

CHROMBIO. 2586

Note

Confidence limits for a gel-permeation system

MARION P. CULLEN\* and R.W.R. BAKER

*Evelina Children's Hospital, United Medical and Dental Schools of Guy's and St. Thomas' Hospitals, Department of Paediatrics, Guy's Hospital, London SE1 9RT (U.K.)*

(First received November 27th, 1984; revised manuscript received February 7th, 1985)

A semi-automatic gel-permeation system designed for the study of human glomerular filtration of dextrans has recently been described by Cullen et al. [1]. These authors used two columns of TSK chromatography media (Toyo Soda, Tokyo, Japan) in series under pressure for the separation of discrete dextran fractions. Calibration of the assembly was effected with the use of no less than seventeen narrow-cut dextran fractions, each of which is closely characterised in terms of weight-average molecular weight (MW). For each fraction, a mean elution volume ( $V_e$ ) was found from generous (at least ten-fold) replications over a range of column loads. Thus an almost unique collection of data was available for the statistical study of probable errors. Dextrose and dextran of weight-average  $MW > 2 \cdot 10^6$  daltons were used as end-markers and in the first instance, the mean elution volumes of these were taken as the total separation volume ( $V_t$ ) and the void volume ( $V_0$ ), respectively, in evaluating partition coefficients  $K_{av} = (V_e - V_0)/(V_t - V_0)$ .

We found the relationship between  $-\ln K_{av}$  (ordinates) and  $m$  ( $= MW^{2/3}$ ) was sigmoidal in nature, with slopes increasing sharply as  $m$  declined towards zero or increased finally at the upper end of its range. Over part of the elution range, however, there was an indication of linearity.

EXPERIMENTAL

Data used in the present study are shown in Table I and were calculated from the original observations of Cullen et al. [1]. In an attempt to realize a closer approach to the classical linear form (Porath [2]; Hjertén [3]) a series of triads was used to derive values for  $V_0$  and for  $V_t - V_0$  as described earlier by Baker [4]. A triad is here defined as any group of three points (Table I), each

TABLE I  
ELUTION DATA

| Point No. | Substance        | Mean MW          | No. of replicates | $V_e$ (ml, mean $\pm$ S.E.M.) |
|-----------|------------------|------------------|-------------------|-------------------------------|
| 1         | Dextrose         | 180              | 19                | 38.56 $\pm$ 0.10              |
| 2         | Dextran fraction | 1170             | 10                | 35.13 $\pm$ 0.16              |
| 3         | Dextran fraction | 2500             | 10                | 33.95 $\pm$ 0.12              |
| 4         | Dextran fraction | 4100             | 11                | 33.37 $\pm$ 0.14              |
| 5         | Dextran fraction | 4500             | 10                | 32.75 $\pm$ 0.16              |
| 6         | Dextran fraction | 5250             | 12                | 32.66 $\pm$ 0.11              |
| 7         | Dextran fraction | 7900             | 10                | 31.57 $\pm$ 0.10              |
| 8         | Dextran fraction | 8825             | 12                | 31.67 $\pm$ 0.10              |
| 9         | Dextran fraction | 11500            | 12                | 31.01 $\pm$ 0.12              |
| 10        | Dextran fraction | 14700            | 10                | 30.02 $\pm$ 0.13              |
| 11        | Dextran fraction | 21975            | 12                | 29.06 $\pm$ 0.14              |
| 12        | Dextran fraction | 31900            | 10                | 28.12 $\pm$ 0.17              |
| 13        | Dextran fraction | 39200            | 10                | 27.63 $\pm$ 0.15              |
| 14        | Dextran fraction | 42150            | 12                | 27.47 $\pm$ 0.10              |
| 15        | Dextran fraction | 73625            | 12                | 25.95 $\pm$ 0.09              |
| 16        | Dextran fraction | 104450           | 12                | 24.93 $\pm$ 0.09              |
| 17        | Dextran fraction | 144960           | 12                | 24.28 $\pm$ 0.14              |
| 18        | Dextran fraction | 239825           | 12                | 23.23 $\pm$ 0.11              |
| 19        | Dextran fraction | $> 2 \cdot 10^6$ | 11                | 18.29 $\pm$ 0.05              |

characterised by the crude data, i.e. mean elution volume and molecular weight. We thus obtained means for the above parameters from which a significantly linear correlation of  $-\ln K_{av}$  with  $m$  was obtained over about half of the range of elution volumes from  $V_0$  to  $V_t$  as given by the end-markers. With the calibration equation now determined, variabilities could be found by standard statistical procedures. Next, taking the partition coefficients as variables not subject to significant error, it was possible to assess quantitatively the errors to be expected when mean elution volumes were used to estimate molecular weights.

Programs written for a small computer (Sinclair ZX81) enabled the solution of simultaneous exponential equations to be carried out very rapidly, and provided the print-out of trial plots of calibration graphs.

## RESULTS

In a preliminary trial, seven well spaced triads from the range of points 2–18 (Table I) were used to derive seven values for  $V_0$  and for  $V_t - V_0$ . Those for  $V_0$  ranged from 22.3 to 26.6 ml with a mean of 24.3 ml, while for  $V_t - V_0$ , the mean was 11.2 and the range was 10.2–12.4 ml. Using these means to calculate  $Y = -\ln K_{av}$ , a linear relationship of high significance was found as  $Y = 0.001441m + 0.004881$  ( $r = 0.9972$ ; eight degrees of freedom) for the points 4–13 in Table I. The mean  $\bar{Y}$  corresponds to an elution volume of 30.39 ml, with  $\bar{m}$  equivalent to a MW of 13 623 daltons. The 90% confidence interval (CI) for the above regression coefficient was defined by the limiting values 0.001357 and 0.001526. Points 2 and 3 (Table I) when added to the graph fell well within these limits, while point 14, plotted similarly, fell within the 90% CI for the variation of  $Y$  about the regression.

With this much established, the series of triads was extended arbitrarily to a sample of size 20 covering the range of points 2–14 inclusive. Now using the better estimates of means for  $V_0 = 25.65$  ml (standard error = 0.15 ml) and for  $V_t - V_0 = 11.00$  ml ( $\pm 0.12$  ml), values for  $-\ln K_{av}$  were computed for the points 1–15 and plotted. Correlation for points 2–14 inclusive showed the relationship  $Y = 0.001490m + 0.00200$  ( $r = 0.9991$ , eleven degrees of freedom). The mean  $\bar{Y} = 0.82431$  corresponds to  $V_e = 30.47$  ml, with  $\bar{m} = 554.62$  equivalent to a MW of 13 060 daltons. The 90% CI for the regression coefficient was defined by limiting values of 0.001456 and 0.001524, while the standard deviation (S.D.) for  $\bar{Y}$  at 90% confidence was  $\bar{Y} \pm 0.0439$ . Graphical representation of these results (Fig. 1) shows that variation of  $Y$  for a given molecular weight depends largely on the latter factor, since that induced by the error of the slope is comparatively small.

Using the same methods as above, it was found, also for points 2–14 (Table I), that the regression of  $m$  to  $Y$  could be expressed as  $m = 670.0Y + 2.330$ . The limiting values of the regression coefficient for 90% confidence were 654.7 and 685.3, with the 90% CI  $\bar{m} \pm 29.44$ . Again, the error in  $m$  at any given value for  $Y$  consists mainly of the component of variance arising from the scatter about the regression. Graphical representation of the errors (90% CI) in the regression of  $m$  to  $-\ln K_{av}$  took a form similar to that of Fig. 1.

In the original work of Cullen et al. [1] calibration was effected by use of

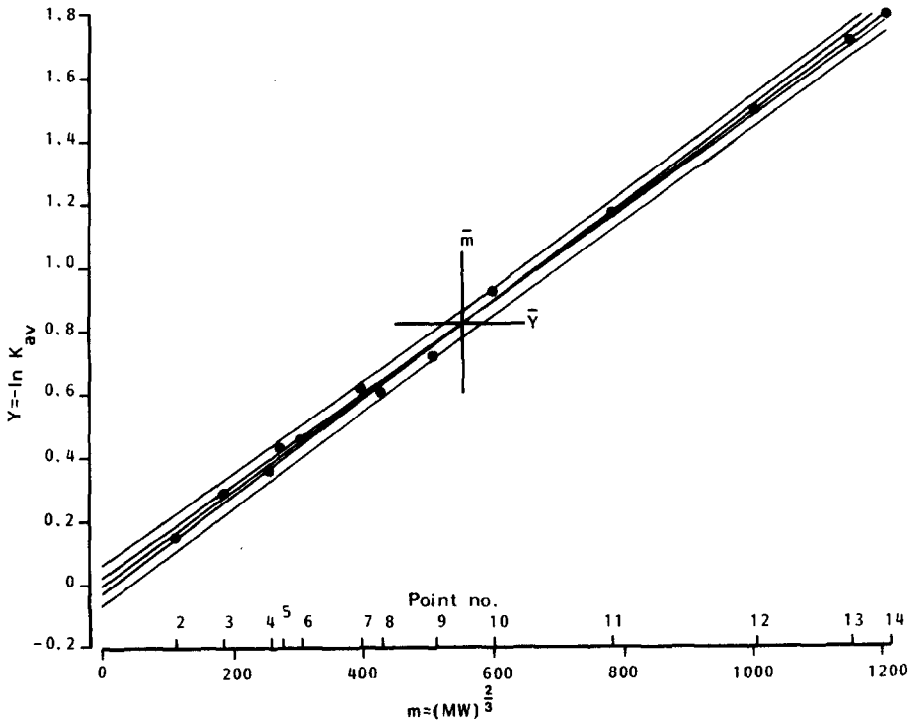


Fig. 1. Regression of  $Y = \ln K_{av}$  to  $m = MW^{2/3}$ . Mean slope with 90% CI for slope and for variation of  $Y$  about the slopes as outer limits.  $K_{av} = (V_e - 25.65)/11.00$ ;  $V_e$  values (ml) as means, points 2–14 (Table I) inclusive.

the significant linear relationship between  $K_{av}$  and  $\ln MW$ , and this method (B) was compared with the above (A). With data from Table I, using points 1 and 19 for  $V_0$  and  $V_t$ , respectively, the regressions (a)  $K_{av} = 1.7050 - 0.1178 \ln MW$  and (b)  $\ln MW = 14.35 - 8.4165 K_{av}$  ( $r = 0.9959$ , fifteen degrees of freedom) were established. With the same methods as used above with method A, 90% confidence intervals were calculated and a corresponding pair of graphs was now constructed for method B. The four graphs served as charts from which the ambits, i.e. intervals between maximum and minimum, could be found as 90% confidence intervals both for the mean elution volumes with given molecular weights and for the molecular weights estimated when the true mean weights were known. Results of these computations of ambits are shown in Fig. 2, from which it is clear that method A gives valuably greater precision for elution volumes than does method B. As estimators of molecular weights, both methods gave the same limits when the true molecular weight was approximately 2000 daltons, but with increasing molecular weights method A showed progressive superiority: only in the range 1000–2000 daltons did method B give an indication of greater precision.

Errors in the estimations of molecular weights from elution volumes may be expressed also as  $\pm$  percentages of the known mean weights since it was found that the limits were almost completely symmetrical about the means. In these terms, errors (90% CI) of method B are between  $\pm 34\%$  (2000 daltons) and  $\pm 30\%$  (40 000 daltons), contrasting with those of method A, which range from  $\pm 34\%$  (2000 daltons) to  $\pm 5.9\%$  (40 000 daltons) (Fig. 2). When method B was applied to points 2–14 (Table I) only, the status of the procedure in comparison with method A was found to have deteriorated, while application of method B with the derived values  $V_0 = 25.65$  and  $V_t - V_0 = 11.00$  ml gave no improvement.

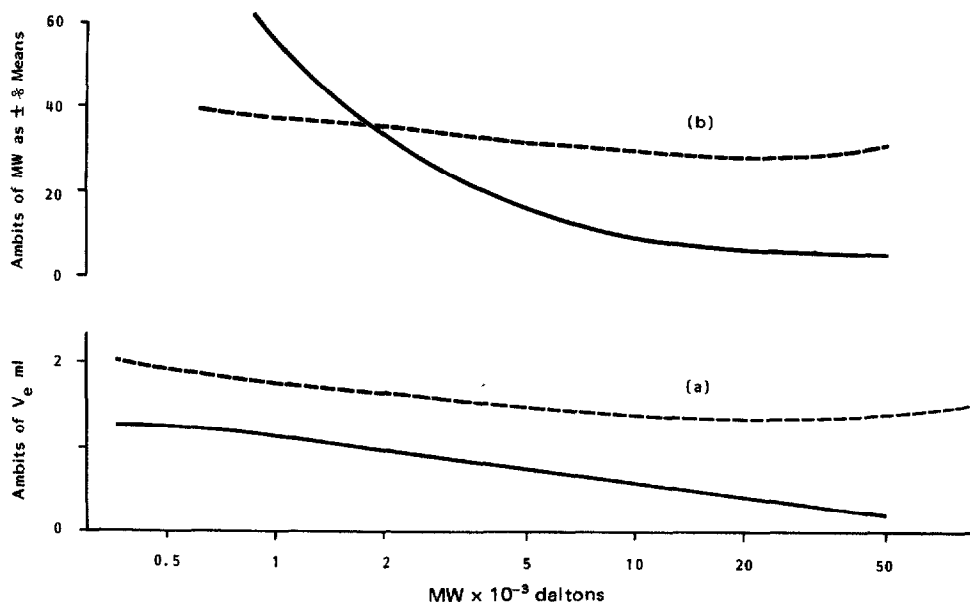


Fig. 2. Ambits (90% CI) for elution volumes as functions of molecular weights (a) and for molecular weights, as  $\pm$  percentage MW, estimated from elution volumes (b). Solid lines show results for method A, broken lines for method B.

## DISCUSSION

The results obtained show that the rectilinear calibration holds good for the system studied over a useful span of the interval between  $V_0$  and  $V_t$  as indicated by dextrose and highly polymerised dextran. This range allows, with calibration method A, successful treatment of dextrans having molecular weights from 1500 to 50 000 daltons. Near the centre of the working range 90% of mean elution volumes will fall within an ambit of  $\pm 0.20$  ml. With lower molecular weights, e.g. 1000 daltons, 90% of these volumes should be spread over  $\pm < 0.6$  ml, while at the other extreme (50 000 daltons) means for  $V_e$  should exceed an ambit of  $\pm 0.10$  ml from the central volume in only 10% of cases, one in twenty giving higher volumes and with the same proportion falling lower. These conclusions cannot be expected to apply to other than the iso-chemical series treated here.

Whereas the confidence limits for mean elution volumes may be of interest, as e.g. in deciding that volumes need be measured only to the nearest 0.1 ml, it is perhaps of greater importance to know how precisely molecular weights may be deduced once the mean for  $V_e$  has been estimated.

It is clear from Fig. 2 that calibration by method A is preferable, although the useful range is smaller than that with method B. It is also seen that only rough estimates are possible with the system studied if true molecular weights are much below 5000 daltons.

The formal use of confidence intervals in this study tends to give a disconcerting impression for two reasons. Firstly, in truth, fewer than 5% of cases will be values reaching limits for both sources of variation in the same sense. A proportion of the results that are high because of variation about the regression will be modified by a counter-balancing diminution due to the variability of the regression coefficient. Secondly, the central tendency will ensure that precision is influenced favourably in a preponderance of cases. Admittedly the statistical methods used, although robust, are only good approximations. In particular, it is known that although the standard molecular weights are truly invariate, the mean elution volumes are estimated means, with known small standard errors. None of the standard dextran fractions was homogenous; unfortunately, pure compounds are not available. It must be emphasized that mean elution volumes have been used throughout, so it is important that data be obtained by replication with rigid control of instrumental conditions, pumping rate and temperature being particularly pertinent.

To summarize the comparison of methods A and B treated here, it may be said that method A is preferable, but requires at least five to six standards to furnish an adequate sample of triads. Method B is quicker and more convenient to apply, but lacks precision, although the working range is wider. Its adequacy is equally dependent on the amount of data from standards. If method B is used in preliminary work, results of sufficient interest can then be re-examined by the use of the better but more laborious procedure.

It is noticeable that  $V_0$  forms a far greater proportion of  $V_t$  than the 25–30% that would be expected. This may reflect that the gel may contain a proportion of highly permeable particles. Whatever the cause, it is tempting to speculate that the inductive use of triads may explain the practical superiority

of method A. Failure of the calibration at lower molecular weights as  $V_e$  tends towards  $V_t$  may be ascribed to the fact that the shorter molecules cannot assume a random coil configuration.

In the system studied, two TSK gels (3000 PW and 5000 PW) were used in series. The above considerations might well apply to systems in which other grades of gel of similar manufacture are to be employed for investigations with different ranges of molecular weights in an isochemical series.

#### ACKNOWLEDGEMENTS

The authors express their appreciation to Dr. George B. Haycock for his consideration and assistance. The financial support of the Special Trustees of Guy's Hospital is appreciated and special thanks are due to Ms. Pauline Spink for her preparation of this manuscript.

#### REFERENCES

- 1 M.P. Cullen, C. Turner and G.B. Haycock, *J. Chromatogr.*, 337 (1985) 29.
- 2 J. Porath, *Pure Appl. Chem.*, 6 (1963) 233.
- 3 S. Hjertén, *J. Chromatogr.*, 50 (1970) 189.
- 4 R.W.R. Baker, *J. Chromatogr.*, 154 (1978) 3.