Biophysical Chemistry 20 (1984) 299–304
Elsevier

BPC 00885

A SIMPLE METHOD OF ANALYZING PROFILES IN TIME-DIFFERENCE DIRECT OPTICAL SCANNING GEL CHROMATOGRAPHY

JOHN R. GROTHUSEN * and JAMES K. ZIMMERMAN **

Department of Biochemistry, Clemson University, Clemson, SC 29631, U.S.A.

Received 6th February 1984
Accepted 11th May 1984

Key words: Gel chromatography; Time difference; Large zone; Direct optical scanning

A simple method is presented for the analysis of time-difference, large-zone chromatography profiles. The method is derived from basic theory and is applicable to multicomponent as well as single-component systems. Simple computer simulations are used to demonstrate the inaccuracies of earlier, more empirical methods. This method has been tested on several proteins using an inexpensive, semi-automated, data acquisition and control system.

1. Introduction

Direct optical scanning gel chromatography was introduced by Brumbaugh and Ackers [1] in 1968. The general theory and advances within the field are covered in several reviews by Ackers [2–4]. A variety of experimental protocols are used in direct optical scanning gel exclusion chromatography [1–4]. One technique often used is termed the large-zone method in which the sample volume is comparable to the column volume, so that a concentration plateau is reached. Scans taken during the loading process are called leading-edge scans.

Time-difference chromatography for the direct optical gel scanner was introduced by Brumbaugh et al. [5] to minimize experimental errors introduced during the time course of large-zone experiments. This method has been used by Saffen and Chun [6] and Oeswein and Chun [7] to analyze the behavior of oxyhemoglobin A and the interaction of human serum apolipoprotein B with sodium deoxycholate. Zimmerman [8] examined the technique of time-difference chromatography using computer simulations and suggested some modifications in the method of analysis.

In this paper we demonstrate that both of the previously suggested methods have inherent errors that can be manifested under certain conditions. We also introduce a method that is extremely simple and based on the fundamental properties of gel chromatography.

The rate of movement of the centroid of large-zone profiles has been shown to be a measure of the partition coefficient and therefore a measure of the molecular size of the sample [1]. The determination of an equivalent sharp boundary (centroid) is limited by the experimental noise in the baseline, in the plateau, in the concentration profile, and in the problems associated with the numerical integration of this type of data.

A much simpler method of determining the rate of centroid movement, called the time-difference method, has been proposed by Brumbaugh et al. [5]. The procedure is to subtract successive large-zone profiles from each other, resulting in $n-1$ difference profiles, where n is equal to the number of large-zone scans. Brumbaugh et al. [5] stated that this reduces the problem of calculating

* Current Address: Department of Entomology, Fisheries, and Wildlife, Clemson University, Clemson, SC 29631.
** To whom correspondence should be addressed.

centroid movement to one of determining the rate of movement of the peaks obtained in time-difference scans. This novel data analysis technique was used in a study of the dimer-tetramer equilibrium of oxyhemoglobin A [6]. A later paper using computer simulations [8] claimed that it is the center of mass of the time-difference profiles, and not the peaks, that are of interest in determining the rate of large zone profile movement. A recent paper by Oeswein and Chun [7] uses the time-difference peak procedure to study human serum apolipoprotein B.

In this paper we present a new analytical procedure for large-zone centroid rate determinations. This method, which we call the 'time-difference area technique,' is based on fundamental parameters. Computer simulations have been generated to highlight the difference of this method with the other two semi-empirical methods.

In addition to the computer simulations we have analyzed experimental data collected on a semi-automated system. This system is controlled and processed by an inexpensive Hewlett-Packard 85 desk-top computer [9].

2. Materials and methods

Sephadex G-200, Blue dextran, DNP-phenylalanine, myoglobin, ovalbumin, bovine serum albumin and alcohol dehydrogenase were purchased from Sigma Chemical Co. Other common laboratory chemicals were purchased from Fisher Scientific. The precision bore quartz columns used for scanning (inner diameter 0.935 cm, outer diameter 1.1 cm, length 25.5 cm) were purchased from W.A. Sales, Chicago, IL. The pump used with the scanner system was an ISCO model 1612 peristaltic pump purchased from ISCO. The microcomputer was a Hewlett-Packard 85 with the 00085-15003 input/output ROM option. The microcomputer was connected via a Hewlett-Packard 82937A interface bus (HP-IB) to a Hewlett-Packard 3497A data acquisition/control unit with the digital voltmeter and the 44428A 16 channel actuator output options.

The basic scanner system used for this work has been described in detail elsewhere [1,9–11]. In our instrument the column is mounted on a drive mechanism that moves the column through a double-grating monochrometer at a precise speed. The signal current from the phototube is converted to a voltage and read by the Hewlett-Packard 3497A data acquisition/control unit. For all experiments using this system the time increment between successive readings was set so that 500 data points were collected in the 2.15 min it takes to scan the 17.6 cm of column used. The control unit is equipped with the actuator option which allows programmable control over switching operations. The data are collected in memory, then transferred to magnetic tape. A baseline scan can be stored in memory so that at any time between actual data collections one can view the most recent scan with baseline subtracted.

The stored data can be analyzed on the HP-85 using software we have written in Hewlett-Packard Extended BASIC. Alternatively, the raw data can be transferred to the Clemson University IBM 3081 computer.

All of the techniques used, as well as the software developed for this system, are described in detail elsewhere [9].

3. Results

A simplest-case approach was used to investigate some of the assumptions used in the analysis of time-difference profiles. A simulation was written in which leading-edge, large-zone profiles were represented as straight lines with decreasing slopes. This is shown in fig. 1.

In this simulation, successive profiles were generated that had their centroids separated by 10 data points per profile. The concentration gradient for each profile was represented as a straight line to simplify calculations and to maximize analytical errors. The height of each concentration plateau was arbitrarily set at 1 absorbance unit. The point of intersection of the sloping line with the plateau for each scan was determined by subtracting the product of the spread factor and the profile number from the centroid position. The baseline intercept was determined by adding the product of the spread factor and the profile number to the

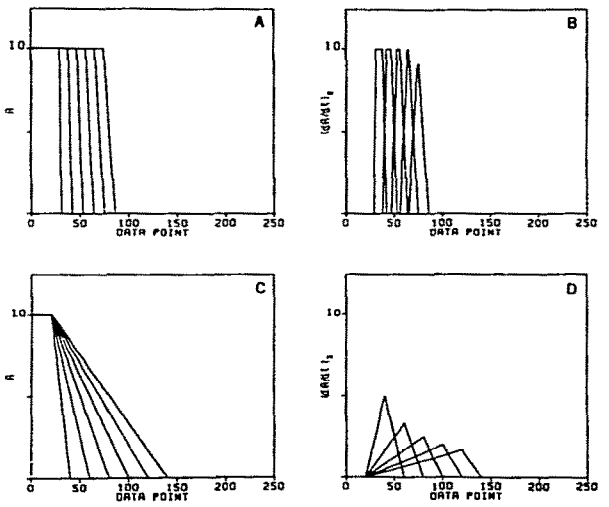

Fig. 1. Simplest-case leading-edge profiles. Profiles in A and C have centroids at 30, 40, 50, 60, 70 and 80 data points. Panel B shows the corresponding time-difference profiles of panel A. Panel D shows the corresponding time-difference profiles of panel C. Parameters for these two cases are indicated in table 1.

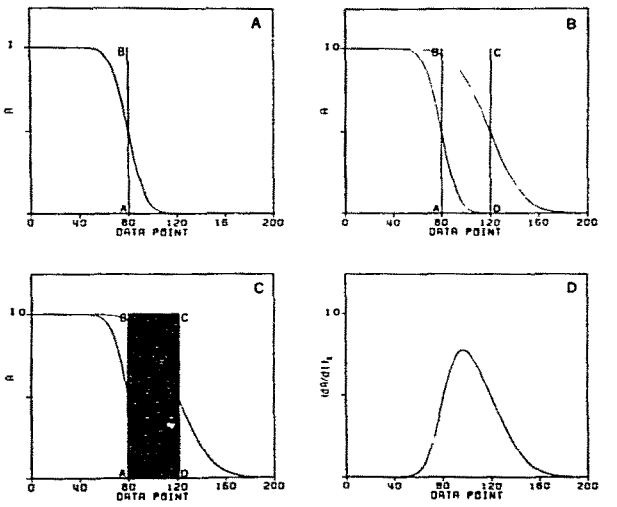

Fig. 2. Diagrammatic representation of the time-difference area procedure. Panel A depicts a single leading edge and the equivalent sharp boundary, AB (centroid). Panel B shows the equivalent sharp boundaries of two successive profiles (CD and AB). The area of the rectangle ABCD (shaded) in panel C is equivalent to the area experimentally obtained (panel D) when the first profile is subtracted from the second profile. Therefore, the centroid movement AD is equal to the area in panel D (also equal to AB × AD), divided by the plateau value, AB.

centroid position. The straight line connecting the two intersection points passes through the centroid and describes a simplistic front representing either a homogeneous solute, a heterogeneous one or a 'reaction profile' having the assigned centroid. The slope of this straight line is $-1/(2 \times$ profile number $\times$ the spread factor). The actual shape of the concentration profile will not affect the results if the basis of the analysis is theoretically correct. The time-difference profiles were created by subtracting successive simulated profiles.

The time-difference peak movement was calculated from a linear regression of the position of the peak of each time-difference profile vs. the profile number (time) as described in ref. 5. The time-difference center of mass movement was similarly calculated using the position of the center of mass of each time-difference profile vs. profile numbers as described in ref. 8.

The time-difference area was calculated by first determining the area under the time-difference profile and then dividing this value by the plateau height of the large zone. This procedure is illustrated in fig. 2. By definition, lines AB and CD describe the position of equivalent sharp boundaries at two successive times. The rate of movement of the centroid is then the distance AD divided by the difference in time. The area of rectangle ABCD describes the mass transported between successive scans. The area under the curve in fig. 2D is equal to the shaded area in fig. 2C and therefore also equal to the area of the rectangle ABCD. Dividing the area under the experimentally determined curve in fig. 2D by the plateau height (line AB or CD in fig. 2B) gives the distance AD, or the movement of the centroid in the time span between the two scans. For a particular experiment this value is determined $n-1$ times, where n is the number of concentration profiles collected.

The three methods of calculating centroid movement from time-difference profiles were compared. The rate determined by the movement of the time-difference peaks bears little relation to

Table 1

A comparison of time-difference analyses using a simplest-case simulation

Spread factor	Peak movement rate (data points/profile)	Center of mass rate (data points/profile)	Reduced area rate (data points/profile)	Input centroid movement (data points/profile)
1 [a]	10.00	10.01	10.00	10
2	12.00	10.23	10.00	
5	15.00	11.34	10.00	
7	17.00	12.39	10.00	
10 [b]	20.00	14.35	10.00	
1	5.80	5.11	5.00	5
2	7.00	5.45	5.00	
3	8.00	5.91	5.00	
4	8.80	6.49	5.00	
5	10.00	7.17	5.00	
5	25.00	20.75	20.00	20
10	30.00	22.67	20.00	
15	35.00	25.37	20.00	

[a] Corresponds to panels A and B of fig. 1.
[b] Corresponds to panels C and D of fig. 1.

the true centroid movement. This is shown in table 1 where comparisons are made at different rates of movement, as well as with different spread factors.

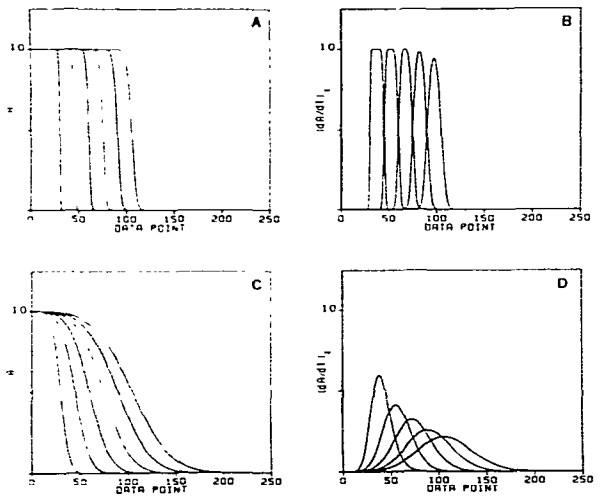

Fig. 3. Time-difference using Gaussian simulations. Profiles in A and C have centroids at 30, 45, 60, 75, 90 and 105 data points. The profiles in panel A (corresponding to set A of table 2) have smaller spread factors than those in panel C (corresponding to set D of Table 2). The profiles on the right are the corresponding time-difference profiles; (B corresponds to A, D corresponds to C).

The rates determined by the movement of the centers of mass of the time-difference profiles are closer to the correct values, but also increase with increased overlap between scans (increased spread factor). In all cases, even though the rate of movement calculated may be grossly incorrect, the correlation coefficient of the linear least-squares fit is very high (0.999). In each case tested, the area method gave the correct value.

Similar results were obtained when simulations were made using the integrated form of the Gaus-

Table 2

A comparison of time-difference analyses using a Gaussian simulation

Rate of centroid movement = 15 points/profile. Relative spread value: relative to set A. The spread values are proportional to the variance in the Gaussian equation. For a fixed time this value is a function of the flow rate and the dispersion coefficient which itself is a function of several parameters [3,12]. Profiles for sets A and D are given in fig. 3.

Set	Relative spread value	Peak movement data (points/profile)	Center of mass (data points/profile)	Area (average) (data points/profile)
A	1	15.20	15.01	15.00
B	3	15.50	15.34	15.00
C	6	16.50	16.28	15.00
D	9	17.90	17.68	15.01

Table 3

Comparison of rates of movement using calculated equivalent boundaries and time-difference areas

	Leading edge equivalent sharp boundary (data points/min)	Time-difference area (data points/min)	Stokes radius [13] (Å)	M_r
Blue dextran	3.09	2.99	–	$>10^6$
Alcohol dehydrogenase	2.23	2.24	41.7	141000
Bovine serum albumin	2.04	2.07	37.0	67000
Ovalbumin	1.79	1.88	27.6	45000
Myoglobin	1.57	1.65	18.6	18600
Phenylalanine	1.39	1.42	–	165

sian equation to describe leading-edge profiles as done by Halvorson and Ackers [12] (fig. 3). The rates of movement of the time-difference peak positions, the rates of movement of the centers of mass, and the time-difference areas were again calculated and are shown in table 2. Again the method using time-difference area returns the input parameters while the other two methods fail.

The time-difference area technique was also tested with experimental data. The absorbance at each point of a baseline-subtracted, leading-edge scan was first divided by the absorbance of the plateau at that point (obtained from saturation values). The resulting profiles were then subtracted one from another to yield the time-difference profiles. The area for each member of a particular set of time-difference profiles was determined by trapezoidal integration. The individually determined areas for each profile set were then averaged.

In these runs, equally spaced data points were collected over the 17.6 cm of the gel column scanned at a wavelength of 220 nm. The time between the initiation of scans in all cases was 10 min. The void volume marker was Blue dextran and the internal volume marker was DNP-phenylalanine. The buffer used in all cases was 0.1 M potassium phosphate (p. 7.6). The results are shown in table 3 where they are compared with values obtained from equivalent sharp-boundary analyses of the leading edges [1]. The average difference in values (always using the lower value of the rate for the division) is 2.9%. In no case is there a difference greater than 5%. It can be seen that the time-difference method and equivalent sharp-boundary measurements agree quite well

4. Discussion

The time-difference method for large-zone analysis is useful since each scan becomes a baseline for the next scan, thus eliminating the major sources of noise that develop during the course of an experiment. However, as seen in the simple cases presented, the more overlap between successive profiles, the more the time-difference peak movement increases with respect to the equivalent sharp-boundary rate of movement. The same holds true for the movement of the center of mass of the difference curve. Large molecules have greater dispersion than small molecules [12], but also faster centroid movement. Faster centroid movement results in less overlap of profiles and therefore the time-difference peak movement more closely resembles the movement of the centroid. The time-difference peaks reflect the amount of dispersion as well as the rate of centroid movement, dispersion being more pronounced for smaller molecules. In the case where successive profiles are not completely separated, the result will be erroneously high rates.

The method of time-difference area has several distinct advantages. Values calculated are dispersion independent and are determined by the rate of centroid movement only. Random noise can be reduced, since the areas obtained can be averaged, while examination of the individual areas can be

used to detect any interaction of the sample with the gel matrix. This technique will also give weight-average partition coefficients for multicomponent or interacting systems which previous time-difference techniques cannot justify on a theoretical basis.

References

1 E.E. Brumbaugh and G.K. Ackers, J. Biol. Chem. 243 (1968) 6315.
2 G.K. Ackers, Adv. Protein Chem. 24 (1970) 343.
3 G.K. Ackers, in: The proteins, 3rd edn. 1, ed. H. Neurath (Academic Press, New York, 1975) p. 2.
4 G.K. Ackers, in: Methods of protein separation, ed. N. Catsimpoolas (Plenum Press, New York, 1976) p. 1.
5 E.E. Brumbaugh, E.E. Saffen and P.W. Chun, Biophys. Chem. 9 (1979) 299.
6 E.E. Saffen and P.W. Chun, Biophys. Chem. 9 (1979) 329.
7 J.Q. Oeswein and P.W. Chun, Biophys. Chem. 18 (1983) 35.
8 J.K. Zimmerman, Biophys. Chem. 11 (1980) 173.
9 J.R. Grothusen, Ph.D. Thesis, Clemson University, U.S.A. (1983).
10 E.E. Brumbaugh, Ph.D. Thesis, University of Virginia, U.S.A. (1971).
11 B.B. Brown, Masters Thesis, Clemson University, U.S.A. (1977).
12 H.R. Halvorson and G.K. Ackers, J. Polym. Sci. A-2 9 (1971) 245.
13 J.M. Brewer, A.J. Pesce and R.B. Ashworth, Experimental techniques in biochemistry (Prentice-Hall, New York, 1974) pp. 355.