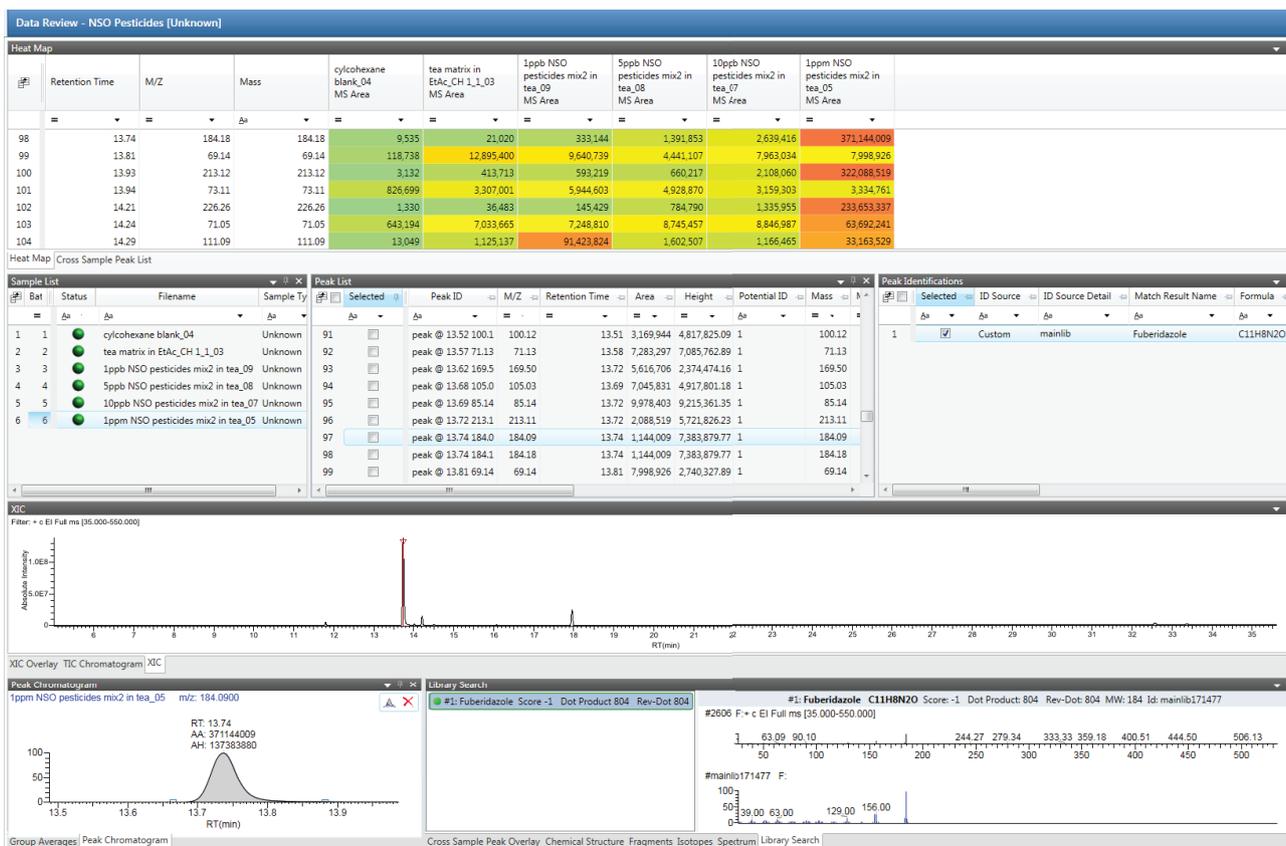
Figure 13. Unknowns screening view of TraceFinder software with data from the deconvolution plug-in software

Conclusion

This technical note illustrates the features and workings of the Thermo Scientific deconvolution plug-in software for GC-MS data. It shows the ability of the software to deconvolve a single chromatographic peak composed of two or more co-eluting compounds. The software also has the ability to identify compounds at low levels whose mass traces would otherwise be part of the noise without showing a well-defined chromatographic peak. In addition, it has many options that add to the usability of the software for unknown screening. Those include filtering functions, the ability to export the binned mass traces to NIST format to create user-defined libraries, and an exclusion option to filter peaks associated with the matrix. When combined with the unknown screening function of TraceFinder software, version 4.1, the deconvolution plug-in software allows detection of components that are well below the full-scan detection limits of the unit-resolution mass spectrometer.

References

1. Sparkman, O. D.; Penton, Z. E.; Kitson, F. G. Gas Chromatography and Mass Spectrometry – A Practical Guide; Academic Press, 2011.
2. Hübschmann, H.-J. Handbook of GC-MS; Wiley-VCH, 2015.

Find out more at **thermofisher.com/unstoppableGCMS**

Thermo Fisher
S C I E N T I F I C