Molecular Mass Determination by Sedimentation Velocity Experiments and Direct Fitting of the Concentration Profiles

Joachim Behlke and Otto Ristau
Max Delbrück Center for Molecular Medicine, 13122 Berlin, Germany

ABSTRACT A new method for the direct molecular mass determination from sedimentation velocity experiments is presented. It is based on a nonlinear least squares fitting procedure of the concentration profiles and simultaneous estimation of the sedimentation and diffusion coefficients using approximate solutions of the Lamm equation. A computer program, LAMM, was written by using five different model functions derived by Fujita (1962, 1975) to describe the sedimentation of macromolecules during centrifugation. To compare the usefulness of these equations for the analysis of hydrodynamic results, the approach was tested on data sets of Claverie simulations as well as experimental curves of some proteins. A modification for one of the model functions is suggested, leading to more reliable sedimentation and diffusion coefficients estimated by the fitting procedure. The method seems useful for the rapid molecular mass determination of proteins larger than 10 kDa. One of the equations of the Archibald type is also suitable for compounds of low molecular mass, probably less than 10 kDa, because this model function requires neither the plateau region nor a meniscus free of solute.

INTRODUCTION

Analytical ultracentrifugation is a powerful tool for molecular mass determination of macromolecules (e.g., see monographs of Harding et al., 1992; Schuster and Laue, 1994). Two main methods can be distinguished, the meniscus depletion sedimentation equilibrium technique proposed by Yphantis (1964) and the sedimentation velocity variant. Whereas the first method can be used for direct determination of the molecular mass (M), the latter yields at least sedimentation coefficients (s). To calculate the molecular mass, this parameter has to be combined with the diffusion coefficient (D) of the sample, which is usually obtained from overlay experiments using a synthetic boundary cell (Behlke et al., 1986) or measuring the time-dependent boundary spreading of concentration gradients in a special device (Muramatsu and Minton, 1988). Other methods to get D directly from sedimentation velocity experiments were developed by Attri and Lewis (1992) and Stafford (1996). Attri and Lewis have used an empirical sigmoid function to fit the concentration profiles to locate the radial position of the square root of the second moment of the concentration data. The approach of Stafford is based on the time derivative dc/dt, which is converted to an apparent sedimentation coefficient distribution function. Both methods have no direct basis as solutions of the Lamm equation. More accurate results for s and D are expected by direct fitting of sedimentation velocity concentration profiles using approximate solutions of the Lamm equation. Holladay (1979a) used Eq. 2.265 of Fujita (1962; also see Materials and Methods), an approximate Archibald-type solution, but without explicit consideration of the bottom region. The results of these experiments were unsatisfactory partially because of the very slow convergence of Fujita's solution in the form of an infinite series. In an effort to get more reliable results, Holladay (1979b, 1980) developed another approximate solution of the Lamm equation. Comparing it with the solution of Fujita and MacCosham (1959) and the half-height method, it was found that the approach given by Holladay (1979b) yields correct values for $\epsilon < 0.02$ or high–molecular mass compounds of greater than 50 kDa. Philo (1994, 1997) presented a method in which multiple raw data sets of concentration profiles, taken at various times during the run, were simultaneously fitted by a nonlinear least squares technique to appropriate solutions of the differential equation of the ultracentrifugation. These model functions derived by Fujita (1962, 1975) are approximate solutions of the Lamm equation (Lamm, 1929) of the Faxén type (Faxén, 1929). For the conventional double-sector cell, Philo used Eq. 2.94 (Fujita, 1975; see Materials and Methods), and for the synthetic boundary cell he used Eq. 2.127 (Fujita, 1975). Because Eq. 2.94 (Fujita, 1975) used by Philo (1994) does not fulfill the boundary condition at the meniscus, the first records of sedimentation experiments have to be omitted. That means it is difficult to estimate molecular masses of less than ~15 kDa. To overcome these difficulties, it is necessary to use model functions that fulfill the initial condition of the cell as well as the boundary condition at the bottom. For conventional cells, one of the equations given by Fujita (Eq. 2.280; Fujita, 1962) fulfills both conditions. It is the aim of the present communication to compare the efficiency of the model functions described by Fujita (1962, 1975) of the Faxén type as well as the Archibald type to estimate the sedimentation and diffusion coefficients of low–molecular mass proteins, especially those between 10 and 20 kDa. A computer program, LAMM, was written using the five model functions discussed above to analyze experimental sedimentation pro-

Received for publication 26 June 1996 and in final form 4 November 1996.
Address reprint requests to J. Behlke, Max Delbrück Center for Molecular Medicine, Robert-Rossle-Strasse 10, Berlin D13122, Germany. Tel.: 49-30-9406-2205; Fax: 49-30-9406-2802; E-mail: behlke@mdc-berlin.de.
© 1997 by the Biophysical Society
0006-3495/97/01/428/07 $2.00

files. The approach was checked on calculated curves obtained by the "finite-element method" (Claverie et al., 1975; Cox and Dale 1981) as well as using sedimentation velocity profiles of proteins with molecular masses of greater than 10 kDa. Using these model functions, which include up to six error functions, it is possible to obtain more reliable results also from traces with a low signal-to-noise ratio. Application of these equations for the analysis of sedimentation velocity runs will be demonstrated.

MATERIALS AND METHODS

Sperm whale myoglobin and hen egg lysozyme were obtained from Serva (Heidelberg, Germany), and cytochrome *c* was from Merck (Darmstadt, Germany).

Sedimentation velocity runs were performed with an XL-A ultracentrifuge (Beckman Instruments, Palo Alto, CA) equipped with UV absorption optics. Experiments were carried out in conventional double-sector or synthetic boundary cells (part 331,431), respectively.

In addition to the experimental curves, noise-free data obtained by the finite-element method (Claverie et al., 1975; Cox and Dale, 1981) were also used for the simultaneous determination of sedimentation and diffusion coefficients. Radial concentration profiles calculated for a sedimentation coefficient of 2 S and a diffusion coefficient of 1×10^{-6} cm^2/s considering 42,000 or 50,000 rpm (synthetic boundary cells) and 50,000 rpm (conventional cells), respectively, were obtained by 800 or 1600 data points between $r_m = 6.4$ cm and $r_b = 7.2$ cm. The accuracy of the curves is given by a simulation time, *dt*, and the radial step length, *dr*. A time step of $dt = 1$ s and a step length of $dr = 0.0005$ cm was formed to be optimal.

To estimate sedimentation and diffusion coefficients, sedimentation velocity concentration profiles were fitted directly based on the five different equations given by Fujita (1962, 1975). A computer program, LAMM, was written in Turbo Pascal 7, running under MS-DOS 6.0 on a Pentium personal computer. It is able to read up to 18 XL-A data files and fit them simultaneously (global fit). The necessary information such as elapsed time, $\omega^2 t$, and rotor speed are taken from the data header. All parameters of the model functions and, in addition, a baseline (offset) can be estimated or held constant at the initial value. The initial boundary values and the parts of the profiles to be included in the fitting procedure were determined graphically on the monitor using cross-hairs. The initial values for all other parameters were calculated by the program. The program accepts only absorbance values of less than then 3.2 absorbance units. The five model functions for sector-shaped cells (according to the monograph of Fujita, 1962; 1975) are:

1) Equation 2.127 (Fujita, 1975) for synthetic boundary (also used by Philo, 1994):

$$c = \frac{c_0 e^{-\tau}}{2}[1 - \mathrm{erf}(\xi)]$$

$$\xi = \frac{z - \tau}{2\sqrt{\epsilon\tau}} \qquad x = \left(\frac{r}{r_0}\right)^2 \qquad \tau = 2\omega^2 s t \qquad (1)$$

$$\epsilon = \frac{2D}{s\omega^2 r_0^2} \qquad z = \ln(x)$$

In this equation (denoted Eq. 1 here), c_0 is the loading concentration, r_0 is the liquid-liquid meniscus position, erf( ) is the error function, D is the diffusion constant, ω is the angular velocity, and s is the sedimentation constant.

2) Equation 2.191 (Fujita, 1962) is also applicable for synthetic boundary cells but includes an additional parameter for the concentration dependence of the sedimentation constant:

$$c = \frac{c_0 e^{-\tau}}{1 - \lambda} \frac{(1 - \mathrm{erf}(p))e^{p^2}}{(1 - \mathrm{erf}(p))e^{p^2} + \sqrt{1 - \lambda}(1 + \mathrm{erf}(\xi))e^{\xi^2}}$$

$$\xi = \frac{1 - \sqrt{x(1 - \zeta}}{\sqrt{\epsilon\zeta}} \qquad p = \frac{\xi - \alpha\sqrt{\zeta/\epsilon}}{\sqrt{1 - \lambda}} \qquad \lambda = \alpha\xi \qquad (2)$$

$$\zeta = 1 - e^{-\tau} \qquad \epsilon = \frac{2D}{s_0\omega^2 r_0^2} \qquad s_c = s_0\left(1 - \alpha\frac{c}{c_0}\right)$$

In the monograph of Fujita (1975), only the concentration gradient of this function is described. The symbol x has the same meaning as in Eq. 1. s_0 is the Svedberg constant at zero concentration. If the concentration-dependent parameter α becomes zero, this function is identical to that in Eq. 2.128 (Fujita, 1975) for synthetic boundary experiments.

3) Equation 2.167 (Fujita, 1975) or 2.108 (Fujita, 1962) is applicable to standard double-sector cells and was used by Philo (1994); it accounts for neither a liquid-liquid nor bottom boundary:

$$c = \frac{c_0 e^{-\tau}}{2}\left\{1 - \mathrm{erf}(\xi) + \frac{\sqrt{2\epsilon\sinh(\tau/2)}}{x^{1/4}[1 + (xe^{-\tau})^{1/4}]\sqrt{\pi}} e^{-\xi^2}\right\} \qquad (3)$$

All symbols have the same meaning as in Eq. 2. This function accounts for a special case of the initial condition for the synthetic boundary cell. If the boundary moves sufficiently away from the meniscus, Eq. 3 can be used for the analysis of experiments in conventional double-sector cells.

4) Equation 2.137 (Fujita, 1975) or 2.163 (Fujita, 1962; Fujita and MacCosham 1959) is useful for concentration profiles in standard double-sector cells if taking into account the liquid-air boundary. To improve this equation, a new empirical parameter, H, is introduced, following the idea of Holladay (1979b):

$$c = \frac{c_0 e^{-\tau}}{2}\left\{1 - \mathrm{erf}(p) - 2\sqrt{\frac{\tau}{\pi\epsilon}}e^{-p2} + \left(1 + \frac{\tau + \ln x}{\epsilon}\right)\right.$$
$$\left.\cdot\left[1 - \mathrm{erf}\left(\frac{\tau + \ln x}{\sqrt{4\epsilon\tau}}\right)\right]e^{((\ln x)/\epsilon)}\right\}$$

$$p = \frac{\tau - \ln x}{\sqrt{4\epsilon\tau}}; \qquad (4)$$

in the improved case ϵ may be modified to $\frac{\epsilon}{H}$

The exact validity of Eq. 4 is limited to the earlier steps of sedimentation experiments in which the ratio $x = (r/r_0)^2$ in the Lamm equation (r_0 means the air-liquid meniscus position) should be nearly 1. This restriction also applies to the equations developed for the synthetic boundary cell. In the case of Eq. 4, the limitation can be reduced by introducing a time-dependent parameter. Written in dimensionless variables, the Lamm equation reads

$$\frac{\partial\theta}{\partial\tau} = \epsilon e^{-z}\frac{\partial^2\theta}{\partial z^2} - \frac{\partial\theta}{\partial z} \qquad \theta = e^{\tau}\frac{c}{c_0} \qquad z = \ln(x) \qquad (4a)$$

In the presentation of Fujita and MacCosham (1959) the factor $\exp(-z)$ is replaced by unity. However, to a first approximation, this factor must fall below 1.0 during the time of sedimentation; therefore, replacing $\exp(-z)$ by 1.0 leads to an overestimation of ϵ. After having solved Eq. 4a, ϵ has to be corrected by time-dependent factor H, which approximately compen-

sates for the decrease of e^{-z}.

$$H = \frac{1}{1 + (e^{\tau} - 1)p} \tag{4b}$$

The expression $(e^{\tau} - 1)$ in the denominator in Eq. 4b describes the shift of the midpoint of the moving boundary caused by to the relation $(r^*/r_0)^2 = e^{\tau}$ (Fujita, 1975, Eq. 2.107).

5) Equation 2.280 (Fujita, 1962), not included in the new monograph (Fujita, 1975), takes into account both boundaries of the conventional cell. (Note that Eq. 5 contains a misprint. At the exponent of the last term in the fifth row, the factor λ is neglected in Fujita's 1962 monograph.)

$$\begin{aligned}\frac{c}{c_0} = {} & \frac{e^{-\tau}}{2}[\mathrm{erf}(p_{22}) - \mathrm{erf}(q_2)] \\ & + \frac{1}{2a}\{2(1 + a)e^{-\lambda(X-x)} - 2e^{-\lambda(1-a)(X-x) - \tau} \\ & -(1 + a)[(1 - \mathrm{erf}(q_{11}))e^{-\lambda(X-x)} \\ & - (1 - \mathrm{erf}(p_1))e^{-\lambda(1-x)}] - e^{-\tau}[(1 - \mathrm{erf}(p_2))e^{\lambda(1-a)(x-1)} \\ & - (1 - \mathrm{erf}(q_{22}))e^{-\lambda(1-a)(X-x)}]\}\end{aligned}$$

$$\begin{aligned} p_1 &= \frac{(1 + a)\sqrt{\tau}}{2\sqrt{a}} + \frac{\lambda\sqrt{a}}{2\sqrt{\tau}}(x - 1) \qquad p_2 = \frac{(1 - a)\sqrt{\tau}}{2\sqrt{a}} \\ &\quad + \frac{\lambda\sqrt{a}}{2\sqrt{\tau}}(x - 1) \qquad p_{22} = \frac{(1 - a)\sqrt{\tau}}{2\sqrt{a}} + \frac{\lambda\sqrt{a}}{2\sqrt{\tau}}(X - x) \\ q_2 &= \frac{(1 - a)\sqrt{\tau}}{2\sqrt{a}} - \frac{\lambda\sqrt{a}}{2\sqrt{\tau}}(x - 1) \qquad q_{11} = \frac{(1 + a)\sqrt{\tau}}{2\sqrt{a}} \\ &\quad - \frac{\lambda\sqrt{a}}{2\sqrt{\tau}}(X - x) \qquad q_{22} = \frac{(1 - a)\sqrt{\tau}}{2\sqrt{a}} - \frac{\lambda\sqrt{a}}{2\sqrt{\tau}}(X - x) \\ a &= \frac{2\epsilon}{1 + X} \qquad \lambda = \frac{1}{\epsilon} \qquad X = \left(\frac{r_2}{r_0}\right)^2 \qquad x = \left(\frac{r}{r_0}\right)^2 \end{aligned} \tag{5}$$

Here, r_0 is the air-liquid position, and r_2 is the bottom position. All other symbols have the same meaning as before. Eq. 5 is not of the Faxén type, because the cell length is finite (Archibald type). Here, the term x of the Lamm equation is approximated by the expression $(1 + X)/2$. Therefore, when applying this formula, the best results for the fitting procedure can be expected in the middle part of the concentration distribution curve or plateau region. This equation also considers the data points at the bottom region. Because of the low precision of the radial steps of the XL-A ultracentrifuge, not all data points can be taken for the analysis. Our program used all data points until a maximal slope of the concentration profile was reached.

From the estimated sedimentation and diffusion coefficients, the molecular mass of the macromolecule under investigation can be calculated using Eq. 6:

$$M = \frac{sRT}{D(1 - \rho\bar{\nu})} \tag{6}$$

where R is the gas constant, T is the absolute temperature, ρ is the solvent density, and $\bar{\nu}$ is the partial specific volume.

The accuracy of the fits was represented by the residuals, which were amplified by the mentioned factor due to the concentration profiles in the lower part of the figure.

Numerical methods

The time-consuming part of the fitting procedure is the numerical evaluation of the error functions. Fortunately a very effective method developed by Gautschi (1969) can be used to reduce the computing time, especially for the error functions. The derivatives of the fitting functions are calculated algebraically to save computer time. The fitting algorithm is a "damped least square" procedure according to Levenberg (1944) in the version given by Wynne and Wormell (1963). The progress of the fit can be followed on the display. Standard deviations of the parameters were calculated in the simple linear model version. Usually this seems sufficient, because the results differ often more between different experiments than the confidence interval of one scan set allows. The fitting procedure to analyze 12 data files (3300 data points) took 25 s of processing time on a Pentium personal computer (100 MHz).

RESULTS

Fitting of noise-free synthetic data

To compare the efficiency of the five model functions to obtain accurate sedimentation and diffusion coefficients, we have fitted Eqs. 1–5 to calculated concentration distribution curves. These were calculated for sedimentation and diffusion coefficients of 2 S and 1×10^{-6} cm^2/s using the finite-element method. In the fitting procedure the sedimentation and diffusion coefficients, the loading concentration (c_0), and the radius position at the meniscus or bottom (Eq. 5) were estimated. The baseline parameter (zero offset, common to all data sets) was held at zero.

Fitting of synthetic boundary data sets

For the fitting procedure, Eq. 1 and 2 have been used. Both model functions yield a very good fit, with only very small deviations given by the residuals in eightfold amplification (Fig. 1, *A* and *B*). The parameters obtained from the fit agreed with the expected values. In addition, Eq. 2 offers the possibility of analyzing the concentration dependence of sedimentation coefficients. An appropriate Claverie simulation with concentration-dependent sedimentation coefficients (Cox and Dale, 1981) demonstrates the usefulness of Eq. 2, which results in excellent values for s and D, whereas the parameter α deviates considerably (see Fig. 2, *A* and *B*, and Table 1).

Curve fitting of conventional double-sector cell data

In contrast to the synthetic boundary data, the curves for conventional double-sector cells were calculated for a larger period of 12–324 min after attaining the maximal speed of 50,000 rpm. Curves with the moving boundary at different radial regions have been involved in the fitting procedure for the three model functions.

When using Eq. 3, which does not account for the boundary at the meniscus or bottom of the cell, we do not observe an optimal fit to the calculated curves (Fig. 3). In particular, gradients without meniscus depletion show stronger deviations reflected in the fourfold-amplified residuals. Even when omitting the earlier data sets, a reasonable fit to the

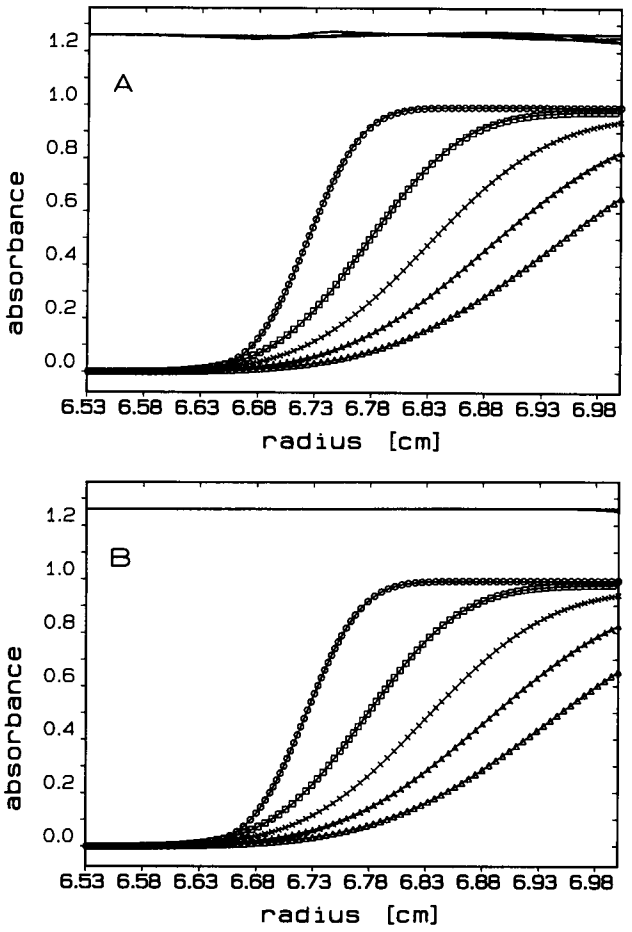

FIGURE 1 Claverie simulation for $s = 2$ S and $D = 1 \times 10^{-6}$ cm^2/s and 50,000 rpm (synthetic boundary). (*A*) Fit using Eq. 1; (*B*) fit using Eq. 2 with $\alpha = 0$. Residuals are given in eightfold amplification. Estimated values: (*A*) $c_0 = 1.0005$, $s = 1.991$ S, $D = 9.852 \times 10^{-7}$ cm^2/s; (*B*) $c_0 = 0.9999$, $s = 1.992$ S, $D = 10.001 \times 10^{-7}$ cm^2/s.

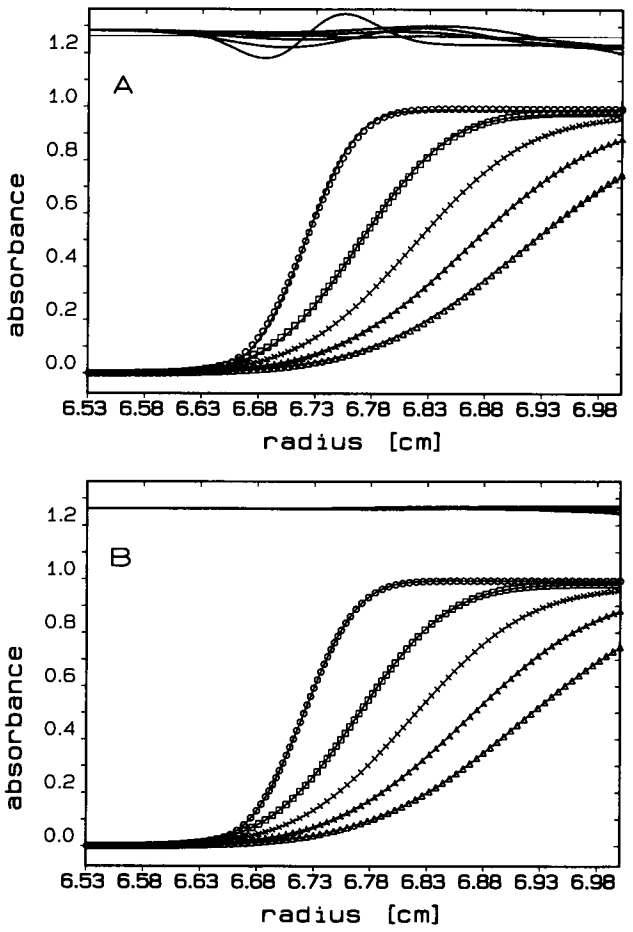

FIGURE 2 Claverie simulation (Cox and Dale, 1981) considering a concentration-dependent sedimentation with $s = s_0 (1 - \alpha c/c_0)$ for $s = 2$ S, $D = 1 \times 10^{-6}$ cm^2/s, 50,000 rpm, and $\alpha = 0.2$. (*A*) Fitting procedure by Eq. 2 with α held at 0; (*B*) Fitting procedure by Eq. 2 with additional estimation of α. Estimated values: (*A*) $c_0 = 0.9942$, $s = 1.794$ S, $D = 8.500 \times 10^{-7}$ cm^2/s, $\alpha = 0$ (held at 0); (*B*) $c_0 = 1.0016$, $s = 2.014$ S, $D = 10.09 \times 10^{-7}$ cm^2/s, $\alpha = 0.1114$.

later data sets could not be obtained. However, despite the poor fit, realistic sedimentation coefficients could be estimated. In contrast, the diffusion coefficients were found to be up to 7% too low (see Table 2), resulting in considerably higher molecular masses than expected.

Equation 4, which accounts for the boundary effect at the meniscus, fits the data very well, as shown in Fig. 4 *A*. To demonstrate the quality of the fit, the residuals were presented in an eightfold amplification relative to the scale of the concentration profiles. Despite the close fit, only the sedimentation coefficient agreed with the expected value, whereas the diffusion coefficient was found to be ~5% too low. Surprisingly, both parameters did not change significantly when skipping data points close to the meniscus position. In contrast to the original Eq. 4 the "improved" function yields the expected parameters more closely (Fig. 4 *B*). In the first line this is attributable to the more accurate determination of D. The parameter p does not depend on the molecular mass but slightly on the distance of the moving boundary from the meniscus. The optimal choice for p is ~0.5.

Equation 5 fits the data relatively well over a large radial region. Although this model function accounts for both boundary conditions of the cell, there are certain deviations near the meniscus (Fig. 5). Both the estimated s and D values were found to be slightly too low. Nevertheless, the

TABLE 1 Estimated parameters from Claverie simulation

Eq.	s (S)	D (10^7 cm^2/s)	c_0	r_0 (cm)	α
1	1.99_1	9.85_2	1.0005	6.6972	—
2	1.99_2	10.00_1	0.9999	6.6972	—
2*	1.79_4	8.50_0	0.9942	6.6971	0#
2*	2.01_4	10.09_0	1.0016	6.6974	0.1114

The following predetermined parameters were used: sedimentation coefficient, $s = 2$ S; diffusion coefficient, $D = 1 \times 10^{-6}$ cm^2/s; loading concentration of solute, $c_0 = 1$; initial boundary position, $r_0 = 6.6975$ cm; the concentration-dependent parameter, $\alpha = 0$ or 0.2 (for indicated equation); speed, 50,000 rpm (synthetic boundary), using different equations.
*$\alpha = 0.2$ (predetermined).
#α held at 0.

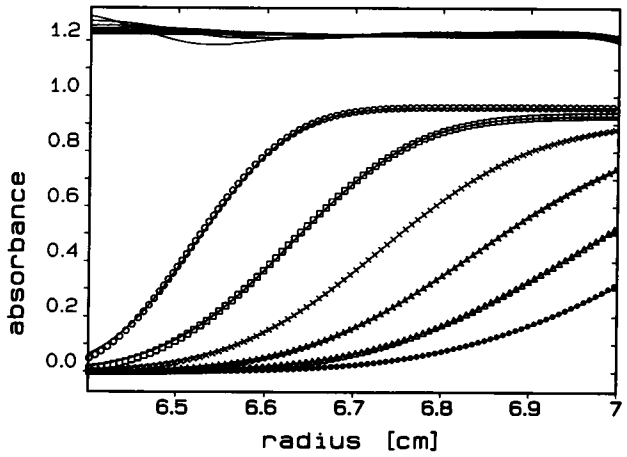

FIGURE 3 Claverie simulation for $s = 2$ S and $D = 1 \times 10^{-6}$ cm^2/s at 50,000 rpm (conventional double-sector cell) and fit of the data using Eq. 3. As pointed out earlier (Fujita, 1962), the fit to this function appears insufficient near the meniscus region (residuals amplified fourfold). Estimated values: $c_0 = 1.001$; $s = 2.012$ S; $D = 9.334 \times 10^{-7}$ cm^2/s; $r_0 = 6.398$ cm.

molecular mass obtained by these parameters is very close to the expected value. Omitting the first data points near the meniscus leads to a slight increase of the estimated sedimentation as well as diffusion coefficients reaching the expected values. The influence of the procedure on the molecular mass determination seems small. In contrast to the meniscus region, which has little effect on the results, the data points near the cell bottom up to 3.2 absorbance units seem critical for the accuracy of the results. Here the limit is set by the XL-A ultracentrifuge. The accurate radial steps of the Claverie simulation allowed one to involve all data points in the fit procedure independent of the slope.

Altogether, comparing the five model functions using synthetic data sets under various conditions, we can show that the most accurate values for sedimentation and diffusion coefficients can be obtained when fitting the data sets using Eq. 1 or 2 and for standard cells with small restrictions by Eq. 4 and 5. Eq. 3 results in reliable sedimentation data, but the diffusion coefficients were found to be underestimated. To obtain precise values for both parameters, strict meniscus depletion is necessary, which requires a larger column length and a higher speed of at least 60,000 rpm or smaller diffusion coefficients.

Analysis of experimental data sets

Because of our experience derived from the Claverie simulation with respect to the accuracy of the model functions, we have analyzed several proteins with molecular masses of greater than 10 kDa. In most cases the substances were tested under nearly identical conditions (protein concentration, buffer composition, speed, and column height in the cells). On average, 10 different traces of each experiment were used for the calculation of sedimentation and diffusion coefficients.

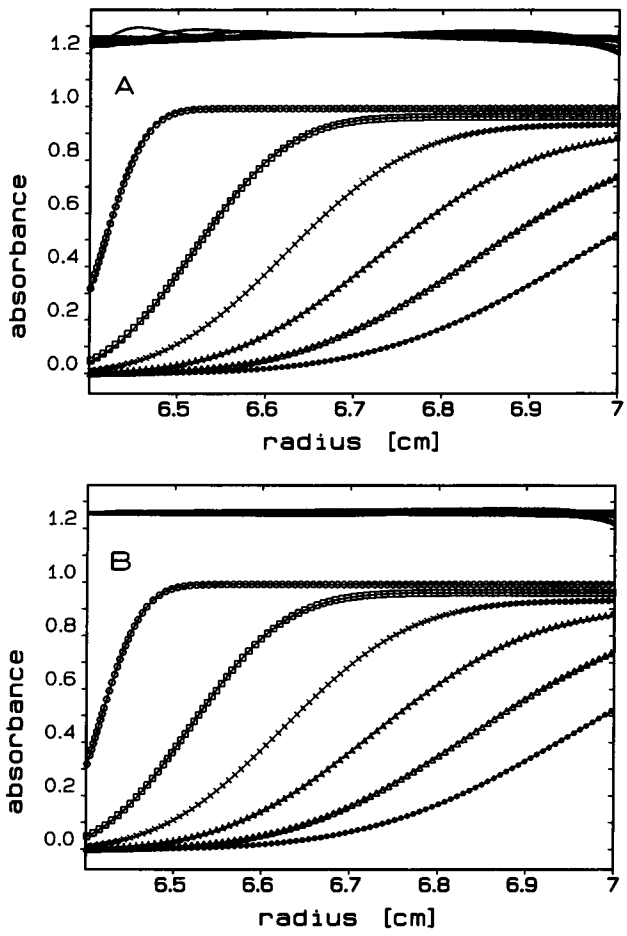

FIGURE 4 Fits of the simulation as shown in Fig. 3, by Eq. 4 (*A*) and considering the improvement as given by Eq. 4a and 4b (*B*). P is the "improving" parameter in Eq. 4b. Residuals are given in eightfold amplification. Estimated values: (*A*) $c_0 = 1.000$, $s = 1.991$ S, $D = 9.477 \times 10^{-7}$ cm^2/s, $r_0 = 6.3998$ cm, $p = 0$ (held at 0); (*B*) $c_0 = 0.999$, $s = 1.988$ S, $D = 9.880 \times 10^{-7}$ cm^2/s, $r_0 = 6.4001$ cm, $p = 0.4119$.

Synthetic boundary experiments

Although Claverie data sets could be fitted very well by Eq. 1 and 2 in the experiments, the first traces often have to be omitted, because these curves are disturbed during the overlay procedure. Despite this disadvantage, reliable parameters could be obtained with respect to s, D, and accordingly for the molecular mass (see Table 3). Usually the experiments were carried out at a very low protein concentration at which the concentration dependence of s can be neglected. Therefore, it is not useful to use the parameter α, especially if the experimental curves are noisy.

Conventional double-sector cell experiments

As shown in Table 4 the use of model Eq. 3 and 4 yields nearly correct sedimentation coefficients, but the diffusion coefficients are underestimated. The modification of the original Eq. 4 reduces these deviations. As mentioned in Materials and Methods, part Eq. 5 fits the total radial

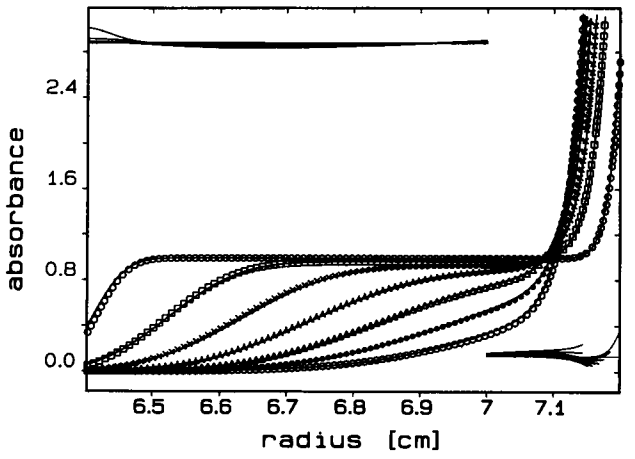

FIGURE 5 Fit of the Claverie simulation for $s = 2$ S and $D = 1 \times 10^{-6}$ cm²/s at 50,000 rpm using Eq. 5. Seven of the 14 data sets obtained for a long period are presented. Although Eq. 5 accounts for both boundaries, the fit in the meniscus region is not optimal (residuals in twofold amplification). Estimated values: $c_0 = 0.996$; $s = 1.957$ S; $D = 9.770 \times 10^{-7}$ cm²/s; $r_0 = 6.3917$ cm; $r_2 = 7.1973$ cm. When omitting the first data points up to 6.49 cm, the following results were obtained: $c_0 = 0.995$; $s = 1.998$ S; $D = 9.997 \times 10^{-7}$ cm²/s; $r_0 = 6.3827$ cm; $r_2 = 7.1976$ cm.

concentration profile and implies some problems concerning the limited accuracy of the radial steps. Therefore, the program allows one to cut all data points above a profile slope higher than 0.05 *A*/mm. This value is derived empirically. In a first run the slope is calculated, and in a second run the reduced data set is fitted again. This curve-fitting procedure enabled us to obtain reliable *s*/*D* ratios and, therefore, suitable molecular masses. However, the isolated parameters *s* and *D* are slightly too low. Because Eq. 5 does not allow an optimal fit for data near the meniscus suboptimally, it was found to be useful to remove the first part of the traces. This procedure led to a slight increase of both parameters *s* and *D* to the expected values (also see Table 4).

DISCUSSION

The possibility of direct molecular mass determination from sedimentation velocity runs based on approximate solutions of the Lamm equation (Lamm, 1929) is attributable to Fujita (Fujita and MacCosham, 1959; Fujita, 1962; Fujita, 1975), who, a long time ago, developed different model functions describing radial concentration profiles of high-speed sedimentation experiments. For many years the fitting procedure of such traces, using Fujita's equations with up to six error functions, seemed computationally too difficult for practical applications. However, more powerful computer hardware and effective methods for numerical evaluation of error functions given by Gautschi (1969) allow to use all the five equations from Fujita (1962, 1975) for the simultaneous estimation of sedimentation and diffusion coefficients routinely. These procedures permit rapid molecular mass determination for substances that are unstable and would

TABLE 2 Estimated parameters from Claverie simulation

Eq.	s (S)	D (10^7 cm²/s)	c_0	r_0 (cm)	r_2 (cm)	p
3	2.012	9.334	1.001	6.398_0	—	—
4	1.991	9.477	1.000	6.399_8	—	0*
4a	1.988	9.880	0.999	6.400_1	—	0.4119
5	1.957	9.770	0.996	6.391_7	7.1973	—
5#	1.998	9.997	0.995	6.382_7	7.1976	—

The following predetermined parameters were used: sedimentation coefficient, $s = 2$ S; diffusion coefficient, $D = 1 \times 10^{-6}$ cm²/s; loading concentration of solute, $c_0 = 1$; initial boundary position, $r_0 = 6.4$ cm; bottom radius, $r_2 = 7.2$ cm; speed, 50,000 rpm (synthetic boundary), using different equations.

*p held at 0.

#Omitting data points, $r < 6.49$ cm.

TABLE 3 Sedimentation and diffusion coefficients of cytochrome *c* and lysozyme obtained by fitting of the radial concentration profiles using Eq. 1 or 2 and the calculated molecular mass determined by Eq 6 with $\rho = 1.003$ g/cm³ (cytochrome *c*) or $\rho = 1.0005$ (lysozyme), respectively

Protein	Eq.	s (S)	D (10^7 cm²/s)	M_{sD}*	ΔM (%)#
Cytochrome *c*	1	$1.69_5 \pm 0.00_3$	11.82 ± 0.10	12,269	−0.5
Cytochrome *c*	2	$1.69_2 \pm 0.00_2$	11.73 ± 0.13	12,342	+0.01
Lysozyme	1	$1.90_6 \pm 0.00_2$	11.90 ± 0.15	14,478	+1.1
Lysozyme	2	$1.89_9 \pm 0.00_2$	11.80 ± 0.18	14,547	+1.6

Mean values and SDs are calculated from six different runs. Errors are calculated by the fitting program (in the simple linear model) are of nearly the same order.

*$\bar{\nu} = 0.713$ ml/g for cytochrome *c* (Timchenko et al., 1981); $\bar{\nu} = 0.730$ ml/g for lysozyme (Schausberger and Pilz, 1977).

#Deviations from the expected molecular mass of 12,330 for cytochrome *c* (Margoliash et al., 1961) or 14,316 for lysozyme (Canfield, 1963).

TABLE 4 Sedimentation and diffusion coefficients for cytochrome *c*, lysozyme, and myoglobin derived by fitting the concentration distribution profiles using Eq. 3–5, as well as molecular mass data and deviations from the expected values

Protein	Eq.	s (S)	D (10^7 cm²/s)	M_{sD}*	ΔM (%)#
Cytochrome *c*	3	$1.69_5 \pm 0.00_5$	11.01 ± 0.17	13,172	+6.8
Cytochrome *c*	4	$1.67_0 \pm 0.00_4$	11.35 ± 0.20	12,589	+2.1
Cytochrome *c*	4a§	$1.67_2 \pm 0.00_4$	11.66 ± 0.19	12,269	−0.5
Cytochrome *c*	5	$1.62_1 \pm 0.00_3$	11.60 ± 0.08	11,956	−0.3
Cytochrome *c*	5¶	$1.66_9 \pm 0.00_5$	11.75 ± 0.11	12,153	−0.15
Lysozyme	3	$1.88_3 \pm 0.00_4$	10.57 ± 0.21	16,112	+12.5
Lysozyme	4	$1.86_3 \pm 0.00_2$	10.82 ± 0.10	15,563	+8.7
Lysozyme	4a	$1.86_5 \pm 0.00_1$	11.42 ± 0.11	14,762	+3.1
Lysozyme	5	$1.79_4 \pm 0.00_2$	11.25 ± 0.03	14,414	+0.7
Lysozyme	5¶	$1.85_0 \pm 0.01_2$	11.63 ± 0.11	14,379	+0.4
Myoglobin	3	$1.99_0 \pm 0.00_9$	10.02 ± 0.29	18,925	+6.0
Myoglobin	4	$1.95_2 \pm 0.00_3$	10.14 ± 0.18	18,438	+3.2
Myoglobin	4a	$1.95_8 \pm 0.00_2$	10.48 ± 0.15	17,803	−0.3
Myoglobin	5	$1.88_7 \pm 0.00_2$	10.07 ± 0.11	17,856	−0.02
Myoglobin	5¶	$1.98_6 \pm 0.00_5$	10.69 ± 0.14	17,703	−0.9

*$\rho = 1.003$ g/ml; $\bar{\nu} = 0.742$ ml/g (Behlke and Wandt, 1973) for myoglobin.

#Molecular mass: 17,860 for myoglobin (Edmundson, 1965). All other data are given in Table 3.

§Eq. 4a is the improved model function (Eq. 4).

¶Omitting data points, $r < 6.49$ cm.

not survive in sedimentation runs for a long time. Because of possible microdisturbance in overlay experiments, we prefer sedimentation velocity runs with the conventional double-sector cells. With respect to the application of the different model functions to obtain reliable results, Eq. 3–5 can be recommended. From our experience, application of Eq. 3 requires longer columns. In contrast, the improved Eq. 4 and, in particular, Eq. 5 are suitable for experiments with smaller column lengths. Furthermore, Eq. 5 does not require a plateau region and fulfills the boundary conditions and is therefore suitable for substances with lower molecular masses. In future work we hope to eliminate the missing capability of Eq. 5 to fit the meniscus region of the concentration profiles by a small empirical alteration, as suggested for the improvement of Eq. 4.

The authors are grateful to Dr. Walter Stafford for the Claverie simulations.

REFERENCES

Attri, A. K., and M. S. Lewis. 1992. A fitting function for the analysis of sedimentation velocity concentration distribution. *In* Analytical Ultracentrifugation in Biochemistry and Polymer Science. S. E. Harding, A. J. Rowe, and J. C. Horton, editors. The Royal Society of Chemistry, Cambridge, United Kingdom. 138–166.

Behlke, J., U. A. Bommer, G. Lutsch, A. Henske, and H. Bielka. 1986. Structure of initiation factor eIF-3 from rat liver. *Eur. J. Biochem.* 157:523–530.

Behlke, J., and I. Wandt. 1973. Determination of partial specific volumes of hemoglobins and myoglobins. *Acta biol. Med. Ger.* 31:383–388.

Canfield, R. E. 1963. The amino acid sequence of egg white lysozyme. *J. Biol. Chem.* 238:2698–2707.

Claverie, J. M., H. Dreux, and R. Cohen. 1975. Sedimentation of generalized systems of interacting particles. I. Solution of systems of complete Lamm equations. *Biopolymers.* 14:1685–1700.

Cox, D. J., and R. S. Dale. 1981. Simulation of transport experiments for interacting systems. *In* Protein-Protein Interaction. C. Frieden and L. W Nichol, editors. John Wiley & Sons, Inc., New York. 173–201.

Edmundson, A. B. 1965. Amino acid sequence of sperm whale myoglobin. *Nature.* 205:883–887.

Faxén, H. 1929. Über eine Differentialgleichung aus der physikalischen Chemie. *Ark. Mat. Astr. Fys.* 21B:1–6.

Fujita, H. 1962. Mathematical Theory of Sedimentation Analysis. Academic Press, New York.

Fujita, H. 1975. Foundations of Ultracentrifugal Analysis. John Wiley & Sons, Inc., New York

Fujita, H., and V. J. MacCosham. 1959. Extension of sedimentation velocity theory to molecules of intermediated sizes. *J. Chem. Phys.* 30: 291–296.

Gautschi, W. 1969. Algorithm 363, complex error function. *Commun. ACM.* 12:635.

Harding, S. E., A. J. Rowe, and J. C. Horton, editors. 1992. Analytical Ultracentrifugation in Biochemistry and Polymer Science. The Royal Society of Chemistry, Cambridge, United Kingdom.

Holladay, L. A. 1979a. Molecular weights from approach-to-sedimentation equilibrium data using nonlinear regression analysis. *Biophys. Chem.* 10:183–185.

Holladay, L. A. 1979b. An approximate solution of the Lamm equation. *Biophys. Chem.* 10:187–190.

Holladay, L. A. 1980. Simultaneous rapid estimation of sedimentation coefficient and molecular weight. *Biophys. Chem.* 11:303–308.

Lamm, O. 1929. Die Differentialgleichung der Ultrazentrifugierung. *Ark. Mat. Astr. Fys.* 21B:1–4.

Levenberg, K. 1944. Method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* 2:164–168.

Margoliash, E., E. L. Smith, G. Kreil, and H. Tuppy. 1961. Amino acid sequence of horse heart cytochrome *c*. *Nature* 192:1121–1127.

Muramatsu, N., and A. P. Minton. 1988. An automated method for rapid determination of diffusion coefficients via measurements of boundary spreading. *Anal. Biochem.* 168:345–351.

Philo, J. S. 1994. Measuring sedimentation, diffusion, and molecular weights of small molecules by direct fitting of sedimentation velocity concentration profiles. *In* Modern Analytical Ultracentrifugation. T. M. Schuster and T. M. Laue, editors. Birkhäuser, Boston. 156–170.

Philo, J. S. 1997. An improved function for fitting sedimentation velocity data for low-molecular-weight solutes. Biophys. J. 72:434–443.

Schausberger, A., and I. Pilz. 1977. Präzionsmessungen zur Bestimmung spezifischer Volumengrößen von Makromolekülen in einem Vielkomponentensystem, im besonderen von Lysozym und Papain in wäßrigen Saccharoselösungen. *Makromol. Chem.* 178:211–225.

Schuster, T. M., and T. M. Laue, editors. 1994. Modern Analytical Ultracentrifugation. Birkhäuser, Boston.

Stafford, W. F. 1996. Rapid molecular weight determination by sedimentation velocity analysis. *Biophys. J.* 70:231*a*. (Abstr.)

Timchenko, A. A., A. I. Denesyuk, and B. A. Fedorov. 1981. Wide-angle X-ray scattering comparison of the structure of crystalline cytochrome *c* and cytochrome c in solution. *Biofizika.* 26:32–36.

Wynne, C. G., and P. M. J. H. Wormell. 1963. Lens design by computer. *Appl. Optics.* 2:1233–1238.

Yphantis, D. A. 1964. Equilibrium ultracentrifugation in dilute solutions. *Biochemistry.* 3:297–317.