Journal of Chromatography, 103 (1975) 33–48

CHROM. 7806

COMPUTER ANALYSIS OF AMINO ACID CHROMATOGRAMS

NEIL BUCHAN*

Computer Science Laboratory, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA (Great Britain)

(Received April 9th, 1974)

SUMMARY

A procedure is described for the automatic off-line analysis of amino acid chromatograms of protein hydrolysates, using a small computer. The data requirements are basic, and, unlike previous programs, the present system allows the separation and identification of bands, as well as the quantitative determination of composition. With minor modification, the program could be extended for use with most types of chromatographic data. The validity of the application of the program to experimental data is discussed.

INTRODUCTION

In the last decade, the automation of amino acid analysis and the development of multi-column high-speed chromatographic techniques have rendered the manual reduction of data a restrictive factor. A relatively large number of automatic analytical procedures have appeared in the literature[1–15] in recent years, but the majority of these are limited in their application (and, surprisingly, in the results they produce) and are often written for specialised custom-built data collection devices, and special-purpose computers. It is also strange to note that, although many authors make a point of favourably comparing their methods with the corresponding manual techniques, in several instances the apparent aim of the project has been lost, the input requirements of some of these programs being so tedious that it is hard to perceive any advantage in automation.

The most recent publication in this field[15] offers a distinct improvement over all previous procedures, but it would appear that the successful application of the quoted program is effectively confined to routine analysis, the method being limited both in terms of degree of automation (and necessity of user interaction) and of handling of non-ideal chromatograms.

In the present paper a program is described which is designed to meet the analytical requirements of most amino acid analysis groups having a small computer available for off-line analysis. Written in FORTRAN IV it has been developed in the author's laboratory on a Honeywell DDP-516, but with minor modifications would

* Present address: "The Lindens", Duffield Road, Darley Abbey, Derby DE3 1JB, Great Britain.

be compatible with any computer capable of accepting FORTRAN and having approximately 16K of store.

The program has the advantage over most previous examples in that (a) it is a single, multi-purpose program, capable of interpreting chromatograms of standards and unknowns; (b) the data requirements are basic (digitised plotter voltages) and minimal in view of the flexibility of the procedure, and the method of data collection can be applied to the majority of commercial analysers; (c) buffer changes, sloping or negative baselines, and overlapping bands are readily accommodated; (d) band identification in a sample chromatogram is automatic; (e) the output is comprehensive (and could be more so if desired); (f) a plotting option is included, which yields a detailed graphical record of the analytical procedure. Point (b) is particularly important, for previous authors seem to have concentrated more on automating the final assessment and tabulation of results, rather than also performing the actual analysis of the chromatogram trace by machine. Indeed many earlier programs employ peak areas as input data, derived from integrators[8,11,14] or even manually[6,8] the former, while affording accurate results, eliminating the possibility of critical examination and separation of overlapping bands, and the latter approach at best only partially eliminating the tedium involved in the general manual method.

EQUIPMENT

The computer used in this work was a Honeywell DDP-516 with 16K of store, fitted with a Digitronics (Model 2500) 300 ch sec^{-1} high-speed tape reader; a Teletype Corporation (BRPE) 110 ch sec^{-1} high-speed tape punch; a Data Printer Corporation (V-132-C) 10 ln sec^{-1}, 132 ch ln^{-1} line printer; a Honeywell 1 Mbyte disc; a Data Dynamics ASR 33 teletype; a Calcomp 565 drum plotter; and a Tektronix 611 storage oscilloscope. However, with appropriate sacrifice of facilities, and modifications to the program, the need for the disc, the plotter, and the oscilloscope could be eliminated, and indeed the line printer is not essential, as the output could be produced less efficiently on the teletype. Effective program operation requires some initial intervention by an operator. In the present case this was achieved by the use of sense switches, and a set of these was made in the laboratory, as the computer software disallowed the use of the console switches. The program is written in FORTRAN IV compatible with DDP computers, but this varies generally only in minor details from the standard language.

Amino acid analysis was carried out on a Beckman 120C amino acid analyser equipped with a multi-record strip chart null type transmittance recorder. The outputs from the three-unit flow colorimeter, housing two 0.66-cm and one 0.22-cm path length cuvettes, are fed to the dotting recorder, which records the output from the three photocells sequentially, at approximately 2-sec intervals. One photocell monitors the transmittance at 440 nm, while the other two monitor that at 570 nm, at different levels according to the cuvette path length; digitisation was applied only to the higher-level member of this latter pair. This procedure introduced the problem of low accuracy in the proline peak, and a future modification of the digitisation procedure is envisaged, allowing the output of the 440-nm channel to be recorded over at least this section of the chromatogram. In the present work, a record from the single channel was obtained by digitising a voltage taken from a slave slide wire placed in parallel

with the recorder slide wire, digitisation being initiated by a reed relay actuated magnetically by the rotatory recorder head. A stabilised 10 V was applied across the slave slide wire, the voltage corresponding to the transmittance value being measured by a digital voltmeter (Solartron LM1450) attached through an interface to an Addo punch, on which the voltage was digitised. (In other systems, an equally satisfactory approach would be to monitor and digitise subsequently the output of one or more of the photocells.) For convenience, the output to the punch was retained within the limits 000 to 999 and was produced on tape as three numbers followed by a carriage return character. Negative absorbances returned figures in continuity with the positive values, commencing with 999 (corresponding to zero absorbance) and progressing downwards. Blank leader and trailer could be produced manually when required, and an automatic facility was included for implanting end-of-run, and other characters on the tapes. The linear voltage-against-chart-absorbance relationships was checked at regular intervals, but was found to remain unchanged within the limits of accuracy of the readings.

The analyser was equipped with four columns as follows: columns 1 and 4 were 69.0 × 0.9 cm; column 2 was 11.0 × 0.9 cm; and column 3 was 29.0 × 0.9 cm. All columns had a buffer flow-rate of 1.13 cm^3 min^{-1}, and their temperature was held constant at 50°F throughout the run. Maximal pressure within the columns was 400, 100 and 50 p.s.i., respectively, the long columns containing Beckman PA-28 separating resin, and the shorter columns Beckman PA-35. Buffer solutions at pH 3.25 and 4.25 for the long columns and 5.28 for the shorter columns were made by the standard procedures, were boiled and filtered, and stored at 4°C. In use they were maintained at slightly above ambient temperature to reduce the release of air bubbles during runs. The growth of micro-organisms in the buffer solution was inhibited by the use of phenol and thiodiglycol in the presence of Brij-35. The ninhydrin flow-rate was 0.57 cm^3 min^{-1}. Beckman concentrated standards were used in all the experiments, 0.5 cm^3 aliquots, corresponding to 0.125 μmole of each amino acid, generally being used in a calibration run.

Protein hydrolysis was carried out according to the method of Moore and Stein[16], using constant boiling hydrochloric acid at 108°C for 24 or 72 h in sealed evacuated tubes.

PROGRAM

The program operates in either of two modes, corresponding to the interpretation of a standard, or of an unknown chromatogram respectively, the appropriate mode being selected by a sense switch. Data point numbers are used for the abscissa units, and not minutes, allowing the arithmetic to be performed in integer form, thus conserving storage space. A flow chart of the procedure is not included here in view of its complexity, but this and a listing of the program is available on request.

Standard analysis mode

In either mode the initial input requirement is for an identifying name, NAM, and for digitised voltage data (see Table I). In the case of a standard run this latter will correspond to the chromatogram of a known mixture, with, hopefully, at least all the peaks that are likely to occur in the ensuing unknown curves. (Fig. 1 depicts a

TABLE I

TYPICAL INPUT DATA FOR THE ANALYSIS OF A STANDARD AMINO ACID MIXTURE

Primary data tape

```
1*   5    10   15   20   25   30   35   40
Long column 4 standard 17.1.73 data 4B
021
021
022
023
022
.
.
.
.
.
038
037
037
038
037
-99
```

Secondary data tape

```
1        5        10       15      20
16
 1      20.0     CYA     0.1345**
 2      46.9     ASP     0.125
 3      56.3     THR
 4      59.7     SER
 5      70.5     GLU
 6      79.3     PRO
 7      96.0     GLY
 8     103.2     ALA
 9     118.3     CYS
10     127.0     IMP***  0.1
11     132.2     VAL
12     150.0     MET
13     157.8     ILE
14     162.2     LEU
15     183.1     TYR
16     189.3     PHE
       13.0      142.8
```

Teletype input

```
1        5        10
   10    6
```

* These figures denote punch columns.

** The absence of a figure in this column indicates that the concentration is to be internally defined (in this example as 0.125 μmoles).

*** This peak corresponds to an impurity added deliberately in order that a full description of the capabilities of the program could be given by a single example.

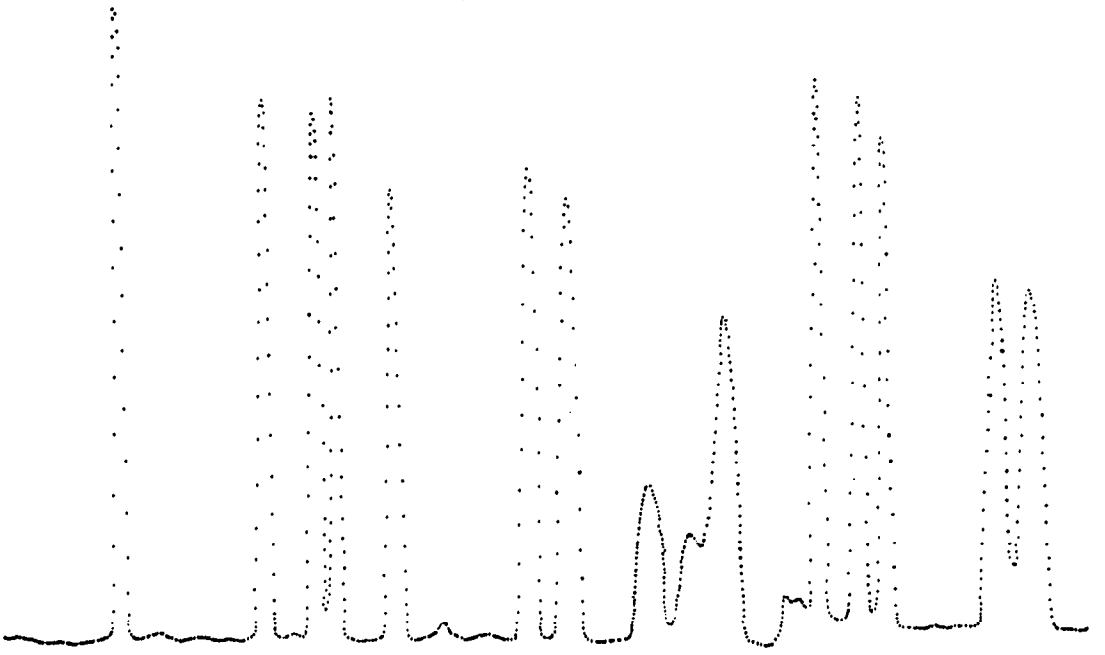

Fig. 1. Representative standard neutral/acidic amino acid chromatogram.

standard chromatogram, sections of which are used in the remaining figures.) Having read the primary data tape, the execution will pause, to allow a standard/sample choice to be made through a sense switch. With the sense switch set, re-execution then introduces a demand for a secondary data tape (see Table I) containing approximate retention times, RTM(*I*) (obtained from the recorder chart by inspection), and mnemonics, ALPH(*I*), for all the peaks of interest, together with the various amino acid concentrations, CONC(*I*). As this latter is usually a constant (in the present work at 0.125 μmoles), a figure need be included in the data set only if the value differs from the accepted normal level, QONC. Experimental constants ZLEDM and BUFFM must also be included on this secondary tape, the former being the interval between the start of the experiment and the initiation of digitisation, and the latter being the period elapsed before the appearance on the chart of the buffer change peak. Both times are in minutes. Although the necessity for this tape is somewhat undesirable, such standard runs take place relatively infrequently, and this approach has been found for standard traces to be superior to that of the more common automatic peak search method[2,5,6]. The final input requirement is in terms of the five analytical threshold limits, LIM1 to LIM5, which are defined in the forthcoming paragraphs, and depend on the characteristics of the recorder and sample, and is satisfied via the teletype (for convenience) based on experience of values required in previous analyses. In most cases of pure routine analysis it was found that only two of those limits, LIM1 and LIM3, needed to be defined at this point. However, for wider applications of the procedure, it may be necessary to define more of the limits at run time. Following the input procedure the chromatogram data are converted to

the absorbance equivalent using the constants derived from the calibration of recorder voltage in terms of absorbance. Data corresponding to negative absorbance values are assessed in terms of magnitude, and all data are subsequently incremented by an amount sufficient to produce a completely positive data set. This procedure does not, of course, influence the subsequent numerical manipulations. At the end of the input section, an option is incorporated, involving a second sense switch, which allows the production of a plot only, execution bypassing the analytical stages. This can be useful for assessing the experimental data, following the production of unusual or suspect results, when the original chart is not available.

Peak maximum detection then commences with the identifying name being written as a heading to possible forthcoming messages. In this laboratory these messages were produced on a storage oscilloscope for convenience, but a teletype could be used with equivalent results. Detection of the maximum for each peak proceeds by comparison of absorbance levels over a region (the width of which is determined by LIM1) centred on the retention time given in the input list. If the approximation on this value has been too great, a request will be made for a revised time to be entered on the teletype for the particular peak. When the maximum is detected (LIM2 ensures that enough steps have been taken beyond the maximum to be sure of its authenticity) the retention time is corrected accordingly, giving CRTM(I), and the process is repeated until all named peaks have been examined. LIM5 is the minimum acceptable peak size limit, and is equal to half the number of points necessary to define a peak.

The next section determines the band limits by successive absorbance comparison from the band centre outwards. These limits, ISTLIM(I) and IFNLIM(I) for band commencement and termination respectively, are assigned when a point is reached where the overall change in absorbance for a given number (defined by LIM4) of points is less than a predetermined threshold limit (LIM3). A summary of the procedure thus far is depicted in Fig. 2.

Satisfactory integration of overlapping peaks can only be realised by their isolation, and the first step in this process is taken in the next section. The nature (whether single or multiple) of each peak in a chromatogram is determined by assessment of the adjacent minima. If the absorbance values at both minima remain effectively constant for more than a given number of points (determined by LIM4) then the peak is deemed to be single. If, however, this criterion is not met at one or both minima, the peak is considered to be one of a multiple set. The limits of each set are determined by assessment of the characteristics of each member of the set, those two having a leading and a trailing static minimum respectively being defined as the initial and final members. For separation purposes a matrix is then set up holding details of the groups of peaks within a trace; Fig. 3 gives an example of this procedure.

The area under a band envelope, AREA(I), is obtained by the standard method of strip summation between the limits determined and for isolated peaks is a straightforward matter of scanning the single-peak matrix, and thence summing for the indicated bands (trapezoidal baseline corrections are applied as in Fig. 4a). For overlapping bands, however, a preliminary separation technique must be applied, and this stage involves the categorising of overlap types in order that the appropriate procedure may be implemented. Taylor and Davies[15] consider that perpendicular band separation is satisfactory even for strongly overlapping bands, and have had

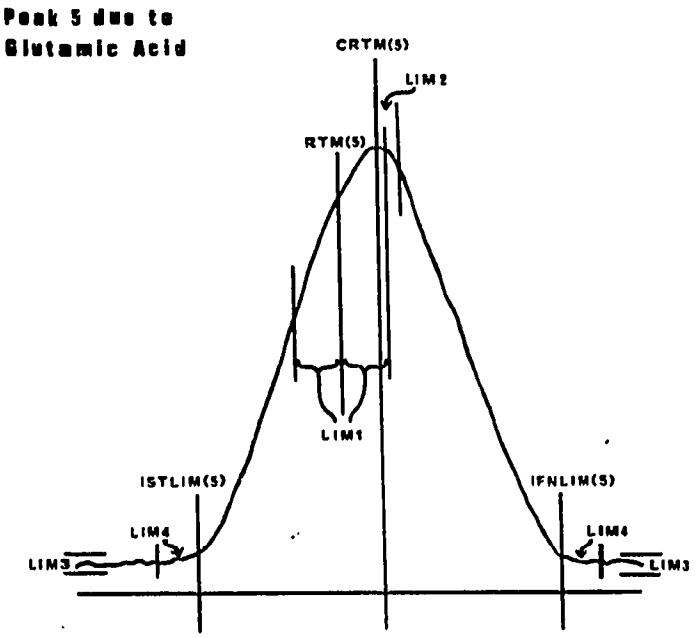

Fig. 2. Detection of peak limits. The abscissa has been expanded in the interests of clarity.

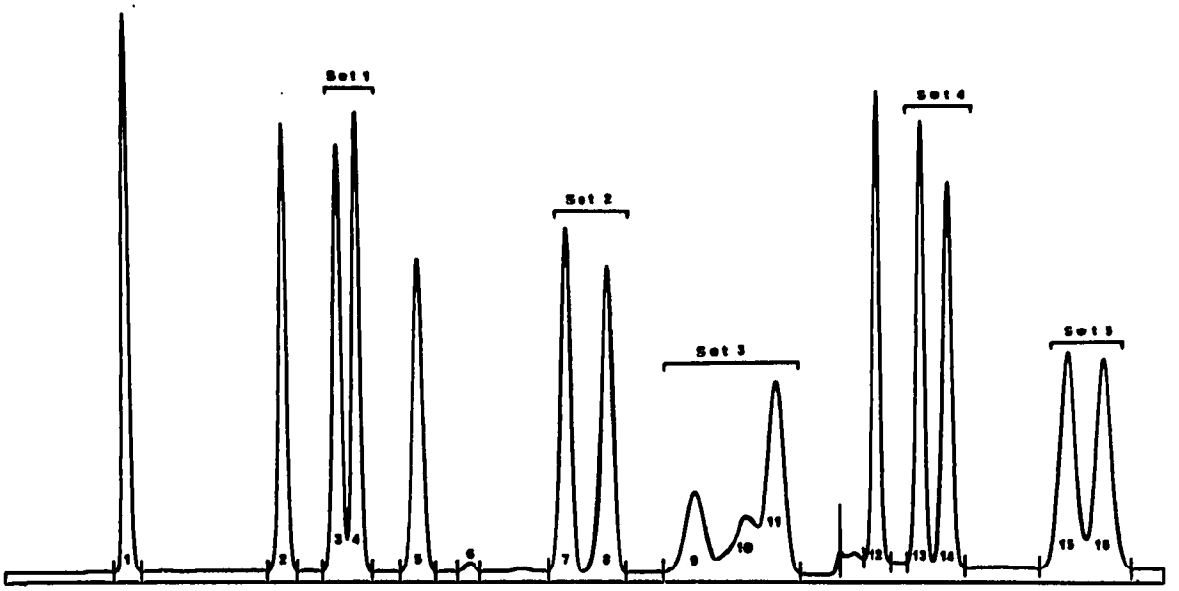

Fig. 3. Peak type assignment. Main integration limits are depicted on the contour by vertical dashes.

Single peak matrix, ISNGPK(*I*):

Peak number, *I*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Matrix value	1	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0

Multiple peak matrix, IMLTPK(*I*):

Peak number, *I*		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Matrix value	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
Set number, *J*	3	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0
	4	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1

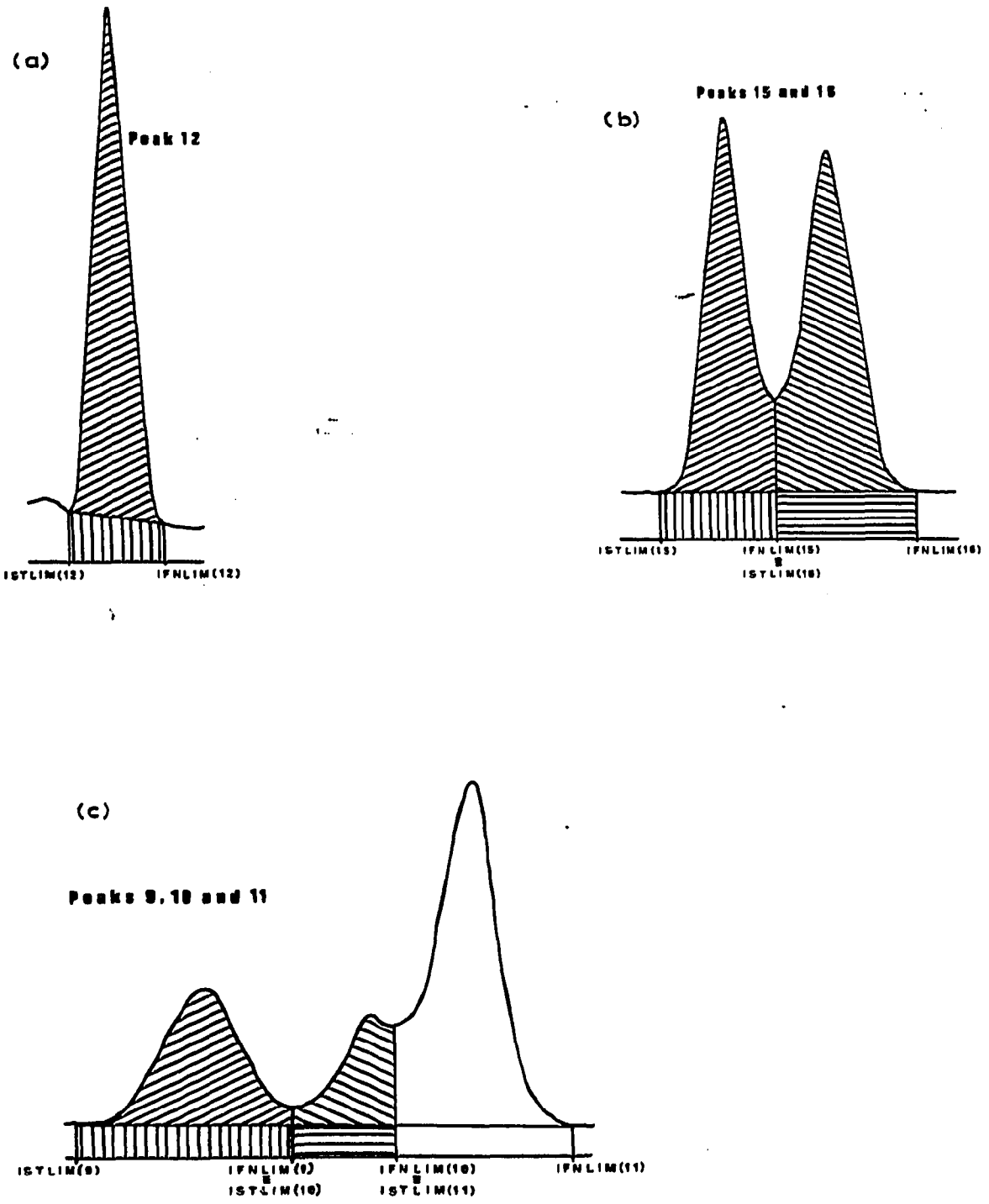
(a)
Peak 12
ISTLIM(12)
IFNLIM(12)
(b)
Peaks 15 and 16
ISTLIM(15)
IFNLIM(15)
ISTLIM(16)
IFNLIM(16)
(c)
Peaks 9, 10 and 11
ISTLIM(9)
IFNLIM(9)
ISTLIM(10)
IFNLIM(10)
ISTLIM(11)
IFNLIM(11)

Fig. 4.

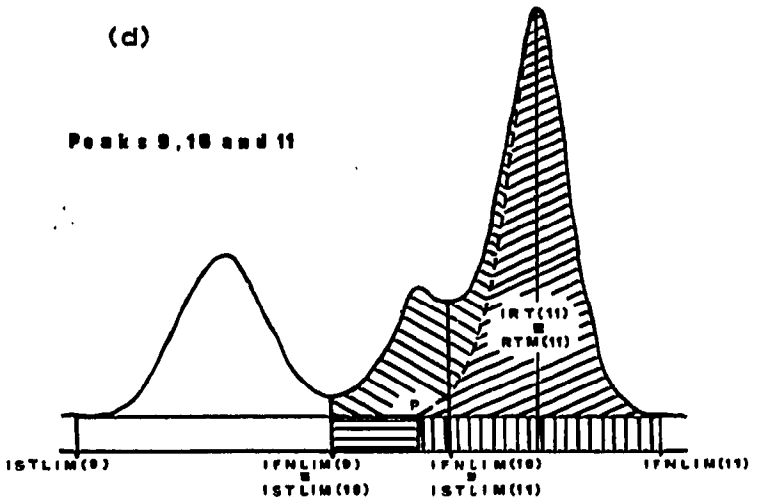

Fig. 4. (a) Derivation of area under a single peak. Peak area obtained from total area under contour between peak limits, minus area of baseline trapezium. (b) Derivation of areas of a weakly overlapping pair. Areas of peaks obtained by perpendicular separation, and correction for baseline, defined by limits of the overlapping set. (c) and (d) Derivation of areas of strongly overlapping bands. In (c) areas of peaks 9 and 10 are derived by perpendicular separation, with the baseline defined by the limits of the overlapping set. In (d) areas of peaks 10 and 11 are derived by subtracting twice the area under the contour between IRT(11) and IFNLIM(11) from the area between the limits of the two peaks. The former area is defined as the uncorrected area for peak 11 and the remainder as that for peak 10. The correction trapezia are defined from ISTLIM(10) to P, and from P to IFNLIM(11), respectively, where P is obtained from IRT(11)−P=IFNLIM(11)−IRT(11), the baseline being defined as before.

considerable success in applying this approach to routine data. However, even with their inclusion of proportional trapezoidal baseline corrections, the areas derived under certain circumstances could be grossly inaccurate.

In the present work, each overlapping set (which may consist of several bands) is inspected in pairs of peaks, commencing from the left, the degree of overlap within each pair determining the method of separation. After much experiment, the following types of overlap have been categorised, each demanding a different separation procedure:

(a) A pair in which both true maxima are at least a certain factor times the mutual true minimum —This critical height factor has been determined empirically as 2.4 by prolonged studies in IR spectroscopy[17], being a figure at and above which separation may be legitimately achieved simply by dropping a perpendicular from the minimum and defining limits accordingly. Integration is then straightforward with correction being applied as in the case of the single peaks (see Fig. 4b).

(b) A pair in which one true maximum is at least 2.4 times the mutual true minimum while that of the other is not —This is not a common feature of routine amino acid analysis, but any complete procedure must be capable of accommodating such an occurrence. One important example of this type of band merging in amino acid analysis involves the overlap of the band corresponding to homoserine with the leading wing of the (usually) much larger band of glutamic acid. In whichever orienta-

tion such a pair exists the analysis proceeds by setting twice the area under the contour from the pair limit remote from the small peak to the major peak maximum, equal to the area of the large peak. Subtraction of this area from the total area under the contour yields the effective area of the small peak. Baseline correction is then applied by trapezium subtraction, limits for this procedure being determined by reflection of the large peak limit through the maximum perpendicular, the position of the experimental baseline being determined by the initial and final limits of the overlapping set. This method of correction, although open to some criticism, with the experimental baseline adequately close to zero incurs in practice very little error in peaks which are in any case invariably minor. An example of this procedure is depicted in Fig. 4d.

(c) A pair of bands in which both true maxima are less than 2.4 times the mutual true minimum —This is a relatively abnormal occurrence, but in such circumstances a request is made to the operator to define the separation technique from those already outlined, through a code number, on the teletype.

Two points arising in the preceding paragraphs warrant expansion. Firstly, the assignment of areas to bands within a composite involves all bands but the first and the last being assessed twice for type by inspection of both neighbours. Typical of this procedure is peak 10, detailed in Figs. 4c and 4d. The final area assigned to peak 10 will derive from the procedure embodied in Fig. 4d, as opposed to that from Fig. 4c. (In the circumstances where the relative position of the bands is reversed the procedure causes the former assessment of the area to be assigned.) The second point involves a peak of minimal height which falls between two other peaks, each exceeding the critical height criterion. Although the area of such a peak will be derived via method (b) above, it will be in error by virtue of having been corrected only for single neighbour peak overlap. As such peaks are likely to be of limited importance and, by any technique, of doubtful area, the peak area derived in such circumstances is simply modified by FACTOR, which at the time of writing is set arbitrarily at 0.5. However, it has not, to date, proved possible to test the validity of this value practically, as no chromatograms have yet been obtained containing such diminutive peaks.

In the final analytical section of the program, area-to-concentration conversion factors, CONFAC(I), are determined for each peak, expressed for convenience in terms of concentration per unit area, and control is then passed to the output section.

The initial output of results takes place via the line printer, a list of concentrations, accurate retention times, peak areas and conversion factors for each amino acid being produced for inspection, and assessment of the analysis. A typical example of such a listing is given in Table II. A plot of the chromatogram may then be obtained if required, containing information on the analytical path taken, in particular as regards peak separation. Fig. 5 depicts the plot obtained from the analysis of the example shown in Fig. 1. The final stage of the output section involves the listing on the storage oscilloscope of the elements of the two band type matrices, ISNGPK(I) and IMLTPK(I,J), and the production of a control tape for the program when run in its alternative mode, containing amino acid retention times, mnemonics and conversion factors, plus values for limits LIM1 and LIM3.

Sample analysis mode

The general analytical path of the program in its second mode is very similar to that followed during the analysis of a standard chromatogram, but there are some

TABLE II

REPRESENTATION OF AN OUTPUT FROM THE ANALYSIS OF THE STANDARD AMINO ACID MIXTURE

Tabulation

Peak number	*Amino acid*	*Concentration (μmoles)*	*Retention time (min)*	*Band area (min)*	*Conversion factor (μmoles per unit area)*
1	CYA	0.1345	20.1	1.019	0.1320
2	ASP	0.1250	46.9	0.918	0.1362
3	THR	0.1250	56.2	0.946	0.1321
4	SER	0.1250	59.5	0.971	0.1288
5	GLU	0.1250	70.3	0.893	0.1400
6	PRO	0.1250	78.6	0.025	4.9769
7	GLY	0.1250	95.5	0.949	0.1317
8	ALA	0.1250	102.7	0.968	0.1292
9	CYS	0.1250	117.7	0.437	0.2858
10	IMP	0.1000	126.5	0.320	0.3127
11	VAL	0.1250	131.5	0.872	0.1434
12	MET	0.1250	149.0	0.980	0.1275
13	ILE	0.1250	156.8	1.007	0.1241
14	LEU	0.1250	161.3	0.983	0.1271
15	TYR	0.1250	182.1	0.973	0.1285
16	PHE	0.1250	188.2	0.964	0.1296

LIMITS 10 3 6 8 5 ZLEDM 13.0 BUFFM 142.8

Control tape

```
1         5         10        15        20
c.r.
16
          20.1  CYA  0.1320
          46.9  ASP  0.1362
          56.2  THR  0.1321
          59.5  SER  0.1288
          70.3  GLU  0.1400
          78.6  PRO  4.9769
          95.5  GLY  0.1317
         102.7  ALA  0.1292
         117.7  CYS  0.2761
         126.5  IMP  0.2853
         131.5  VAL  0.1331
         149.0  MET  0.1300
         156.8  ILE  0.1241
         161.3  LEU  0.1271
         182.1  TYR  0.1285
         188.2  PHE  0.1296
        10   6
```

important differences. The initial input requirement of a named data tape is identical to that of mode 1 (see Table III), but the second requirement is for a control tape produced from the analysis of the appropriate standard, separated under the necessarily similar experimental conditions. The final input requirement of the present program is for the punch start time ZLEDM, and the buffer change time, BUFFM,

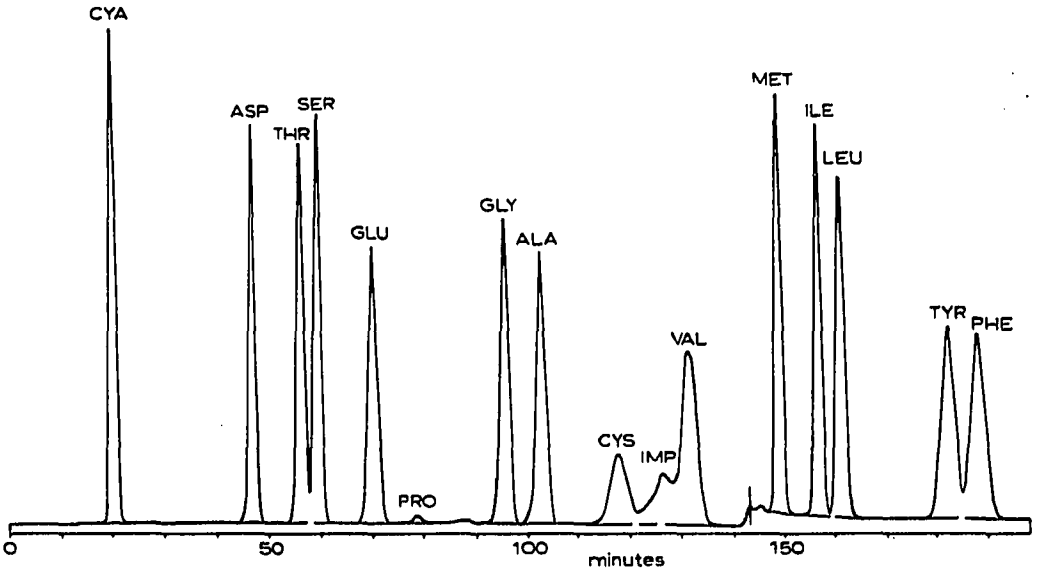

Fig. 5. Graphical output from standard analysis.

TABLE III

TYPICAL INPUT DATA FOR THE ANALYSIS OF AN UNKNOWN AMINO ACID MIXTURE

Primary data tape

1 5 10 15 20 25 30 35 40 45

```
Long column 4 29.1.73 D/4 A(1) data 5,4 B
016
016
016
017
016
.
.
.
.
046
046
045
046
045
-99
```

Control data tape

See Table II.

Teletype input

1 5 10 15

```
  25.0 143.0
```

both in minutes. In a more routine environment these two parameters could be defined internally.

In the following section, the number of peaks present in the chromatogram, and their accurate retention times, are determined, by scanning the whole contour, and defining peak maxima by the change in absorbance difference method outlined in the standard mode description. Peaks below a minimum size (defined by LIM5) are ignored, and care is taken to avoid interpretation of the buffer change as a peak by blanking the buffer change region with internally defined limits LIM6 and LIM7.

The limits of all the bands detected are determined by the same procedure as that employed in the standard mode, hence leading to assignment of band types and areas, again by identical paths to those described previously.

In the final analytical sections, bands in the unknown chromatograms are identified by comparison of retention times with those introduced on the standard control tape, any band not having a retention time corresponding to a value in the standard table being assigned as unknown. Retention times are used in this work, with appropriate leeway (LIM1), this method having proved as satisfactory as any more complicated alternative. Exss *et al.*[11] made an interesting study of the reproducibility of retention times and corresponding time ratios, and, although the results were somewhat startling, the latter function, to which these authors resorted for peak identification, did not show itself to be indisputably superior. Furthermore, it is felt that, certainly in routine analysis, the reproducibility normally to be expected in either parameter is somewhat better than that found by these authors, although it is as well to bear in mind their findings on the sensitivity of retention times and time ratios of particular amino acids to the varying concentrations of other acids. The validity of the identification procedure can be judged for individual analyses by inspection of the tabulated and graphical results. Failure of this section of the procedure results in the inability to identify bands, although their areas will still be produced, and incorrect identification should rarely occur. The final stage in this section involves the calculation of amino acid concentrations for all identifiable bands derived from band areas and conversion factors, the latter, again, being introduced on the control data tape.

The output procedure is similar to that followed in the standard mode, the tabulation containing retention times, peak areas, conversion factors, and concentrations of amino acids for all identified bands (see Table IV). Peak number, retention time, and area only are listed for unidentified bands. Again a plotting option is available (see Fig. 6), and the final stage is the tabulation of the band type matrices on the storage oscilloscope. Such a table affords a useful guide to the validity of the analysis.

DISCUSSION

The experimental assessment of the program was conducted in terms of the comparison of results with those produced manually. No attempt has been made to investigate the accuracy of the analytical method, and it is not, at this time, the intention so to do. However, the author has found evidence similar to the findings of Exss *et al.*[11], in terms of the variation of conversion (colour) factors, as well as retention times, derived for a given amino acid band in the presence of various concentrations of other amino acids, and it would certainly seem that a comprehensive assessment

TABLE IV

OUTPUT FROM THE ANALYSIS OF THE UNKNOWN AMINO ACID MIXTURE

Peak number	*Amino acid*	*Retention time (min)*	*Band area (min)*	*Conversion factor (μmoles per unit area)*	*Concentration (μmoles)*
1	UNK	40.8	0.217	—	—
2	ASP	45.8	0.536	0.1362	0.0731
3	THR	55.1	0.312	0.1321	0.0413
4	SER	58.4	0.324	0.1288	0.0417
5	UNK	67.4	0.644	—	—
6	GLY	94.5	0.324	0.1317	0.0427
7	ALA	101.7	0.597	0.1292	0.0771
8	VAL	130.5	0.582	0.1331	0.0774
9	ILE	156.8	0.230	0.1241	0.0286
10	LEU	161.3	0.482	0.1271	0.0612
11	UNK	167.7	0.327	—	—
12	TYR	182.7	0.190	0.1285	0.0244
13	PHE	188.4	0.169	0.1296	0.0218

LIMITS 10 3 6 8 5 ZLEDM 25.0 BUFFM 143.0

of the range(s) of validity of the area/concentration equivalence method would not be out of place.

Table V shows the results obtained for the manual and automatic analyses of five typical amino acid mixtures, and it will be noted that the agreement between concentrations derived from the two methods is very good (figures for proline must be treated with reserve for the reasons outlined earlier). It should be noted that absolute agreement of band areas between the two data reduction techniques is not expected * in view of the differences in the integration procedures employed. The data tabulated contain figures obtained from peaks with amplutides throughout the range 0.02 to 1.2 absorbance units, and from peaks overlapping in pairs and triplets (poor resolution was deliberately induced in some mixtures), and there is no evidence to suggest that

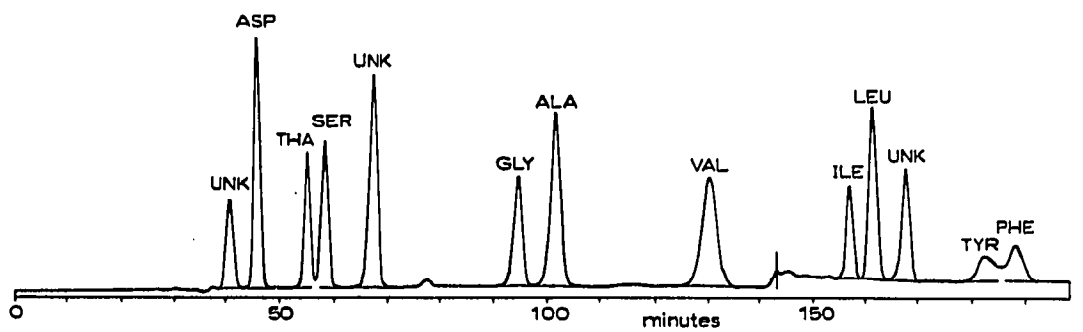

Fig. 6. Graphical output from sample analysis.

* It can be readily shown that there is a constant relationship between true band areas and areas derived by the height–width method, independent of half-width or absorbance maximum. If the bands are assumed to be of Gaussian shape then the true area, A_T, is related to the half-width area, A_{HW}, by $A_T = 1.0645\ A_{HW}$.

TABLE V

COMPARISON OF MANUAL AND AUTOMATIC ANALYSES OF NEUTRAL/ACIDIC AMINO ACID MIXTURES

Amino acid	*Mixture 1*		*Mixture 2*		*Mixture 3*		*Mixture 4*		*Mixture 5*	
	Manual	*Auto.*	*Manual*	*Auto.*	*Manual*	*Auto.*	*Manual*	*Auto.*	*Manual*	*Auto.*
Asp	0.172	0.171	0.033	0.032	0.024	0.024	0.076	0.073	0.163	0.165
Thr	0.069	0.070	0.018	0.018	0.015	0.016	0.042	0.042	0.122	0.120
Ser	0.061	0.060	0.022	0.025	0.022	0.022	0.042	0.042	0.023	0.023
Glu	0.141	0.142	0.045	0.045	0.024	0.024	0.091	0.090	0.101	0.099
Pro	0.070	0.064	0.018	—	0.048	—	0.094	0.097	0.046	0.039
Gly	0.091	0.090	0.037	0.038	0.032	0.033	0.042	0.043	0.040	0.040
Ala	0.086	0.086	0.061	0.062	0.027	0.027	0.078	0.078	0.061	0.061
Cys	0.012	0.012*	Nil	Nil	Nil	Nil	0.015	0.016	Nil	Nil
Val	0.069	0.069	0.029	0.029	0.054	0.051	0.079	0.079	0.048	0.049
Met	0.013	0.014	0.006	0.005	Nil	Nil	Nil	Nil	0.022	0.021*
Ile	0.044	0.045	0.016	0.016	0.014	0.014	0.029	0.029	0.020	0.020
Leu	0.125	0.127	0.033	0.034	0.027	0.026	0.062	0.060	0.060	0.061
Tyr	0.055	0.056	0.009	0.007	0.011	0.011	0.025	0.025	0.030	0.028
Phe	0.056	0.056	0.015	0.015	0.025	0.026	0.026	0.022	0.030	0.029

* Interpolated from area, and colour factor from standard.

general lack of agreement between the two methods of data reduction occurs under certain conditions. However, the more sophisticated approach to multiple band separation employed in the program yielded in some instances results not in agreement with the manually derived counterpart, but in close agreement with the true figures when these were known.

The system has now been tested on over fifty mixtures of widely varying compositions, involving some 700 peaks and in only two instances has misidentification occurred due to non-reproducibility of retention times. Fourteen cases of lack of identification were noted, but this situation presented no problems, manual determination of acid concentrations being afforded by use of the area produced and the appropriate colour factor.

Assessed in terms of agreement with the manual alternative, the procedure has proved most successful, and in the cases of artificially prepared mixtures of known concentration (some 20% of the total) has produced solution composition values consistent with the accuracy of the analytical method.

ACKNOWLEDGEMENTS

The author wishes to thank Dr. M. J. Crumpton for his guidance and criticism as a potential user throughout this work. Gratitude must also be expressed to Mr. B. F. Stenning for his tireless construction and modification of the data logging devices.

REFERENCES

1 D. H. Spackman, *Fed. Proc.*, 22 (1963) 244.
2 W. L. Porter and E. A. Talley, *Anal. Chem.*, 36 (1964) 1692.
3 R. D. B. Fraser, A. S. Inglis and A. Miller, *Anal. Biochem.*, 7 (1964) 247.

4 A. Yonda, D. L. Filmer, H. Pate, N. Alonzo and C. H. W. Hirs, *Anal. Biochem.*, 10 (1965) 53.
5 M. I. Krichevsky, J. Schwartz and M. Mage, *Anal. Biochem.*, 12 (1965) 94.
6 G. N. Graham and B. Sheldrick, *Biochem. J.*, 96 (1965) 517.
7 A. J. Robins, R. A. Evans, J. A. de S. Siriwardene and A. J. Thomas, *Biochem. J.*, 99 (1966) 46p.
8 W. C. Starbuck, C. M. Mauritzen, C. McClimans and H. Busch, *Anal. Biochem.*, 20 (1967) 439.
9 K. Ozawa and S. Tanaka, *Anal. Biochem.*, 24 (1968) 270.
10 J. F. Cavins and M. Friedman, *Cereal Chem.*, 45 (1968) 172.
11 R. E. Exss, H. D. Hill and G. K. Summer, *J. Chromatogr.*, 42 (1969) 442.
12 J. J. T. Gerding, *Int. J. Protein Res.*, 1 (1969) 169.
13 H. L. Back, P. J. Buttery and K. Gregson, *J. Chromatogr.*, 68 (1972) 103.
14 H. D. Spitz, G. Henyon and J. N. Siverton, *J. Chromatogr.*, 68 (1972) 111.
15 R. Taylor and M. G. Davies, *Anal. Biochem.*, 51 (1973) 180.
16 S. Moore and W. H. Stein, *Methods Enzymol.*, 6 (1963) 819.
17 N. Buchan, *Ph.D. Thesis*, University of London, 1971.