Integrated Analysis of High Resolution GC-MS Data Using Hard and Soft Ionization to Identify Trace Impurities

Author: Robert Cody, Principal Scientist, JEOL USA, Inc.

OVERVIEW

Non-targeted analysis of complex mixtures by GC-HRMS should make use of all of the available data to identify unknowns. An automated data analysis software package combining chromatographic deconvolution with integrated analysis of high-resolution mass spectra for electron ionization (EI) and soft ionization measurements is applied to the identification of trace impurities in a fine chemical (triphenylphosphine).

INTRODUCTION

Gas chromatography combined with mass spectrometry (GC-MS) is well known as a powerful analytical tool for non-targeted analysis of complex mixtures such as environmental samples, quality control for feedstocks and fine chemicals, contaminants and impurities in commercial products, and metabolomic studies. GC-MS with a pyrolysis inlet is also widely used for materials characterization.

Interpretation of GC-MS data is heavily reliant upon searching electron ionization (EI) mass spectra against databases such as the NIST and Wiley mass spectral databases. Problems arise if the molecular ion is weak or absent, making it difficult to discriminate between compounds with different compositions, but similar mass spectra. A further complication is that many compounds do not have an entry in the databases.

A high-resolution time-of-flight mass spectrometer (HRTOFMS) is ideal for non-targeted analysis because it provides accurate mass and isotope data and is compatible with fast chromatography. Soft ionization methods[1] such as chemical ionization (CI), field ionization (FI) and photoionization (PI) provide molecular weight and elemental composition data to complement the database search results for electron ionization. Manual interpretation of the data is time-consuming. Here, we present a simple example of how an automated data interpretation program ("*msFineAnalysis*") combines all of the information from the high-resolution mass spectra for both electron ionization and chemical ionization mass spectra[2] to identify trace impurities in triphenylphosphine (*1*), a fine chemical used as a reagent for organic synthesis.
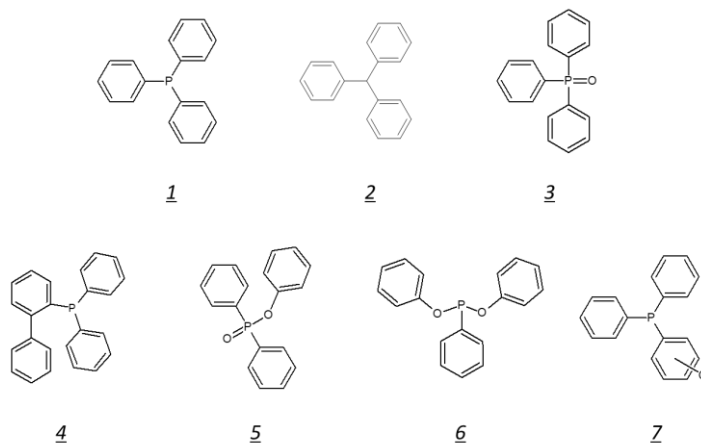
EXPERIMENTAL

A sample of triphenylphosphine from our chemical stockroom was dissolved in dichloromethane at a concentration of 250 ng μL^{-1} and analyzed by GC with a 20:1 split injection. Mass spectra were acquired using a JEOL AccuTOF-GCX Plus mass spectrometer. Separate injections were made for EI and CI measurements. The chemical ionization reagent was 5% ammonia in methane and perfluorotributylamine was used as an external mass calibration standard. Column bleed peaks were used as a drift correction for GC-MS accurate mass measurements. The measurement conditions are summarized in Table 1.

GC	MS
• Agilent 7890B	• Separate EI and CI measurements
• Column: 30 m DB-5ms	• EI, 70 eV, 300 µA
• Split injection (25:1)	• CI: 5% ammonia in methane
• Oven temperature program:	• Resolving power = 10,000
• Initial temperature = 100°C (hold 1 min)	• m/z 35-800 at a spectrum s^{-1}
• Heat to 300°C at 15°C min^{-1} (hold 20 min)	• Solvent delay: 1.8 minutes

Table 1. Measurement Conditions

The results were analyzed by using JEOL msFineAnalysis version 2 software. The chromatographic deconvolution function detected low-level peaks that were not immediately evident from the total ion chromatograms (TIC) for the EI and CI data. Each detected peak in the EI data was searched against the NIST

2020 mass spectral database. Molecular weight and elemental composition information in the CI data was correlated with the EI database matches, using both accurate masses and isotopic abundances. A fragment coverage score was calculated for the EI mass spectra by determining elemental compositions for all fragments assuming the elemental composition for the candidate molecule. Compounds that have high database search scores with confirming elemental compositions and high fragment coverage are color-coded in pale blue to indicate reliable assignments. Data analysis is fully automated, with the results presented in a summary table Using a color-blind friendly palette, compounds that have multiple matches are color-coded in pale orange, and compounds that do not have any reliable assignments are color-coded in white. Clicking on any entry brings up details of the analysis for that peak: database search results, molecular weight and elemental composition information from the soft ionization measurement, and a table summarizing all of the data for that peak assignment. A list of fragment ion compositions and a list of neutral losses is created from all detected peaks. This list can be used to perform group analysis to identify "families" of related compounds that have common fragments or neutral losses.

RESULTS

Compound identification using multiple pieces of information

To look for impurities, a high concentration of triphenylphosphine was injected with a 25:1 split ratio. Because the peak for triphenylphosphine eluting at 11.17 minutes is off scale, the mass spectrum is clipped, distorting the spectral pattern. The top database match is (2-Carboxyethyl)-triphenyl-phosphonium bromide which has the composition $C_{21}H_{12}BrO_2P$ and a molecular weight of 414. However, the CI data identify the correct formula as $C_{18}H_{15}P$ consistent with the second-best database match, triphenylphosphine. Although Br was not included in the list of elements used to calculate elemental compositions, the program used "feedback" from the database search matches to ensure that all elements in the database matches are considered for the elemental composition calculation. The two impurities with the highest relative abundances are readily identified as triphenylmethane (*2*) at 10.12 minutes and triphenylphosphine oxide (*3*) at 13.14 minutes. (Figure 1).
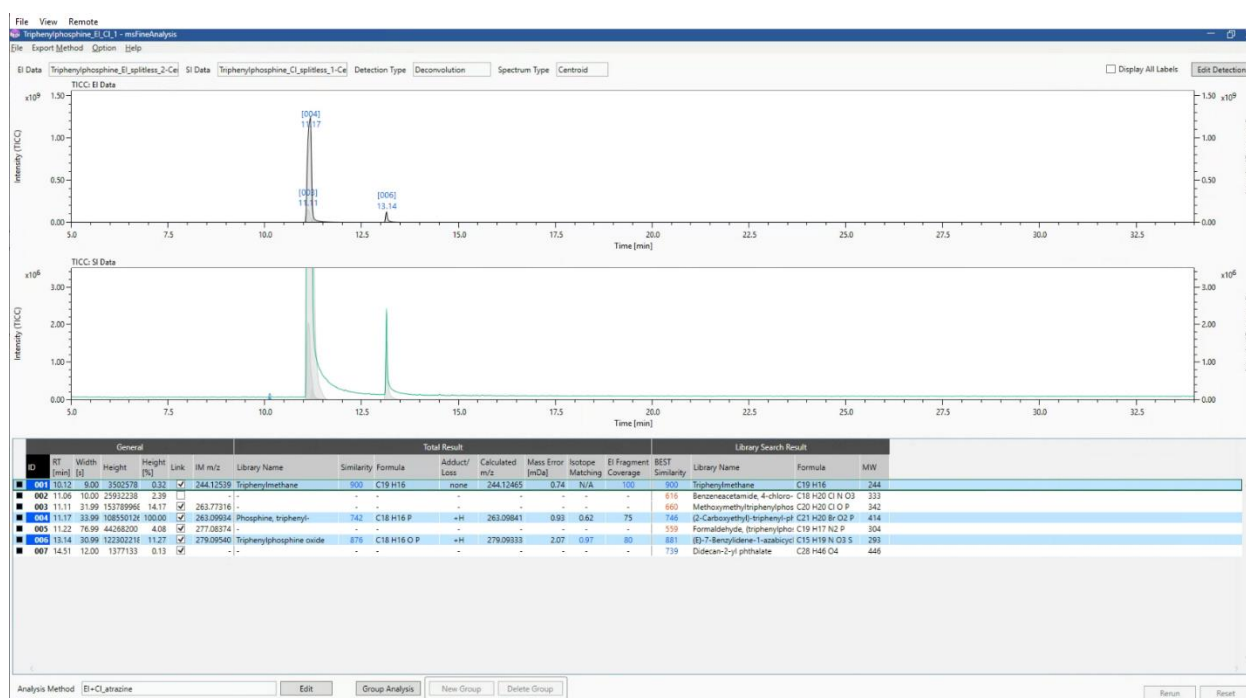

Figure 1. Assignments for the three most abundant peaks in the chromatogram

The triphenylphosphine oxide assignment will be used as an example to show of how the information is used by the program. The highest-scoring database match the peak eluting at 13.14 minute is **not** triphenylphosphine oxide – the oxide is the second-best match (Figure 2).
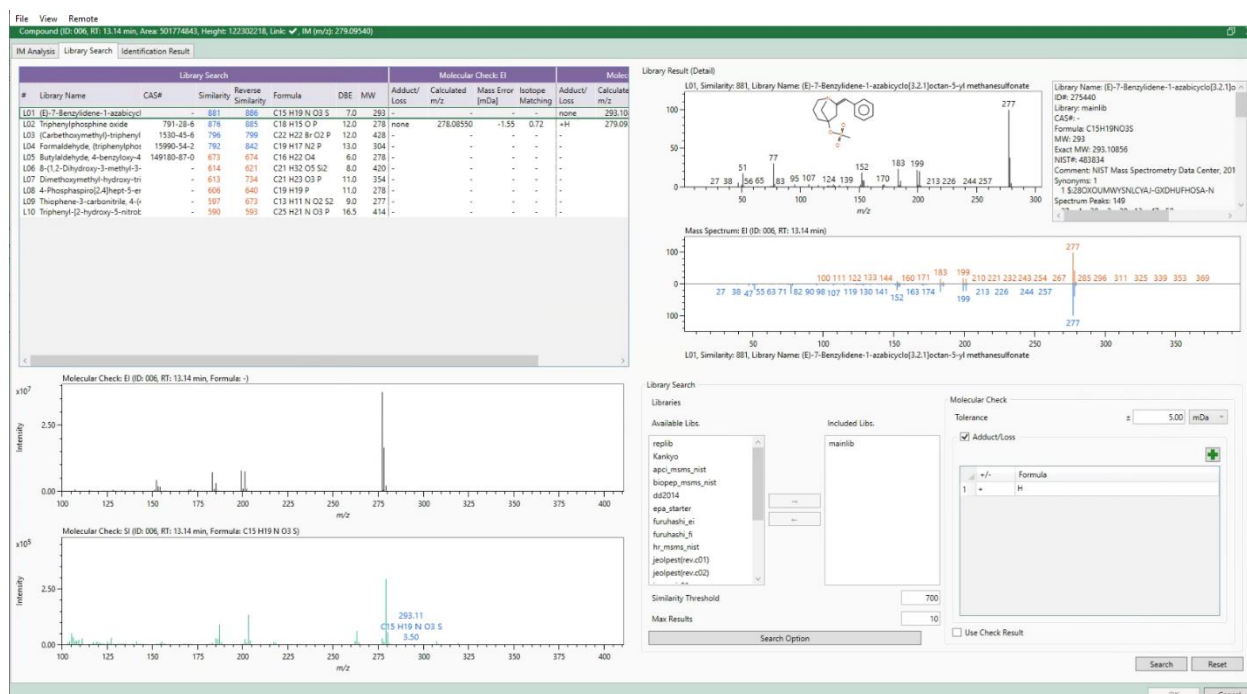

Figure 2. Database search results for the peak eluting at 13.14 minutes

However, from the CI data, we know the protonated molecule is detected at *m/z* 279.095. One of the calculated elemental compositions is $C_{18}H_{16}OP$, which is consistent with the second-highest scoring database match, triphenylphosphine oxide (Figure 3).
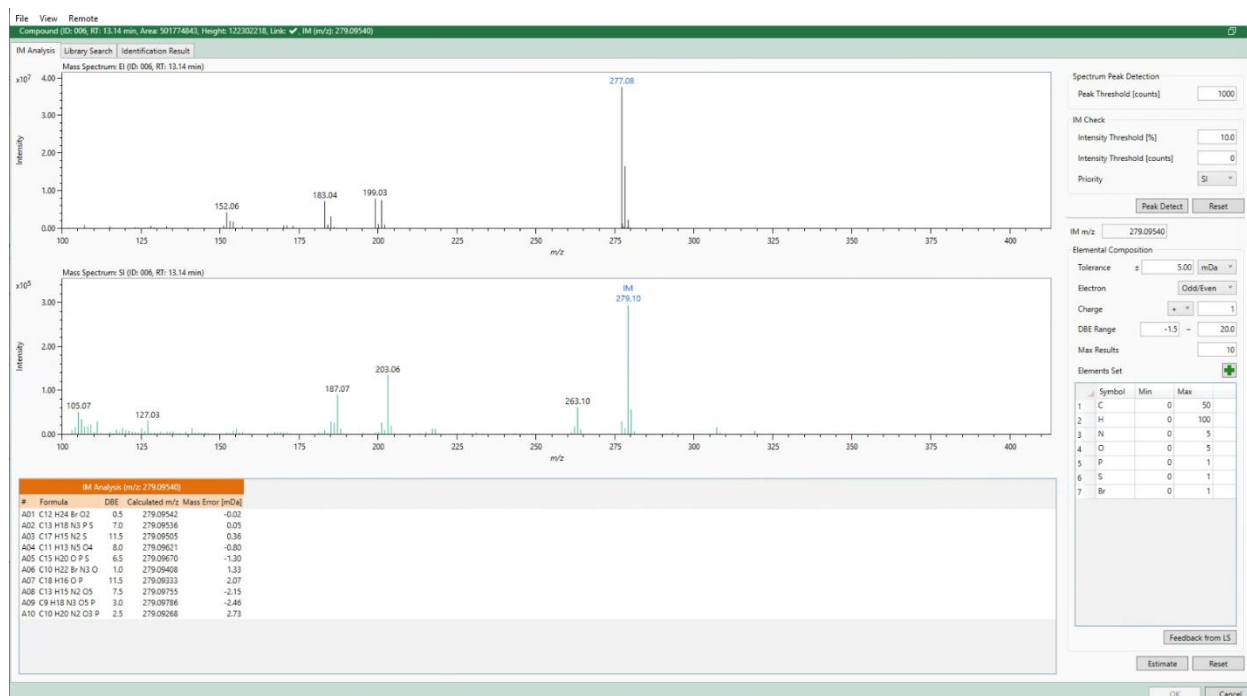

Figure 3. Elemental compositions from the CI mass spectrum for the peak at 13.14 minutes

Triphenylphosphine oxide is the only assignment that has correlated data for all of the information at that retention time: high database match scores, accurate mass and isotope data in the CI mass spectrum with a consistent elemental composition, and high fragment ion coverage (Figure 4). Therefore, msFineAnalysis reports triphenyl phosphine oxide as the correct assignment. The presence of triphenylphosphine oxide is not surprising, given that triphenylphosphine undergoes slow oxidation in air.
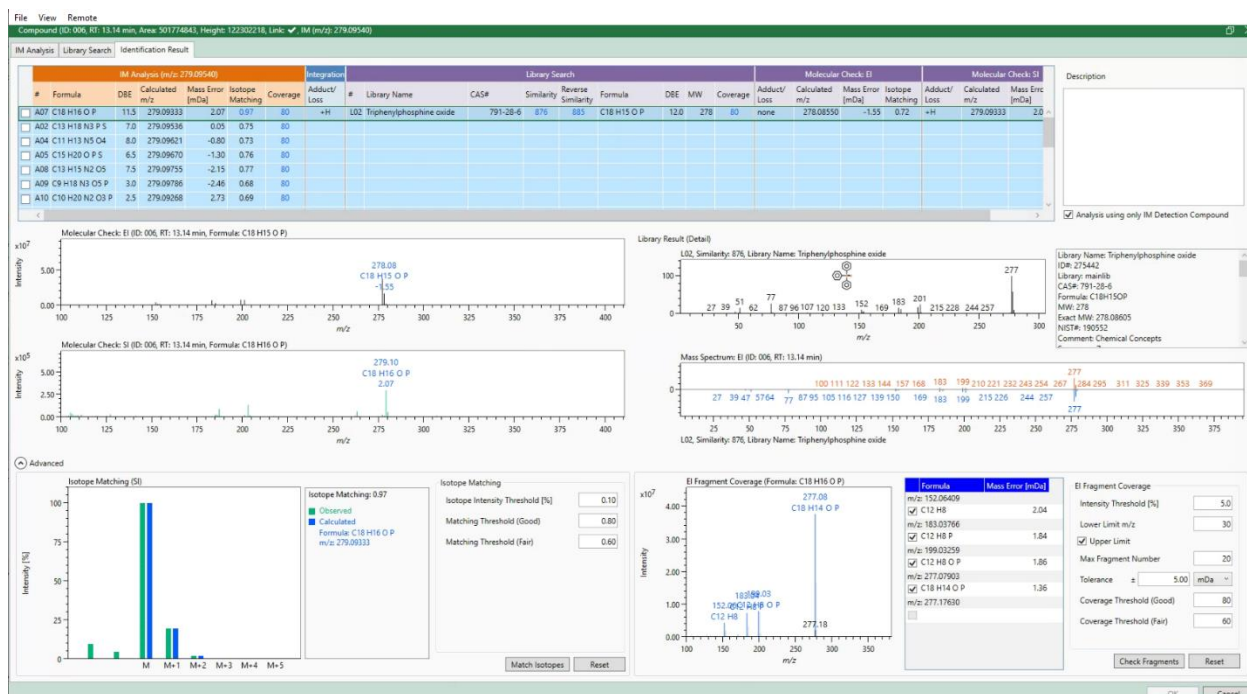

Figure 4. Summary page showing details of triphenylphosphine oxide assignment.

Another compound eluting at 12.64 minutes is identified as $C_{18}H_{15}O_2P$. The highest-scoring database match is diphenylphosphinic acid, phenyl ester (5) although the second-best match phosphonous acid, phenyl-, diphenyl ester (6) has a nearly identical mass spectrum and cannot be ruled out without a standard sample.

Using Group Analysis to find related compounds

We can use the group analysis functions to look for families of related compounds. From the list of fragment ions (Figure 5), we can see that 4 compounds have a diphenyl fragment $C_{12}H_8^+$.
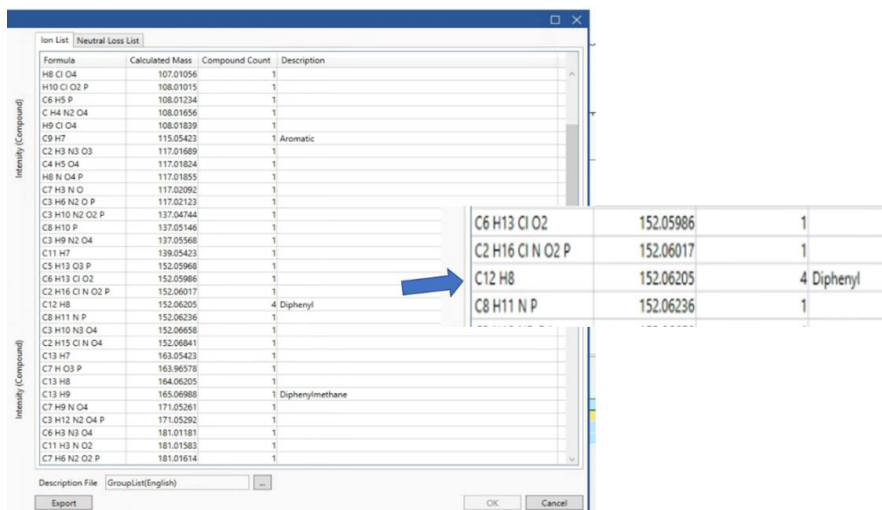

Figure 5. Group Analysis tab showing fragment ions detected in the dataset

Triphenylmethane and triphenylphosphine oxide are identified with high confidence in this group along with 2 (Diphenylphosphino)-biphenyl (_4_), which elutes at 13.99 minutes. A small peak eluting at 12.07 minutes is highlighted in yellow (Figure 6) because it does not have a confirmed composition with matching database entries and elemental composition.

	General							Total Result							Library Search Result		
ID	RT [min]	Width [s]	Height	Height [%]	Link	IM m/z	Library Name	Similarity	Formula	Adduct/ Loss	Calculated m/z	Mass Error [mDa]	Isotope Matching	EI Fragment Coverage	BEST Similarity	Library Name	Formula	MW
003	10.12	9.00	3902774	0.29	✓	244.12539	Isomer	-	C19 H16	none	244.12465	0.74	N/A	100	916	Triphenylmethane	C19 H16	244
013	12.07	11.00	538759	0.04	✓	296.05237	-	Multi Hits	-	-	-	-	-	-	513	erythro-8-(1,2-Dibromo-2-ph	C16 H16 Br2 N4 C	454
019	13.14	30.99	14315900	10.55	✓	279.09540	Triphenylphosphine oxide	910	C18 H16 O P	+H	279.09333	2.07	0.97	86	920	(E)-7-Benzylidene-1-azabicyc	C15 H19 N O3 S	293
024	13.99	18.00	1176418	0.09	✓	339.13194	2-(Diphenylphosphino)-biphe	900	C24 H20 P	+H	339.12971	2.22	N/A	100	900	2-(Diphenylphosphino)-biphe	C24 H19 P	338

Figure 6. Compounds detected with a $C_{12}H_8^+$ fragment

Double-clicking on this entry brings up detailed information from this analysis. The molecular ion occurs at nominal _m/z_ 296 which is confirmed by the _m/z_ 297 peak for the protonated molecule in the CI mass spectrum. From the accurate mass and isotope data and fragment coverage (Figure 7), it is clear that the elemental composition is $C_{18}H_{14}ClP$ or (2,3, or 4-chlorophenyl) diphenylphosphine (_7_). Although database searching did not return a match because there is no entry in the NIST or Wiley databases for any compounds with this elemental composition, the accurate mass and isotope data allowed us to assign a composition, and the fragment ion elemental compositions are consistent with the proposed structure.
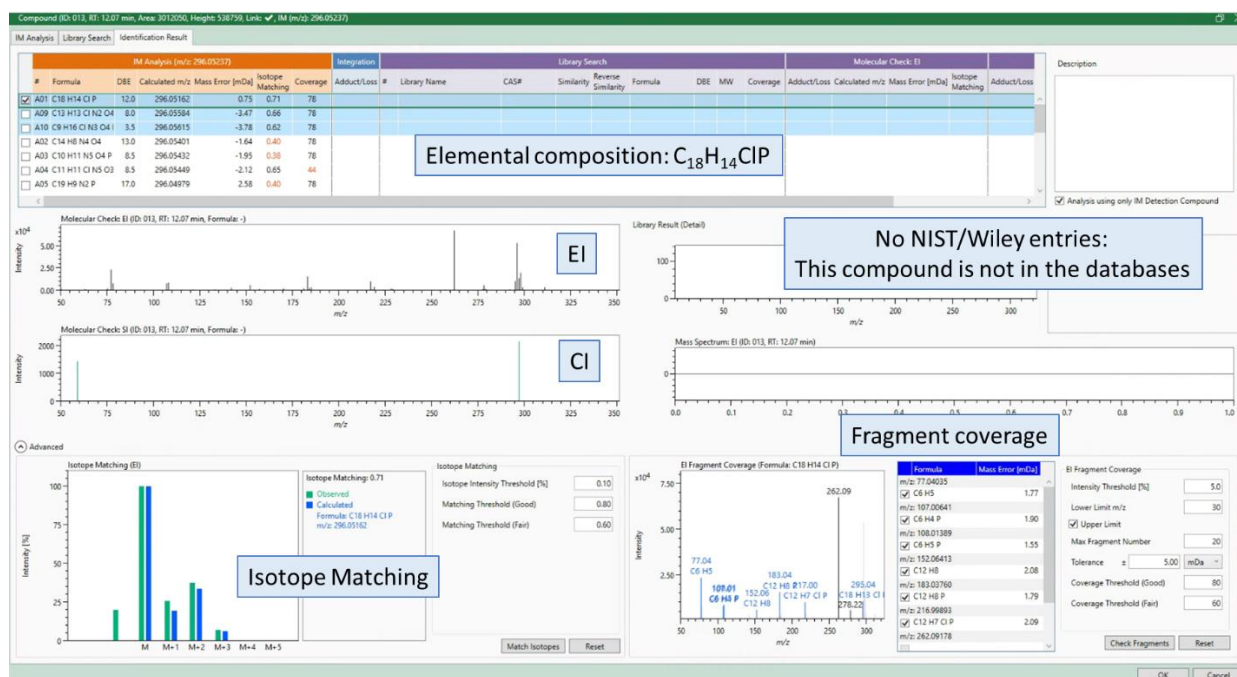

Figure 7. Summary page for identification of $C_{18}H_{14}ClP$

Conclusions

The AccuTOF-GCX Plus GC-HRTOFMS system is shown to be a useful method for identifying trace impurities. The msFineAnalysis version 2 software combines chromatographic deconvolution with integrated data analysis of both soft (EI) and hard (CI) ionization measurements, including accurate mass and isotope data. The resulting report assigns high confidence to compounds that have high-scoring database search matches and accurate mass/isotope data that support the search result. High-resolution mass spectral data is summarized to aid in interpreting mass spectra for compounds that do not have database entries. Although a simple example was shown here, the method is valuable for more complex mixtures, as shown in the examples in references 2 and 3.[2, 3]

References

1. T. Furuhashi, K. Okuda. Application of GC/MS Soft Ionization for Isomeric Biological Compound Analysis. *Critical Reviews in Analytical Chemistry*, **2017**, 1.
2. M. Ubukata, A. Kubo, K. Nagatomo, T. Hizume, H. Ishioka, A. J. Dane, R. B. Cody, Y. Ueda. Integrated qualitative analysis of polymer sample by pyrolysis–gas chromatography combined with high-resolution mass spectrometry: Using accurate mass measurement results from both electron ionization and soft ionization. *Rapid Communications in Mass Spectrometry*, **2020**, *34*, e8820.
3. R. B. Cody. Analysis and Identification of Mezcal and Tequila Aromas by Ambient Ionisation MS, GC-MS, and GCxGC-MS. *Chromatography Today*, **Aug-Sep 2020**.