Determination of Sedimentation Coefficients for Small Peptides

Peter Schuck,* Cait E. MacPhee,# and Geoffrey J. Howlett#

*Molecular Interactions Resource, Biomedical Engineering and Instrumentation Program, ORS, National Institutes of Health, Bethesda, Maryland 20892 USA and #Russell Grimwade School of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria 3052, Australia

ABSTRACT Direct fitting of sedimentation velocity data with numerical solutions of the Lamm equations has been exploited to obtain sedimentation coefficients for single solutes under conditions where solvent and solution plateaus are either not available or are transient. The calculated evolution was initialized with the first experimental scan and nonlinear regression was employed to obtain best-fit values for the sedimentation and diffusion coefficients. General properties of the Lamm equations as data analysis tools were examined. This method was applied to study a set of small peptides containing amphipathic heptad repeats with the general structure Ac-YS-$(AKEAAKE)_n$GAR-NH_2, $n = 2$, 3, or 4. Sedimentation velocity analysis indicated single sedimenting species with sedimentation coefficients ($s_{20,w}$ values) of 0.37, 0.45, and 0.52 S, respectively, in good agreement with sedimentation coefficients predicted by hydrodynamic theory. The described approach can be applied to synthetic boundary and conventional loading experiments, and can be extended to analyze sedimentation data for both large and small macromolecules in order to define shape, heterogeneity, and state of association.

INTRODUCTION

The rate of migration of macromolecules in a centrifugal field is determined by a number of factors including the shape, density, and molar mass of the sedimenting species. Sedimentation velocity measurements therefore permit examination of the gross conformation of proteins and peptides and some of the changes which may be induced under different solution conditions. Interpretation of sedimentation coefficients in terms of model structures has been limited historically to rather simple shapes such as spheres, ellipsoids, and cylindrical rods where a simple relationship exists to describe the dependence of sedimentation velocity on the frictional coefficients that measure the resistance to movement. The ability to predict sedimentation coefficients for multisphere assemblies (Bloomfield et al., 1967; Kirkwood, 1949, 1954) has led recently to the development of new approaches to the interpretation of sedimentation velocity behavior where known three-dimensional protein structures are treated as equivalent bead models (Byron, 1997). This approach has been used to successfully predict the hydrodynamic parameters for several multisubunit proteins and has the potential to model the sedimentation velocity behavior of a range of different macromolecules including small peptides.

Recently, increasing attention has been paid to the development of sedimentation methods for the analysis of small proteins and peptides (Behlke and Ristau, 1997; Philo, 1997). This has been stimulated in part by the discovery of a number of important small regulatory and immunologically active proteins (Philo, 1997), and the increased availability of isolated protein domains and synthetic peptides as models for protein folding. Application of sedimentation analysis techniques to the study of small peptides is relatively straightforward in the case of sedimentation equilibrium experiments where the high rates of diffusion ensure rapid attainment of equilibrium. Sedimentation velocity analysis for small peptides is more difficult. Current analysis techniques are based on approximate analytical solutions of the Lamm equation (Holladay, 1979; Behlke and Ristau, 1997; Philo, 1997; Stafford, 1992) describing movement of a sedimentation boundary under conditions where accumulation of material at the bottom of the cell is small or negligible. However, for small peptides, the sedimentation rates are so small that clearly resolved and sedimenting boundaries are not observed. Even using synthetic boundary cells, the rate of boundary diffusion is sufficiently rapid that solvent and solution plateaus defining the boundary are quickly lost, and the influence of the finite length of the solution column becomes dominant.

In the present study we have overcome this limitation of traditional sedimentation velocity analyses and exploited the increased computing power now available even on PCs to utilize numerical integration of the Lamm equation for data analysis. There are a number of advantages. First, the influence of the meniscus and bottom on the sedimentation profiles can be accurately taken into account. Second, due to the generality of this approach, no assumptions about the initial distribution of the solute at the start of centrifugation are used. Instead, any experimentally measured distribution can serve as an initial condition for calculation of its evolution in the centrifugal field. Third, the approach taken is, in principle, extendible to the analysis of heterogeneous and interacting solutes at finite reaction rates, where analytical solutions of the Lamm equation are not available.

Received for publication 17 July 1997 and in final form 25 September 1997.

Address reprint requests to Dr. Peter Schuck, BEIP, ORS, Bldg. 13, Rm 3N17, National Institutes of Health, Bethesda, MD 20892-5766. Tel.: 301-435-1950; Fax: 301-496-6608; E-mail: pschuck@helix.nih.gov.

0006-3495/98/01/466/09 $2.00

We use these techniques to analyze the sedimentation behavior of several peptides, including a family of synthetic peptides containing heptad repeat sequences (Mulhearn et al., 1995) and interpret the data in terms of predicted sedimentation parameters calculated for model structures.

MATERIALS AND METHODS

Experimental

Peptide synthesis and purification

Peptide C, (ApoC-II$_{19-39}$ Ac-KESL SSYWESAKTAAQDLYEK-NH$_2$; M_r = 2475) corresponding to an amino acid sequence from mature human apolipoprotein C-II (Jackson et al., 1977), was synthesized using standard procedures for solid-phase peptide synthesis on a Rink resin, using an Applied Biosystems model 431A peptide synthesizer. Crude peptide C was purified on a Brownlee 10 × 250 mm semi-preparative Aquapore RP-18 reversed phase column. A linear gradient of 0.5% min^{-1} from 20% to 40% acetonitrile containing 0.1% TFA at 4 ml/min was used to elute the peptide. Desired fractions were detected by MALDI-TOF mass spectrometry, pooled, and lyophilized three times from pure water. Three peptides Ac-YS(AKEAAKE)$_n$GAR-NH$_2$; n = 2, 3, or 4 (referred to in the following as H2, H3, and H4, respectively) were kindly provided by Dr. R. Anders (Walter and Eliza Hall Institute of Medical Research, Melbourne). These peptides were synthesized by solid-phase peptide synthesis using procedures similar to those described previously (Mulhearn et al., 1995). Peptide purity was assessed by HPLC on a Brownlee 2.6 × 220 mm analytical RP-18 HPLC column and by MALDI-TOF mass spectrometry.

Circular dichroism spectroscopy

Circular dichroism (CD) spectra were recorded at 20°C on an AVIV 62DS spectrometer, using a 1 mm pathlength quartz cuvette. The instrument was routinely calibrated with an aqueous solution of *d*-10-camphorsulfuric acid. Ellipticities are reported as mean residue ellipticity (MRE) in deg cm^2 dmol^{-1}. The α-helicity of peptide H3 in 30% (v/v) 2,2,2-trifluoroethanol (TFE; obtained from Sigma, St. Louis, MO) was calculated using the mean residue ellipticity at 222 nm. The maximum ellipticity at 222 nm for peptide H3 when in the fully α-helical form was calculated to be 35,596 deg cm^2 dmol^{-1} using the peptide length-dependent equation described by Chen et al. (1974).

Sedimentation velocity

Sedimentation velocity experiments were conducted using a Beckman XL-A analytical ultracentrifuge equipped with absorption optics, using an An60-Ti rotor with cells containing quartz windows and either conventional or synthetic boundary double-sector charcoal-filled epon centerpieces. The peptides were redissolved in 0.1 M phosphate buffer, pH 7.4, to a final concentration of 0.5 mg/ml. Synthetic boundary cells were loaded with 150 μl solute with the reference sector filled with 400 μl buffer. Samples were initially centrifuged at 3000 rpm to allow determination of the appropriate wavelength for data acquisition. Samples were then centrifuged at 40,000 rpm and scans were taken across the length of the solution column at predetermined intervals. Absorbance data were obtained at the wavelengths indicated and at radial increments of 0.003 cm, each data point being an average of three measurements. Data were acquired at intervals over a period of 24 h, by which time the system had attained equilibrium. Equilibrium distributions were acquired at 0.001 cm radial increments, each data point being the average of five measurements. Equilibrium distributions were fitted for the molar mass of the solute by standard numerical analysis using the program SEDEQ1B (kindly provided by Dr. Allen Minton, National Institutes of Health, Bethesda).

Analysis of centrifugation data acquired in the presence of TFE requires knowledge of the variation in density and viscosity of the solvent mixtures. Density measurements of aqueous mixtures of TFE and water were performed at 20°C using an Anton Paar DMA 02C density meter. The variation in density with increasing TFE content was found to be described by:

$$\rho = 0.9982 + 0.50242 \times v - 0.1096 \times v^2 \qquad (1)$$

where ρ denotes the solution density, and v denotes the volume fraction of TFE per total volume of added TFE and water (C. MacPhee, M. Perugini, W. Sawyer, and G. Howlett, submitted for publication). The variation in solvent viscosity η with TFE content was found to be described by the fourth-order polynomial:

$$\eta = 0.9355 + 2.04 \times v + 1.355 \times v^2 - 7.532 \times v^3 + 4.464 \times v^4 \qquad (2)$$

Calculation of theoretical sedimentation coefficients

Peptide structures were modeled using the program HYPERCHEM. Values for ϕ, φ, and ω angles of 180°, 180°, and 180°, respectively, were used to compute the spatial coordinates of extended structures, and values of −58°, −47°, and 180° were used to compute the corresponding α-helical conformations. Two methods were used to generate the corresponding bead models for these structures. The first was based on the algorithm AtoB (Byron, 1997) using a resolution of 3 Å and the option to create equally sized beads. The second method, referred to as the α-carbon method, involved taking the coordinates of the α carbons and placing equally sized beads at these locations so that the weight and density of the resulting model equaled that calculated from the amino acid composition of the peptide. These bead models were used to calculate a theoretical value for the sedimentation and translational diffusion coefficients of the peptides using the method described by García de la Torre et al. (1994) and the program HYDRO. Values used for the molar mass, partial specific volume, and hydration of the peptides were calculated as described by Laue et al., 1992.

Numerical Methods

Calculation of solutions to the Lamm equation

With the following description of the diffusional fluxes $j_D(r)$ and the sedimentation fluxes $j_S(r)$ from a volume element at r into a neighboring volume element at $r + dr$

$$j_D(r) = -D\frac{dc}{dr}$$
$$j_S(r) = -s\omega^2 rc(r) \qquad (3)$$

where D denotes the diffusion coefficient, s the sedimentation coefficient, ω the angular velocity of the rotor, and taking into account the radial dilution of the concentration distribution as it moves in the sector-shaped solution column in the analytical ultracentrifuge, the well-known Lamm equation can be derived:

$$\frac{dc}{dt} = \frac{1}{r}\frac{d}{dr}\left[rD\frac{dc}{dr} - s\omega^2 r^2 c\right] \qquad (4)$$

(Lamm, 1929; Fujita, 1962). Given any initial distribution $c(r, t_0)$, this partial differential equation describes in a general way the evolution of the concentration distribution $c(r, t > t_0)$ between the meniscus m and the bottom b of the solution column in the centrifugal field.

Finite difference (Cox, 1965; Dishon et al., 1966; Sartory et al., 1976) and finite element approaches (Claverie et al., 1975) have been described

for numerical solution of the Lamm equations, and both types have been used in the present study (implemented in a Windows program called *sedfit*, which is available from the authors on request). They have in common the approximation of the concentration distribution $c(r)$ using a grid of equidistant radial points $r_1, \ldots, r_N$. In the case of the finite difference method, $c(r)$ is expressed as a vector of concentrations, $\mathbf{c}$, in a series of compartments of uniform concentrations; in the case of the finite element approach it is approximated by a superposition of geometric elements with amplitudes $\mathbf{c}$. It can be shown that in the finite difference approach, a discretized analog of Eq. 3 leads to a matrix equation describing the propagation of $\mathbf{c}(t)$ during the time interval Δt

$$\mathbf{c}(t + \Delta t) - \mathbf{c}(t) = \Delta t \mathbf{E} \mathbf{c}(t) \tag{5a}$$

where $\mathbf{E}$ denotes a tridiagonal matrix with

$$\begin{aligned}
E_{i,i-1} &= \frac{1}{r_i \Delta r}\left[\omega^2 s r_{i-1}(r_i - 0.5\Delta r) + D\frac{(r_i - 0.5\Delta r)}{\Delta r}\right] \\
E_{i,i} &= -\frac{1}{r_i \Delta r}\left[\omega^2 s r_i(r_i + 0.5\Delta r) - D\frac{2r_i}{\Delta r}\right] \\
E_{i,i+1} &= \frac{1}{r_i \Delta r}\left[D\frac{(r_i + 0.5\Delta r)}{\Delta r}\right] \\
E_{1,1} &= -\frac{1}{r_1 \Delta r}\left[\omega^2 s r_1(r_1 + 0.5\Delta r) + D\frac{(r_1 + 0.5\Delta r)}{\Delta r}\right], \\
E_{N,N} &= -\frac{1}{r_N \Delta r}\left[D\frac{(r_N - 0.5\Delta r)}{\Delta r}\right]
\end{aligned} \tag{5b}$$

An equation of similar tridiagonal structure has been derived by Claverie et al. (1975) for the finite element approach, using hat functions as elements of first order

$$\mathbf{B}[\mathbf{c}(t + \Delta t) - \mathbf{c}(t)] = \Delta t[\omega^2 s \mathbf{A}^{(2)} - D\mathbf{A}^{(1)}]\mathbf{c}(t) \tag{5c}$$

where the matrices $\mathbf{A}$ and $\mathbf{B}$ denote integrals over the elements, as described in Eq. 11 in Claverie et al. (1975). These tridiagonal systems can be very efficiently solved (Press et al., 1992).

To increase the stability of Eq. 5 for larger time steps, which is essential for the rapid simulations needed for the fitting of experimental data, it can be used in a Crank-Nicholson scheme (Crank and Nicholson, 1947). This is accomplished by evaluating concentrations in the right-hand side of Eq. 5a or Eq. 5c in the middle during the time step Δt, approximated as $(\mathbf{c}(t) + \mathbf{c}(t + \Delta t))/2$. Insertion of this into Eq. 5a leads, for the finite difference method, to

$$\mathbf{c}(t + \Delta t) = (2\mathbf{I} - \Delta t \mathbf{E})^{-1}(2\mathbf{I} + \Delta t \mathbf{E})\mathbf{c}(t), \tag{6}$$

or, for the finite element method, to

$$\begin{aligned}
\mathbf{c}(t + \Delta t) = &[2\mathbf{B} - \Delta t(\omega^2 s \mathbf{A}^{(2)} - D\mathbf{A}^{(1)})]^{-1} \\
&\cdot [2\mathbf{B} + \Delta t(\omega^2 s \mathbf{A}^{(2)} - D\mathbf{A}^{(1)})]\mathbf{c}(t).
\end{aligned} \tag{7}$$

Both the finite difference (Eq. 6) and finite element approach (Eq. 7) have been implemented, using the evaluations of the integrals $\mathbf{A}$ and $\mathbf{B}$ given by Cox and Dale (1981). For a given spatial grid size N, the finite difference simulations introduce an error that can be empirically described as an artificially increased diffusion coefficient; it can be shown that it is larger than the true diffusion coefficient by a term of the order of $\sim\omega^2 s\Delta r$, where Δr is the size of a single compartment. However, for small sedimentation coefficients, this error is small, and can be additionally reduced by fitting the data with an adaptive grid size (see below).

The initial condition $c(r, t_0)$ is obtained from the first of a series of experimental concentration profiles selected for the analysis. In order to prevent the noise in this scan from propagating into errors of the subsequent calculated concentration distributions, the simulation is started with a number of small time steps Δt. This is followed by an adaptive step size Δt, which keeps the maximum relative change $\max(dc_i/c_i)$ of all compartments within an adjustable, predefined range. This procedure has the advantage of using a small Δt value when fluxes are high, while making use of the high stability of Eqs. 6 and 7 in using larger Δt values with approach to equilibrium. (Smaller Δt's are used, if necessary, to report $\mathbf{c}(t_i)$ at the times t_i of the experimental scans.)

The accuracy of the simulation was tested against existing simulation software using the program Svedberg (Philo, 1994), and it was tested to produce correct sedimentation equilibrium profiles. The accuracy and convergence of the simulated concentration profiles with respect to the grid size N and the maximal relative change $\max(dc_i/c_i)$ controlling the time steps was verified. For example, finite element simulations for a small solute under synthetic boundary conditions (similar to those in Fig. 1) with a grid size $N = 1000$ and $\max(dc_i/c_i) \leq 0.1$ were equal, within an rms deviation of <0.0007 OD, to either finite difference or finite element simulations with $N = 200$. Under these conditions, the value of $\max(dc_i/c_i)$ governing the time step did not affect the results significantly for all $\max(dc_i/c_i) < 10$. (However, for calculations at higher $s\omega^2$, N had to be increased and $\max(dc_i/c_i)$ values had to be reduced.) Calculating with several hundred compartments, one simulation for small s spanning 20000 s of real time generally could be performed on a 75 MHz PC in <1 s.

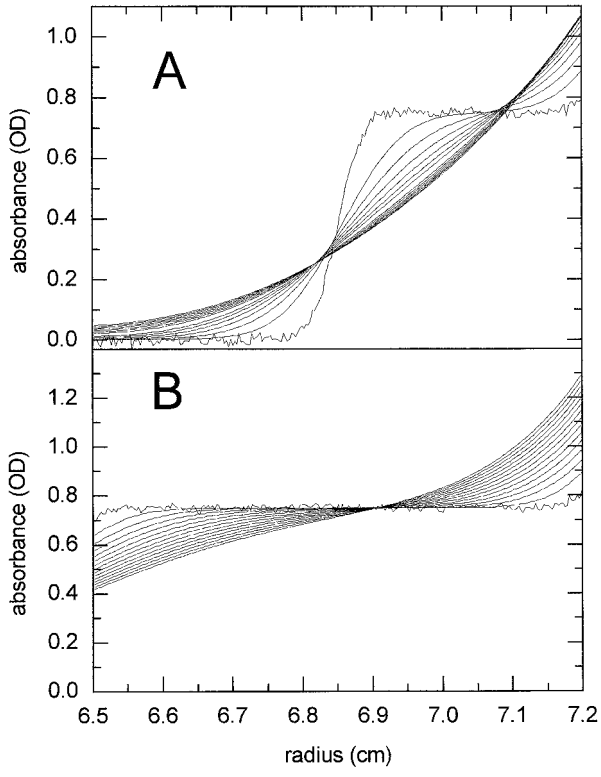

FIGURE 1 Simulated solutions of the Lamm equation showing the decay of noise in the initial conditions. Calculated concentration distributions for a small peptide ($s = 0.5$ S, $D = 2 \times 10^{-7}$ cm^2/s), starting from a simulated initial distribution with 0.01 OD Gaussian noise, under synthetic boundary conditions (*A*) and conventional loading conditions (*B*). Finite element simulations for concentration distributions with a solution column from 6.5 cm to 7.2 cm, at a rotor speed of 40,000 rpm, from a simulated sedimentation time of 200 s in 1000 s intervals to 15,200 s.

Fitting of experimental distributions

In our implementation, the values for the meniscus and bottom of the solution column can either be numerically entered, graphically determined, or can be treated as fitting parameters. To avoid fitting of data in the region of optical artifacts near the bottom and meniscus, the radial range of the data to be fitted was set separately. This point requires special attention for the initial scan, since the concentration distribution $c(r, t_0)$ has to be known over the entire solution column in order to calculate its time evolution. However, as long as the initial distribution does not contain high gradients, this problem can be solved by using polynomial extrapolation. In the implementation used, the order of the polynomial can be preset and the assumptions made by this extrapolation can be controlled and varied. A constant baseline offset absorbance can be added to the theoretical curves and optionally treated as unknown fitting parameter.

Starting values for D and s are optimized by using the Simplex algorithm, minimizing the sum of squared differences of all experimentally measured absorbance distributions $A_{exp}(r, t_i)$ and the theoretical absorbance distributions $A_{calc}(r, t_i)$. Optionally, the optimization cycle can be repeatedly restarted, doubling N each time, until the changes of the derived parameters values are within the desired tolerance. This adaptive grid is especially useful in conjunction with the finite difference calculations, compensating for its lower accuracy and reducing its error below the experimental error in data acquisition.

It can be useful to map the parameter space in s and D using the Svedberg equation

$$M = \frac{RTs}{D(1 - \bar{\nu}\rho)} \tag{8}$$

(where M denotes the solute molar mass, R the gas constant, T the temperature, $\bar{\nu}$ the partial specific volume of the solute, and ρ the solvent density) into a space of s and M as independent parameters to be optimized. The buoyant molar mass $M(1 - \bar{\nu}\rho)$ is available from analysis of sedimentation equilibrium profiles or from amino acid composition (Laue et al., 1992). This can be used to constrain the buoyant molar mass (i.e., effectively the ratio s/D) to an independently determined value.

RESULTS

Characterization of the data analysis

To examine the reliability of the fitting procedure and the conditioning of the inverse problem to obtain accurate sedimentation and diffusion coefficients, we have first generated and then re-analyzed sets of theoretical concentration distributions for solutes of different size. To mimic the analysis of experimental data, Gaussian distributed noise of 0.01 OD was added to the theoretical curves, and data points within 0.02 cm from meniscus, 0.03 cm within the bottom, as well as points with concentrations higher than 1.5 OD were excluded from analysis. This simulates the presence of optical artifacts in real data near the ends of the solution column.

In order to avoid assumptions on the initial concentration distribution (such as $c(r, 0)$ be a stepfunction in synthetic boundary loading experiments) the data analysis was started by taking the calculated distribution at 200 s as the initial condition (Fig. 1 *A*). A first-order polynomial extrapolation was used to estimate the initial distribution in the regions of optical artifacts. As is obvious from Fig. 1, the noise in the data sets used as the initial conditions for both synthetic boundary and for conventional loading is not propagated into the subsequent calculated distributions. Since the experimental noise corresponds to relatively steep local concentration gradients, this noise rapidly vanishes in the calculated evolution within a few small time steps due to the simulated diffusional fluxes between neighboring compartments. Over the time interval considered (4.2 h) the results in Fig. 1 show a more rapid approach to equilibrium for synthetic boundary loading than for conventional loading.

Table 1 presents best-fit values and error estimates for a range of solutes. The best-fit parameter values were always found to closely resemble the values generating the simulations, with small estimates for the statistical errors, low correlation between s and D, and nearly symmetrical error surfaces. This demonstrates that the information about s, D, and M included in such sedimentation profiles can be extracted in a well-conditioned analysis. It should be noted that this is true also for conditions where virtually no moving boundary is formed, and no solution or meniscus plateau is exhibited.

Under the simulated conditions, the estimated error for the sedimentation coefficient was constant at 0.01–0.02 S over a wide range of s values. In contrast, for the diffusion coefficient, the relative error remained approximately constant at ~2–3%. The use of the buoyant molar mass of the solute as prior knowledge in the analysis slightly improved the accuracy of the determination of s. If, on the other hand, the buoyant molar mass of the solute was treated as unknown, it could be determined with an accuracy of between 2% and 5% (Table 1). Under improved experimental conditions, such as a longer solution column and higher loading concentrations, the error in the determination of s could be as low as 0.01 S for a solute of 11 S (Table 1). This is comparable to the accuracy of 0.1% experimentally achieved with difference sedimentation velocity experiments with aspartate transcarbamylase using Schlieren optics (Howlett and Schachman, 1977).

Small deviations of the best-fit value from the true generating value of s can be introduced by the extrapolation of the initial distribution near the meniscus and bottom. However, if later scans that did not show a high curvature in these regions were taken as initial distributions, or if the assumption of constant distribution at $t = 0$ s was used, the accuracy improved significantly. Overall, the simulations demonstrate that the influence of the artifacts at meniscus and bottom regions on the initial conditions is small. Unlike the parameters s and D, an unknown baseline offset was found to be correlated with s, leading to slightly higher errors in s if the baseline offset was treated as an additional unknown.

It should be noted that sets of concentration distributions obtained during the approach to equilibrium in conventional low-speed experiments with proteins and other macromolecules of high molar mass can be used to extract s values (Table 1). Such experiments have sedimentation profiles which are similar in shape to those for the sedimentation of small peptides at high rotor speeds (results not shown).

TABLE 1 Best-fit values and calculated error estimates of for direct analysis of sedimentation distribution

	True s (S)	True D (10^{-7} cm^2/s)	Best-Fit s (S)	σ_s (S)	Best-Fit D (10^{-7} cm^2/s)	σ_D (10^{-7} cm^2/s)	σ_s With Prior Knowledge of M (S)	σ_M/M (%)
Synthetic boundary pure diffusion	0	20	0.000	0.010	19.69	0.5	—	—
Synthetic boundary (ω = 40,000 rpm)	0.5	20	0.504	0.016	19.65	0.6	0.013	1.7
Conventional loading (ω = 40,000 rpm)	0.5	20	0.504	0.008	19.93	1.1	0.008	4.0
Conventional loading (ω = 40,000 rpm)	1	15	1.013	0.01	15.27	0.35	0.009	2.1
Conventional loading (ω = 40,000 rpm)	3	10	3.04	0.02	10.20	0.1	0.012	1.7
Conventional loading (ω = 40,000 rpm)	5	7	5.07	0.02	7.14	0.09	0.02	3
Conventional loading (ω = 40,000 rpm)	13	3	13.34 (12.99)*	0.05	2.96 (3.04)*	0.05	0.05	5.4
Conventional loading longer solution column[#] (ω = 40,000 rpm)	11	3	10.995	0.01	2.969	0.11	0.01	8.0
Conventional loading equilibrium experiment[§] (ω = 8000 rpm)	3	2.71	2.99	0.03	2.66	0.25	0.02	2.0

Finite element simulations of solute distribution in a solution column of 6.5 cm to 7.2 cm, with data points saved at radial increments of 0.003 cm. The loading concentration was 0.75 OD. All calculated curves had 0.01 OD normally distributed noise added. Data were analyzed from 6.52 cm to 7.17 cm, with absorbancies <1.5 OD. The simulated data at 200 s were taken as an initial condition, with linear extrapolation outside the analysis range to the meniscus and bottom. The simulated data were fit from 1200 to 15,200 s in 1000-s intervals. The error estimate was based on *F*-statistics (Bevington et al., 1992).

*Initial data at 1200 s.

[#]Solution column from 6.2 cm to 7.2 cm, loading concentration 1 OD, noise 0.01 OD, scans from 100 s in 500-s intervals up to 8100 s.

[§]Experimental conditions similar to sedimentation equilibrium run of solute with molar mass 100,000 at 8000 rpm, solution column 6.7 cm to 7.2 cm, with 20 scans taken at intervals of 1 h.

Analysis of experimental sedimentation velocity data

A series of small peptides of different length and conformation was investigated in sedimentation velocity experiments in order to test the analytical procedures. Sedimentation equilibrium analysis indicated that under the conditions used, these peptides were primarily monomeric. Fig. 2 shows sequential concentration distributions obtained with peptide C in a synthetic boundary experiment, well described by a direct fit of the Lamm equation. In the sedimentation velocity data analysis, a small unknown baseline offset was taken into account, and the buoyant molar mass from the equilibrium data was taken as prior knowledge. The residuals are small and randomly distributed in all scans at all times ranging from shortly after the start of the experiment until equilibrium was attained. The fitted sedimentation coefficient of 0.46 S falls between the values for two structural extremes calculated from hydrodynamic theory for the peptide in completely extended and in α-helical structure (Table 2). As a test to determine the statistical accuracy of the derived sedimentation coefficient of 0.46 S, this experiment was performed in triplicate, with a standard deviation of below 0.01 S, confirming the high statistical accuracy predicted by the analysis of theoretical data (Table 1).

The sedimentation coefficients obtained for the series of heptad repeats using synthetic boundary loading conditions were 0.36, 0.45, and 0.52 S for the H2, H3, and H4 peptide, respectively (Table 2). These results were obtained by constraining the buoyant molecular mass of the peptides to the values calculated from their composition. Essentially identical values for the sedimentation coefficients were obtained by assigning buoyant molecular mass values derived from sedimentation equilibrium analysis. The reproducibility of the derived sedimentation coefficients for each peptide of the heptad family was verified in a conventional loading experiment (Fig. 3) immediately following the synthetic boundary experiment after mixing the contents of the centrifuge cell. The difference in the s values obtained in these experiments was smaller than 0.02 S for all heptads (Table 2).

The sedimentation coefficients obtained for the three heptad-containing peptides (Table 2) clearly reflect the increasing size of the molecules. The values lie between the limits of s-values calculated from hydrodynamic theory for these peptides in completely extended and in α-helical structure. The presence of a structure in between these extremes is confirmed by CD spectroscopy. Fig. 4 shows the CD spectrum for peptide H3 in aqueous solution (*solid line*). The strong minimum at 200 nm, and the lack of a

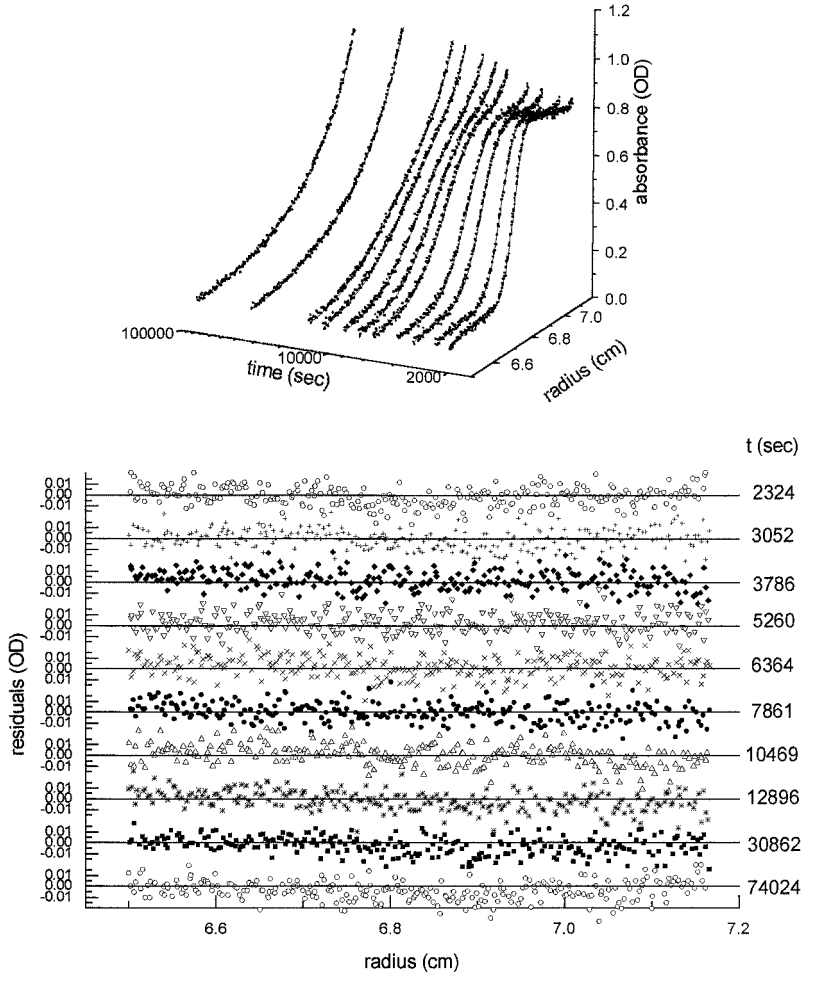

FIGURE 2 Experimental sedimentation distribution of peptide C and the best-fit finite element solution of the Lamm equation. *Upper panel:* Measured concentration distributions in the analytical ultracentrifuge at different times in a synthetic boundary experiment (*symbols*) and best-fit (*solid line*) using the buoyant molar mass from the sedimentation equilibrium as prior knowledge, while fitting for *s* (0.46 S) and a small constant baseline offset (0.012 OD). The first scan (1932 s) was taken as an initial condition. The lower panel shows the residuals at different times.

maximum at 190 nm, are typical of a random coil conformation. Data obtained for peptides H2 and H4 were essentially identical. The CD spectrum for peptide C (Fig. 4) also indicates a predominantly random coil structure in aqueous solution.

The accuracy of the sedimentation coefficients achieved in the experiments suggests the possibility of detecting shape differences in small peptides. Therefore, as an exemplary potential application of this method, preliminary experiments were performed with peptide H3 in the presence of TFE. This solvent has been widely used to examine the helical propensity of peptides. Titration with TFE showed that 30% (v/v) TFE was sufficient to induce the maximum change in the CD spectra. The mean residue ellipticity at 222 nm of the peptide in this concentration of TFE (Fig. 4) indicated a helical content of ~80%. To avoid assumptions regarding the partial specific volume of peptide H3 in 30% TFE, the sedimentation velocity data were analyzed with an unconstrained buoyant molar mass and sedimentation coefficient. The value obtained for the reduced molecular mass $M(1 - \bar{\nu}\rho)$ of 478 compares with a value of 444 calculated from the peptide composition and solution density. The sedimentation coefficient of the peptide in 30% (v/v) TFE in synthetic boundary experiments was 0.20 S, which gives a value, corrected for the density and viscosity of the TFE solution, of 0.52 S. Constraining the analysis to a reduced molecular mass of 444 (calculated from peptide composition and solution density) yielded a corrected value for the sedimentation coefficient of 0.49 S. Although due to the large correction factors these values are subject to a higher error than reported above, they are consistent with the higher sedimentation coefficient predicted by hydrodynamic theory for a more compact helical peptide (Table 2).

DISCUSSION

Direct fitting of the Lamm equation allows, in principle, the determination of sedimentation or flotation coefficients from any set of nonequilibrium analytical ultracentrifugation data. This eliminates the traditional requirement of measuring boundary movement, instead only relying on measurable sedimentation of the solute. The validity of this approach has been verified in the present paper using both computer simulations and experimental data. The principal intentions to develop this technique were, on the one hand, to obtain the ability to measure sedimentation coefficients and hydrodynamic shapes of small peptides which do not

TABLE 2 Theoretical and experimental sedimentation coefficients for the heptad family peptides and peptide C

	H2	H3	H4	Peptide C
Calculated hydration (g/g)	0.61	0.64	0.65	0.51
Molar mass	2049	2777	3505	2475
Calculated *s* (Byron, 1997)*				
s Extended	0.32	0.351	0.375	0.366
s α-Helix	0.455	0.521	0.571	0.512
Calculated *s* (α-carbon method)				
s Extended	0.327	0.353	0.374	0.381
s α-Helix	0.48	0.540	0.603	0.567
Experimental results#				
Synthetic boundary	0.36	0.45	0.52	—
Conventional loading	0.38	0.45	0.51	0.46
30% TFE§	—	0.52 (0.49)	—	

*Sedimentation coefficients are given in Svedberg units (10^{-13} s^{-1}).
#Sedimentation coefficients were determined from fits of the experimental data taking into account a possible small unknown baseline offset and using the known buoyant molar mass as prior knowledge. The solution conditions were 0.1 M potassium phosphate buffer, pH 7.4 and 20°C. The results are corrected for the effects of solution density and viscosity ($s_{20,w}$).
§Synthetic boundary data for peptide H3 in 0.1 M potassium phosphate buffer, pH 7.4 containing 30% TFE (v/v) and fitted with an unconstrained buoyant molar mass. The value in parenthesis is the fitted value using a constrained buoyant molar mass (444) calculated from the peptide composition and solution density.

exhibit sedimentation boundaries, and on the other hand, to explore the properties and potential of the Lamm equation when used as a general analytical tool.

The analysis procedure presented is well-conditioned and yields sedimentation coefficients with high statistical accuracy. In part, this appears to be due to the large number of data points that can be included in the direct analysis. The application to the heptad repeat peptides and to peptide C revealed *s* values with high reproducibility, increasing with peptide size, and centered between those calculated by hydrodynamic theory for fully extended and compact α-helical structure. That these peptides indeed assume a structure well in between those extremes in aqueous solutions appears reasonable from the CD spectra, which indicate the presence of a random coil. These findings demonstrate that the obtained *s*-values can be interpreted in the context of the size and the shape of peptides. This conclusion is consistent with the result from preliminary experiments in TFE, which suggests that the stabilization of α-helical structure in this solvent, as observed by CD, is accompanied by an increased *s*-value. The increase of the sedimentation coefficients of peptides in TFE solutions as a result of self-association will be described in MacPhee, C., M. Perugini, W. Sawyer, and G. Howlett (submitted for publication).

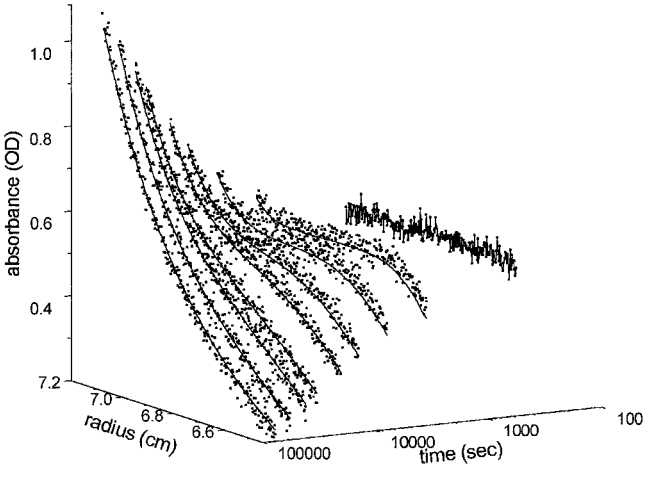

FIGURE 3 Concentration distributions of peptide H2 at different times in a conventional loading experiment (*symbols*). Best-fit distributions (*solid line*) calculated using finite element solutions of the Lamm equation based on the scan at 482 s as an initial condition. For results see Table 2.

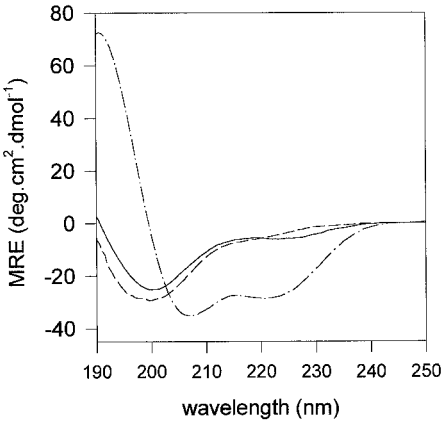

FIGURE 4 CD spectra of peptide H3 (—) and peptide C (– –) in aqueous solution (0.1 M phosphate buffer, pH 7.4); and peptide H3 (– · –) in 30% (v/v) TFE containing 0.1 M phosphate buffer, pH 7.4.

Analysis of the data in Table 2 assumed a single sedimentation coefficient for the peptides. Given the expected heterogeneity of conformations represented by the random coil state, the excellent fit to the data (Figs. 2 and 3) is noteworthy, suggesting a narrow range of sedimentation coefficients with the best-fit value representing a weight-average estimate. The observation that the values obtained fall within the limits calculated for a completely extended conformation and for a compact helical conformation raises the prospect of using the measurements to estimate the fraction of helical or compact structure. There appear to be two limitations to this approach. The first is the variation in the predicted values of the sedimentation coefficients using the stacked cube approach (Byron, 1997) compared to the α-carbon method (Table 2). These differences may reflect the dependence of the calculated values on the dimensions of the cube lattice (Byron, 1997). The second limitation is the uncertainty in treating hydration effects. The strategy used in the present work was to calculate hydration based on amino acid composition (Laue et al., 1992) and to swell the beads uniformly to give the correct expansion according to the volume occupied by the water of hydration. This method assumes that all of the amino acids are exposed to the solvent, which seems reasonable for the small peptides considered. Extension to the case of solutions containing

TFE introduces the additional problems of contributions of TFE to the solvent shell and higher statistical errors in the experimentally obtained sedimentation coefficients due to larger correction factors. This may require the combined use of other hydration-dependent techniques such as nuclear magnetic resonance and fluorescence anisotropy (García de la Torre et al., 1997).

Comparing the two experimental procedures employed, conventional and synthetic boundary loading, we found similar results with both techniques (Table 2). For equivalent column lengths, the synthetic boundary design reaches equilibrium more rapidly. This facilitates the independent determination of the buoyant molar mass for use as a constraint in the fitting procedure. Conversely, the conventional uniform loading experiment has the advantage of a greater choice of column lengths given the current restricted choice of synthetic boundary cells. The ability to observe initially a solute-free baseline near the meniscus in synthetic boundary experiments may be an advantage for detecting small baseline offsets, caused, for example, by imperfections in the absorbance optical data acquisition system. While not yet implemented in the present version of the program, the effects of such optical imperfections using either absorbance or interference optics could be minimized by fitting *dc/dt* versus *t* data obtained by subtracting sequential scans in a manner similar to that described by Stafford (1994). A further feature of the synthetic boundary design, important in the case of interference optics, is the capacity to directly measure the initial solute concentration in the appropriate refractometric units.

The advantage of using numerical versus approximate analytical solutions of the Lamm equation include the greater flexibility in the experimental design such as choice of rotor speed, length of the solution column, and the nature of the initial concentration distribution. The numerical approach avoids empirical correction factors (Philo, 1997; Behlke, 1997) or assumptions of an infinite solution column (Stafford, 1992). Additionally, since it accurately takes into account the end effects of the solution column, it allows the use of a lower rotor speed, diminishing the potential problems of the finite time necessary to take a scan with the current commercial absorbance optical system. Even for conventional low-speed sedimentation equilibrium experiments, estimates for s and D can be derived if the time course of approach to equilibrium is analyzed. The information contained in this part of the experiment so far has been neglected. On the other hand, the costs of the numerical approach are, first, the explicit need to specify meniscus and bottom position of the solution column, and second, the need for a concentration distribution $c(r, t_0)$ over the entire solution column initializing the analysis. It should be noted, however, that some of the more recent and advanced analytical methods (Philo, 1997; Behlke, 1997) also need meniscus and bottom position; and most techniques make assumptions $c(r, 0) = \text{const}$ over the entire solution column.

With respect to the above limitations, the meniscus can usually be determined experimentally with sufficient accuracy in the absorbance optical scans. The position of the bottom is not as easy to define. It may be extracted from an intensity scan observing the drop in the incident light in the reference sector when the detector reaches the shadow of the bottom of the cell, or it may be determined optically with the use of a synthetic bottom (e.g., using a fluorocarbon oil). On the other hand, since the Lamm equation implies mass conservation, the position of meniscus and bottom can be treated as additional unknowns to be determined. We observed that the exact bottom position can be slightly correlated with the sedimentation coefficient in the case of small solutes. Although this was not critical, the experimental determination seems advantageous.

The initialization of the analysis with a concentration distribution $c(r, t_0)$ becomes important if the analysis is started with an experimental scan at $t_0 > 0$ s, as is advantageous, for example, in synthetic boundary experiments to avoid transient mixing disturbances impairing the accuracy of the analysis. The concentration distribution within the regions of optical artifacts can be extrapolated with a first- or second-order polynomial. We found in computer simulations that the best-fit parameters for s and D are robust with respect to different extrapolations, as long as the curvature in $c(r, t_0)$ in the regions of the extrapolation is small. This suggests that the best data for initialization are obtained either from scans taken early in the experiment, or from those where the meniscus already has cleared to some visible extent, and accumulation of material at the bottom of the cell is visible. Another important feature of the initialization with an experimental concentration distribution, $c(r,t_0)$, is the ability to analyze sequential sets of concentration profiles during centrifugation. This feature has potential applications in the analysis of time-dependent conformational changes and corresponding time-dependent sedimentation coefficients, which might be expected, for instance, in protein refolding studies.

The careful study of the specific properties and problems of directly fitting the Lamm equation for a single component is an essential first step, and it is also unique with respect to its ease of application to a wide variety of starting conditions, justifying its separate examination. The numerical approach has potential extensions to a wide variety of experimental systems and to analyses where the concentration dependence of the sedimentation coefficient and radial variation in the solution density and viscosity are significant factors. However, from a practical viewpoint, the main limitation of the present approach is the limitation to a single component. Extensions of this method for noninteracting multicomponent systems are currently in progress, and are, in principle, straightforward by simple superposition of the independent sedimentation of components. Multiple components require independent concentration distributions as initial conditions. Except for the conventional loading technique at the start of the centrifuge, where all species can be assumed to be homogeneously distributed, these initial conditions might be difficult to obtain. This problem might be overcome, however, using chromophore

labels and multiwavelength techniques. Different methods of calculating solutions of the Lamm equation for multiple interacting species have been described (Claverie et al., 1975; Cohen and Claverie, 1975; Cox, 1971), for example via weight average sedimentation and gradient average diffusion coefficients assuming infinite reaction rates. Using the increased computational power recently available, they should be useful also when implemented as analytical tools. At present, it is not clear up to what level of complexity of interactions the inverse problem of extracting the sedimentation and diffusion coefficients, and equilibrium constants, from experimental data remains well-conditioned. We have shown here that the numerical approach to sedimentation velocity data analysis is not only feasible, but gives accurate values, and it has the potential to accurately take into account the reaction/diffusion processes of mixtures of solutes in the centrifugal field.

The authors acknowledge Dr. Emilios Dimitriadis, Dr. Marc Lewis, and Dr. Allen Minton for inspiring discussions and for critically reading the manuscript. The authors also thank Dr. Walter Stafford and Dr. Olwyn Byron for their advice and assistance with the bead model calculations.

This work was supported by a grant to GJH from the Australian Research Council.

REFERENCES

Behlke, J., and O. Ristau. 1997. Molecular mass determination by sedimentation velocity experiments and direct fitting of the concentration profiles. *Biophys. J.* 72:428–434.

Bevington, P. R., and D. K. Robinson. 1992. Data reduction and error analysis for the physical sciences. Mc-Graw-Hill, New York.

Bloomfield, V., W. O. Dalton, and K. E. Van Holde. 1967. Frictional coefficients of multi-subunit structures. I. Theory. *Biopolymers*. 5:135–148.

Byron, O. 1997. Construction of hydrodynamic bead models from high-resolution x-ray crystallographic or nuclear magnetic resonance data. *Biophys. J.* 72:408–415.

Chen, Y. H., T. Y. Yang, and K. M. Chau. 1974. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry*. 13:3350–3359.

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.

Cohen, R., and J.-M. Claverie. 1975. Sedimentation of generalized systems of interacting particles. II. Active enzyme centrifugation—theory and extensions of its validity range. *Biopolymers*. 14:1701–1716.

Cox, D. J. 1965. Computer simulation of sedimentation in the ultracentrifuge. *Arch. Biochem. Biophys.* 112:259–266.

Cox, D. J. 1971. Computer simulation of sedimentation in the ultracentrifuge. VI. Monomer-tetramer systems in rapid chemical equilibrium. *Arch. Biochem. Biophys.* 146:181–195.

Cox, D. J., and R. S. Dale. 1981. Simulation of Transport Experiments for Interacting Systems. *In* Protein-Protein Interactions. C. Frieden and L. W. Nichol, editors. Wiley, New York.

Crank, J., and P. Nicholson. 1947. A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Proc. Cambridge Philos. Soc.* 43:50–67.

Dishon, M., G. H. Weiss, and D. A. Yphantis. 1966. Numerical solutions of the Lamm equation. I. Numerical procedure. *Biopolymers*. 4:449–455.

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

García de la Torre, J., S. Navarro, M. C. Lopez Martinez, F. G. Diaz, and J. J. Lopez Cascales. 1994. HYDRO: a computer program for the prediction of hydrodynamic properties of macromolecules. *Biophys. J.* 67:530–531.

García de la Torre, J., B. Carrasco, and S. E. Harding. 1997. SOLPRO: theory and computer program for the prediction of SOLution PROperties of rigid macromolecules and bioparticles. *Eur. Biophys. J.* 25:361–372.

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

Howlett, G. J., and H. K. Schachman. 1977. Allosteric regulation of aspartate transcarbamylase. Changes in the sedimentation coefficient promoted by the bisubstrate analog *N*-phosphonacetyl-L-aspartate. *Biochemistry*. 16:5077–5083.

Jackson, R. L., H. N. Baker, E. B. Gilliam, and A. M. Gotto. 1977. Primary structure of very low density apolipoprotein C-II of human plasma. *Proc. Natl. Acad. Sci. USA*. 74:1942–1945.

Kirkwood, J. G. 1949. The statistical mechanical theory of irreversible processes in solutions of macromolecules (visco-elastic behaviour). *J. Chem. Phys.* 68:649–660.

Kirkwood, J. G. 1954. The general theory of irreversible processes in solutions of macromolecules. *J. Polym. Sci.* 12:1–14.

Lamm, O. 1929. Die Differentialgleichung der Ultrazentrifugierung. *Ark. Mat. Astr. Fys.* 21B:1–2.

Laue, T. M., B. D. Shah, T. M. Ridgeway, and S. L. Pelletier. 1992. Computer-aided interpretation of analytical sedimentation data for proteins. *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. 90–125.

Mulhearn, T. D., G. J. Howlett, G. E. Reid, R. J. Simpson, D. J. McColl, R. F. Anders, and R. S. Norton. 1995. Solution structure of a polypeptide containing four heptad repeat units from a Merozoite surface antigen of *Plasmodium falciparum*. *Biochemistry*. 34:3479–3491.

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. Birkhauser, Boston. 156–170.

Philo, J. S. 1997. An improved function for fitting sedimentation velocity data for low-molecular-weight solutes. *Biophys. J.* 72:435–444.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical recipes in C. 2nd ed (corrected 1994). University Press, Cambridge.

Sartory, W. K., H. B. Halsall, and J. P. Breillatt. 1976. Simulation of gradient and band propagation in the centrifuge. *Biophys. Chem.* 5:107–135.

Stafford, W. F. 1992. Boundary analysis in sedimentation transport experiments: a procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile. *Anal. Biochem.* 203:295–301.

Stafford, W. F. 1994. Boundary analysis in sedimentation velocity experiments. *Methods Enzymol.* 240:478–501.