CHROM. 7793

# MOLECULAR-WEIGHT DISTRIBUTION DETERMINATION OF CLINICAL DEXTRAN BY GEL PERMEATION CHROMATOGRAPHY

GÖRAN NILSSON

*Department of Applied Mathematics, Pharmacia AB, Uppsala (Sweden)*

and

KJERSTIN NILSSON

*Department of Quality Control, Pharmacia AB, Uppsala (Sweden)*

(Received June 13th, 1974)

## SUMMARY

The method described utilizes polydisperse standards for the molecular-size calibration of columns. The only information on these standards required is an accurate measure of the mean of the molecular-size parameter of interest; size can be expressed in terms of weight, chain-length, hydrodynamic volume, etc. The principle for this calibration is described and discussed. Emphasis is placed on calibration according to molecular weight, using the $\bar{M}_w$ of the standards as a basic parameter.

Details are given on the systematic and random errors of the estimates of molecular-weight distributions, obtained by the method described.

The significance of different sources of error is briefly discussed.

## INTRODUCTION

When manufacturing clinical dextrans, accurate and reliable methods are needed for the molecular-weight characterization of the products. The successful application[1] of gel permeation chromatography (GPC) with Sephadex to the determination of molecular-weight distributions (MWDs) prompted us to introduce such methods as a complement to the conventional[2] fractionation methods used.

A general method for MWD determination, based on GPC with Sephadex or Sepharose, and suitable for different applications, had previously been developed in this company. The work reported in this paper is based on that method and developed with special regard to analytical applications. The principle of molecular-size calibration has also been dealt with further.

### Calibration

GPC of a polymer primarily gives its elution volume distribution. The elution volume of a molecular species is a function of

(1) its size or some monotonous function thereof; size is a fictitious variable, related to the mechanism of separation;

(2) operational variables, such as column size, gel, solvent, temperature, loading, flow-rate;

(3) variables other than (2), influencing, for instance, curve broadening.

The aim of the molecular-size calibration is to establish an expression that relates elution volume to molecular size at given values of operational and other variables. The parameters governing GPC and the relationship between elution volume and molecular size have been dealt with in many papers. The calibration methods applied may, according to their principles, be| classified as follows ("known average molecular size" should be understood as the size measured using an absolute method):

(1) A number of monodisperse or narrow polymer fractions with known averages of, for instance, molecular weight, limiting viscosity number, chain-length, degree of polymerization, of the polymer of interest, are used. As the fractions are narrow, it is assumed that the mode or mean of the volume distribution corresponds to the average of the molecular-size variable. For many polymers, sufficiently narrow fractions are not readily available.

(2) The column is calibrated according to, for instance, method (1), with an auxiliary polymer, and a so-called universal calibration variable, $u$, is used as molecular size. The variable $u$ is a function of molecular weight with one or more known or readily measurable parameters characteristic of each polymer. It is assumed that the relationship between $u$ and the elution volume is independent of the type of polymer used. As $u$ is a known function of the molecular weight, it is possible to calculate the molecular-weight distribution from the distribution of $u$.

(3) The column is calibrated according to, for instance, method (1), with an auxiliary polymer. A number of fractions, not necessarily narrow, with known average molecular sizes of the polymer of interest are available. It is assumed that the relationship between the molecular size of the polymer of interest $(m)$ and that of the auxiliary polymer $(z)$, corresponding to the same elution volume, can be expressed as a mathematical function, usually $m = a \cdot z$ or $m = a \cdot z^{b}$. The parameters $a$ and $b$ are determined in such a way that the agreement between the average molecular weights given and those calculated is as good as possible according to some criterion.

(4) A number of (broad) fractions with known average molecular sizes are available. It is assumed that the molecular-size distributions can be expressed as a mathematical function, with the number of unknown parameters not exceeding the number of known average molecular sizes for each fraction. From the known average molecular sizes, the parameters can be calculated and, from the percentiles of this theoretical molecular size distribution and the corresponding percentiles of the elution volume distribution observed, the calibration curve is obtained.

(5) A number of (broad) fractions with known average molecular sizes are available. It is assumed that the calibration curve can be expressed as a mathematical function with two or more unknown parameters. The parameters are determined so that the agreement between the known average molecular sizes and the average molecular sizes calculated from the GPC curves is as good as possible according to some criterion.

(6) A number of (broad) fractions with known average molecular sizes are available. A preliminary calibration curve is successively improved by trial and error

until an "acceptable" agreement is obtained between the average molecular sizes given and those calculated from the chromatograms.

(7) One or more broad fractions with a completely known molecular-size distribution are available, and the calibration curve is obtained from the corresponding percentiles of the molecular size distributions given and the elution-volume distributions observed. As the molecular-size distributions of the calibration fractions have to be estimated by some other method, the errors of these estimates, both systematic and random, will cause systematic errors in all subsequent GPC results.

In addition, various methods have been proposed for molecular-size calibration, in which the calibration curves are deduced from assumptions or measurements of the gel pore-size distribution and theories of the separation mechanism. So far, however, these methods have, in general, been used less for estimating MWD than for elucidating GPC mechanisms.

In the applied calibration methods, ln $M$ ($M$ = molecular weight) is often assumed to be a linear function of the elution volume, but polynomials of higher degrees have also been used. In order to define the so-called universal calibration curve, ln $([\eta]M)$ as a function of elution volume ($[\eta]$ = limiting viscosity number), Shultz et al.[3] used a third-degree polynominal, and Ambler[4] used fourth- and fifth-degree polynomials.

The calibration method described in this paper refers to principle (5) (see above). In relating elution volume to molecular weight, we have found acceptable flexibility by expressing $M$ as the sum of a constant and an exponential function of a third-degree polynominal of $K_{\text{av}}$ (see *Notations*). The parameters of the function are determined so that maximal agreement, according to the method of least squares, is obtained betwemn $\bar{M}_w$ values, estimated by light scattering for at least six standard fractions, and the corresponding values estimated from GPC curves. A number of calibration methods, referring to principle (5), are described in the literature. A few examples are given below.

Frank et al.[5] assumed that the relationship between molecular weight ($M$) and elution volume ($V$) within the range of each standard can be approximated by

$$\ln M = a - b\, V \qquad (1)$$

The parameters $a$ and $b$ are not assumed to be the same for all standards, and if both $\bar{M}_n$ and $\bar{M}_w$ are known, $a$ and $b$ can be determined for each calibration standard.

By assuming a model for the curve broadening, a mathematical expression of the form of the chromatograms of standards, and also assuming that the relationship between $M$ and $V$ follows eqn. 1, Almin[6] was able to estimate the parameters of the assumed functions from a number of standards with known GPC curves and known number- and weight-average molecular weights.

One or two broad standards with two measured average molecular weights (number-, weight- or viscosity-), were used by Balke et al.[7] to estimate the parameters in eqn. 1. This was achieved by a Rosenbrock search[8]. Balke et al. suggested that the method could be modified so as to determine the parameters of a non-linear calibration curve, or the existing method might be used to approximate a non-linear curve by straight-line segments.

Whitehouse[9] utilized standards with known $\bar{M}_w$ and $\bar{M}_n$ values and determined

the parameters $a$ and $b$ in eqn. 1 by varying initial estimates in a stepwise manner, keeping first the intercept constant.

*Capability*

According to accepted practice, clinical dextrans are controlled by light-scattering determination of the $\bar{M}_w$ of the total distribution of $M$, and the $\bar{M}_w$ of a 10% low-molecular-weight and a 10% high-molecular-weight fraction of the distribution[2]. This is later referred to as the "current method".

When introducing a new method, which is supposed to be superior, the capability of the current and the new method must, of course, be compared with one another and with the capability of the production process of the samples to be measured. For the GPC method, a capability study was performed and systematic errors and sources of variation are discussed. For the current method[2] and the production process, the standard deviations of the weight-average molecular weights of the total distributions are reported.

*Quality control of the GPC method*

Direct control of the errors in analytical results is, of course, impossible when the samples are unknown. Hence the control has to be concentrated upon variables which vitally affect the measuring process. The principal features of the current quality control of the GPC measurement process are briefly described.

EXPERIMENTAL

*Notations*

   *Theoretical variables and parameters*
   $w_i$ = weight proportion of molecular species $i$;
   $n_i$ = number proportion of molecular species $i$;
   $m_i$ = molecular weight of species $i$;
   $\mu_w$ = weight-average molecular weight

   $$= \sum_{i=1}^{\infty} w_i\, m_i;$$

   $\mu_n$ = number-average molecular weight

   $$= \sum_{i=1}^{\infty} n_i\, m_i = \frac{1}{\sum_{i=1}^{\infty} \dfrac{w_i}{m_i}};$$

   $\sigma_w$ = standard deviation of the MWD

   $$= \sqrt{\sum_{i=1}^{\infty} w_i\, (m_i - \mu_w)^2}.$$

   *Observed variables*
   $V_p$ = pre-fraction volume;
   $V_f$ = effluent fraction volume;

$$V_e = V_p + j \cdot V_f$$

= elution volume after $j$ effluent fractions;

$V_0$ = void volume of the column;

$V_t$ = total volume of the column;

$$K_{av} = \frac{V_e - V_0}{V_t - V_0}$$ (as an argument in a function denoted by $v$ for short);

$k$ = number of effluent fractions;

$v_j$ = $K_{av}$ corresponding to the mid-point of the $j$th effluent fraction;

$x_j$ = dextran concentration of the $j$th effluent fraction;

$$f_h = \frac{x_h}{\sum\limits_{j=1}^{k} x_j}$$

= weight proportion of dextran in the $h$th effluent fraction, $h = 1, 2, \ldots \ldots k$;

$$F_h = \sum\limits_{j=1}^{h} f_j.$$

*Calibration curve*

$m(v)$= molecular weight ($m$) as a function of $v$ (see *GPC method, calibration technique*).

*Estimates.*

$$\hat{\mu}_w = \sum\limits_{j=1}^{k} f_j \, m \, (v_j)$$

= estimate of $\mu_w$;

$$\hat{\sigma}_w = \sqrt{\sum\limits_{j=1}^{k} f_j \, [m(v_j) - \hat{\mu}_w]^2}$$

= estimate of $\sigma_w$;

$$\hat{\mu}_n = \frac{1}{\sum\limits_{j=1}^{k} \dfrac{f_j}{m(v_j)}}$$

= estimate of $\mu_n$;

$K_p$ = estimate of the $p$ percentile of the $K_{av}$ distribution, obtained by linear interpolation of the observed distribution $\{f_j\}$;

$M_p = m(K_{100-p})$

= estimate of the $p$ percentile of the molecular weight distribution;

$\bar{M}_w = \mu_w$ estimated from light-scattering measurement (see *Material*);

$\bar{M}_n = \mu_n$ estimated from end-group analysis (see *Material*).

*Equipment*

The light-scattering instrument was a Sofica, Type 42.000. The GPC column was a Pharmacia Fine Chemicals AB (Uppsala, Sweden) Type K 16/70, 70 × 1.6 cm, provided with a thermostat jacket and a sample applicator. A Stålprodukter (Uppsala, Sweden) Type SP 1 peristaltic pump was used, and the fraction collector was a Stålprodukter, Type Fractomin FR 4, with 2 × 80 cups.

The AutoAnalyzer system consisted of: (1) a Technicon AB Type II pump; (2) a Stålprodukter Sampletron sampler; (3) a multichannel absorptiometer (LKB Produkter AB, Bromma, Sweden), Type 5901 A, fitted with two 625-nm filters and two Technicon continuous flow cells; (4) two Yokogawa 3046/C laboratory recorders; (5) a Technicon Type I heater with two 13-m delay coils.

An Eppendorf flame photometer was employed.

A Hewlett-Packard disc-operating HP 2120 computer system was used consisting of: (1) a central unit, Type HP 2100 A, 16K words; (2) an input device, Type HP 2892 A, card reader; (3) an output device, Type HP 2610, line printer; (4) a Tektronix Type 4010/4010-1 computer display terminal, with Tektronix Type 4610/4610-1 hard copy unit; (5) a console typewriter, Type HP 2752 A; (6) a disc memory, Type HP 12960 A.

*Material*

Twelve fractions of dextran were used for the calibration (see Table I). $\bar{M}_w$ values were estimated by light-scattering measurements. The solvent was water, and measurements were performed at 436 nm at 45, 90 and 135 angular degrees to the incident beam for fractions with $\bar{M}_w < 200,000$ and at six angles, ranging from 30 to 150 degrees, for fractions with higher $\bar{M}_w$. The temperature was 20°. The instrument was calibrated with benzene, distilled in a Fischer fractionating column, packed with glass helices. The following physical constants were used in the calculations: refractive index increment for dextran at 436 nm, 0.152 ml/g; Rayleigh ratio 90° for benzene at 436 nm, $48.5 \cdot 10^{-6}$; refractive index for benzene at 436 nm, 1.523.

TABLE I

DEXTRAN FRACTIONS USED FOR CALIBRATION

| Dextran fraction No. | $\bar{M}_w$ | | | $\bar{M}_n$ | | |
|---|---|---|---|---|---|---|
| | *n* | *Mean* | *S.D.* | *n* | *Mean* | *S.D.* |
| 1 | 3 | 3,370 | 60 | 4 | 1,950 | 45 |
| 2 | 2 | 10,900 | 140 | 3 | 5,870 | 210 |
| 3 | 4 | 21,400 | 250 | 4 | 14,900 | 485 |
| 4 | 3 | 27,800 | 175 | 3 | 19,800 | 60 |
| 5 | 6 | 39,900 | 365 | 3 | 26,000 | 305 |
| 6 | 3 | 50,400 | 305 | 5 | 37,500 | 505 |
| 7 | 4 | 59,500 | 425 | 1 | 34,900 | — |
| 8 | 9 | 71,200 | 2,200 | 9 | 38,900 | 965 |
| 9 | 3 | 100,900 | 8,210 | 7 | 66,000 | 2,815 |
| 10 | 5 | 146,100 | 4,345 | 7 | 93,700 | 8,455 |
| 11 | 5 | 252,300 | 7,825 | 6 | 112,000 | 4,070 |
| 12 | 4 | 484,800 | 7,500 | 8 | 182,000 | 11,175 |

$\bar{M}_n$ values were estimated by end-group analysis according to the Somogyi copper phosphate method[10].

## GPC procedure

The bed material was Sepharose 4B (Pharmacia). Before use, the column was eluted with the eluent (see below) for 1 week.

A solution, containing 0.1 % of the dextran sample, 0.3 % of sodium chloride and 0.01 % of potassium chloride, was prepared as the sample solution. The solution for the determination of $V_0$ contained 0.5 % of Dextran 2000 ($\bar{M}_w \approx 2 \cdot 10^6$) (Pharmacia) and 0.3 % of sodium chloride.

The eluent was a 0.3 % solution of sodium chloride, containing 0.001 % of phenyl mercury(II) acetate.

After application of the sample to the column, the first effluent, of volume at least about $V_0$, is free from dextran. The approximate amount of dextran-free effluent was calculated for the individual sample from assumptions of its MWD. This volume, reduced by a safety margin, was collected and measured separately as $V_p$.

The flow-rate used was 6–8 ml/h and the temperature was 20°.

*General analysis procedure.* The column was run with a downwards flow. Eluent was fed from a Mariotte flask, the flow-rate being adjusted using a peristaltic pump at the outlet of the column. The sample (or solution for the determination of $V_0$ or dextran fraction for calibration) and eluent were applied to the column via a three-position valve. Precautions were taken to reduce evaporation during the fractionation.

Effluent volumes were determined by weighing. The weight in grams was taken as the volume in millilitres. The mean of the volume of the last five fractions was taken as $V_f$.

The content of dextran in the effluent fractions was determined automatically in an AutoAnalyzer system[11]. A colorimetric method of analysis was used, based on the reaction between glucose and anthrone[12]. A calibration curve was fitted to the same mathematical function as described for the molecular weight calibration (see *GPC method, calibration technique*).

*Determination of $V_0$ of the column.* About 0.5 ml of solution for the determination of $V_0$ (see above) was applied to the column. A volume corresponding to ca. $V_0 - 3$ ml was collected separately as $V_p$ and weighed. Thirty effluent fractions, containing 0.20–0.25 ml, were collected. $V_f$ was determined as described above.

Before the determination of dextran in the effluent fractions, 0.50 ml of water was added to every fraction with an Eppendorf pipette. The dextran content was plotted against $V_e$. The volume, corresponding to the inflexion point of the ascending part of the dextran content curve, was defined as $V_0$. $V_0$ was determined once every fortnight.

*Determination of MWD.* About 1.5 ml of sample solution (see above) was applied to the column. After the separate collection and weighing of $V_p$, 50–60 effluent fractions, containing 1.5–2.5 ml, were collected. $V_f$ was determined as described above.

The content of potassium in about 15 effluent fractions at and around the expected total volume of the column was determined flame photometrically. The potassium content was plotted against $V_e$. The volume corresponding to the mean of the potassium content curve was defined as $V_t$. Thus, $V_t$ was determined simultaneously at every determination of a MWD.

The content of dextran in the effluent fractions was determined as described above.

*Quality control of the GPC method*
The control of the dextran concentration measurements embraces: (a) goodness of fit of the dextran content calibration curve; (b) changes in the blank level; (c) interaction between samples; (d) concentration range of the samples in relation to the range of the calibration; (e) irregularities in the AutoAnalyzer curves.

The following operational variable and parameters are included in the control plan: $V_f$: mean volume in the interval 1.5–2.5 ml; the difference between the first five and last five effluent fractions shall be less than 4%; $V_t$: the deviation from the mean of the calibration runs shall be less than 1 ml; $V_0$: the deviation from the preceding determination shall be less than 0.5 ml.

The $K_{av}$ distribution is controlled with regard to: (a) percentage of dextran in the first and last fractions; this shall be close to zero; (b) the range; not more than 1% outside 0.1–0.9 is accepted.

*GPC method, calibration technique*
It is assumed that the calibration curve can be expressed as

$$m(v) = b_5 + \exp(b_4 + b_1 v + b_2 v^2 + b_3 v^3) \qquad (2)$$

where $b_1, b_2, b_3, b_4$ and $b_5$ are constants that are determined according to the method of least squares.

In principle the calibration technique is as follows:

(1) a number ($n \geqslant 6$) of dextran fractions is selected; the $\bar{M}_w$ values have to cover the interval of molecular weights for which the calibration curve is to be valid;

(2) a GPC analysis is performed and the set of $v_j$ and $f_j$ for each fraction is calculated;

(3) those values of $b_1 \ldots b_5$ in eqn. 2 for which

$$\sum_{i=1}^{n} \left[ \frac{\hat{\mu}_w(i) - \bar{M}_w(i)}{\bar{M}_w(i)} \right]^2$$

has its minimum are found (the summation variable $i$ refers to the standard fraction number).

The minimizing is carried out by using Hartley's modification[13] of the Gauss–Newton method.

*Capability*
*GPC method, systematic errors.* Six columns were calibrated with twelve standard fractions each (see Table I). This was denoted as the complete calibration of the column in question. In order to investigate the influence of the selection of standard fractions on the calibration curves, the calibration procedure for each column was also performed with the following six combinations of the twelve standard fractions:

| Combination | Standard fractions |
|---|---|
| I | All except number 1 |
| II | All except number 12 |
| III | All except numbers 1 and 12 |
| IV | All except numbers 7 and 8 |
| V | Numbers 1, 3, 5, 7, 9, 11 |
| VI | Numbers 2, 4, 6, 8, 10, 12 |

In this way, seven different calibration curves were obtained for each column. The curves were compared and the differences are analyzed and discussed (see Results and discussion, *GPC method, systematic errors*).

From the complete calibrations (see above), $\hat{\mu}_w/\bar{M}_w$ and $\hat{\mu}_n/\bar{M}_n$ were calculated for the twelve standard fractions and the six columns. Means and standard deviations of these quotients, obtained for the fractions on the different columns, were calculated. Deviations from unity were tested by the $t$-test. Possible causes of the deviations are discussed (see Results and discussion, *GPC method, systematic errors*).

*GPC method, components of variance.* Two test substances, one dextran 40, $\bar{M}_w \approx 40,000$, and one dextran 70, $\bar{M}_w \approx 70,000$, were analyzed three times on each of the six columns. Between each analysis on a column, a new determination of $V_0$ was performed. Mean values, standard deviations and percentiles of the $K_{av}$ and molecular weight distributions were determined (see Results and discussion, *GPC method, components of variance*). The complete calibration curves (see above) were used for the calculation of the MWDs. Of the 36 GPC analyses of this series, three were rejected and repeated on account of the quality control specifications for the GPC method (see Experimental, *Quality control of the GPC method*).

*Variances of measurement and production processes.* During the last 4 years, 71 batches of dextran 40 and 25 batches of dextran 70 have been analyzed according both to the current[2] and the GPC method. The weight-average molecular weights of the total distributions were estimated and compared (see Results and discussion, *Variances of measurement and production processes*).

## RESULTS AND DISCUSSION

An experimental result has no quantitative meaning unless the measurement process is in a state of statistical control and a characterization of its accuracy with respect to systematic errors and components of variance is given. The concepts and techniques of statistical process control are today generally applied to measurement processes in the laboratory. An excellent exposition of this topic was given by Eisenhart[14], and in a review by Currie *et al.*[15] a comprehensive survey of the subject is presented.

For the GPC method described in this paper, the causes and magnitudes of systematic errors are discussed and the components of variance are estimated from the capability study. In order to be able to improve the accuracy and to determine whether or not the measurement process is in a state of statistical control, it is necessary to identify the different sources of variation and to investigate their contribution to systematic errors and the components of variance.

*GPC method, systematic errors*

For the complete calibration curves and the calibration curves based on combinations I–VI of the twelve standard fractions (see Experimental, *Capability*), the molecular weights corresponding to $K_{av} = 0.1, 0.2, \ldots, 0.9$ were calculated for all columns. For each $K_{av}$ and combination of standards, the mean and standard deviations of the differences in molecular weight from the complete calibration curves were calculated and expressed as a percentage of the mean of the complete calibration curves. The importance of different combinations of standards was considerable at $K_{av} = 0.1$ and 0.9 and minor at $K_{av} = 0.2$–0.8 (see below).

At $K_{av} = 0.1$ and 0.9, considerable differences occurred for some combinations of standards. Exclusion of the extreme standard fractions resulted in immense changes of the corresponding ends of the calibration curves. Exclusion of standard fractions in the middle of the range, however, had minor effects. A comparison by analysis of variance between combinations V and VI, which have no standard fractions in common, showed significant effects of both columns and the selection of standard fractions.

In order to ensure good separation conditions, however, the use of a column has been restricted to samples with $\geqslant 99\%$ of their distribution of $K_{av}$ in the range 0.1–0.9. The corresponding figure both for dextran 40 and 70 is $> 99.5\%$.

At $K_{av} = 0.2$–0.8, corresponding to $> 90\%$ of the distribution of dextran 40 and $> 95\%$ of the distribution of dextran 70, none of the standard deviations of differences exceeded 5%, and most of them were less than 2%. Two of the mean differences were between 5 and 10%, but most of them were less than 2%.

The effects of the mathematical form chosen for the calibration curve, the calibration fitting method, errors in estimates of average molecular weights used in the calibration, and broadening of MWDs by GPC, can best be shown by a comparison of average molecular-weight estimates obtained by light-scattering measurements or end-group analysis and by GPC on different columns.

For each of the twelve standard fractions, means and standard deviations for $\hat{\mu}_w/\bar{M}_w$ and $\hat{\mu}_n/\bar{M}_n$ for the six different columns were calculated (see Experimental, *Capability*, and Table II).

If the mathematical form of the calibration curve is not flexible enough to fit the true relationship between $K_{av}$ and molecular weight, systematic errors will appear. This is probably the case outside the range of the average molecular weights of the standard fractions. A bad fit may, of course, also be due to errors in the molecular-weight estimates used in the calibration, *i.e.*, errors in $\bar{M}_w$.

Not only the mathematical form of the curve, but also the method of fitting the curve to the observed data and the criteria for the optimal solution may give rise to systematic errors. Many methods for the correction of broadening of MWDs by GPC have been proposed[16–18]. However, such methods have not been applied in the GPC method described. A problem is how to estimate the effect on systematic errors. A study of the quotients $\hat{\mu}_n/\bar{M}_n$ may give some information.

Values of the quotients close to unity are a necessary but not sufficient condition for negligible broadening. The random variation of $\hat{\mu}_n/\bar{M}_n$ (see Table II) seems, in most instances, to be more essential than a systematic error. A significant deviation from unity may, of course, also be explained by errors in the estimate of $\bar{M}_n$.

TABLE II

MEANS AND STANDARD DEVIATIONS FOR $\hat{\mu}_w/\bar{M}_w$ AND $\hat{\mu}_n/\bar{M}_n$ FOR THE SIX COLUMNS

Significant deviations from unity (t-test) are denoted by: $^* = p < 0.05$; $^{**} = p < 0.01$; $^{***} = p < 0.001$.

| Standard fraction | $\bar{M}_w$ | $\hat{\mu}_w/\bar{M}_w$ | | $\bar{M}_n$ | $\hat{\mu}_n/\bar{M}_n$ | |
|---|---|---|---|---|---|---|
| | | Mean | S.D. | | Mean | S.D. |
| 1 | 3,370 | 1.011** | 0.006 | 1,950 | 1.176 | 0.186 |
| 2 | 10,900 | 1.005 | 0.012 | 5,870 | 1.096 | 0.182 |
| 3 | 21,400 | 0.995 | 0.006 | 14,900 | 0.974 | 0.049 |
| 4 | 27,800 | 0.993 | 0.016 | 19,800 | 0.990 | 0.038 |
| 5 | 39,900 | 1.015* | 0.011 | 26,000 | 1.006 | 0.032 |
| 6 | 50,400 | 1.015** | 0.005 | 37,500 | 1.011 | 0.019 |
| 7 | 59,500 | 0.983 | 0.011 | 34,900 | 0.914*** | 0.022 |
| 8 | 71,200 | 0.992 | 0.018 | 38,900 | 0.976 | 0.031 |
| 9 | 100,900 | 0.983** | 0.007 | 66,000 | 0.966* | 0.031 |
| 10 | 146,100 | 1.026** | 0.014 | 93,700 | 0.993 | 0.049 |
| 11 | 252,300 | 0.989 | 0.018 | 112,000 | 1.009 | 0.040 |
| 12 | 484,800 | 1.003 | 0.008 | 182,000 | 0.999 | 0.046 |

*GPC method, components of variance*

From the analyses of the two test substances, the components of variation, within and between columns, were estimated for the determined variables. We have assumed the following model:

$x_{ij} = \mu + a_i + e_{ij}$;

$x_{ij}$ = determination number $j$ on column number $i$ of variable $x$;

$\mu$ = total mean;

$a_i$ = bias of column $i$, including the calibration error;

$e_{ij}$ = error due to the variation of measurements within columns.

Both $\{a_i\}$ and $\{e_{ij}\}$ are regarded as random variables with Var $a_i = \sigma_a^2$ and Var $e_{ij} = \sigma^2$. $\sigma_a$ and $\sigma$ are estimated from analysis of variance, and the hypothesis $\sigma_a = 0$ is tested by the $F$-test. The results are given in Table III. $s_a$ and $s$ are the estimates of $\sigma_a$ and $\sigma$, respectively.

The total standard deviation of the method is estimated by

$$s_T = \sqrt{s_a^2 + s^2}$$

Obvious differences between the columns were found for all variables in the $K_{av}$ distribution. Significant differences in the molecular-weight distributions remain, but the quotients $s_a/s$ are smaller, and for $\hat{\mu}_w$ the differences between columns are not significant. It is not surprising that the smallest differences between columns are found in the $\hat{\mu}_w$ values, as this variable is used in the calibration procedure. A rough study of the individual deviations from the means did not contradict symmetrical distributions of the errors within and between the columns.

*Sources of variation of the GPC method*

From Table III, it is evident that the dominating component of variance for the estimates of percentiles of an MWD is between columns. Consequently, the most

## TABLE III

## RESULTS OF THE CAPABILITY STUDY

The hypothesis $\sigma_a = 0$ is tested by the $F$-test and significant results are denoted by: $* = p < 0.05$, $** = p < 0.01$, $*** = p < 0.001$.
The selection of the rather odd percentiles 2.5, 22.5, etc., has historical reasons. When the average molecular weights were calculated manually, the mid-point of each 5% class of the distribution was calculated by interpolation and the mean of these mid-points was used as an estimate of the weight-average molecular weight. These percentiles are still calculated (by the computer) and are used to describe the distribution. Observe that $K_{2.5}$ corresponds to $M_{97.5}$, and so on.

| Dextran | Variables and estimates | Mean | $s_a$ | $s$ | $s_T = \sqrt{s_a^2 + s^2}$ | Coefficient of variation[†] $c_a$ (%) | $c$ (%) | $c_T$ (%) | $\dfrac{s_a}{s}$ |
|---|---|---|---|---|---|---|---|---|---|
| 40 | $V_0$ | 30.68 | 2.24*** | 0.10 | | 7.3 | 0.3 | | 23.6 |
| | $V_t$ | 113.29 | 2.80*** | 0.42 | | 2.5 | 0.4 | | 6.6 |
| | $K_{av}$ distribution | | | | | | | | |
| | Mean | 0.689 | 0.007*** | 0.001 | | | | | 6.5 |
| | S.D. | 0.076 | 0.001*** | 0.001 | | | | | 1.7 |
| | $K_{2.5}$ | 0.527 | 0.004*** | 0.002 | | | | | 2.6 |
| | $K_{22.5}$ | 0.634 | 0.007*** | 0.001 | | | | | 6.3 |
| | $K_{47.5}$ | 0.689 | 0.007*** | 0.001 | | | | | 6.5 |
| | $K_{47.5}$[‡‡] | 0.689 | 0.007*** | 0.004 | | | | | 2.0 |
| | $K_{77.5}$ | 0.746 | 0.008*** | 0.001 | - | | | | 5.6 |
| | $K_{97.5}$ | 0.826 | 0.009*** | 0.002 | | | | | 4.6 |
| | $MWD$ | | | | | | | | |
| | $\hat{\mu}_n$ | 25500 | 910*** | 380 | 990 | 3.6 | 1.5 | 3.9 | 2.4 |
| | $\hat{\mu}_w$ | 40200 | 260 | 340 | 430 | 0.6 | 0.8 | 1.1 | 0.8 |
| | $\hat{\sigma}_w$ | 27300 | 860*** | 300 | 910 | 3.1 | 1.1 | 3.3 | 2.9 |
| | $M_{2.5}$ | 8200 | 650*** | 220 | 690 | 8.0 | 2.7 | 8.4 | 3.0 |
| | $M_{22.5}$ | 20500 | 540*** | 300 | 620 | 2.6 | 1.4 | 3.0 | 1.8 |
| | $M_{52.5}$ | 34800 | 330* | 350 | 470 | 1.0 | 1.0 | 1.4 | 1.0 |
| | $M_{77.5}$ | 53800 | 330 | 390 | 510 | 0.6 | 0.7 | 1.0 | 0.8 |
| | $M_{97.5}$ | 112800 | 2690*** | 1140 | 2920 | 2.4 | 1.0 | 2.6 | 2.4 |
| 70 | $V_0$ | 30.71 | 2.27*** | 0.07 | | 7.4 | 0.2 | | 32.9 |
| | $V_t$ | 113.38 | 2.87*** | 0.41 | | 2.5 | 0.4 | | 7.0 |
| | $K_{av}$ distribution | | | | | | | | |
| | Mean | 0.624 | 0.007*** | 0.001 | | | | | 6.6 |
| | S.D. | 0.093 | 0.002*** | 0.001 | | | | | 2.4 |
| | $K_{2.5}$ | 0.436 | 0.004*** | 0.002 | | | | | 1.8 |
| | $K_{22.5}$ | 0.553 | 0.006*** | 0.001 | | | | | 4.8 |
| | $K_{47.5}$ | 0.621 | 0.008*** | 0.001 | | | | | 6.9 |
| | $K_{47.5}$[‡‡] | 0.621 | 0.008*** | 0.004 | | | | | 2.2 |
| | $K_{77.5}$ | 0.675 | 0.009*** | 0.002 | | | | | 6.2 |
| | $K_{97.5}$ | 0.796 | 0.010*** | 0.003 | | | | | 3.7 |
| | $MWD$ | | | | | | | | |
| | $\hat{\mu}_n$ | 40700 | 1400*** | 740 | 1590 | 3.5 | 1.8 | 3.9 | 1.9 |
| | $\hat{\mu}_w$ | 69200 | 200 | 570 | 610 | 0.3 | 0.8 | 0.9 | 0.4 |
| | $\hat{\sigma}_w$ | 48800 | 1570*** | 720 | 1730 | 3.2 | 1.5 | 3.5 | 2.2 |
| | $M_{2.5}$ | 11900 | 850*** | 390 | 940 | 7.2 | 3.2 | 7.9 | 2.2 |
| | $M_{22.5}$ | 33000 | 960*** | 430 | 1050 | 2.9 | 1.3 | 3.2 | 2.3 |
| | $M_{52.5}$ | 59200 | 640** | 490 | 810 | 1.1 | 0.8 | 1.4 | 1.3 |
| | $M_{77.5}$ | 95100 | 860* | 760 | 1150 | 0.9 | 0.8 | 1.2 | 1.1 |
| | $M_{97.5}$ | 197100 | 5230*** | 2440 | 5780 | 2.7 | 1.2 | 2.9 | 2.1 |

[†] $c_a$, $c$ and $c_T$ are $s_a$, $s$ and $s_T$ as a percentage of the mean.
[‡‡] Calculated from means of $V_0$ and $V_t$.

important sources of variation are those which contribute to this component of variance. We shall now discuss different sources of variation and how they contribute to the variation in estimates of percentiles, suitably expressed in terms of the coefficient of variation ($c$). For the variation within columns, $c$ can be obtained from the standard deviation of $K_{av}$ by multiplication by $m'(v)/m(v)$. The term $m'(v)/m(v)$ is the derivative of the calibration curve divided by the corresponding molecular weight. For the six columns in the capability study, the curves for $m'(v)/m(v)$ are very close to each other, and an idea of their appearance can be gained from Table IV.

TABLE IV

$m'(v)/m(v)$ VALUES

| $K_{av}$ | $m \cdot 10^{-3}$ | $m'(v)/m(v)$ |
|---|---|---|
| 0.1 | 2000 | −10 |
| 0.4 | 250 | − 6 |
| 0.6 | 70 | − 7 |
| 0.8 | 11 | −12 |
| 0.9 | 2 | −25 |

The sources of variation considered to be the most important are discussed below.

(1) Errors of estimates of column parameters supposed to be constant for each column, i.e., $b_1, \ldots, b_5$ in the calibration curve (eqn. 2). These errors contribute only to the variation between columns and are probably the main source of this component of variation. For a more thorough investigation of the effect of these errors, two or more independent calibrations of each column are required.

(2) A real variation of the parameters $b_1, \ldots, b_5$. If this variation is random, a contribution to the variance within columns is obtained. Systematic trends in the parameters cause contributions both to the variance within columns and systematic errors. As the variance within columns is less than the variance between columns, the real variation of $b_1, \ldots, b_5$ seems to be negligible in relation to the errors in their estimates.

(3) A real variation of $V_0$ or error of the estimate of $V_0$, a column parameter considered to be constant during fixed time intervals. If the difference between the estimate and the actual value of $V_0$ is $\Delta V_0$, we obtain an error in $K_{av} \approx \Delta V_0 \cdot \partial K_{av}/\partial V_0 = - \Delta V_0 \cdot (1 - K_{av})/(V_t - V_0)$. This source of variation affects only the variance within columns.

(4) An error of the estimate of the column parameter $V_t$. As $V_t$ is estimated in every GPC analysis, an error contributes only to the variance within columns. An error in the estimate equal to $\Delta V_t$ corresponds to an error in $K_{av} \approx \Delta V_t \cdot \partial K_{av}/\partial V_0 = - \Delta V_t \cdot K_{av}/(V_t - V_0)$. In the capability study, the 47.5 percentiles of the $K_{av}$ distributions were also calculated, using the means of $V_t$ and $V_0$ instead of the current values. As can be seen in Table III, the standard deviation within columns is increased (3–4 times). This implies that the errors of the estimates are small in relation to the actual variation, and confirms earlier findings[19] that corrections for the current values of at least $V_t$ reduces the variance within columns.       .

(5) Errors of the estimates of the operational variables $V_p$ and $V_f$ only have an influence on the variance within columns. If the errors in the estimates are $\Delta V_p$ and $\Delta V_f$, respectively, the corresponding errors in $K_{av}$ are $\Delta V_p \cdot \partial K_{av}/\partial V_p = \Delta V_p \cdot (1 - K_{av})/(V_t - V_0)$ and $\Delta V_f \cdot \partial K_{av}/\partial V_f = \Delta V_f \cdot (V_0 - V_p)(1 - K_{av})/V_f (V_t - V_0)$. Note that the effect of an error in the estimate of $V_f$ is dependent on $V_p$. To minimize this effect one should choose $V_p \approx V_0$.

(6) Errors in estimates of dextran concentrations. A change of all concentrations from a GPC analysis with a constant factor has no effect on the MWD estimates. Errors not proportional to the concentration level and errors correlated with the order of the effluent fractions may, however, be serious and three causes of this kind of errors will be discussed, *viz.* interaction between samples in AutoAnalyzer measurements, evaporation of effluent fractions, and errors or changes in the blank level.

Let $\hat{x}_j$ be the dextran concentration of the $j$th effluent fraction after interaction or evaporation. For the interaction, we assume the following model:

$$\hat{x}_j = (1 - I)x_j + I x_{j-1}$$

where $I$ is the coefficient of interaction. We have

$$\sum_{j=1}^{k} \hat{x}_j \approx \sum_{j=1}^{k} x_j$$

and

$$\hat{F}_h = \frac{\sum\limits_{j=1}^{h} \hat{x}_j}{\sum\limits_{j=1}^{k} \hat{x}_j}$$

$$\approx \frac{\sum\limits_{j=1}^{h} x_j + (1 - I) x_h}{\sum\limits_{j=1}^{k} x_j}$$

$$= F_h - I f_h$$

If the effect of interaction is neglected, *i.e.*, $I$ is assumed to be zero, the result is an error in $K_{av} \approx V_f \cdot I/(V_t - V_0)$. If $I$ is constant, the error in $K_{av}$ is about the same for all GPC analyses and the error is compensated in the calibration of the columns. A variation in $I$ results in a contribution to the variance within columns.

For the evaporation, we assume:

$$\hat{x}_j = \frac{1}{1 - \dfrac{k-j}{k-1} \cdot A} \cdot x_j$$

*i.e.*, a linear decrease from an evaporation of a proportion $A$ in the first effluent fraction to no evaporation in the last fraction (number $k$). If $A$ is small, we have

$$\hat{x}_j \approx (1 + \frac{k-j}{k-1} \cdot A) x_j .$$

and, if we further assume that $k$ is even,

$$x_j = x_{k+1-j}$$

and that $x_1 \leqslant x_2 \leqslant \ldots \ldots \leqslant x_{\frac{k}{2}}$, we obtain

$$\sum_{j=1}^{k} \hat{x}_j \approx (1 + \frac{A}{2}) \sum_{j=1}^{k} x_j$$

and, for $h \leqslant \dfrac{k}{2}$

$$\hat{F}_h \approx \frac{\sum_{j=1}^{h} f_j (1 + \frac{k-j}{k-1} \cdot A)}{1 + \frac{A}{2}}$$

$$= F_h + \frac{A \sum_{j=1}^{h} f_j \left(\frac{k-j}{k-1} - \frac{1}{2}\right)}{1 + \frac{A}{2}}$$

$$\leqslant F_h + \frac{h(k-h)}{2(k-1)} \cdot A \cdot f_h$$

As $k$ is large (50–60), we have

$$\frac{h}{k-1} \approx \frac{h}{k} \approx \frac{V_e - V_p}{V_t - V_p} < K_{av}$$

and

$$F_h \leqslant F_h + \frac{K_{av}(V_t - V_e)}{2V_f} \cdot A \cdot f_h$$

$$= F_h + \frac{(V_t - V_o) K_{av} (1 - K_{av})}{2V_f} \cdot A \cdot f_h$$

which, when the effect of evaporation is neglected, corresponds to a negative error in $K_{av} < K_{av} (1 - K_{av}) \cdot A/2$. For symmetry reasons, this inequality is valid also for $h > k/2$.

As for the interaction, the effect of a constant value of $A$ may be compensated in the calibration of the columns and variations in $A$ contribute to the variance within columns.

As regards errors and changes in the blank level, it is obvious that the effect is highly dependent on the shape of the MWD, especially the ends of the distribution. Consequently, the effects of these sources of variation have to be related to a specified MWD, but this is beyond the scope of this paper.

*Variances of measurement and production processes*

For the estimates of weight-average molecular weight from the production batches, we assume the following models:

Current method:

$$x_i = \mu_c + b_i + e_i$$

GPC method:

$$y_i = \mu_g + b_i + d_i$$

where

$x_i$ = weight-average molecular weight of batch number $i$ estimated with the current method;

$y_i$ = weight-average molecular weight of batch number $i$ estimated with the GPC method;

$\mu_c$ = mean of the production process according to the current method;

$\mu_g$ = mean of the production process according to the GPC method;

$b_i$ = deviation from the process mean;

$e_i$ = error of the current method;

$d_i$ = error of the GPC method.

Let

$\sigma_b^2 = \mathrm{Var}\{b_i\} =$ variance of the production process;

$\sigma_e^2 = \mathrm{Var}\{e_i\} =$ variance of the current measurement process;

$\sigma_d^2 = \mathrm{Var}\{d_i\} =$ variance of the GPC measurement process.

With these notations, we have

$$\mathrm{Var}\{x_i\} = \sigma_b^2 + \sigma_e^2$$
$$\mathrm{Var}\{y_i\} = \sigma_b^2 + \sigma_d^2$$
$$\mathrm{Var}\{x_i - y_i\} = \sigma_e^2 + \sigma_d^2$$

and from the observed variances of $x_i$, $y_i$ and $x_i - y_i$, $s_x^2$, $s_y^2$ and $s_{x-y}$, we can consequently calculate estimates of $\sigma_b$, $\sigma_e$ and $\sigma_d$, denoted by $s_b$, $s_e$ and $s_d$ (see Table V). The standard deviations obtained this way, $s_d$, for the GPC method, are greater than the corresponding estimates in the capability study, $s_T$. The estimates $s_d$ were calculated from current analysis of batches produced during an interval of about 4 years. As several improvements to the method have been introduced during this interval, it may be expected that $s_d > s_T$.

TABLE V

STANDARD DEVIATIONS OF MEASUREMENT AND PRODUCTION PROCESSES

| Standard deviation | Dextran 40 (No. of batches, 71) | Dextran 70 (No. of batches, 25) |
|---|---|---|
| $s_x$ | 1402 | 2252 |
| $s_y$ | 1273 | 2160 |
| $s_{x-y}$ | 1325 | 1406 |
| $s_b$ | 957 | 1970 |
| $s_e$ | 1025 | 1092 |
| $s_d$ | 840 | 960 |
| $s_T$ (from the capability study) | 427 | 607 |

CONCLUSION

Gel permeation chromatography with Sepharose 4B provides a useful technique for the MWD determination of clinical dextrans. Compared with conventional methods, it yields superior information in that not only average estimates of the molecular weight of a dextran sample and of a few prepared fractions of it are obtained, but also the total distribution of the molecular weight.

The method is extremely rapid and convenient. Once a column has been packed and calibrated, an analysis takes only a few hours of effective working time.

The precision of the method, expressed as the standard deviation of the $\bar{M}_w$ values, is about 400 and 600 for dextrans of $\bar{M}_w$ 40,000 and 70,000 respectively. Corresponding figures for the production process of these dextrans over a period of 4 years are about 1000 and 2000, respectively.

Efforts to effect further improvements in the GPC method should be directed towards a decrease in the variance between columns. Use of a few well investigated standard fractions, if available, for the molecular-weight calibration could decrease the variance between columns and the number of fractions necessary.

REFERENCES

1 K. A. Granath and B. E. Kvist, *J. Chromatogr.*, 28 (1967) 69.
2 *American Military Medical Purchase Description*, Stock No. 1-161-890, U.S. Navy Supply Depot (Code 1051), Philadelphia, Pa., 24 May 1951, pp. 3–6.
3 A. R. Shultz, A. L. Bridgman, E. M. Hadsell and C. R. McCullough, *J. Polym. Sci., Part A-2*, 10 (1972) 273.
4 M. R. Ambler, *J. Polym. Sci.*, 11 (1973) 191.
5 F. C. Frank, I. M. Ward and T. Williams, *J. Polym. Sci., Part A-2*, 6 (1968) 1357.
6 K. E. Almin, *Polym. Prepr., Amer. Chem. Soc., Div. Polym. Chem.*, 9 (1) (1968) 727.
7 S. T. Balke, A. E. Hamielec and B. P. Leclair, *Ind. Eng. Chem.*, 8 (1969) 54.
8 D. J. Wilde, *Optimum Seeking Methods*, Prentice-Hall, Englewood Cliffs, N. J., 1964, p. 151.
9 B. A. Whitehouse, *Macromolecules*, 4 (1971) 463.
10 H. S. Isbell, C. F. Snyder, N. B. Holt and M. R. Dryden, *J. Res. Natl. Bur. Stand.*, 50 (1953) 81.
11 H. Jenner, *Automation in Analytical Chemistry, Technicon Symp.*, 2 (1967) 203.
12 J. R. Burt, *Anal. Biochem.*, 9 (1964) 293.
13 H. O. Hartley, *Technometrics*, 3 (1961)
14 Ch. Eisenhart, *J. Res. Natl. Bur. Stand., C, Eng. Instrum.*, 67C (1963) 161.
15 L. A. Currie, J. J. Filliben and J. R. de Voe, *Anal. Chem.*, 44 (1972) 497R.
16 L. H. Tung, *J. Appl. Polym. Sci.*, 10 (1966) 375.
17 L. H. Tung, J. C. Moore and G. W. Knight, *J. Appl. Polym. Sci.*, 10 (1966) 1261.
18 L. H. Tung, *J. Appl. Polym. Sci.*, 10 (1966) 1271.
19 K. A. Hansson, private communication, 1968.