CHROM. 5957

A NEW METHOD FOR THE COMPUTATION OF AMINO ACID CHROMATOGRAMS

H. L. BACK, P. J. BUTTERY AND K. GREGSON

Departments of Applied Biochemistry and Nutrition and Physiology and Environmental Studies, University of Nottingham School of Agriculture, Sutton Bonington, Loughborough, Leics. LE12 5RD (Great Britain)

(Received November 22nd, 1971)

SUMMARY

A system for the automatic integration of amino acid chromatograms is described. The computer program employed differs from those previously published in that the critical points of the chromatogram are detected by orthogonal linear and quadratic components through sets of absorbance readings. The program can also accommodate changes in baseline due to buffer changes and to adjustments on the photometers. A very high correlation was observed between the areas computed using the program and those calculated using the conventional height times width at half the height method.

INTRODUCTION

To realize the full potential of the rapid amino acid analyser systems now available, analysis of the chromatograms must be automated either by using an integrator directly connected to the machine or by storing the data from the machine on paper or magnetic tape for processing in a computer. This paper describes the equipment and a new computer program used to integrate the peak areas of chromatograms. The program differs from those previously published[1,2] by using orthogonal linear and quadratic components to detect the critical points of the chromatogram and in the safeguards to prevent fluctuations in the baseline affecting accuracy of the computation.

EXPERIMENTAL AND DISCUSSION

An Evans Electroselenium Model 294 amino acid analyser using the lithium citrate buffer system described by ATKIN AND FERDINAND[3] was used. The output from a 570-nm photocell was fed to a digital voltmeter (Model 1420.2, Solatron Electronic Group Ltd., Hampshire, Great Britain) and then by means of a data transfer unit (Solatron Electronic Group Ltd.) to a paper tape punch (Model 4070, Facit, Sweden). The output from the photocell was recorded on the paper tape every 10 sec following a command from the clock in the data transfer unit. This interval could easily be altered but 10 sec was found to be most appropriate for our analyser.

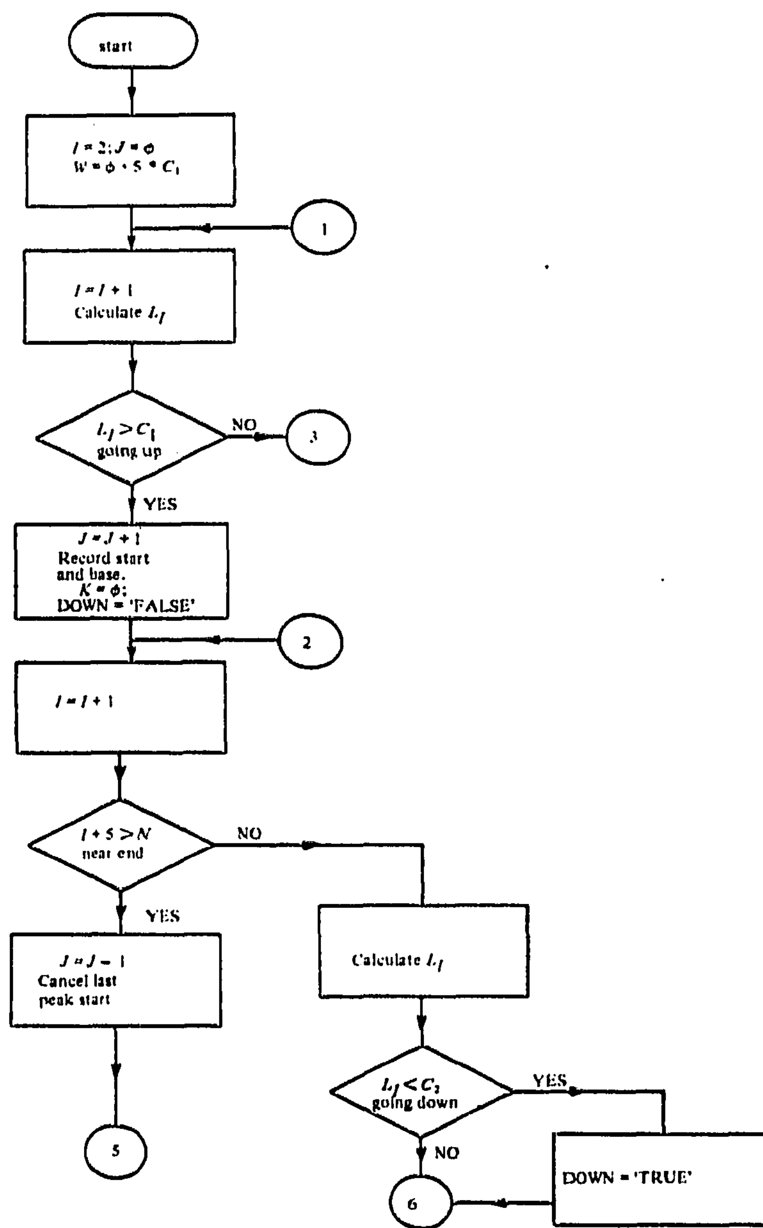

Fig. 1. Flow diagram of the analysis procedure. X = ARRAY of the output of the photocell converted to absorbance readings with bounds 1 to N; N = number of absorbance readings; C_1, C_2, C_3, C_4 = constants supplied by the user (see Table I); $W = {}^1/_2C_1$; I = counter through data points; J = counter of peaks found; K = counter of number of peaks found in a multiple peak; L_I = linear component centred on X_I; Q_I = quadratic component centred on X_I; N_1 and Z = internal stores for the position and value of the maximum reading in a peak; DOWN is a BOOLEAN: 'TRUE' when $L_I < C_2$ 'FALSE' otherwise.

The data transfer unit was modified so that the output of the paper tape was in the form of five characters for each reading, namely four digits followed by a space. The characters "carriage return line feed" could be placed on the tape by the closure of two relay contacts. This facility was employed to indicate where one chromatogram ended and the next began.

The interaction between the selector switch of the multipoint recorder (Speedomax W, Leeds and Northrup, Birmingham, Great Britain) and the digital

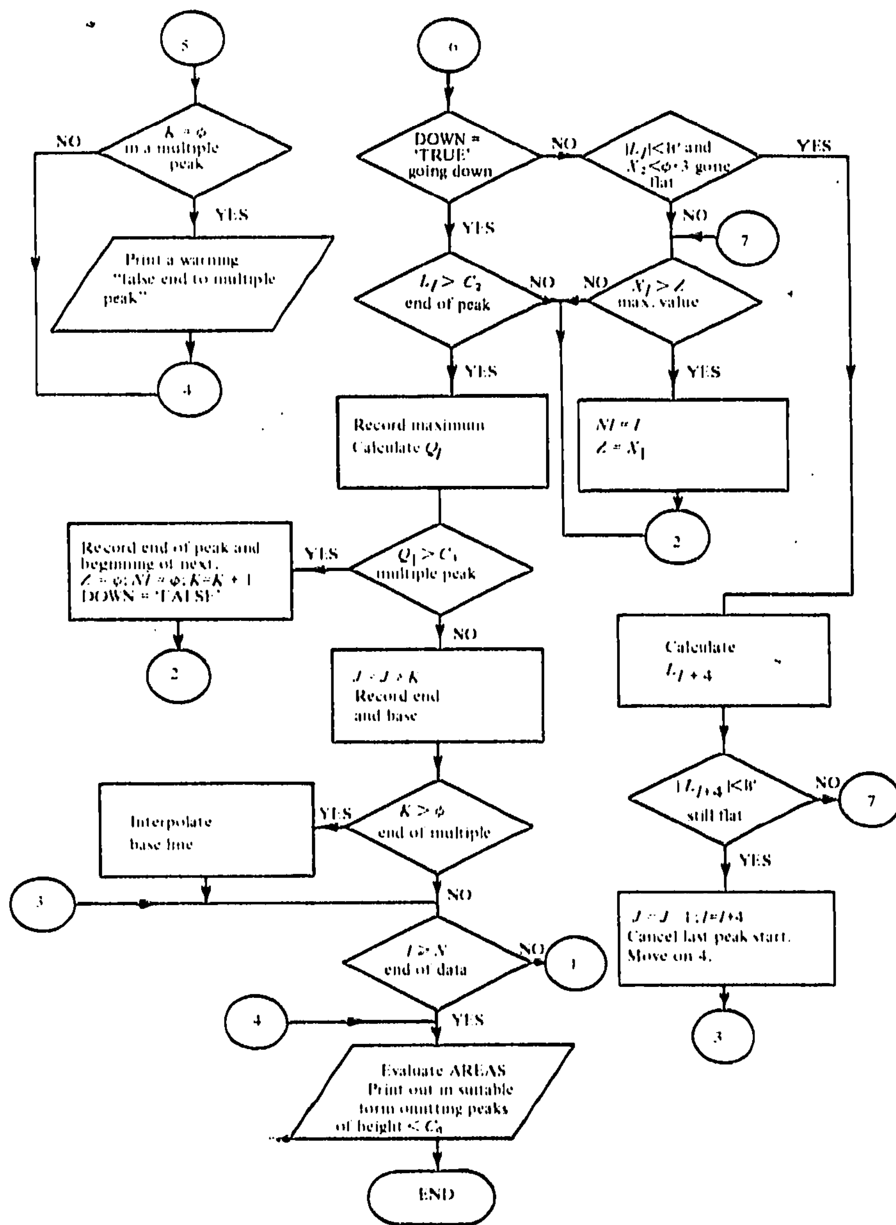

Fig. 1 (*continued*).

voltmeter was minimal thus removing the need to complicate the system by the fitting of a transmitting potentiometer to the multipoint recorder (see ref. 1).

Initially the computer program converts the millivolt readings from the photocell into absorbance so that the readings in the computer and on the chromatogram chart are identical. A block diagram of the analysis procedure is shown in Fig. 1. The various criteria used to determine the start, summit and end of a peak and the presence of a multiple peak (see Table I) require explanation.

Let the absorbance readings be denoted by $X_1, X_2 \cdots X_n$. Then the values can be calculated when required which are proportional to the orthogonal linear (L_i) and quadratic (Q_i) components of curves through a set of readings on either side of X_i. Even without smoothing acceptable results have been obtained using five points when the equations are:

$$L_i = (-2X_{i-2} - X_{i-1} + X_{i+1} + 2X_{i+2})/10$$

$$Q_i = (2X_{i-2} - X_{i-1} - 2X_i - X_{i+1} + 2X_{i+2})/14$$

The behaviour of L_i and Q_i through typical single and multiple peaks is shown in Fig. 2. The critical points are readily located by comparison of L_i and Q_i with constants C_1, C_2 and C_3 (see Table I). The values of these constants depend on the machine and the sensitivity required. In our particular system they are supplied with each paper tape on a data card.

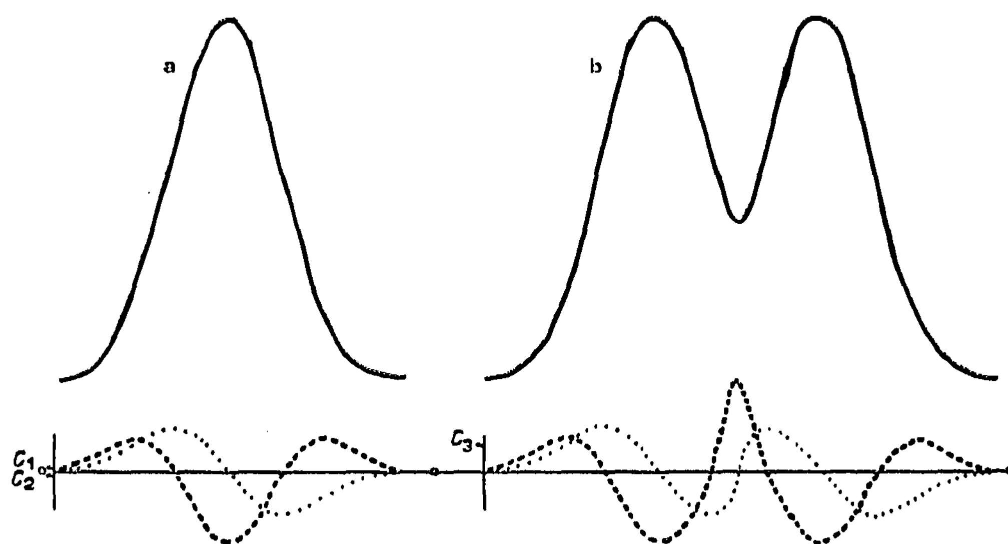

Fig. 2. Behaviour of L_i (linear component through five consecutive points) (· · · · · ·) and Q_i (quadratic component through five consecutive points) (– – – – – –) in the region of (a) a single peak and (b) a double peak. The original trace is also shown (———). C_1, C_2 and C_3 are constants which depend upon the operating condition of the amino acid analyser and the data capture system. See also Table I. (Traces are not to any particular scale.)

TABLE I

CRITERIA USED TO DETECT CRITICAL POINTS ON CHROMATOGRAMS

See also Fig. 2.

Critical point	L_i	Q_i	$\|L_i\|$
Start	$> C_1$	—	—
End single peak	$> C_2$	$< C_3$	—
Trough between multiple peaks	$> C_2$	$> C_3$	—
Descent of peak (*i.e.*, after maximum)	$< C_2$	—	—
False start			$< {}^1/_2\, C_1$ and $\|L_{i+4}\| < {}^1/_2\, C_1$

It is advisable to set C_1 and C_2 numerically small so that peaks are not unduly curtailed. This is especially important on the basic physiological fluids runs[3], where the peaks tend to be very broad. The numerous small peaks that result from a chromatogram with a bad baseline when C_1 and C_2 are set low can be suppressed from the output by using the criterion of a minimum change in absorbance between the summit and either tail of a peak. We have found it useful to indicate on the output the number and position of the suppressed peaks.

The false start test is a precaution against the inclusion of a sharp upward baseline change into a genuine peak. Such changes occur, on our machine at least, at a buffer change, when a photometer bulb is about to fail, or at the start of a run before the pulsation suppression on the buffer line has settled down. This false start test is inadmissible in the region of multiple peaks and also at high absorbance readings where there is a tendency for the tops of peaks to be flattened.

Should the data acquisition system stop in the middle of a peak, care must be exercised in interpreting the computed peak areas if the preceding peak formed a multiple peak with the partially recorded peak. Its baseline will be incorrectly calculated (see Fig. 3), and it is essential to print a warning if this happens.

We normally record the following data on our output statement: peak number omitting suppressed peaks, peak number including suppressed peaks, the value of I (see Fig. 1) and the absorbance reading at the start, summit and end of a peak and the area of the peak. The results from each separate chromatogram are printed on a

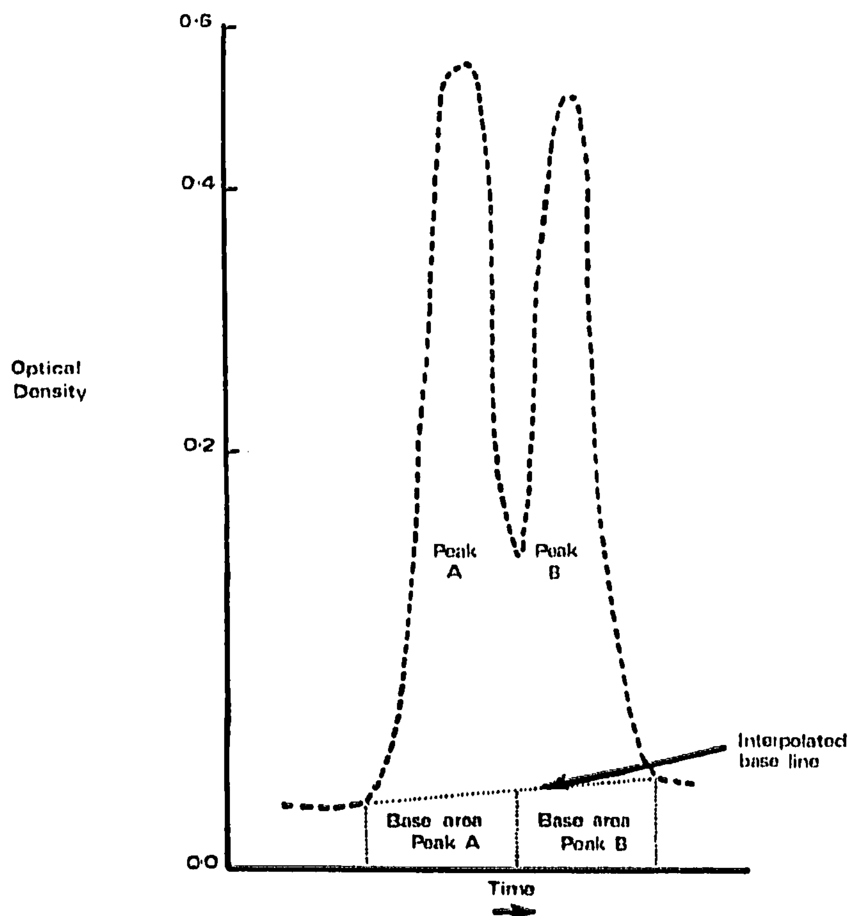

Fig. 3. Base areas deducted through a multiple peak.

new page headed by the details of the run which are supplied on data cards which accompany each paper tape.

The accuracy of the system was tested by two types of experiments, varying quantities of a standard amino acid mixture (Evans Electroselenium Ltd.) were assayed and the peak area of each acid on each chromatogram computed. The correlation coefficients between the quantity of each amino acid assayed and the computed peak areas were determined (see Table II). The slightly lower value for valine was probably due to the tendency for the tail end of this peak to be distorted by emergence of the buffer change. Second, the area of 50 peaks ranging in height from 0.02 to 0.80 optical density units were calculated manually using the height times width at half height method[4] and the results compared with the value computed using the program described above. A correlation coefficient of +0.995 was obtained.

TABLE II

EVALUATION OF COMPUTER CALCULATION OF AREA OF PEAKS

A series of neutral protein hydrolysate and acidic amino acid analysis runs[3] were conducted using varying amounts of a standard amino acid mixture (Evans Electroselenium Ltd.). The amounts of each amino acid analysed were 0.100, 0.050, 0.025 and 0.010 μmole. In no case did the intercept of the linear regression differ significantly from zero.

Amino acid	*Linear correlation coefficient between computed area and quantity of amino acid applied to column*
Aspartic acid	0.999
Threonine	1.000
Serine	1.000
Glutamic acid	1.000
Glycine	0.998
Alanine	0.998
Valine	0.993
Methionine	0.998
Isoleucine	1.000
Leucine	1.000
Nor-leucine	1.000
Tyrosine	1.000
Phenylalanine	1.000

Previously reported programs[1,2] for the computation of the areas of peaks on amino acid chromatograms have used the difference between successive absorbance readings to detect the critical points of the chromatogram. When applied to our analysers we found this system unsatisfactory since it was too sensitive to slight unsteadiness in the output of the photometers, even when the absorbance readings were smoothed by the computer. There was a tendency for single peaks to be split into several peaks. This splitting of peaks could be avoided by increasing the difference required between two values to indicate the start and end of a peak but this tended to make the computation of the peak areas inaccurate, especially when they were broad. In addition, our system was much more efficient in the use of computer time.

With the present system it is possible to use the computer for all the calculations involved in amino acid analysis with the exception of introducing the appropriate

colour factor. Results are normally expressed as area/g of tissue or millilitre of sample for each peak. The appropriate colour factor is employed at a latter stage after inspection of the chromatogram. Although a method has been reported[5] for introducing the colour constants into computer programs for the calculation of amino acid chromatograms it is only suitable for simple runs where the chromatographic conditions are always constant, *e.g.*, protein hydrolysate analyses.

The program is written in Algol and Usercode and run under the Egdon system on an English Electric KDF9 computer at the Cripps Computing Centre, Nottingham University. A copy can be obtained from the authors.

ACKNOWLEDGEMENT

The authors wish to acknowledge the help of the Cripps Computing Centre for the help with the paper tape reading routine.

REFERENCES

1 A. J. Robins, R. A. Evans, J. A. de Siriwardene, and A. J. Thomas, *Biochem. J.*, 99 (1966) 46p.
2 M. I. Krichevsky, J. S. Schwartz and M. Mage, *Anal. Biochem.*, 12 (1965) 94.
3 G. E. Atkin and W. Ferdinand, *Anal. Biochem.*, 38 (1970) 313.
4 D. H. Spackman, W. H. Stein and S. Moore, *Anal. Chem.*, 30 (1958) 1190.
5 R. E. Exss, H. D. Hill and G. K. Summer, *J. Chromatogr.*, 42 (1969) 442.