Journal of Chromatography, 314 (1984) 83–92
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 17,122

THE MASS SPECTROMETER AS A CHLORINE-SELECTIVE CHROMATOGRAPHIC DETECTOR

I. DESCRIPTION AND EVALUATION OF THE TECHNIQUE

J. L. LaBROSSE* and R. J. ANDEREGG*
Department of Chemistry, University of Maine, Orono, ME 04469 (U.S.A.)
(Received July 31st, 1984)

SUMMARY

A computer program to selectively locate the mass spectra of chlorinated compounds in a gas chromatography–mass spectrometry data set is described. The method is based on a search for chlorine isotope patterns and functions as a chlorine-selective "detector". Its performance has been evaluated using the EPA/NIH Mass Spectral Data Base of 31,333 mass spectra. Ninety-five percent of the chlorinated compounds in the library were "found"; and mass spectra of chlorinated compounds were assigned scores which averaged 18 times higher than those of non-chlorinated species. Positive and negative interferences are discussed.

INTRODUCTION

Since the introduction of gas chromatography (GC), the separation of complex mixtures of organic compounds has been greatly facilitated. An increased desire to detect low levels of specific components in these mixtures has led to the development of a variety of selective GC detectors[1]. Chlorine-selective detectors have been a particularly important part of that development. Chlorinated compounds have been found to cause many long-term problems both in the environment and in biological species. Among the more prominent chlorinated pollutants are a large number of pesticides and polychlorinated biphenyls that are found throughout the world[2,3].

The most common chlorine-selective detector presently employed is the electron-capture detector. The electron-capture detector responds to electrophilic compounds; the signal produced is related to the quantity and type of compound detected. Chlorinated compounds are strong electrophiles and produce a relatively large response, even when introduced in sub-picogram quantities[4]. Unfortunately, many non-chlorinated electrophilic compounds also produce large responses in the electron-capture detector and cause a substantial interference in the detection of trace

* Present address: Fein-Marquart Associates, Inc., Baltimore, MD 21212, U.S.A.

0021-9673/84/$03.00

levels of chlorinated compounds. Other chlorine-selective detectors are based on electrochemical[5] or spectroscopic principles[6,7].

It is important to note that these detectors only provide information as to whether or not a molecule contains a chlorine atom. Little, if any, additional structural evidence is provided. The coincidence of chromatographic retention time and the probable presence of chlorine are often the entire evidence for a compound's identification. These are examples of selective data collectors. To gain more structural information about unknown compounds, more spectroscopic evidence must be collected. For example, repetitive scanning gas chromatography–mass specrometry (GC–MS) could be used. In this case, data collection is non-selective: any molecules eluting from the chromatograph are ionized and detected. Because of the tremendous amounts of data generated in a continuously scanning GC–MS experiment, a variety of selective data reduction methods have been employed to limit the amount of data that the analyst must interpret. One approach to selective data reduction is to have the computer search for patterns of ions in the mass spectra, such as are produced by the natural isotopic distribution of chlorine or bromine atoms. Computer programs to locate and identify specific isotope cluster patterns have been described by ourselves and others[8–10]. In this paper, we report a computer program which searches a mass spectral data set successively for each of the possible chlorine-isotope cluster patterns in a range of one to ten atoms per cluster. Each mass spectrum is assigned a score based on the likelihood that it contains chlorine clusters and on the intensity of those clusters. A normalized plot of these scores produces a "chlorine-selective chromatogram", which indicates where in the data set chlorinated compounds are most likely to occur. The selectivity of the program has been evaluated using the U.S. Environmental Protection Agency (EPA)/National Institutes of Health (NIH) Mass Spectral Data Base, a collection of 31,333 mass spectra. Several types of interferences, both positive and negative, were identified, and their potential interference is discussed. A subsequent paper will describe several useful applications of the method.

EXPERIMENTAL

The instrument used in both the computer programming and the library analysis was a Hewlett-Packard 5985 B GC–MS system. The programs were written in BASIC (main program) and FORTRAN (subprograms), and are available through one of the authors (R.J.A.).

The chlorine-selective detector is based on the isotope cluster chromatography algorithm described elsewhere[10]. The spectrum to be searched is broken down into a series of spectral fragments. Each of these spectral fragments which satisfies certain pre-search conditions is compared to the isotope cluster of interest. A similarity index[11], which is a measure of the closeness of a match, is calculated for each frag...ent; 0.0 being a bad match and 1.0 being a perfect match. Although the Biemann similarity index[11] was used in this work, other investigators are evaluating alternative similarity comparisons[12]. If the similarity index of a spectral fragment exceeds a user-set threshold, a contribution to the total score is computed as:

$$S_i = (I_{\text{sim}})^a (A)^b N_{\text{Cl}} \tag{1}$$

where S_i is the score contribution, I_{sim} is the similarity index, A is the abundance of the cluster, N_{Cl} is the number of chlorine atoms present, and a and b are exponents set by the user. When all the contributions from all the relevant spectral fragments are summed, a total score is obtained for the mass spectrum. A normalized plot of those scores *vs.* scan number generates the chlorine-selective chromatogram. We have made several improvements in the program previously described[10]. Rather than calculating the chlorine clusters each time the program is executed, the clusters are now stored at time of compilation and used as needed. In addition, a set of pre-searches checks to be sure important ions of the cluster are present and in approximately the correct proportions before a similarity index is calculated. These improvements eliminate a number of time-consuming calculations and result in about at 67% reduction in the time required for searching over our previous isotope cluster search. The time required for searching is a function of the range of chlorine clusters sought and of the complexity of the mass specrum; but is typically 2.5 sec per spectrum for a search of one of ten chlorine atoms.

To evaluate the performance of the searching routine, scores were calculated for all of the mass spectra in the EPA/NIH Mass Spectral Data Base. This version of the data base contains 31,333 mass spectra of many diverse compounds, including 2350 chlorinated species. Ideally, chlorinated compounds would exhibit mass spectra which would receive high scores. Non-chlorinated compounds would be assigned low scores because their mass spectra would not display the isotope patterns of one or more chlorine atoms.

RESULTS AND DISCUSSION

Fig. 1 shows the isotope cluster patterns that would be expected for 1–10 chlorine atoms. Any mass spectral fragment ion containing chlorine would reflect one of these patterns; and it is these clusters for which a computer search is conducted. Because one does not know *a priori* what other elements may occur with the chlorine (*e.g.*, carbon, hydrogen, nitrogen, oxygen), it is impossible to predict at what masses the cluster patterns will occur. Therefore it would be impossible to detect the clusters by use of mass chromatograms[13] or sets of "subset masses"[14]. The isotope cluster chromatography program, however, can detect these patterns at any mass.

The EPA/NIH Mass Spectral Data Base contains a wide variety of spectra collected on different instruments in different laboratories under different conditions. It was, therefore, a useful collection of mass spectra to use for evaluation of the chlorine searching program. Of the 2350 mass spectra from chlorinated species, 58 were eliminated from consideration because the spectra were (a) abbreviated, and so did not exhibit the isotope cluster patterns they should have; or (b) severely distorted due to instrumental or operating conditions, so the isotope patterns were unrecognizable. These problems did not occur in mass spectra of chlorinated compounds generated on our instrument. The fact that our chlorine selective program was able to help us locate these "errors" in the data base suggests a potentially useful application of the technique in data base evaluation and improvement.

As described in the Experimental section, there are three search parameters over which the user has some level of control: the similarity index threshold and the two values, a and b, the exponents of the similarity and abundance terms, respec-

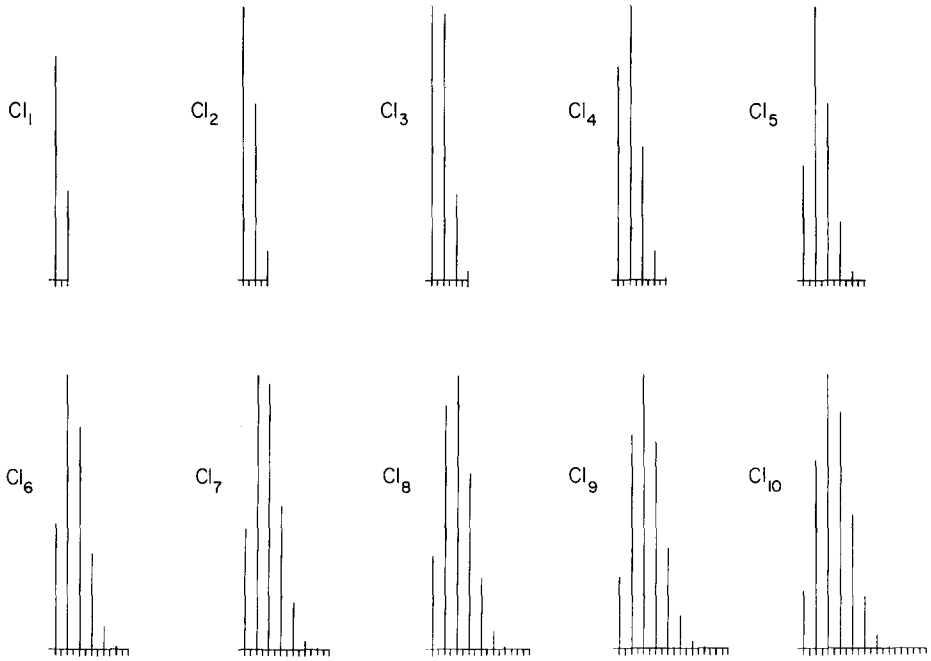

Fig. 1. Isotope-cluster patterns calculated for one to ten chlorine atoms.

tively. By setting the threshold value high (*e.g.*, 0.9), one can eliminate all but the very best matches. Usually this causes the rejection of virtually all non-chlorinated compounds, but also rejects a good many chlorinated species. A low threshold value would be expected to locate all the mass spectra of chlorinated molecules, but would find many "false positives" as well. Clearly, some compromise must be reached. Because one- and two-chlorine cluster patterns are relatively simple, the threshold value was set higher for these clusters than for the more complex patterns of three or more chlorine atoms. Initially a value of 0.85 was set for the Cl_1 and Cl_2 patterns, a value of 0.7 for all other chlorine clusters. With these threshold values, scores were calculated for each mass spectrum in the data base. Of the 2292 chlorinated compounds represented in the data base, 2218 (97%) were assigned non-zero scores. The overall average relative score for mass spectra of chlorinated compounds in this search was 32,600. (These scores only have meaning in a relative sense, *i.e.*, scores calculated using a particular set of search parameters may only be compared to other scores calculated using the same search parameters and the same data set). Of the 28,983 non-chlorinated compounds in the data base, 15,484 (53%) were assigned non-zero scores with an overall mean socre of 4000. If one defines a selectivity factor as the ratio of the means of the scores for chlorinated to non-chlorinated compounds, this corresponds to a selectivity of 8.2 (32,600/4000).

It was suspected that the large number of false finds (non-chlorinated compounds whose mass spectra were assigned a non-zero score) was primarily due to similarities with the simple one-chlorine pattern. This doublet in a 3:1 ratio could happen easily by chance. To confirm this suspicion, a search was made as before, but

without including any Cl_1 contributions. The percentage of non-chlorinated compounds assigned non-zero scores dropped from 53% to 13.5%. Obviously, such a search also eliminates all monochloro compounds and many di- and trichloro compounds; and so this cannot be recommended as a selectivity enhancement unless one knows that only polychlorinated compounds are to be expected in the analysis.

To increase the selectivity toward chlorinated compounds over non-chlorinated compounds, an increase in the similarity index thresholds was investigated. The similarity index thresholds were set to 0.90 for the Cl_1 search, 0.85 for the Cl_2 search, and 0.7 for Cl_3–Cl_{10} searches. This substantially improved the selectivity of the overall search (see below). Additional increases in the similarity index thresholds to greater than 0.9, 0.85, and 0.7 for the Cl_1, Cl_2, and Cl_3–Cl_{10} searches, respectively, decreased the efficiency of locating chlorinated compounds. Therefore, these values were chosen as the best set of values to conduct subsequent analyses. The values of the parameters a and b (eqn. 1) affect the assigned scores in a very predictable fashion. The exponent a of the similarity index allows a greater weighting of the "goodness of fit" aspect of the calculation. If a spectral fragment has a similarity index of 0.9 and a second fragment has a similarity index of 0.7, changing a from 1 to 2 will change the ratio of the scores by a factor of 1.3 (0.9/0.7). As a increases further, this factor will increase exponentially. The net effect of changing a on the overall score is similar to that of changing the similarity index threshold; namely to give greater weight to better matches. Although several values of a were explored ($1 \leqslant a \leqslant 32$), we found it most effective to leave a at 2 and vary the similarity index threshold as described above.

The parameter b (eqn. 1) is the exponent of the abundance term. Increasing this value causes a greater weighting of clusters with a high abundance of ion current. This was found to be a very useful parameter, since it allowed one to discriminate against false positive finds that occurred randomly in the low-abundance baseline "noise". A value of $b = 2$ enhanced the score contribution of real clusters which occurred with appreciable intensity in a spectrum.

The EPA/NIH library was searched a second time using the parameters: similarity index threshold $Cl_1 = 0.9$, $Cl_2 = 0.85$, Cl_3–$Cl_{10} = 0.70$; $a = 2.0$, $b = 2.0$. The overall average score for chlorinated compounds in this search was $2.2 \cdot 10^8$; the overall average score for non-chlorinated compounds was $0.12 \cdot 10^8$. Note that all of the scores are much higher in this search, primarily as a consequence of squaring the abundance terms ($b = 2.0$). As before, the scores are intended to be used in a relative sense, so scores calculated with one set of parameters should not be compared with scores calculated using different search parameters. The selectivity factor of this second search was found to be 18 ($2.2 \cdot 10^8/0.12 \cdot 10^8$). As can be seen by the discussion above, the chlorine selective detector program has an adjustable selectivity, controlled by the choice of values for the similarity index threshold and for the exponents a and b.

It is not surprising that the success of the search improves as the complexity of the isotope cluster pattern increases. The complex patterns arising from 6–10 chlorine atoms are unlikely to occur by chance in a series of unrelated fragment ions. These patterns are so distinctive that they are virtually never missed by the computer. Table I shows a breakdown of the score results based on the number of chlorine atoms present in the molecule. As can be seen, the percentage of chlorinated com-

TABLE I

CHLORINE-SELECTIVE SEARCH OF THE EPA/NIH MASS SPECTRAL DATA BASE

Results as a function of the number of chlorine atoms present. (Search parameters: similarity index thresholds $Cl_1 = 0.9$, $Cl_2 = 0.85$, $Cl_{3-10} = 0.7$, $a = 2$, $b = 2$.)

Number of chlorine atoms	*Number of spectra "found" (score ≠ 0)*	*Number of spectra "missed" (score = 0)*	*% "found"*	*Average score ($\times 10^8$)*
1	1231	73	94	0.89
2	559	26	96	2.1
3	154	10	94	4.3
4	107	3	97	6.2
5	32	1	97	9.3
6	60	0	100	9.6
7	9	0	100	16
8	16	0	100	13
9	6	0	100	18
10	5	0	100	14

pounds "found" (assigned non-zero scores) and the mean of the scores both generally increase as the number of chlorine atoms increases.

There are two classes of interferences for an element-selective "detector" such as the one described above: *negative* interference, that is, chlorine-containing compounds whose mass spectra are assigned low or zero scores and *positive* interference, in the form of non-chlorinated compounds the mass spectra of which are assigned relatively high scores. We have identified two main reasons why chlorinated compounds would be assigned a low or zero score by our searching routine. The first, or Type I, is if the mass spectrum, due to the specific fragmentation of a given compound, does not exhibit a recognizable chlorine isotope cluster pattern. Fig. 2 shows

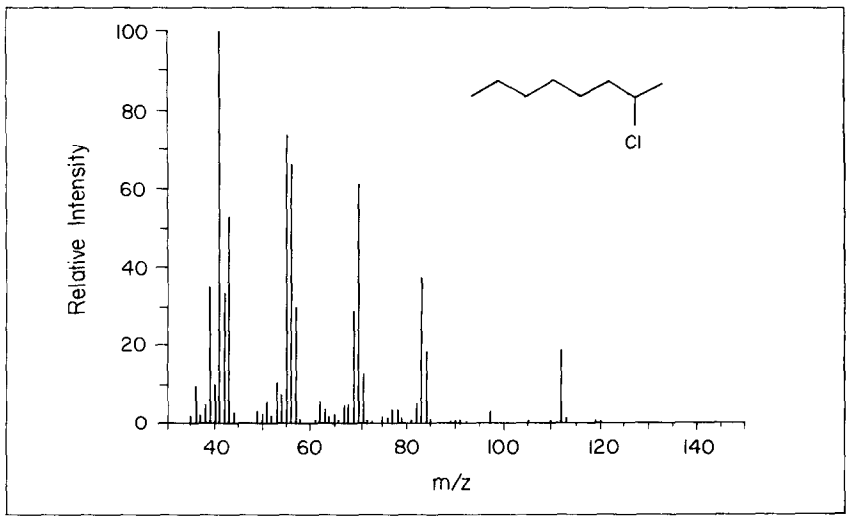

Fig. 2. Mass spectrum of 2-chlorooctane, a Type I interference. Reproduced with permission from ref. 15.

the mass spectrum of 2-chlorooctane. Upon electron impact ionization, the molecule readily loses HCl; and no fragment ions containing chlorine are observed. Since no chlorine isotope patterns are present, the computer would assign this spectrum a score of zero. Of course, a human interpreter would also be unlikely to identify this spectrum as coming from a chlorinated compound. The detection of Type I interferences is frequently enhanced by means of a less energetic ionization method such as chemical ionization. Because fragmentation is suppressed, a chlorinated molecule is more likely to display its true isotope pattern.

A second type of negative interference, Type II interference, can result if the chlorine isotope cluster pattern is distorted by the presence of other elements which display their own isotopic distributions. The mass spectrum of 1-bromo-1-chlorocyclobutane (Fig. 3) serves as an example. The isotope pattern beginning at *m/z* 140 is a result of the natural abundances of ^{35}Cl, ^{37}Cl, ^{79}Br and ^{81}Br. This cluster was assigned a similarity index of zero by the computer due to the low similarity between a normal Cl cluster and this distorted ClBr cluster. Fortunately, this type of interference is less severe than might be expected. Frequently, because of fragmentation of the molecule, some ions are seen that contain the chlorine atom(s), but not the interfering elements as at *m/z* 84 (Fig. 3). These ions will display the undistorted isotope patterns of chlorine; and they will be found by computer. The Type II interferences are only made worse by less energetic ionization, which suppresses fragmentation thereby keeping the chlorine and the distorting elements together in the same molecular fragments.

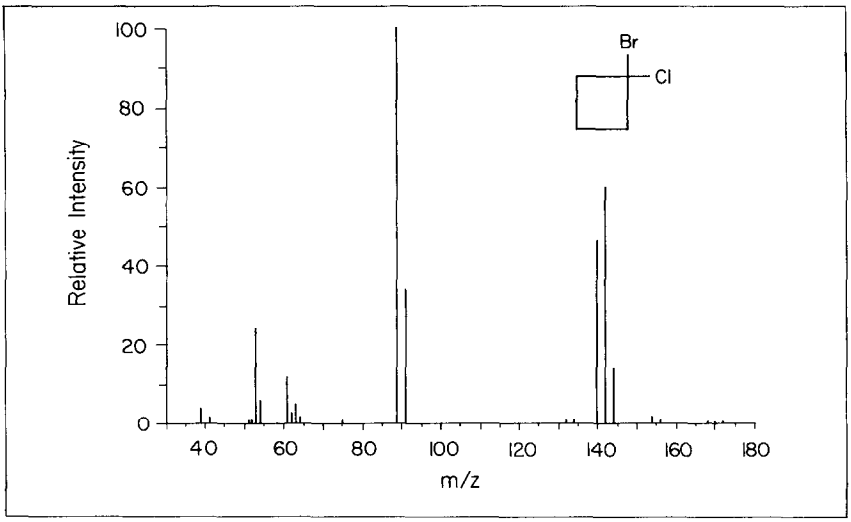

Fig. 3. Mass spectrum of 1-bromo-chlorocyclobutane, a Type II interference. Reproduced with permission from ref. 15.

Our computerized detector exhibits two types of positive interferences as well. In these cases, the mass spectra of non-chlorinated compounds are assigned relatively high scores. Type III interferences occur when, because of the fragmentation of a particular molecule, a group of ions coincidentally occurs in a pattern that resembles

the cluster being sought. An example of such a case is the spectrum of 4,4-dimethyl-3-piperidino-2-cyclobuten-1-one (Fig. 4). Although this compound contains only carbon, hydrogen, oxygen, and nitrogen, there is a pattern of ions beginning at *m/z* 83 which strongly resembles a Cl_2 cluster pattern (inset, Fig. 4). This spectrum was assigned a relatively high score because of a good match with the Cl_2 cluster, and high abundance of the "cluster". Once again, the computer should be forgiven for this error, since a human interpreter could easily fall victim to the same error in the absence of further information. It is only upon further examination of the entire spectrum that one would be likely to rule out the presence of chlorine atoms. As with the Type I interferences, Type III are minimized by utilizing a less energetic ionization method. Suppressed fragmentation results in less likelihood of the chance occurrence of a group of ions which are not isotopically related. Type III interferences were by far the most serious problem in our library evaluation, particulary for groups of ions occurring at low abundance in the spectral noise. As described above, by judicious choice of the abundance exponent parameter, such interferences can be reduced.

Type IV interferences occur when non-chlorinated compounds contain other elements with distinctive isotope distributions. Occasionally, the isotope pattern of these other elements can be mistaken for that of a chlorine cluster. Type IV interferences are quite rare, at least in the systems we have examined. As in the case of Type II interferences, the use of the chemical ionization does not improve the detector's performance toward Type IV interferences. The isotopic patterns responsible for the interferences are not suppressed as fragmentation is reduced. Table II summarizes the types of interferences described above.

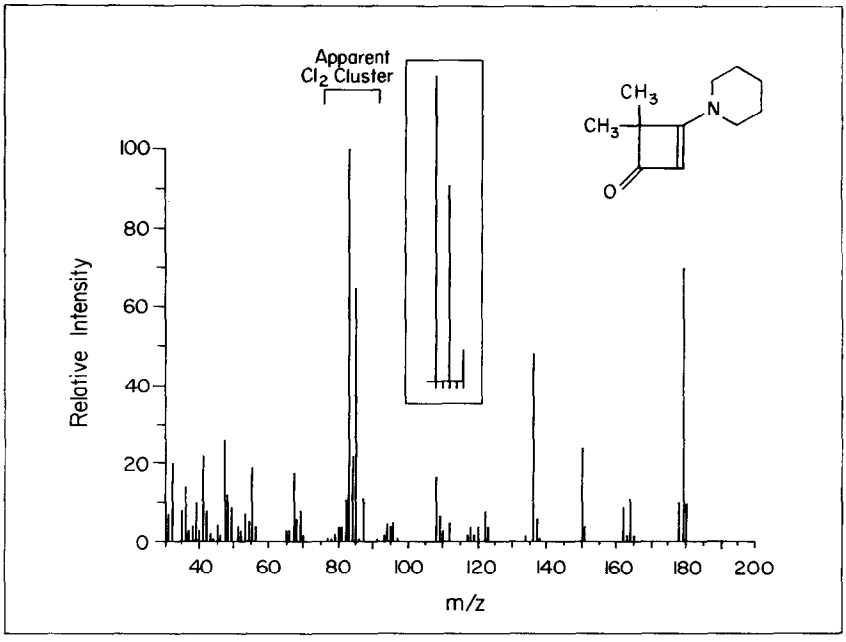

Fig. 4. Mass spectrum of 4,4-dimethyl-3-piperidino-2-cyclobuten-1-one, a Type III interference. Reproduced with permission from ref. 15.

TABLE II

TYPES OF INTERFERENCES IN THIS CHLORINE-SELECTIVE DETECTOR

	Type	*Description*	*Improved by chemical ionization*
Negative (Cl-containing species)	I	No isotope pattern	Yes
	II	Distorted isotope pattern	No
Positive (non-Cl-containing species)	III	Apparent Cl-isotope pattern from unrelated ions	Yes
	IV	Apparent Cl-isotope pattern from other elemental patterns	No

It should once again be emphasized that the purpose of this technique is not to replace element selective chromatographic detectors in all situations; nor is it designed to automatically interpret mass spectra. Its function is to increase the efficiency of analysis of GC–MS data by allowing the analyst to concentrate interpretive efforts on that portion of the data that the computer has deemed most likely to contain chlorine atoms. The final decision about the presence of a chlorine isotope pattern is ultimately made by the analyst.

The "selectivity factors" of this detector also deserve comment. In conventional element selective detectors, selectivity is defined in terms of the ratio of detector response to equal amounts of chlorinated and non-chlorinated material[6]. One must strive for very high selectivity, since the data will either be collected or rejected based on this selectivity. The situation with our data reduction technique is quite different. All of the data are collected and stored regardless of the presence or absence of chlorine. The software detector we describe aids in locating relevant spectra. The analyst may accept or reject the spectra based on subsequent examination of the spectra. In this situation, a much lower degree of selectivity can be tolerated.

Indeed, the whole definition of selectivity factor is sufficiently changed that comparisons are probably meaningless. The response of our detector (the score calculated for a particular mass spectrum) is extremely dependent on compound type. The majority of non-chlorinated compounds receive a score of zero; while the vast majority of chlorinated compounds receive a non-zero score. Taking a ratio of these would lead to a selectivity factor (as defined in the usual sense) that approaches infinity. We have chosen, therefore, to adopt a more conservative estimate, using the averge scores calculated on all compounds in the data base.

CONCLUSION

A chlorine-selective detector has been developed to selectively locate chlorinated compounds in samples analyzed by GC–MS. The computer program is able to search a GC–MS data set for isotope patterns indicative of any number (1–10) of chlorine atoms. Each mass spectrum in the data set is assigned a score, reflecting the likelihood that it contains isotope clusters: chlorine-containing compounds receiving higher scores than non-chlorinated species. In evaluation of the program using the

EPA/NIH Mass Spectral Data Base, the scores assigned to spectra of chlorinated compounds averaged 18 times higher than those of non-chlorinated compounds.

The mass spectra of chlorinated compounds are identified unless the chlorine isotope pattern is missing (due to fragmentation) or distorted by the presence of other elements with their own unusual isotope patterns. The most serious interference is from non-chlorinated species the mass spectra of which display an apparent isotope cluster, even though no chlorine is present.

The use of this program in no way alters the original GC–MS data, so the investigator may perform other data reduction methods, library searches, etc., on the same data. Unlike most other selective GC detectors, full mass spectra are still available for subsequent interpretation.

ACKNOWLEDGEMENTS

The financial support of the National Science Foundation (Grant No. CHE-8209056) is gratefully acknowledged. The spectra in Figs. 2–4 were reprinted with permission from the EPA/NIH Mass Spectral Data Base, courtesy of Dr. L. H. Gevantman.

REFERENCES

1 E. R. Adlard, *CRC Crit. Rev. Anal. Chem.*, 5 (1975) 13.
2 O. Hutzinger and A. A. M. Roof, in J. Albaiges (Editor), *Analytical Techniques in Environmental Chemistry*, Pergamon Press, New York, 1980, p. 167.
3 H. S. Stoker and S. L. Seager, in J. O'M. Bockris (Editor), *Environmental Chemistry*, Plenum Press, New York, 1977, p. 412.
4 J. E. Lovelock, *Anal. Chem.*, 35 (1963) 474.
5 P. T. Holland and R. Greenhalgh, in H. A. Moye (Editor), *Analysis of Pesticide Residues*, Wiley, New York, 1981, p. 51.
6 M. C. Bowman, M. Beroza and G. Nickless, *J. Chromatogr. Sci.*, 9 (1971) 44.
7 J. W. Carnahan, K. J. Mulligan and J. A. Caruso, *Anal. Chim. Acta*, 130 (1981) 227.
8 D. C. Canada and F. E. Regnier, *J. Chromatogr. Sci.*, 14 (1976) 149.
9 I. K. Mun, R. Venkataraghavan and F. W. McLafferty, *Anal. Chem.*, 49 (1977) 1723.
10 R. J. Anderegg, *Anal. Chem.*, 53 (1981) 2169.
11 H. S. Hertz, R. A. Hites and K. Biemann, *Anal. Chem.*, 43 (1971) 681.
12 D. Peterson, Hewlett-Packard Company, personal communication.
13 R. A. Hites and K. Biemann, *Anal. Chem.*, 42 (1970) 855.
14 J. W. Eichelberger, L. E. Harris and W. L. Budde, *Anal. Chem.*, 46 (1974) 227.
15 S. R. Heller and G. W. A. Milne, *EPA/NIH Mass Spectral Data Base*, U.S. Government Printing Office, Washington, DC, 1978.