DETECTION, CHARACTERIZATION AND CLASSIFICATION OF BIOLOGICAL COMPONENTS

IN AEROSOLS BY TIME-RESOLVED LASER PYROLYSIS MASS SPECTROMETRY

AD-A179 865

FINAL REPORT

HENK L. C. MEUZELAAR, WILLEM WINDIG

MARCH 20, 1987

DTIC
SELECTE
APR 2 8 1987
S   D
D

U. S. ARMY RESEARCH OFFICE

CONTRACT NO. DAAG29-84-K-0009

UNIVERSITY OF UTAH

BIOMATERIALS PROFILING CENTER

20030121161

TABLE OF CONTENTS

ADA179865

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
	ARO 21079.6-LS

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
University of Utah Biomaterials Profiling Center		U. S. Army Research Office

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
391 S. Chipeta Way, Suite F, Research Park Salt Lake City, UT 84108	P. O. Box 12211 Research Triangle Park, NC 27709-2211

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
U. S. Army Research Office		*DAAG29-84-K-0009*

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
P. O. Box 12211 Research Triangle Park, NC 27709-2211	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

Detection, Characterization and Classification of Biological Components In Aerosols by Time-Resolved Laser Pyrolysis Mass Spectrometry.

12. PERSONAL AUTHOR(S)
Henk L. C. Meuzelaar, Willem Windig

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
FINAL	FROM 12/83 TO 11/86	3/20/87	70

16. SUPPLEMENTARY NOTATION

The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	biodetection, aerosols, CO_2 laser pyrolysis, Curie-point pyrolysis, mass spectrometry, simultaneous ion detection, ion trap detector, nucleic acid, albumin, glycogen,

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The original objective of the work reported here was to develop computerized pyrolysis mass spectrometry techniques for detection of biological aerosol components in general and time-resolved CO_2 laser pyrolysis MS methods in particular. Since laser pyrolysis methods require MS instruments with simultaneous ion detection capabilities, availability of a special miniature Mattauch Herzog MS system with electrooptical ion detector was an important element of the original proposal. Unfortunately, this system (funded by an independent DOD/University Instrumentation Program Grant) was never delivered due to bankruptcy of the manufacturer. Consequently, the work plan was focussed more strongly on the development of advanced chemical pattern recognition methods. Nevertheless, successful preliminary experiments were performed with CO_2 laser pyrolysis using a quadrupole mass filter. Moreover, in the third year of the project a Finnigan MAT Ion Trap Detector (ITD) was obtained on loan from CRDEC. This type of MS instrument combines simultaneous ion detection potential with CI and tandem MS capabilities. Two special

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR
83 APR edition may be used until exhausted.
All other editions are obsolete.

(19)

Curie-point pyrolysis inlets were built for the ITD and successfully tested with model polymers. The high sensitivity and dynamic range of the ITD were found to offer excellent prospects for biodetection purposes. Development of advanced chemical pattern recognition methods focussed on a nonsupervised method for mixture analysis, the so-called Variance Diagram (VARDIA) method, and a method for spectral time series analysis (VARDIA-S). Qualitative and quantitative VARDIA analysis results on standard mixtures of biopolymers (deoxyribonucleic acid, albumin, glycogen) as well as complex biological materials (wood, grass leaves) are reported. Finally, successfull use of VARDIA-S for "event detection" in a simulated time series of samples representing different mixtures of microorganisms (<u>Bacillus anthracis</u>, <u>Streptococcus</u>) and Dugway Soil is demonstrated.

(18)

microorganisms, <u>Streptooccus</u>, <u>Bacillus anthrax</u>, Dugway soil, intereferents, multivariate statistical analysis, factor analysis, chemical pattern recognition

LIST OF REPORTS

1. H.L.C. Meuzelaar, W. Windig, A.M. Harper, S.M. Huff, W.H. McClennen and J.M. Richards, "Pyrolysis Mass Spectrometry of Complex Organic Materials" Science, 226, (1984), 268-274.

2. S.M. Huff, J.M. Matsen, W. Windig and H.L.C. Meuzelaar, "Pyrolysis Mass Spectrometry of Bacteria From Infected Human Urine" Biomedical and Environmental Mass Spectrometry, 13, (1986), 277-286.

3. W. Windig, W.H. McClennen, H. Stolk and H.L.C. Meuzelaar, "Unsupervised Chemical Pattern Recognition in Complex Mass Spectra" Optical Engineering, 25, (1986), 117-122.

4. W. Windig, E. Jakab, J.M. Richards and H.L.C. Meuzelaar, "Self-Modeling Curve Resolution by Factor Analysis of a Continuous Series of Pyrolysis Mass Spectra" Anal. Chem., 59, (1987), 317-323.

5. W. Windig, W.H. McClennen and H.L.C. Meuzelaar, "Determination of Fractional Concentrations and Exact Component Spectra by Factor Analysis of Pyrolysis Mass Spectra of Mixtures" Chemometrics and Intelligent Laboratory Systems, In press.

6. W. Windig, T. Chakravarty, J.M. Richards and H.L.C. Meuzelaar, "Multivariate Analysis of Time-Resolved Mass Spectral Data" Accepted for publication in Anal. Chem. Act.

7. A.P. Snyder, J.H. Kremer, H.L.C. Meuzelaar and K. Taghizadeh, "Curie-point Pyrolysis Atmospheric Pressure Chemical Ionization Mass Spectrometry" Preliminary performance data for three polymers, Submitted to Anal. Chem. (1986).

8. W.H. McClennen, J.M. Richards and H.L.C. Meuzelaar, "Direct Qualitative Characterization of Solid Rubber Samples by Laser Pyrolysis Mass Spectrometry (Py-MS)" ASMS 34th Ann. on MS and Allied Topics, 817-818 pg, June 8-13 1986, Cinn, OH.

9. J.M. Richards, H.T. Stolk, W.H. McClennen and H.L.C. Meuzelaar, "Development and Testing of a Curie-point Pyrolysis Inlet for the Finnigan Ion Trap Detector" ASMS 34th Ann. on MS and Allied Topics, 1071 - 1072 pg, June 8-13 1986, Cinn, OH.

10. H.L.C. Meuzelaar and W. Windig, "Biological Pattern Recognition by Pyrolysis Mass Spectrometry and Multivariate Analysis" Proc. Second ARO biodetection workshop, July 13-15, (1982).

11. H.L.C. Meuzelaar and W. Windig, "Development of a Biodetection Method Based on Computerized Pyrolysis Mass Spectrometry", Proc. Third ARO biodetection workshop, April 25-26, (1984).

12. H.L.C. Meuzelaar and W. Windig, "New Mass Spectrometric Techniques for Biodection in Aerosols", Proc. 1985 Scientific Conference of Chemical Defense Research, Nov. 19-22, (1985).

I. STATEMENT OF PROBLEM

1. The Need for New Analytical Approaches

The development of suitable techniques for the detection and identification of biological agents is among the most challenging analytical problems faced by biomedical and physicochemical scientists today. In recent years, outbreaks of previously unrecognized infectious syndromes, e.g., Legionaires Disease [1], have highlighted the difficulties encountered in the identification of unknown infectious agents by highly specific, conventional microbiological techniques geared towards detection and identification of known organisms.

In comparison with agents of natural origin, the detection and identification problems associated with biological warfare agents are even more complex. Because of rapid developments in the field of genetic engineering during the past few years, the likelihood that hitherto unknown organisms or toxins will be encountered under biological warfare conditions is ever increasing. Moreover, under warfare conditions biological agents may be expected to be deliberatory or accidentally concealed among high background levels of other materials of biological origin.

Obviously, reliable detection and subsequent identification of biological warfare agents will require novel analytical approaches which combine high sensitivity to a broad range of organisms and toxins with sufficient specificity to distinguish between biological aerosols of different nature and origin.

2. A Universal Approach to Biological Pattern Recognition

Ideally, one would like to use a rapid, sensitive analytical technique with almost universal applicability to complex biological materials and capable of providing detailed information on the biochemical composition of any aerosol. Figure 1 provides a schematic illustration of an idealized analysis approach to biological pattern recognition.

Assuming that the composition of biological compounds in intentionally dispersed aerosols will generally be quite distinctive from aerosols produced by natural causes, e.g., forest fires, or unintentionally dispersed by war related events, e.g., explosions in swamp areas, then the approach outlined in Figure 1 should be able to distinguish between these different origins. The ability to establish the presence of intentionally dispersed biological aerosols with confidence is obviously of great military, medical and political significance, even if the precise identity of the biological agent involved would not be known immediately. General protective measures could be taken and decontamination and quarantine procedures started while collected aerosol samples could be investigated further in central analytical facilities. Fortunately, it is quite likely that the abovedescribed universal analytical approach will also provide important clues with regard to the nature and possible identity of the biological agents involved. Such clues would mainly be provided by the absence or presence of specific compounds or classes of compounds known to be characteristic for certain groups of organisms. For instance, the absence of nucleic acids would rule out the presence of any organisms where as the absence of muramic acid moieties would exclude the presence of bacteria [2], etc.
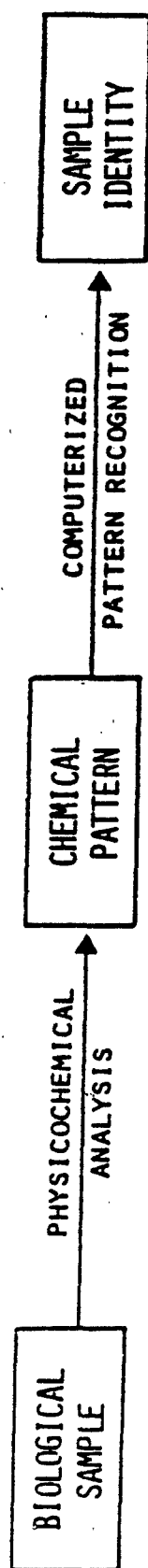
2.

Figure 1. Generalized approach to biological pattern recognition.

BIOLOGICAL SAMPLE

PHYSICOCHEMICAL ANALYSIS

CHEMICAL PATTERN

COMPUTERIZED PATTERN RECOGNITION

SAMPLE IDENTITY

3.

A preliminary example of a hierarchical aerosol identification scheme based on this approach is shown in Figure 2. It should be pointed out perhaps that hierarchical "decision tree" approaches are often less reliable than nonhierarchical multivariate classification techniques such as SIMCA, a multiclass principal component modelling techniques developed by Wold et al. [3]. Therefore, Figure 2 is only intended as an illustrative example of a universal approach to identification of biological aerosols based on recognition of specific classes of biochemical compounds.

high organic —no→ ash, dust, etc.

yes

DNA —no→ RNA —no→ protein —no→ amino sugars —no→ PNA's[1] —no→ olefins

yes (DNA) | yes (RNA) | yes (protein) | yes (amino sugars) | yes (PNA's) | yes (olefins)

RNA —no→ protein —no→ toxin? ←yes— (lipids)

smoke, char, etc. shale, plant odor, gasoline

yes (RNA) | yes (protein)

(protein)   pure virus   wood, grass

yes

cholesterol —no→ muramic acid —no→ lignin —no→ cellulose —no→ chitin —no→ ?

yes (lignin, to wood, grass)

yes (cholesterol) | yes (muramic acid) | | yes (cellulose) | yes (chitin)

(choline)   KDO[2] —no→ gram positive bacteria   chlorophyll —no→ fungi

yes (choline) | yes (KDO) | | yes (chlorophyll)

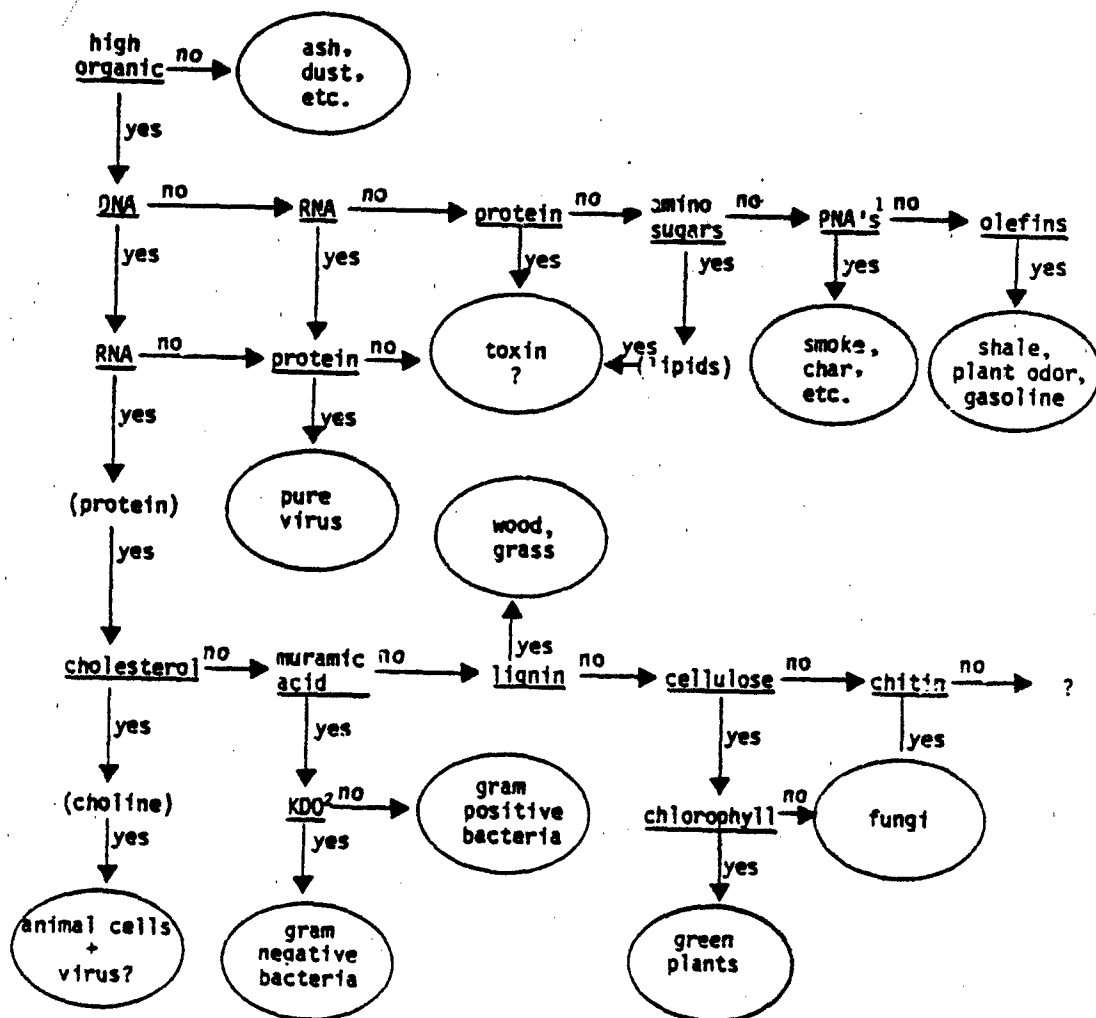animal cells + virus?   gram negative bacteria   green plants

Figure 2. Tentative, hierarchical "decision tree" for classification of biological aerosols (questions-underlined; check questions-hyphenated; conclusions-encircled). [1]Polynuclear aromatic compounds. [2]Ketodeoxyoctonic acid.

5.

II. SELECTED APPROACH

1. Selection of a Suitable Analytical Technique

Only relatively few analytical methods are potentially capable of distinguishing between the various classes of biochemical compounds shown in Figure 2. Among these are chromatographic methods such as gas chromatography (GC) or liquid chromatography (LC) and spectrometric techniques such as infrared spectrometry (IR) or mass spectrometry (MS) as well as various combinations of these methods (e.g., GC/MS, GC/IR, GC/IR/MS, LC/MS, MS/MS, GC/MS/MS, etc.). The latter group of so-called "hyphenated" methods [4] includes many of the most powerful analytical methods for complex organic materials presently available.

Of the basic methods, only IR can be applied more or less directly to all biological materials. LC will often require "solubilization" of insoluble materials (e.g., by hydrolysis) whereas GC and MS require "volatilization" of nonvolatile materials (e.g., by pyrolysis). Other important differences exist between these techniques with regard to sensitivity, specificity, speed and computer compatibility.

2. Potential Applicability of Pyrolysis Mass Spectrometry

Based on our experience with different analytical methods for complex biological materials, pyrolysis MS appears to be superior to GC, LC and IR with regard to sensitivity, speed and computer compatibility. Figure 3 provides a schematic overview of the fully automated and computerized pyrolysis MS system developed at the FOM Institute for Atomic and Molecular Physics in Amsterdam [5] and capable of analyzing up to 36
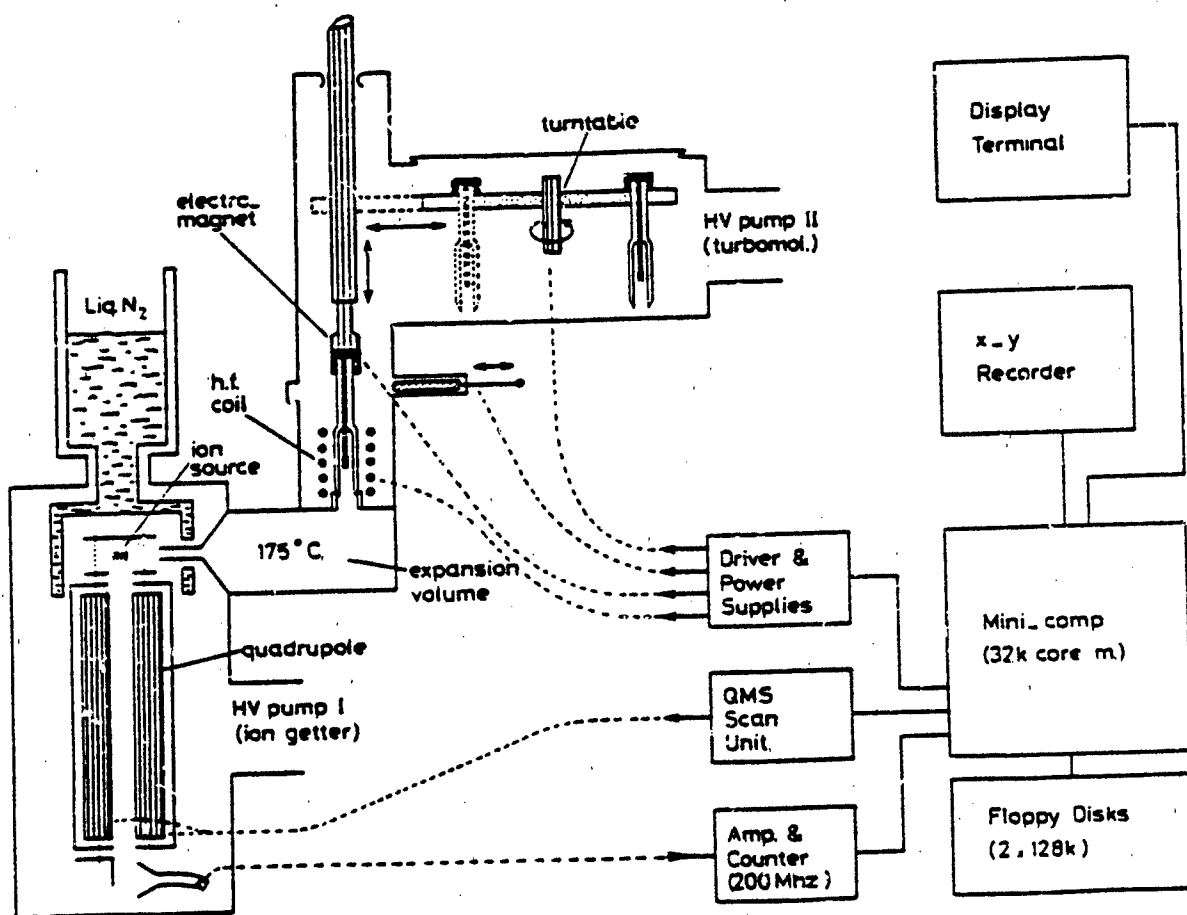
6.

Figure 3. Schematic diagram of an automated Curie-point Py-MS system. Note the 36-position turntable the liquid nitrogen-cooled screen completely the ion source and the high-speed ion counting channel. Typical time needed to analyze one batch of 36 samples is 1 hour.

7.

samples per hour. A somewhat modified pyrolysis MS system with automated sample inlet has been constructed in our laboratory at the University of Utah [6].

Pyrolysis MS systems such as shown in Figure 3 have been successfully applied to the char icterization of a broad range of biological compounds, including bacteria, viruses, mammalian cells and tissues, body fluids, biopolymers, plant tissues, humic compounds coals and shales [7]. Furthermore, Voorhees et al. [8] have reported the characterization of air particulates of biological origin. An overview of these applications is given in a monograph by Meuzelaar, Haverkamp and Hileman [7] which also contains an atlas of reference spectra from biological compounds.

Although hyphenated methods such as GC/MS and MS/MS enable a far more detailed ..mical analysis than single stage MS, single stage pyro.ysis MS if often quite successful in identifying structural moieties, individual compounds or compound classes in complex biological materials [9]. This is largely due to the exceptional degree of computer compatibility of the pyrolysis MS technique.

3. Selected Pyrolysis Mass Spectrometry Approach

In the original proposal to ARO the instrumental approach selected to achieve time-resolved Py-MS conditions consisted of CO_2 laser pyrolysis in combination with a miniature Mattauch Herzog MS system with simultaneous ion detection capabilities by means of electro-optical ion detection. This so-called MMH-EOID system was expected to be obtained from Nuclide Inc. (State College, PA) with DOD/University Instrumentation funding (Grant No. DAAG29-83-G-0070) shortly after the start of our ARO grant. Unfortunately, Nuclide Inc. proved unable to deliver the

o

instrument due to financial problems which resulted in a bankruptcy filing under Chapter 11 in 1985. As soon as it became clear that Nuclide Inc. would not be able to deliver the system on time, the P.I.'s started to look for alternative instrumental solutions. The Finnigan Ion Trap Detector (ITD) was found to be a promising candidate system and in September 1985 a Finnigan ITD was obtained on loan from CRDEC (Dr. Peter Snyder). After suitable modifications, the ITD system was demonstrated to be a highly sensitive instrument for Py-MS [10] and Py-GC/MS [11] studies, enabling Py-MS analysis of polymer samples in the low nanogram range as well as identification of GC peaks in the low picogram range [12].

Recently, DOD permission was obtained to use some remaining funds from the DOD/University Instrumentation grant to purchase an infrared microscope accessory capable of focussing a CO_2 laser beam in the vacuum system of the ITD, in order to pursue time-resolved CO_2 laser pyrolysis experiments with the ITD technique. These experiments will be carried out in 1987 with expected continuation funding from CRDEC.

In conclusion, although major delays were experienced the development of a suitable CO_2 laser MS system due to bankruptcy of the manufacturer of the MS instrument, this part of the project is now scheduled for continuation in 1987 with CRDEC funding. As a result of the delays in the laser MS equipment tasks, the main emphasis of the work carried out under the ARO contract reported here shifted to the development of advanced chemical pattern recognition methods for pyrolysis mass spectra of biological materials, as will be discussed in the next few paragraphs.

4. The Role of Pattern Recognition Methods

The key challenge in the Py-MS approaches described above is how to retrieve the chemical information of the samples as given in Figure 2. The logical choice for assigning a certain type of aerosol to a spectrum seems to be a library search approach. A library with all possible spectra for biodetection purposes is not feasible. Since the spectra often represent complex mixtures, data reduction by mixture analysis is a promising approach. The extreme complexity of the spectra that will be obtained in most biodetection situations and the lack of suitable reference spectra make mixture analysis on a single spectrum impossible. Mixture analysis approaches that do not require reference spectra are available [13,14]. Among these approaches the variance diagram technique has been developed especially for complex spectra, such as obtained by pyrolysis mass spectrometry [15]. These type of techniques, however, require sets of spectra. Consequently, the so-called "unsupervised" mixture analysis methods can be applied if a set of spectra from a single sample, analyzed under different experimental conditions, is available.

One of the most promising methods is time-resolved analysis. During desorption and/or pyrolysis the composition of the product mixture will change, as a result of different volatilization/pyrolysis behavior of the components in the sample. The spectra of these components can be extracted mathematically and assigned to a certain class of compounds using a relatively small library.

The mixture data obtained can be used in automated and/or expert systems to assign the sample to a certain class of aerosols or microorganisms.

10.

III. WORK PERFORMED

1. CO$_2$ Laser Py-MS

As a model for laser pyrolysis studies, rubber samples were chosen. The reason is that with the proper selection of the analysis conditions, separate spectra of the relatively volatile additives and the actual polymer can be obtained. The use of mass spectrometry in thermal and solvent extraction methods for analysis of polymers and their additives has recently been reviewed [16,17]. Previously published work at the Biomaterials Profiling Center has focussed on the determination of the polymers in various rubber compounds by Curie-point Py-MS [18,19]. Although this technique was also a "direct" method applied to the whole sample, it involved a significant effort to grind the vulcanized rubbers into very fine particles for solvent suspensions. Lasers have previously been used for polymer pyrolysis with electron ionization MS, but without detection of additives [20,21]. Lasers are also used as the combined pyrolysis and ionization source in LAMMA experiments where polymer analysis has given extensive fragmentation with poor molecular ion yield and again virtually no additive information [22,23].

A. <u>Laser Py-MS Technique</u> - The mass spectrometer consists of a basic electron ionization quadrupole instrument (Extranuclear 5000-1) in a modified vacuum housing with a cryopump, solids probe and sodium chloride IR window. Low electron energies (typically 15 eV) were used to minimize ion fragmentation and simplify spectra. The carbon dioxide laser was a tunable continuous wave model capable of 5 watt operation on several lines. Most of the spectra were obtained on a strong 10.21 um line. The

11.

laser beam was focussed from its original 1.6 mm diameter to a spot ca. 0.5 mm diameter on the sample 26 cm from the 30.5 cm focal length lens. The relative positions of the laser beam, sample and ion source are shown schematically in Figure 4 to indicate how products were volatilized/-pyrolyzed directly from the sample into the ionization region.

The laser Py-MS experiments were run by pulsing the laser beam with an electro-mechanical shutter. The shutter was opened for 50 to 250 ms at 1 s intervals for 9 consecutive pulses. The sample was moved across the beam path between pulses to permit exposure of fresh sample to each shot. During this time the mass spectrometer was repetitively scanned at 1000 amu/s through the desired mass range 100 to 150 times and the data signal was averaged to give a single spectrum. The multiple laser pulses increased sensitivity and averaged the variations which might occur in a single exposure.

B. <u>Qualitative Laser Py-MS Analysis</u> - The spectra in Figures 5a and b were obtained from a preliminary setup in which the samples were more than 2 cm from the center of a very open ion source. The laser pyrolysis spectra of the SBR sample (Figure 5a) and a SBR and natural rubber (NR) blend (Figure 5b) show predominantly the fragment and molecular ions of monomers and small oligomers from the thermaily degraded polymers. The spectra are very similar to those obtained by Curie-point Py-MS [18,19] where most of the less volatile additives were either condensed out before reaching the ion source or were simply too low in relative concentration to be seen. One additive which can be observed is the sulfur, as indicated by peaks at m/z 34 and H_2S and m/z 48 and 64 from SO_2 (and possibly S_2 for 64).
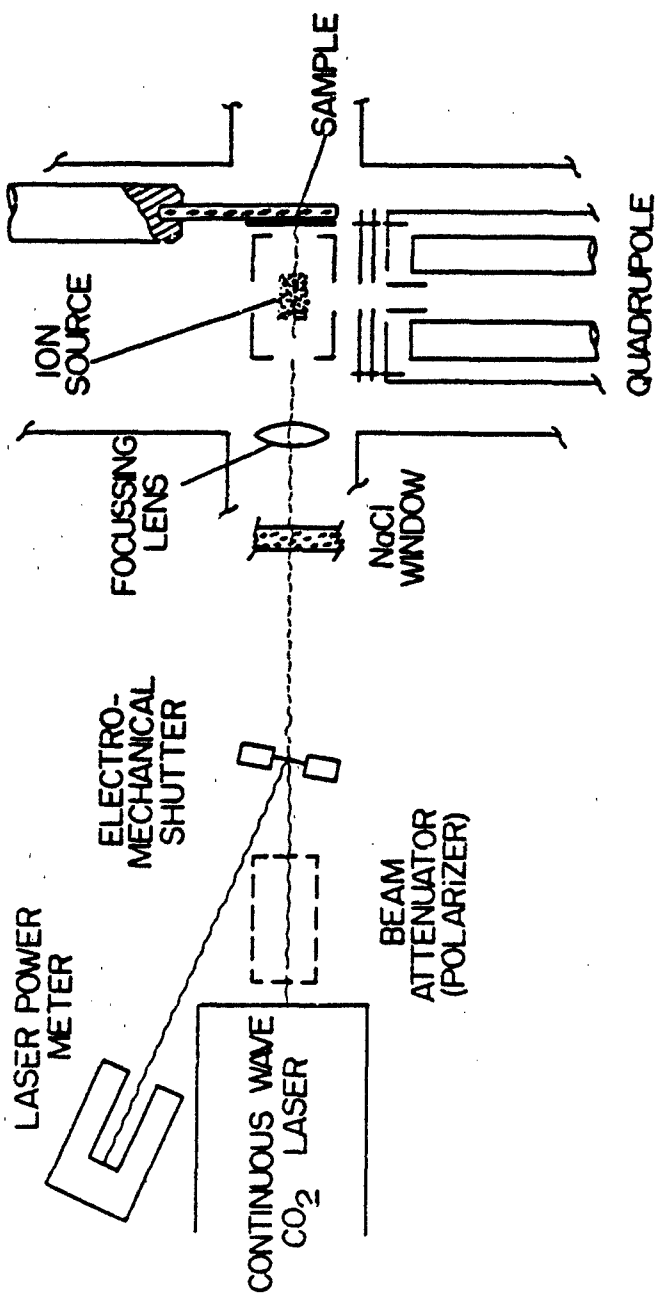
12.

Figure 4. Schematic diagram of the experimental setup for direct pyrolysis of solid rubber samples next to the ion source. Note the polarizer attenuated beam is reflected off the shutter and constantly read by the power meter between pyrolysis experiments. The beam scanner is able to move the focussed beam through a 2.5 mm. diameter area on the sample.
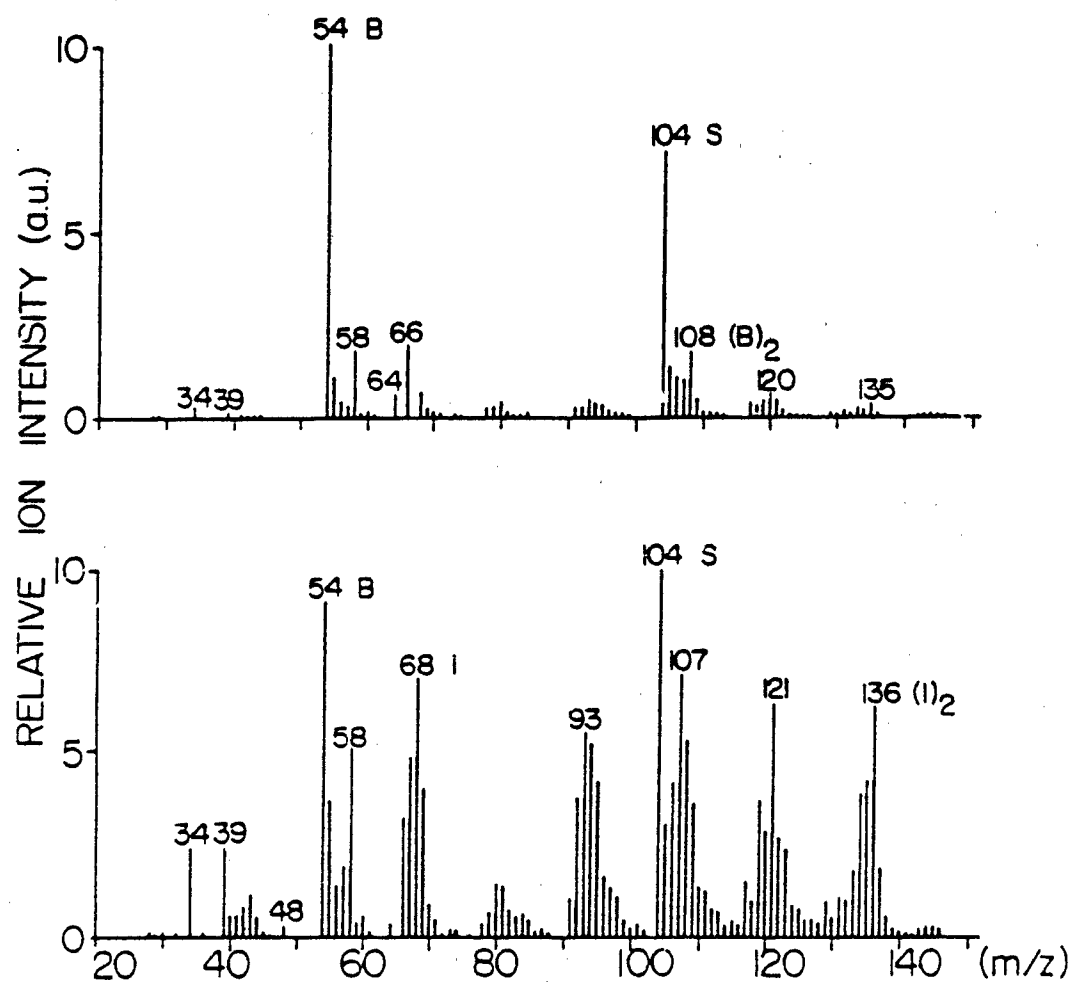
13

Figure 5.   Laser pyrolysis mass spectra of vulcanized SBR (a) and SBR-NR
            (b) rubber compounds. Spectra consist primarily of fragment
            and molecular ions from monomers and dimers of styrene (S)
            butadiene (B) and isoprene (I). For these preliminary samples
            most of the additives condensed out before reaching center of
            ionizer.

Figure 6 shows a Curie-point (volatilization) mass spectrum of the dichloromethane extract from the SBR compound. The molecular and fragment ions from several additives are indicated with the most intense signal from the HPPD antiozonant molecular ion at m/z 268 and a fragment ion cluster around m/z 211 from the loss of C_4H_9. The portion of the spectrum above m/z 300 is expanded x5 to indicate the TMDQ dimer ions (m/z 346 and 331) and the DODPA peak at m/z 393. The accelerator decomposes during vulcanization and is thus observed only as two of its decomposition products; mercaptobenzothiazole at m/z 167 and benzothiazole at m/z 135. The m/z 135 ion could also be the fragment ion with the structure shown in Figure 6 from a t-octylphenol-formaldehyde resin pyrolysis product with molecular weight 206. However, the low voltage ratio of m/z 206 to m/z 135 is normally much larger than that observed here. The small peaks at m/z 129, 185 and 284, which were more intense in other samples, are due to stearic acid. The ion at m/z 256 could represent unvulcanized sulfur S_8 or the impurity palimitic acid in stearic acid [17].

The laser pyrolysis spectrum in Figure 7 was obtained with the sample closer to a less open ion source which directs more of the pyrolysis/-volatilization products through the ionization region. This analysis was also run with a weaker laser beam power (ca. 3 to 4 watts) and longer pulse times (150 ms). Additive ions as indicated in Figure 6 are clearly visible with relatively weak signals from the SBR monomers at m/z 54 and 104. In addition to these peaks, many other fragment ions can be seen resulting largely from the higher molecular weight (MW 400 to 600) hydrocarbon processing oil.
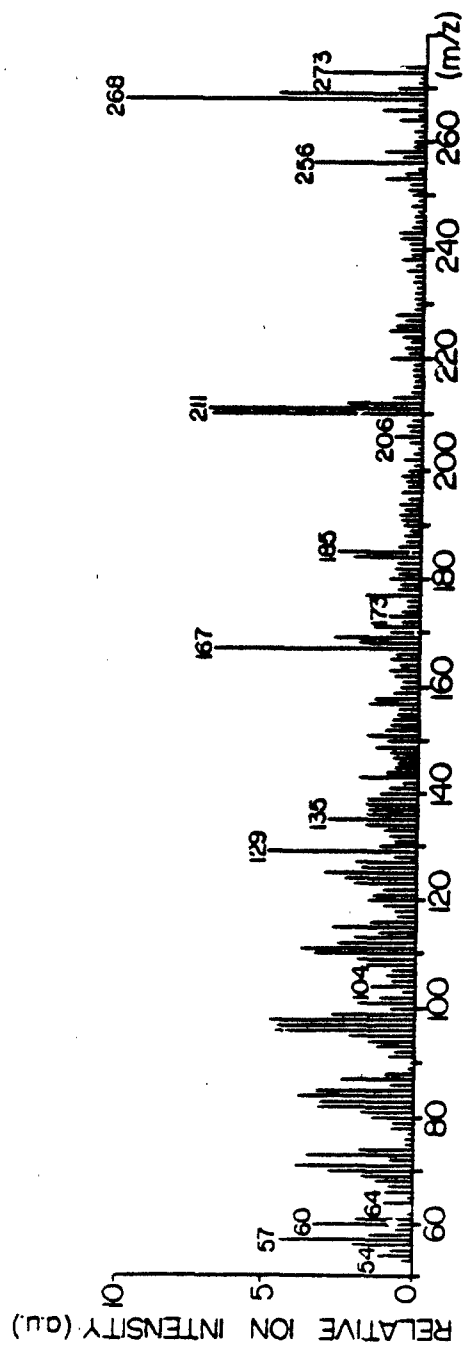
15.

Figure 6. Curie-point vaporization mass spectrum of the dichloromethane extract from the SBR sample. Several of the rubber additives are indicated for comparison to peaks in Figure 7.

16.

Figure 7. Laser pyrolysis/volatilization mass spectrum of the SBR sample at lower laser power with the ion suorce modified from the configuration used in Figure 4. Note obvious presence of additive ions as identified in Figure 6.

17.

Figures 8 a-c show triplicate laser pyrolysis spectra at 4.5 Watts and 100 ms pulse times of the SBR compound. Note the very good reproducibility of the spectra which show much stronger monomer and oligomer molecular ions containing styrene and butadiene units (labeled as S and B respectively). Also note that, other than the abundant fragment ions which may be from the oil as well as polymer pyrolysis, the spectra are relatively simple with very little evidence of the additive ions seen in Figures 6 and 7.

C. <u>Quantitative Laser Py-MS Analysis</u> - Butadiene rubber (BR) and natural rubber (NR) samples and three mixtures of the two were pyrolyzed using the laser beam scanning technique (see Figure 4). Factor and discriminant analysis were performed on the replicate spectra of these five samples and the scores of the main discriminant function are plotted versus their relative composition (Figure 9). The non-linearity of the bivariate plots is apparently due largely to the lower decomposition temperature for the polyisoprene (NR) than the polybutadiene.

2. <u>Curie-point Py-MS with Ion Trap Detector</u>

Though the Finnigan Ion Trap Detector (ITD) is designed and used primarily as a detector system for gas chromatography, it has several features which make it attractive for development as a stand-alone MS, e.g., for pyrolysis mass spectrometry (Py-MS) studies. Among attractive features are its ease of operation, high scanning speed, ability to detect all ions produced in a single, transient event (e.g., flash pyrolysis), and possibilities for CI and tandem MS (MS/MS) operation
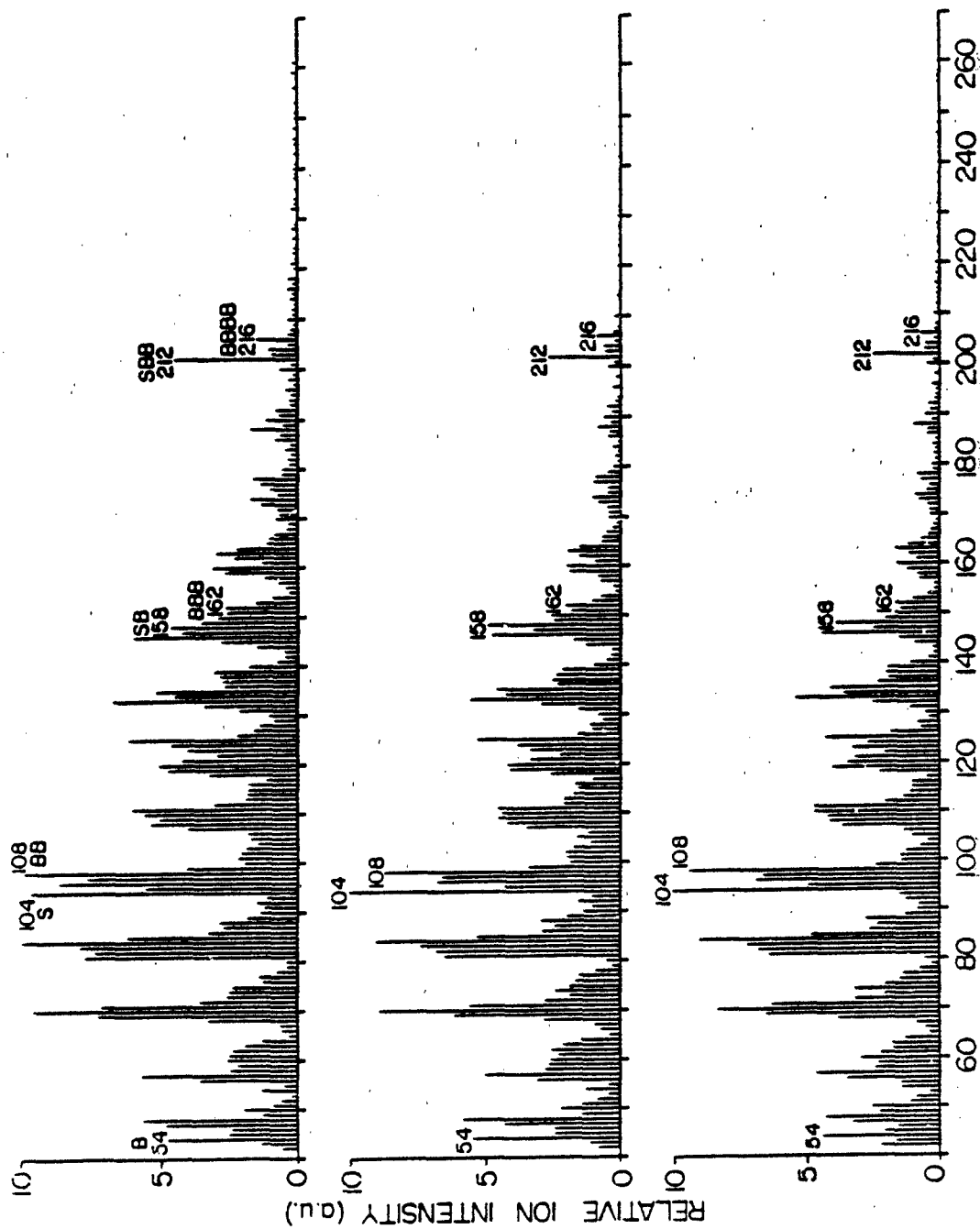
18.

Figure 8.   Laser pyrolysis/volatilization mass spectra of the SBR sample with the instrument as in Figure 4 except for higher laser power and shorter pulse duration. Spectra now composed mainly of polymer fragment and molecular ions with many from less volatile products than could be observed in the instrument configured for Figures 5a and b. Note the resonable reproducibility for these three separate analyses.
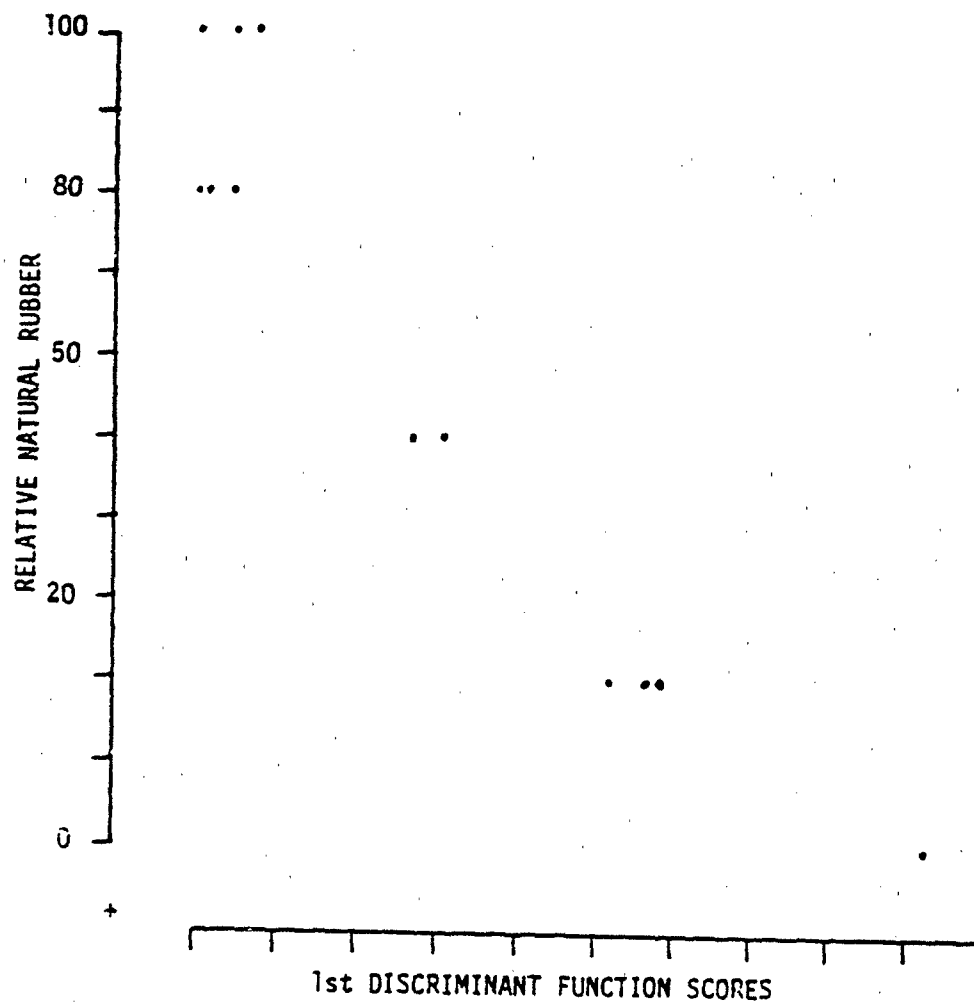
Figure 9.   Plot of discriminant function scores versus % NR composition.
Two or three spectra were run for each of the 5 samples.   The
lack of discrimination of the 80 and 100% NR samples seems to
be predominantly due to the lower decomposition temperature of
the NR compared to the BR.

20.

modes. Moreover, the ITD's small size, low price and its user-friendly operating software make it well suited for environmental monitoring (viz., pollutants or aerosols), industrial process monitoring, biomedical and forensic applications. By interfacing a Curie-point pyrolysis inlet to the ITD, a bench top Py-MS system can be developed which can analyze nonvolatile organic materials in several different Py-MS modes including Py-EIMS, Py-CIMS and Py-MS/MS.

Initial experiments involved connecting a Curie-point pyrolysis inlet directly to the heated, open-split GC interface of the ITD. Pyrolysis in this configuration was carried out in an inert (He) atmosphere, and the products of pyrolysis were swept into the ITD by the He carrier gas. Analyses of synthetic polymers (e.g., poly(methlmethacrylate)) gave spectra with characteristic molecular and fragment ions of the respective monomers and other pyrolysis products. The time information about the relative thermal stabilities of these polymers was also preserved. However, the elution of the pyrolysis products into the ITD was accompanied by extensive peak broadening and tailing. This resulted in the pyrolysis products from a ca. 2 s heating profile (temperature rise time ca. 2 s, total heating time 10 s, equilibrium temperature of the filament 610°C) giving a total ion current curve that was approximately 12 s wide at half maximum. Analysis of biopolymers (deoxyribonucleic acid DNA, bovine serum albumin BSA, and glycogen), however, resulted in the loss of polar pyrolysis products. These losses were apparently due to condensation in the transfer line or the pyrolysis inlet. Moreover, the long residence time of the pryolyzate in the hot (250°C) transfer line may have led to undesirable rearrangements or secondary reactions among the pyrolysis products [24].

To overcome these difficulties, the pyrolysis reactor was modified in order to connect directly to the ITD via a 3 inch glass lined tube. This configuration is shown in Figure 10 (Py-ITD). Analysis of a polymer mixture of nitrocellulose, poly(styrene) and poly(-methylstyrene) showed significant reduction in peak broadening and tailing and preservation of time (temperature) resolved information. The direct interface also resulted in improved sensitivity as was seen in the analysis of 5 ng of a poly(ether urethane urea) (Biomer [25]) under standard ionization conditions. Analysis of Biomer yielded spectra which contain peaks characteristic of the diisocyanate and polyol components [26] and refect the temperature-dependent nature of the pyrolysis process (Figure 11). Analysis of DNA, BSA, and glycogen produced spectra which compare favorably with low voltage Py-EIMS and Py-APCIMS. These differences are apparently due to the different ionization and mass analysis techniques used in the ITD and are not clearly understood at this time.

Future studies on this system will investigate the sensitivity and reproducibility of the system using Py-EIMS and Py-CIMS. Also the information content of the spectra will be studied.

3. Ion Trap Dynamic Range and Sensitivity Testing

We evaluated the new Automatic Gain Control (AGC) ITD software which provides automatic variation of ionization time [12]. AGC maximizes sensitivity for low levels of analyte, and prevents saturation of the ion trap at high levels of analyte. The results are impressive, with full scan EI mass spectra easily obtained at .ow picogram levels, while linear response is maintained up to low microgram levels.
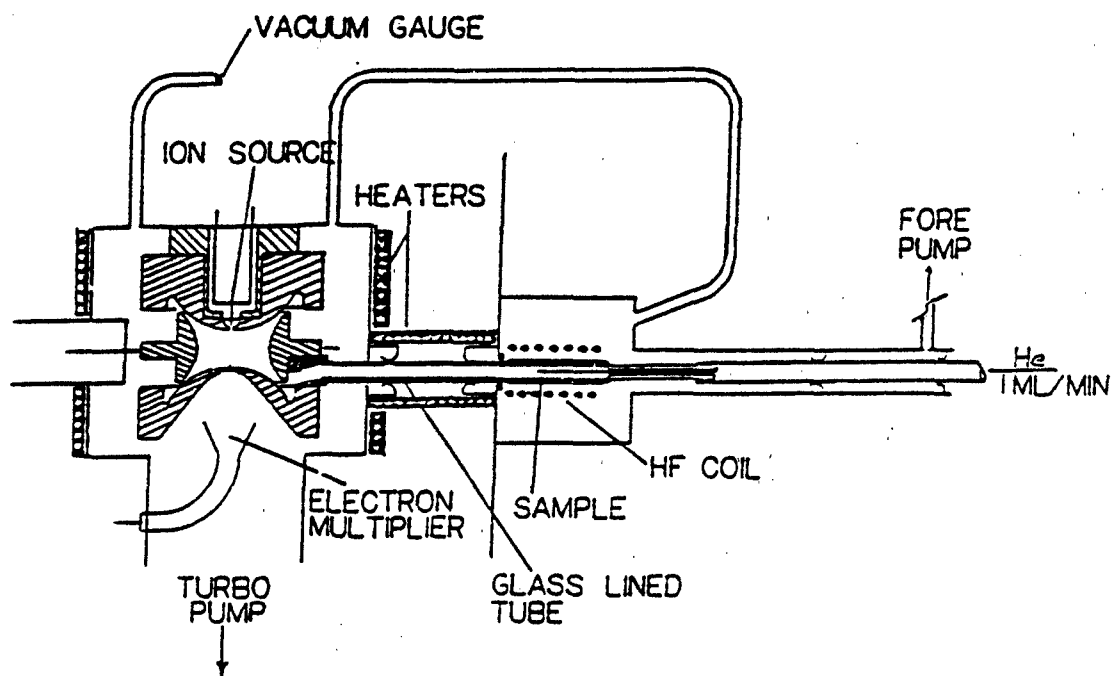
22.

Figure 10. Schematic of the Curie-point pyrolysis Ion Trap Detector system
(Py-ITD). Sample is coated onto the tip of a ferromagnetic
filament which is inductively heated to its Curie-point by the
high frequency coil. Helium is introduced through the sample
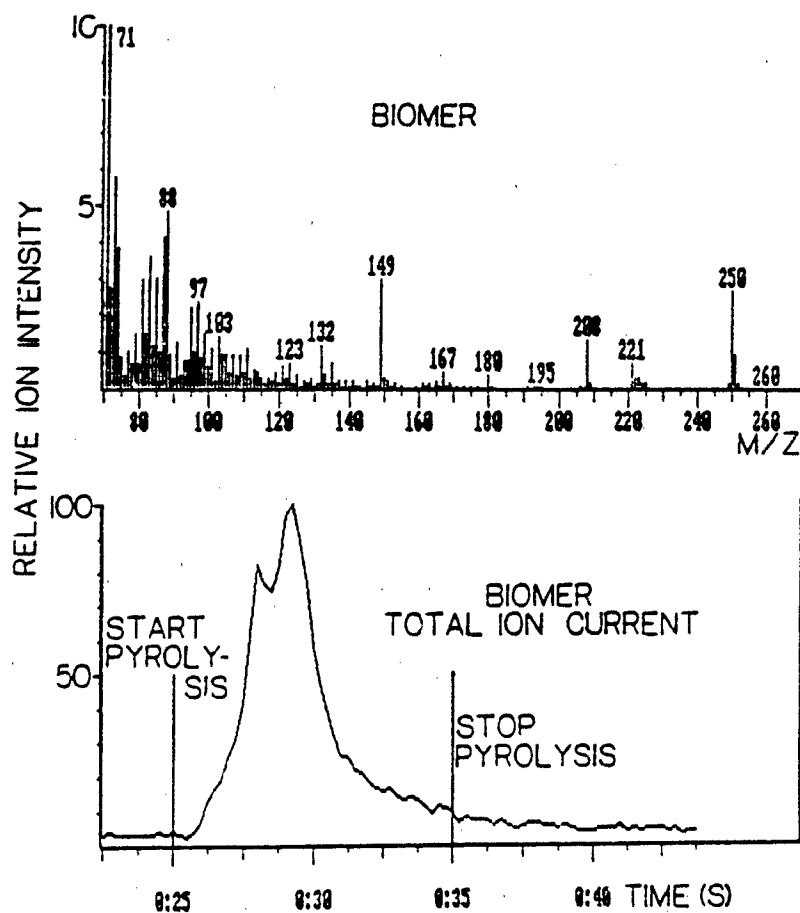probe at a rate of 1 ml/min.

23.

Figure 11. Spectrum and total ion current curve from the analysis of 5 ng of Biomer. In the mass spectrum, note the peaks which are characteristic of the diisocyanate (m/z 208, 221, and 250) and polyol (m/z 71 and 73) components. The peaks which are characteristic of the diamine chain extender (m/z 86) and the polyol dimer fragment (m/z 143 and 145).

24.

The dynamic range of any device which converts analytes into charged particles (e.g., flame ionization detectors or mass spectrometers) will be limited by the onset of space charge. These space-charge effects lead to saturation of instrument response when the density of charged particles (ions) rises so high that ion-ion repulsions become significant. These effects are especially evident in devices which trap or store ions. In the ion trap, these effects are manifest when too large an analyte concentration leads to loss of mass resolution and degradation of mass spectral quality (e.g., self-chemical ionization to form $(M + 1)^+$ ions).

The new ITD AGC software provides a solution to this problem. Whereas earlier versions of ITD software have used a fixed (1 ms) ionization time, the new version automatically selects an ionization time from 0.078 ms to 25 ms (a range of 320) for each microscan, depending on the amount of analyte in the trap. For low levels (e.g., for baseline or small GC peaks) a maximum ionization time of 25 ms is selected, offering an improvement of approximately 25 times in the sensitivity over operation without AGC. As the amount of analyte increases, the ionization time is automatically decreased in order to prevent overfilling the trap with ions. The ion signals in each microscan are automatically scaled to correct for the variation in ionization time. The new scan function is shown in Figure 12. A short "prescan", consisting of a 0.2 ms ionization provides an estimate of the number of ions formed in the trap. This value is then used to calculate the optimum ionization time for the actual mass scan which follows.

In order to evaluate the performance of this new software, a set of serial dilutions of exo-tetrahydrodicyclopentadiene the major component of
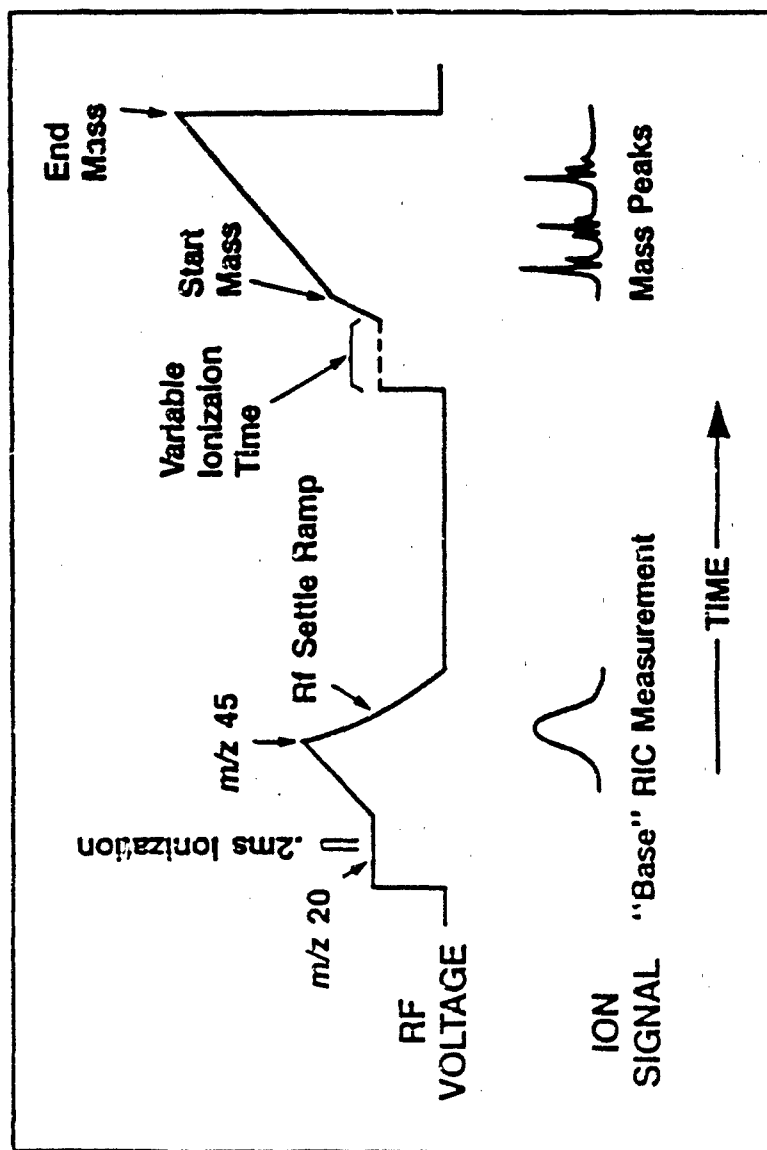
25.

Figure 12. New ITD AGC scan function with automatic variation of
ionization time.

26.

JP-10 jet fuel in n-hexane were analyzed by GC/EIMS on the ITD. This component is a tricyclodecane of formula $C_{10}H_{16}$ and molecular weight 136. Tricyclodecane 1 was analyzed with a 15 m x 0.32 mm ID, 0.25 m film thickness, DB-5 open tubular column (programmed from 30°C to 130°C at 20°C min^{-1} 4.4 psig inlet pressure, 1 ul on-column injections), scanning from m/z 40 to 160. Standard solutions ranged form 15 pg l^{-1} to 1.5 ug l^{-1} for a solvent blank and the 15 pg solution are compared in Figures 13 and 14. The tricyclodecane 1 peak at a retention time of 4:20 is clear at the 15 pg level in the recon- structed ion chromatogram (RIC), as well as in the mass chromatograms for the molecular ion (m/z 136) and the most abundant fragment ion (m/z 67). The background-subtracted mass spectrum for the 15 pg GC peak is shown in Figure 15a. This spectrum compares favorably with that obtained for 150 ng (10,000) times more analyte), shown in Figure 15b.

The clearest indication of the superior perfomance of the AGC software is the linear dynamic range indicated in the calibration curves for tri- cyclodecane 1 shown in Figure 16. Over 6 decades of concentration, the calibration curves for the RIC, m/z 67, and m/z 136 (M^+) show excellent agreement and linearity (slope of the log-log plots = 1.0.). The decreases in slope in the picogram range are those typically observed in all GC/MS calibration curves just above the detection limit.

4. Aerosol Precipitation Experiments

For field sampling purposes, aerosol precipitation was studied. One of the first studies was with a homebuilt system, which could be used for precipitation on wires or quartz slides, see Figure 17. After successful
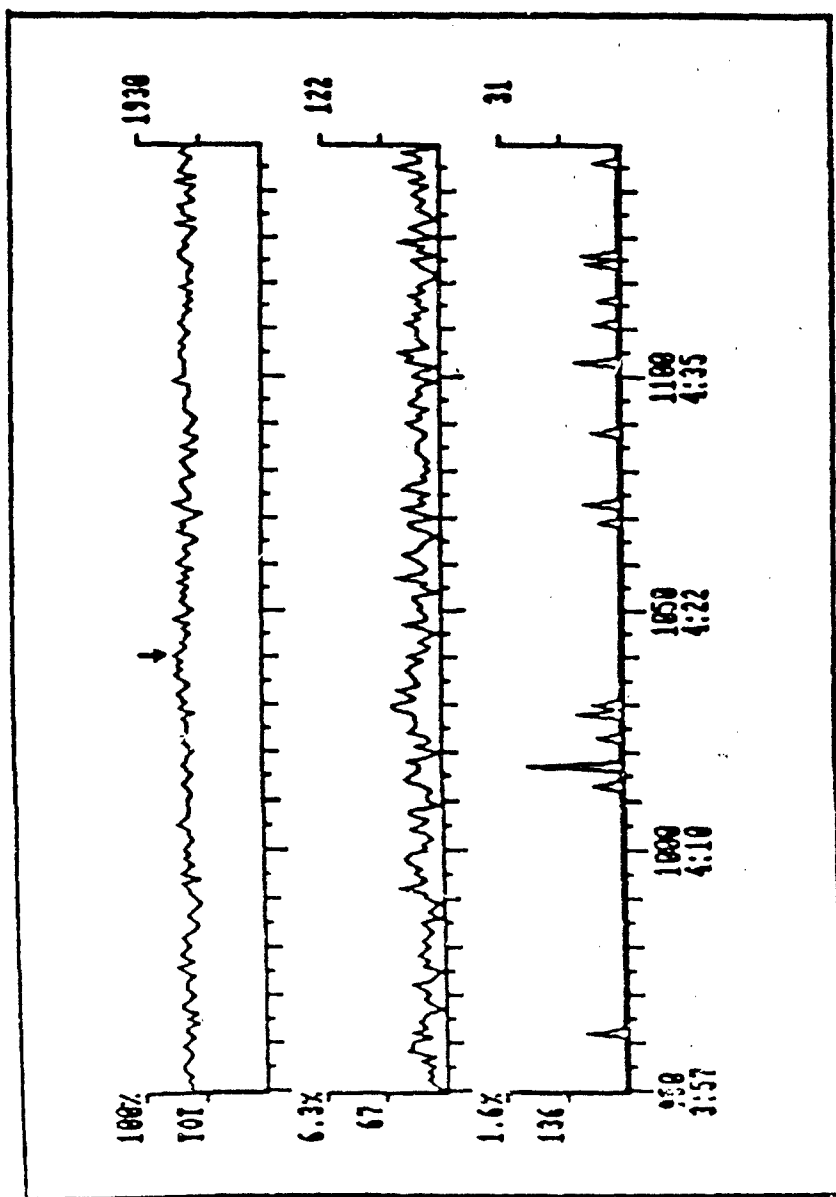
Figure 13. Portion of the chromatogram for a solvent blank. The arrow shows the retention time for tricyclodecane 1.
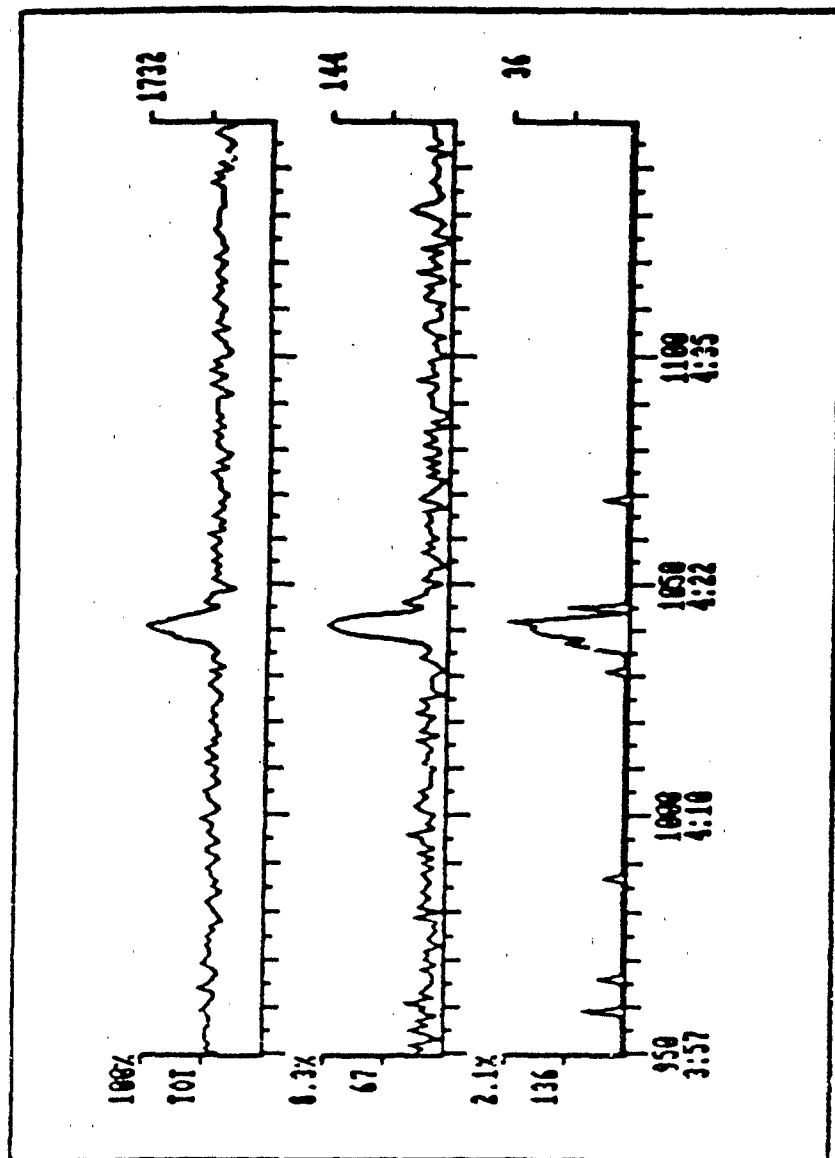
28.

Figure 14. Portion of the chromatogram for the lowest level standard
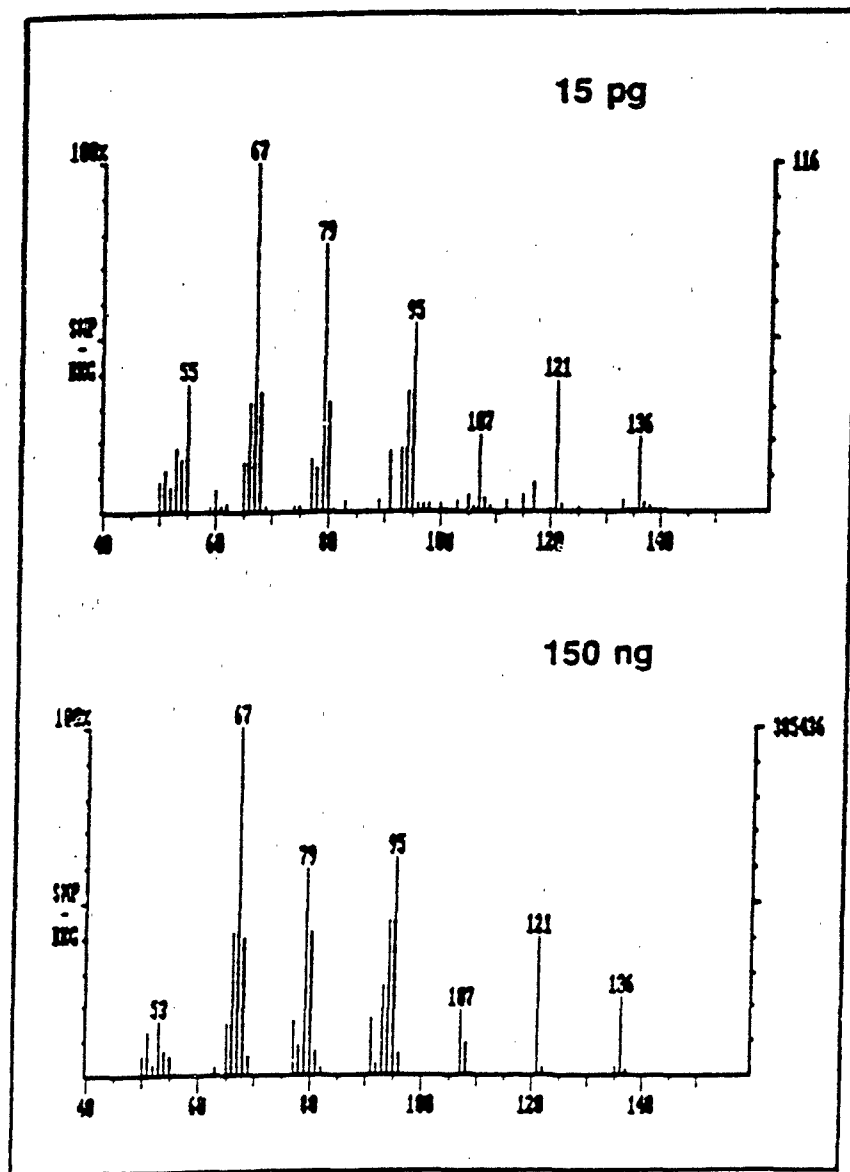solution (15 pg of tricyclodecane 1 injected).

29.

Figure 15. Background-subtracted mass spectra from GC peaks of
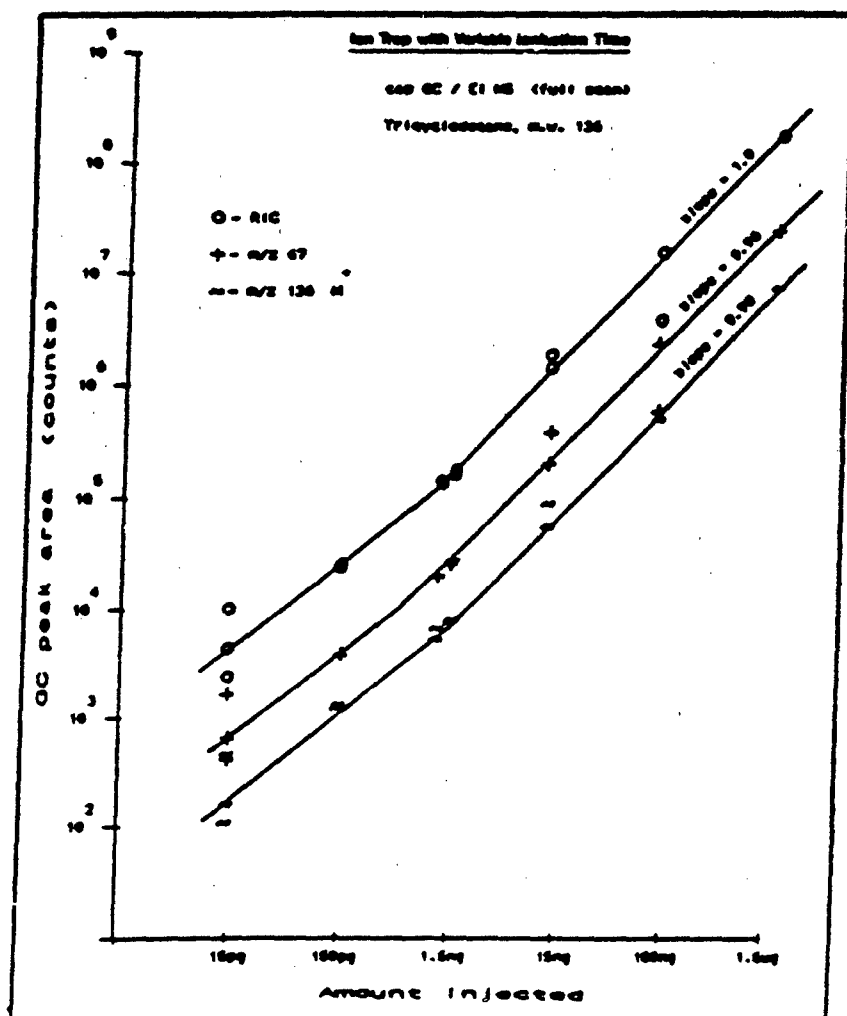tricyclodecane 1: (a) 15 pg; (b) 150 ng.

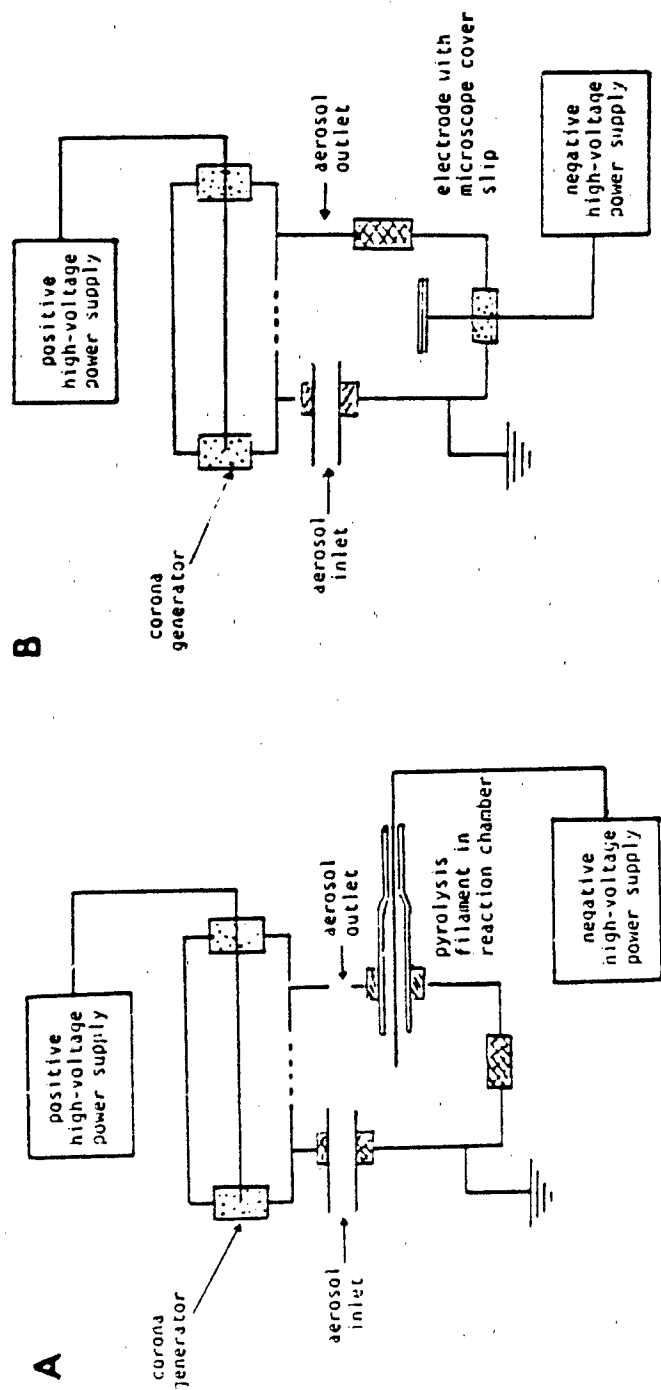Figure 16. Calibration curves for tricyclodecane 1.

Figure 17. A positive corona is generated in a shielded system. In this way direct contact of the aerosol with the corona is avoided. The positive ions generated by the corona are drawn out of the generator towards the negative electrode. The negative electrode is either a pyrolysis wire, mounted on a reaction tube (a), or a metal plate with a microscope cover slip on it (b). The positive ions are attached to the aerosol particles, resulting in a coating of the particles on the pyrolysis wire or the microscope cover slide.

feasibility studies with this approach, commercially available instruments (TSI 3100) and custom built instruments were bought for aerosol generation and precipitation. Successful precipitation on a Curie point wire followed by analysis by Py-MS was achieved. Currently, the aerosol sampling is being tested under field conditions.

5. Development of Advanced Chemical Pattern Recognition Methods

The techniques developed for the evaluation of the data are all based factor analysis based techniques followed by the VARDIA technique. Therefore, an overview of this technique will be given with examples for regular and time resolved data. The last part consists of an adaptation of the VARDIA technique for time-resolved data, which basically is an approach that can be done in a more automated way and can be used as a monitoring tool.

A. **The Variance Diagram Technique** - The Variance Diagram technique needs to be preceeded by multivariate analysis: i.e. factor analysis and/or discriminant analysis.

Factor analysis (principal component analysis) applied to a data set results in independent linear combinations of the original variables according to the equation:

$$F_j = \alpha_{1j}Z_1 + \alpha_{2j}Z_2 + \ldots \alpha_{2j}Z_n, \tag{1}$$

where F_j = factor j and α_{1j} = the loading (= correlation coefficient) of variable Z_i with F_j [27,28].

This generally results in a substantial data reduction, due to correlated behavior of the mass variables. In order to compare the linear combination with the orginal data, the loadings are multiplied by the standard deviation of the mass variable involved. The intensities obtained can be represented in the form of bar plots ("factor spectra" [29]). Since a limited number (e.g., 100) variables have to be selected for multivariate analysis due to limitations of the computer software and/or hardware, chemical interpretation of the factor spectra may be severely hindered. In order to involve the other mass varables in the factor spectrum, their contribution is determined by calculation of the covariance of the variables with the standardized scores. The factor scores are the relative contribution of the orginal spectra in the facots. The scores are standardized; i.e., the average intansity equals 0 and the standard deviation is 1.

Discriminant analysis is a related technique. The independent linear combinaticn resulting from discriminant analysis, however, describe the maximum of the ratio of between-group to within-group variance [27,28,30] where the between-group variance is the differencebetween groups of data (in our applications a group consists of replicate analysis of the same sample) and the within-group variance represents the differences within the groups ("noise"). The results are presented in the form of discriminant spectra and scores, similar to those resulting from factor analysis. After factor analysis, the VARDIA technique [15] can be applied on the data. The VARDIA method shows the strength of the correlated behavior of the mass variables in all directions in two-dimensional factor or discriminant subspaces. The equation for this procedure is:

$$\text{var}(W = \beta)_\gamma = \sum_{i=1}^{n} a_i^2, \quad \text{for } a_i \geq (\alpha_{i1}^2 + \alpha_{i2}^2)^{1/2} \cos \frac{\beta}{2}, \qquad (2)$$

where $a_i = \alpha_{i1}\cos\gamma + \alpha_{i2}\sin\gamma$ and var$(W = \beta)_\gamma$ is the variance i the direction in the space at an angle of γ degrees with the first factor (or the discrimiant function) using a window W of β degrees. Furthermore, a_i is the loading of mass variable i on the rotated function, n is the total number of mass variables, and α_{ij} are the loadings of the mass variables in the unrotated factor j. Expressed in terms that are easier to visualize, expression (2) accomplishes the following task: in a two-dimensional system the sum of the squares of the lenghts of all mass axes present in a pie-shaped window of β degrees (generally $10°$ or $20°$) is calculated while the window "scans" the whole two-dimensional space in discrete steps (generally $10°$).

8. <u>Examples of Variance Diagram Applications</u> - As an example, the results obtained on a data set of grass leaves are given [31]. Because of the complex biochemical composition of recent plant materials, spectra of the grass leaves show a complex pattern, as can be seen in Figure 18. The diagram after normalization correction (Figure 19) reveals a complex pattern between $60°$ and $140°$ with three local maxima. Examination of the discriminant spectra (not shown) and the loadings showed a weak protein-like pattern in this direction in the discriminant space. This possibly indicates changes in several partially correlated proteins or amino acids. Another explanation is that protein is involved in intermoleuclar interactions, such as shown previously in a mixture analysis study on Py-MS data [32].
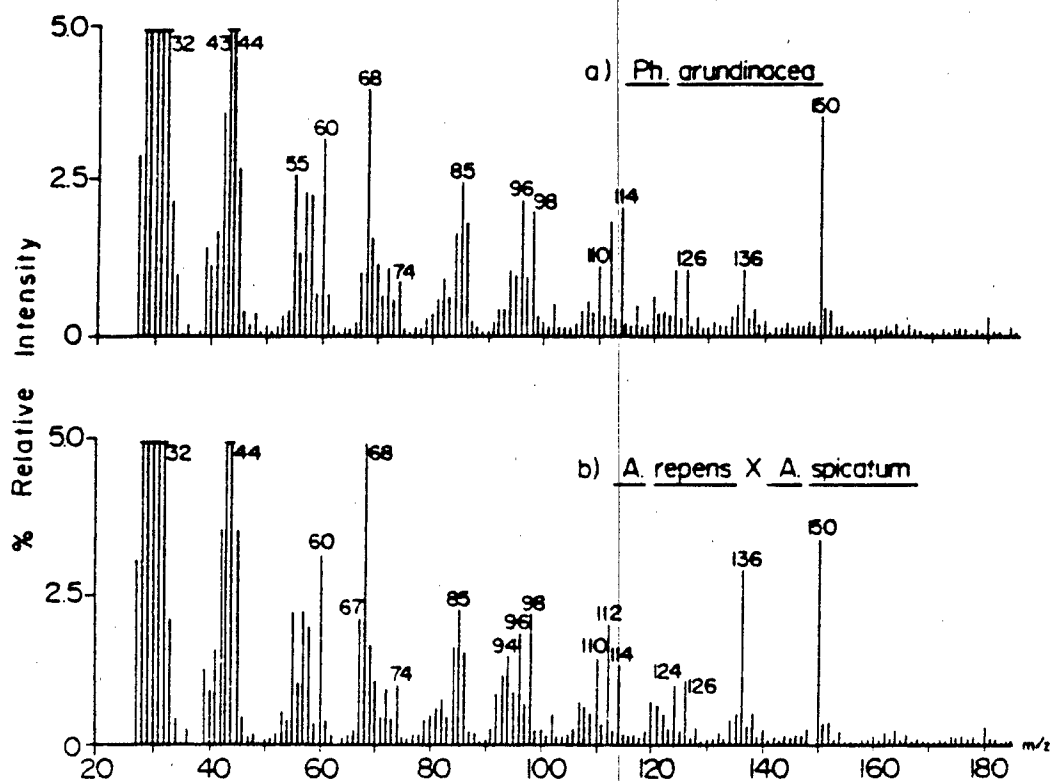
35.

Figure 18. Typical pyrolysis mass spectra form the grass leaves data set.
The spectra show a complex mixture of (poly)hexoses, e.g., at
m/z 31,32,43,60,74,96,98.110,112, and 126; (poly)pentoses,
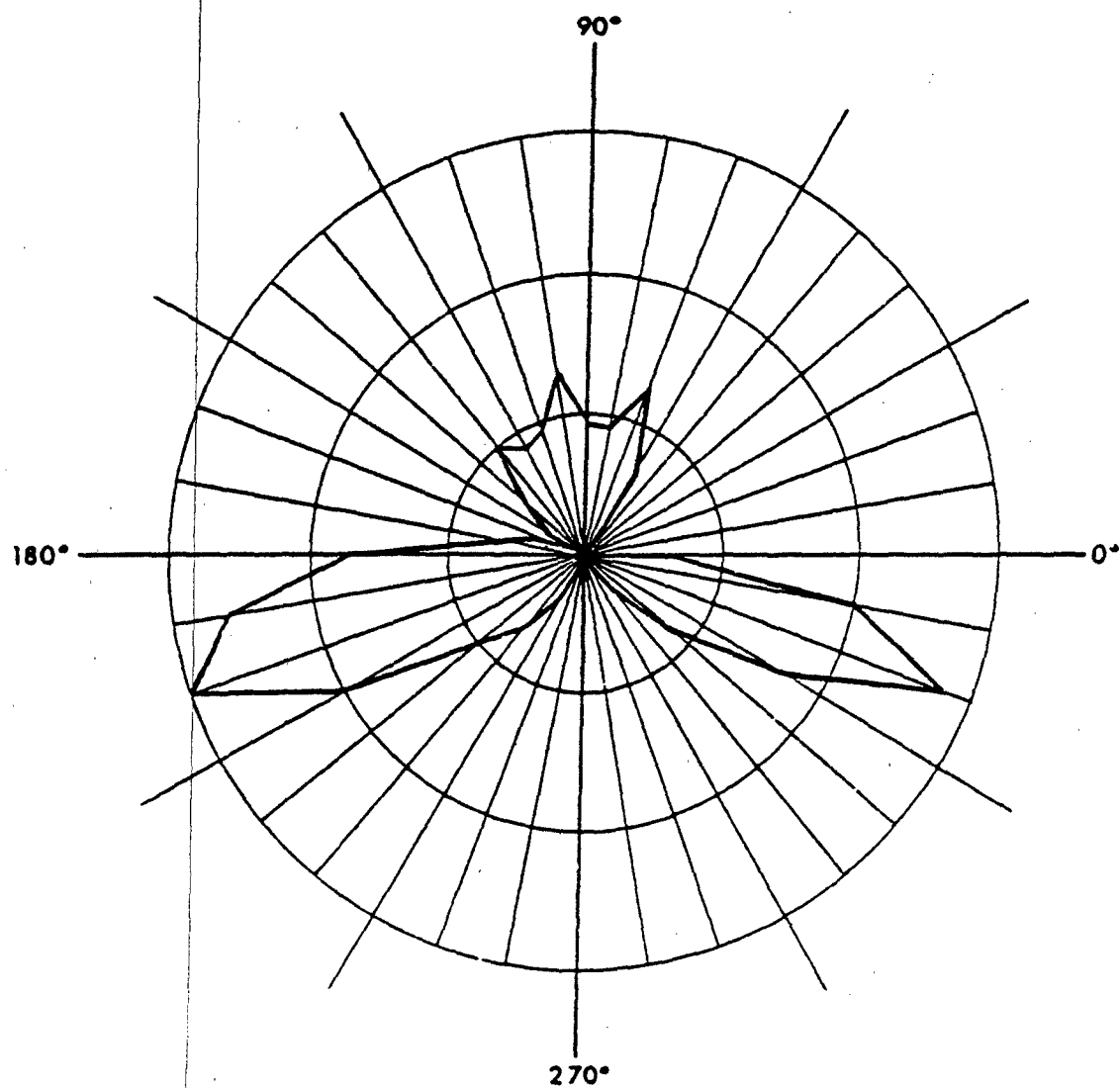e.g., at m/z 85 and 114 and several other component classes.

Figure 19. Smoothed variance diagram obtained from the grass leaves data
set: (a) before and (b) after normalization correction.

The component axis at 200° represents a change in (poly)hexose (e.g., cellulose) content, as can be deduced from Figure 20. The discriminant spectrum of this component axis also shows m/z 85 and 114, which indicates a partial correlation with a change in a (poly)pentose (e.g., hemicellulose) component as deduced form the loadings of these and other pentose related mass values.

The component axis at 340° shows the optimal representation of a relatively pure (poly)isoprenoid pattern, as can be deduced from Figure 20 c,d. From these results it is obvious again that factor analysis techniques can greatly simplify complex spectroscopic patterns.

The component axes obtained were the same as those found with graphical rotation [30].

This procedure can easily be extended to more than two dimensions. For ins   ، let us assume that a certain component is described by three functions. Then the first step is to find the maximum of this component axis in the space of the first two functions L_1 and L_2. If the local maximum of component axis is found at an angle of α degrees with L_1, the next step is to study the variance diagram of a the space described by $((\cos \alpha)L_1 + (\sin \alpha)L_2)$ and L_3. A variance diagram determined in the above described way in a more-dimensional space can be found elsewhere [33]. Our experience with Py-MS data is that rarely ever more than five or six linear combinations are needed to describe the bulk of the data.

As an example of time-resolved analysis, data from analysis of a biopolymer mixture are presented below. For details on this see reference 34.

From preliminary experiments and literature [29,35,36] it is known that DNA has a lower pyrolysis temperature than glycogen which, in turn, has a lower pyrolysis temperature than BSA (bovine serum albumin). Figure 21 shows the total ion current profile. The two unresolved maxima correspond to glycogen and BSA, respectively. The DNA maximum is not visible due to a dominant glycogen change. Although the biopolymer mixture has only three components, the number of factors with an eigenvalue ≥ 1 was 6 (i.e., 64.0, 24.0, 4.9, 2.7, 1.7, 1.3), describing 96.3% of the total variance. This indicates the complex pyrolysis behavior of these polymers. Since the eigenvalues clearly level off after the third factor, only the first three factors were used for further evaluation by the VARDIA technique.

The VARDIA plot of F1 vs. F2 (not shown) did not give a separation of the three components. Since the major change in this data set is the TIC (total ion current), the TIC represented by the first factor. Due to the large variance described by F1 (64%), tendencies other than those described by the F1 are not visible in the VARDIA plot. The VARDIA plot of F2 vs. F3, however gave a clear separation of the three components, as can be judged from Figure 22. Subsequently, the VARDIA technique was used to calculate the linear combination describing the three components.

For the final determination of the linear combinations to describe the component axes, F1 was also included. This was done by taking the 5-10 variables with the highest loadings from each of the maxima in the VARDIA plot in the F2-F3 space. The loadings of these peaks on the first three factors were averaged. The resulting averages were used as the directions of the component axes in the factor subspace. The scores of the
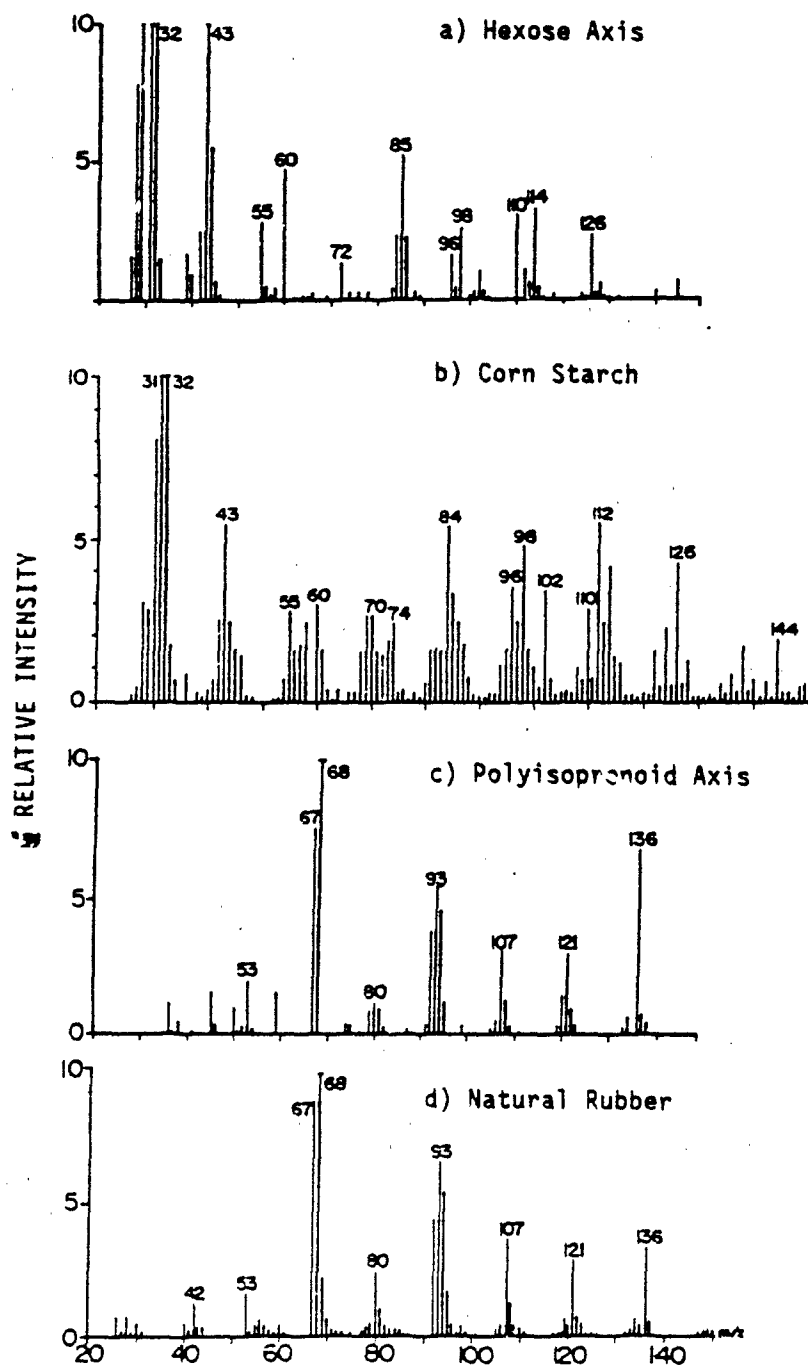
39.

Figure 20. Discriminant spectra of grass leaf component axes and relevant model
compound spectrum: (a) discriminant spectrum of the component axes
at 200° which optimally represents a hexose pattern (compare with
typical hexose pattern of gelatin in Figure 4o); (b) discriminant
spectrum of the component axis at 340° which optimally represents
a (poly)isoprenoid pattern (compare with the typical pattern of
natural rubber in Figure 11c).

components are presented in Figure 23 and show a complete deconvolution of the three time-resolved curves (TRC's) of the components, despite the overlapping behavior as shown in the total ion current curve (Figure 21).

The spectra associated with the component axes were calculated by using the inverse transformation matrix used for the scores. As can be seen from a comparison with the spectra of the pure components in Figure 24, there are clear similarities although some differences, discussed below, are also obeserved. The DNA component axis shows a significantly higher (on the scale used, the intensity of m/z 98 is 17.6) m/z 98 peak than the orignal DNA component. The single ion curve (SIC) of m/z 98 (not shown) shows that it reaches its maximum even before the DNA TRC. The reason for this is not yet clear.

The glycogen component spectrum shows the presence of m/z 64 (SO_2). From the SIC of this ion (not shown) it is obvious that this ion indeed arises early in time. It is very likely that the SO_2 signal originates from BSA.

The lack of an m/z 98 peak in the factor spectrum of glycogen can only be partly due to the fact that the representation here is orthogonal to the DNA axis. If m/z 98 arises from DNA as well as glycogen, the mass variable representing m/z 98 in the factor spacelies between the two components and, consequently, should contribute to the spectrum shown in Figure 18. Thus m/z 98 shows unexpected behavior in DNA as well as in glycogen. A possible explanation may be a nonlinearity in the ion counting system for high count rates. The last spectrum, that of the BSA component axis, shows a high degree of similarity with its model spectrum in Figure 24.
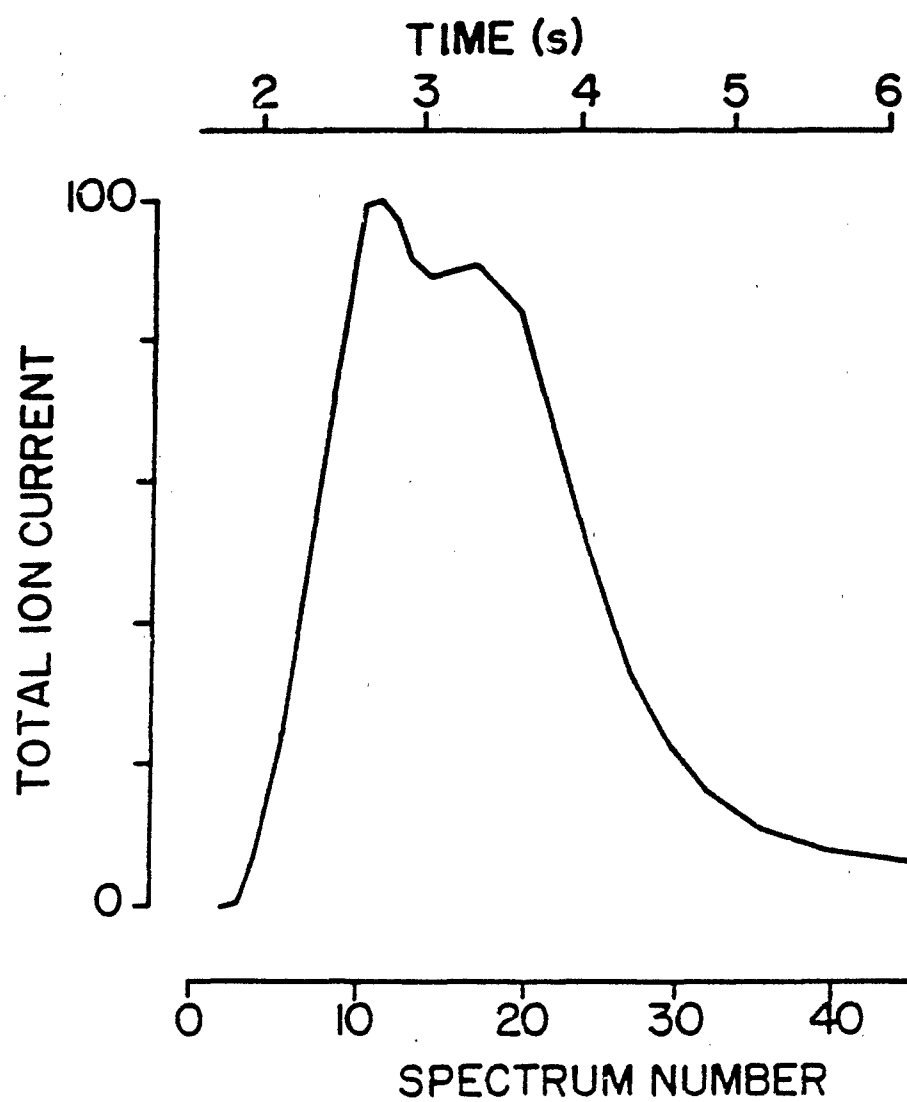
41.

Figure 21. Total ion current (TIC) obtained from time-resolved analysis of the biopolymer mixture.
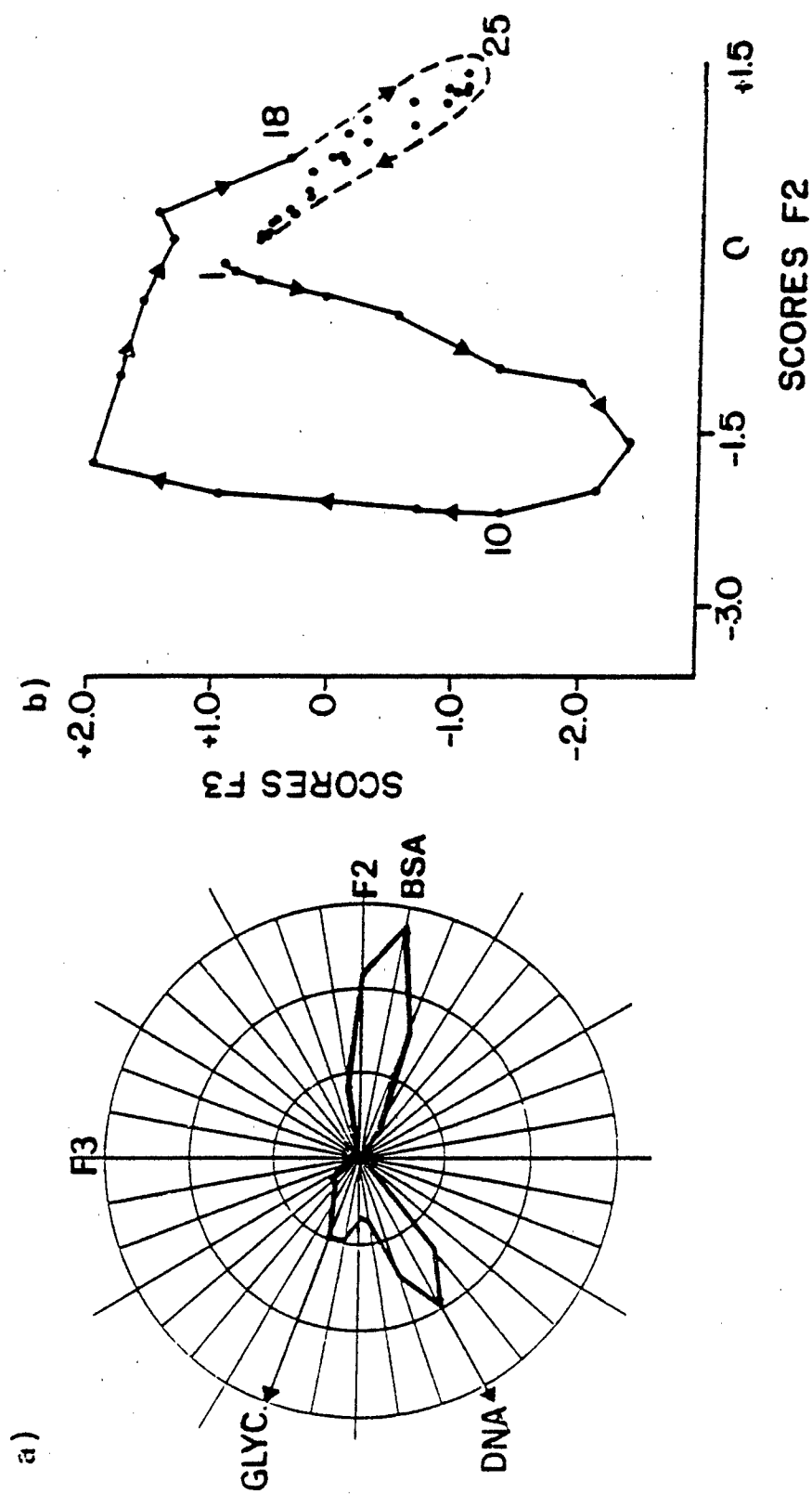
42.

b)

a)

Figure 22. (a) VARDIA plot in the F2-F3 space of the biopolymer data set clearly shows three maxima, corresponding to the three components. (b) Factor scores in the F2-F3 space. The spectral points are interconnected to show the pattern in time. Since the last part of this time series is intertwined, only the tendency is indicated by the dashed line.
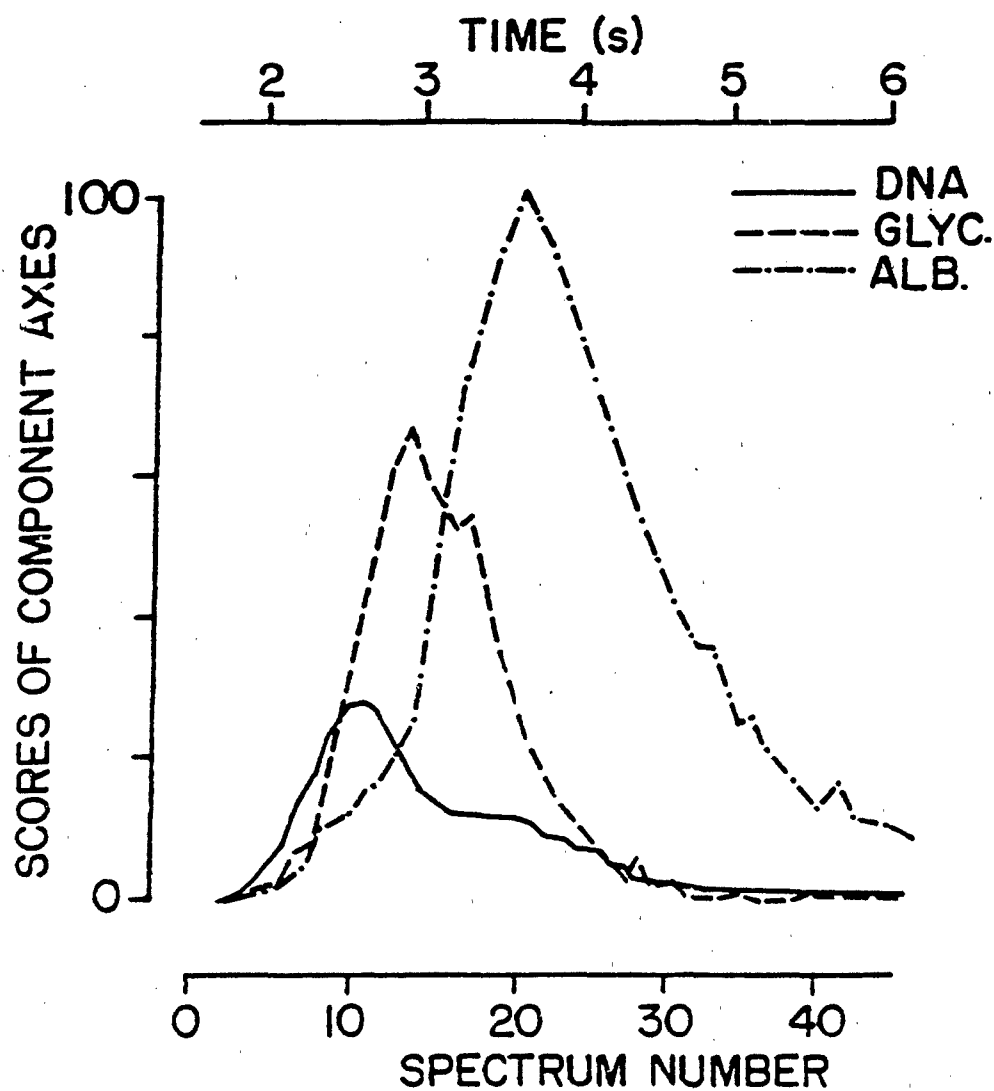
Figure 23. TRC's of the three components of the biopolymer mixture. Despite the heavy overlap of the three components as revealed by the TIC (Figure 4), a complete deconvolution has been obtained.
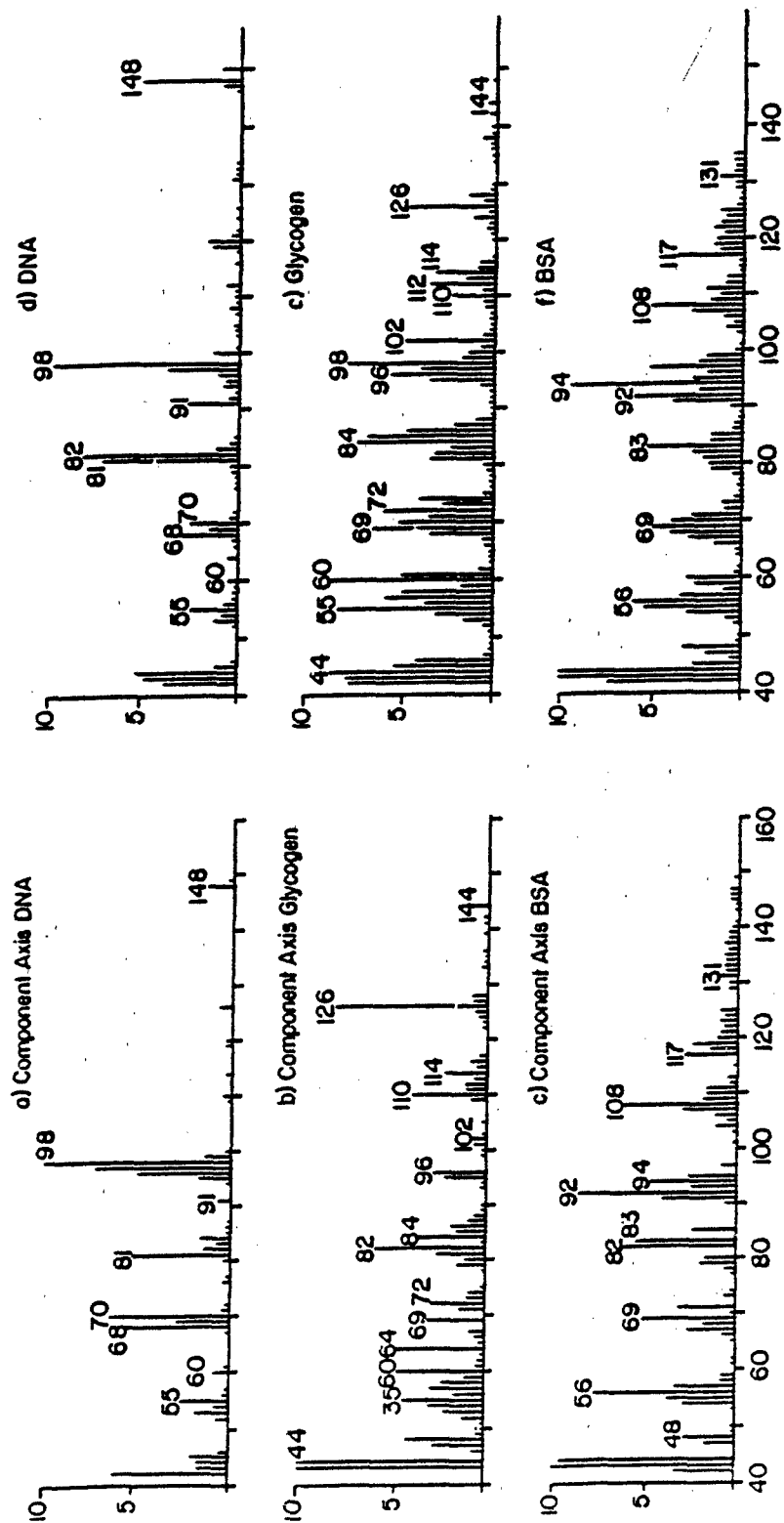
44.

Figure 24. Mathematically extracted spectra from the biopolymer mixture data set and reference spectra: a,b, and c are the component axes; d,e, and f are the model spectra of DNA, GLYC, and BSA, respectively.

45.

C. Use of the Variance Diagram as a Monitoring Tool - Although the VARDIA technique as described above is a powerful tool to find the component axes (scores), presentation of the occurrence of these component axes as a function of time gives a more appropriate representation of the data. This is possible by calculating the variance not in 10° rotational steps in a 2-dimensional subspace but in the directions of the sequential spectra in n-dimensional factor space [37].

Normally, the calculated average spectrum of a data set is used as the reference point (the "origin") in n-dimensional space for rotational procedures such as graphical rotation and the VARDIA technique. This is done because the interest is generally in the relatively minor differences between spectra rather than in the absolute intensities of the signals. In the case of time-resolved spectra, however, the absolute contribution of each spectrum to the data set is of interest. Thus the "permanent" background spectrum seen by the mass spectrometer is a more appropriate reference point. Thus for the time-resolved VARDIA method, the variance as calculated by equation 3, in the direction of the sample spectra with a background spectrum as the reference point can be used. In contrast to the variance values calculated by equation 2, which only considers two dimensional (sub)spaces, the variance in the directions of the sample spectra can be calculated in an n-dimensional space containing all significant factors. The equation describing this procedure reads as follows:

$$Var_t = \sum_{i=1}^{m} a_{i,t}2 \text{ for } a_{i,t} \geq \sum_{j=1}^{n} \alpha_{i,j} * L \tag{3}$$

$$\text{Where } a_{i,t} = \sum_{j=1}^{n}((S_{t,j} - S_{tr,j})/(\sum_{j=1}^{n}(S_{t,j} - S_{tr,j})2) * \alpha_{i,j}$$

46.

Where: m is the number of variables.

n is the number of factors,

$a_{i,t}$ is the loading of mass variable i from spectrum t on the rotated function.

$\alpha_{i,j}$ is the loading of mass variables i on the original factor j,

$S_{t,j}$ score on factor j of spectrum t

$S_{tr,j}$ is score on factor j of reference spectrum

L is the cosine of the angle between the mass-axis and the rotation under consideration.

Expressed in words, the procedure accomplishes the following tasks:

(a) subtraction of the scores of the reference spectrum from the n scores of each spectrum, (b) normalization of the new scores (i.e., the sum of the squares of then scores of each spectrum are made equal to 1.0); (c) calculation of linear combinations of factors for each spectrum; (d) selection of all mass axes that are within a certain angle (angle = $\cos^{-1}$ L) of the rotated functions; and (e) summation of the squares of the selected loadings of each factor to give the variance as defined by the equation.

Thus, only mass variables whose projected direction in n-dimensional factor space makes a relatively small angle (15-20 degrees) with the direction of the selected linear combination of factors are included in this adapted variance calculation. A high variance value indicates a relatively high loading of included mass variables on the selected linear combination of factors, thus indicating which mass variables are responsible for the observed correlated behavior.

Initial application of this technique showed that the results improved considerably after smoothing the scores with the following triangular window:

$$S_{t,j}{}^* = 1/4 * S_{(t-1),j} + 1/2 * S_{t,j} + 1/4 S_{(t+1),j} \qquad (4)$$

where $S_{t,j}{}^*$ is the smoothed score.

Since the time-resolved VARDIA uses a subspace of the original factor space, the question may arise if the overlap between the two spaces is sufficiently high to regard the subspace as representative of the original space. In order to check this, the % variance covered by the serial subspace was calculated with the following equation:

$$\text{Vars} = \sum_{i=1}^{m} (\max(a_{i,1}, a_{i,2} \ldots a_{i,k}))^2 * 100\% \qquad (5)$$

where k is the number of spectra.

The maximum of a loading $a_{i,j}$ (j = 1,k) for all linear combinations in the direction of the spectra in the factor subspace may be considered as the "length" of that mass variable in the subspace (the standardized length of each mass variable in the orginal factor space is 1.0). Since the communality [38] is the square of the length of a variable, the above formula calculates the sum of all the communalities from which the relative variance can be calculated by dividing this sum by the number of variables involved.

As an example, a time-resolved data set of a wood sample will be shown. This data set has been studied before with the VARDIA technique [34,39]. For the study, the 155 spectra recorded between 300° and 500° were used. Prior to multivariate analysis, sequences of five spectra were averaged in order to reduce noise and the size of the data matrix. Three significant factors (describing 97% of the total variance) resulted in three maxima using the VARDIA technique, which allowed a complete deconvolution into three processes. From this study [34] it was concluded that the first process was dominated by hemicellulose pyrolysis, the second by cellulose pyrolysis and the third by lignin pyrolysis. The

pyrolysis of lignin, however, does not only take place in the third pyrolysis step process. It appeared that there were three separate lignin pyrolysis processes, of which the first two were highly correlated with the hemicelluloses and cellulose pyrolysis steps [34]. These findings were in agreement with a time-resolved MS study on biomass samples by Evans et al. [40].

The TIC curve of the time-resolved wood data set is given in Figure 25 whereas the time-integrated spectrum of this data set is shown in Figure 26. The VARDIA-S technique resulted in the curve given in Figure 27. As can be seen, the VARDIA-S technique gives a considerably improved resolution. The contribution of the hemicellulose and lignin components were only minor. Partial (25%) subtraction of the first factor, however, showed the significance of these two components in the VARDIA-S curve. The mathematically extracted spectrum of describing the main lignin pyrolysis process and a model spectrum of lignin are presented in Figure 28.

6. Computer-Enhanced Py-MS Analysis of Simulated Biological Agent/- Interferent Mixtures

The purpose of this experiment was to demonstrate the feasibility of detecting and quantifying two different microorganims (B. anthracis and Streptococcus type B) in the presence of varying concentrations of a natural background material (Dugway soil) by means of pyrolysis mass spectrometry and multivariate analysis techniques.

In view of the model experiment nature of the tests, no attempt was made to obtain multidimensional MS data. Obviously, greatly improved
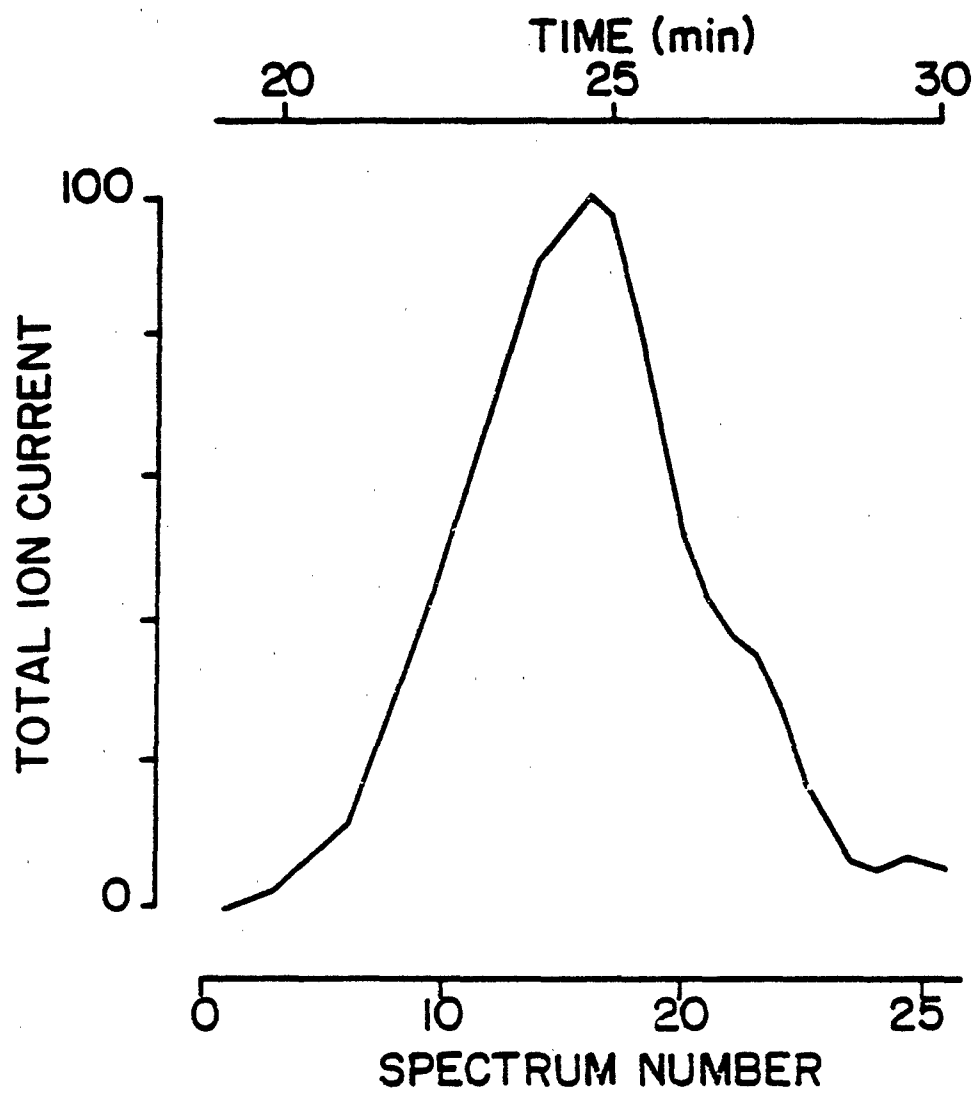
Figure 25. TIC-curve of the time-resolved TG/MS analysis of wood, which does not show clearly resolved processes.
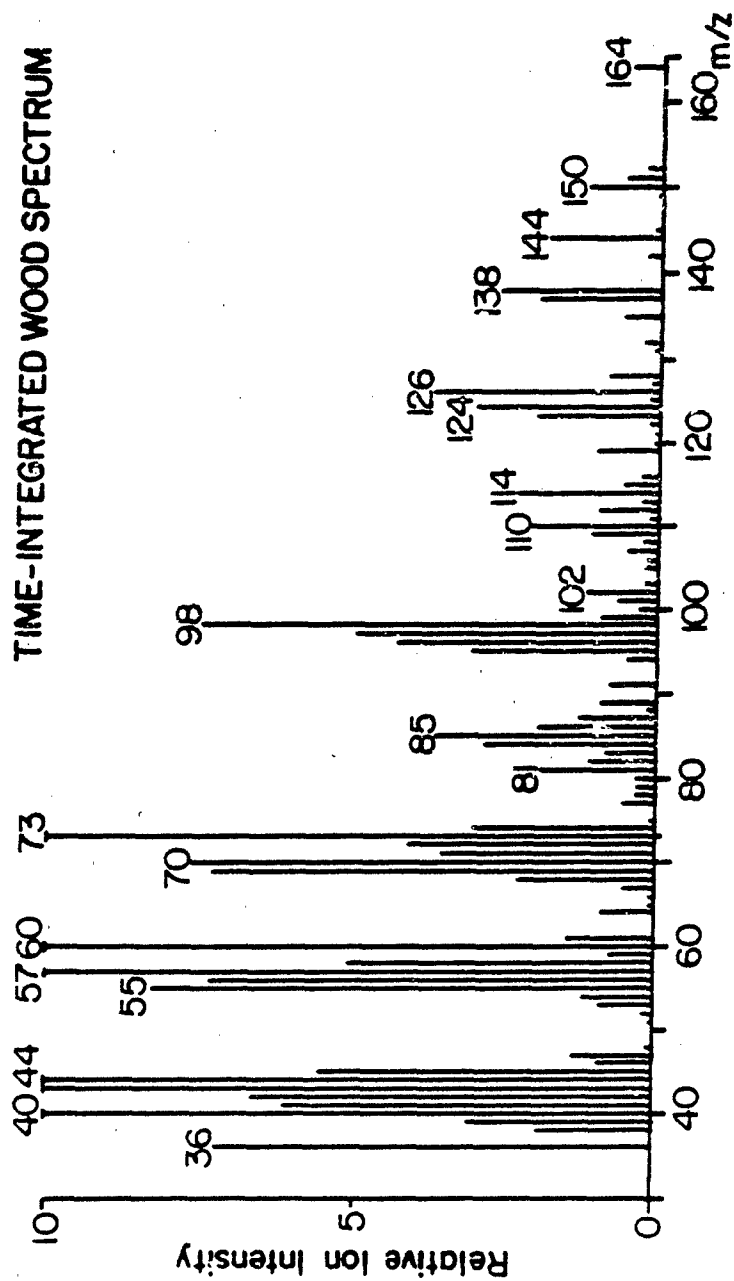
50.

TIME-INTEGRATED WOOD SPECTRUM

Figure 26. The time-integrated spectrum of the wood sample. Contributions of the three major wood components are clearly present: hemicellulose (viz, m/z 85 and 114), cellulose (viz. m/z 98, 102, 110, 126 and 144) and lignin (viz. m/z 124, 138, 150 and 164).
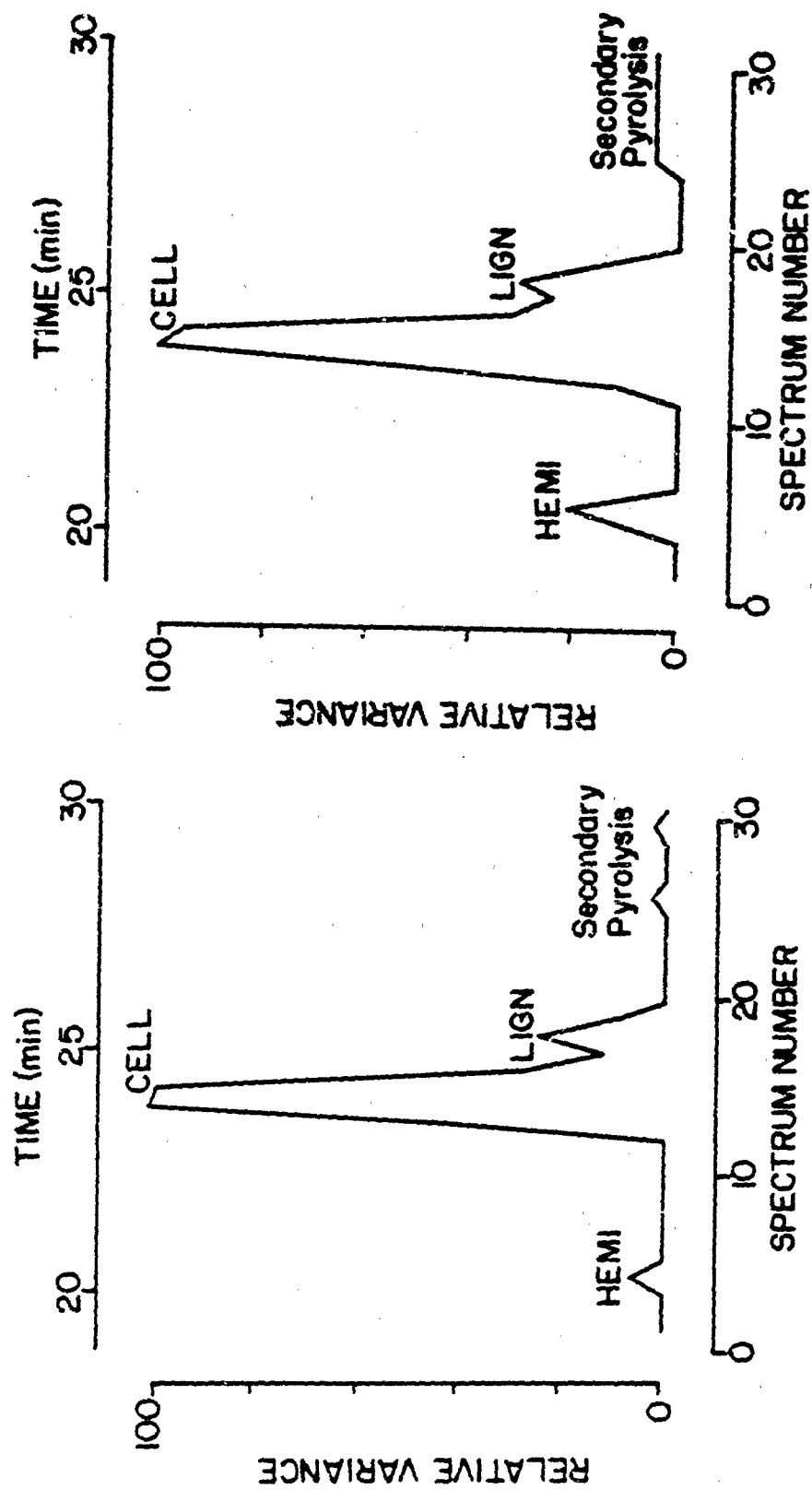
51.

Figure 27. The VARDIA-S curve obtained from the first three factors of the data set of time-resolved analysis of wood (a) clearly shows three processes, thus achieving a better resolution than shown by the TIC curve in Figure 25). Although the maxima of the hemicellulose and lignin components appear minor, their significance is obvious after subtracting 25% of the first factor, as shown in (b). The overlap of curve (a) with the original space is 92.4%, the overlap of (b) with the original space is 90.9%.
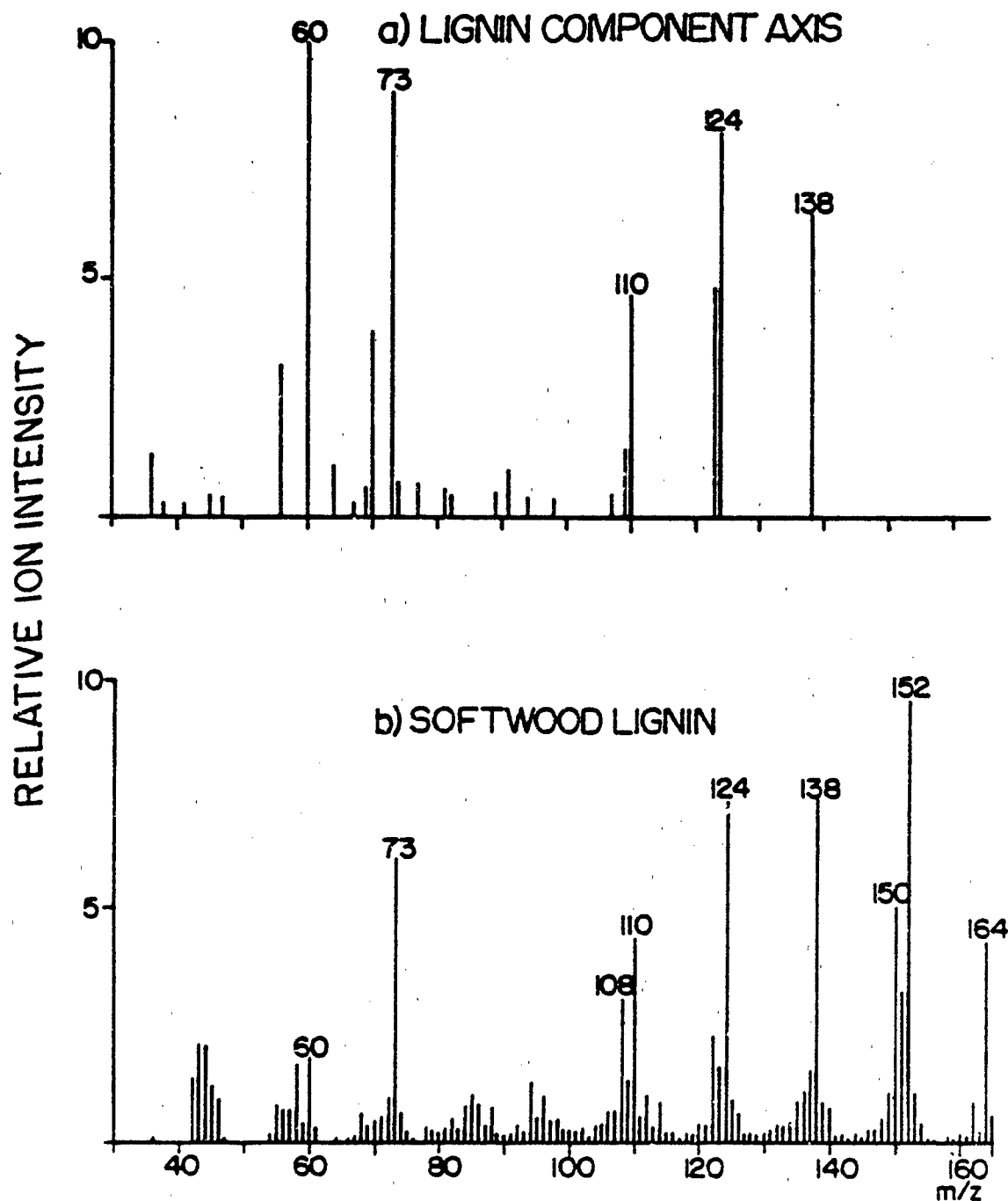
52.

Figure 28. The mathematically extracted spectrum a of the wood data set shows typical lignin peaks at m/z 110, 124 and 138 (a). The softwood lignin spectrum (b) shows additional lignin peaks at m/z 150, 152 and 164. The absence of the latter peaks in (a) is due to the complex behavior of lignin.

detection and identification performance is possible when using MS^n, GC/MS or time-resolved pyrolysis MS techniques. Moreover, a relatively narrow mass range was used, thereby excluding most lipid peaks. Finally, the present data were obtained on a conventional quadrupole MS system rather than on an ion trap MS system with Curie-point pyrolysis inlet. In the latter configuration sensitivity improvements by 2 or 3 orders of magnitude have been demonstrated [10] thereby allowing the use of sample sizes in the low nanogram range. However, one-dimensional spectral data obtained on a conventional quadrupole Py-MS system were judged adequate to test the performance of the pyrolysis techniques and of the multivariate analysis methods.

A. <u>Experimental Procedure</u> - Samples of the two microorganisms and the Dugway soil were received from CRDEC. Samples were prepared by suspending in methanol (1 mg/ml) followed by mixing various proportions of each suspension for a total of 10 different mixtures. As shown in Figure 29 the various concentrations have been chosen in such a way as to simulate systematic variations along a hypothetic time axis. Thus, the 10 mixtures provide a model for time-dependent changes in concentrations, e.g., in consecutively collected aerosol samples. Then ul aliquots of each suspension (10 ug dry weight were analyzed by means of direct pyrolysis MS using an Extranuclear 5000-1 quadrupole system with Curie-point pyrolysis inlet developed at UUBPC. Pyrolysis MS conditions were as follows: Curie-point temperature 610°C, heating rate approx. 100°K/s, total heating time 10 s, electron energy 14 eV, mass range m/z 20-140, scanning rate approx. 500 amu/s, total scanning time 20 s. Multivariate data analysis was performed on an IBM 9000 microcomputer system using the SIGMA programs developed at UUBPC.
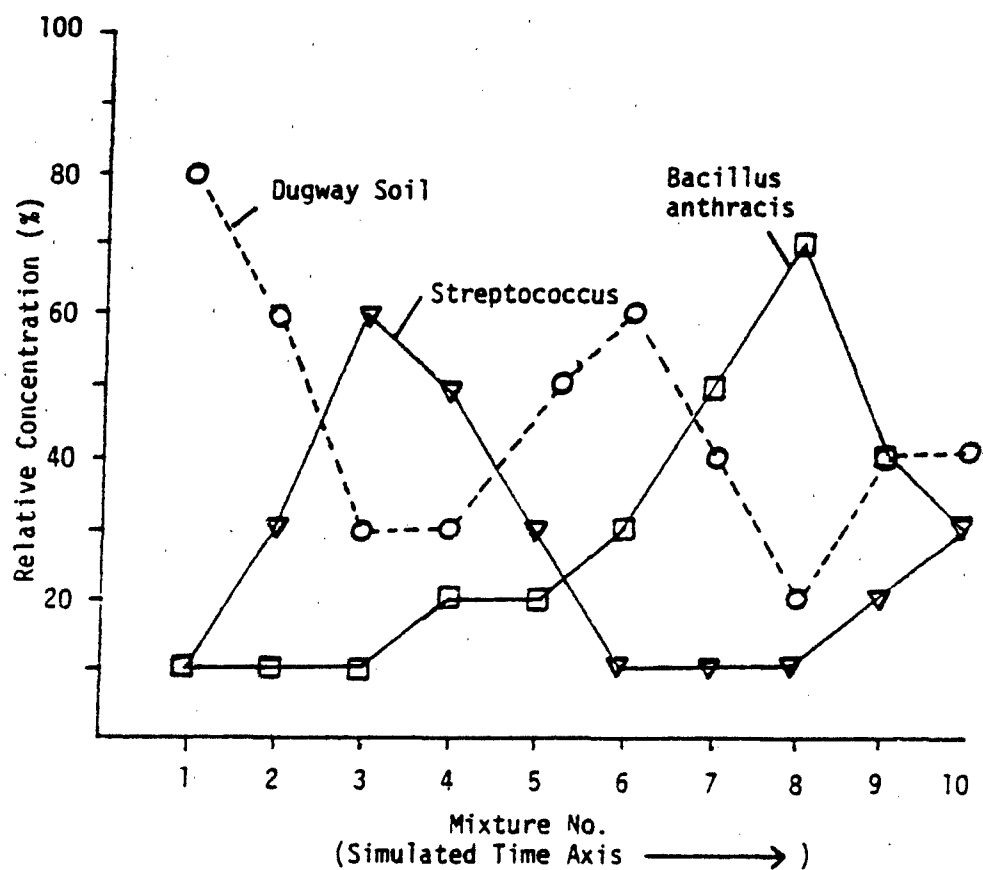
Figure 29. Relative concentrations of 10 ternary mixtures containing simulated biological agents' Streptococcus group B + Bacillus anthracis) and interferents (Dugway Soil).

B. **Results and Discussion** - Pyrolysis mass spectra of the three pure materials are shown in Figure 30. Notice the differences between the pyrolysis patterns of the two bacteria (Figures 30a and b) as well as the major differences between the soil sample and both bacteria. The streptococcus B pattern in Figure 30a is characterized by strong signals representing neutral sugars from cell wall constituents, e.g., at m/z 57, 74, 82, 84, 96, 98, 102, 110, 114 and 128. Especially the signal at m/z 128 is noteworthy, since it is likely to represent rhamnose, a biomarker for microorganisms in general and a characteristic constituent of certain streptococcal group and type antigens. The identity of the peak at m/z 128, which corresponds to the rhamnose "monomer" minus H_2O can be confirmed by MS/MS.

By contrast, the Bacillus anthracis pattern in Figure 30b is dominated by protein signals, e.g., at m/z 92/104 (phenylalanine fragments) and m/z 94/108 (tyrosine fragments). In addition, the relatively inconspicuous peak at m/z 79 can be confirmed by MS/MS to contain pyridine, a decarboxylation product of the characteristic spore wall component di-picolinic acid. The Dugway soil pattern is characterized by relatively high $HCl^{+}\cdot$ peaks at m/z 36/38, apparently representing the high NaCl content of the Utah desert (see also the marked Na+ peak at m/z 23). Furthermore, the high peaks at m/z 64 (SO_2^{+}) and 94 (probably phenol) are likely to represent sulfates and soil humic acids, respectively.

Two different multivariate analysis approaches, namely supervised and unsupervised analysis, were followed in order to demonstrate several possible application of multivariate techniques to agent interferent mixture data.
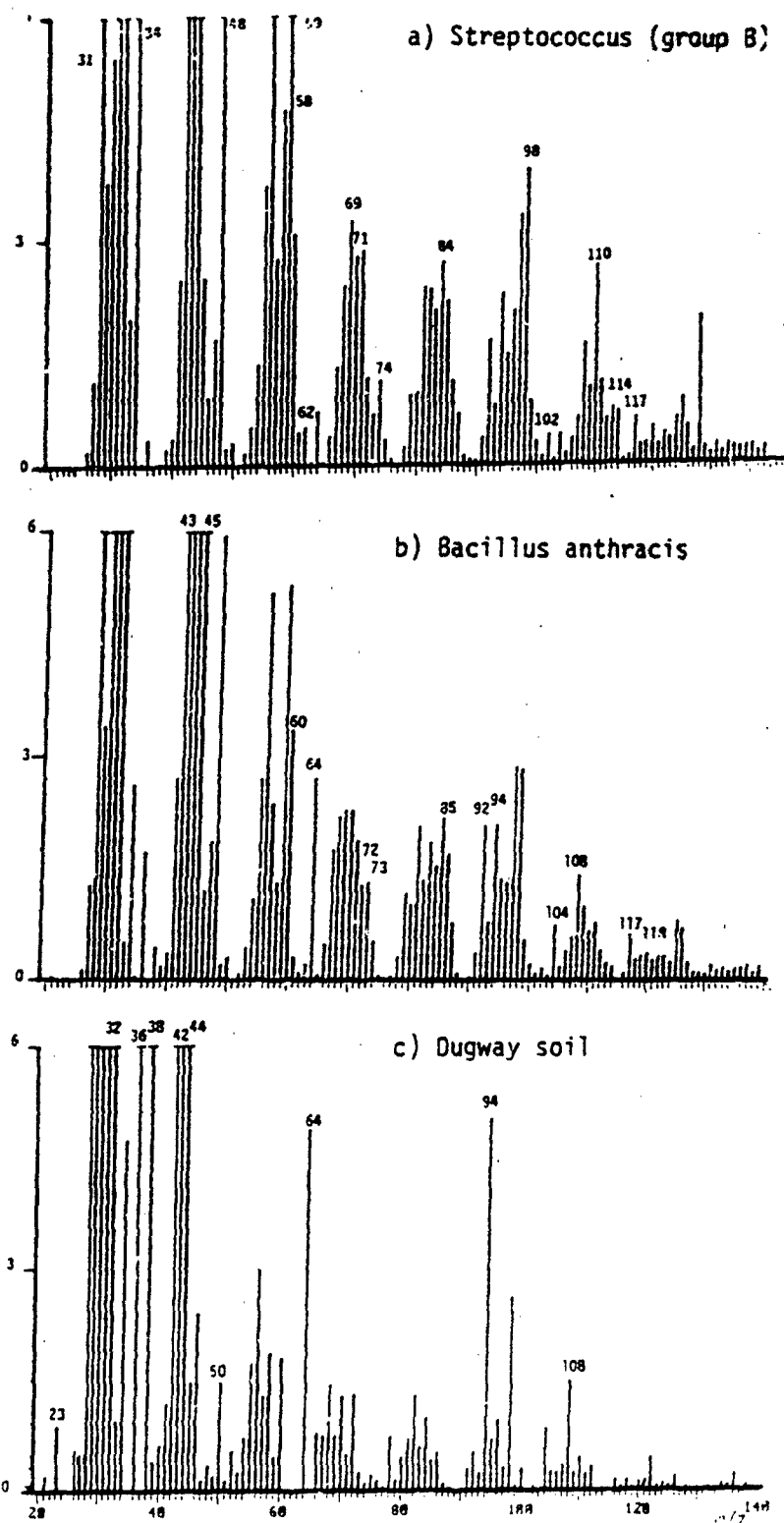
Figure 30. Curie-point pyrolysis mass spectra of the "pure" mixture
components; a) Streptococcus B, b) Bacillus anthracis, and
c) Dugway soil.

57.

In the supervised analysis approach, the mixture data set was subjected to factor analysis while including the spectra of the three pure compounds. This enables numerical extraction of component patterns and estimation of relative concentrations by target rotation and can be compared to a situation where the key chemical patterns of known biological threats and interferents are already present in a computer library.

Figure 31 shows the relationships between the 10 mixture and 3 pure ("target") spectra in F1/F2 space. As expected, all mixtures lie within a triangle spanned by the three target spectra. The triangle represents a so-called "ternary diagram" and is due to the normalization of the total ion signal in each spectrum to 100%. As shown in previous studies [41], ternary diagrams obtained by factor analysis allow a direct estimate of the relativel concentrations of the three components.

Calculated relative concentrations are shown in Figure 32. A direct comparison between Figure 32 and 29 reveals a high degree of correspondence with regard to the relative concentrations of the three components, thereby confirming the high degree of linear additivity of pyrolysis mass spectra obtained by the Curie-point pyrolysis technique. Moreover, the factor analysis approach shown in Figure 31 allows direct visualization of the chemical parameters responsible for the observed relationships between the mixture spectra and the pure target compounds.

Finally, a numerically "extracted" spectrum of the chemical component responsible for the high score of mixture sample #3 on the streptococcus axis in Figure 31 is shown in Figure 33a, revealing nearly all of the mass peaks found to be characteristic for streptococcus group B in Figure 30a.
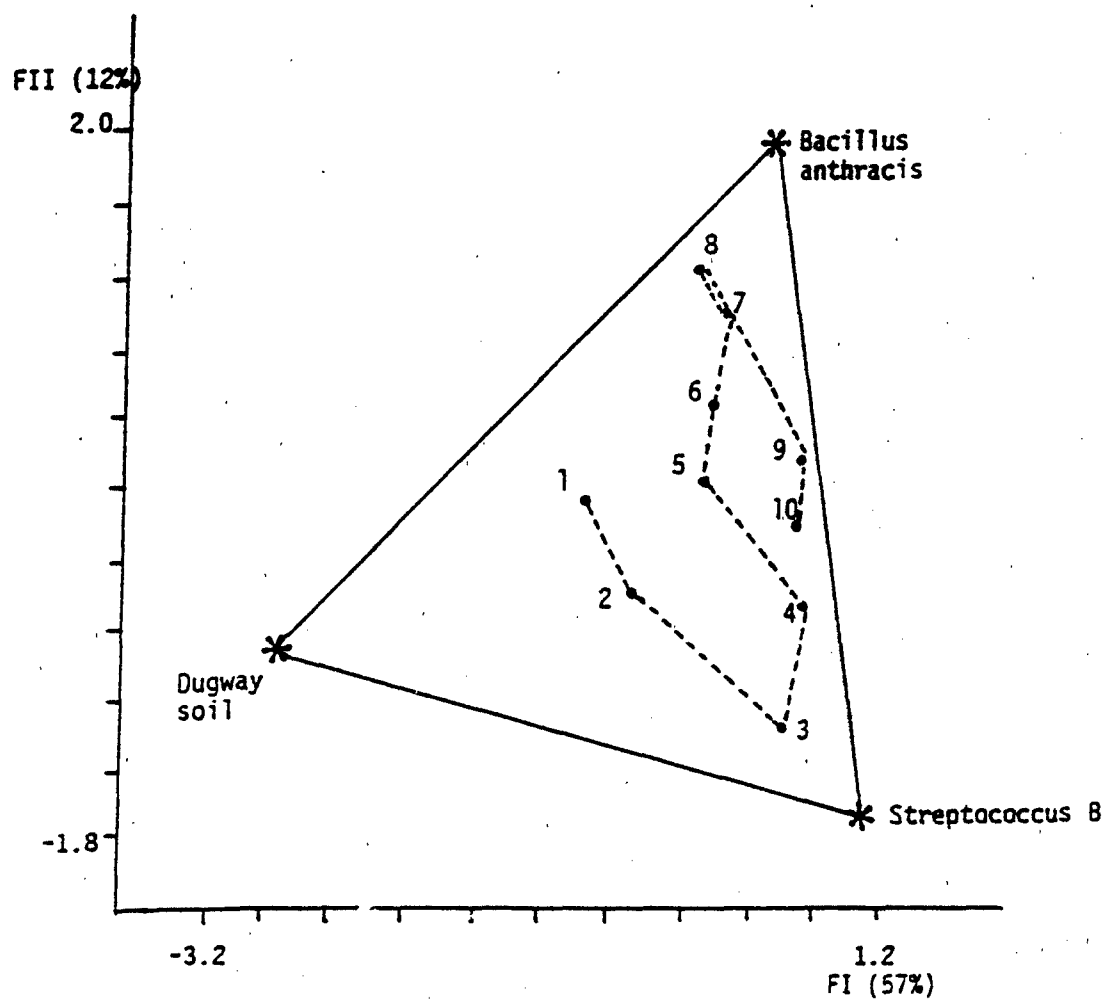
Figure 31. Factor scores of the mixture spectra showing the ternary diagram configuration. The three library ("target") spectra are indicated by *. Note that all 10 mixture spectra lie within the boundries of the triangle formed by the pure target spectra.
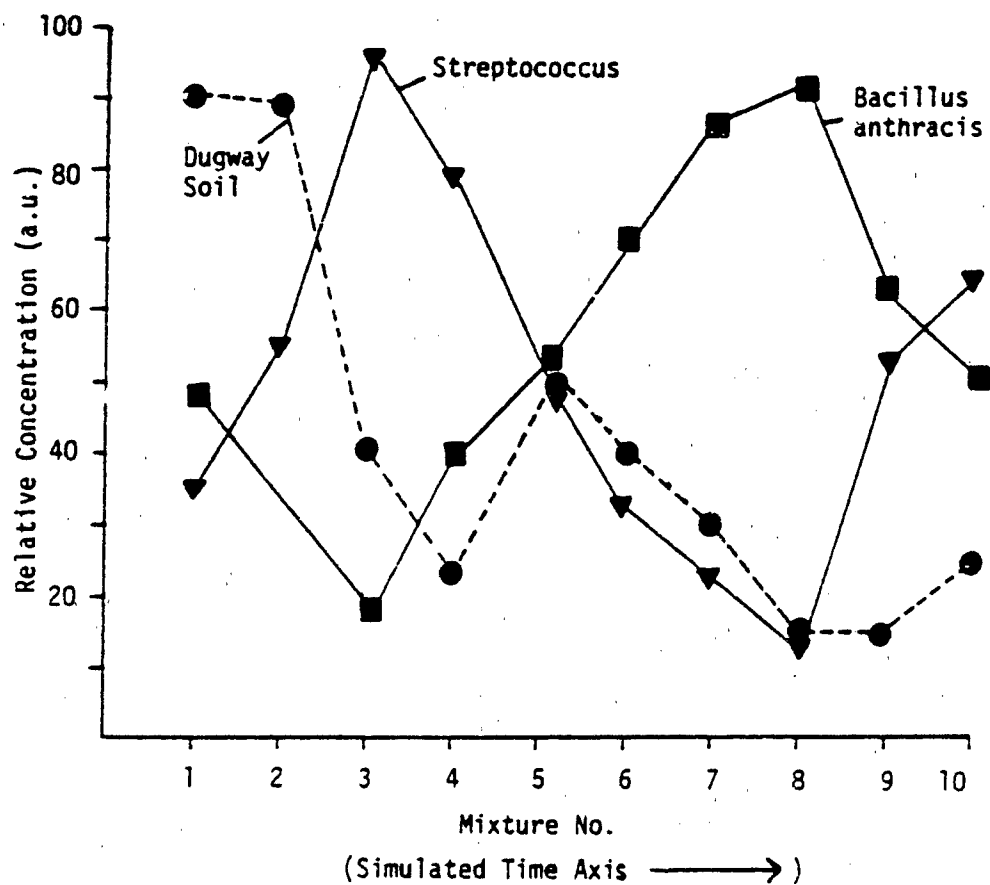
Figure 32. Relative concentrations of mixture components calculated from library target rotations.
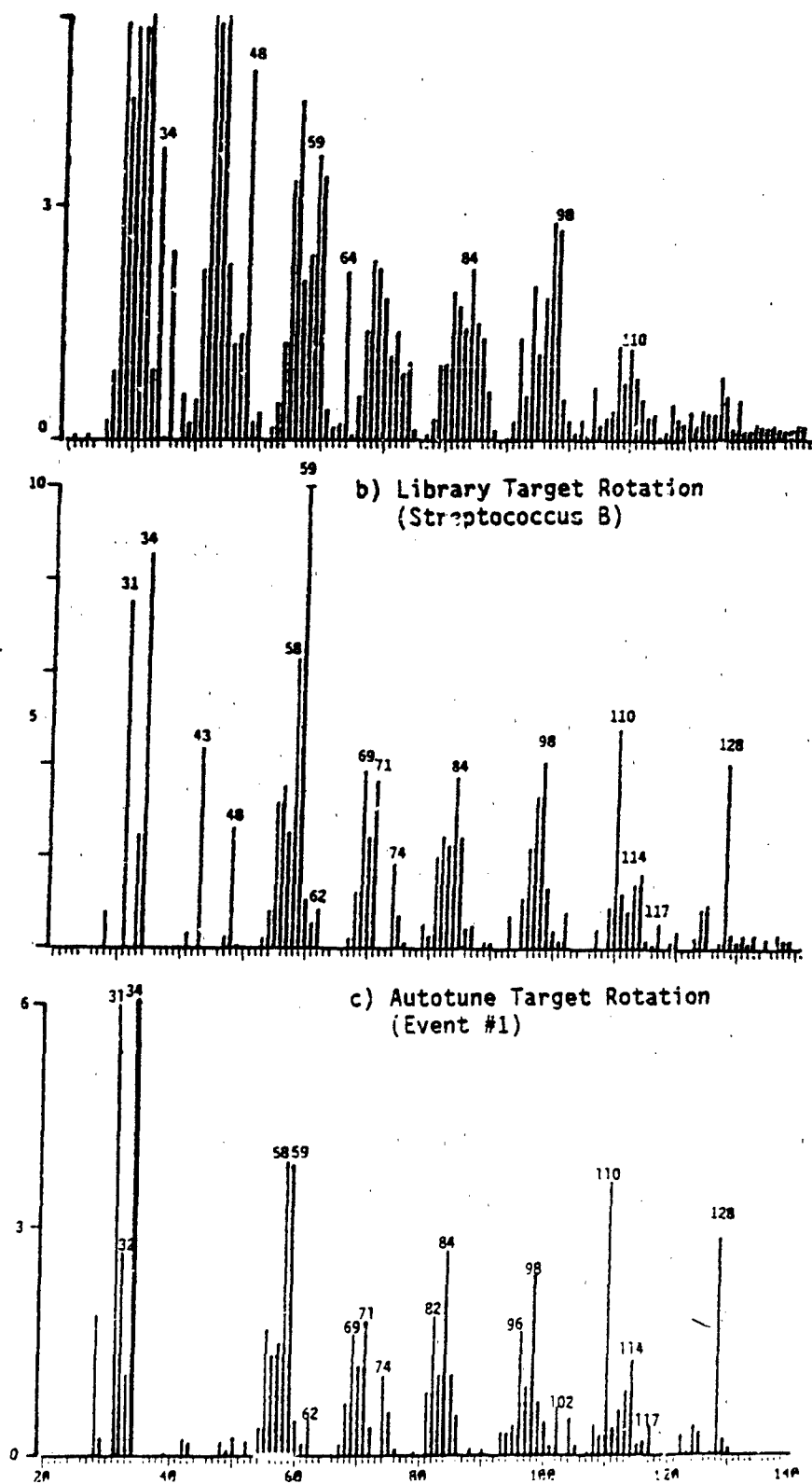
Figure 33.   Example of a typical mixture spectrum (a - Mixture #6), and of two
numerically extracted component spectra (b - Supervised Rotation;
c - Unsupervised Rotation).

61

How can the supervised target spectrum approach be used in a practical
biodetection situation? First, let us assume that our biodetection system
computer contains library of known threat agents and background
interferents and that this library is a true chemical library (i.e.,
containing information about biomarkers and other characteristic chemical
components) rather than a library of "fingerprints" (i.e., patterns
without chemical meaning). In that case the normalized chemical
signatures ("spectra") of pure threat agents and interferents will lie
inside the hypertetrahedron or on its various vertices and triangular
surfaces.

If the chemical pattern of an actual aerosol sample is found to lie
within (or on) the boundaries of the hypertetrahedron its composition can
be readily and totally explained by known agents and/or interferents.
Alternatively, if the new chemical pattern lies outside the boundaries of
the hypertetrahedron it must contain some unknown component(s). The
relative concentration of the unknown fraction can be determined by
subtracting the contributions of known components (which can be measured
through orthogonal projections on the surfaces of the hypertetrahedron).
In general, it can be stated that the greater the distance between the
unknown pattern and the hypertetrahedron, the smaller the contribution of
known library components will be.

If so desired, the position of the "pure" unknown component can be
calculated, according to a method described by Windig et al. [42], thus
forming a new apex which enlarges the volume of space spanned by the
hypertetrahydron and serves to help recognize future occurrences of
similar components. This makes the recognition process self-learning.

The mathematics of this mixture analysis approach as outlined in several UUBPC publications are straightforward and can be applied to any set of mixture patterns. However, the accuracy and reliability of this procedures depends entirely on the precision with which the locations of the hypertetrahedron apices are known in multidimensional space. Since each apex represents a biological agent or interferent material which is highly variable by nature, a so-called "fingerprinting" approach is doomed to fail. Use of biochemical marker signals, however, serves to "anchor" the positions of the apices more firmly in space.

The abovedescribed methods are primarily "supervised" in nature (i.e., dependent on prior knowledge) but flexible enough to be adaptable to a self-learning mode. A typical result of such a supervised analysis will be a list or histogram showing the estimated concentrations of known biological agents and/or interferents as well as any "unexplained residue". Generally, the latter will then be incorporated into the library for future reference purposes.

As discussed in the previous paragraphs, the supervised method will fail to identify library components if the distance between the aerosol pattern and the hypertetrahedron is too large (i.e., the orthogonal distance to the nearest triangular surface is large in comparison to the size of the triangle). In this case, unsupervised methods may produce more reliable results, especially if a series of new aerosol patterns is found to form a more or less coherent cluster well outside the hypertetrahydron. This could be due to the occurrence of an unknown (or unanticipated) class of agents or interferents as well as to a loss of instrument calibration or other systematic malfunctions. Also,

unsupervised methods can be used to detect and characterize unknown aerosol components if the system library has been lost or damaged.

The first challenge for an unsupervised approach is to provide reliable "event" detection, where an "event" is defined as any significant change in the relative concentration of aerosol components over time.

When using the so-called Sequential Variance Diagram (VARDIA-S) method described in the previous section [37], two distinct events are observed, as shown in Figure 34. Since the VARDIA-S method can be used in a "moving time window" fashion, Figure 34 represents the results when using the first mixture (80% soil) as a reference point and the entire set of 10 spectra as the window.

Events are detected as soon as correlated changes occur in mass peak intensities of spectra within the window. Consequently, most noise contributions are filtered out. Moreover, as soon as correlating tendencies are detected and characterized (e.g., variations in soil background concentrations), such components can be effectively subtracted out. Since the "event" peaks observed in Figure 34 represent summed correlations ("loadings") rather than summed peak intensities ("concentrations"), Figure 35 shows the calculated relative concentrations using co-variance as a measure of peak intensity, according to a procedure described by Windig et al. [15,42].

Each of the two concentration maxima in Figure 35 was calculated by using the opposite maximum (as defined by the event detection procedure in Figure 30) as a reference point, thus effectively deconvoluting any remaining overlap between the two components. Finally, after detecting the two events and determining the relative concentrations of the
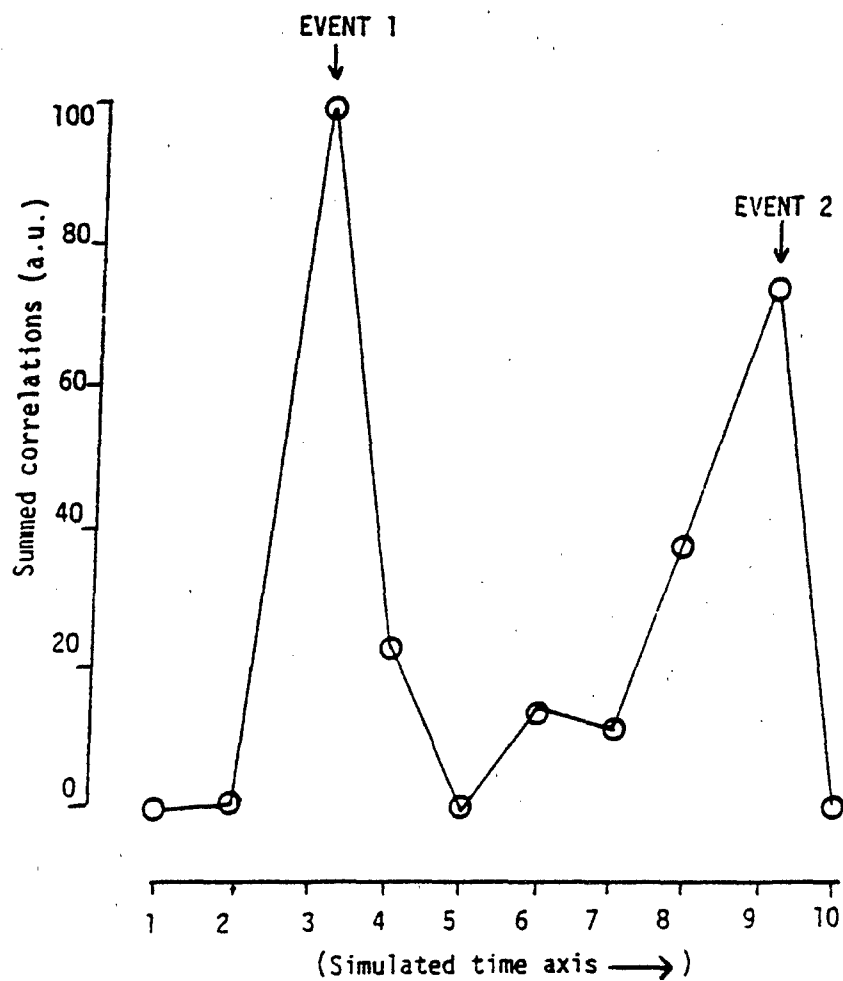
Figure 34.   Event detection approach using the VARDIA-S method
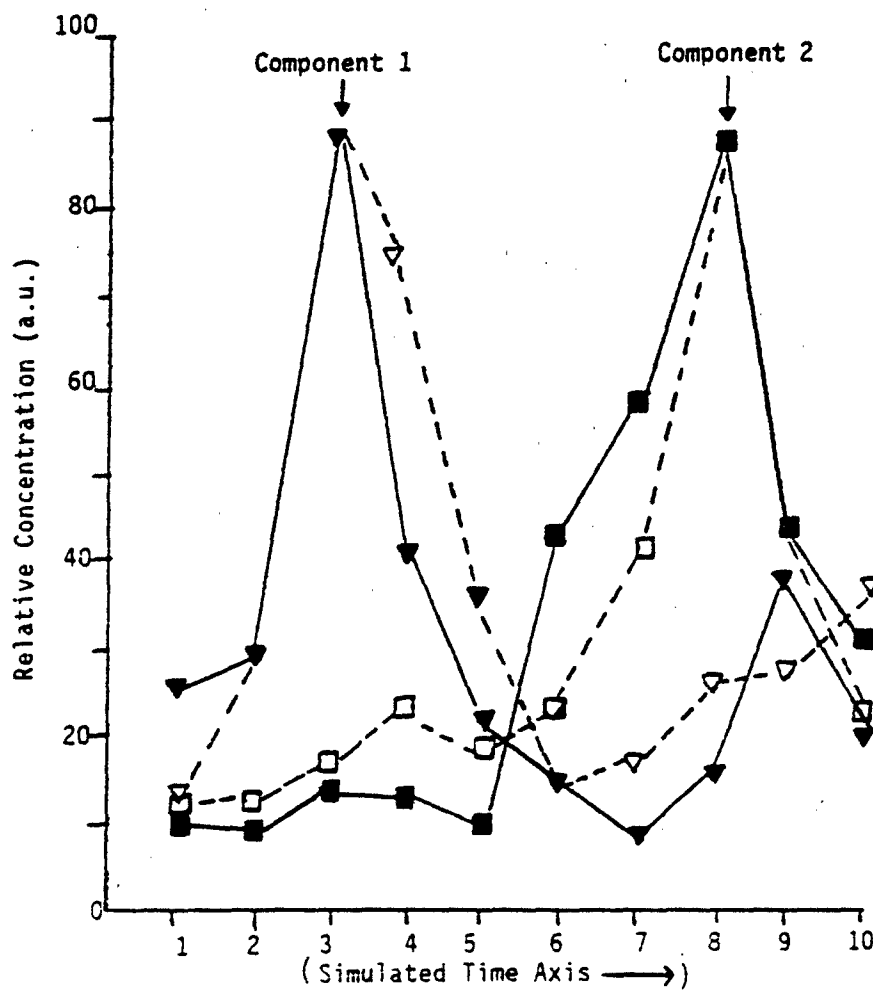(see text).   Note maxima at samples 3 (event 1)
and 9 (event 2).

65.

Figure 35. Relative concentrations of component 1 (Streptococcus B) and component 2 (Bacillus anthracis) calculated from unsupervised ("autotuning") rotations. Dashed lines indicate actual con-centration curves (relative to Dugway soil concentrations).

components involved (absolute concentrations can be determined only if the weight and pyrolysis yield of each sample are known) the remaining question is: what can be learned about the identity of the components detected and measured by this unsupervised approach?

Although there are no library spectra in this case, it is possible to extract the spectra of the unknown components represented by the various events from the series of mixture spectra using graphical rotation methods. The results of this approach for the first maximum in Figures 34 and 35 (caused by the increased concentration of the streptococcus B component) are shown in Figure 33c. Not only is the spectrum so obtained nearly identical to the spectrum in Figure 33b obtained by target rotation to the library spectrum of streptococcus B, the numerically extracted spectrum retains all of the most characteristic peaks.

Thus numerical extraction of chemical patterns from unknown mixtures can provide valuable indications for more specific chemical analyses, e.g., by MS^n in combination with special CI methods, in order to achieve a more complete chemical characterization. By following generalized biochemical classification rules such as outlined in the hierarchical scheme presented in Figure 2 a significant level of identification of unknown aerosol components may be achieved, even if library patterns are completely unavailable. In previous submissions our team members at SRI have provided a detailed discussion of the role of advanced AI-based classification methods in such situations. The algorithmic approaches described in the previous paragraphs provide the means of optimizing the spectral data for final AI-based evaluation by removing redundant data, deconvoluting overlapping events, subtracting background interferences and maximizing signal-to-noise ratio's.

References

1. Marymont, J.H., and Alexander, H., _Laboratory Medicine_, 11, 10 (Oct. 1980) 664.

2. Morgan S.L. and Fox, A., Paper presented at the 2nd ARO Workshop on Biodection, July 1982, Raleigh, NC.

3. Wold, W., and Sjostrom, M., "SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy" in _Chemometrics: Theory and Application,_ B.R. Kowalski (ed.), American Chemical Society Symp. Series No. 52, Washington, D.C. 1977, pp. 243.

4. Hirschfield, T., _Anal. Chem.,_ 52 (1980), 297A.

5. Meuzelaar, H.L.C., Kistemaker, P.G., Eshuis, W., Boerboom A.J.H., "Automated Pyrolysis Mass Spectrometry, Application to the Differentiation of Microorganisms," in _Advanced in Mass Spectrometry_. Vol. 7B, N.R. Daly (ed.), Heyden and Son, London, 1976, pp. 1452.

6. Meuzelaar, H.L.C., McClennen, W.H., Metcalf, G.S., Hill, G.R., "A Curie-point MS/MS system for Analysis of Biomaterials and Fossil Fuels," Proc. 29th ASMS Conf. on Mass Spectrom. All. Topics, Minneapolis, 1981, pp. 673.

7. Meuzelaar, H.L.C., Haverkamp, J., Hileman, F.D., _Pyrolysis Mass Spectrometry of Recent and Fossil Biomaterials: Compendium and Atlas_, Elsevier Scientific Publishing Co., Amsterdam 1982.

8. Voorhees, K.J., Kunen, S.M., Durfee, S.L., Currie, L.A. and Klouda, G.A., _Anal. Chem._, 53 (1981), 1463.

9. Windig, W., de Hoog, G.S., Haverkamp, J., _J. Anal. Appl. Pyrol.,_ 2 (1981) 213.

10. Richards, J.M., Stolk, H.T., McClennen, W.H., Meuzelaar, H.L.C., "Development and Testing of a Curie-point Pyrolysis Inlet for the Finnigan Ion Trap Detector", Proc. 34th Annual Conf. Mass Spec. All. Topics, 1986, Cincinnati, Ohio, 1071.

11. Meuzelaar, H.L.C., McClennen, W.H., Yun, Y., "A Micro-Volume Curie-point Pyrolyzer for Rapid, High Resolution Capillary GC/MS" Proc. ASMS, Denver, CO, 1987 (in press).

12. Yost, R.A., McClennen, W.H., Meuzelaar, H.L.C., "Enhanced Full Scan Sensitivity and Dynamic Range in the Finnigan MAT Ion Trap Detector with the new Automatic Gain Control Software", Application Report #209, Finnigan Corp.

13. Knorr, F.J., Futrell, J.H., _Anal. Chem._ 51 (1979) 1236.

14. Malinowski, E.R., _Anal. Chim. Acta_, 134 (1982) 129.

15. Windig, W., Meuzelaar, H.L.C., _Anal. Chem._ 56 (1984) 2297.

16. Schulten, H.R., Lattimer, R.P., *Mass Spectrom. Rev.*, 3 (1984) 231.

17. Lattimer, R.P., Harris, R.E., *Mass Spectrom. Rev.*, 4 (1985).

18. Pausch, J.B., Lattimer, R.P., Meuzelaar, H.L.C., *Rubber Chem. Technol.* 56 (1983) 1031.

19. Lattimer, R.P., Schur, K.M., Windig, W., Meuzelaar, H.L.C., *J. Anal. Appl. Pyrol.*, 8 (1985) 95.

20. Kistemaker, P.G., Boerboom, A.J.H., Meuzelaar, H.L.C., *Dynamic Mass Spectrometry*, 4 (1975) 139.

21. Lum, R.M., *Thermochim. Acta.*, 18 (1977) 73.

22. Gardella, J.A., Hercules, D.M., Heinen, H.J., *Spectroscopy Letters*, 13 (1980) 347.

23. Gardella, J.A., Hercules, D.M., *Fresenius Z. Anal. Chem.*, 308 (1981) 297.

24. Israel, S.C., Yang, W.C., Bechard, M., *J. Macr. Sci. Chem.* 22 (5-7) (1985) 779.

25. Biomer is a trademark of Ethicon Inc.

26. Richards, J.M., McClennen, W.H., Meuzelaar, H.L.C., Gregonis, D.E., Reichert, W.M., Helle, M.A., *Macromolecules* 18 (1985) 496.

27. Tatsuoka, M.M., *Multivariate Analysis*, Wiley, New York (1971).

28. Cooley, W.W., Lohnes, P.R., *Multivariate Data Analysis*, Wiley, New York (1971).

29. Windig, W., Kistemaker, P.G., Haverkamp, J., Meuzelaar, H.L.C., *J. Anal. Appl. Pyrol.* 2 (1980) 7.

30. Windig, W., Haverkamp, J., Kistemaker, P.G., *Anal. Chem.* 55 (1983), 81.

31. Windig, W., Meuzelaar, H.L.C., Haws, B.A., Campbell, W.F., Asay, K.H., *J. Anal. Appl. Pyrol.* 2 (1983), 183.

32. van de Meent, D., de Leeuw, J.W., Schenck, P.A., Windig, W., Haverkamp, J. *J. Anal. Appl. Pyrol.* 2 (1982) 56.

33. Metcalf, G.S., Windig, W., Hill, G.R., Meuzelaar, H.L.C., "Characterization of U.S. Lignites by Pyrolysis Mass Spectrometry and Multivariate Analysis" *Coal Geology*, (1987) in press.

34. Windig, W., Jakab, E., Richards, J.M., Meuzelaar, H.L.C., "Self Modelling Curve Resolution by Factor Analysis of a Continuous Series of Pyrolysis Mass Spectra", *Anal. Chem.* 56 (1987) 2292.

69.

35. Huff, S.M., Matsen, J.M., Windig, W., Meuzelaar, H.L.C., <u>Biomed. Mass Spectrom.,</u> 13 (1986) 277.

36. Meuzelaar, H.L.C., Huff, S.M., <u>J. Anal. Appl. Pyrol.,</u> 3 (1981) 111.

37. Windig, W., Chakravarty, T., Richards, J.M., Meuzelaar, H.L.C., "Multivariate analysis of time-resolved Mass Spectral Data" <u>Anal. Chem. Act.</u> (1987) in press.

38. Harman, H.H., <u>Modern Factor Analysis</u>, The University of Chicago Press, Chicago, 1976.

39. Windig, W., McClennen, W.H., Stolk, H., Meuzelaar, H.L.C., <u>Optical Engineering,</u> 25 (1986) 117.

40. Evans, R.J., Milne, T.A., Soltys, M.N., <u>J. Anal. Appl. Pyrol.</u> 9 (1986) 207.

41. Lattimer, R.P., Schur, K.M., Windig, W., Meuzelaar, H.L.C., <u>J. Anal. Appl. Pyrol.</u> 8 (1985) 95.

42. Windig, W., McClennen, W.H., Meuzelaar, H.L.C., "Determination of Fractional Concentrations and Exact Component Spectra by Factor Analysis of Pyrolysis Mass Spectra of Mixtures", <u>Chemometrics and Intelligent Laboratory Systems</u>, (1987) in press.