Simplifying Non-Target Analysis: Improving Peak Match Confidence with Gas Chromatography, High Resolution-TOFMS and Improved Data Processing Strategies

Todd Richards, Joseph E. Binkley, Lorne M. Fell | LECO Corporation, Saint Joseph, MI USA

LECO
EMPOWERING RESULTS

Introduction

Historically, targeted analysis has been the only practical route to evaluate complex environmental samples. Recent improvements in detection and data processing capabilities allow scientists to more fully evaluate these same samples using non-targeted techniques. Often these evaluations are driven by a lab's available equipment and are heavily influenced by current workflows and equipment rather than a thoughtful evaluation of available technologies. As a result, GC-MS has been largely overlooked despite being a well-developed technology that is easily paired with extensively curated databases containing hundreds of thousands of mass spectra and retention index data. These advantages can obviate the need for hyper-complex data processing routines or custom scripting, while simultaneously returning high confidence peak identifications.

- GCxGC dramatically improves chromatographic resolution and peak detection
- Industry-leading deconvolution and non-target detection
- Multiplexing mass analyzer increases sensitivity 10X
- High Resolution Accurate Mass (HRAM) data allows for molecular and fragment ion formula calculations and verification
- ChromaTOF® brand software – A single software platform for hardware control and data processing

LECO Pegasus® GC-HRT+ 4D

Method Conditions

As part of the EPA's ENTACT multi-lab trial we were provided with 10 blind samples containing between 100 to 400 compounds. These samples were analyzed in EI and CI modes for both GC and GCxGC separations using the settings described below. The CI data were utilized to confirm the identification for EI spectra with either a low abundance or non-existent molecular ion.

HRAM GC-TOF MS	LECO *Pegasus* GC-HRT+ 4D
Ion Source Temperature	250 °C (EI) 200 °C (CI)
Acquisition Mode	High Resolution, ≥ 25K @ m/z 219 (FWHH)
Ionization Mode	EI and CI (Reagent Gas CH_4 + 5% NH_3)
Mass Range (m/z)	29 to 1000 (EI); 60 to1000 (CI)
Acquisition Rate	200 spectra/s (GCxGC); 6 spectra/s (GC)
Gas Chromatograph	**LECO GCxGC**
Injection	1 µL (diluted 10:1 in DCM) Split 10:1, Inlet Temp 250 °C (Splitless for CI)
Carrier Gas	He at 1.4 mL/min, Constant Flow
Columns	Primary 30 m x 0.25 mm x 0.25 µm Rxi-5MS (Restek, Bellefonte PA) Secondary 0.6 m x 0.25 mm x 0.25 µm Rxi-17Sil MS (Restek, Bellefonte PA)
Oven Program	Primary Oven 40 °C (1 min), 10 °C /min to 330 hold 30 min Secondary Oven +15 °C Offset
Modulation Period (GCxGC)	4 seconds

GCxGC Resolution Improves Peak Identification

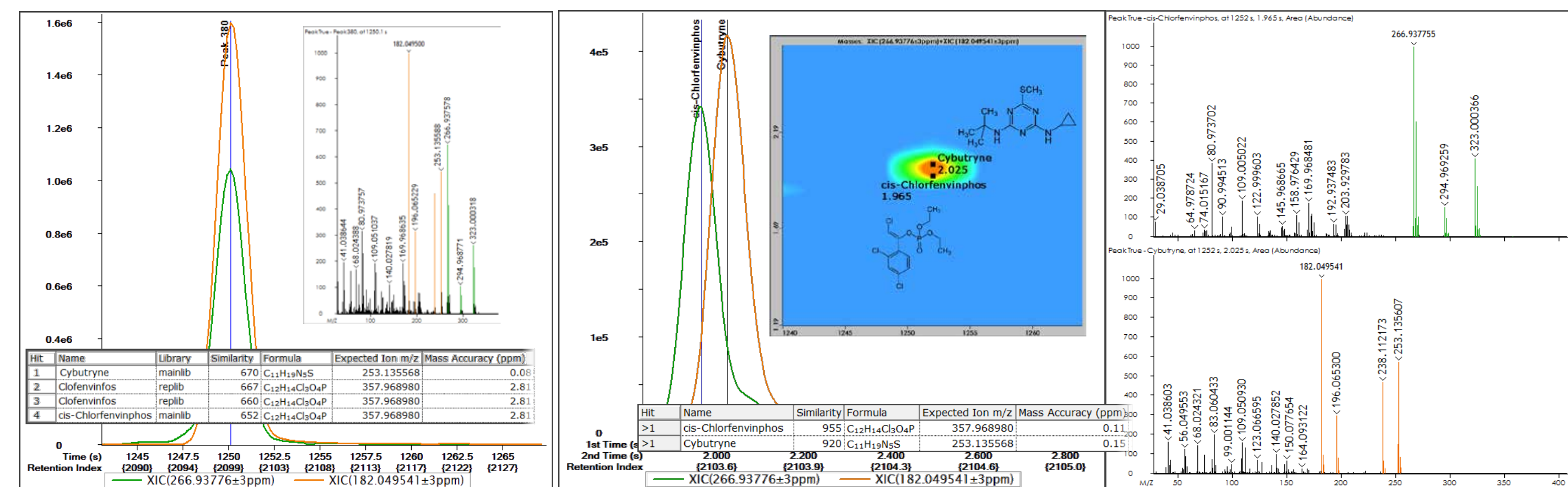1D Separation   GCxGC Separation and Deconvolution

Figure 1. Comparison of traditional 1D vs GCxGC separations.

Figure 1 demonstrates how the nearly identical chromatographic profiles and perfect 1D coelution makes deconvolution of these two peaks impossible, regardless of the MS resolving power, resulting in a combined spectrum with poor match scores. With GCxGC, the two peaks are separated by only 0.06 s, but that difference is enough to allow *ChromaTOF* to effectively deconvolve the two compounds, dramatically improving similarity scores, M+ mass accuracy, and overall match confidence. In this work GCxGC played a critical role in providing clean mass spectra for processing and interpretation.

HRAM CI – Working Beyond Libraries

While reviewing the data from the standard mixtures, chemical ionization (CI) was used to confirm molecular ions and occasionally led the analyst to select a compound not appearing in the library match list. In Figure 2 you can see a deconvoluted spectrum (left) and the best library match with a Similarity score of 792, but with a considerable mismatch at the presumptive molecular ion. *ChromaTOF's* Formula Calculator returns a more likely formula and a search of NIST 17 shows it does not contain a spectra for the likely compound. Using the CI data (right) allowed the suspected molecular ion to be confirmed. A ChemSpider search of the formula lists the suspected structure as the #1 hit.
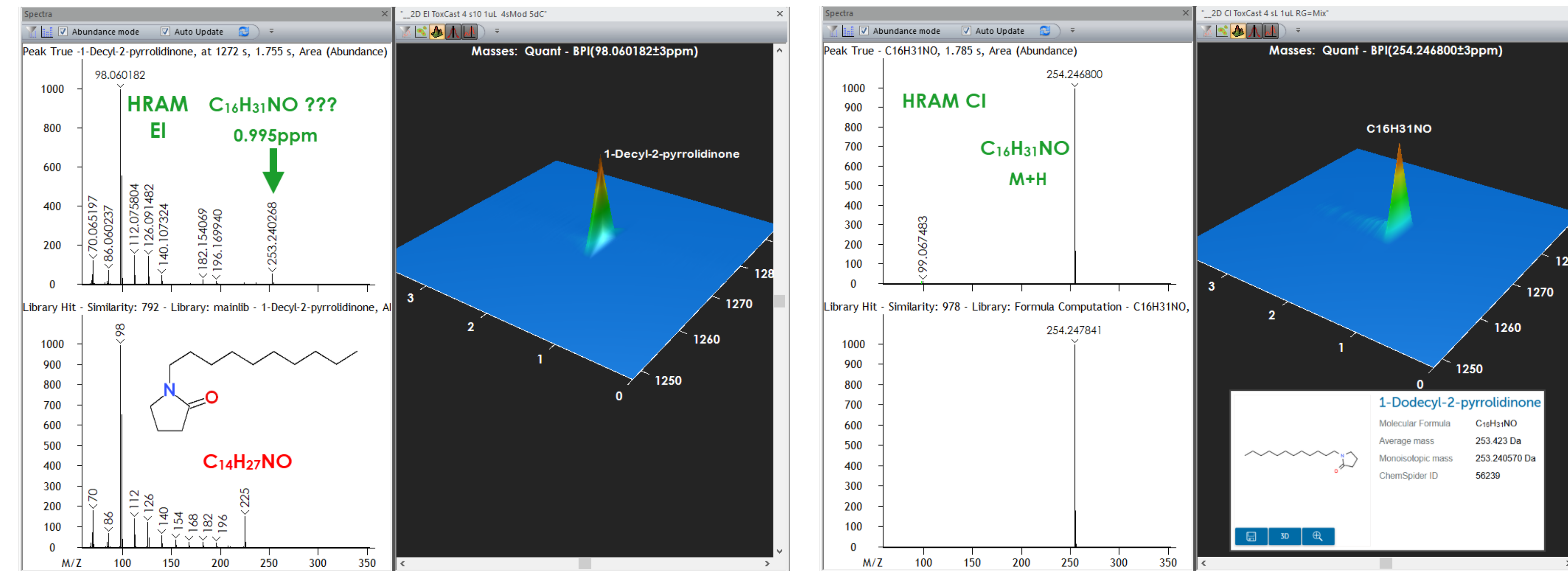

Figure 2. GC-EI spectra (left) and its best library match compared to the GC-CI review of the same analyte. The CI data was instrumental in correctly identifying the peak without a corresponding GC-EI library spectrum.

Data Analysis and Identification Confidence

As part of the sample evaluation, chromatographic peaks were assigned a confidence score using a HRAM GC-MS specific scoring system, which was developed by LECO and accepted by the EPA for this project. Each reported peak was assigned an identification confidence score (A, B, or X) based on the following criteria.

Tier A – All of the following are true:

- Forward spectral similarity score ≥700
- Molecular ion present and within 5 ppm of the expected m/z; may be confirmed with CI data
- Masses w/abundance ≥ 30% of base peak are within ±5 ppm based on library match's formulae
- RI value ±50 compared to NIST (semi-standard-non-polar)
- The reviewing analyst must be confident with the peak deconvolution and identification

Tier B – An "A" with some failing criteria; typically missing M+ or too many isomers to make definitive ID

Tier X – ID was made/changed based on unblinded review

All match filtering, similarity, and mass accuracy calculations were performed by *ChromaTOF* based on the selected library match formulas. Without automatic fragment mass accuracy calculations, that identification step would be tedious and time prohibitive. Identifications with confidence scores of A or B were reported to the EPA during the blinded phase. After the initial review, the EPA released the list of compounds present in each mix, and allowed for reevaluation of the data. Any identifications that were changed as a result of the target list were scored as "X". The match scoring system proved to be so useful that it was added to a subsequent version of *ChromaTOF* software. This new feature is now referred to as the Identification Grading System (IGS).
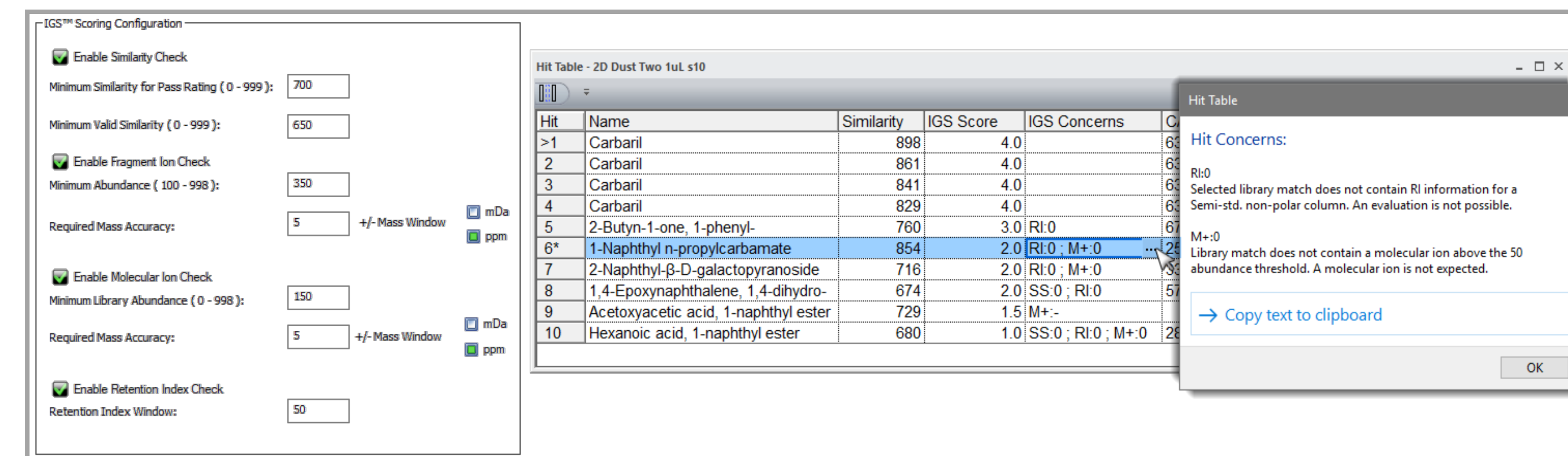

Figure 3. Image of the Identification Grading System (IGS) tool setup in the data processing method (left). Any of the evaluation criteria and limits may be enabled/altered at the user's discretion. Each criteria will award either +1 (pass), -1.5 (fail), or 0 (null result) points. The total score is used to rank the library matches in the Hit Table (right). If a library match does not receive a perfect score (4.0), the IGS Concerns column details why 0 or -1.5 points were awarded. In the image above, the highlighted match has a high similarity score, but because there was no RI information in the library and the library spectra did not include a molecular ion, these two criteria were awarded 0 points and the match was ranked lower in the Hit Table.

Streamlining Sample Review with IGS

Using the IGS Score as a filter you can quickly screen for the matches with the highest confidence score and then move on to peaks with weaker matches. Figure 3 shows the same sample at two different points in the review process. The top after the peak find and library search was run vs. the bottom which was automatically filtered to remove bleed peaks and retain peaks with IGS Score ≥ 2.
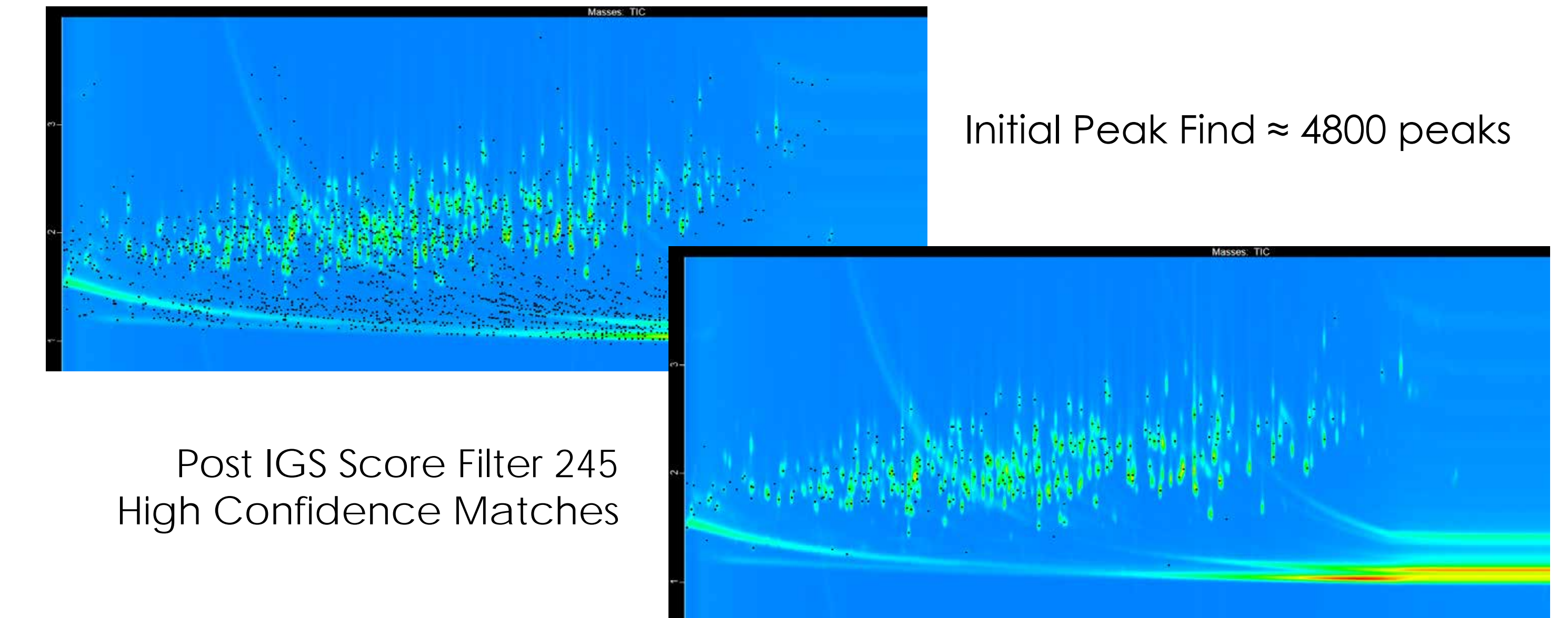

Initial Peak Find ≈ 4800 peaks

Post IGS Score Filter 245 High Confidence Matches

Figure 4. Within seconds the IGS Score filter can identify the highest confidence matches, a process that could take a trained chemist hours do manually.

Peak Detection and Identification Efficiency

Comparing the list of target compounds present in NIST 2017 against the number of found matches in each sample gives a good representation of how well the *Pegasus* HRT+ 4D performed in both the blinded and unblinded phases of this work. In Phase 1 (blinded), the *Pegasus* HRT+ 4D and *ChromaTOF* found, on average, ~85% of the spiked compounds that have spectra in NIST 17. The success rate increased to ~92% once the target list was revealed. An additional point worth consideration is the likelihood that at least some of the spiked compounds reacted in solution, and were therefore no longer present at the time of analysis.
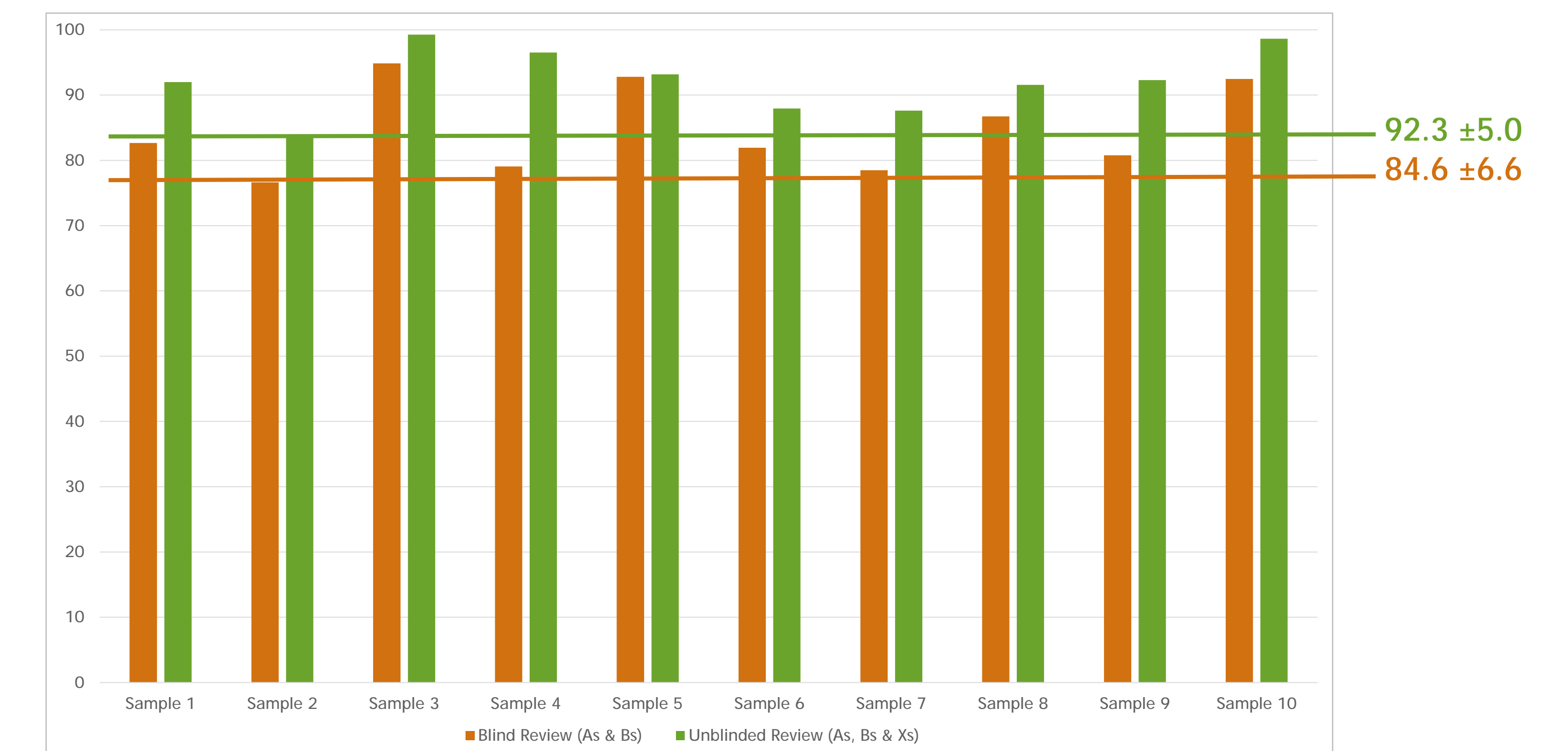

92.3 ±5.0

84.6 ±6.6

Figure 5. Summary of success rate per sample. There does not appear to be much, if any, correlation between success rate and sample complexity. Many of the increases between the blind and unblinded review can be attributed to updates of initially identified compounds to a related isomer.

Conclusions

- GCxGC dramatically improved chromatographic peak resolution leading to superior deconvolution and pure spectra for identification of non-target compounds in complex standard mixtures.
- LECO's industry-leading High Resolution Deconvolution® (HRD®) software feature provides clean mass spectra with unsurpassed spectral fidelity for library searching and spectral interpretation.
- LECO's Identification Grading System (IGS) can minimize time spent reviewing data while also adding confidence to compound characterization.