

Characterization of Metabolites From *Medicago truncatula* Using Gas Chromatography Mass Spectrometry—Addressing the Knowns and Unknowns Using High Resolution Time-of-Flight

Jeff Patrick¹, Joe Binkley¹, David Alonso¹, David V. Huhman², Feng Qiu², Dennis Fine², and Lloyd W. Sumner² | ¹LECO Corporation, St. Joseph, MI, ²The Samuel Roberts Noble Foundation, Ardmore, OK

Introduction

One of the principle challenges to metabolomic analyses is the identification of unknown analytes. Many studies report in excess of 60% of the analytes as unknown. Gas chromatography with mass spectrometry (GC-MS) has been a standard for metabolite analysis and historically relies on established databases for analyte identification. Nominal mass information, poor spectral quality, and limited databases have left many metabolites as poorly identified or unidentified. Here, metabolites from *M. truncatula* extracts are characterized using GC with high resolution time-of-flight (HRT) mass spectral detection. The combination of high resolution/accurate mass with either EI-MS (library matching and classification) or CI-MS (molecular formula assignment) was used to elucidate or confirm metabolite identities. This ultimately improves the biological context of the metabolomics experiments and provides the key information to determine **what else is in your samples**.

Methods

Plant tissues were homogenized by pulverization after lyophilization. Metabolites were extracted from plant tissue using four parts methanol in water. Portions of the extracts were dried under vacuum and lyophilized to remove the residual water. The residue was derivatized using standard protocols including methoxylamine HCl and MSTFA. The derivatized samples were then analyzed by GC-MS (see below). Both EI (70 eV) and CI (5% ammonia in methane) data were acquired.



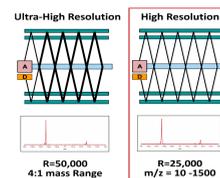
High Resolution GC-TOFMS

Chromatographic Parameters

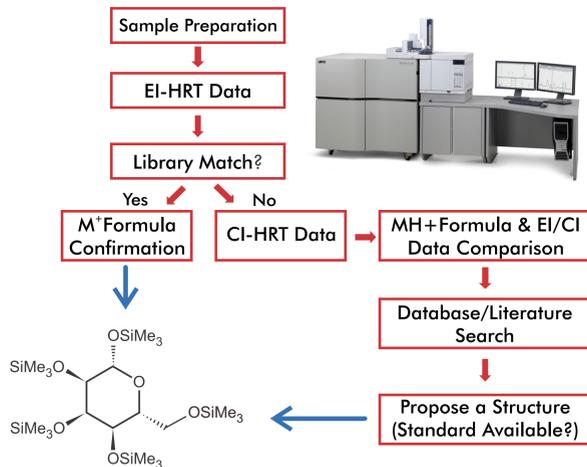
- GC: Agilent Technologies 7890 and 7693 Autosampler
- Column: DB-5MS (60 m x 0.25 mm x 0.25 μm) + 5m Guard
- Carrier/Flow: He, 1.0 μL/min
- Injection: 1 μL, 50:1 split (CI: 1 μL, Splitless)
- Inlet Temp.: 250°C
- Temp. Prog.: 50°C (1 min hold) to 315°C (10°C/min, 5 min hold)

Mass Spectrometry Parameters

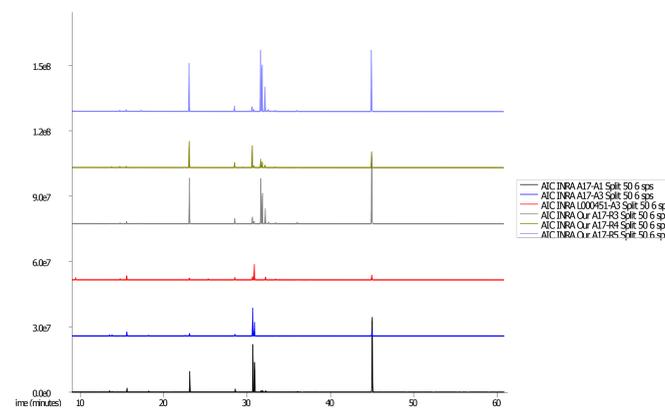
- Transfer Line Temp.: 260°C
- Ion Source Temp.: 250°C (CI: 200°C)
- Electron Energy: EI (70 eV); CI (140 eV)
- Range (m/z): 50 to 650 (CI: 60 to 1000)
- Flight Path: High Resolution (R = 25,000)
- Mass Calibration: PFTBA (Internal)
- HRCI Reagent Gas: 5% Ammonia in Methane



Electron Ionization/Chemical Ionization Workflow

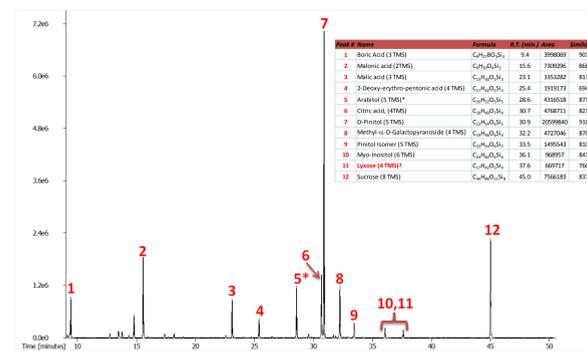


A representative workflow (above) for the use of EI and CI data in the elucidation of unknown metabolites. This workflow forms the basis of unknown identification using GC-HRT. System performance attributes are key to its success. The combination of EI fragmentation for database searching and structural characterization, together with CI for molecular formula determination, are used for confident identification of metabolites.



Above are representative chromatograms from the analysis of the *M. truncatula* cell extracts. Many peaks are common, but each has analytes unique to the extract. Chromatograms using CI were similar, but relative intensities were different owing to differences in ionization responses of the metabolites.

Confirmation of Knowns



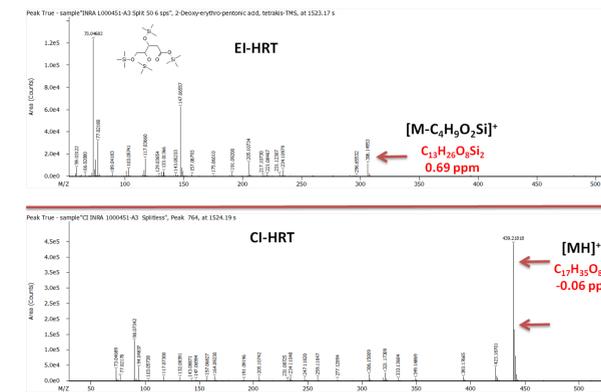
Focusing on one of the samples, 12 peaks can be readily detected and identified using EI-MS library matches (NIST/Fiehn Libraries). These are shown above (labeled 1-12) and are provided in the inset table. Spectral match scores for all but two of these analytes were in excess of 800. Two analytes (4. 2-deoxy-erythro-pentonic acid and 11. lyxose) showed lower scores and will be addressed later. These are candidates for evaluation using high resolution chemical ionization (HR-CI) to establish formulae from quasi molecular ions and/or adducts. The high spectral match scores support HRT-MS as producing relative fragment ion abundances consistent with traditional EI-MS spectra. This makes confident identification of common analytes straightforward.

Beyond the spectral match scores, mass accuracies can be applied to the EI spectra to further support identifications and, more generally, elemental composition. The molecular ions for the analytes in the table above were either not detected or of insignificant intensity. Using the ions assigned to "known" fragments one can calculate mass accuracies and further characterize the metabolites. As is shown below the observed accuracies are between ±1 ppm. In general, using the GC-HRT a mass error of greater than 2-3 ppm is indicative of an issue in the proposed identity for ions of reasonable intensity (ca. S/N > 25).

Peak	Name	Species	Formula	Expected Ion m/z	Observed Ion m/z	MA (ppm)
1	Boric Acid (3 TMS)	[M-CH ₃] ⁺	C ₉ H ₁₂ BO ₃ Si ₃	263.11208	263.11233	0.94
2	Malonic acid (2TMS)	[M-CH ₃] ⁺	C ₇ H ₁₀ O ₅ Si ₂	233.06599	233.06573	-1.11
3	Malic acid (3 TMS)	[M-C ₂ H ₅ O ₂ Si] ⁺	C ₁₃ H ₁₈ O ₅ Si ₃	233.10237	233.10218	-0.83
4	2-Deoxy-erythro-pentonic acid (4 TMS)	[M-C ₂ H ₅ O ₂ Si] ⁺	C ₁₃ H ₁₈ O ₅ Si ₂	306.14932	306.14953	0.69
5	Arabitol (5 TMS)	[M-C ₂ H ₅ O ₂ Si] ⁺	C ₁₃ H ₁₈ O ₅ Si ₃	307.15755	307.15743	-0.4
6	Citric acid, (4TMS)	[M-C ₂ H ₅ O ₂ Si] ⁺	C ₁₃ H ₁₈ O ₅ Si ₄	363.14738	363.14695	-1.18
7	D-Pinitol (5 TMS)	[M-C ₂ H ₅ O ₂ Si] ⁺	C ₁₃ H ₁₈ O ₅ Si ₃	318.14973	318.1497	-0.07
8	Methyl-α-D-Galactopyranoside (4 TMS)	[M-C ₂ H ₅ O ₂ Si] ⁺	C ₁₃ H ₁₈ O ₅ Si ₂	217.10746	217.10736	-0.47
9	Pinitol Isomer (5 TMS)	[M-C ₂ H ₅ O ₂ Si] ⁺	C ₁₃ H ₁₈ O ₅ Si ₃	318.14973	318.14968	-0.14
10	Myo-Inositol (6 TMS)	[M-C ₂ H ₅ O ₂ Si] ⁺	C ₁₃ H ₁₈ O ₅ Si ₃	318.14973	318.14946	-0.83
11	Lyxose (4 TMS)?	?	C ₉ H ₁₂ O ₅ Si ₂	217.10746	217.10721	-1.15
12	Sucrose (8 TMS)	[M-C ₂ H ₅ O ₂ Si] ⁺	C ₁₃ H ₁₈ O ₅ Si ₃	361.16812	361.16796	-0.44

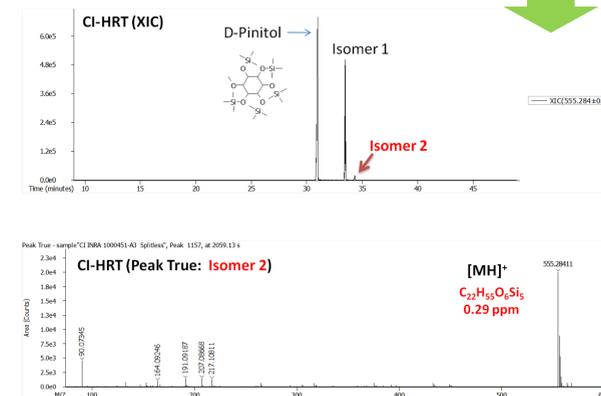
Ave. = 0.69

Dealing With the Unknowns

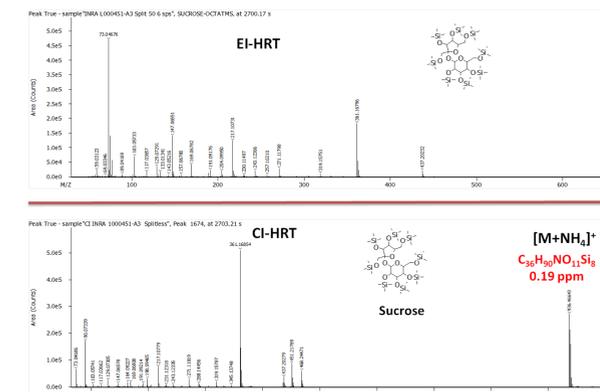


2-Deoxy-erythro-pentonic acid was identified and showed a marginal spectral match of 694. The EI spectrum produced no molecular ion to confirm the ID. Analytes in compound classes often share common fragment ions so the identification is not confirmed, but the low mass error observed in the fragment ions supports a formula consistent with the pentonic acid. To address this, the sample was analyzed using HR-CI on the GC-HRT. The resultant CI spectrum clearly shows a prominent ion at m/z 439.21818 which matches the formula for pentonic acid (TMS deriv.) to 60 ppb. This allows for confident identification of the analyte based on both structural features (EI) and molecular formula (CI).

Alternatively, HR-CI can be useful in extracting information from analyses in a selective fashion. Here three analytes are highlighted. One was identified as D-Pinitol (match >900). This was confirmed using HR-CI. In the same analysis two other analytes with the same exact mass were detected eluting shortly after pinitol. Below is the XIC for the accurate mass (555.284) using a narrow mass window for selectivity. The exact identification of the analytes (isomers 1 and 2) are not known, but the formulas are identical to, and fragmentation similar to, D-Pinitol. These are likely stereoisomeric forms of the cyclic polyol.



Larger Molecules



Many analytes of metabolic interest are conjugates or compounds with labile bonds which are fragmented during ionization with EI. This creates opportunities for misidentification or restricted identification to a class of compounds. In addition, GC-MS in general is perceived to be mass range limited. Above, both of these issues are addressed. The EI spectrum from a compound identified as sucrose (8TMS) is shown above. The absence of a molecular ion leaves the identity as uncertain with numerous disaccharides and substituted monosaccharides as possible IDs (scores >700). By using HR-CI with 5% ammonia in methane, the CI spectrum provides a strong molecular ion with a mass error of 0.19 ppm compared to the theoretical m/z for sucrose. Similar mass accuracies observed for the fragments in the HR-CI spectrum provide further confirmation for this derivatized disaccharide. The observed m/z at 936 also shows that larger metabolite derivatives are readily analyzed with high quality data using GC-HRT. The combination of EI and CI for the identification of labile compounds of higher molecular weight is of significant benefit in these studies.

Conclusions and Summary

Analyses of the *M. truncatula* extracts were performed, and data processing focused on qualitative analyses. This provided the opportunity for known analyte verification and unknown identification in a meaningful matrix and the demonstration of the key attributes and benefits of GC-HRT. Specifically the data here shows:

- The capabilities of GC-HRT are indispensable for analysis of complex plant extracts
- The analysis of labile biologics is facilitated by CI
 - Provides additional critical information for identification when paired with HRMS
 - Facilitates a different level of selectivity by using accurate mass, narrow window signal extraction
- GC-HRT provides high quality, accurate mass data for:
 - Spectral similarity searches (NIST, Wiley, etc.)
 - Formula determination (fragment, molecular, and adduct ions)
 - Confirmation of analyte molecular by utilization of CI