United States
Environmental Protection
Agency

Environmental Research
Laboratory
Athens GA 30613

Research and Development

EPA-600/S4-84-072 Sept. 1984

EPA

Project Summary

Application of an Analysis Protocol to Identify Organic Compounds Not Identified by Spectrum Matching

Joan T. Bursey

Industrial wastewater survey samples were analyzed for organic compounds not identified by spectrum matching. Analysis of the samples proceeded from an initial packed column GC/MS analysis for Priority Pollutants, through computerized spectrum matching for other organic compounds to the present capillary column GC/MS analysis of a chosen set of sample extracts. Attention was focused on the spectra seen to occur frequently yet not tentatively identified by spectrum matching.

A plan for systematic study of these sample components was devised that included, in step-wise fashion, the use of high resolution gas chromatography (HRGC), high resolution mass spectrometry (HRMS), chemical ionization mass spectrometry (CIMS) with positive and negative ion detection (NCI), and Fourier transform infrared spectroscopy (FT-IR). Sample cleanup was used at all levels to mitigate interference.

For 55 extracts in which components of interest were observed, accurate mass measurement was successfully used to generate chemical formulas in 35 cases. Of these, the results for 16 could be narrowed to one or two possibilities each. Tentative structures were proposed in six cases. Since the proposed compounds were not commercially available and the costs of synthesis were prohibitive, no further confirmation was made.

Conclusions were: 1) that this type of compound isolation/identification effort is very time and labor intensive, 2) that the labor costs are high because highly trained and experienced personnel are required, and 3) that the amount of definitive information that can be obtained by application of any one of the analytical techniques discussed above (or of several of the techniques in succession) ranges from minimal to very high, but integration of all the information available is often not as simple as the analyst might wish.

This Project Summary was developed by EPA's Environmental Research Laboratory, Athens, GA, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

In June 1976, a Consent Decree was signed by EPA and several public interest groups that set deadlines for EPA to promulgate regulations for treating industrial wastewater. Beginning in August 1977, some 4,000 samples of plant influents and effluents were taken in 21 industrial categories to gather data on treatment processes. The analytical work was centered on 114 organic compounds and 15 inorganic substances called Priority Pollutants.

The sampling and analysis effort continued through August 1980. In 1979, an amendment to the Consent Decree directed EPA, among other things, to determine, by mass spectrum matching, the compounds, in addition to the Prior-

ity Pollutants, that might be present, chemically confirm those compounds of highest concentration and frequency of occurrence, and attempt to identify by other means significant components not identified by spectrum matching. Computerized spectrum matching of the raw GC/MS data was performed by the Athens Environmental Research Laboratory (ERL). The chemical confirmation of matching results was presented in *Analysis of Industrial Wastewater for Organic Pollutants in Consent Decree Survey Samples* (EPA-600/4-83-028).

The original computerized analysis aimed at deconvoluting and identifying all of the peaks of the chromatogram. A number of compounds, however, could not be identified by the computer but were observed frequently in several industrial categories. The objective of the second part of this program was to endeavor to identify these compounds or at least arrive at a point where a structure could be postulated on the basis of available information.

There were several steps in the screening process used to select the extracts for analysis. Because a primary criterion was frequency of occurrence of a characteristic mass spectrum that could not be identified, a compilation was made of these mass spectra, ordered by frequency of occurrence. The spectra were screened by the Project Officer and the Principal Investigator, both experienced mass spectroscopists, a few identifications that appeared obvious were made, and some of the mass spectra were deleted from consideration for further analytical effort.

Another major criterion for the selection of a mass spectrum for further analytical effort was an obvious indication of the presence of halogen, that is, the presence of the isotopic clusters characteristic of the occurrence of halogen. The ultimate and deciding criterion was the availability of an extract in which the unknown compound was present in a concentration sufficient to produce a mass spectrum of reasonable quality. If several industrial effluent extracts containing the unknown mass spectrum were available, the highest priority was given to the extract in which the component was present at the highest concentration.

Experimental Procedures

Capillary gas chromatography with accurate mass measurement was performed under the same conditions as those shown in Table 1 for the Finnigan 4021 GC/MS system. The accurate mass measurements were performed on a VG Micromass 7070HS, at a resolving power of 2000, with a 5% valley. Ionizing energy was 70 eV, 200 μA trap current, a source temperature of 210°C, and a transfer line temperature of 275°C. The scan speed was 700 msec/decade, from *m/z* 550-20, with a 500 msec reset time. C_2I_4 was used as a secondary reference. Data were processed on a VG 2035 data system.

Chemical ionization analyses were performed in the pulsed positive/negative chemical ionization mode. The instrument used was a Finnigan 4500 GC/MS/DS. Capillary columns and conditions correspond to those used on the Finnigan 4021 and reported in Table 1. The reagent gas was methane, at a source pressure of 0.7 torr and an analyzer pressure of 2×10^{-5} torr. The source temperature was 120°C, with an electron energy of 100 eV and an emission current of 0.2 mA. The scan cycle was 1 sec, with a scan range of 90-550 daltons. The multiplier was operated at 1300 V, with the amplifier at 10^{-7} A/V.

The Unacon 810A is a commercial injection system that uses a preconcentration technique to increase injection volume and thereby decrease analytical detection limits. In operation, the Unacon removes solvent before injection of analytes onto the chromatographic column. The Unacon 810A system was in-

Table 1. *Instrumental Parameters*

	Finnigan 3300	*Finnigan 4021*
	Gas Chromatography Parameters	
Column	*60 m OV-101 wide-bore thick film fused silica capillary.*	
Initial Temperature	*30°C*	
Minutes Isothermal	*6 minutes at 75°C*	
Programming Rate	*4°C/min*	
Maximum Temperature	*270°C*	
Column Flow	*1.2 mL/min He*	*1.6 mL/min He*
Injector Temperature	*250°C*	*270°C*
Injection Mode	*Splitless for 60 sec, then ~10:1 split*	*Splitless for 24 sec, then 7.5:1 split*
Column	*60 m Carbowax 20 M wide-bore thick film fused silica capillary.*	
Initial Temperature	*30°C*	
Minutes Isothermal	*6 minutes at 75°C*	
Programming Rate	*4°C/min*	
Maximum Temperature	*220°C*	
Column Flow	*1.2 mL/min He*	
Injector Temperature	*250°C*	*230°C*
Injector Mode	*Splitless for 60 sec, then ~10:1 split*	*Splitless for 24 sec, then 2.5.1 split*
	Mass Spectrometer Parameters	
Transfer Line/Separator Oven	*270°C*	
Ionization Energy	*70 eV*	
Emission Current	*0.5 mA*	
Mass Range	*40-500 daltons*	
Scan Speed	*2 sec/cycle*	*1 sec/cycle*

terfaced to a Varian 3700 gas chromatograph using a fused silica capillary column. The capillary column was attached directly to the Unacon 810A's transfer line, which was wrapped with heating tape and heated to 165°C to prevent condensation of the less volatile organics during transfer. The Unacon 810A system was equipped with a large-bore trap (6.35 mm O.D, 2.54 mm ID x 20.3 cm in length) containing Tenax GC® and glass beads and a small-bore trap (6.35 mm ID, 0.9 mm ID x 20.3 cm in length) also containing Tenax and glass beads as sorbent material.

A Nicolet 7199 Fourier transform infrared spectrometer was used with a Model 7001A Michelson interferometer, a high energy cooled glow bar source, a heated 0.3 mm ID x 42 cm (3 cm^3) gold-plated light pipe, and a 4000 to 800 cm^{-1} range liquid nitrogen cooled Mercury-Cadmium-Telluride detector, with related optics. The entire system was enclosed within a sealed nitrogen-purged enclosure.

The spectrometer was controlled by a Nicolet 1180 data system with 40K solid state memory interfaced to the spectrometer by a high precision 15 bit analog-to-digital converter. Twenty-bit spectral words were acquired and stored on a Diablo 44B dual density, dual disk memory system. Long-term data storage was accomplished using a Kennedy 9000 9-track, 800 bits-per-inch, 37 inches-per-second tape deck.

Results and Discussion

An analysis protocol developed by the Athens ERL was adhered to in this phase of the program. A flow chart of the protocol is presented in Figure 1. Progress was evaluated in relation to the points in the flow chart labeled A, B, etc.

Point A ("capillary GC/low resolution EIMS analysis. Pertinent component still observed?") was the starting point for each industrial effluent extract. Three outcomes were possible from this analysis: 1) the component of interest was observed (progress to points B and C); 2) the component of interest was not observed (analysis of the pertinent industrial effluent extract was abandoned and another extract was substituted, if available); or 3) a mass spectrum only similar to (but not exactly corresponding to) the component of interest was observed (same outcome as [2], above). At point B ("sufficient GC separation?"), the GC separation did not prove crucial to the capability of observing a mass spectrum for the component of interest. Using computerized background subtraction techniques, it was possible to deconvolute mass spectra even in the face of severe coelution problems. Several extracts (see point E) were subjected to cleanup procedures, however, because of their extreme complexity.

At point C ("Is the component still unidentified?"), the answer was almost invariably that the component was still unidentified. At point D ("Manual interpretation of low resolution EI mass spectrum: tentative identification?"), some structural features could be eluci-

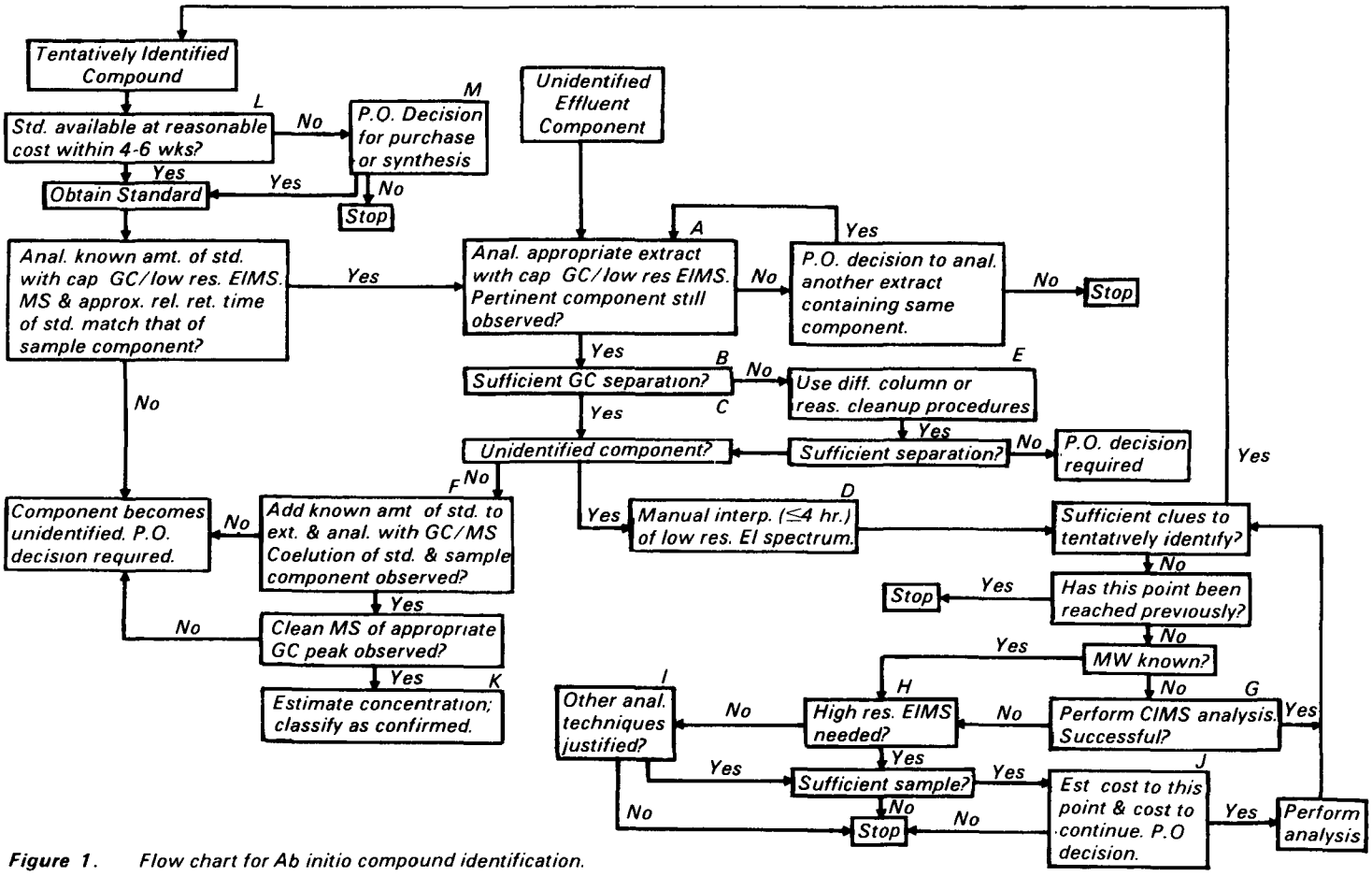


Figure 1. *Flow chart for Ab initio compound identification.*

dated but only rarely could a postulation of a complete structure be made. For all practical purposes, point F ("Add standard and reanalyze. Coelution of standard and sample component?") was never reached. Only once was a standard obtained commercially corresponding to a postulated structure; in this instance, elution times of standard and component of interest did not correspond. At point L ("Standard available at reasonable cost in 4-6 weeks?"), the standards to correspond to tentative structures were found to be unavailable commercially. At point M ("Project Officer decision for purchase or synthesis."), synthesis was considered at one time when the postulated structure seemed to be narrowed down to one possible isomer, but the projected cost of $2,000-3,000 for a single compound was prohibitive, and no synthesis was performed.

At point G ("Perform CIMS. Successful?"), chemical ionization mass spectrometry (in both positive and negative mode) was performed for a number of samples. One of the primary instances of application of the chemical ionization technique was the occurrence of components of interest that appeared to be highly halogenated alkanes and alkenes. These compounds do not produce molecular ions (or any species characteristic of molecular weight) under electron ionization conditions. The mass spectrometric literature does not show instances of successful application of chemical ionization techniques to these types of compounds, but an effort was made (in both positive and negative mode) to see whether any information could be obtained.

Point H ("High resolution CIMS analysis needed?") proved to be crucial in the decision-making process, and the high resolution EIMS technique was applied to most of the extracts. Attempts to use high resolution techniques without the benefits of the preceding capillary separation were uniformly unsuccessful.

At point I ("Other analytical techniques justified?"), a major decision needs to be made. The criterion for deciding whether other analytical techniques are justified is a cost-benefit analysis: relative to the cost in time and dollars of performing the proposed analysis, is the structural information to be gained from the application of the technique proportionate?

Results of the most successful attempts at compound identification are shown in Table 2. Structure postulation was not made in most cases due to the lack of functional group information that might have been obtained from a more sensitive GC/FT-IR analysis. Formulae were not postulated in cases where no molecular ion could be produced by CIMS. No standards for any of the postulated compounds were obtainable commercially, and the expense of custom synthesis was deemed too great.

If the objective of the analytical procedures is to confirm a structural postulation by co-injection of a standard, five conditions must be met:

1) micrograms of sample material must be available for analytical methods that can provide the most definitive structural information;
2) extensive resources must be available from the point of view of accessibility to major analytical instrumentation;
3) extensive financial resources must be available in order to make the most effective use of analytical techniques that require major instrumentation. Few people are sufficiently specialized in interpreting data from the wide range of analytical techniques that must be used;
4) a major financial commitment could be required to obtain analytical standards that corresponds to the structures postulated. Where it is impossible to decide between two or three possible isomers on the basis of available spectral information, extensive custom synthesis could be required;
5) an overall success rate is very likely to be low despite effective use of all available techniques. It would be possible, for example, to

Table 2. *Summary of Selected Analysis Results*

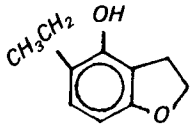
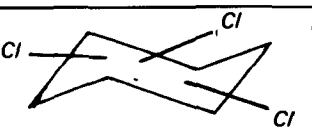
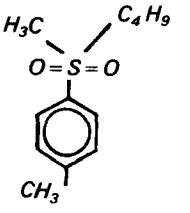
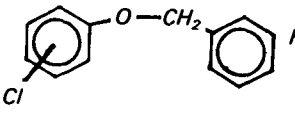
Extract # (Fraction)	*Industrial Category*	*Number of Occurrences*	*Postulated Formula(s)*	*Postulated Structure*	*Instrumental Techniques Used Successfully*
584 (ACI)	*Pulp and Paper*	*9*	$C_8H_7O_3CL$ $C_5H_{11}SO_3Cl$	*None*	*HRMS*
7367 (B/N)	*Amusements and Athletic Goods*	*10*	$C_9H_{18}N_4O$	*None*	*HRMS*
13230 (ACI)	*Mechanical Products*	*6*	$C_9H_{19}Cl_2PO_2S$	*None*	*HRMS, CI*
19436 (ACI)	*Plastics and Synthetics*	*9*	$C_9H_{20}O_3$	*None*	*HRMS*
21258 (B/N)	*Aluminum*	*5*	$C_{16}H_{16}O_2$ $C_{15}H_{16}ON_2$ $C_{11}H_{16}O_4N_2$	*None*	*HRMS*
5855 (ACI)	*Pesticides*	*5*	$C_{10}H_{12}O_2$		*HRMS, FT-IR*
11338 (ACI)	*Pesticides*	*29*	$C_6H_4O_3F_2$	*None*	*HRMS*
11536 (B/N)	*Inorganic Chemicals*	*5*	*None*	*None*	
18959 (ACI)	*Pulp and Paper*	*5*	$C_8H_{16}O_4$	$CH_3(CH_2)_2CH(OH)$- O-$(CH_2)_3CO_2H$	*HRMS, CI*
20250 (BN)	*Explosives*	*16*	$C_{15}H_{24}O_4$	*None*	*HRMS*

Table 2. *(Continued)*

Extract # (Fraction)	Industrial Category	Number of Occurrences	Postulated Formula(s)	Postulated Structure	Instrumental Techniques Used Successfully
2389 (ACI)	Inorganic Chemicals	6	$C_6H_{19}Cl_3$		HRMS
2390 (ACI)	Inorganic Chemicals	54	Series of halo alkyls	None	
5694 (B/N)	Organic Chemicals	17	$C_{19}H_{13}SO_2N$		HRMS, CI
5849 (ACI)	Pesticides	5	$C_5H_{10}N_2O_5$	None	HRMS
5859 (ACI)	Pesticides	7	$C_4H_4N_2S_2$	$NC\text{-}CH_2\text{-}S\text{-}S\text{-}CH_2CN$	HRMS
17245 (B/N)	POTW	11	$C_{13}H_{11}OCl$		HRMS

have several viable possibilities after accurate mass measurement, to use infrared spectrometric techniques to pinpoint a number of functional groups, and still to be unable to postulate a definitive structure for a component of interest.

Conclusions

Several conclusions can be reached on the basis of evaluation of data obtained in the course of this phase of the program. The ability to perform capillary gas chromatographic analysis in conjunction with accurate mass measurement is invaluable. This combined technique should be more widely available and much more widely applied. For an effort that is proportionately not much more than that required to perform the low resolution analysis, information about the atomic composition of the component of interest can be obtained and used in making further decisions in the chemical analysis.

An analytical technique that appeared very promising was the coupled technique of capillary gas chromatography Fourier transform infrared spectrometry, prefaced by the Unacon automatic sample concentrator. In the one successful case, structural features delineated by the infrared data integrated smoothly with the other information available (low resolution electron ionization mass spectrum and accurate mass measurement) to produce a structural postulation that narrowed the possibilities to one of two isomers, one more probable than the other.

Most of the analytical methods employed in this program have been found to be very labor intensive, especially in the area of highly trained technical personnel (i.e., experienced degree-holders at Bachelors, Masters, and Ph.D. levels). The most efficient and cost-effective way to take advantage of the available mass spectrometric techniques is to have direct access to a single mass spectrometer where the original capillary gas chromatography/low resolution electron ionization mass spectrometric analysis can be performed. This analysis would then logically be followed immediately by re-injection of the sample and performance of the accurate mass determination, with subsequent re-injection(s) of the sample to take advantage of available chemical ionization techniques, if appropriate.

Recommendations

The single major recommendation for optimizing this type of analysis is to have all applicable analytical techniques readily accessible, so that all information can be obtained in a short time and integrated quickly. It might also be helpful to other analysts doing similar types of analyses to gather the mass spectra of these components of interest that have been observed many times, in several types of industrial effluents, even though the compounds might be unidentified. If such a data base indicates that occurrence of a particular mass spectrum is frequent and that this mass spectrum is observed through a wide range of industrial effluents, then pursuit of definitive structural information would reasonably be assigned a higher priority, and perhaps then collection of sufficient samples would allow isolation of the component in relatively pure form by preparative analytical techniques. Access to micrograms of a relatively pure material would permit analysis by nuclear magnetic resonance and would certainly allow successful FT-IR analysis, and it should then be easier to formulate structural assignments.

Joan T. Bursey is with Research Triangle Institute, Research Triangle Park, NC 27709.
***Walter M. Shackelford** is the EPA Project Officer (see below).*
The complete report, entitled "Application of an Analysis Protocol to Identify Organic Compounds Not Identified by Spectrum Matching" consists of two parts:
Part 1: Text (Order No. PB 84-229 715; Cost 23.50, subject to change).
Part 2: Appendices (Order No. PB 84-229 723; Cost 22.00, subject to change).
The above reports will be available only from:
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650
The EPA Project Officer can be contacted at:
Environmental Research Laboratory
U.S. Environmental Protection Agency
Athens, GA 30613

United States
Environmental Protection
Agency

Center for Environmental Research
Information
Cincinnati OH 45268

Official Business
Penalty for Private Use $300