Towards More Robust, Accurate, and Automated GC/MS Compound ID-
Use of Quadrupole GC/MS for Accurate Mass Identification and Fragment Elucidation of Unknown Compounds

Yongdong Wang, Don Kuehl, Cerno Bioscience, Norwalk, CT, USA

The Objective: Improve and automate GC/MS search and compound ID

GC/MS library search is a mainstay application for the identification, confirmation and quantitation of known and unknown compounds. In spite of its huge popularity, library search results are sometimes ambiguous and in some instances can fail to produce any useful results. This is due to a number of reasons including unresolved chromatographic peaks (mixtures), column bleed and background interferences, or simply the absence of the target compound in the library. While there are manual methods to work around some of these limitations[1], they tend to require a high degree of user expertise and can be tedious and time consuming.

Recent advances in MS calibration technology[2,3,4] (TrueCal™) have demonstrated that commercial single quad GC/MS systems are capable of accurate mass measurements to ±5 mDa, and more importantly the isotope lineshape profiles (Spectral Accuracy) to greater than 99.0%. This novel calibration method enables these instruments to perform reliable formula ID comparable to expensive high resolution MS instrumentation (CLIPS - Calibrated Line-shape Profile Search[2]). In this poster, we will explore utilizing these technologies to improve library search and unknown compound ID, automate mixture and background correction, and most importantly provide orthogonal metrics to provide more accurate and robust compound ID.

The Solution: A multi-pronged and systematic approach

To improve the accuracy and robustness of GC/MS compound ID, a multi-pronged approach is proposed to mitigate the limitations mentioned previously including:
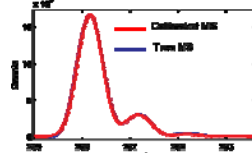
- Utilize CLIPS formula ID as an orthogonal metric to cross validate the library search
- Automatically flag unresolved chromatographic peaks and interferences using Spectral Accuracy
- Deconvolve unresolved peaks using multivariate methods enabled from TrueCal calibrated profile data
- Perform extended EI fragment elucidation on calibrated spectra to ensure spectral integrity and assist in the ID of unknowns not in the MS library
- Enhance current nominal mass library search using accurate mass filters
- Explore the use of fully calibrated accurate mass libraries

CLIPS formula ID and Spectral Accuracy purity confirmation

Calibrate profile mode MS to a defined line-shape gaining 100X in mass accuracy and >99% Spectral Accuracy
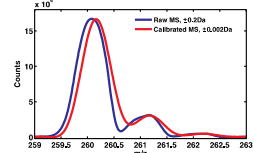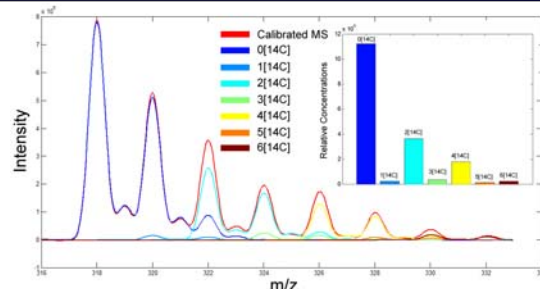

Search all formula within accurate mass window and return best matches by Spectral Accuracy

An automatic, comprehensive calibration using the built-in PFTBA calibration gas of the GC/MS provides for accurate mass formula search to ID formula candidates. The calibrated MS can be accurately matched to the "True" calculated MS since the line-shape is defined from the calibration to provide accurate formula ID. The formula ID results are cross compared with the library search results to further confirm the match. In addition, if the MS is not pure (a mixture) the Spectral Accuracy value will be severely degraded, flagging the problem[5].

Formula	Mass Error (mDa)	Spectral
C16H22NO2	-1.9	99.71
C14H20N4O	-0.5	99.40
C12H18N7	0.8	99.15
C17H24S	3.3	98.99

Deconvolve unresolved chromatographic peaks using multivariate analysis


The plot above shows the deconvolution of a complex (7 component) isotope mixture which is chromatographically nearly impossible to resolve. The series of deconvolved spectra are shown under the calibrated mass spectra (red) and can be quantitatively measured below the 1% level[6] (bar graph). The Spectral Accuracy of the complete fit is a direct indication of the accuracy of the analysis.

If a chromatographic peak is flagged as "impure", the principal component analysis can be used to estimate the number of components in the peak down to the 1% level. From there an iterative multivariate analysis algorithm is used to mathematically separate out the pure spectra from the mixture, which can then be resubmitted to CLIPS formula ID and the library search. The method is dependent on the comprehensive mass and Spectral Accuracy calibration and the excellent linearity of most single quadrupole GC/MS instruments.

EI fragment elucidation is used to further qualify the spectral data and provide assistance to true unknown identification

GC/MS is a powerful tool for compound ID by virtue of the characteristics of the EI fragments. Like MS/MS, the fragments can provide significant insight into the chemical structure of a true unknown (a compound not present in the library). With the ability to use CLIPS to provide accurate formula ID of each fragment, it is possible to calculate the Spectral Accuracy for a logical series of commonly recognized ion fragments derived from the molecular ion. The Spectral Accuracy across all fragments serves as a quantitative evaluation of correctness of the proposed fragments, thus providing insight into the possible molecular structures.

The graph at right shows a fragment analysis based off the molecular ion. The Spectral Accuracy including each and every fragment is calculated across the whole mass range based on the proposed fragment loses (Repeat Units) and provides a quantity and overall quality metric for the proposed fragments or fragment series.
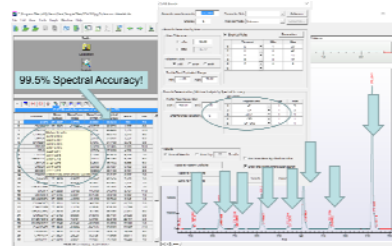
Accurate mass library search improves the quality of the match
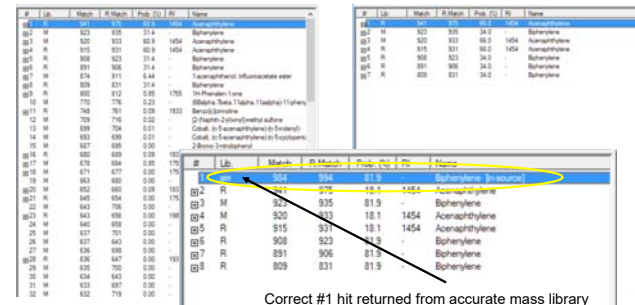
Correct #1 hit returned from accurate mass library

The search results on the left show a conventional NIST library search of nominal mass spectra which returns a long list of candidates with the #2 hit being the correct match. The search results on the right apply the accurate mass filter of +/- 10mDa, limiting the hit candidates for consideration and improving confidence but still lists the correct compound at #2. A fully calibrated accurate mass library was created[7] which provided superior match quality and identified the correct compound as #1 (lower insert) with higher forward or reverse match scores.

Summary

These preliminary results show that the mass accuracy and Spectral Accuracy of the TrueCal comprehensive calibration can be used to potentially provide dramatic improvements to the current conventional GC/MS library search. This systematic approach could largely eliminate the common pitfalls of GC/MS library search and provide a fully automated approach to accurate and robust compound ID by GC/MS. Furthermore, for true unknowns, CLIPS formula ID and fragment analysis can provide useful information leading to not only elemental composition but also possible molecular structures.

Acknowledgement

References

1. NIST MS search and NIST AMDIS for co-elution deconvolution.
2. Y. Wang, M. Gu., Anal. Chem., **2010**, 82, 7055.
3. R.J. Abel, Canadian Journal of Forensic Science, **2014**, 47, 74.
4. S. Moldoveanu & K. Kilby, ASMS Annual Meeting, 2014.
5. J.P. Chen, D. Sparkman, M.Gu, ASMS Annual Meeting, 2009.
6. M. Gu, Y. Wang, D. Kuehl, 29th NA Chapter Meeting of International Isotope Society, 2016.
7. D. Sparkman, M. Gu, Y. Wang, ASMS Annual Meeting, 2016.