

CHROM. 9475

GAS CHROMATOGRAPHIC-MASS SPECTROMETRIC-COMPUTER METHODS IN THE STUDY OF DRUG METABOLISM

R. N. STILLWELL and W. G. STILLWELL

Institute for Lipid Research, Baylor College of Medicine, Houston, Texas 77030 (U.S.A.)

SUMMARY

The computer-assisted interpretation of repetitively scanned mass spectra can be of great value in the analysis of a complex mixture even in the absence of a library of reference spectra. We have written a set of interpretation programs for a PDP11/45 laboratory computer. A file structure is used that combines economy of storage with ease of creation and retrieval. User interaction is by a simple command language implemented by an interpreter. The program structure is such that additional functions can be added easily.

INTRODUCTION

The study of the metabolic fate of exogenous compounds, including drugs, is of importance in obtaining information about both the activity and the actual or potential toxicity of these compounds. Of particular interest to us and to others is the obtaining of evidence for the metabolism of unsaturated and aromatic compounds via epoxide intermediates. Epoxidation has been shown to be a common pathway of metabolism, and whereas many epoxides appear to be harmless (at least on limited exposure), some have been implicated in mutagenesis, carcinogenesis, teratogenesis, and cytotoxicity^{1,2}.

Methods have been developed in our laboratory for the isolation, derivatization and gas chromatographic (GC)-mass spectrometric (MS)-computer analysis of metabolites including dihydrodiols, which are produced by the action of epoxide hydase enzymes on epoxides and are the normal excretory products of compounds metabolized by the epoxide pathway. These methods have enabled us to detect a great number of dihydrodiols and hydroxylated dihydrodiols from a variety of compounds³⁻⁸. In this paper we report the computer-assisted analysis program used in this work^{9,10}.

Mini-computers dedicated to MS data acquisition and display are available in many laboratories. Many of these systems provide a great variety of output options such as library searching, reconstruction of mass chromatograms, etc. In general, however, they permit only one function at a time, so that new samples cannot be run while old data are being reviewed. Effective analysis of data obtained by rapid

repetitive scanning of the mass spectrometer requires an interactive system so that the experimenter can combine his knowledge of chemistry and MS, much of which is "intuitive" in that he cannot state it in algorithmic form, with the number-processing capacity of the computer. Such an interactive system requires the passing of large amounts of information in graphic form to the experimenter's terminal at high data rates. The high data rates and large amounts of processing time required make the use of a university or commercial time-sharing system economically infeasible. The decline in cost of computer components, particularly of processors and memory, has, however, made it possible for many laboratories to install central laboratory computers of sufficient capacity to support one or more users running such interactive programs. Our laboratory uses a PDP11/45 computer (Digital Equipment Corp.) for this purpose.

EXPERIMENTAL

Apparatus

Mass spectra were obtained on an LKB Model 9000 gas chromatograph-mass spectrometer interfaced to a PDP12 computer. The interface and programs were those developed by Holmes *et al.*¹¹. A 9 ft. \times 2 mm 1% SE-30 column was used. The computer was fitted with the TC12F option, which allows conversion from the LINCtape format normally used on the PDP12 to DECTape format; the standard conversion program supplied by DEC was modified to copy PDP12 LINCtape to PDP11-format DECTape. Since the basic DECTape word length is 18 bits, of which the PDP11 controller ignores two, one 12-bit LINCtape word is copied into one DECTape word; hence $1\frac{1}{2}$ DECTapes are required for one LINCtape.

The data analysis was performed with a PDP11/45 computer equipped with 64 K words of core memory, memory management, three 1.2 M-word disks, and DECTape. This computer runs under the RSX-11D operating system. The programs used for the analysis are described below. The user's terminal is an Imlac PDS-1 graphic display terminal; it is programmed to simulate a Tektronix 4010 storage scope terminal with a few additional features; the analysis programs do not use these features and hence would run without modification on a Tektronix. A Houston Instruments DP-1 incremental plotter is the hard-copy device; it is interfaced to the PDP11 by the standard DEC XY11 interface and is driven by a handler task that also simulates a Tektronix terminal with the addition of third-order X and Y coordinate control bytes, extending its "display" field to $2^{15} \times 2^{15}$ increments. This allows the data analysis program to use the same code to produce the hard copy as the original display. The plotter output is spooled to disk.

Sample extraction and derivatization

The urine sample was collected from a hospitalized patient 4 days after ingestion of an overdose of methaqualone. The drug and its metabolites were isolated from 20 ml of urine using an ammonium carbonate-ethyl acetate procedure previously described⁸. The ethyl acetate extract was evaporated to dryness under a nitrogen stream, and the residue was dissolved in pyridine and converted to the trimethylsilyl derivatives for analysis by GC and GC-MS-computer.

Data acquisition

After injection of the derivatized sample on to the GC column, the oven temperature was programmed from 150° at 2°/min. Shortly before the first peak of interest (unchanged methaqualone) was due to come off, the switch on the computer was set to start acquiring data automatically. The mass range from m/e 10 to 550 was scanned at a speed setting of 7 (5 sec/scan) at a scan interval of 8 sec. The mass error display was monitored on the scope to guard against mass marker drift. After 260 spectra had been obtained and stored, the tape was copied to two PDP11-format DECtapes.

Analysis programs

Analysis of the repetitive scan data is carried out in two stages. The first is to read the tapes and store the data on disk in a specially constructed file format; this is done by a program called SAVE. The second is the analysis proper, which is carried out by LEXER.

SAVE. This program when loaded queries the user for the input and output file names, etc., then reads the tape directory to find the input file. Reading spectra from the tape is for the most part straightforward decoding of the bitmap and intensity format in which they are stored. An attempt is made, however, to follow the trend of the mass marker in order to correct the integral mass in cases where the sum of the mass marker error and the mass defect exceed 0.5 a.m.u.

Block-inverted file format. SAVE creates in a single pass a file that allows for rapid retrieval of either a given mass spectrum or a given mass chromatogram (Fig. 1). Although the structure of the file is optimized for the PDP11 disk, it is created and read using FORTRAN direct-access READ and WRITE statements of the type that is an almost universal enhancement to ANSI FORTRAN. The file is a direct access file of 256-word blocks, numbered from 0 to N . Block 0 is the index block. Blocks 1 through N are treated as arrays of 16 mass numbers of 16 spectrum numbers; each word contains in its low-order 13 bits the intensity associated with the corresponding mass number and spectrum number. For purposes of exposition, the arrays are described here as if they were numbered from (0,0) to (15,15). In the actual program, subscripts start at (1,1), according to the FORTRAN convention. Mass numbers 0 through 5 are reserved for the total ion current, etc., while mass number 6 corresponds to the least mass stored in the file (usually 40); this number (LOMASS) is placed in the index block along with the total number of blocks (TBLKS) and the total number of spectra (TSPEC) in the file. The file is also considered to be organized into groups of blocks, each group (except for the last) containing 16 spectra. Since spectra vary in length throughout the run, space is saved by terminating each group after all 16 of its spectra are complete. This requires that the index block also contain a pointer to the first block of each group (Fig. 1). Of course, this requires one extra disk access per spectrum or mass chromatogram, or else that the index block be retained in memory during the analysis. In order to access a given mass number or spectrum number, the given number is divided by 16, the quotient being either the starting block number less 1 (for mass chromatograms) or the index into the pointer array (for spectra). The remainder is the row or column, respectively. Fig. 1 illustrates the method of addressing. In practice, of course, since both FORTRAN and users number things starting from 1, not 0, the given number is decremented by 1 before the division, and the quotient and remainder are incremented by 1 afterwards.

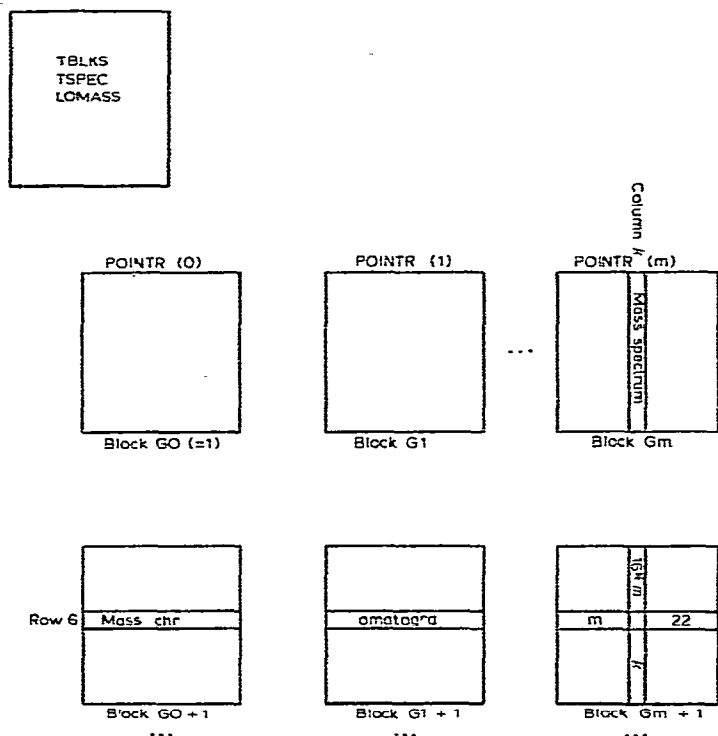


Fig. 1. Block-inverted file format used by LEXER. Each row is a mass chromatogram; each column is a mass spectrum. Block 0, the index block, contains pointers to the beginning of each spectrum group.

Reading a spectrum or a mass chromatogram with a range of 500 mass or spectrum numbers takes about 2 sec with a single user on the system.

A file of this type, with the addition of another file containing compound names, etc., is also in use as a spectrum library. Performing a mass search on this file is equivalent to reading a mass chromatogram from a repetitive scan data file. The library file was not used in the work described here.

LEXER. This program, the interactive analysis program proper, derives its name from the way in which it handles user input. Instead of going through a question-and-answer procedure for each operation, which is all very well for the novice user but soon gets tiresome, the program incorporates an interpreter for a simple command language. Commands are typed in response to the prompt "ENTER COMMAND". A user command consists in its complete form of a keyword (e.g. "DISPLAY") followed by one or more PL/I-like assignment statements (e.g. "spectrum = 147;"). In this article, keyword names are in capitals and parameter names are in italics. Both, of course, are typed in capitals on the keyboard. Parameters which are not to be changed need not be re-entered. Each user's parameters are stored in a file under his account number and hence are available for successive runs of the program. Parameter types include integers, floating point numbers, strings ("mode = normal; file = 'DK1:RUNB';"), and arrays ("mass = 147, 191, 207;"). In general, keywords correspond to program overlays; new functions can be easily added to LEXER. The

principal functions used for interactive data analysis are MAXIMA, DRGC, and DISPLAY.

MAXIMA scans the data file by groups of 16 mass chromatograms and flags local maxima by setting the high-order bit^{9,10}. A similar program has been reported by Biller and Biemann¹². A maximum is defined by the parameters *halfwidth*, *threshold*, and *factor*. A point satisfies the criteria for a local maximum if its intensity exceeds *threshold*, is greater than or equal to that of the preceding spectrum number (at the same mass) and greater than that of the succeeding spectrum number, and is greater than *factor* times the intensity of at least one point within *halfwidth* spectra on either side. In addition, MAXIMA accumulates a total intensity of flagged peaks for each spectrum number; these are stored in one of the reserved mass number rows.

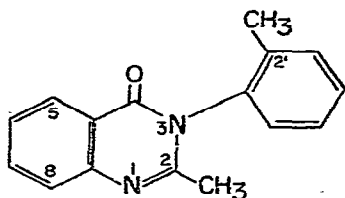
DRGC (for "display reconstructed gas chromatogram") displays on the CRT screen one or more superimposed mass chromatograms. The parameter *mass* determines which mass chromatograms are to be displayed: for example, "DRGC file = RUN22; mass = 191, tic, tmax;" causes the display of the mass chromatogram for *m/e* 191 together with the total ion current trace and the total of the maxima for the file named RUN22. *Tic* and *tmax* are parameters permanently equated to 0 and 1, respectively, and are simply mnemonic acids. Other parameters control the size of the plot and the vertical spacing between traces.

Once a mass chromatogram or spectrum has been requested, the first frame (about 200 masses or spectrum numbers) is shown on the screen. The prompt "#" then invites the user to request a hard copy or to modify the display. Hard copy is requested by typing COPY; the whole spectrum or chromatogram (not merely the current frame) is spooled to disk for later output on the incremental plotter. Other display requests are R: next frame; L: previous frame; U: vertical expansion ($\times 2$); D: vertical contraction; >: horizontal expansion; <: horizontal contraction; X: exit to "ENTER COMMAND". Requests can be ganged: RRRRRUU displays the last (or at least the 6th) frame of the display, expanded vertically by a factor of 4.

DISPLAY causes the display of a single spectrum. *Spectrum*, specifies the spectrum number; *mode* ("normal" or "maxima") specifies whether all peaks or only those flagged are to be displayed; *norm* specifies the peak to which the spectrum is to be normalized ("*norm* = 147;"), or peaks which are not to be used as the base peak ("*norm* = -73, -75;"). Requests to modify the display or produce hard copy are the same as for DRGC.

RESULTS AND DISCUSSION

The techniques described above were applied to a study of the metabolism of methaqualone in the human. Urine was obtained from a patient who had been hospitalized with an overdose of methaqualone. The drug and its metabolites were



Methaqualone.

extracted and derivatized as described previously⁸. Mass spectra were obtained at 8-sec intervals during the relevant portion of the temperature-programmed GC separation. The spectra on tape were converted into a block-inverted file on the PDP11/45 disk as described above, and the maxima were flagged using the program default values for the relevant parameters (the command was "MAXIMA file = 'DK1:M66C';"; the defaulted parameters were equivalent to "halfwidth = 4; threshold = 5; factor = 1.4;"). The total ion current and the total of the maxima were then displayed by giving in turn the commands "DRGC mass = tic;" and "DRGC mass = tmax;". As each display appeared on the screen, the display request "COPY" was given. This caused a plot file to be created and spooled to disk. On termination of LEXER, the plotter was made ready and plotting was initiated; the results are shown in Figs. 2 and 3. The *tmax* plot provides a chromatographic trace of considerably enhanced resolution. One might expect an effect of infinite resolution; *i.e.* a bar graph with each component represented by a single vertical line, since all the ions due to a component should reach a maximum intensity at the same time. That this is not observed is partly because spectra are not taken instantaneously but scanned over a period of time, and hence are biased, and partly due to noise (especially ion statistical noise) in the system, which displaces the maxima of certain ions, particularly the weaker ones.

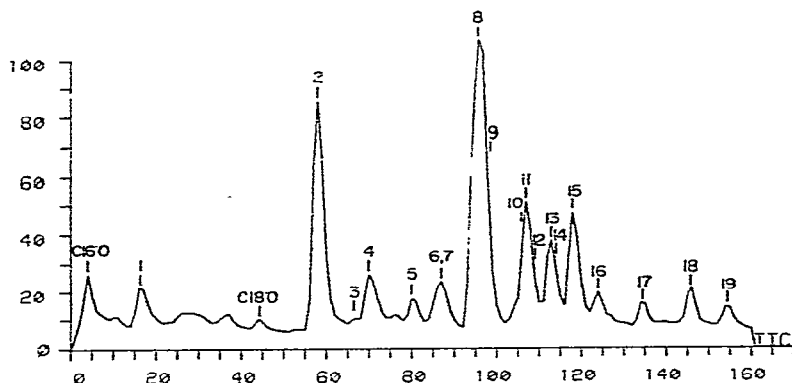


Fig. 2. Total ion current. Numbers above the trace indicate the position of elution of methaqualone and its metabolites, as listed in Table I. C16:0 and C18:0 are trimethylsilyl palmitate and trimethylsilyl stearate, respectively.

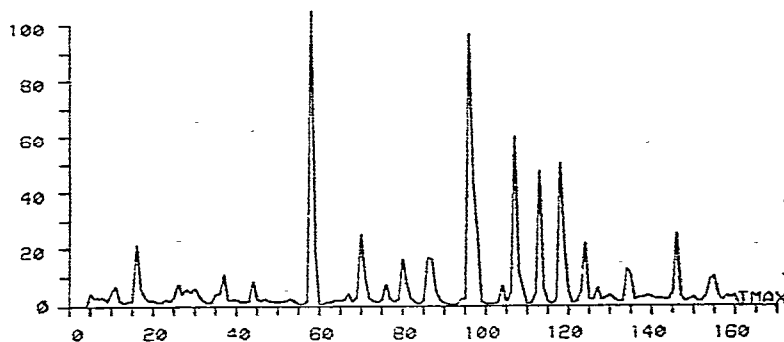


Fig. 3. Total of flagged peaks. The effect is that of a chromatogram of markedly enhanced resolution.

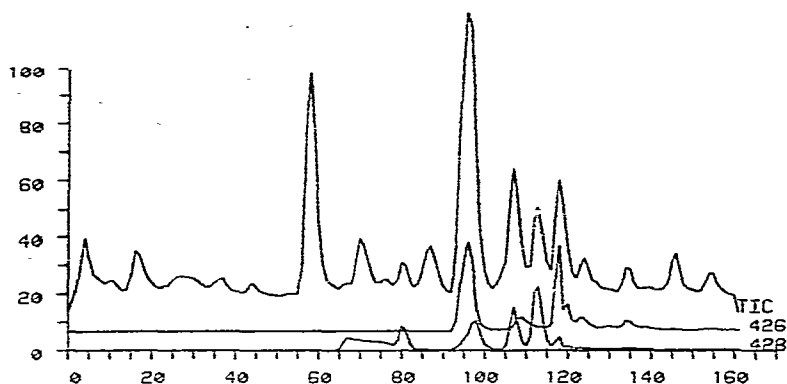


Fig. 4. Mass chromatograms for m/e 426 and 428 and total ion current. The peaks at spectrum numbers 97 and 107 contain unresolved mixtures. The peak at m/e 426 corresponds to M^+ for di-OTMS derivatives of methaqualone, m/e 428 to M^+ for dihydro-di-OTMS derivatives.

Fig. 4 shows the reconstructed mass chromatograms for m/e 426 and 428, which are the molecular weights of trimethylsilyl (TMS) derivatives of diol and dihydrodiol metabolites, respectively. A peak in the mass chromatogram is not, of course, proof of the presence of the compound of corresponding molecular weight, as the ion may be a fragment ion (inspection of spectrum 135 showed that m/e 426 at that point was due to trimethylsilanol elimination from a hydroxydihydrodiol derivative, for example) or an isotope peak (the peak at m/e 428 at spectrum 117 is simply due to the second isotope peak of m/e 426). In most cases, inspection of the spectrum is sufficient to confirm or refute the assignment. Where the chromatographic peaks overlap, however, the evidence of maxima for other ions must be used. Fig. 5 shows that m/e 191 has maxima corresponding to the four principal peaks for m/e 428. The peak at m/e 191 corresponds to the well known ion of structure $\text{TMSO}-\text{CH}=\overset{\oplus}{\text{O}}\text{TMS}$, which is diagnostic for trimethylsilyloxy groups on vicinal saturated carbon atoms. More complete information, and in many instances virtually

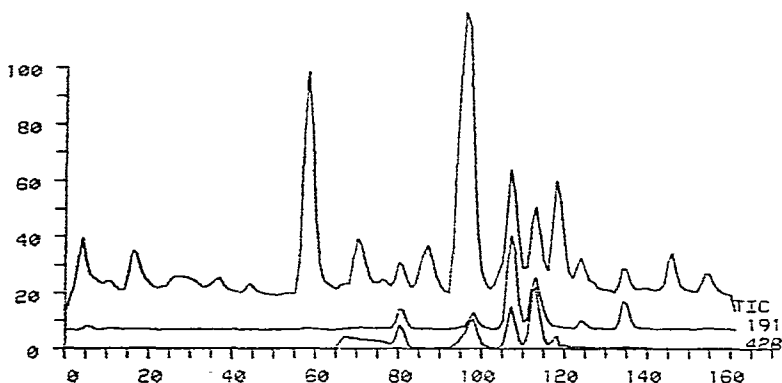


Fig. 5. Mass chromatograms for m/e 191 and 428 and total ion current. The peak at m/e 191 confirms the presence of a vicinal diol TMS structure in each of the four major compounds possessing a peak at m/e 428.

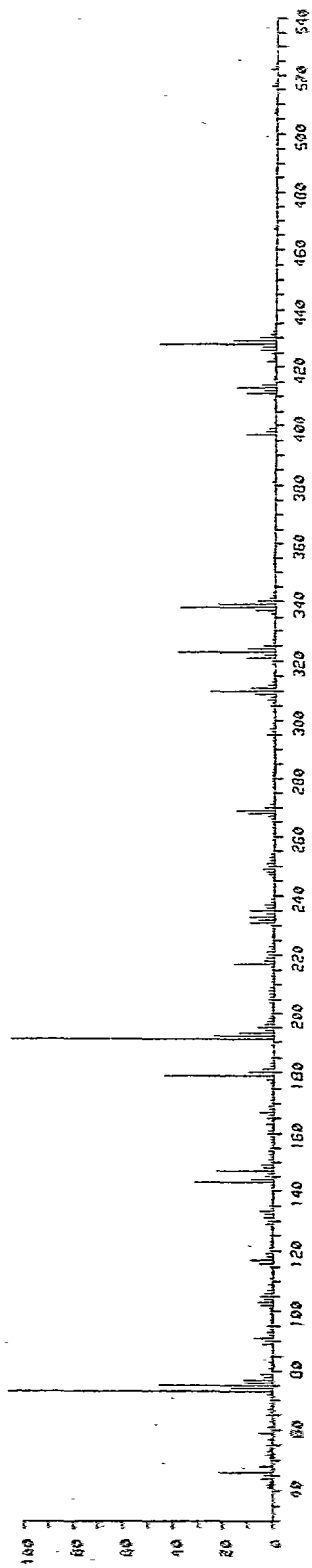


Fig. 6. Mass spectrum at scan 107. The presence of peaks at m/e 411 and 413 and at m/e 426 and 428 suggest the presence of a mixture of methaqualone diol TMS and dihydrodiol TMS derivatives.

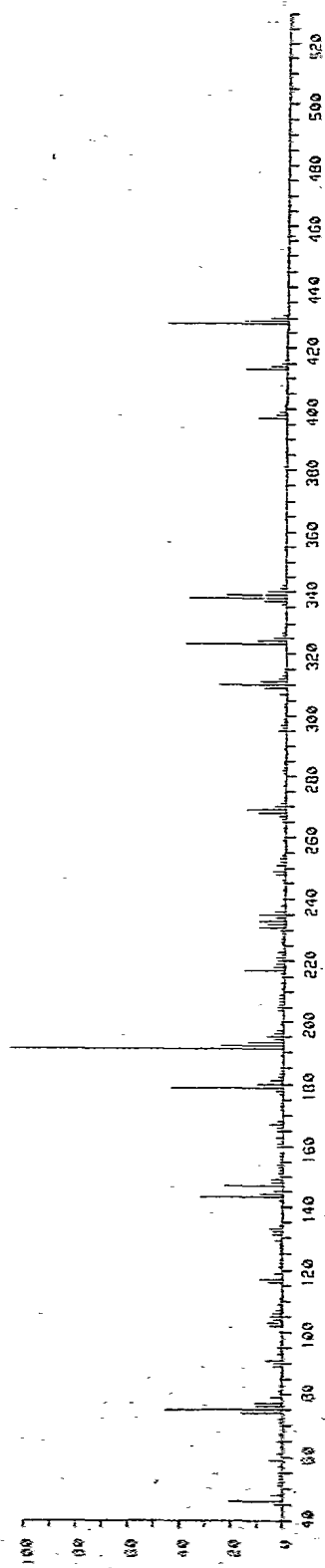


Fig. 7. Reduced spectrum at scan 107. A clean spectrum of a methaqualone dihydrodiol TMS derivative. The base peak at m/e 191 shows that the -OTMS groups are vicinal. The peak at m/e 143 shows that the tolyl ring has been oxidized.

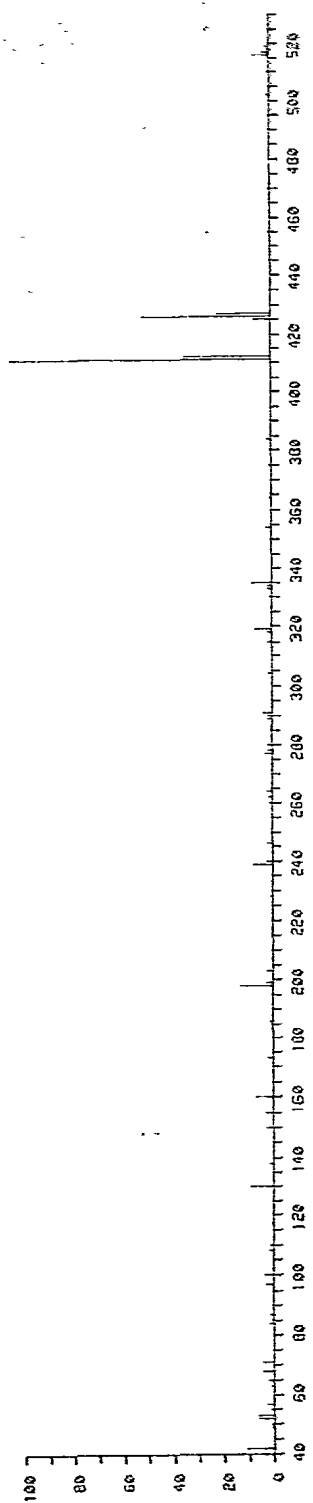


Fig. 8. Reduced spectrum at scan 109. The molecular ion shows up at m/e 516, indicating that the compound is the TMS derivative of a hydroxydihydrodiol metabolite of methaqualone. The peaks at m/e 426 and 411 represent loss of trimethylsilanol and loss of trimethylsilanol and methyl, respectively, from the molecular ion. This compound has not previously been reported.

complete separation of the spectra of overlapping compounds, may be obtained by plotting the spectra of only the flagged masses. Fig. 6 is the normal spectrum number 107, corresponding to the top of the peak in the total ion current chromatogram; clearly it contains more than one compound. It was obtained from LEXER by the command "DISPLAY spectrum = 107;". The reduced spectrum at that point was obtained by typing "DISPLAY mode = maxima;". The parameter *spectrum* was unchanged and so did not need to be repeated; the word "maxima" was not a legal parameter (although it is a keyword), and therefore was accepted by the program simply as a character string. The resulting spectrum is shown in Fig. 7. It is essentially the spectrum of the pure dihydrodiol TMS derivative. The command "DISPLAY spectrum = 109;" produced the reduced spectrum of the diol TMS derivative (Fig. 8). Since this is the minor component of the mixture, many fragments that it has in common with the major component disappear since they peak at the elution time of the major component. However, the spectrum is clearly that of the TMS derivative of a hydroxydihydrodiol metabolite of methaqualone. It should, in theory, be possible to reconstruct each spectrum completely by deconvoluting each mass chromatographic peak, perhaps using an ion unique to each component to determine the peak shape. In biological work, however, where the sample size is small and the scan time must be kept fairly long relative to the peak width in order not to lose sensitivity, it would be difficult to gain much additional information with any reliability.

The components which could be identified in this mixture are indicated by numbers in Fig. 2 and listed in Table I. With the exception of compounds 10 and 12 all of the metabolites in the table have been previously reported^{8,13,14}.

TABLE I

METHYLENE UNIT (MU) VALUES AND MOLECULAR WEIGHTS (TMS DERIVATIVES) OF METHAQUALONE AND ITS METABOLITES

Numbers refer to the peaks in Fig. 2. Hydroxyl groups labeled as 2 and 2' are on the methyl groups at those positions. Q and T refer to positions on the quinazolinone and tolyl nuclei, respectively.

No.	Compound	MU 5% SE-30	MW
1	Methaqualone	21.1	250
2	2'-Hydroxy	23.3	338
3	Q-Dihydrodiol	23.8	428
4	8-Hydroxy	23.9	338
5	T-Dihydrodiol	24.5	428
6	3'-Hydroxy	24.7	338
7	4'-Hydroxy	24.7	338
8	2,T-Diol	25.1	426
9	Q-Dihydrodiol	25.1	428
10	Hydroxydihydrodiol	25.6	516
11	T-Dihydrodiol	25.6	428
12	Hydroxydihydrodiol	25.7	516
13	T-Dihydrodiol	25.9	428
14	T-hydroxy-T-methoxy	25.9	368
15	2',T-Diol	26.1	426
16	2'-Hydroxy-T-dihydrodiol	26.5	516
17	2'-Hydroxy-T-dihydrodiol	26.9	516
18	Mixed triols	27.5	514
19	2,T,T-Triol	28.0	514

ACKNOWLEDGEMENTS

This work was supported by Grant GM-16216 from the National Institutes of Health. We thank Stephen S. Bates for programming assistance and Carol S. Myron for running the mass spectrometer.

REFERENCES

- 1 D. M. Jerina and J. W. Daly, *Science*, 185 (1974) 573, and references therein.
- 2 P. G. Wislocki, A. W. Wood, R. L. Chang, W. Levin, H. Yagi, O. Hernandez, D. M. Jerina and A. H. Conney, *Biochem. Biophys. Res. Commun.*, 68 (1976) 1006.
- 3 M. G. Horning, C. Stratton, A. Wilson, E. C. Horning and R. M. Hill, *Anal. Lett.*, 14 (1971) 537.
- 4 D. J. Harvey, L. Glazener, C. Stratton, D. B. Johnson, R. M. Hill, E. C. Horning and M. G. Horning, *Res. Commun. Chem. Pathol. Pharmacol.*, 4 (1972) 247.
- 5 D. J. Harvey, L. Glazener, C. Stratton, J. Nowlin, R. M. Hill and M. G. Horning, *Res. Commun. Chem. Pathol. Pharmacol.*, 3 (1972) 557.
- 6 W. G. Stillwell, M. Stafford and M. G. Horning, *Res. Commun. Chem. Pathol. Pharmacol.*, 6 (1973) 579.
- 7 W. G. Stillwell, *Res. Commun. Chem. Pathol. Pharmacol.*, 12 (1975) 25.
- 8 W. G. Stillwell, P. A. Gregory and M. G. Horning, *Drug Metab. Dispos.*, 3 (1975) 287.
- 9 R. N. Stillwell, *22nd Annual Conference on Mass Spectrometry and Allied Topics, Philadelphia, Pa., 1974*, paper V-11.
- 10 W. G. Stillwell, P. A. Gregory and R. N. Stillwell, *23rd Annual Conference on Mass Spectrometry and Allied Topics, Houston, Texas, 1974*, paper J-4.
- 11 W. F. Holmes, W. H. Holland and J. A. Parker, *Anal. Chem.*, 43 (1971) 1806.
- 12 J. E. Biller and K. Biemann, *Anal. Lett.*, 7 (1974) 515.
- 13 R. Bonnichsen, C.-G. Fri, C. Negoita and R. Ryhage, *Clin. Chim. Acta*, 40 (1972) 309.
- 14 R. Bonnichsen, Y. Marde and R. Ryhage, *Clin. Chem.*, 20 (1974) 230.