Journal of Biochemical and Biophysical Methods, 7 (1982) 25–34
Elsevier Biomedical Press

Some observations on a new type of point average molecular weight

J. Michael Creeth [1] and Stephen E. Harding [2]

[1] *Department of Medicine, University of Bristol, Bristol Royal Infirmary, Bristol BS2 8HW, and*
[2] *Department of Biochemistry, University of Bristol, Bristol BS8 1TD, U.K.*

(Received 14 May 1982)
(Accepted 10 July 1982)

Summary

A new type of (reduced) point average molecular weight, A^*, is described. Several interesting properties are developed: (i) A^* (cell base) = reduced weight average molecular weight over the whole cell, A°_w; (ii) A^* (meniscus) = A_w (meniscus); (iii) A^* (zero concentration) = reduced number average molecular weight, A_n (meniscus). In addition, its usefulness in extracting the meniscus concentration, J(a), and in examining heterogeneous systems such as mucus glycoproteins, are discussed. The evaluation and application of A^* requires only simple computational facilities, without the use for large-scale multiple data acquisition and recycling techniques.

Key words: point average molecular weight; meniscus concentration; heterogeneous system; mucus glycoprotein.

Introduction

In an earlier report [1] a new type of point average molecular weight, $M^*(r)$, for analysing macromolecular distributions in sedimentation equilibrium experiments was introduced. It was stated that M^*, or more conveniently the reduced form A^* (see below), can be determined relatively simply provided that a reliable estimate for the meniscus concentration, $J(a)$, could be found, and a relation was derived equating $A^*(r)$ at the cell base (b) with the reduced weight average molecular weight over the whole cell, A°_w. We present here simulations which show that the method of extrapolating $A^*(r)$ to the cell base in order to find A°_w represents an improvement on methods used hitherto. Two new relations are derived. The first equates $A^*(r)$ at the meniscus ($A^*(a)$) to the reduced weight average molecular weight at the meniscus ($A_w(a)$) and simulations are given which confirm this. Secondly we equate the hypothetical quantity $A^*(r)$ at zero concentration ($A^*(J=0)$) to the reduced number average molecular weight at the meniscus, $A_n(a)$. Although the requirements on the precision of experimental data are still severe, extrapolation of $A^*(r)$ as a

0165-022X/82/0000–0000/$02.75

procedure for obtaining $A_n(a)$ is shown to represent an improvement on other methods. It follows necessarily that the approximation $A^*(a) \approx A_n(a)$ is only valid for single-solute or pseudo-single-solute systems. Nonetheless, the properties of the $A^*(r)$ function are seen to be advantageous for analysing systems that are both polydisperse and non-ideal, characteristics pre-eminent amongst polysaccharides and mucus glycoproteins.

The $A^*(r)$ function

In the mathematical analysis of ultracentrifugation, molecular weights are often 'reduced' to a standard form, facilitating comparison irrespective of the conditions of each particular experiment. Here we follow the convention of Rinde [2] by defining the relation between the reduced molecular weight, A_i, and its true value, M_i, by the equation

$$A_i = \frac{M_i(1 - \bar{v}_i\rho)\omega^2}{2RT} \tag{1}$$

where the symbols have their usual significance.

If it is assumed (as is often justifiable) that the partial specific volumes, $\bar{v}_i$, of the various solute components are all equal, then direct interconversion between A and M is possible. It must be emphasised, however, that the quantity A_i alone is determinable from the ultracentrifuge.

Denoting absolute concentrations (in terms of fringes) as J and concentrations relative to that at the meniscus as j (thus $J(r) = J(a) + j(r)$), the fundamental differential equation of sedimentation equilibrium may be manipulated to give

$$\frac{J(r)}{A_n(r)} - \frac{J(a)}{A_n(a)} = 2\int_a^r rJ\mathrm{d}r = J(a)(r^2 - a^2) + 2\int_a^r rj\mathrm{d}r \tag{2}$$

where A_n is the reduced number average molecular weight. We define the new type of reduced average molecular weight A^* by the expression

$$\frac{j(r)}{A^*(r)} \equiv \frac{J(r)}{A_n(r)} - \frac{J(a)}{A_n(a)} \tag{3}$$

Thus the formal definition of $A^*(r)$ is

$$A^*(r) \equiv \frac{A_n(r) \cdot A_n(a)}{[A_n(a) \cdot J(r) - A_n(r)J(a)]} \cdot j(r) \tag{4}$$

and hence Eqn. 2 becomes

$$\frac{j(r)}{A^*(r)} = J(a)(r^2 - a^2) + 2\int_a^r rj\mathrm{d}r \tag{5}$$

$A^*(r)$ is thus easily determinable once $J(a)$ is found, although it is clear that experimental precision must diminish at low values of j.

Extraction of $J(a)$

For associating or very polydisperse systems such as mucus glycoproteins, in which the $A^*(r)$ function is likely to be of most use, a meniscus depletion type of experiment where $J(a) \simeq 0$ is at best unsuitable [3], or more commonly, impossible to carry out. This is because at speeds where $J(a)$ can be made to approach zero, a significant proportion of the macromolecular component is concentrated at the cell base beyond the resolving power of the optical system [4,5]. Accordingly, to ensure that optical registration extends to the cell base, a speed must be chosen at which $J(a)$ remains finite and significant: a method for the general determination of $J(a)$ must therefore be devised. Although a variety of methods exist, they depend on factors such as conservation of cell contents, auxiliary initial concentration measurements or the use of boundary-forming cells, which may be unreliable, inconvenient or unavailable: simple mathematical manipulations (summarised by Creeth and Pain [6]) are strictly applicable only to single solutes. A method that depends only on the data necessarily recorded in the experiment itself is required: such an approach is possible in terms of the A^* function, although the limitations will be apparent. Eqn. 5 has the limiting form

$$\lim_{(r \to a)} \left(\frac{j}{r^2 - a^2} \right) = A^*(a) J(a) \qquad (6)$$

A graph of $j/(r^2 - a^2)$ vs. $\int_a^r r j \mathrm{d}r/(r^2 - a^2)$ therefore has a *limiting* slope of $2A^*(a)$ and an intercept of $A^*(a)J(a)$: hence both $J(a)$ and $A^*(a)$ are in principle determinable from the slope and intercept [7,8]. However, the slope at $r > a$ is greater than the value given (except for single solute systems) and the equation gives no guidance as to the functional dependence. Nonetheless the advent of on-line fringe data processing [9] and advanced statistical packages (NAG) [10] may enable reasonably accurate values of $J(a)$ and $A^*(a)$ to be obtained. However, at present the most common method of data collection is still that of manual plate reading using a microcomparator. Simulations of model mixed-solute systems shows that over the upper one-third of the cell at an appropriate speed a linear extrapolation gives $J(a)$ to better than 15%. As speeds are such that $J(a)$ is approximately 0.5 ($A_i \approx 0.6$ cm^{-2}), this is quite adequate for most purposes. For mucus glycoproteins, a parabolic fit is often better. Teller et al. [11] list other manipulations of the integral which similarly yield $J(a)$.

Determination of the concentration of the cell base

Eqn. 5, written for the cell base, b, becomes

$$\frac{j(b)}{A^*(b)} = J(a)(b^2 - a^2) + 2\int_a^b rj\mathrm{d}r \tag{7}$$

The expression for the (reduced) weight average molecular weight over the whole cell, denoted A°_w, is

$$A^\circ_w = \frac{J(b) - J(a)}{J_0(b^2 - a^2)} = \frac{j(b)}{J_0(b^2 - a^2)} \tag{8}$$

Since $J(b) - J(a)$ is simply $j(b)$, and the usual conservation of mass equation is $J_0(b^2 - a^2)/2 = \int_a^b rJ\mathrm{d}r$, it follows simply that

$$A^*(b) = A^\circ_w \tag{9}$$

It was noted earlier [1] that extrapolation of $A^*(r)$ to b would probably be a more satisfactory way of finding A°_w than the hitherto used extrapolation of ln J and the application of Eqn. 8 [12]. The difficulty in applying the latter method is a reflection of the fact that no simple function is known that accurately reproduces the behaviour of the actual curve (cf. Donnelly [13]).

Because $A^*(r)$ varies relatively little with r, it provides a much less sensitive extrapolation function than j, J or ln J. Since measurement of concentration cannot be made at the base itself, some form of extrapolation is essential, and a least-squares parabolic fit of $A^*(r)$ vs. $(b^2 - r^2)$, extrapolated to zero, appears to be most suitable. This method has been adopted for several examples in Table 1. It is seen to bear out the contention that M°_w is found significantly more accurately than by extrapolation of ln J. The rather extreme case III, where the concentration rises very sharply near the base of the cell (cf. the values quoted in Ref. 1) is seen to be beyond the powers of either extrapolation procedure.

It is also worth noting that it is then unnecessary to determine J_0 (the cell loading concentration) as a separate procedure: this procedure is not only time consuming but gives a value which may not reflect entirely the total concentration of species

TABLE 1

COMPARISON OF M°_w VALUES OBTAINED VIA (ln J or M^*) EXTRAPOLATIONS

	I	II	III	IV	V
ln J	$6.59 \cdot 10^5$	$4.73 \cdot 10^6$	$2.45 \cdot 10^6$	$2.15 \cdot 10^6$	$4.99 \cdot 10^5$
M^*	$6.64 \cdot 10^5$	$4.87 \cdot 10^6$	$2.77 \cdot 10^6$	$2.63 \cdot 10^6$	$5.00 \cdot 10^5$
Theoret.	$6.67 \cdot 10^5$	$4.84 \cdot 10^6$	$3.02 \cdot 10^6$	$3.02 \cdot 10^6$	$5.00 \cdot 10^5$
System I	2 comp. mixt., $M_2 = 3M_1$; $c_2^\circ = 0.167\ c_1^\circ$ ($M_1 = 5 \cdot 10^5$)				
System II	Very polydisperse (log-normal: $\sigma/\log M_w = 0.044$; $M = 1 - 15 \cdot 10^6$)				
System III	Isodesmic assn. $c = 1$, $k = 2$ ($M_1 = 5 \cdot 10^5$)				
System IV	As III but with fringe data with $\pm 2\mu$ standard error				
System V	Single solute. $M = 5 \cdot 10^5$, fringe data with $\pm 2\mu$ standard error				

resolved by the optical system. The integration $\int J\mathrm{d}r^2$ however is free from this objection and has already been performed, of course, during the data processing.

Determination of the weight average molecular weight at the meniscus

Another useful property of the A^* function is apparent if we consider its behaviour as $r \to a$. From Eqn. 4, the relation tends to 0/0, but from l'Hôpital's rule, differentiation gives the limiting value of A^* as

$$\lim A^*_{j\to 0} = \frac{[A_n(a)]^2}{A_n(a) - J(a)\mathrm{d}A_n(r\to a)/\mathrm{d}j} \tag{10}$$

Now, from Eqn. 27 of Wales [14]:

$$M_w(r) = M_n(r) + M_w(r) \cdot \frac{\mathrm{d}\ln M_n(r)}{\mathrm{d}\ln J}$$

Rearranging and using the simple identities

$x(\mathrm{d}\ln x/\mathrm{d}\ln y) = y\mathrm{d}x/\mathrm{d}y$ and $\mathrm{d}j = \mathrm{d}J$,

$$M_n(r) = \frac{M_n^2(r)}{M_w(r)} + \frac{J\mathrm{d}M_n(r)}{\mathrm{d}j}$$

or, in terms of the corresponding reduced quantities,

$$A_n(r) = \frac{A_n^2(r)}{A_w(r)} + \frac{J\mathrm{d}A_n(r)}{\mathrm{d}j} \tag{11}$$

and hence

$$\frac{\mathrm{d}A_n(r)}{\mathrm{d}j} = \frac{A_n(r) - \dfrac{A_n^2(r)}{A_w(r)}}{J} \tag{12}$$

Thus

$$\left(\frac{\mathrm{d}A_n(r)}{\mathrm{d}j}\right)_{r\to a} = \frac{A_n(a) - \left(A_n^2(a)/A_w(a)\right)}{J(a)} \tag{13}$$

Substituting into Eqn. 10 we find

$$A^*(a) = A_w(a) \tag{14}$$

It is seen therefore that extrapolation of $M^*(r)$ to the meniscus provides a method of estimating $M_w(a)$. The method outlined earlier for determining an estimate for $J(a)$ also therefore yields a (less accurate) estimate for $M_w(a)$. Eqn. 14 has been verified by computer simulation for single solute, monomer–trimer and isodesmically associating systems. The value of $M_w(a)$ can be used to obtain an estimate for the z average molecular weight for the cell, M_z° [6].

Determination of the number average molecular weight at the meniscus, $M_n(a)$

Teller [15] has described two independent methods for obtaining $M_n(a)$, the number average molecular weight at the meniscus. The simplest is to approximate $M_n(a)$ to the weight average molecular weight at the meniscus, $M_w(a)$ [16]. Our simulations have shown, however, that this is only applicable to systems of limited heterogeneity and may well be very poor for ideal isodesmically associating systems (the discrepancy can be of the order of 25%). The assumption may, however, be applicable to certain non-ideal isodesmically associating systems where the effects of non-ideality and association roughly counteract to give a system that is a pseudo-ideal single solute. Such a system has apparently been found for a bronchial glycoprotein from a cystic fibrotic [17].

Alternatively, Teller [15,18] has shown by finite difference calculus that a graph of $\int_{a^2}^{r^2} J\mathrm{d}r$ vs. J can be extrapolated to $-J(a)/A_n(a)$ at $J=0$ (although J is nowhere zero in an actual 'low speed' experiment and can never be exactly zero in a 'high speed' experiment).

It can be simply shown that

$$\int_{a^2}^{r^2} J\mathrm{d}r^2 = 2\int_a^r rJ\mathrm{d}r$$

and hence there follows from [5]:

$$\int_{a^2}^{r^2} J\mathrm{d}r^2 = \frac{j}{A^*(r)} \tag{15}$$

Equivalently, from Eqn. 1, a plot of $j/M^*(r)$ vs. J should extrapolate to $-J(a)/M_n(a)$ at $J=0$. It is clearly evident from Fig. 1, however, that an extrapolation of $j/M^*(r)$ must provide an estimate of $M_n(a)$ which is very susceptible to error.

However, rearranging Eqn. 15,

$$A^*(r) = \frac{j}{\int_{a^2}^{r^2} J\mathrm{d}r^2}$$

and hence it is also evident that, since $j(J=0) = -J(\mathrm{a})$ and

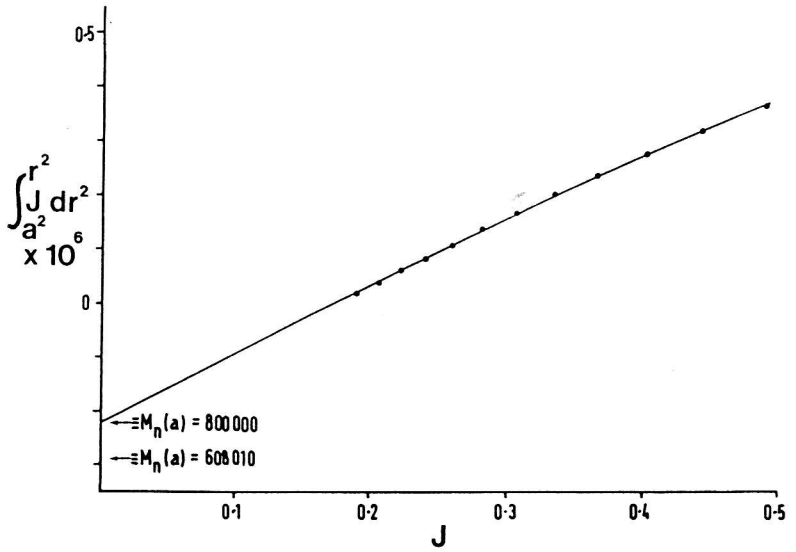

Fig. 1. Use of Teller's [11] method for estimating the number average molecular weight at the meniscus, $M_n(a)$ (see text for details). The Rayleigh fringe data from which the ordinate values were obtained were for a synthetic isodesmically associating system (system III of Table 1), precise to 0.1μ. The extrapolation is seen to be insensitive to $M_n(a)$. True $M_n(a)=608010$.

$$\int_{a^2}^{r^2} J\mathrm{d}r^2 \text{ at } J=0 \rightarrow J(a)/A_n(a)$$

we find $A^*\ (J=0)=A_n(a)$, or equivalently $M^*(J=0)=M_n(a)$.

$M^*(r)$ is plotted as a function of J for an isodesmic association in Fig. 2. It is

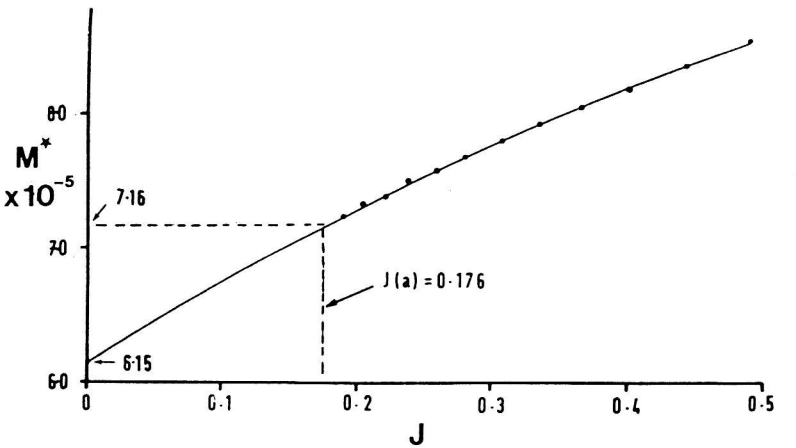

Fig. 2. Use of the M^* function for estimating the number and weight average molecular weights at the meniscus. The Rayleigh fringe data are the same as that used for Fig. 1. True $M_n(a)=608010$. True $M_w(a)=716969$.

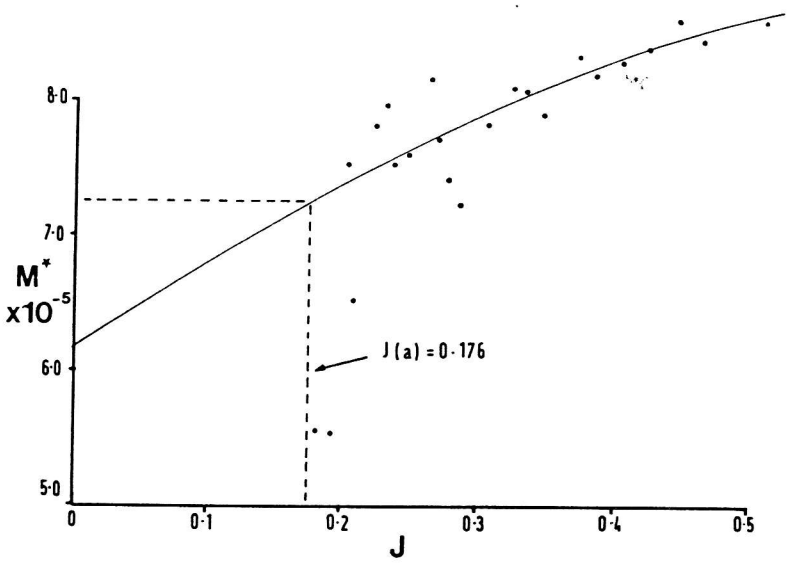

Fig. 3. As Fig. 2 but the fringe data used had synthetic normal pseudorandom error of 1μ standard deviation.

immediately apparent that the extrapolation of $M^*(r)$ to $J = 0$ for obtaining $M_n(a)$ may have some advantage over the method of Teller. The figure clearly shows that $M^*(J = 0) = M_n(a)$ and also $M^*(a) = M_w(a)$. In Fig. 3 we present an extrapolation from the same simulated data but with synthetic error of 1 μ standard deviation on the fringe data points (about the precision of an automatic plate reader) [19,20].

The extrapolation, like that of Teller's is, however, strongly dependent on an

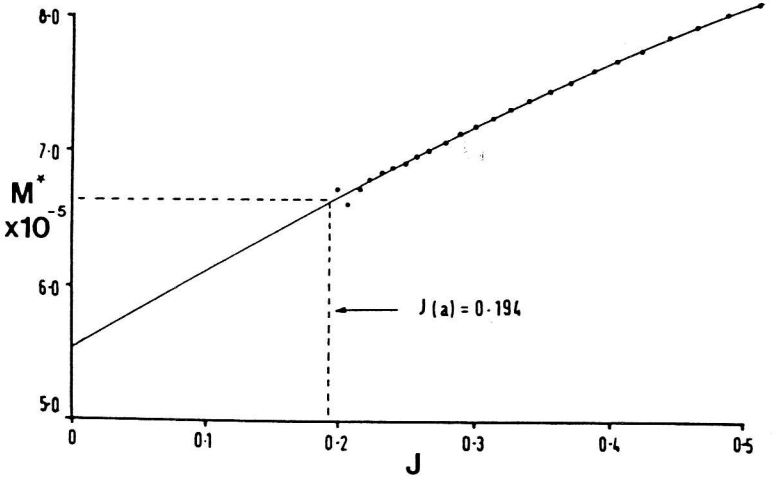

Fig. 4. As Fig. 2 but with an assumed error of +10% in the estimate for the meniscus concentration $J(a)$.

adequate estimate for the meniscus concentration $J(a)$. Fig. 4 reveals that errors in $J(a)$ of ca. 10% ($\equiv 6\,\mu$) can lead to a corresponding error in the estimate for $M_n(a)$ of ca. 12%. Precise determinations of number average molecular weights for associating systems will depend therefore on accurate determinations of $J(a)$ [11].

Discussion

In an ideal world one would like to have the facilities for reading clearly defined-laser generated fringe patterns on line (or off line via magnetic tape) into a large computer [9,19,20]. Multiple data acquisition (M.D.A.), appropriate averaging and statistically weighted smoothing are the only real means of isolating signal from noise. M.D.A. is virtually impossible with manual methods and simple 'desktop' computing facilities, and one purpose of this communication is to attempt to improve the limits on the kind of systems which may usefully be handled by the simplest kind of measuring procedures and data processing. One can reasonably say that the A^* function does represent an improvement on hitherto used methods for the extraction of M_w°, $M_w(a)$, M_z°, $M_n(a)$ and hence point number average molecular weight, $M_n(r)$, data.

In analysing systems that are polydisperse, non-ideal thermodynamically or associating, various diagnostic relationships and plots involving $M_n(r)$ and $M_w(r)$ prove most useful [15,21]. We are currently investigating the applicability of these methods to mucus glycoproteins: these substances are apparently polydisperse, non-ideal *and* self-associating, and hence the properties of the A^* function described in this study will prove invaluable. Ultimately the extraction of accurate A^* values will depend on improved procedures for obtaining $J(a)$, and we are currently investigating iterative methods based on the constraints that $A^*(a) = A_w(a)$ and $A^*(r) < A_w(r)$.

Simplified description of the method and its advantages

In this study some interesting and important properties of a new type of point average molecular weight have been described. Relations have been derived relating this new point average with (i) the cell weight average molecular weight, (ii) the weight average molecular weight at the meniscus and (iii) the number average molecular weight at the meniscus. Extensive simulations have shown that it provides improved estimates for these parameters and hence also the number and z cell average molecular weights. Its use will prove helpful in characterising heterogeneous and non-ideal systems without necessarily large computational facilities.

Acknowledgement

The authors are grateful to Dr. P.A. Charlwood of the N.I.M.R. Laboratories, Mill Hill, for helpful advice. S.E.H. is grateful to the Lister Institute of Preventive Medicine for a Postdoctoral Award.

References

1 Creeth, J.M. (1980) Biochem. Soc. Trans. 8, 520–521
2 Rinde, H. (1928) Dissertation, University of Uppsala, Sweden
3 Yphantis, D.A. (1964) Biochemistry 3, 297–317
4 Creeth, J.M. and Knight, C.G. (1967) Biochem. J. 105, 1135–1145
5 Richards, E.G., Teller, D.C. and Schachman, H.K. (1968) Biochemistry 7, 1054–1076
6 Creeth, J.M. and Pain, R.H. (1967) Progr. Biophys. Mol. Biol. 17, 217–287
7 Creeth, J.M. and Holt, J.C. (1971) Unpublished work
8 Gratzer, W.B., Creeth, J.M. and Beaven, G.H. (1972) Eur. J. Biochem. 31, 505–509
9 Laue, T.M. and Yphantis, D.A. (1979) Biophys. J. 25, 164a
10 'NAG' FORTRAN Library Manual (1978) Numerical Algorithms Groups, Oxford
11 Teller, D.C., Horbett, J.A., Richards, E.G. and Schachman, H.K. (1969) Ann. N.Y. Acad. Sci. 164, 66–101
12 Creeth, J.M., Bhaskar, K.R., Donald, A.S.R. and Morgan, W.T.J. (1974) Biochem. J. 143, 159–170
13 Donnelly, T.H. (1969) Ann. N.Y. Acad. Sci. 164, 147–155
14 Wales, M. (1948) J. Phys. Coll. Chem. 52, 235–248
15 Teller, D.C. (1973) Methods Enzymol. 27, 346–441
16 Aune, K.C. and Timasheff, S.N. (1971) Biochemistry 10, 1609
17 Harding, S.E. and Creeth, J.M. (1982) IRCS Med. Sci. 10, 474–475.
18 Teller, D.C. (1965) Ph.D. Dissertation, University of California, Berkeley, Calif.
19 DeRosier, D.J., Munk, P. and Cox, D.J. (1972) Anal. Biochem. 50, 139–153
20 Carlisle, R.M., Patterson, J.I.H. and Roark, D.E. (1974) Anal. Biochem. 61, 248–263
21 Kim, H., Deonier, R.C. and Williams, J.W. (1977) Chem. Rev. 77, 659–690