Journal of Chromatography, 290 (1984) 339–350
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 251

THE APPLICATION OF COLACHROM, A NEW COMMAND LANGUAGE FOR CHROMATOGRAPHIC DATA PROCESSING

ENGELBERT ZIEGLER* and GERHARD SCHOMBURG
Max-Planck-Institut für Kohlenforschung, Mülheim/Ruhr (F.R.G.)

SUMMARY

The COLACHROM command language is tailored to the needs of interactive processing of chromatographic data. Single-line commands operate on "raw data" files of unreduced digitized chromatograms as well as on report data files resulting from a previous peak analysis. Examples of the application of this command language are given.

INTRODUCTION

At the Max-Planck-Institut in Mülheim, computer processing of chromatographic data has been in routine usage since 1969. At that time all instruments were connected directly to a DEC10 time-sharing computer system that was operated as the Institute's computing centre. The type of computer used and its operation provided the background for the development of very powerful and flexible software for laboratory data processing, especially for chromatography[1,2]. For the special environment of the Institute this approach was the most economical one. During the lifetime of the DEC10 system *ca.* 40,000 chromatograms, covering a large variety of analytical problems, had been processed annually.

This first computer system was in operation at Mülheim until 1981. Over the years the evolution of mini and micro computers, accompanied by a much improved price/performance ratio, meant that the original advantages of the old system became outdated. Therefore, in 1981, it was replaced by a new system[3] (Fig. 1) which is structured hierarchically. The central hardware, which is again the computing centre of the institute, consists of two VAX11/780 machines, so-called supermini-computers. In this system the real-time tasks, data sampling and instrument control, are delegated to local mini-computer satellites in centralized chromatography laboratories. For insulated, decentralized chromatographs a local analog-to-digital (A/D) converter with buffer memory (Nelson 861) connected via a RS232 serial line to the central system is used. The evaluation of the acquired data and especially all interactive processing, however, are tasks for the central VAX system. All dialog terminals are connected to it.

REQUIREMENTS FOR CHROMATOGRAPHIC DATA PROCESSING

The following major characteristics are required from a computer system for chromatography:

0021-9673/84/$03.00

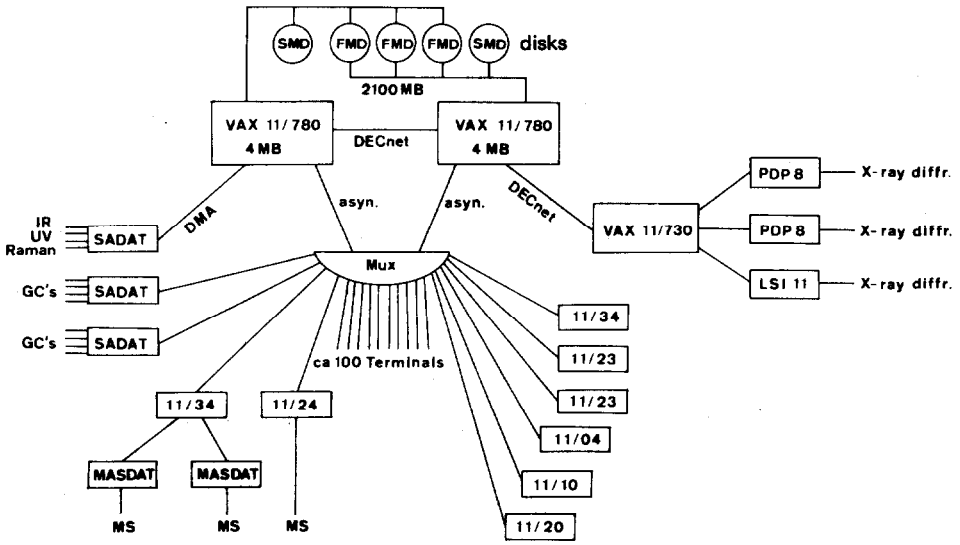

Fig. 1. Hierarchically structured system: mini computer (*e.g.* SADAT- and MASDAT) satellites and *ca.* 100 terminals are connected to the central VAX11/780-machines. Connections are via Direct-Memory-Access (DMA) or asynchroneous serial (multiplexed) lines.

(1) The computer system should not limit the capabilities of the chromatographic instrumentation. Most chromatographic detectors can handle a large dynamic range of signal intensities and therefore require appropriate digitizing hardware. The selected A/D hardware provides for signal voltages between 2.5 μV and 10 V by using automatic gain ranging. No manual preselection of amplification ranges is necessary. Besides providing a large dynamic range for relative signal heights within a chromatogram, the need to repeat chromatograms due to wrong amplification ranges or to improperly chosen injection volumes is avoided.

(2) Repetition of chromatographic separations consequent upon improper choise of data evaluation parameters is not acceptable. In order to avoid such repetition all digitized chromatograms ("raw data") are retained in an external memory (disk storage) in complete, unreduced form. No real-time reduction of any kind is applied. For this reason the peak search and area determination for a once-recorded chromatogram may be repeated with complete freedom of selection of modified peak analysis parameters and without requiring a repetition of the separation.

(3) The software should be able to handle a diversity of analytical tasks originating from many areas of chemistry. Therefore flexible, interactive methods for the processing of chromatograms and the preparation of customer reports have to be provided.

(4) No computer expertise should be asked of the laboratory staff. Even unskilled personnel should be able to use the computer system effectively. Therefore, a good human interface with the software system is essential. Comfortable hardware and software for interaction with the computer system are necessary.

(5) The data evaluation process should be automated as far as possible. On the other hand the user in the laboratory should also be able to execute each step of the data evaluation process interactively. Furthermore, the interactive mode has to

be flexible enough to be adaptable to chromatographic data evaluation and report preparation for a variety of analytical problems as is encountered in a large chemistry research institute. For this purpose COLACHROM, an interactive command language for chromatography, has been developed.

ORGANIZATION OF THE COLACHROM SYSTEM

The data acquisition part of the system has been described elsewhere[3,4]. A few seconds after the end of a chromatogram the analyst will find the result of an automatic evaluation step in two data files stored on his part of the disk memory: one file containing the unreduced "raw data" and another containing extended report data (comprising peak-specific data, such as positions, integration limits, peak areas, heights, half widths, etc., and supplementary information, such as time of measurement, date, sample identification, noise, dynamic range, etc.) generated according to proviously specified peak detection parameters. Furthermore, a peak list will be printed by a ("spooled") laboratory printer.

The analyst may then initiate the COLACHROM system for interactive processing through his dialog terminal. COLACHROM reads the report data file and/or the unreduced data file of the chromatogram to be processed into its internal data section. The latter is subdivided into a main buffer, where the final report is prepared, and several "compare-buffers" that may hold either additional chromatograms from which partial information can be extracted, or tables of reference information. Formally, no distinction is made between genuine chromatographic reports and any tables or artificial reports generated interactively. The structure of the report data file is independent of its contents and its previous processing.

A typical COLACHROM command consists of a single line of input typed by the user at his terminal. The command specifies the kind of action to be executed together with objects, *e.g.*, peaks, and values assigned to objects. Most of the existing commands operate either on the report data or the ADC data; some of them, *e.g.*, in connection with graphical presentation, operate on both types of data.

In general a command line consists of four different fields: command name/buffer/object/value.

Any of the command names may be abbreviated by truncating letters as long as the remaining keyword is unique. (The first three letters of a command name are always unique.) The second field specifies the number of the compare-buffer containing the data to be operated on and is omitted for operations in the main buffer. The object field normally identifies (by sequential numbers or by time ranges) peaks or ADC data points to be treated by the command. The value field assigns numerical values or character-strings to objects. (In special cases, however, the terms "object" and "value" may be interpreted differently.)

The following examples illustrate typical commands:

```
FACTOR 5 = 1.22
```

assigns response factor 1.22 to peak 5.

```
NAME/1 13 = CYCLOHEXANE
```

assigns a compound name to peak 13 of compare-buffer 1.

```
PRANGE 7, 10, 14–20, 23
```

restricts further processing to the specified list of peaks.

SUPPRESS 16, 17

excludes peaks 16 and 17 from further processing.

THRESHOLD = 3%

defines a threshold for report printout.

One of the more complicated commands is

SEARCH/2 INDEX/0.5%

resulting in a comparison of Kováts indices in the report contained in the main buffer with the table of indices in compare-buffer 2. Matches are found if indices are identical within a window of 0.5%. Omission of the "%" would result in an absolute window of 0.5 index units.

CANCEL IND

will erase all indices obtained from previous actions.

CANCEL 12–16/IND

will erase only indices for peaks 12 to 16.

GROUP 1 = 4–8

combines the areas of peaks 4–8 in group number 1.

SAMPLE = fraction 3 of E705

specifies a sample ID to be printed in the report header.

MOVE/1 4 = 7

copy peak 7 of compare-buffer 1 into peak 4 of the main buffer.

DISPLAY

generates a plot of the raw data chromatogram on the display screen.

TABLE I

AVAILABLE COLACHROM COMMANDS FOR REPORT PREPARATION, PEAK SELECTION AND GENERAL INFORMATION

SELECT	Read raw data and report data from disk
NSELECT	Read next chromatogram within a series
THRESHOLD	Set percentage threshold for peak areas
VTHRESHOLD	Set voltage threshold for peak areas
SUPPRESS	Suppress peaks from further processing
PRANGE	Limit processing to specified peak no.s
MRANGE	or to retention time range
PURGE	Retain *n* largest peaks only
GROUP	Combine peaks into group
MERGE	Merge two reports into one
SAMPLE	Assign sample ID
SIGEL	Assign customer ID
NAME	Assign compound name to peak
CANCEL	Invalidate a previous action
RESET	Clear all data buffers
INFO	Type out general information on buffer contents
HELP	Type help information on specified commands
HILFE	Type help information in german language
GRAPHICS	Provides information on use of the cross-hair cursor
WINDOWS	Information on graphical windowing
NEWS	Latest news for chromatography laboratory
VERSION	Type out current version no. of COLACHROM
EXIT	Exit from COLACHROM, saving modified data
QUIT	Exit without saving modified data

TABLE II

QUANTITATIVE EVALUATION AND RETENTION TIMES

STANDARD	Quantitative analysis with added reference
ISTAND	Quantitative analysis with internal reference
EXSTAND	Quantitative analysis with external reference
EXNORM	Quantitative analysis with standard addition method
PSTAND	Normalization relative to specified peak
FACTOR	Assign response factor to one or more peaks
RESPONSE	Calculate response factor relative to reference
INDEX	Calculate Kováts indices
LINDEX	Calculate linear indices
RRTIME	Calculate retention times relative to reference peak
METHANE	Specify methane peak for dead-time calculation
PETERSON	Dead-time after Peterson–Hirsch
PLATES	Calculate effective and theoretical plates, HETP and k' values
SEPARATION	Separation numbers for specified n-alkanes

DIAGRAM HETP/K'

generates a diagram of HETP values *versus* k' values.

SPLIT 3/30, 30, 40

subdivides the display screen vertically into three plotting areas with relative heights of 30, 30 and 40 for the subdivisions.

TABLE III

PEAK ANALYSIS AND OTHER OPERATIONS ON RAW DATA

ANALYZE	Peak analysis based on raw data file
NOISEFACTOR	Peak detection sensitivity coupled to average noise
BASEPAR	Baseline detection parameter
DRIFTPAR	Allowable drift between baseline sections
SOLVENTSUB	Subtract solvent peak for peak analysis
ISOTHERM	Baseline detection as function of type of chromatogram
LCDAT	Converts "negative" LC peaks into "positive" ones
TANGENT	Tangent method for shoulder peaks
ANBEGIN	Position where peak search is to be started
ANEND	Position where peak search is to be ended
NULL	Redefine baseline sections (also by cross-hair cursor)
REANALYZE	Peak analysis with user-defined baseline
CRUDE	Simulated distillation according to ASTM D2887
ADD	Sum two raw data curves
SUBTRACT	Subtract raw data curve
XNULL	Extract reconstructed baseline as artificial chromatogram
SMOOTH	Smooth raw data with specified filter functions
FILTER	Apply exponential filter function for smoothing
DERIVATIVE	Calculate derivative of raw data curve
AUTOCORR	Autocorrelation of raw data
CROSSCORR	Cross-correlation of raw data
FOURIER	Fourier transformation of raw data
SIMULATE	Simulate peaks (profiles: Gauss, Lorentz, tailing)
ADC	Read or modify single data points
RNOISE	Overlay raw data with random noise.
ZERO	Erase all raw data

TABLE IV

COMPILATION AND COMPARISON OF TABLES

CREATE	Create a table or a command procedure
PINDEX	Assign index value to table entry
FACTOR	Assign response factor
AREA	Assign absolute area
MOVE	Copy peak from report into table
NMOVE	Copy with peak area corrected relative to reference
VARIAN	Calculate average values and standard deviation
SEARCH	Compare peak list of report with other report or table
SSEARCH	Suppress peaks not matched in search
PROCEDURE	Manipulate command procedure
@	Execute the command procedure stored on disk

ANALYZE

evaluates the raw data in the main buffer, resulting in a peak list.

The existing commands are summarized in Tables I–V. (A more detailed description of the available commands is given in refs. 4 and 5.) They cover a large variety of applications and allow for variations in working practices. Laboratory practice has shown that normally a given user concentrates on a subset of commands relevant to his special needs. Nevertheless, the large number of commands necessitates on-the-spot availability of help in form of explanations. Therefore a HELP command provides approximately one display screen of explanatory text to each command.

EXAMPLES FOR THE APPLICATION OF COLACHROM

Some typical applications of the COLACHROM language are illustrated by the following examples.

Quantitative analysis

Several methods of quantitative evaluation are implemented as distinct

TABLE V

REPORT OUTPUT AND GRAPHICAL PRESENTATIONS

TYPE	Type out current report to user terminal
REPORT	Print report on spooled laboratory printer
PRINT	Print report on computing centre printer
STORE	Store report on disk memory
SAVE	Dump contents of internal buffers to disk
PRECISION	No. of digits of fixed-precision output
DISPLAY	Graphical presentation of raw data
DDISPLAY	Draw raw data as sequence of single points
SPLIT	Split display screen into vertical sections
SCREEN	Specifies section for next DISPLAY command
REFRESH	Draw fresh picture of original chromatogram
RETURN	Return from graphics dialog into normal dialog
DIAGRAMM	Draw diagram of specified report information

COLACHROM commands: if one or more standard compounds have been added to the sample, the peak numbers and weights are specified through the STAND command. If the reference compound is to be part of the sample, the command ISTAND is used; an external standard is characterized through the EXSTAND command, the "standard addition method" through EXNORM. A corresponding command sequence could be as follows (all user input being in capitals or underlined):

Example 1:

```
SELECT 102                    ; read last report and raw data file from instrument 102
TYPE                          ; type out present contents of report (type out not shown here)
ISTAND 14/3.52 mg             ; internal standard calculation with peak 14 as reference
  weight of sample: 105.6     ; user is asked for unknown sample weight
TYPE                          ; check result
REPORT                        ; print result onto central reporting device of the laboratory
```

In the case of the "standard addition method" the sample chromatogram and the chromatogram which contains an additional amount of the interesting component are loaded into the main buffer and a compare-buffer respectively:

Example 2:

```
SELECT 102/5                                ; load data from run 5 on instrument 102 into main buffer
SELECT/1 102/6                              ; run 6 into compare-buffer 1
EXNORM                                      ; enter dialog:
  Reference peak in main report: 3          ; correct for irreproducible
                 in com report: 3           ; injection volume via reference peak present in main chromatogram and in compare chromatogram
  Peak of interest in main report: 12       ; component to be determined
                   in com report: 13
  Additional weight in com report: 12.4 mg
  Resulting weight for peak 12 in main report is 9.37 mg
```

Calculation of Kováts indices

The calculation of Kováts indices requires net retention times; therefore, the dead-time must be determined, either by specifying the methane peak (METHANE command) or by a Peterson–Hirsch calculation (PETERSON).

Example 3:

```
SELECT 105                          ; read data files from instrument 105
TYPE                                ; look at current contents of report
METHANE = 1                         ; identify methane peak
INDEX 7, 12, 18/600, 700, 800       ; identify reference peaks
TYPE                                ; look at result
THRESH = 0.5%                       ; exclude small peaks from printout
REPORT                              ; send report to laboratory printer
```

"Simulated distillation"

Simulated distillation for petroleum fractions according to the ASTM "Proposed Test Method for Boiling Range Distribution of Crude Petroleum by Gas Chromatography" is possible by loading the sample chromatogram (raw data and peak list from a normal peak search) into the main buffer and the corresponding calibration chromatogram (*n*-alkanes) into a compare-buffer. The CRUDE command identifies the *n*-alkane peaks of the latter, subtracts the extracted baseline from the sample chromatogram and determines eluted percentages as a function of the temperature by continuously integrating the raw data of the sample chromatogram and interpolating, the boiling temperatures of the *n*-alkanes. For the case of incompletely eluting samples a known amount of an internal standard may be added.

Example 4:

```
SELECT/1 118/101                    ; load run 101 (n-alkanes) of instrument 118
                                      into compare-buffer 1
SELECT 118/102                      ; load sample chromatogram into main buf-
                                      fer
TYPE                                ; look at peak list
TYPE/1                              ; peak list of n-alkane chromatogram
METHANE = 4                         ; define common time scale via artificial
METHANE/1 = 3                       ; net retention times
SUPPRESS 4                          ; discard solvent peak
CRUDE/1                             ; perform actions described above
  Specify peak no., carbon no.: 29, 12  ; n-C12 peak and
  Specify peak no., carbon no.: 46, 14  ; n-C14 peak for calibration
  sample weight (mg): 1225          ; for the case of incomplete
  standard weight: 20.5               elution a standard is introduced
  peak no. of standard: 8
  its respective factor: 1
DISPLAY CRUDE                       ; display sample chromatogram with tem-
                                      peratures and percentages of eluted mate-
                                      rial as a function of temperature (see Fig.
                                      2).
REPORT CRUDE                        ; print report onto laboratory printer
```

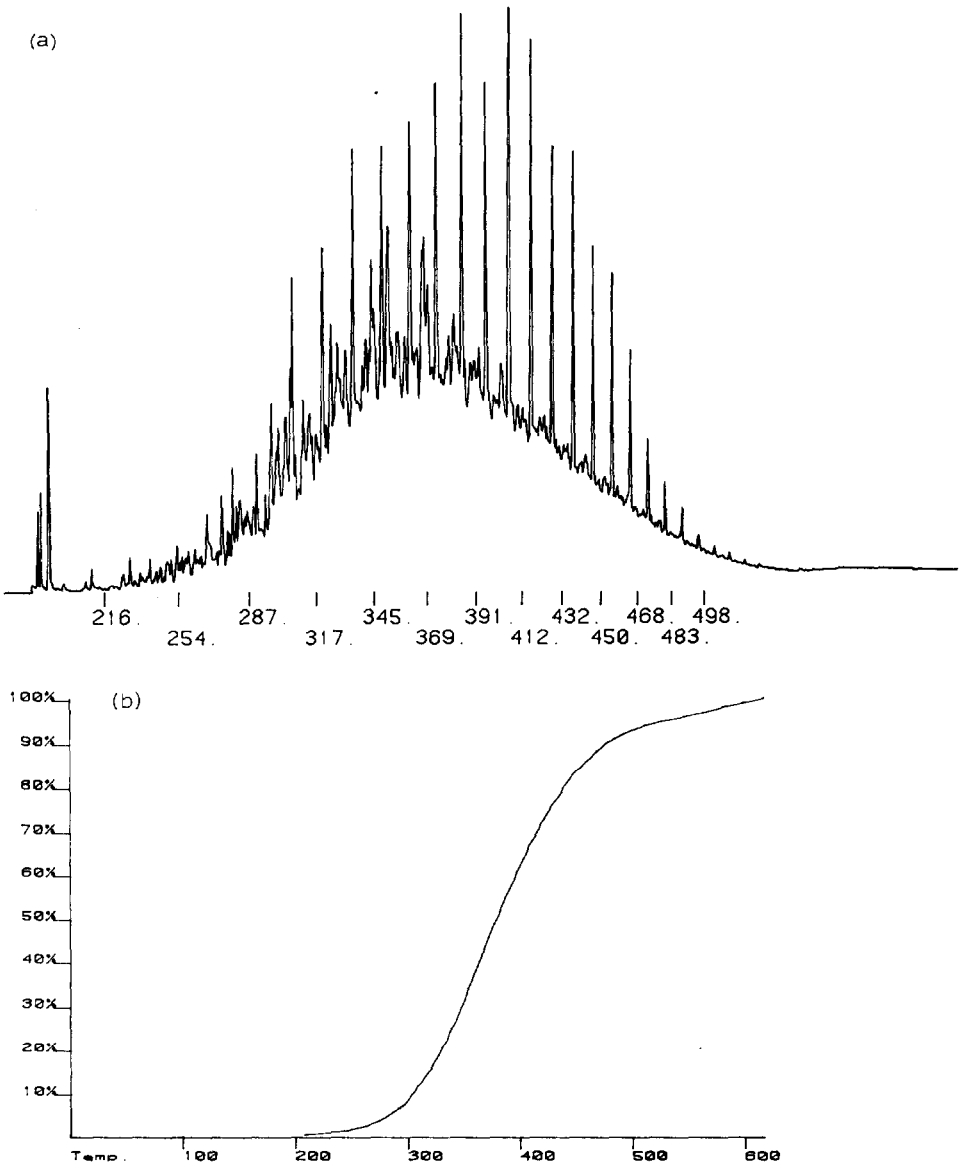

Fig. 2. Result of DISPLAY CRUDE command in example 4: a, sample chromatogram with temperatures; b, percentages of eluted material as a function of temperature.

External baseline subtraction

For some cases of temperature-programmed chromatography a baseline compensation by blank run subtraction is a useful technique[2]:

Example 5:

SELECT 108/2 ; read sample chromatogram

```
SELECT/1 108/1          ; read blank run into compare-buffer 1
SPLIT = 2               ; split display screen into two sections
SCREEN = 1              ; next DISPLAY command into section 1
DISPLAY/1               ; display blank run
SCREEN = 2              ; next into section 2
DISPLAY                 ; display sample chromatogram, use cross-
                          hair cursor to find displacement between
                          the two curves
SUBTRACT/1/125          ; subtract blank run after right-shift
  factor: 1.0             of 125 data points
ANALYZE                 ; perform peak search on corrected sample
                          chromatogram
TYPE                    ; look at result
SAVE                    ; save data files on disk storage
```

Peak search with user-defined baseline

In chromatograms with complicated baseline characteristics the peak analysis program may have difficulties in reconstructing the baseline. The user is able to control the reconstructed baseline by asking for the baseline to be drawn in after having issued the DISPLAY command. In the "graphics dialog mode" he may then erase or redefine baseline points via cross-hair cursor input.

Example 6:

```
SELECT 108              ; read data of instrument 108
DISPLAY                 ; draw raw data on display, with cross-hair
                          cursor dialog ask for reconstructed base-
                          line and correct for bad baseline recon-
                          struction
REANALYZE               ; new peak analysis based on user-defined
                          baseline
REPORT                  ; print result of peak search
```

Procedures for semi-automatic processing

If a series of chromatograms is to be treated with the same sequence of commands, the repeated typing-in of the commands would be tedious and time-consuming. This kind of work is made easier by storing a command sequence in a disk file and by providing means for applying such a command procedure ("method") to subsequently selected chromatograms. All commands executed earlier can be saved in a procedure file, *e.g.*, in KINETIC.PRO, by:

SAVE/PROC KINETIC

A CREATE/PROC enters a dialog, where the commands of a procedure may be typed in. Another way of creating a procedure file would be by use of a standard text-editor, outside the COLACHROM program. Such saved procedures can be executed by a PROC/EXECUTE, abbreviated as "@" command, *e.g.*:

@ KINETIC

A procedure may be built that ends with a call to itself, resulting in a loop.

Example 7:

The following procedure "SERIES" has been created to contain the commands:

```
SAMPLE = KIN.SERIES ABC          ; sample ID for report
SIGEL = ZGL-HB-101               ; customer ID
PURGE 6                          ; retain the six largest peaks only
SEARCH/1 TIME/0.1%               ; identify the components by comparing the
                                   retention times with a table of reference
                                   compounds
REPORT                           ; print the resulting report
NSELECT                          ; proceed to the next chromatogram and
@ SERIES                           restart the procedure
```

Once the first report data file of a series of sequential chromatograms has been selected, this procedure may be executed. The three commands:

```
SELECT/1 rtime.tab               ; read reference table into compare buffer 1
SELECT 10/1                      ; read first chromatogram of series
@ SERIES                         ; execute procedure stored in disk file
                                   SERIES.PRO
```

are then sufficient to process the entire series of chromatograms measured with in352strument 10. The execution comes to an end if the NSELECT command fails to find the next chromatogram, *i.e.*, at the end of the series.

CURRENT IMPLEMENTATION OF COLACHROM

In the current implementation COLACHROM allows for six internal storage buffers: the main buffer and five compare-buffers.

All COLACHROM commands are integrated into one single program comprising many modules. The program was written in 1981 in FORTRAN-77 for operation with a standard VAX/VMS operating system. The virtual address space required depends on the sizes and numbers of arrays for ADC data and report data. In the current version 350 kB of virtual address space is needed to allow for up to 26,000 data points and 400 peaks in the main buffer, up to 13,000 points and 200 peaks in compare-buffer 1, up to 1000 points and 25 peaks in compare-buffers 2, 3, 4, 5.

The program is operational on any VAX- (micro-, mini-, midi-) computer and should be transferable without major difficulties to any machine with adequate address space for this large program (thus excluding most 16-bit mini- and micro-computers!) and with an adequate FORTRAN-77 compiler. The interface with other data systems is given by the format of the data file for unreduced ADC data.

CONCLUSIONS

COLACHROM has been used in the chromatography laboratories of the two Max-Planck-Institutes in Mülheim since 1981. It is a powerful tool especially in a "production-type" environment, where users perform chromatography work as part of their daily routine. Because of the variety of available commands the use of

COLACHROM is not restricted to chromatography; it is applied as well to other analytical data, such as optical spectra or ESR signals. The language can easily be expanded by adding new commands.

ACKNOWLEDGEMENTS

Dr. B. Weimann, I. Wronka, U. Häusig, F. Weeke and K. Schnettberg participated in the design and implementation of COLACHROM.

REFERENCES

1 G. Schomburg, F. Weeke, B. Weimann and E. Ziegler, *J. Chromatogr. Sci.*, 9 (1971) 735.
2 G. Schomburg, F. Weeke, B. Weimann and E. Ziegler, *Chromatographia*, 7 (1974) 477.
3 E. Ziegler, *Anal. Chim. Acta*, 147 (1983) 77.
4 E. Ziegler, B. Weimann, I. Wronka, G. Schomburg and U. Häusig, *Anal. Chim. Acta*, 147 (1983) 91.
5 COLACHROM — User Manual, Max-Planck-Institut für Kohlenforschung, Mülheim, 1983.