Confident Unknown Identification of SVOC Compounds by Combining NIST Library Search with Elemental Composition Determination

Jianping Chen[1], O. David Sparkman[2], and Ming Gu[3]

[1]Connecticut State Department of Environmental Protection, Windsor, CT; [2]University of the Pacific, Stockton, CA; [3]Cerno Bioscience, Danbury, CT

Overview

➢ A new method based on mass spectral peak shape calibration was developed for the identification of unknown Semivolatile Organic Compounds (SVOCs).

➢ On a unit mass resolution quadrupole GC/MS instrument, 122 SVOC compounds were identified and/or confirmed by NIST library search and MassWorks formula determination.

➢ Complementary approaches to SVOC unknown ID by NIST Mass Spectra Database and CLIPS search were discussed.

Introduction

Using GC/MS for the rapid identification of releases of hazardous or toxic chemicals requires confirming the analytes identity by a second method. High resolution MS such as TOF may provide this confirmation, but it is not a option due to high cost and limited space available in this mobile lab. Alternatively, a novel method which provides a primary and secondary confirmation with a single analysis using a well established EPA procedure for SVOCs has been developed. The method is based on formula determination by advanced data processing of integer-mass-only results produced by the GC-MS in conjunction with a conventional library search. This dual process dramatically improves the accuracy of chemical identification.

Methods

➢ **Data Acquisition:** All data were acquired in an Agilent 6890-5973N GC/MS system in a raw (profile) scanning mode and followed by NIST MS Database Search and elemental composition determination by MassWorks. A total of 122 SVOC compounds were investigated. About 0.1 mg of each compound were injected and separated by GC/MS according to EPA methods. At the end of run, PFTBA spectra were acquired for 20 seconds to built calibration by MassWorks. The system can hold the calibration for days if not weeks before recalibration.

➢ **NIST MS Database Search:** By an Agilent ChemStation data processing script, the profile spectra were automatically converted into centriod spectra for NIST MS Database Search.

➢ **MassWorks processing:** MassWorks calibrates both the mass position and the mass spectral peak shape function, a key for achieving high mass accuracy. When applied to raw mass spectral data, the raw mass spectrum can be transformed into its calibrated version with mass spectral peaks located at accurate mass positions. Furthermore, the mass spectral peak shape would also be transformed in the same process to a mathematically definable function, a key for achieving high spectral accuracy and CLIPS formula ID.

➢ **Performance Evaluation:** Among the 122 SVOC, 13 compounds with a mass range from 93 to 276 were used as routine performance check before running unknown substance identification . High spectral accuracy of 98.5% and top 3 search ranking are required performance criteria for those 13 compounds before applying it to identify real unknown compounds.

Results and Discussion

MassWorks Calibration+CLIPS Formula ID
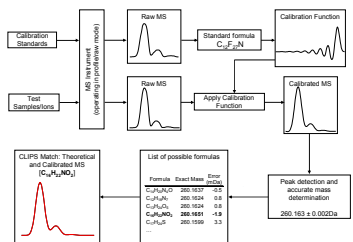
For each known ion fragments from the PFTBA standard, both mass and peak shape are adjusted as part of the calibration to enable highly accurate mass determination at unit mass resolution in the presence of isotope interferences, as shown from the $C_3F_5^+$ fragment ion (accurate mass 130.9907 vs. exact mass 130.9920Da).

Fig 1. Mass and Peak Shape Calibration with PFTBA
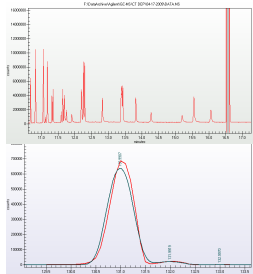
This calibration is then applied to the data file of GC/MS of SVOC. First, the 13 selected performance checking standards were analyzed. As shown in Table 2, all the standards achieved better than expected spectral accuracy of 98.5% and expected ranking of top 3 on search list. As an example shown in Fig 2., Acenaphthylene ranked as top 1 has spectra accuracy of -0.3 mD and spectral accuracy 99.7%. This great performance is attributed to the novel calibration technology that corrects unsymmetrical peak shape (Fig. 2a) and provides effective isotope pattern matching (Fig2 2b) between calibrated and theoretically calculated spectra.

Secondly, accurate mass measurements and formula identification were conducted on all the SVOC. Formula search conditions listed on the Table 1 include a wide range of elements and very open lower and upper limits. The electron state of "odd" electron was selected to focus on the analysis of the compounds having molecular ions which accounts for 91% of the SVOC under investigation. Fig 3. shows the ranking distribution of both NIST MS Database and CLIPS search. 94% of the SVOC appeared on top 3 from CLIPS search lists including 76% 1st, 16% 2nd, 2% 3rd hits for the compounds with molecular ions. NIST library search generated positive ID for 92% of SVOC regardless the

Table 1. CLIPS Search Parameters

		Element	Min	Max
Accurate Mass	130.99			
Charge	1	C	0	50
Mass Tolerance (mDa)	20	H	0	100
Electron State	Odd	N	0	15
DBE Min	-10	O	0	15
DBE Max	50	S	0	10
Profile Mass Start (Da)	-0.5	F	0	12
Profile Mass End (Da)	4.5	P	0	12
		Br	0	10
		Cl	0	12

Table 2. Mass and Spectral Accuracy for Performance Checking Standards

Target Compounds	Formula	Retention Time	m/z	Rank	Spectral Accuracy	Mass Accuracy	Total Formu
Aniline	C6H7N	4.82	93	1	99.18	-4.1	27
1,3-Dichlorobenzene	C6H4Cl2	5.07	146	1	99.60	2.6	186
Benzyl alcohol	C7H8O	5.45	108	2	99.66	-7.4	62
3+4-Methylphenol	C7H8O	5.51	108	1	99.85	-4.9	62
O-Toluidine	C7H9N	5.78	107	1	99.61	-5.5	39
N-Nitrosopiperidine	C5H10ON2	6.18	114	1	99.75	7.3	72
2,4-Dichlorophenol	C6H4OCl2	6.36	162	1	99.41	-0.1	288
Naphthalene	C10H8	7.34	128	1	97.48	0.2	116
2,4,5-Trichlorophenol	C6H3OCl3	7.94	196	2	98.63	-6.3	479
Acenaphthylene	C12H8	9.38	152	1	99.74	-6.3	222
3,3'-Dichlorobenzidine	C12H10Cl2N2	12.19	252	2	98.66	-1.8	2273
Benzo(a)pyrene	C20H12	13.91	252	1	99.73	-4.4	1469
Benzo[ghi]perylene	C22H12	16.07	276	1	99.49	3.7	2195

Fig 2. Raw, Calibrated, and Calculated Spectra of m/z 152
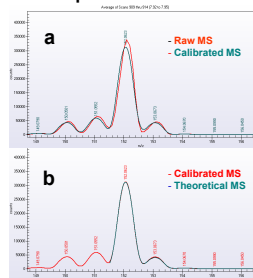
Table 3. Top 10 of Search for m/z 152

Row	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (PPM)	Spectral Accuracy	DBE
1	C12H8	152.0626	-0.3184	-2.0936	99.7377	9
2	C9H6F2	152.0438	18.5253	121.8271	98.9747	6
3	C9H9OF	152.0637	-1.4612	-9.6093	98.8539	5
4	C9H13P	152.0755	-13.205	-86.8398	98.6873	4
5	C8H8O3	152.0473	14.9378	98.2346	98.6661	5
6	C8H9FN2	152.075	-12.6946	-83.4829	98.5936	5
7	C7H9N2P	152.0503	11.9471	78.5871	98.564	5
8	C7H8O2N2	152.0586	3.7044	24.361	98.5006	5
9	C6H5FN4	152.0498	12.4575	81.9238	98.4565	6
10	C9H12S	152.066	-3.6892	-24.2609	98.3343	4

existence of the molecular ions. Poor results from NIST library search most likely were due to complex mass spectra caused by co-eluting compounds. For example, the major peaks of the spectra of N-Nitrosopyrrolidine should be m/z 100 as molecular ion and a dominant fragment at m/z 41 (Fig 4b). However, this compound was co-eluted with other molecules and its spectra appears to be much more complicated (Fig 4a). This resulted in poor match and reverse match values at 392 and 593 respectively by NIST MS Database Search. Instead of matching multiple ions fragmentation pattern by a NIST MS Database Search, MassWorks performs formula determination by analyzing individual ions and is able to deal with co-eluted problems as long as the interesting ions are not overlapped by another unknown ion.

Fig. 5 shows that the elemental composition of N-Nitrosopyrrolidine was found to be the 2nd hit by CLIPS search with mass accuracy of 4.5 mD and spectral accuracy of 99.3%

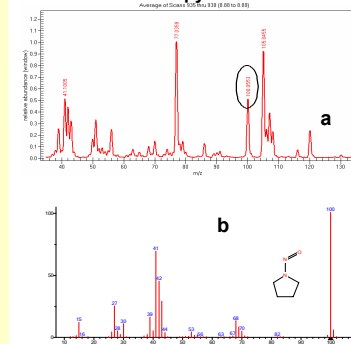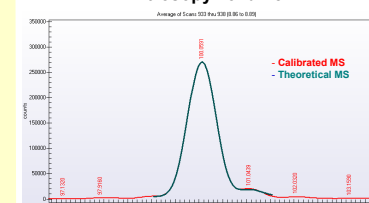### Fig 4. Complex and pure spectra of N-Nitrosopyrrolidine

Fig 3. NIST Library and CLIPS Search Rank Distribution

CLIPS 94% — Top 3 Hits / Hit #4 to #10

NIST 92% / 6% — Failed / Broadline / Confirmed

Fig 5. Isotope Pattern Match for N-Nitrosopyrrolidine

Row	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (PPM)	Spectral Accuracy	RMSE	DBE
1	C5H8O2	100.0524	-6.6705	-66.6656	99.3219	835	2
2	C4H8N2O	100.0637	4.5628	45.6019	99.3192	838	2
3	C6H9F	100.0688	9.7285	97.2274	98.9866	1,247	2
4	C7H4N6	100.0497	-9.3558	-93.5031	98.4509	1,906	3
5	C3H7F3	100.05	-9.1152	-91.0979	98.4372	1,923	-1

Conclusions

➢ With PFTBA calibration performed to each GC/MS run, mass accuracy better than 10mDa and Spectral Accuracy better than 99% can be achieved for reliable formula determination.

➢ Based on both NIST library and CLIPS search, confident identification of unknown SVOCs can be achieved.

➢ When NIST library search results in multiple compounds with similar Match Value (MV) and Reverse Match Value (RMV), MassWorks can provide unique formula determination for unknowns.

➢ NIST MS Database Search can provide positive identification regardless the existence of molecular ions.

References

1. Tyler, A. N. et al, *Anal. Chem.* **Exact Mass Measurement of Polar Organic Molecules at Low Resolution Using Electrospray Ionization and a Quadrupole Mass Spectrometer,** 1996, 68, 3561.
2. Fiehn, O. et al, *Anal. Chem.* **Identification of Uncommon Plant Metabolites Based on Calculation of Elemental Compositions Using Gas Chromatography and Quadrupole Mass Spectrometry,** 2000, 72, 3573.
3. Gu, M et al, *RCM* **Accurate mass filtering of ion chromatograms for metabolite identification using a unit mass resolution liquid chromatography/mass spectrometry system**, 2006, 20, 764.
4. Wang, Y.; Prest, H. *Chromatography (Japan)* **Accurate Mass Measurement on Real Chromatographic Time Scale with a Single Quadrupole Mass Spectrometer**, 2006, 27, 135.