A New Approach to the Analysis of Unresolved Chromatogram Peaks in GC/MS

Yongdong Wang, Stacey Simonoff
Cerno Bioscience, Norwalk, CT, USA

TP 816

The Objective: Improve and automate analysis of unresolved chromatogram peaks

Unresolved chromatographic peaks are commonly quoted as one of the more common problems in GC/MS analysis. For target compounds in known matrices unique fragment ions can be used to provide another dimension of separation and are commonly used in routine analysis. However, for unknown compounds and/or unknown matrices the problem is significantly more challenging. Unresolved compounds can seriously degrade library search results and provide ambiguous results. Several approaches are used to address this issue, mainly reverse library search and deconvolution applied in spectral space such as the NIST AMDIS program[1]. In this paper we will demonstrate an innovative approach that provides for deconvolution in chromatogram space and can provide significant advantages over traditional methods

Test Method

A test solution based on the EPA method for volatile organic compound (VOC) analysis (EPA 8260) was prepared for evaluating the method. The test solution was created by combining commercially available mixtures of compounds of environmental concern in methanol for a total of 77 compounds at approximately 20 ug/L for each component. The VOC test mixture was introduced with a 1 uL injection into an Agilent 5975 GC/MS. All data were collected in Raw Scan (profile or continuum) mode. The GC temperature programming was adjusted to provide runs with good and poor peak separations for testing. At the end of each run, the Agilent PTFBA was turned on briefly to provide MS calibration data for the analysis software[2].

Methodology

There are a number of chromatogram peak deconvolution methods, many of which are proprietary to the MS vendor software. NIST publishes the AMDIS software for this operation and it is likely commercial approaches work similarly to AMDIS or other published methods. NIST provides a good overview of other published work in the AMDIS documentation[3]. Briefly, selected ion chromatograms are generated over the chromatogram peak of interest and inspected. Groups of the “sharpest” peaks are considered likely candidates for the “pure” peaks. Peaks that are broadened are likely to contain ions from multiple “pure” components and are ignored. Once the peak locations are identified, a simple least squares method is used to extract the “pure” spectrum.

There are multiple challenges to these spectrum-based approaches. First, only a limited number of ions are used, which many times are not the most abundant in the spectra. This reduces the signal-to-noise in comparison to the TIC and limits the sensitivity of the technique. For example, in most cases, the “pure” spectra must be separated by about 50% of the FWHM, and it typically only works well for binary mixtures. Second, the interface is highly interactive and requires a well trained user and the proper setting of multiple parameters.

In our method, which we call TrueChrom MX, instead of operating on a limited number of ions, we instead use the TIC as the basis for determining the underlying peaks. This results in improved sensitivity due to the improved signal-to-noise. Furthermore, the method is essentially parameter free and can be automatically performed on the entire run with no user interaction. The approach is summarized here:

Step 1: Use a statistical based peak picker to locate all the chromatogram peaks in the run and produce a t-value for each peak indicative of the signal-to-noise.

Step 2: Perform Principle Component Analysis (PCA) on the set of mass spectra across each chromatogram peak. The number of factors will indicate the number of “pure” component spectra in each chromatogram peak (Figure 1)

Step 3: Pure peaks with high t-values are located throughout the run. These pure peaks are used as models for analyzing the remaining mixture peaks. If the run is so complex that an adequate number of pure peaks are not present, an external “standard” run with well resolved peaks may be used instead.

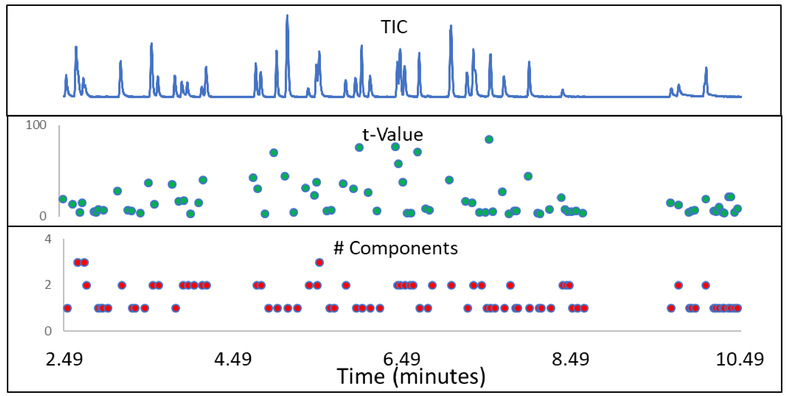


Figure 1. Peaks in the TIC are located, a t-value is calculated for each peak to indicate statistical significance, and a PCA is performed to obtain the number of components in each peak.

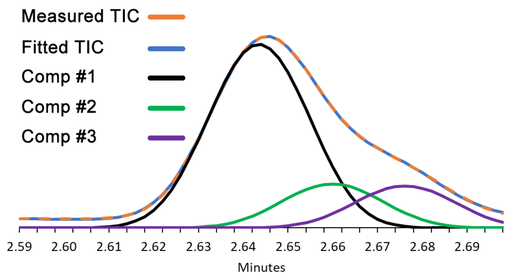


Figure 2. Plot of measured and fitted mixture peak and the underlying “pure” chromatogram peaks

Step 4: Using a modified, constrained Simplex and the “pure” peak model, each mixture peak is fit to determine the exact location and amplitude of each component in the mixture peak (Figure 2).

Step 5: Once the underlying peak positions and concentrations are known so we can use a Classic Least Squares Fit (CLS)[4] to solve for the matrix of pure component spectra.

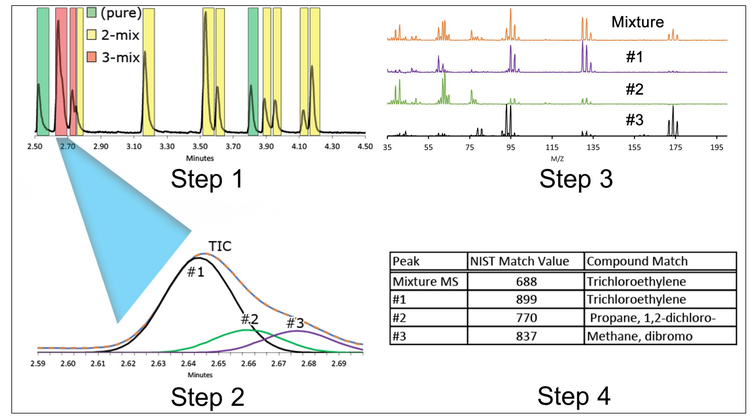
Once we have obtained the pure component spectra, we can then perform direct analysis on the data in the normal fashion including library search and spectral calibrations to obtain accurate mass data of high Spectral Accuracy. The later is used to determine the formula ID of the molecular ion or fragment ions to aid in the analysis[5].

Results and Discussion

The VOC test data contained a significant number of chromatogram peaks with unresolved mixtures of 2 and 3 components (Figure 1). The entire run was automatically post-processed as described and all “pure” peaks were automatically searched against the NIST 2017 library by directly calling the NIST search software APIs. In addition, the PTFBA gas was used to calibrate the mass spectra to accurate mass and high Spectral Accuracy. This allows further validation of the library search results by confirming the identity of the molecular ion (if present) from the NIST match.

All of the results were saved in raw format and a report of the results was generated in PDF format for review.

The results from the unresolved mixture peaks were validated against the fully resolved GC/MS run by comparing the relative retention times of each compound between the two runs. In general, the correct compound was the top match from the NIST search for the majority of the mixture peaks. One such example peak is shown in Figure 3. In this peak analysis, we also compared the search results using the averaged MS against those using the “pure” deconvoluted peaks and in all cases the search results improved significantly. Direct comparative measurements show that the upper limit of the number of components to be deconvoluted is 3-4 and that the separation between “pure” peaks can be as small as 10% of the FWHM. While computationally intensive, the complete processing of the run took about 1 minute, including library search.



Peak	NIST Match Value	Compound Match
Mixture MS	688	Trichloroethylene
#1	899	Trichloroethylene
#2	770	Propane, 1,2-dichloro-
#3	837	Methane, dibromo

Figure 3. Typical processing flow and results for a mixture peak.

Conclusion

A new approach to analysis of unresolved chromatogram peaks is demonstrated. Significant improvements in sensitivity and performance can be gained by operating in the TIC domain instead of the spectral domain . This is primarily due to the increased signal-to-noise of the signal(s) being processed (the TIC vs a smaller number of ion chromatograms). Typical performance limits were up to 4 components at 10% FWHM resolution as compared to 2 components at 50% FWHM resolution.

The downside of the approach is that it is significantly more computationally intensive than other methods. However, thanks to the high level of computing power available today, an entire run of moderate complexity can be analyzed in about 1 minute.

Finally, and perhaps more importantly, the method can be run fully automatically which can save the analyst significant time in the review and reprocessing of GC/MS runs.

References

1. MassWorks Rx, Cerno Bioscience, Norwalk, CT USA.
2. NIST MS search and NIST AMDIS for co-elution deconvolution.
3. Peter D’Arcy and W. Gary Mallard. AMDIS User Guide, National Institute for Standards.
4. D.M. Haaland and R.G. Easterling, Appl. Spec., 34, 539 (1980).
5. Y. Wang, M. Gu., Anal. Chem., **2010**, *82*, 7055.