Biophysical Chemistry 10 (1979) 183–185

MOLECULAR WEIGHTS FROM APPROACH-TO-SEDIMENTATION EQUILIBRIUM DATA USING NONLINEAR REGRESSION ANALYSIS

Leslie A. HOLLADAY
Department of Biochemistry, Vanderbilt University, Nashville, Tennessee 37232, USA

Received 23 October 1978

Data obtained from early times during the transient period of sedimentation equilibrium experiments are analyzed using an approximate solution to the Lamm equation to estimate s/D. The C_T versus r data obtained at several times during approach-to-equilibrium are analyzed using a nonlinear least squares algorithm and Fujita's approximate solution. This procedure was tested using D-Ser^{13}-somatostatin, ribonuclease, and ovalbumin. The results obtained demonstrate that for monodisperse samples s/D may be rapidly and reliably estimated using this method.

1. Introduction

Since the advent of the analytical ultracentrifuge many improvements in instrumentation, theory, and methodology have accumulated. The use of sedimentation equilibrium has become the standard method for estimating the molecular weights of proteins since the work of Van Holde and Baldwin [1] describing the use of short columns for the rapid attainment of equilibrium. For smaller proteins in aqueous solutions, sedimentation is reached in only a day. For large, asymmetric molecules or for unstable proteins, the time-to-equilibrium can be so large as to prohibit the use of the sedimentation equilibrium technique. Archibald [2] showed that estimates for M, the molecular weight, could be obtained during the transient period. The Archibald method requires extrapolation of the data to the ends of the solution column, and makes little use of data in the center of the solution column. The extrapolations required lead to some uncertainty in the final result.

The exact solution of Archibald to the Lamm equation for a sector-shaped cell with varying field is rather complex and cumbersome [3]. Yphantis and Waugh showed that by simple transposition the equations of Mason and Weaver [5] were applicable to ultracentrifugation by assuming the solution column to be rectangular and that a uniform field exists. This approximation appears to be quite good even for 1 cm columns [4], and entirely satisfactory for using short column (0.3 cm) data in estimating diffusion constants [1,6]. The Mason-Weaver solution contains an infinite series, but this series converges rapidly for the conditions generally used in low speed equilibrium experiments. Fujita [7] obtained an approximate solution which retains the correct boundary conditions for sector-shaped cells.

This study is concerned with examining the usefulness of using data from the transient period and Fujita's solution to provide estimates of M in only one-tenth of the time required to reach sedimentation equilibrium.

2. Materials and methods

2.1. Nonlinear least squares algorithm.

The algorithm estimates three parameters: s, s/D, and C_0. The method is based on the algorithm of Marquardt (1963) as implemented in the XZSSQ routine of the International Mathematical and Statistical Library [8]. This implementation uses a finite difference method for estimation of the required partial derivatives [9]. Convergence was considered attained when the residual sum of squares for two successive iterations had a relative difference of less than 2×10^{-4} or when the gradient norm was less than 2×10^{-4}. In

Table 1
Molecular weights from approach-to-sedimentation equilibrium using a nonlinear least squares fit to Fujita's approximate solution

Sample	rotor speed (rpm)	c_0 (mg/ml)	times (s)	OD_{res} [a]	$s_{20} \times 10^{13}$ (s)	$M_{(s/D)20}$	M_{equil} [b]
D-ser^{13}-Somatostatin [c]	48000	0.20	660–3450	0.005	0.41	1578	1685
ribonuclease [d]	20000	0.32	2827–7950	0.004	1.53	13535	13455
ribonuclease [d]	20000	0.50	1368–5226	0.004	1.43	13772	13390
ovalbumin [d]	13000	0.50	1905–6705	0.009	2.77	46630	44649

[a] $OD_{res} = ((\Sigma_{i=1}^{n}(OD_{exp} - OD_{calc})^2)/(n-p))^{1/2}$.
[b] M as calculated from data at sedimentation equilibrium (20°C).
[c] in 0.1 M KCl 10 mM potassium phosphate, pH 7.0.
[d] in 0.15 M KCl 20 mM potassium phosphate, pH 7.0.

most cases, both criteria were met with the final iteration.

The computation of the function as defined by the Fujita equation involves computing the sum of an infinite series. Successive terms of the series were summed until the next term was less 0.001 of the sum of the previous terms. Tests of this method using more restrictive criteria demonstrated that this criterion introduced negligible error in the computation of $C(r, t)$.

2.2. Standard proteins and reagents

Three times recrystallized bocine ribonuclease (code R-0G-B) and two times recrystallized ovalbumin (code 0A-35P709) were purchased from Worthington Biochemical Corporation. D-ser^{13}-somatostatin was the generous gift of Dr. Jean Rivier. All salts were reagent grade.

2.3. Ultracentrifugation.

Data were obtained using a Model E ultracentrifuge equipped with electronic speed control and a photoelectric scanner. The monochromator was set at 280 nm, the monochromator slit at 1.0–2.0 mm, and the photomultiplier slit at 0.1 mm. An AN-H rotor with a 12 mm Kel-F centerpiece was used with a solution column height of 2.8 mm. Fluorocarbon oil FC-43 (Beckman Instruments) was used to provide a visible lower meniscus. In calculations, times for the scans of C versus r were corrected for lower angular velocities during acceleration as previously described [1]. In all cases, the rotor was accelerated using a 12 amp current until the desired speed was reached to ensure correctness of the times used in the calculations. The value for the partial specific volume of ribonuclease was taken to be 0.695 [10]. The value for the partial specific volume of ovalbumin was taken to be 0.748 [11]. For D-ser^{13}-somatostatin, the partial specific volume was calculated from the amino acid composition [12]. Densities of solvents were calculated assuming additivity of components [13]. All runs were at 20°.

At equilibrium, the OD_{280} versus r data were analyzed by fitting the entire $\ln(OD_{280})$ versus r^2 data to a parabola in r^2 and testing to determine if significant curvature existed in the plot. Molecular weights were estimated from equilibrium data by using the least squares slope of the $\ln(OD_{280})$ versus r^2 plot.

For approach-to-equilibrium scans, ten points of OD_{280}, r were taken for each time. The points were equally spaced, and covered the central 90–92 per cent of the column height. The true solution baseline for runs with ribonuclease and ovalbumin was determined by overspeeding the rotor at 56000 for three hours, then decelerating to the equilibrium speed, where several scans were made to determine the correct baseline position.

3. Results

The estimates of M as computed from approach to equilibrium data and from equilibrium data are given in table 1. The plots of $\ln(OD_{280})$ versus r^2 at sedimentation equilibrium appeared to be linear; the coefficient of the $(r^2)^2$ term was not significantly different ($p > 0.05$)

Table 2
Variation of $M_{S/D}$ for individual time points

Time of scan	OD_{res}	$s_{20} \times 10^{13}$ (s)	$M_{S/D}$
2827	0.003	1.34	10912
3750	0.003	1.56	15586
4730	0.004	1.55	13280
6030	0.004	1.48	12907
6990	0.002	1.67	11982
7950	0.003	1.44	15556
Mean result ± S.D.		1.51±0.11	13370±1891

from zero for the standard proteins when the data was fitted to a parabola in r^2. The agreement between M(transient) and M(equilibrium) varies between one and eight per cent.

Table 2 examines the question as to how many time points are required to reliably estimate M. A single scan (ten pairs of OD_{280}, r) appears to estimate M with an accuracy of about fourteen per cent. Six time points appear to estimate M to within about three to five per cent of the value obtained at sedimentation equilibrium. About thirty pairs of OD_{280}, r (two replicate scans) were used to estimate M at sedimentation equilibrium. The lack of fit of the data to the least-squares $C(r, t)$ versus r curve (OD_{res}) averages around 0.007 OD_{280}, close to the experimental error in the original scans.

4. Discussion

Some thirty years ago Archibald proposed using a solution to the Lamm equation to estimate molecular weights without waiting for sedimentation equilibrium [14]. This study demonstrates the feasibility of his suggestion. The use of the approximate solution of Fujita rather than Archibald's exact solution greatly shortens the computational time required for the algorithm to estimate s/D. Of interest is the observation that the estimates obtained for the parameter s are substantially in error, being lower than published values for ovalbumin [15] and lower than that obtained for the same sample of ribonuclease using the conventional sedimentation velocity method (1.80×10^{-13} s). In examining Fujita's approximate solution the parameter τ ($= 2D/(s\omega^2 r_a^2)$ occurs much more frequently than the parameter τ ($= 2s\omega^2 t$). In addition, the parameter τ contains the value for the corrected time of centrifugation. Both these factors may contribute to the unreliable estimation of s.

The principal experimental difficulty with Archibald's method is that $(dc/dr)_a$ is required. This means that (dc/dr) must be extrapolated to its value at the upper meniscus. As Fujita [16] points out, there is "no generally accepted guiding principal" available for this extrapolation. The method tested in this study uses data over the entire length of the cell column, requires only c versus r data, and involves no extrapolations. These considerations suggest that the method discussed herein should generally be more reliable than Archibald's method.

Acknowledgement

This investigation was supported by the Vanderbilt Diabetes-Endocrinology Center (AM-17026) and AM-15838. The excellent services of the Vanderbilt University Computer Center are gratefully acknowledged.

References

[1] K.E. Van Holde and R.L. Baldwin, J. Phys. Chem. 62 (1958) 734.
[2] W.J. Archibald, J. Phys. Colloid Chem. 51 (1947) 1204.
[3] W.J. Archibald, Ann. N.Y. Acad. Sci. 43 (1942) 211.
[4] D.A. Yphantis and D.F. Waugh, J. Phys. Chem. 60 (1956) 623.
[5] M. Mason and W. Weaver, Phys. Rev. 23 (1924) 412.
[6] F.E. LaBar and R.L. Baldwin, J. Phys. Chem. 66 (1962) 1952.
[7] H. Fujita, Mathematical theory of sedimentation analysis (Academic Press, New York, 1962) pp. 113–114.
[8] T.J. Aird and N.L. Darilek, in: International mathematics and statistics library, ZXSSQ1-ZXSSQ-6, IMSL, Houston (1975).
[9] K.M. Brown and J.E. Dennis, Jr., Numerische Mathematik 18 (1972) 289.
[10] W.F. Harrington and J.A. Schellman, Compt. rend. trav. Lab. Carlsberg, Ser. Chim. 30 (1956) 21.
[11] M. Dayhoff, G. Perlmann and D. MacInnes, J. Amer. Chem. Soc. 74 (1952) 2515.
[12] E.J. Cohn and E.T. Edsall, Proteins, amino acids, and peptides (Nafner Publishing Co., New York, 1943) pp. 370–381.
[13] R.C. Weast, Handbook of chemistry and physics (Chemical Rubber Publ. Co., Cleveland, 1975) pp. D242–D246.
[14] W.J. Archibald, J. Appl. Physics 18 (1947) 362.
[15] O. Lamm and A. Polson, Biochem. J. 30 (1936) 528.
[16] H. Fujita, Foundations of ultracentrifugal analysis (John Wiley & Sons, New York, 1975) p. 216.