A New General Method for the Assessment of the Molecular-Weight Distribution of Polydisperse Preparations

ITS APPLICATION TO AN INTESTINAL EPITHELIAL GLYCOPROTEIN AND TWO DEXTRAN SAMPLES, AND COMPARISON WITH A MONODISPERSE GLYCOPROTEIN

By RICHARD A. GIBBONS, STEPHEN N. DIXON and DAVID H. POCKOCK
 Agricultural Research Council Institute for Research on Animal Diseases, Compton, Newbury, Berks., U.K.

(Received 14 May 1973)

A specimen of intestinal glycoprotein isolated from the pig and two samples of dextran, all of which are polydisperse (that is, the preparations may be regarded as consisting of a continuous distribution of molecular weights), have been examined in the ultracentrifuge under meniscus-depletion conditions at equilibrium. They are compared with each other and with a glycoprotein from *Cysticercus tenuicollis* cyst fluid which is almost monodisperse. The quantity $c^{-\frac{1}{3}}$ (c = concentration) is plotted against ξ (the reduced radius); this plot is linear when the molecular-weight distribution approximates to the 'most probable', i.e. when $M_n:M_w:M_z:M_{z+1}:\dots$ is as 1:2:3:4: etc. The use of this plot, and related procedures, to evaluate qualitatively and semi-quantitatively molecular-weight distribution functions where they can be realistically approximated to Schulz distributions is discussed. The theoretical basis is given in an Appendix.

In principle the radial macromolecular-concentration distribution at equilibrium in the ultracentrifuge contains all the information necessary to characterize the molecular-weight distribution function and methods have been described for doing this (Billick *et al.*, 1967; Provencher, 1967; Sundelöf, 1968; Scholte, 1968, 1969; Magar, 1970; Wiff & Gehatia, 1972; Williams, 1972). These techniques require a good deal of experimental and computational effort; the result, moreover, is often not unique. It would be valuable to have a simpler method for assessing approximately the broadness of a molecular-weight distribution of a biochemical preparation known to be homogeneous but polydisperse in the sense described by Gibbons (1963). This condition does not commonly arise with proteins, but polysaccharides, many glycoprotein and some nucleic acid preparations do present this obstacle to proper biophysical characterization. In these cases the extrapolation necessary to obtain the required parameters at the cell base is difficult, as the usual logarithmic plot may be quite curved and a function more nearly linear is desirable.

This investigation originated from the theoretical considerations given in the Appendix. With symbols as defined in Table 1, for the so-called 'most-probable' distribution for which the ratios of successive molecular-weight averages are in arithmetic progression, a plot of $c^{-\frac{1}{3}}$ against ξ will be linear provided that c , all its derivatives with respect to ξ and the two integrals $\int c d\xi$ and $\int (\int c d\xi) d\xi$ have values that are negligibly small at the meniscus compared with the value at the base of the cell. This in effect is the widely used meniscus-depletion condition (Yphantis, 1964)

where $c_b > 10^3 \times c_m$. However, the relevant function $c = 2q^2/(q\xi + k)^3$ gives, for the second integral above, the values $1/k$ and $1/(q+k)$ at the cell base and the meniscus respectively, and for the latter to be negligible compared with the former, q must be much greater than k . This imposes a very much more stringent restriction on the ratio of c_b/c_m than does the Yphantis (1964) condition, for if $c_b/c_m = 10^3$, the ratio of the two quantities $1/k$ and $1/(q+k)$ is 10. Nevertheless, even with values of this order of magnitude, linearity is very good for materials having molecular-weight averages close to the most-probable distribution, the only serious discrepancy being that of M_n/M_w , which becomes 0.54 in place of the ideal 0.5.

Two samples of dextran that are considered to have continuous molecular-weight distributions close to the most-probable one have been examined to explore the usefulness of this proposal. The results are compared with those obtained with an essentially monodisperse glycoprotein and with those found for an intestinal glycoprotein preparation in which the distribution is broader than the most-probable one.

Experimental

Materials

Glycoproteins. (i) *Tenuicollis* glycoprotein. This material was isolated from the cyst fluid of the tapeworm *Cysticercus tenuicollis* in the goat (Dixon *et al.*, 1973). (ii) Porcine intestinal glycoprotein. This material was isolated by squeezing the contents of 1 m lengths of the intestinal tracts (duodenum and

Table 1. *Definition of symbols used*

c = concentration	R = gas constant
r = radius of rotation	T = temperature ($^{\circ}\text{K}$)
$\bar{v}$ = partial specific volume of macromolecular species	q, k = empirical constants as defined in the Appendix [eqn. (9)]
ρ = solvent density	p, h = empirical constants as defined in the Appendix [eqn. (10)]
ω = angular velocity	$\alpha = 1/(h+2)$

M = molecular weight, the subscripts $n, w, z, (z+n)$ indicating the number, weight, z , or $(z+n)$ average respectively, n being any positive integer. The subscripts m, b , and o appended to the variables c, r and ξ , and to derivatives of these quantities, imply the value of the subscripted variable at the meniscus, at the base of the cell and before the commencement of the equilibrium experiment respectively. Two derived quantities are also used: the experimental constant $\lambda = [(1-\bar{v}\rho)(r_b^2 - r_m^2)\omega^2]/2RT$ and the reduced radius of rotation $\xi = (r_b^2 - r^2)/(r_b^2 - r_m^2)$.

Table 2. *Ultracentrifugal analysis of glycoproteins and dextrans by the method of Chervenka (1970)*

The initial concentration was that introduced into the ultracentrifuge cell before formation of the boundary. $M_w(\text{app.})$ was calculated as described in the text. All runs were carried out at 293°K .

	Initial concn. (mg/ml)	Mean speed of centrifugation (rev./min)	Time (h)	$\bar{v}$	$M_w(\text{app.})$
<i>C. tenuicollis</i> glycoprotein	0.25	18 300	43	0.74	96 800
Pig intestinal glycoprotein	3.0	5500	100	0.66	3.31×10^5
	0.6	5500	100	0.66	5.34×10^5
	0.3	5500	100	0.66	8.54×10^5
Dextran 1	0.5	16 600	76	0.613	35 200
Dextran 2	0.25	18 300	100	0.613	66 900

jejunum) of pigs into saline at 4°C , then removing the bulk of the food particles by filtration through muslin. It was purified by digestion with crystalline pepsin (0.05 mg/ml) at pH 3.5 and 37°C for 48 h, followed by extraction with 90% (w/v) phenol as described by Morgan & King (1943). The aqueous extract from the phenol treatment was dialysed against ten changes of water at 4°C and freeze-dried. The solid was then fractionated by gel filtration on a column (2.5 cm $\times$ 100 cm) of Sephadex G-200 in 0.05 M-sodium acetate buffer, pH 6.5, containing 0.02% (w/v) NaN_3 . This yielded two major components, an excluded and an included fraction; both fractions were subsequently recovered by dialysis against ten changes of water at 4°C followed by freeze-drying. The higher-molecular-weight material (excluded fraction) was used in the work reported here. (iii) Dextran. Two samples were used; dextran 1 was a partially acid-hydrolysed unfractionated sample and dextran 2 a 'clinical' dextran with a stated molecular weight of 60 000–90 000.

Solvents. (i) 0.01 M-Sodium barbiturate buffer, pH 7.0, containing 0.1 M-NaCl; (ii) 0.15 M-NaCl.

Methods

Preparations of samples. Solutions of *C. tenuicollis* glycoprotein, 0.02% in solvent (i), of porcine intestinal glycoprotein, 0.3% in solvent (i), and of the two dextrans, 0.05 and 0.025% respectively in solvent (ii), were prepared. Each solution was dialysed against 10 vol. of the appropriate solvent for 24 h at 4°C and subsequently centrifuged at 20 000 rev./min (r_{av} 2.5 cm) for 16 h to eliminate any gelatinous aggregates or undissolved material. Dilutions of 1:5 and 1:10 of the porcine intestinal glycoprotein solutions were also made.

Analytical ultracentrifugal examinations. All six solutions were examined in the synthetic-boundary cell of the Spinco model E ultracentrifuge as described by Chervenka (1970) by using 0.15 ml of solution plus 0.02 ml of silicone oil and 0.4 ml of solvent plus 0.04 ml of silicone oil. Conditions of each run are given in Table 2. Equilibrium was checked before termination of each run by the comparison of two plates exposed at intervals of 8 or 12 h. The cell base was located after running speed was attained by making an exposure using schlieren optics with the

phase-plate angle at 90° . Exposures using light of wavelength 541.1 nm were made immediately after running speed was reached as well as at the end of the run. Exposures with white light were also made at these times and the comparison of the position of the zero-order fringe at the meniscus gave an additional check that this identifiable fringe had

not moved perceptibly at the meniscus, i.e. that the meniscus concentration was actually zero.

Solvent densities were measured in a 50 ml pyknometer. Partial specific volumes were calculated from the sugar and amino acid composition; values for the appropriate residues were taken from Cohn & Edsall (1943) and Gibbons (1972). Plates were

Fig. 1. Equilibrium distributions in the ultracentrifuge under meniscus-depletion conditions

The method of Chervenka (1970) was used; for conditions see Table 2. Concentrations, c , are in arbitrary units, r is in cm and $\xi = (r_b^2 - r^2)/(r_b^2 - r_m^2)$. ■, $\ln c$ against r^2 ; ▲, c^{-1} against ξ . The broken line indicates linearity where the observed plot is not linear. (a) *C. tenacollis* glycoprotein, 0.02%; (b) dextran 1, 0.05%; (c) dextran 2, 0.025%; (d), (e) and (f) pig intestinal glycoprotein examined at 0.3%, 0.06% and 0.03% respectively.

measured with a tool-maker's micro-comparator (Precision Grinding Instrument Co., Mitcham Junction, Surrey, U.K.) and aligned by using the air fringes. Measurements along the r axis were made at intervals of 0.02 cm. The data were processed by using an Olivetti P 101 calculator. The values of M_w (app.) (not corrected for non-ideal behaviour, Table 2) were calculated as $c_b/\lambda c_0$ [see Appendix, eqn. (2)]; c_0 was calculated as $\int_0^1 c d\xi$ both from the initial and final green-light exposures, and c_b was calculated by extrapolation of either the plot of $\ln c$ against r^2 or of $c^{-\frac{1}{2}}$ against ξ , whichever was most linear.

Results and Discussion

The 'most-probable' distribution of molecular weights is a Schulz distribution with $h=1$ (Schulz, 1944) and is of peculiar significance, as it is the molecular-weight distribution resulting from the random break-up of an indefinitely large molecule (Pathria & Nanda, 1959) or from macromolecule synthesis by random chain-lengthening (Tanford, 1961). For material having a continuous molecular-weight distribution, with one maximum and two points of inflexion only, it is often valuable to know whether the distribution is broader or narrower than the most-probable one. Under meniscus-depletion conditions, the amount of and direction of curvature in a plot of $c^{-\frac{1}{2}}$ against ξ will immediately give information about this point with the minimum of extra work. The two dextran preparations, although giving markedly curved plots of $\ln c$ against c/r^2 , gave plots of $c^{-\frac{1}{2}}$ against ξ that were highly linear (Figs. 1b and 1c) and the correlation coefficients for samples 1 and 2 were 0.9996 and 0.9994 respectively. Extrapolation to $\xi=0$ can be made with considerable confidence and from the measured values of q and k any molecular-weight average may be deduced. On the other hand, the epithelial glycoprotein examined gave marked curvature in the plots of $c^{-\frac{1}{2}}$ versus ξ (Figs. 1e and 1f) except at the highest concentration examined (Fig. 1d); this indicates that these materials have a wider distribution of molecular weights than that expected from random synthesis or breakdown, a conclusion already reached by Creeth & Knight (1968) for a similar, human, epithelial glycoprotein. Further, the curve cannot be made satisfactorily linear by selecting a different exponent between $-\frac{1}{2}$ and $-\frac{1}{3}$ (Fig. 2), which implies that a Schulz distribution is inadequate as a representation of the molecular-weight distribution of this glycoprotein. Creeth & Denborough (1970) have shown that their sample of epithelial glycoprotein is polydisperse with respect to partial specific volume as well, a conclusion that most probably applies to the similar glycoprotein studied here. Thus the distribution function is being examined of not M but $M(1-\bar{v}\rho)$; until the distribution in $\bar{v}$ can be further elucidated this point should

Fig. 2. Curves relating $c^{-\alpha}$ to ξ for different values of the exponent α for pig intestinal glycoprotein (0.03%)

Reference to the Appendix, second section, shows that for Schulz distributions wider than the most-probable distribution this will become linear for some value of α between $\frac{1}{2}$ and $\frac{1}{3}$. $\square$, $\alpha=0.03$; $\circ$, $\alpha=0.35$; $\triangle$, $\alpha=0.4$; $\bullet$, $\alpha=0.45$; $\blacksquare$, $\alpha=0.5$.

always be remembered. In addition the effect of non-ideal behaviour is evident for the epithelial glycoprotein; this reveals itself as an apparent narrowing of the molecular-weight distribution as concentration is increased (Figs. 1d, 1e and 1f). At the lowest concentrations, where ideal conditions are approached, a distribution wider than the most-probable one is apparent. Concentration effects must therefore be taken into account in the assessment of molecular-weight distributions, just as for the assessment of monodispersity with a plot $\ln c$ against r^2 . A condition of 'pseudo-most-probable distribution', entirely analogous to the spurious appearance of monodispersity in plots of $\ln c/r^2$, can arise from non-ideal behaviour (Fig. 1d). In view of the relatively low molecular weights of the two dextran samples and of the low concentrations at which they were examined, the disturbances due to the second virial coefficient should be small with this polysaccharide (Berry & Casassa, 1970; Senti *et al.*, 1955; Jeans *et al.*, 1954). The *C. tenuicollis* glycoprotein is included (Fig. 1a) in the data for comparative purposes; it appears to be a globular protein that is very nearly monodisperse.

The assumption that $\int (\int_0^\infty c d\xi) d\xi$ has a negligible value at $\xi=1$ is not justified at the angular velocities used here; nevertheless it appears to cause little disturbance to the linearity of the plots where the materials have a molecular-weight distribution close to the most-probable one. It may be calculated that for dextran 1 successive molecular-weight ratios are 0.566, 0.678, 0.75, 0.8, etc. and for dextran 2 are 0.54,

0.671, 0.75, 0.8, etc. if the graphs are assumed to be strictly linear. Where narrower distributions are encountered, the exponent α (see the Appendix) becomes less than $\frac{1}{2}$ and such distributions may more readily be characterized by linearity of the appropriate plot, because the smaller the exponent the less serious is the assumption that $\int (\int_{\xi}^{\infty} c d\xi) d\xi$ is 0 at $\xi = 1$. For example, with a Schulz distribution having $h = 2$, $c^{-\frac{1}{2}}$ is linear with ξ to a high degree of accuracy if $c_m < 10^{-3} \times c_b$.

References

- Berry, G. C. & Casassa, E. F. (1970) *Macromol. Rev.* **4**, 1–66
- Billick, I. H., Schulz, M. & Weiss, G. H. (1967) *J. Res. Nat. Bur. Stand. Sect. A* **71**, 13–17
- Chervenka, C. H. (1970) *Anal. Biochem.* **34**, 24–29
- Cohn, E. J. & Edsall, J. T. (1943) *Proteins, Amino Acids and Peptides as Dipolar Ions*, pp. 370–381, Reinhold, New York
- Creeth, J. M. & Denborough, M. A. (1970) *Biochem. J.* **117**, 879–891
- Creeth, J. M. & Knight, C. G. (1968) *Solution Properties of Natural Polymers*, pp. 303–313, The Chemical Society, London
- Dixon, S. N., Gibbons, R. A., Parker, J. & Sellwood, R. (1973) *Int. J. Parasitol.* **3**, 419–424
- Gibbons, R. A. (1963) *Nature (London)* **200**, 665–666
- Gibbons, R. A. (1972) in *Glycoproteins, their Structure and Function* (Gottschalk, A., ed.), 2nd edn., vol. A, p. 78, Elsevier, Amsterdam
- Jeans, A., Haynes, W. C., Wilham, C. A., Rankin, J. C., Melvin, E. H., Austin, M. J., Cluskey, J. E., Fischer, H. M., Tsuchiya, H. M. & Rest, C. E. (1954) *J. Amer. Chem. Soc.* **76**, 5041–5052
- Magar, M. E. (1970) *J. Theor. Biol.* **27**, 127–169
- Morgan, W. T. J. & King, H. K. (1943) *Biochem. J.* **37**, 640–651
- Pathria, R. K. & Nanda, V. S. (1959) *J. Chem. Phys.* **30**, 31–34
- Provencher, S. W. (1967) *J. Chem. Phys.* **46**, 3229–3236
- Scholte, T. G. (1968) *J. Polym. Sci. Part A-2* **6**, 111–127
- Scholte, T. G. (1969) *Ann. N.Y. Acad. Sci.* **164**, 156–171
- Schulz, G. V. (1944) *Z. Phys. Chem. (Leipzig) A* **193**, 168–172
- Senti, F. R., Hellman, N. N., Ludwig, N. H., Babcock, G. E., Tobin, R., Glass, C. A. & Lamberts, B. L. (1955) *J. Polym. Sci.* **8**, 527–546
- Sundelöf, L. O. (1968) *Ark. Kemi* **29**, 279–300
- Tanford, C. (1961) *Physical Chemistry of Macromolecules*, pp. 139–147, Wiley, New York
- Wiff, D. R. & Gehatia, M. (1972) *J. Macromol. Sci. Phys.* **B6**, 287–307
- Williams, J. W. (1972) *Ultracentrifugation of Macromolecules*, pp. 3–35, Academic Press, New York
- Yphantis, D. A. (1964) *Biochemistry* **3**, 297–317

APPENDIX

Derivation of an Approximate Function Relating c and ξ for Schulz Molecular-Weight Distributions in the Meniscus-Depletion Condition in the Equilibrium Ultracentrifuge

By RICHARD A. GIBBONS

Agricultural Research Council Institute for Research on Animal Diseases, Compton, Newbury, Berks., U.K.

(Received 14 May 1973)

Most-Probable Case

With the notation from Table 1 of the main paper (Gibbons *et al.*, 1973), successive molecular-weight averages may be written (Fujita, 1962):

$$M_n = \frac{c_o \lambda}{\left[\int_0^1 \left(\int_{\xi}^{\infty} c d\xi \right) d\xi \right]} \quad (1)$$

$$M_w = \frac{1}{\lambda c_o} (c_b - c_m) \quad (2)$$

$$M_z = \frac{1}{\lambda^2 c_o M_w} \left[\left(\frac{dc}{d\xi} \right)_m - \left(\frac{dc}{d\xi} \right)_b \right] \quad (3)$$

and generally:

$$M_{(z+n)} = \frac{1}{\lambda^{(n+2)} c_o M_w M_z \dots M_{(z+n-1)}} \times \left[\left(\frac{d^{(n+1)}c}{d\xi^{(n+1)}} \right)_m - \left(\frac{d^{(n+1)}c}{d\xi^{(n+1)}} \right)_b \right] \quad (4)$$

if n is an even integer; if it is odd then the term within the square brackets becomes

$$\left[\left(\frac{d^{(n+1)}c}{d\xi^{(n+1)}} \right)_b - \left(\frac{d^{(n+1)}c}{d\xi^{(n+1)}} \right)_m \right]$$

The ratios of successive molecular-weight averages simplify to

$$\frac{M_n}{M_w} = \frac{c_o^2}{\left[\int_0^1 \left(\int_{\xi}^{\infty} c d\xi \right) d\xi \right] (c_b - c_m)} \quad (5)$$