Identification and Interpretation of Complexity in Sedimentation Velocity Boundaries

Borries Demeler,* Hashim Saber,# and Jeffrey C. Hansen*
*Department of Biochemistry, The University of Texas Health Sciences Center at San Antonio, San Antonio, Texas 78284-7760, and #Department of Mathematics, The University of Pittsburgh at Bradford, Bradford, Pennsylvania 16701 USA

ABSTRACT Synthetic sedimentation velocity boundaries were generated using finite-element solutions to the original and modified forms of the Lamm equation. Situations modeled included ideal single- and multicomponent samples, concentration-dependent samples, noninteracting multicomponent samples, and reversibly self-associating samples. Synthetic boundaries subsequently were analyzed using the method of van Holde and Weischet, and results were compared against known input parameters. Results indicate that this analytical method provides rigorous diagnostics for virtually every type of sample complexity encountered experimentally. Accordingly, both the power and utility of sedimentation velocity experiments have been significantly expanded.

INTRODUCTION

Most of the information that can be derived from an analytical sedimentation velocity experiment is contained in a sedimenting concentration gradient termed a "moving boundary" or simply a "boundary" (reviewed in Cantor and Schimmel, 1980, and van Holde, 1985). As is the case for any concentration gradient, a sedimentation velocity boundary also diffuses over time. Thus the shape of the boundary minimally will be influenced by the solute sedimentation coefficient and diffusion coefficient, as well as the rotor speed and length of sedimentation. Further complexity arises when more than one species is present, when the sedimenting species are capable of self- or hetero-associations, and when one or more of the respective sedimentation coefficient(s) are concentration dependent (Eisenberg, 1976; Correia et al., 1976). Consequently, although a sedimentation velocity boundary may contain a wealth of biologically relevant structural information, the complexity of the factors involved makes this information difficult to extract.

The single biggest barrier to correct interpretation of a sedimentation velocity experiment arguably is the boundary spreading caused by diffusion. In particular, diffusional spreading can mask the presence of multiple components. To address this problem, van Holde and Weischet (1978) developed a global fitting method that effectively removes the contribution of diffusion from a set of sedimentation velocity boundaries to yield an integral distribution of sedimentation coefficients, $G(s)$. It has been shown empirically that the G(s) profile both provides a useful diagnostic for sample homogeneity/heterogeneity and facilitates rigorous interpretation of the complex boundaries formed by multicomponent systems (reviewed in Hansen et al., 1994).

Coincident with the availability of the new generation analytical ultracentrifuge, the method of van Holde and Weischet (1978) has been increasingly utilized in recent years. Nonetheless, there are a number of important aspects of the $G(s)$ analysis that currently cannot be found in the literature. These include interpretive issues relating to complex multicomponent systems (e.g., limits of resolution, behavior of interacting versus noninteracting systems) as well as a demonstration of how concentration dependence of the sedimentation coefficient influences the $G(s)$ profile. In this article, finite-element solutions to the Lamm equation have been used to numerically generate complex sedimentation velocity boundaries. The synthetic data subsequently have been analyzed by the method of van Holde and Weischet (1978), and the $G(s)$ profiles compared against known input parameters. The results of these studies significantly enhance the interpretability of sedimentation velocity data obtained experimentally in the analytical ultracentrifuge, particularly if the samples contain multiple components and (or) exhibit concentration-dependent nonideal behavior.

MATERIALS AND METHODS

Synthetic sedimentation velocity data were generated and analyzed using a 486-type personal computer configured with both the LINUX (version 1.2.13) and MS-DOS/Windows operating systems, and the UltraScan software package (version 2.81). (UltraScan is a comprehensive data analysis software package, developed by one of the authors (BD), for use with the XL-A analytical ultracentrifuge.) Realistically matched sedimentation and diffusion coefficient combinations were predicted from previously established relationships between molecular mass and hydrodynamic shape (van Holde, 1985) using the XL-A utility module from UltraScan.

Data simulation

Sedimentation velocity data were simulated with the SEDVFIN module from UltraScan, which utilizes finite-element solutions to the Lamm equa-

Received for publication 26 June 1996 and in final form 9 October 1996.

Address reprint requests to Dr. Borries Demeler, Department of Biochemistry, The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Dr., San Antonio, TX 78284-7760. Tel.: 210-567-6592; Fax: 210-567-6592; E-mail: demeler@biocO2.uthscsa.edu.

0006-3495/97/01/397/11 $2.00

tion as described by Claverie et al. (Claverie et al., 1975; Claverie, 1976). (SEDVFIN is a sedimentation velocity data analysis program utilizing the finite-element solutions of the Lamm equation to simulate and fit boundaries to experimental data (B. Demeler and H. Saber, manuscript in preparation).) The Lamm equation (Eq. 1), which describes the sedimentation velocity behavior of a single ideal, noninteracting solute in the ultracentrifuge cell, has been modified to represent various other situations likely to be encountered experimentally. To simulate experimental data, solutions to the Lamm equation must be found. Analytical solutions to the Lamm equation are exceedingly difficult to obtain, especially when additional modifications of the model are included. A more practical approach is to solve each equation using a numerical approximation, such as the finite-element method. This method provides a discrete solution for a partial differential equation such as the Lamm equation. A detailed derivation of each finite-element solution has been described previously by Claverie et al. (Claverie et al., 1975; Claverie, 1976) and by Todd et al. (1981). Detailed information on the finite-element method and its application to discrete solutions of partial differential equations can be found in Zienkiewicz (1971). Listed below is a brief outline of the underlying principles for each modification used in this study.

Ideal noninteracting system

For an ideal noninteracting system the Lamm equation, F_k, is given by

$$\frac{\partial C_k}{\partial t} + \frac{1}{r}\frac{\partial (rJ_k)}{\partial r} = F_k; \quad J_k = s_k\omega^2 r C_k - D_k\frac{\partial C_k}{\partial r}, \tag{1}$$

with boundary conditions

$$J(r_m, t) = J(r_b, t) = 0, \quad 0 \le t \le T; \quad C(r, 0) = C_0(r). \tag{2}$$

In this equation, the subscript k refers to the solute k. For $k = 1$, this model represents the original Lamm equation of a single, ideal component as presented by Fujita (1962). C is the solute concentration, J is the flux, t is time, T is the elapsed time at the end of the experiment, r is the radius from the center of rotation, and $r_m \le r \le r_b$, where r_m and r_b are the radii of the meniscus and the bottom of the cell. ω refers to the angular velocity, and s_k and D_k are the sedimentation and diffusion coefficients of component k, respectively. F_k represents an external force, such as a chemical reaction. For an ideal solute, s_k and D_k are independent of C and are treated as a constant. Multiple noninteracting components were simulated by summing partial concentration profiles C_k (Eq. 3):

$$C_{total} = \sum_{k=1}^{i} C_k, \tag{3}$$

where i was ≤ 5 for cases presented in this study (see figure legends for details).

Solute nonideality/concentration dependence of s

For the case of concentration-dependent (i.e., nonideal) solution behavior, s_k is not a constant, but instead is a function of concentration. As described by Claverie (1976) and Todd et al. (1976), this can be represented in the form

$$s_k = s_{0k}\left(1 - \sigma_k \cdot \sum_{i=1}^{k} C_i\right) \tag{4}$$

where σ_k is a constant parameter that describes the variation of s_k from its value at infinite dilution, s_0. $\sigma = 0$ corresponds to the ideal case where s_k is independent of s. Equation 4 represents a first-order expansion of the Taylor series of the function $s(C)$, which is sufficient for most cases. The parameter σ is dimensionless. The model of concentration dependence of s as shown in Eq. 4 represents the simplest case, where interactions between different components are neglected.

Self-associating monomer-dimer equilibrium

A monomer-dimer equilibrium with monomer M and dimer D can be represented by the following relationship:

$$M + M \underset{k_2}{\overset{k_1}{\rightleftharpoons}} D, \tag{5}$$

with $k_1/k_2 = K = C_D/(C_M)^2$, and $C_D + C_M = C_L$. Here K is the equilibrium constant, and k_1 and k_2 are the forward and backward rate constants, respectively, C_L is the loading concentration, C_M is the monomer concentration, and C_D is the dimer concentration. The equilibrium that existed before sedimentation is constantly disturbed and reestablished during sedimentation, and consequently a weight-averaged sedimentation and diffusion coefficient are observed at each point in the cell. The concentration gradient averaged is dependent on the partial concentrations of each species M and D:

$$s_{avg} = \frac{s_M C_M + s_D C_D}{C_M + C_D} \tag{6}$$

and

$$D_{avg} = \frac{D_M(\partial C_M/\partial r) + D_D(\partial C_D/\partial r)}{\partial C_M/\partial r + \partial C_D/\partial r}. \tag{7}$$

Data analysis

Sedimentation velocity data were simulated using the finite-element solutions to the modified Lamm equations. The simulations were performed with 800 equally spaced radial data points of 0.00175 cm length, spanning a typical double-sector ultracentrifugation cell with the meniscus at 5.8 cm and the bottom of the cell at 7.2 cm. The time discretization step was set to 20 s/iteration for all simulations. All simulations were performed using a concentration proportional to an optical density of 1.0, representing a typical absorption of a ~1.0 mg/ml protein at 280 nm or a DNA concentration of 50 μg/ml at 260 nm, respectively. In Figs. 2 and 8, different loading concentrations were used for the simulations, and the details are listed in the figure legend.

Synthetic sedimentation data were analyzed by the method of van Holde and Weischet (1978) using

$$s_{app,i} = s_{actual} - (D^{1/2} \cdot C_i) \cdot t^{-1/2}, \tag{8}$$

the appropriate UltraScan data analysis software module. This analysis method, which is designed to remove the contribution of diffusion from the boundary shape, is based on the fact that sedimentation is a transport process proportional to the first power of time, whereas diffusional transport is proportional to the square root of time. Hence, in the limit of infinite time, the contribution of diffusion on the boundary shape is negligible. Briefly, the first step in this analysis is to divide each boundary into j equally spaced fractions, f_i, along the concentration axis (with $i = 1, 2, \ldots, j$). An apparent sedimentation coefficient ($s_{app,\,i}$) at each f_i was calculated for each scan according to Eq. 8 (a simplified form of equation 10 of van Holde and Weischet, 1978) and plotted against the inverse square root of time, where C_i is constant for each f_i of any given experiment. It is apparent from this equation that a linear extrapolation of s_{app} to infinite time ($t^{-1/2} = 0$) of equivalent f_i from each scan will yield a diffusion-corrected s_{actual} at the intercept. For a single component system, all regressions will converge with increasing time, leading to a distinctive "fan" shape. The slope of each linear regression is proportional to the diffusion coefficient of the species sedimenting at that position in the boundary

(which is reflected in the constant C_i in Eq. 8). An integral distribution of sedimentation coefficients, $G(s)$, was obtained by plotting f_i versus s_{actual}. All data presented in this paper were analyzed with $j = 20$ boundary divisions. At early sedimentation times, some boundaries did not sufficiently clear the meniscus. That is, if the boundary absorbance near the meniscus was larger than the total absorbance divided by the number of divisions j, the lower fractions f_i could not be calculated. These boundaries were excluded from the analysis. All other boundaries were included in the analysis. A tutorial for analogous editing of experimental data can be found on the Internet (http://bioc09.uthscsa.edu/.biochem/tutor.html). In all cases the baseline was simulated to be zero relative concentration.

RESULTS

Ideal single-component system with different sedimentation and diffusion coefficients

Both the extent to which diffusion is capable of influencing the shape of sedimentation velocity boundaries, as well as the effectiveness of the analysis method of van Holde and Weischet (1978) at overcoming this potential obstacle, are addressed by the data in Fig. 1. For these purposes, we chose to model three different samples that together encompassed a wide but realistic range of sizes and hydrodynamic parameters (see Table 1). To facilitate direct comparison, for each sample simulated sedimentation velocity boundaries are shown under conditions of $\omega^2 t$ that caused each sample to sediment approximately the same distance in the cell (Fig. 1 *A*). From these data it can be seen that the breadth of the boundary was proportional to both the magnitude of the diffusion coefficient and the length of sedimentation. Perhaps most importantly, these results provide a clear indication of how diffusional spreading complicates correct interpretation of a sedimentation velocity experiment, i.e., it was impossible to conclude that only a single ideal species is present in each case simply from visual inspection of the shape of the boundaries.

To remove the contributions of diffusion from the experiments simulated in Fig. 1 *A*, complete sets of boundaries were analyzed by the method of van Holde and Weischet (1978) as described under Materials and Methods. For each boundary fraction, s_{app} was calculated using Eq. 8 and plotted against $t^{-1/2}$ (Fig. 1 *B*). This plot is called an "extrapolation plot" throughout this paper. As predicted by Eq. 8, the width of the "fan" was proportional to the diffusion coefficient (when plotted on the same relative y scale). Furthermore, at each boundary fraction the global linear regressions converged to a single s_{actual} value at the y intercept of the extrapolation plot (Fig. 1 *B*), thereby yielding $G(s)$ plots (f_i versus s_{actual}) that were vertical (Fig. 1 *C*). These results indicate that all three samples were composed of a single ideal species. Importantly, the s_{actual} in each case reproduced the value of s used for the simulation. Thus the data in Fig. 1, *A* and *B*, demonstrated that the analysis method of van Holde and Weischet led to the correct conclusion regarding sample composition and homogeneity, despite the considerable diffusion-dependent differences in boundary shapes.

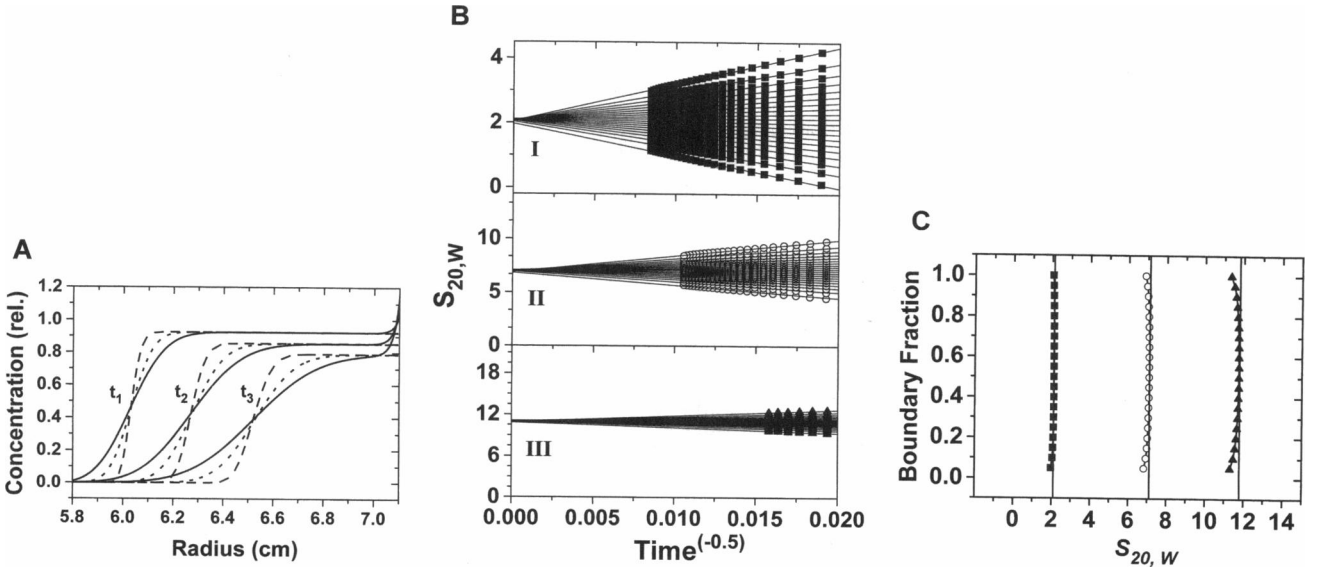

FIGURE 1 Analysis of ideal, single-component samples. (*A*) Simulated boundaries for samples I (——), II (·····), and III (– – –) having the properties described in Table 1. Rotor speeds and sedimentation times were as follows: sample I: 60,000 rpm, $t_1 = 4667$ s, $t_2 = 9333$ s, and $t_3 = 14{,}000$ s. sample II: 48,000 rpm, $t_1 = 3000$ s, $t_2 = 6000$ s, and $t_3 = 9000$ s. sample III: 60,000 rpm, $t_1 = 1333$ s, $t_2 = 2667$ s, and $t_3 = 4000$ s. For clarity, only three of the boundaries used for the extrapolation and $G(s)$ plots are shown. (*B*) Extrapolation plots for samples I (■), II (○), and III (▲). To facilitate comparison, all extrapolations are plotted on the same time scale and on the same relative S scale (~2 times the extrapolated s value). Because of different times used for each sample, not all time points used for the extrapolation are shown in the graph; 25 time points were included for sample I, 30 time points for sample II (22 are shown), and 30 time points for sample III (six shown). In all cases, 20 boundary fractions (f_i) were analyzed. (*C*) $G(s)$ plots for samples I (■), II (○), and III (▲); the vertical lines indicate the input S value used for the simulation.

TABLE 1 Sample properties

Sample	Description	$S_{20,w}$ ($\times 10^{13}$)	$D_{20,w}$ ($\times 10^{7}$)	Shape	M_r(kD)	Axial ratio	v-bar
I	Small protein	2.097	13.1	Sphere	15	1	0.74
II	Medium protein	7.108	6.662	Oblate ellipsoid	100	2	0.74
III	Dimer of (II)	11.79	5.525	Sphere	200	1	0.74
IV	Large protein	15.49	4.826	Sphere	300	1	0.74
V	100-bp DNA	5.662	4.646	Long rod	66	20	0.55
VI	Medium protein	3.328	10.4	Sphere	30	1	0.74
VII	Dimer of (VI)	5.283	8.253	Sphere	60	1	0.74
VIII	Tetramer of (VI)	8.618	6.462	Sphere	120	1	0.74
IX	Hexamer of (VI)	10.99	5.722	Sphere	180	1	0.74
X	Octamer of (VI)	13.31	5.199	Sphere	240	1	0.74
XI	Decamer of (VI)	15.49	4.826	Sphere	300	1	0.74

Single nonideal component system exhibiting concentration dependence of s

Some macromolecules exhibit concentration dependence of s during sedimentation velocity experiments (Johnston and Ogston, 1946). This nonideal behavior can be caused by either steric hindrance or charge interactions between molecules (Eisenberg, 1976) and generally is an indication that the macromolecule is nonglobular. Double-stranded DNA is an example of a molecule whose s is strongly concentration dependent (Demeler, 1992). Boundaries obtained for a simulated linear, double-stranded 100-bp DNA molecule (σ = 0.1) at three different loading concentrations are shown in Fig. 2 *A*. A direct comparison of these boundaries indicates both an asymmetrical boundary sharpening and a reduction in sedimentation velocity at increased sample concentrations. However, it is important to note that the presence of nonideality rarely can be ascertained from examination of the boundary shape per se. In this regard, the analysis method of van Holde and Weischet provides several characteristic diagnostics for the presence of nonideal behavior.

The first diagnostic is associated with unusual convergences in the extrapolation plot. Concentration-dependent samples sediment more slowly at high concentration than at low concentration (Fig. 2 *A*; Johnston and Ogston, 1946; Fujita, 1962; Correia et al., 1976). Hence extrapolations to infinite time originating from the upper boundary fractions (i.e., those corresponding to higher sample concentration) should yield lower s values than extrapolations originating from the lower portion of the boundary. This in turn should cause the regressions to converge at a point to the right of the y axis of the extrapolation plot. As shown in Fig. 2 *B*, precisely such a result was obtained for the representative DNA sample at loading concentrations of 0.55 and 1.0. Furthermore, the intersection point was shifted increasingly to the right with increasing loading concentration (Fig. 2 *B*), consistent with a greater extent of sample nonideality present at the higher concentrations. The second distinctive diagnostic relates to the slope of the $G(s)$ plot. At loading concentrations of 0.55 and 1.0, the $G(s)$ profile of the representative DNA sample had a negative slope, the magnitude of which increased in proportion to the loading concentration (Fig. 2 *C*). Importantly, linear extrapolation of the $G(s)$ profiles to zero concentration (f_i = 0) only approximated the true s_0 value. Finally, the decrease in s measured at the boundary midpoint (i.e., f_i = 0.5), which is known to occur for concentration-dependent systems, also can be seen in Fig. 2 *C*. It should be noted that for any concentration-dependent component in the system, an approximate value for σ can be obtained from experimental data by solving Eq. 4 for σ:

$$\sigma_k = \frac{1 - s_k/s_{0,k}}{C_k}, \tag{9}$$

where the values of s_k, $s_{0,k}$, and C_k (the partial concentration of component k) can be obtained from the distribution plot after converting boundary fractions to absolute concentration. It should be noted that for samples exhibiting significant nonideality, extrapolations of the distribution plot to zero concentration do not yield the input S value. This most likely represents the fact that the correct model for the extrapolation of the distribution plot to zero concentration (i.e., the x axis) has not been identified, and probably does not apply equally well over the entire concentration range. However, the traditional method for checking for concentration dependency by extrapolating the midpoint to zero concentration will result in the true s_0 (Fig. 2 *D*).

The data in Fig. 2 illustrate a situation in which the loading concentration was varied while the extent of sample nonideality, σ, was held constant. Analogous results also have been obtained for simulations in which σ was increased while the loading concentration was held constant, the boundaries became sharper (Fig. 3 *A*), extrapolation plots intersected increasingly to the right of the y axis (Fig. 3 *B*), and the G(s) profiles exhibited increasingly negative slopes with lower midpoint s values (Fig. 3 *C*).

Resolution of multicomponent systems

It has been well established experimentally that the analysis method of van Holde and Weischet provides a powerful tool for studying multicomponent systems (reviewed in Hansen et al., 1994). However, a systematic analysis of the exper-

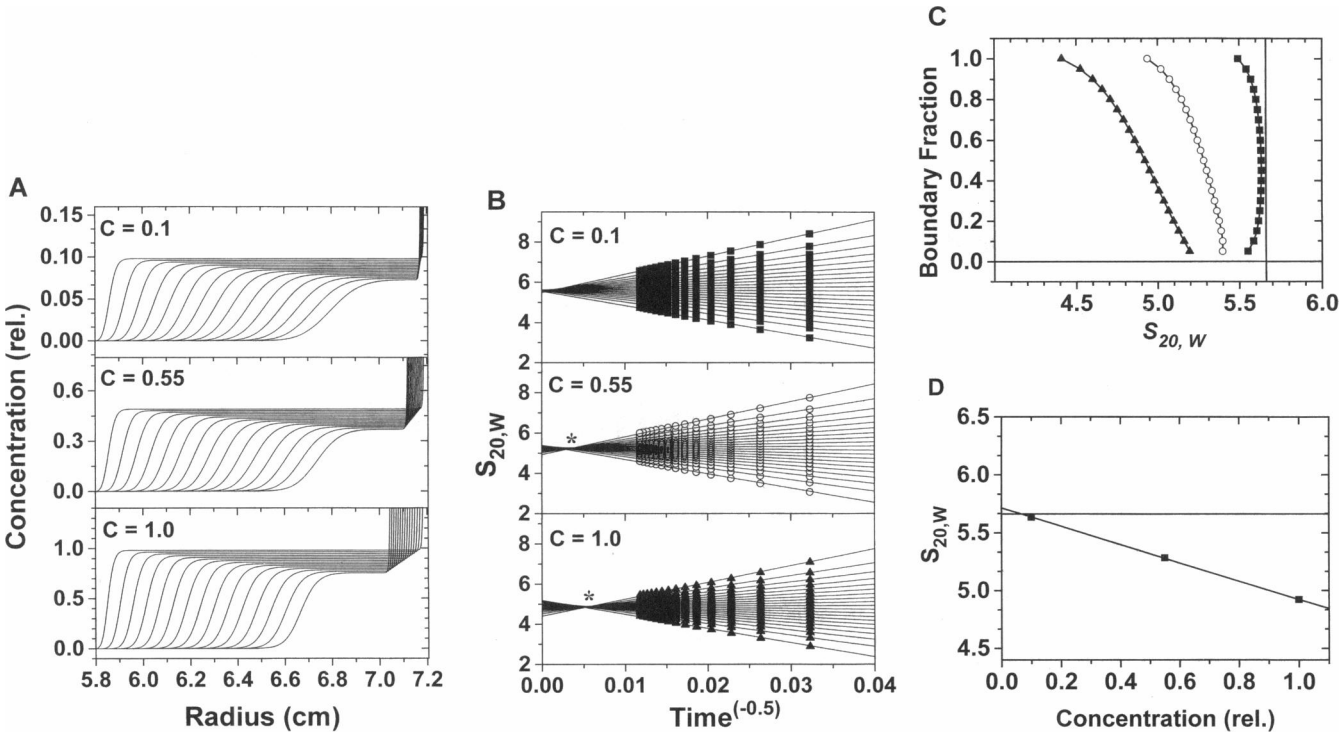

FIGURE 2 Analysis of concentration dependence of s. (*A*) Simulated boundaries for sample V (see Table 1). Rotor speed: 60,000 rpm; sedimentation time: 7200 s; $\sigma = 0.1$. Relative loading concentration C_L: 0.1 (*top*), 0.55 (*center*), and 1.0 (*bottom*). (*B*) Extrapolation plots for sample V: $C_L = 0.1$ (■), $C_L = 0.55$ (○), and $C_L = 1.0$ (▲). The asterisk marks the intersection point of the extrapolation. Note that this point is shifted increasingly to the right with increased loading concentration. (*C*) $G(s)$ plots from *B*: $C_L = 0.1$ (■), $C_L = 0.55$ (○), and $C_L = 1.0$ (▲). The vertical line indicates the input S value used for the simulation. (*D*) ■, Midpoint extrapolation to zero concentration. The horizontal line indicates the input S value used for the simulation. The extrapolation and input S value agree well.

imental conditions under which this method is capable of resolving multiple components has not been reported. Here we focus on issues relating to sedimentation time, rotor speed, and boundary shape. Fig. 4 *A* shows the simulated boundaries obtained after sedimentation of an ideal two-component system composed of equal amounts of a 3.3S and 7.1S species for either 1, 2, or 3 h at 42,000 rpm. A single broad boundary existed after 1 h of sedimentation, whereas two discrete boundaries were present after both 2 and 3 h of sedimentation. However, despite the significant differences in boundary shape, the extrapolation (Fig. 4 *B*) and $G(s)$ plots (Fig. 4 *C*) derived from all three data sets clearly indicated the presence of two distinct species sedimenting at ~3S and ~7S. The only significant difference occurred in the transition region of the $G(s)$ plot ($f_i = 0.5$ in this example); at shorter sedimentation times the region between the resolved 3S and 7S species was less distinct and was spread over a wider range of the $G(s)$ plot than was observed at longer times.

We next simulated the sedimentation of the same two-component system using different rotor speeds but keeping $\omega^2 t$ approximately constant. Conditions were chosen such that the combination of highest speed and shortest sedimentation time produced two clearly defined boundaries, whereas a longer run at a lower speed produced a set of single broad boundaries (Fig. 5 *A*). The shape of the extrapolation and $G(s)$ plots followed the same trend as observed in Fig. 4; in all three cases the boundaries were resolved into distinct species sedimenting around 3S and 7S, albeit with a less distinct transition region in the $G(s)$ plot derived from the set of single broad boundaries formed during the slow speed/long duration experiment.

To determine whether the results shown in Figs. 4 and 5 are also relevant to multicomponent systems that are unable to separate into discrete boundaries, we simulated the sedimentation of an ideal two-component system composed of equal amounts of 2.1S and 3.3S species. In this case, even the optimal combination of sedimentation time/rotor speed produced only single broad boundaries (data not shown). The $G(s)$ plots obtained after sedimentation of this mixture for 1, 2, or 3 h at 60,000 rpm are shown in Fig. 6. Although there is a very small systematic difference centered around $f_i = 0.5$, for all practical purposes all three profiles are identical. The nonvertical $G(s)$ profiles rigorously demonstrate the presence of heterogeneity in s, regardless of sedimentation length. Furthermore, they provide a fairly accurate estimate of the minimum (2.0S at $f_i = 0.05$) and maximum (3.1S at $f_i = 1.0$) sedimentation coefficient in the sample. However, it is important to note that because of the broadness of the profiles, the number and amount of species present cannot be determined quantitatively. Simulations of samples with a combination of larger S values and smaller

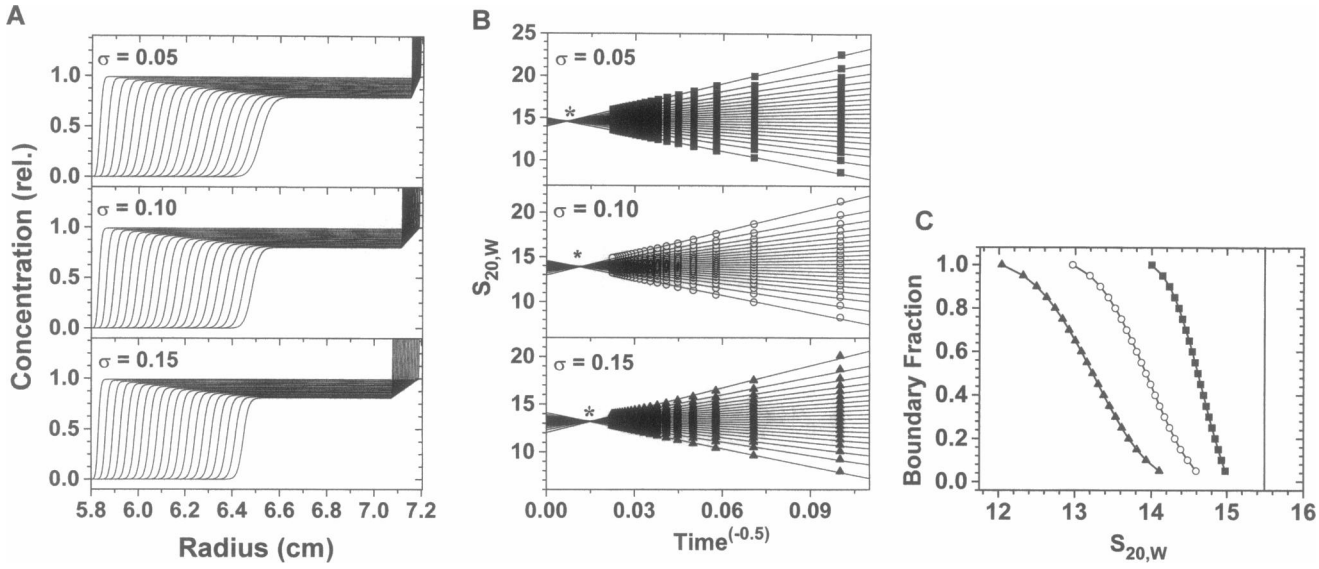

FIGURE 3 Analysis of concentration dependence factor σ. (*A*) Simulated boundaries for sample IV (see Table 1). Rotor speed: 60,000 rpm; sedimentation time: 2000 s; relative loading concentration $C_L = 1.0$, $\sigma = 0.05$ (*top*), $\sigma = 0.1$ (*center*), $\sigma = 0.15$ (*bottom*). (*B*) Extrapolation plots for sample IV: $\sigma = 0.05$ (■), $\sigma =$ (○), and $\sigma = 0.15$ (▲). The asterisk marks the intersection point of the extrapolation. Note that this point is shifted increasingly to the right with increasing σ. (*C*) $G(s)$ plots from *B*: with $\sigma = 0.05$ (■), $\sigma = 0.10$ (○), and $\sigma = 0.15$ (▲). The vertical line indicates the input S value used for the simulation.

diffusion coefficients (e.g., sample 1 with 43.1S and $D = 2.89 \times 10^{-7}$, and sample 2 with 45.2S and $D = 2.82 \times 10^{-7}$) have been performed; under these conditions, heterogeneous samples with a difference in s of as little as 5% can be distinguished (data not shown). Analogous conclusions were originally reported by van Holde et al. (1978).

We frequently encounter the experimental situation in which a sample consists of one major species, as well as a small amount of faster sedimenting material, e.g., stable oligomeric aggregates, large contaminants. Under these conditions there is only a short window of time, occurring relatively early in the run, in which the boundaries are suitable for analysis by the $G(s)$ method, i.e., the boundaries have sufficiently cleared the meniscus and yet the plateau concentration of the scans has not been reduced by sample pelleting. However, in many instances the presence of the

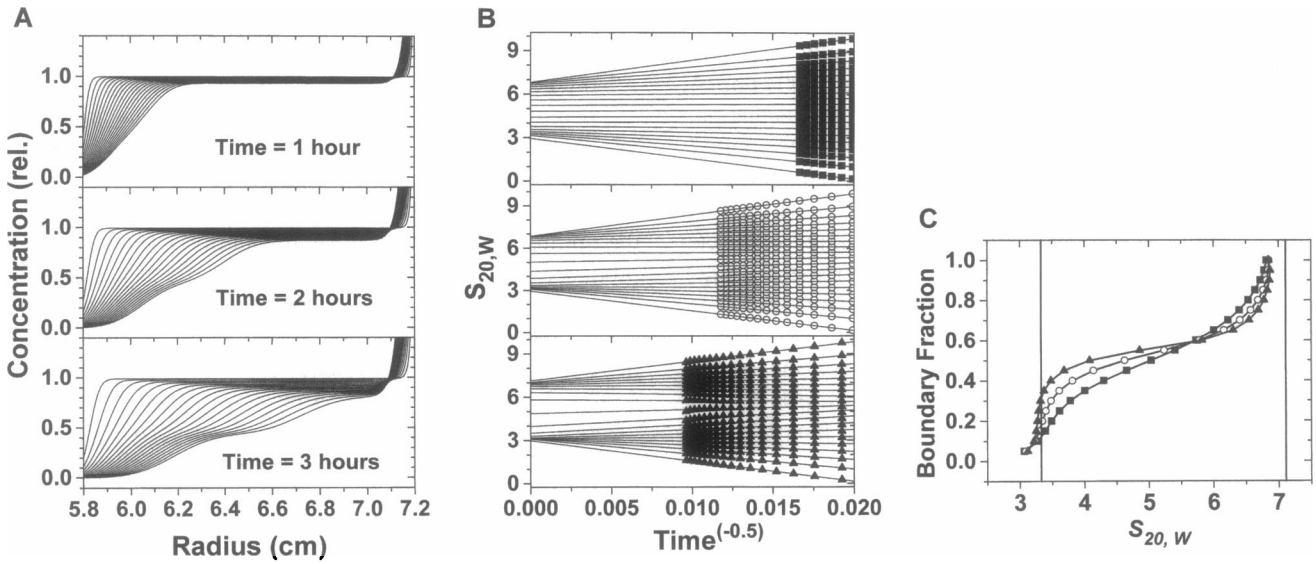

FIGURE 4 Effect of sedimentation time on resolution of multicomponent system at constant speed. (*A*) Simulated boundaries for a 50:50 mixture of samples II and VI (see Table 1). Rotor speed: 45,000 rpm; sedimentation time: (*top*) 1 h; (*center*) 2 h; (*bottom*) 3 h. (*B*) Extrapolation plot for boundaries shown in *A*. Sedimentation times: 1 h (*top*, ■), 2 h (*center*, ○), and 3 h (*bottom*, ▲). (*C*) $G(s)$ plots: 1 h (■), 2 h (○), and 3 h (▲). The vertical lines indicate the input S value used for the simulation. Note that all three panels show best resolution under conditions of transport to the bottom of the cell.

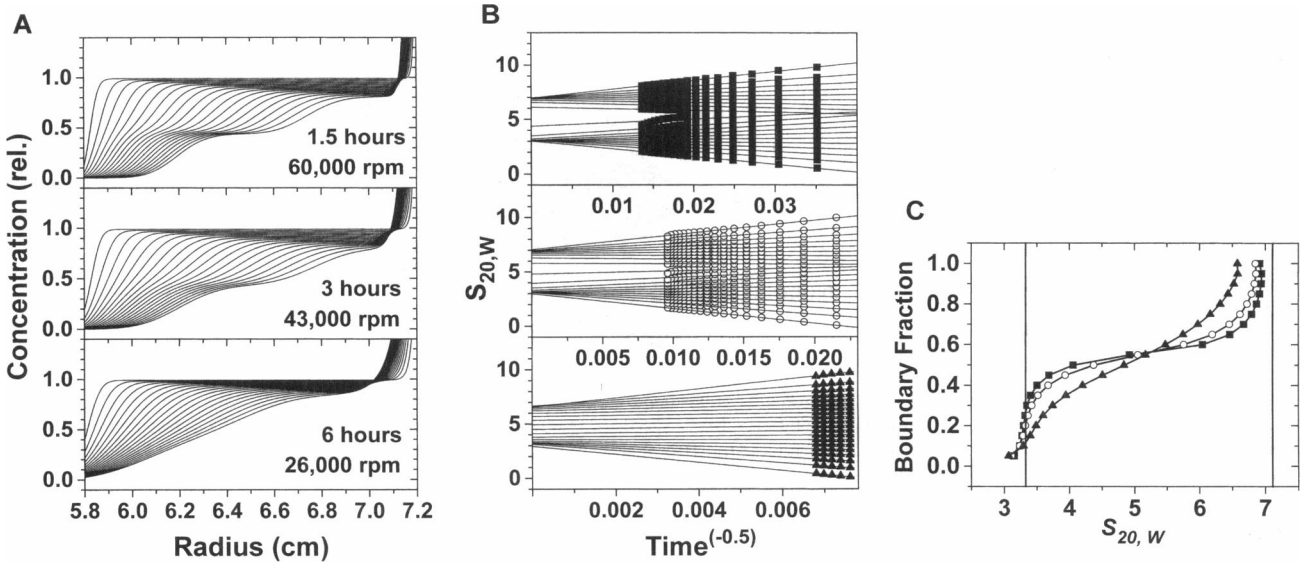

FIGURE 5 Resolution of multicomponent mixtures at approximately constant $\omega^2 t$. (*A*) Simulated boundaries for a 50:50 mixture of samples II and VI (see Table 1). (*Top*) Rotor speed: 60,000 rpm; sedimentation time: 1.5 h. (*Center*) Rotor speed: 43,000 rpm; sedimentation time: 3 h. (*Bottom*) Rotor speed: 26,000 rpm; sedimentation time: 6 h. (*C*) $G(s)$ plots of data shown in *A*: rotor speed: 60,000 rpm; sedimentation time: 1.5 h (■). Rotor speed: 43,000 rpm; sedimentation time: 3 h (○). Rotor speed: 26,000 rpm; sedimentation time: 6 h (trif]). The vertical lines indicate the input S value used for the simulation. Note that all three panels show best resolution at the highest speed and shortest sedimentation time.

faster sedimenting material cannot be visualized against the background of boundary and plateau noise, and some of the later scans are inappropriately included in the analysis. This type of situation is illustrated in Fig. 7. In this example, we simulated boundaries in which ~90% of the sample sedimented at 3.3S, ~5% at sedimented at 5.3S, and the remaining ~5% sedimented between 8.6 and 15.5S (Fig. 7 *A*). Extrapolation and $G(s)$ plots are shown in Fig. 7, *B* and *C*, respectively. The negatively sloping region present at the top of the $G(s)$ plot (Fig. 7 *C*) was reminiscent of that expected for a nonideal sample (Figs. 2 *C* and 3 *C*). However, in this case the result stemmed from the inclusion of scans in the analysis in which some of the larger components had pelleted. A comparison of Figs. 2 *B* and 7 *B* indicated that the distinction between nonideality and sample pelleting can be made on the basis of the data in the upper region of the extrapolation plot; the s_{app} decreased continuously versus $t^{-1/2}$ when the sample was nonideal, but discontinuously when sample pelleting was involved. The discontinuity results from the fact that once a portion of the sample pellets, the boundary fractions are no longer equivalent throughout the experiment and hence the s_{app} are incorrectly aligned in the extrapolation plot. Importantly, when one analyzes only those boundaries collected before sample pelleting (i.e., before the discontinuity), the extrapolations lead to the correct conclusion that ~90% of the sample consists of a homogeneous 3S species, whereas ~10% of the sample sediments much faster (Fig. 7 *B*, *dashed lines*).

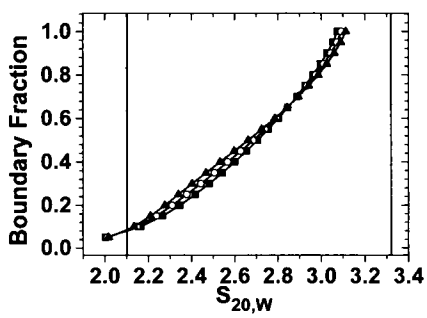

FIGURE 6 Limits of resolution of a multicomponent mixture. Shown is the $G(s)$ plot for a 50:50 mixture of samples I and VI (see Table 1). Rotor speed: 60,000 rpm; loading concentration, C_1 = 1.0; sedimentation times are 1 h (■), 2 h (○), and 3 h (▲). The vertical lines indicate the S value used for the simulation. Note that both minimum and maximum s values are accurately resolved, but the number of components and their relative amounts cannot be determined.

Reversibly associating monomer-dimer system

Many biological systems analyzed in the ultracentrifuge are capable of reversible self-association. In such cases, the distribution of monomeric and oligomeric species present during sedimentation will be influenced by the sample concentration, the rates of association/dissociation, and the magnitude of the equilibrium constant(s) (see Eq. 5). We have simulated the effects of each of these parameters on the sedimentation of the simplest associating system, a monomer-dimer equilibrium. The $G(s)$ plots obtained after analysis of the simulated data by the method of van Holde and Weischet are shown in Fig. 8. For the case in which the sample concentration and rate constants were held constant,

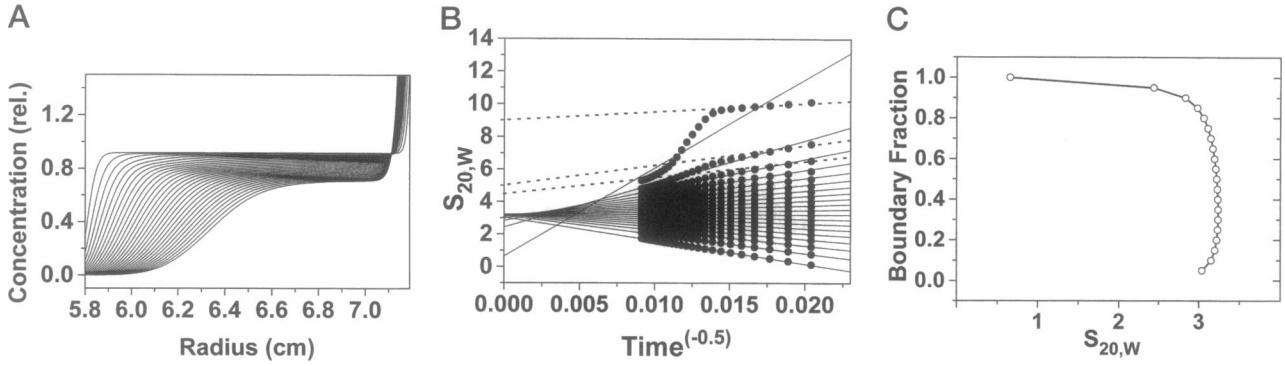

FIGURE 7 Analysis of a complex six-component system. The C_L of each component were: VI, 0.8; VII, 0.05; VIII, 0.025; IX, 0.0125; X, 0.0125; XI, 0.0125. Samples are described in Table 1. (*A*) Simulated boundaries. Rotor speed: 45,000 rpm; sedimentation time: 12,000 s. (*B*) Extrapolation plot of the data shown in *A*. The dashed lines show the extrapolations for the top 10% that would have been observed had the later scans been excluded from the analysis. (*C*) $G(s)$ plot of extrapolation plot in *B*.

a small equilibrium constant caused almost the entire sample to remain in the monomeric 7S form, whereas a large equilibrium constant shifted the distribution mostly to the dimerized 11S form. However, an intermediate equilibrium constant resulted in a set of single unresolved boundaries and a narrow, smoothly increasing plot that essentially was centered between the monomer/dimer limits (Fig. 8 *A*). A similar result was obtained when only the rate constants were changed (Fig. 8 *B*). Finally, under conditions where the combination of equilibrium and rate constants produced intermediate smoothly increasing $G(s)$ profiles, decreases in sample concentration shifted the profile toward the sedimentation coefficient of the monomeric species without changing the overall shape of the plot (Fig. 8 *C*). The behavior shown in Fig. 8 is in marked contrast to that observed for a mixture of noninteracting 7S and 11S components with partial concentrations equivalent to conditions shown in Fig. 8 *B*, which separated into two distinct boundaries and resulted in $G(s)$ plots similar to those in Figs. 4 *C* and 5 *C* at all sample concentrations (Fig. 8 *D*).

It is important to note that the data in Fig. 8 have been modeled to demonstrate general trends. In actual experiments, the conditions necessary to mimic the specific results shown above will be different for each system analyzed, depending on the monomer molecular weight, the extinction coefficient, the sedimentation and diffusion coefficients of both the monomer and dimer, the magnitude of the equilibrium and rate constants for the association reaction, and the

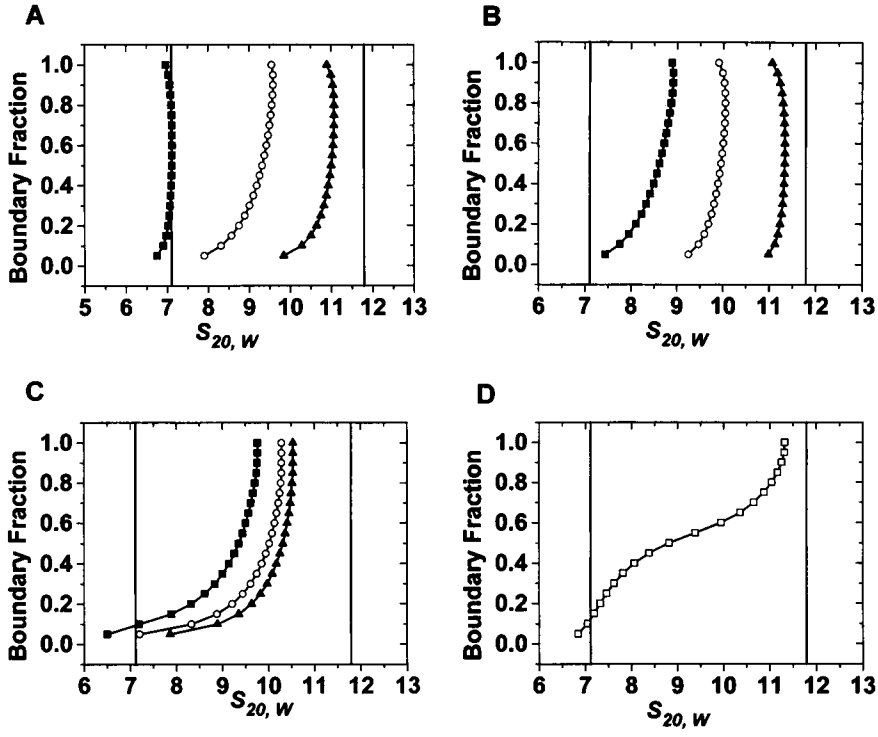

FIGURE 8 Analysis of a reversibly associating system. Shown are $G(s)$ plots for monomer-dimer equilibria involving samples II and III (see Table 1); the vertical lines indicate the input S value used for the simulation. Rotor speed: 45,000 rpm; sedimentation time: 12,000 s. (*A*) Effect of a variable equilibrium constant K on $G(s)$ distribution. Loading concentration C_L = 1.0; forward rate constant k_1 = 1.0; equilibrium constant K: 0.01 (■), 5.0 (○), and 100.0 (▲). (*B*) Effect of forward rate constant k_1 on $G(s)$ distributions. Loading concentration, C_L = 1.0; equilibrium constant, K = 2.0; forward rate constant k_1: 1.0 (■), 2.0 (○), and 10.0 (▲). (*C*) Effect of variable loading concentration on $G(s)$ distributions. Equilibrium constant K = 100; forward rate constant k_1 = 0.2; C_L: 0.5 (■), 1.0 (○), and 1.5 (▲). The vertical lines indicate the input S value used for the simulation. All equilibrium, rate constants, and loading concentrations are in relative concentration units proportional to optical density.

loading concentration used during sedimentation. It should also be noted that in practice it is not feasible to extract the magnitude of the equilibrium or rate constants from the van Holde-Weischet analysis; rather, the method serves as a valuable diagnostic for determining whether the system exhibits reversible association.

Simulated experimental noise

The method's sensitivity toward experimental noise is shown in Fig. 9. Even at levels of noise that exceed those routinely encountered experimentally with the absorption optical system, the results are not in any way changed. The generality of this conclusion has been verified by repeating the other simulations presented in this paper with 5% of random noise added (data not shown). It should be noted that this question will have to be reevaluated in the future for the interference optical system of the XL-I ultracentrifuge.

DISCUSSION

Analytical ultracentrifugation is a powerful technique for characterizing macromolecules and their interactions (Schachman, 1992; Lee and Rajendran, 1994; Hansen et al., 1994; Hensley, 1996). Yet, for almost 20 years it has been increasingly rare to encounter a biological system that is under active study by analytical ultracentrifugation. With the recent availability of the new generation analytical ultracentrifuges, many investigators again are incorporating this technology into their ongoing investigations of macromolecular structure/function relationships. However, because of its long absence, many experimental samples are being studied in the analytical ultracentrifuge for the first time; in these cases investigators must design their ultracentrifuge strategy in the absence of a data base of existing knowledge. This situation is further complicated by the fact that more commonly used techniques such as gel filtration, light scattering, and native gel electrophoresis often fail to identify existing sample complexity, thereby leading to an oversimplified view of the solution-state behavior of the system. Thus when a system is analyzed in the analytical ultracentrifuge for the first time, the initial experiments should focus on determination of basic fundamentals relating to sample composition and behavior, i.e., is the sample homogeneous or heterogeneous, ideal or nonideal, interacting or noninteracting? The data presented in this paper demonstrate that when analyzed by the method of van Holde and Weischet (1978), as few as two or three sedimentation velocity experiments can provide virtually all of the information required to delineate the fundamental solution-state behavior of both simple and complex macromolecular systems. These results significantly advance the utility of the analytical ultracentrifuge as a general-purpose research tool.

Our results demonstrate that the analysis method of van Holde and Weischet provides rigorous diagnostics for sample homogeneity and nonideal solution behavior, even under situations in which moderate noise is present in the data (Fig. 9). Furthermore, this method allows one to determine whether samples exhibiting heterogeneity in s are noninteracting or reversibly interacting systems. Importantly, such information can be obtained simply by examining the extrapolation and $G(s)$ plots obtained over a 5–10-fold range in loading concentration, which generally involves two or three sedimentation velocity experiments performed at one or two different wavelengths. (The monochromator of the XL-A analytical ultracentrifuge has several strong emission maxima (e.g., 230 nm) that usually allow low concentrations of macromolecules to be studied with an acceptable signal-to-noise ratio.) Characteristic diagnostics are as follows: $G(s)$ plots that are vertical at all loading concentrations are indicative of an ideal single-component system

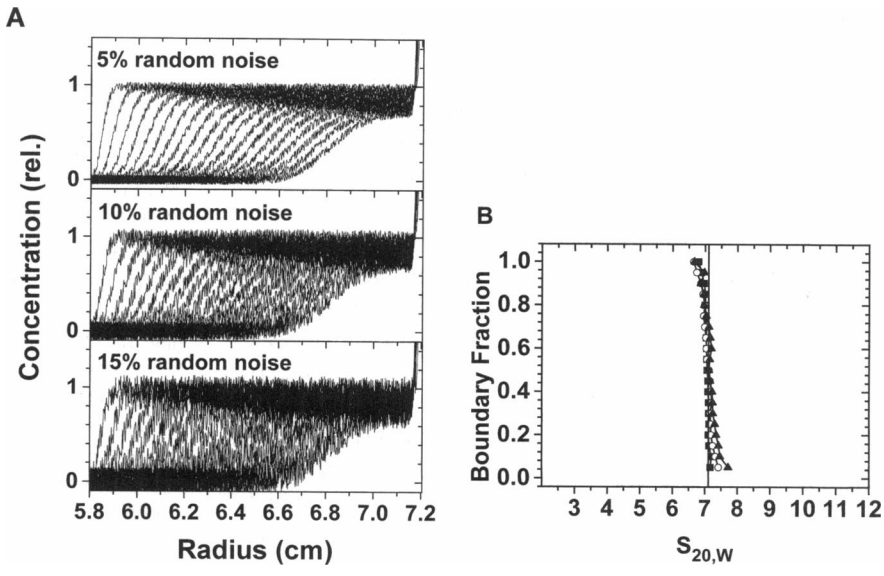

FIGURE 9 Analysis of the effect of experimental noise. (*A*) Simulated velocity data for sample II (see Table 1). (*Top*) Five percent random noise; (*center*) 10% random noise; (*bottom*) 15% random noise. Noise levels are given in percentage of total OD. (*B*) $G(s)$ plots for data in *A*: 5% random noise and 10 points/frame smoothing average (■); 10% random noise and 15 points/frame smoothing average (○); 15% random noise and 20 points/frame smoothing average (▲). The vertical line indicates the input S value used in the simulation. Note that all represented noise levels produce roughly the same result.

(see Fig. 1). (Some distributions shown in the figures may appear not to be truly vertical, and have a very small curvature. This is the result of slight errors introduced by the approximation in the discrete solution. This error can be effectively eliminated by using very short time and space steps. However, computation time becomes prohibitive, at least on the 486 PC used for these studies. The step sizes chosen for these studies keep this error small enough so that in no case presented are the results affected.) $G(s)$ plots in which the slope is negative at higher loading concentrations but becomes less negative at lower concentration are indicative of nonideal samples exhibiting the concentration dependence of s. Furthermore, for a nonideal sample, an increase in s across the entire distribution will accompany the decrease in negative slope at low loading concentration (Fig. 2 *C*). Trends in the slope of the $G(s)$ plot similar to those observed in the nonideal case may be observed for an ideal multicomponent sample in which a small amount of rapidly sedimenting material pellets before the end of the run; in this case the discontinuities in the upper region of the extrapolation plot both prevent misinterpretation of the $G(s)$ plots as nonideality and indicate which scans are appropriate to use during reanalysis of the data (Fig. 7 *B*). In addition, in the case of pelleting of noninteracting systems, there should be no concentration dependence of s across the $G(s)$ profile. Samples whose $G(s)$ profiles are superimposable at all sample concentrations and have either multiple distinct components or broad positive slopes, consist of two or more ideal components (see Figs. 4–6, 8). Interpretation of the $G(s)$ plots of such systems is discussed further below. A situation in which both the positive slope and the s values of the $G(s)$ profile increase with decreasing loading concentration is indicative of the presence of multiple nonideal components (data not shown). Thus, a multicomponent sample that appears homogeneous at high loading concentration due to compensating nonideality and heterogeneity in s will be correctly revealed as heterogeneous at low loading concentration. Simple self-associating systems (e.g., monomer-dimer, monomer-trimer) are identified by relatively narrow, smoothly increasing $G(s)$ distributions that are shifted to lower s values at lower sample concentrations, although such systems may sediment equally well as either stable monomers or oligomers (and thereby exhibit vertical $G(s)$ profiles) under the specific conditions used in sedimentation velocity experiments (Fig. 8). Finally, systems that undergo isodesmic association into large polymeric aggregates will be susceptible to the problems caused by sample pelleting, but will exhibit $G(s)$ profiles that are shifted to lower s at lower loading concentrations. Note that once self-association is suspected, one should use Gilbert theory (Lee and Rajendran, 1994) or sedimentation equilibrium methods (Johnson et al., 1981) to rigorously analyze the system. A tutorial for the analysis and interpretation of experimental sedimentation velocity data using the method of van Holde and Weischet (1975) is available over the Internet at http://bioc09.uthscsa.edu.biochem/tutor.html.

The results obtained in Figs. 4–6 also provide important insight into the interpretation of the $G(s)$ profiles of multicomponent systems. Not surprisingly, boundaries that separate into discrete components during sedimentation produce the most highly resolved $G(s)$ plots (Figs. 4 and 5). Under these conditions, the $G(s)$ profiles provide information regarding the homogeneity of the different boundary components. However, as often as not, multicomponent samples do not resolve into distinct boundaries during sedimentation. Unresolved boundaries tend to be encountered when two components have similar sedimentation coefficients (Fig. 6), when more than two components are present in the sample (Hansen et al., 1994), or when a multicomponent sample cannot be subjected to optimal sedimentation conditions because of the presence of faster sedimenting material (see Figs. 4, 5, 7). Each of these situations is frequently encountered experimentally. Importantly, in all three cases, analysis of the unresolved boundaries yields positive sloping $G(s)$ plots in which s increases gradually across the boundary fraction range (Fig. 6; Schwarz and Hansen, 1994; Hansen et al., 1994). Thus, a broad positively sloping $G(s)$ plot both provides a rigorous indication of sample heterogeneity and yields the maximum and minimum sedimentation coefficients present in the sample, but does not indicate the number of components present or their amounts.

This work was supported by National Institutes of Health grant GM45916 (to JCH) and by the University of Texas Health Science Center at San Antonio Department of Biochemistry.

REFERENCES

Cantor, C. R., and P. R. Schimmel. 1980. Biophysical Chemistry, Part II. Freeman, San Francisco.

Claverie, J.-M. 1976. Sedimentation of generalized systems of interacting particles. III. Concentration dependent sedimentation and extension to other transport methods. *Biopolymers.* 15:843–857.

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

Correia, J. J., M. L. Johnson, G. H. Weiss, and D. A. Yphantis. 1976. Numerical studies of the Johnston-Ogston effect in two-component systems. *Biophys. Chem.* 76:255–264.

Demeler, B. 1992. Ph.D. thesis. Oregon State University, Corvallis, Oregon.

Eisenberg, H. 1976. Sedimentation in the ultracentrifuge and diffusion of macromolecules carrying electrical charges. *Biophys. Chem.* 5:243–251.

Fujita, H. 1962. Mathematical Theory of Sedimentation Analysis. Academic Press, New York. 144.

Hansen, J. C., J. Lebowitz, and B. Demeler. 1994. Analytical ultracentrifugation of complex macromolecular systems. *Biochemistry.* 33: 13155–13163.

Hensley, P. 1996. Defining the structure and stability of macromolecular assemblies in solution: the re-emergence of analytical ultracentrifugation as a practical tool. *Structure.* 4:367–373.

Johnson, M. L., J. J. Correia, D. A. Yphantis, and H. R. Halvorson. 1981.

Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* 36:575–588.

Johnston, J. P., and A. G. Ogston. 1946. A boundary anomaly found in the ultracentrifugal sedimentation of mixtures. *Trans. Faraday Soc.* 42: 789–799.

Lee, J. C., and S. Rajendran. 1994. Studies of macromolecular interaction by sedimentation velocity. *In* Modern Analytical Ultracentrifugation. T. M. Schuster and T. Laue, editors. Birkhäuser, Berlin. 138–155.

Schachman, H. K. 1992. *In* Analytical Ultracentrifugation in Biochemistry and Polymer Science. S. E. Harding, A. J. Rowe, and J. C. Horton, editors. Royal Society for Chemistry, Cambridge, England. 3–15.

Schwarz, P. M., and J. Hansen. 1994. Formation and stability of higher order chromatin structures: contributions of the histone octamer. *J. Biol. Chem.* 269:16284–16289.

Todd, G. P., and R. H. Haschemeyer. 1981. General solution to the inverse problem of the differential equation of the ultracentrifuge. *Proc. Natl. Acad. Sci. USA.* 78–11:6739–6743.

van Holde, K. E. 1985. Physical Biochemistry, 2nd Ed. Prentice-Hall, Englewood Cliffs, NJ.

van Holde, K. E., and W. O. Weischet. 1975. Boundary analysis of sedimentation velocity experiments with monodisperse and paucidisperse solutes. *Biopolymers.* 17:1387–1403.

Zienkiewicz, O. C. 1971. The Finite Element Method in Engineering Science. McGraw-Hill, London.