# MEASURING SEDIMENTATION, DIFFUSION, AND MOLECULAR WEIGHTS OF SMALL MOLECULES BY DIRECT FITTING OF SEDIMENTATION VELOCITY CONCENTRATION PROFILES

**John S. Philo, Protein Chemistry Dept., Amgen Inc., Thousand Oaks, CA**
(from *Modern Analytical Ultracentrifugation*, T.M. Schuster and T.M. Laue, eds., Birkhauser, 1994, pp. 156-170)

## INTRODUCTION

Sedimentation velocity experiments have traditionally been used for proteins with relatively high sedimentation coefficients and low diffusion. Such proteins give sharp boundaries from which it is relatively easy to extract the sedimentation coefficient, and which permit the separation of multicomponent samples into distinct boundaries. However, many proteins of interest for therapeutic purposes, such as cytokines and growth factors, have molecular masses of only 10-40 kDa. Even at 60000 rpm, such small molecules give very broad boundaries which are difficult to analyze by existing techniques.

For example, the approach used in the program XLAVEL (Beckman Instruments) involves numerically differentiating concentration profiles to give $\partial c/\partial r$. The derivative data is then treated like Schlieren data, *e.g.* the natural log of the position of the maximum in $\partial c/\partial r$ may be plotted versus $\omega^2 t$ and then fitted to a straight line to determine $s$. However, the broad boundaries from small molecules give particularly small values of $\partial c/\partial r$. This makes the derivative data particularly sensitive to dirt or scratches on the windows, so typically considerable smoothing of the derivative data is required to even see a peak. In our hands, for small proteins this method gives significantly different $s$ values from the same data depending on the degree of smoothing and the person manually defining the position of the peaks in $\partial c/\partial r$.

Another approach that might seem more suitable for small molecules is to determine the boundary position from the second moment of the concentration profiles (Goldberg, 1953) and/or the diffusion coefficients from the boundary spreading (Muramatsu and Minton, 1988), as is done in the program VELGAMMA (Beckman Instruments). However, these methods require data at times where the meniscus is clear and there is also a plateau region at the bottom of the cell. For small molecules, by the time the meniscus is clear a plateau may no longer exist, due to the large movement of the leading edge of the boundary, as well as the typically 1-2 mm region where there is back diffusion of material accumulated at the bottom of the cell. (This is particularly true if aluminum centerpieces are not used and rotor speed is limited to 42,000 rpm). At best, there is often only a short time in the run during which this method may be applied, and the small boundary movement during this time can limit the accuracy of the analysis.

A more fundamental flaw with both of these approaches is that neither directly fits the raw experimental data. Rather, one or two properties are extracted from each scan, and these values are in turn fitted to an appropriate function of time during the run. These two-step approaches make it virtually impossible to correctly evaluate the confidence limits for the hydrodynamic parameters, and difficult to assess how strongly the results may be affected by factors such as the exact values chosen for the concentration in the plateau region or manual selection of peaks in $\partial c/\partial r$.

Moreover, neither of these approaches is useful when the sample contains more than one species, i.e. exactly the situation where sedimentation velocity can be often be particularly useful. All these considerations, as well as the availability of considerable computing power at low cost, have led us to develop a method where multiple raw data sets of concentration *vs*. radius, taken at various times during the run, are simultaneously fitted to appropriate approximate solutions of the Lamm equation, with $s$, $D$, and the loading concentration as fitting variables. As a bonus, the determination of $s$ and $D$ from the same experiment allows calculation of the molecular weight, and as shown below, this can be surprisingly

accurate. In addition, with this approach it is possible to analyze samples containing multiple non-interacting species.

## METHODS

**Numerical Methods.** In order to directly fit sedimentation velocity concentration profiles by non-linear least squares techniques, an appropriate theoretical fitting function must be chosen. Since there is no general, closed form solution of the Lamm equation, an approximate solution is needed which is accurate enough to represent the data, but not too time consuming to compute. We have chosen initially to ignore the concentration dependence of sedimentation coefficients, since for modern absorbance or refractometric optical systems one can easily work at concentrations of <1 mg/ml where the concentration dependence is generally negligible. Fujita (1975) has summarized a number of approximate solutions of the Faxén type (unbounded at the bottom of the cell), appropriate for different types of cells during different times during the run. For data from the conventional sector-shaped velocity cell, we therefore use equation 2.94 from Fujita (1975):

$$c = \frac{c_o e^{-\tau}}{2}\left\{1 - \mathrm{erf}\left[\frac{1-\left(xe^{-\tau}\right)^{1/2}}{\left[\varepsilon\left(1-e^{-\tau}\right)\right]^{1/2}}\right] + \frac{\left[2\varepsilon\sinh(\tau/2)\right]^{1/2}}{\sqrt{\pi}\,x^{1/4}\left[1+\left(xe^{-\tau}\right)^{1/4}\right]}\exp\left(-\frac{\left[1-\left(xe^{-\tau}\right)^{1/2}\right]^2}{\varepsilon\left(1-e^{-\tau}\right)}\right)\right\}$$

where $c_o$ is the loading concentration, $r_o$ is the meniscus position, erf() is the error function, $\tau \equiv 2s\omega^2 t$, $x \equiv (r/r_o)^2$, and $\varepsilon \equiv 2D/s\omega^2 r_o^2$. This solution does <u>not</u> account for restricted diffusion near the meniscus, and thus may not be appropriate at times very early in the run. It also cannot be applied near the bottom of the cell where solutes accumulate.

For synthetic boundary cells, equation 2.127 of Fujita (1975) is used:

$$c = \frac{c_o e^{-\tau}}{2}\left[1 - \mathrm{erf}\left(\frac{\tau - \ln x}{2\sqrt{\varepsilon\tau}}\right)\right]$$

where the terms are as defined above except that $r_o$ is now the initial position of the boundary. This equation accounts for diffusion of solutes above the initial position of the boundary, but again does not treat accumulation of solutes at the cell bottom.

A program, SVEDBERG, was written using Microsoft Visual Basic for Windows which incorporates these fitting functions into a non-linear least squares fitting routine. This program directly reads up to nine raw data files from the Beckman XL-A and extracts the elapsed time, $\omega^2 t$, and rotor speed from the data header. Baseline files may also be subtracted from the data to correct for poor matching of window absorbance or poor matching of solvent absorbance in sample and reference channels. The user then selects whether the data is from a conventional or synthetic boundary cell, and whether 1, 2, or 3 non-interacting species are to be included in the fit. The program then allows the user to graphically set the position of the meniscus (or the initial position of the boundary for synthetic boundary data) and an inner and outer radius that define the range of data to be included in the fit (constant for all data sets).

For each species, the program can either fit the values of $s$, $D$, and $c_o$, or hold them fixed at values input by the user. For synthetic boundary data the initial position of the boundary is also generally fitted, while for the conventional cell the meniscus position usually is held fixed (but may be fitted as desired). In addition, a zero offset (common to all data sets) may be included in the fit for those situations where the absorbance of the reference channel is not well matched to the sample.

The fitting algorithm uses a modified Gauss-Newton method developed in-house which is functionally identical to the "preferred method" described by Johnson & Faunt (1992; see also this volume). When

desired, SVEDBERG can also evaluate the confidence interval for each fitted parameter (which is generally asymmetric about the "best" value) again using the "preferred method" described by Johnson & Faunt. Evaluating these fitting functions is relatively slow, due primarily to the required numerical approximations of the error function. Therefore the derivatives of the fitting function, which are needed for the Gauss-Newton method, are calculated algebraically rather than by difference methods to avoid extra evaluations of the error function. For the conventional ultracentrifuge cell, in calculating these derivatives the last term in the large brackets in the fitting function (above) is ignored because it is quite small compared with other terms. (This approximation is equivalent to treating $\partial c/\partial r$ as a Gaussian function.)

Numerical simulations of sedimentation velocity experiments were carried out using the finite element Claverie method (Claverie *et al*., 1975), using program code kindly provided by David Cox and Walter Stafford. The sample was divided into 400 radial segments, and the time integration step size was 1 s.

**Experimental.** Centrifugation experiments were carried out with a Beckman XL-A ultracentrifuge using either conventional aluminum centerpieces or aluminum-filled epon synthetic boundary cells. Bovine serum albumin "monomer" was purchased from Sigma (#A1900). Recombinant transforming growth factor α and brain-derived neurotrophic factor were expressed in *E. coli*, refolded, oxidized, and purified to homogeneity by sequential column chromatography. All samples were made up in, or dialyzed into, Dulbecco's phosphate buffered saline (PBS) (Gibco). The light scattering/size exclusion analysis of the BSA sample was kindly done by Jie Wen (Protein Chemistry, Amgen) using methods described previously (Philo *et al*., 1993).

## RESULTS and DISCUSSION

**Single Species.** We have tested this method with a number of proteins with $M_r$ ~ 6,000-70,000. Figure 1 shows results for brain-derived neurotrophic factor (BDNF) in a conventional velocity cell at 60000 rpm. In solution BDNF is a very tightly bound homodimer with a sequence molecular weight of 27,274 per dimer. The derivative data shown in the inset illustrate the difficulty of defining the boundary position

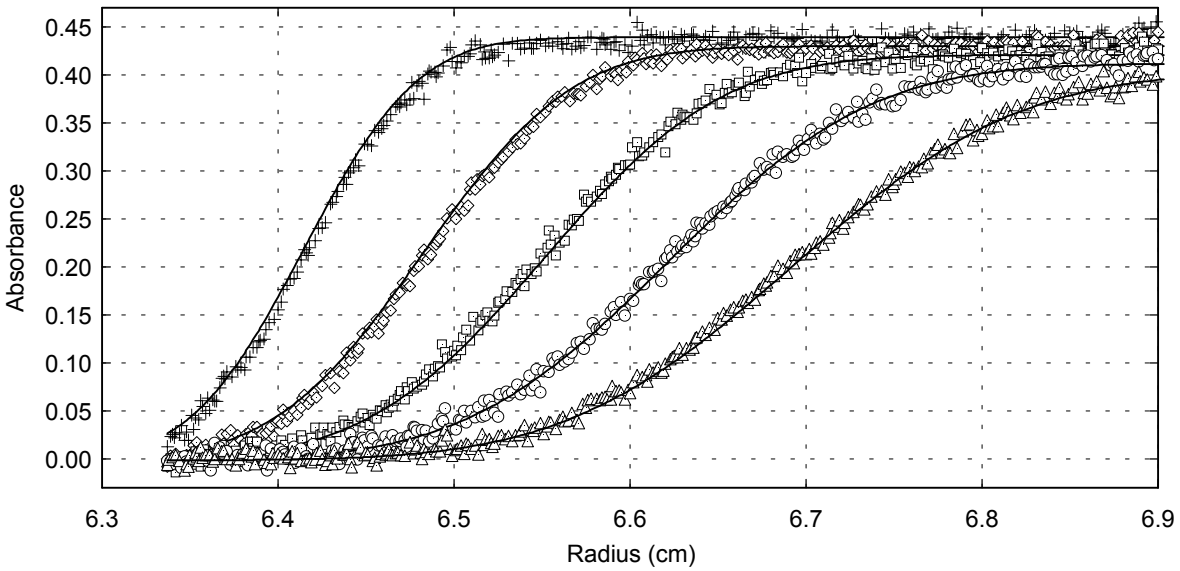


**FIGURE 1.** Sedimentation velocity data (symbols) and fitted data (curves) for 0.23 mg/ml BDNF in PBS at 20° C and 60000 rpm. Nine data sets were used in the fitting, but for the sake of clarity only the odd numbered sets are shown. The inset shows $\partial A/\partial r$ for the last data set after a 13-point smoothing.

from the peak even after considerable smoothing. Using our method, the fit to 9 scans (for clarity only 5 are shown) returns an $s$ value of 2.513 S with a 95% confidence interval of 2.509 to 2.517, a value for $D$ of 8.36 [8.24, 8.47] $\times 10^{-7}$ cm$^2$/s, a loading concentration of 0.4542 [.4529, .4554] absorbance units (AU) (equivalent to 0.23 mg/ml), and a zero offset of -0.0015 [-.0026, -.0005] AU, with an rms residual of .00704 AU. As also shown in Figure 1, the theoretical curves (solid lines) agree very well with the experimental data, and the residuals from the fit (not shown) are approximately randomly distributed.

If this analysis is correct, we should be able to calculate the molecular weight from the relation

$$M = \frac{sRT}{D(1-\bar{v}\rho)}$$

where $R$ is the gas constant, $\bar{v}$ is the partial specific volume, and $\rho$ is the solvent density. Using the measured $\rho$ = 1.00394 g/ml and $\bar{v}$ = 0.7271 ml/g calculated from the amino acid composition (Laue *et al.*, 1992), the fitted values of $s$ and $D$ imply a molecular mass of 27.27 [26.87, 27.71] kDa, in excellent agreement with the sequence molecular weight.

Based on ~1 year of experience to date, for proteins of 15-70 kDa this analysis typically gives molecular weights within ±2% of sequence values *if the sample truly contains only one species.* This accuracy compares quite favorably with that of sedimentation equilibrium, and may be limited in part by inaccuracies in the partial specific volume or density.

Another important point to note is that this fitting function seems to represent the data well even at times prior to complete clearing of the meniscus, even though it does not explicitly account for the restricted diffusion. We generally find it acceptable to include data after the concentration at the meniscus is <20% of the loading concentration. On the other extreme, this function seems to work well to at least the time when the leading edge of the boundary reaches the cell bottom.

**Multiple Species.** In contrast to the excellent fit of the BDNF data, Figure 2 shows a single species fit for a sample of commercial BSA "monomer", which clearly shows systematic deviations. This fit returns $s$ = 4.53 S and $D = 7.94 \times 10^{-7}$ cm$^2$/s, which implies $M$ = 53 kDa, well below the formula weight of 66,268. As shown in Figure 3, the fit is substantially improved if we include a second species, which reduces the variance by a factor of 1.9. For this fit, the first species (90.2%) has $s$ = 4.439 [4.429, 4.449] S, $D$ = 6.11 [5.97, 6.26] $\times 10^{-7}$ cm$^2$/s, and $M$ = 67.6 [65.9, 69.2] kDa, and therefore is an excellent match for the BSA monomer. The second species, contributing (9.8 ± 1)% of the total absorbance, has $s$ = 6.72 [6.63,

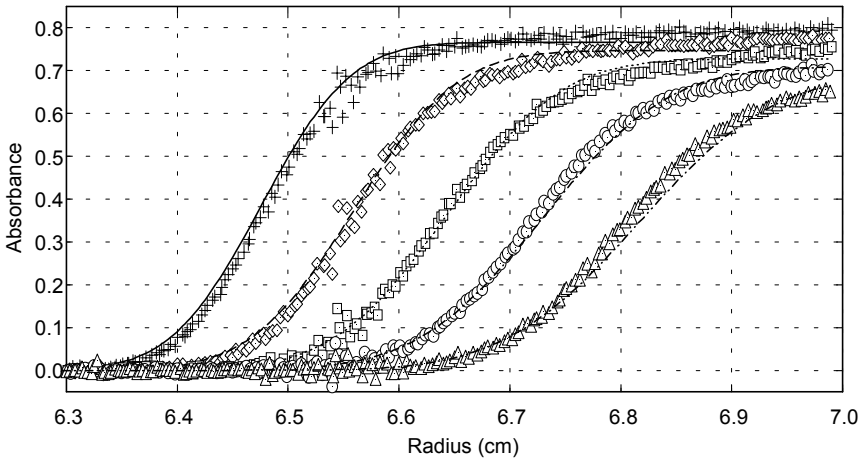


**FIGURE 2.** Sedimentation velocity data and single species fit for 1 mg/ml BSA "monomer" (Sigma) in PBS at 20°C and 60000 rpm. For the sake of clarity only 5 of the 9 data sets are shown.

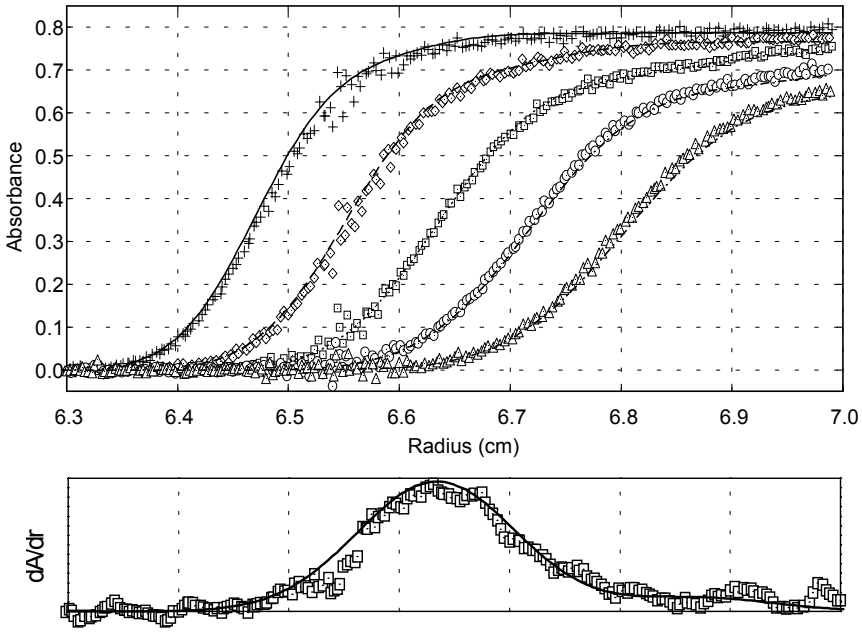


**FIGURE 3.** (upper) Sedimentation velocity data and two-species fit for BSA data from Fig. 2. (lower) $\partial A/\partial r$ for the last data set after a 13-point smoothing (symbols), and computed derivative from the two-species fit (curve).

6.80] S, $D$ = 5.35 [3.88, 7.37] × $10^{-7}$ $cm^2$/s, and $M$ = 117 [85, 163] kDa. Not only does the apparent $M$ suggest a BSA dimer, but (more significantly) the ratio of sedimentation coefficients for the two species is 1.51, nearly a perfect match for the ratio of 1.50 expected for a dimer of 2 hard spheres (van Holde, 1975).

To further validate these results, we analyzed this same material by size exclusion chromatography with light scattering detection. The chromatograms indeed show the presence of 9% dimer plus ~1% of an unresolved peak which appears to be a mixture of trimer and tetramer. If a third species with fixed s and D values corresponding to those expected for a BSA trimer or tetramer is also included, the sedimentation analysis returns about a 2% contribution, but the improvement in the fit is marginally significant. The derivative data in the lower panel of Fig. 3 shows that although a second peak for the dimer is resolved in theory, in practice 9 to 10% dimer would not be detected in a derivative analysis.

These BSA results also illustrate a number of general features of this method of analysis: (1) Treating a heterogeneous sample as a single species will usually give $s$ values close to those of the predominant species, but $D$ values which are too large. (The boundary spreading from the large $D$ can partially mimic the spread due to an additional species.) (2) In a multi-species analysis the $s$ and $c_0$ values are likely to be more reliable than are the $D$ values. (3) The method is fairly insensitive to scratches or dirt on the cell windows, as present near 6.55 cm in Fig. 2. (4) The ability to resolve species with markedly different $s$ values may be limited by the fact that this method requires that the rotor speed is constant and that the slowest sedimenting species is mostly clear of the meniscus, at which time faster sedimenting species may have reached the cell bottom. For example, it would be difficult to get accurate information about the presence of tetramers in the BSA sample, because they are still present in only the first 2 of the 9 data sets. In this case we could, in fact, have included data earlier in the run to get more contribution from faster sedimenting species, but this would compromise our ability to resolve monomers from dimers since this separation is much better later in the run. Such considerations suggest it may never be practical to determine $s$ and $D$ values for 3 species with this method.

**Claverie Simulations.** We have also tested this methodology with data simulated by the Claverie method (Claverie *et al.*, 1975). Figure 4 shows a Claverie simulation for $s = 2$ S and $D = 1 \times 10^{-6}$ cm$^2$/s ($M \approx 18$ kDa) at a loading concentration of 1 AU and 60000 rpm. Overlaid is the fit to these simulated data, which returns $s =$ 2.019 [2.018, 2.020] S, $D = 1.001$ [1.000, 1.002] $\times 10^{-7}$ cm$^2$/s, $c_0 = 1.0009$ [1.007, 1.0010], and a zero offset of -0.00031 [-0.0004, -.0002]. The residuals, shown in the right panel, have an rms value of only $6 \times 10^{-6}$ $c_0$ and the maximum deviation is only ~0.3% $c_0$. Thus the simulation suggests that the approximations in the fitting function are good enough to allow determination of $s$, $D$, and $M$ with an accuracy of 1% or better.

However, this excellent agreement with the Claverie simulations only holds true if the simulation places the initial position of the boundary away from the meniscus (i.e. allowing diffusion inward from the initial boundary, effectively treating the conventional velocity cell like a synthetic boundary cell). A Claverie simulation that does not allow diffusion above the meniscus produces narrower boundaries, especially early in the run. In this case the fit gives the correct $s$, but has much larger residuals and underestimates the true $D$ by about 6%. This is perhaps not surprising, since the approximate solution we are using does not explicitly account for restricted diffusion. Nonetheless, as noted above, the *experimental* data are well represented by this function, even very early in the run, and the fits give the correct $D$ and $M$. At present the source of this apparent conflict between experiment and simulation remains unclear. It is likely that the experimental data show early boundaries broader than predicted by the simulation because diffusion is occurring prior to the time the rotor reaches full velocity, while the simulation assumes instantaneous acceleration (and in fact it is our usual practice to pause at 3000 rpm long enough to take an initial scan). It is also possible that part of the conflict arises from the numerical approximations used in the Claverie method.

Claverie simulations have also been used to test whether it would be possible to resolve two species of much lower molecular weight than BSA. Simulations like that in Fig. 4 were run for $c_0 = 0.5$ AU of a monomeric species with s = 1.9 S, $D = 1.1 \times 10^{-6}$ cm$^2$/s (a ~16 kDa protein) and $c_0 = 0.5$ AU of its dimer, s = 2.85 S, $D = 8.25 \times 10^{-7}$ cm$^2$/s, at 60000 rpm. Random (Gaussian) noise with an rms amplitude of 0.006 AU, corresponding to typical noise in our XL-A, was then added to the simulated data for a more realistic test. The resulting test data, shown in Figure 5, show no resolution of the two species by conventional criteria, and as shown in the inset there is not even a distinct shoulder in the theoretical $\partial A/\partial$

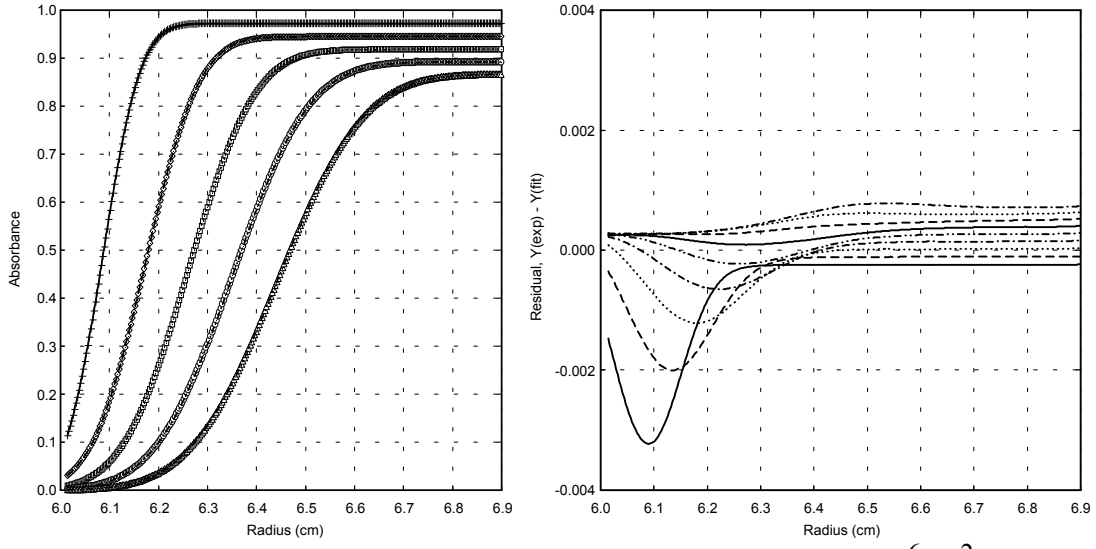


**FIGURE 4.** (left) Claverie simulation and single species fit for $s = 2$ S and $D = 1 \times 10^{-6}$ cm$^2$/s at 20°C and 60000 rpm. Five of 9 data sets are shown. (right) Residuals for all 9 data sets. The largest deviation is for the earliest time.

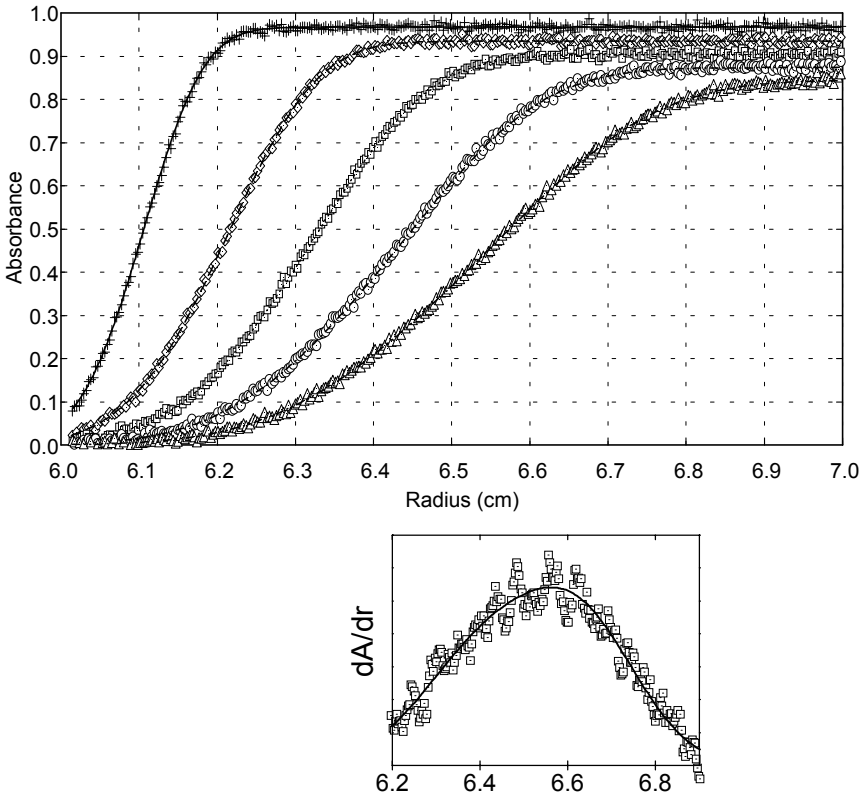


**FIGURE 5.** (upper) Claverie simulation and two-species fit for equal loading concentrations of a 1.9 S, $D = 8.25 \times 10^{-7}$ cm$^2$/s monomer and its dimer at 20°C and 60000 rpm. Random noise with rms amplitude 0.006 AU was added to the simulated data. Five of the 9 data sets are shown. (lower) 13-point smoothed $\partial A/\partial r$ for the last data set and $\partial A/\partial r$ predicted from the fit, showing the lack of resolution by conventional criteria.

*r*. These data were first tested to see if fitting could correctly deduce the proportions of the two species if the *s* and *D* values of the two species were known independently. A zero offset was also included in the fitting, since this is often necessary for experimental data. This fit found .4846 [.4821, .4871] AU of monomer and .5164 [.5144, .5183] of dimer, an error of ~3% from the true values.

If instead the *s* and *D* values are also fitted, .5147 [.4982, .5313] AU of s = 1.936 [1.918, 1.953] S, $D$ = 1.11 [10.76, 11.44] $\times 10^{-6}$ cm$^2$/s and .4861 [.4699, .5034] AU of s = 2.876 [2.863, 2.889] S, $D$ = 8.18 [8.04, 8.32] $\times 10^{-7}$ cm$^2$/s are found (curves shown in Fig. 5). This agreement is remarkably good, but it is also deceptive. An examination of the parameter cross-correlation coefficients (Johnson & Faunt, 1992) shows that the monomer $c_o$ has a -0.999 correlation with the dimer $c_o$, and the monomer *s* has a 0.9815 and -0.9842 correlation with these loading concentrations, respectively. These high cross-correlations mean that we can almost completely compensate for a change in any of these parameters by an adjustment of the others, *i.e.* they are not truly independently determined. (SVEDBERG alerts the user whenever there are cross-correlations > 0.97.) Therefore, for species this small, one would need independent information about *s* or *D* for one of the species in order to have confidence in the results.

**Comparison to the Attri & Lewis Method.** With these results in hand, it is also useful to compare our approach to a somewhat similar one developed independently by Attri & Lewis (1992). Their approach uses a sigmoid function to fit the concentration profiles, a function chosen because it is both rapid to compute and accurately locates the radial position of the square root of the second moment of the concentration data, but which has no direct theoretical basis as a solution of the Lamm equation. The

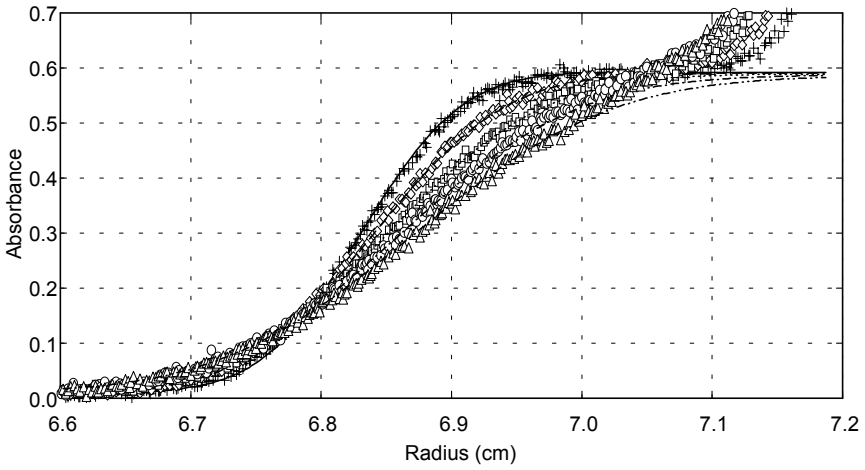


**FIGURE 6.** Synthetic boundary data and fits for TGFα at 20°C and 42000 rpm.

sigmoid function is first fitted to each data set, and the derived position and width parameters are then fitted in turn to an appropriate function of time to derive *s* and *D*. This approach appears to be reasonably accurate for single species. A sigmoid function does not reproduce the shape of experimental sedimentation boundaries or Claverie simulations as well as does the function used in our method, especially for broad boundaries like those in Figs. 1 and 4. The *D* values derived from their approach are therefore probably less accurate, especially when *D* is large. Attri & Lewis demonstrated some capability of resolving two components, but quantitative results for two-species fits were not presented. Further, like all two-step fitting approaches, their method makes it difficult to accurately assess the errors in the fitted parameters. This is particularly true for *D*, since the value of *s* is also explicitly used in the data transformation for determining *D*. Nonetheless, the Attri & Lewis method is clearly a viable approach and well suited to situations where one is primarily interested in rapidly computing sedimentation coefficients and obtaining at least an estimate of *D*.

**Synthetic Boundary Data.** To date we have little experience with applying this methodology to the synthetic boundary cell, but one example is shown in Figure 6, a measurement of a very small (5678 Da) protein, transforming growth factor α (TGFα) at 42000 rpm. The figure shows data at radii beyond the actual fitting region (which stopped at 7 cm) to illustrate a major drawback of this cell: the region where the data are influenced by back diffusion from the cell bottom grows so rapidly that very little boundary movement can be obtained. We also find that the earliest scans must usually be rejected because they show systematic deviations from the expected shape. This is probably due to inhomogeneities in the initial boundary, which later become less significant as they are averaged out by diffusion. The fit to these data gives $s$ = 0.849 [.825, .872] S, $D$ = 1.43 [1.41, 1.45] × $10^{-6}$ $cm^2/s$, the initial position of the boundary as 6.8166 [6.8156, 6.8176] cm, and gives reasonably non-systematic residuals. These results imply $M$ = 5.05 [4.91, 5.18] kDa. This is about 11% below the true value, and this error almost certainly reflects an overestimate of *D* because the initial boundary was not sharp. We have also tested the synthetic boundary fitting function against Claverie simulations, and again find excellent agreeement and that the approximations are good enough to give an accuracy of much better than 1%. These simulations also further emphasize the problem of back diffusion from the cell bottom, since a simulation of the data in Figure 5 shows that back diffusion contributes >1% of the total concentration at radii >7 cm.

While SVEDBERG also implements multi-species fits for the synthetic boundary cell, the smaller separation obtainable probably means that, at best, this may be useful for high molecular weight species. Overall, we have not found the use of the synthetic boundary cell to be very advantageous.

**Limitations and Future Developments.** It is important to emphasize that for multiple species this approach can only give meaningful results *if they are non-interacting.* Thus it is generally inappropriate for self-associating systems, unless the kinetics of redistribution among species is slow compared to the time scale of the experiment. Sample concentrations must also be kept low enough that the Johnston-Ogston effect is negligible.

As mentioned previously, the fact that the fitting functions assume a fixed rotor speed makes it difficult to measure samples containing widely different *s* values. In some cases one is only interested in the more slowly sedimenting species, but one or more larger species affect the data early in the run. In such cases we have found that including a species with variable *s* but a fixed very large value of *D* can be useful to model aggregates with a broad distribution of sizes.

SVEDBERG uses data from the same radial range for all data sets. This means that a large number of data points are included from the plateau region of early data sets, which can exacerbate the problem just discussed. Moreover, we sometimes observe sloping plateau regions in the data from our XL-A. This slope in the plateau remains constant over several data sets, and therefore it cannot be due to the presence of other species. (The origin of this problem is still unclear). Since the fitting functions cannot produce a sloping plateau, such data will be poorly fit and will give an overestimate of *D*.[1] Therefore it may be worthwhile in some cases to confine the data range for each data set to the vicinity of the boundary.[2]

Another practical limit to the accuracy of this analysis is the accuracy of sample temperature control. The temperature dependence of the viscosity of aqueous samples will produce a change in the apparent *s* and *D* of ~2.2% per °C. While the steady-state temperature control in the XL-A appears to be quite good, during acceleration to 60000 rpm the titanium rotor heats by ~.8°C,[3] so some temperature change during velocity runs seems unavoidable. Lastly, this analysis does not account for the movement of the boundary *during the course of a scan*, which could be significant for sharp boundaries and slow scan speeds.

While these fitting functions are relatively slow to compute, this is not a severe limitation. The two-species BSA fit in Fig. 3 (2189 data points, 6 parameters) takes about 2.5 min to converge using a 33 MHz Intel 486 CPU, while single species fits take < 1 min. The calculation of rigorous confidence limits for 2 species can take ~20 min, but this is usually only done for final results. Numerical calculations using Visual Basic are not particularly rapid, and a true 32-bit compiler would probably speed these computations 3-4 fold.

It is possible that the resolution and reliability of multi-species fitting could be improved by the inclusion of more data sets. (SVEDBERG's present limitation to 9 sets is entirely arbitrary). However, we have tried eliminating some of the data at intermediate run times, and this seems to have little effect other than slightly increasing the uncertainty in the parameters. Fundamentally, the resolution of the analysis is limited by the physical separation actually achieved, so as long as data covering the full time span appropriate for this method are included, varying the number of data sets will not have a dramatic effect.

In the future it will also be interesting to try this approach using more accurate approximate solutions of the Lamm equation. As computation speeds continue to increase, it may be practical on a microcomputer to use Claverie simulations as the fitting function, and in principle such simulations could allow for changing the rotor speed during the run.

**[Note added in proof]** It has come to our attention that an analytical approach similar to this one was reported some years ago. Holladay (1980) analyzed multiple sedimentation velocity scans by fitting $c_0$, $s$, and the ratio $s/D$, using an approximate solution to the Lamm equation (Holladay, 1979) which does account for restricted diffusion at the meniscus (and is therefore probably superior at early times in the run). Multiple-species fits were not reported. The fitting function used by Holladay is more complex

[1] A software approach to account for this optical artifact has now been implemented.
[2] The ability to define the fitting range for each data set has now been implemented.
[3] In principle, adiabatic expansion should produce cooling, but we invariably see the reported temperature of the rotor rise.

than that employed here (Fujita, 1975), and would probably require ~3 times longer to compute. In future work we hope to compare this fitting function to the one used here.

## GLOSSARY OF SYMBOLS

| | |
|---|---|
| AU | absorbance units |
| $c_o$ | loading concentration of solute |
| erf($y$) | the error function of y, $\frac{2}{\sqrt{\pi}}\int_0^y e^{-y^2} dy$ |
| $D$ | diffusion coefficient |
| $M$ | molecular mass |
| $R$ | gas constant |
| $r_o$ | meniscus or initial boundary position |
| $s$ | sedimentation coefficient |
| $t$ | elapsed time of centrifugation |
| $\bar{v}$ | solute partial specific volume |
| ρ | solvent density |
| ω | angular velocity |

## REFERENCES


Attri AK and Lewis MS (1992): A fitting function for the analysis of sedimentation velocity concentration distributions. In: *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, Harding SE, Rowe AJ and Horton JC, eds. Cambridge: The Royal Society of Chemistry.

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

Fujita H. (1975): *Foundations of Ultracentrifugal Analysis*. New York: John Wiley & Sons, pp. 64-81.

Goldberg RJ (1953): Sedimentation in the ultracentrifuge. *J. Phys. Chem.* 57: 194-202.

Holladay LA (1979): An approximate solution to the Lamm equation. *Biophys. Chem.* 10: 187-190.

Holladay LA (1980): Simultaneous rapid estimation of sedimentation coefficient and molecular weight. *Biophys. Chem.* 11: 303-308.

Johnson ML and Faunt LM (1992): Parameter estimation by least-squares methods. *Methods Enzymol.* 210: 1-37.

Johnson ML and Straume M (1993): Comments on the analysis of sedimentation equilibrium experiments. In: *Modern Analytical Ultracentrifugation*, Schuster TM and Laue TM, eds. Boston: Birkhauser Publishing Inc., this volume, Chapter 3.

Laue TM, Shah BD, Ridgeway TM and Pelletier SL (1992): Computer-aided interpretation of analytical sedimentation data for proteins. In: *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, Harding SE, Rowe AJ and Horton JC, eds. Cambridge: The Royal Society of Chemistry.

Muramatsu N and Minton AP (1988): An automated method for rapid determination of diffusion coefficients via measurements of boundary spreading. *Anal. Biochem.* 168: 345-351.

Philo JS, Rosenfeld R, Arakawa T, Wen J and Narhi LO (1993): Refolding of brain-derived neurotrophic factor from guanidine hydrochloride: Kinetic trapping in a collapsed form which is incompetent for dimerization. *Biochemistry* 32: 10812-10818.

van Holde KE (1975): Sedimentation analysis of proteins. In: *The Proteins*, 3rd ed., vol. 1, Neurath H and Hill R, eds. New York: Academic Press.