

Sedimentation velocity analysis of highly heterogeneous systems[☆]

Borries Demeler^{a,*}, Kensal E. van Holde^b

^a Department of Biochemistry, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA

^b Department of Biochemistry and Biophysics, Oregon State University, Corvallis, OR 97331, USA

Received 22 July 2004

Available online 28 October 2004

Abstract

This article discusses several improvements to the van Holde–Weischet (vHW) method [Biopolymers 17 (1978) 1387] that address its capability to deal with sedimentation coefficient distributions spanning a large range of s values. The method presented here allows the inclusion of scans early and late in the experiment that ordinarily would need to be excluded from the analysis due to ultracentrifuge cell end effects. Scans late in the experiment are compromised by the loss of a defined plateau region and by back-diffusion from the bottom of the cell. Early scans involve partial boundaries that have not fully cleared the meniscus. In addition, a major refinement of the algorithm for determining the boundary fractions is introduced, taking into account different degrees of radial dilution for different species in the system. The method retains its desirable model-independent properties (the analysis of sedimentation data does not require prior knowledge of a user-imposed model or range of sedimentation coefficients) and reports diffusion-corrected s value distributions, which can be presented either in a histogram format or the traditional integral distribution format. Data analyzed with the traditional vHW method are compared with those of the improved method to demonstrate the benefit from the added information in the analysis.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Sedimentation; Analytical ultracentrifugation; s value distributions; Composition analysis

The method of van Holde and Weischet (vHW)¹ [1] is established throughout the literature as a versatile, model-independent approach to determine sedimentation coefficient distributions by graphically analyzing sedimentation velocity experimental data. The vHW method's strength derives from its ability to differentiate boundary spreading due to diffusion from boundary spreading due to heterogeneity in the sedimentation coefficient. The basis for this ability lies in the fact that sedimentation is a transport process in which flow is proportional to the first power of time, whereas diffusion flow is proportional to the square

root of time. In an extrapolation of apparent sedimentation coefficients corresponding to specific concentration points in the boundary to infinite time, the diffusion contribution to boundary spreading is minimized because as $t \rightarrow \infty$, transport due to sedimentation outweighs transport due to diffusion. The vHW method can be used to reliably distinguish heterogeneous systems from single-species systems as well as self-associating systems from noninteracting systems. It can also be used to identify concentration-dependent nonideality, approximately quantify the partial concentration of different components, and obtain qualitative information about diffusion coefficients. A discussion of applications for the vHW method can be found in [2]. We briefly reproduce here the algorithm for the graphical transformation of sedimentation velocity data according to the original vHW method:

[☆] Supplementary data for this article, including the citations for Refs. [10–12], may be found on ScienceDirect.

* Corresponding author. Fax: +1 210 567 1136.

E-mail address: demeler@biochem.uthscsa.edu (B. Demeler).

¹ Abbreviations used: vHW, van Holde and Weischet.

1. Identify a baseline concentration (the signal can be in fringes or absorbance units, depending on the optical system). This information is customarily obtained from the region near the meniscus in the final scan in the experiment, where most of the material has been sedimented to the bottom of the cell.
2. Discard scans that have not cleared the meniscus and scans for which a stable plateau concentration cannot be identified.
3. For each scan i , identify the plateau concentration.
4. Subdivide the distance between plateau and baseline concentrations of each scan into n equally spaced boundary fractions. Because of radial dilution, this distance decreases with later scans; hence, the spacing of the boundary fractions decreases as well for later scans.
5. Find the intercept for each boundary fraction j with each scan i . Use the radial distance of each intercept and the time of each scan to calculate an apparent sedimentation coefficient, $s_{i,j}^*$, according to the following equation:

$$s_{i,j}^* = \ln \left[\frac{r_j(t_i)}{r_a(t_0)} \right] [\omega^2(t_i - t_0)]^{-1},$$

$$1 \geq i \geq m, \quad 1 \geq j \geq n, \quad (1)$$

where r_a is the radius of the meniscus, r_j is the radius at the intercept of the scan and boundary fraction j , m is the number of scans that have cleared the meniscus and exhibit a stable plateau concentration, ω is the angular velocity, t_i is the time at which scan i was taken, and t_0 is the theoretical start time of the experiment that would have been observed if the centrifuge could have been started at full rotor speed. It is commonly obtained by extrapolating the $\omega^2 t$ integral values recorded by the ultracentrifuge to zero time.

6. Plot all $s_{i,j}^*$ against the inverse square root of the time of each scan i .
7. Extrapolate $s_{i,j}^*$ values from corresponding boundary fractions j from all scans to infinite time (zero on the inverse time scale). The intercepts with the y axis then represent the diffusion-corrected sedimentation coefficients for the components sedimenting at a particular boundary fraction in the moving boundary.
8. Plot the boundary fractions versus the y axis intercepts (extrapolated s values) to obtain an integral s value distribution $G(s)$.

There are a number of problems encountered in attempting to apply the vHW method to the analysis of samples having a wide range of sedimentation coefficients. First, the method does not properly take into account the different radial dilutions experienced by components with very disparate sedimentation coefficients. Consequences of this include: (i) distortions in the $G(s)$ graph in regions where s changes rapidly

(e.g., between two markedly different components) and (ii) uncertainty in the determination of the partial concentration in cases of several components. Other problems arise from the theoretical basis of the method. Unlike numerical or approximate solutions to the Lamm equation [3–5] that incorporate the boundary conditions into the solution, the vHW method is based on the Faxén solution, which assumes an infinitely long cell. Because of this restriction, the boundary effects caused by the meniscus and the bottom of the cell cannot be modeled accurately by the vHW method, and regions near the meniscus and the bottom of the cell need to be excluded from the analysis. Effects caused by the physical limitation of the cell near the bottom include piling up of material during sedimentation. This generates a steep concentration gradient at the bottom of the cell, which in turn causes diffusion back into the cell. The back-diffusion distorts the plateau region near the cell bottom, such that it becomes unsuitable for the measurement of a plateau value in step 2 of the algorithm. Therefore, scans whose plateau concentration is significantly affected by back-diffusion need to be excluded from the analysis. In addition, later scans will eventually lose material due to pelleting, which will further interfere with the measurement of a reliable plateau value required for the determination of the boundary fractions. At the other extreme, slowly sedimenting material will require significant time to clear the meniscus, and only late scans will contain concentration points for all boundary fractions. For heterogeneous systems with a large range of s values, the problem is compounded by the fact that when rapidly sedimenting material already pellets at the bottom of the cell, slowly sedimenting material has barely begun to clear the meniscus and only a few scans both will have a stable plateau and will have cleared the meniscus. In this article, we present modifications to the vHW method that address the problems described above so as to allow the analysis of sedimentation velocity data from samples with a wide range of molecular size with the vHW method. We test these modifications both by numerical simulation and by experiments.

Materials and methods

Both simulated and experimental data were used to test the modified algorithm. All simulations were performed with the finite element solution implemented in UltraScan [3,6] using 0.001 cm radial resolution and 10 s time resolution. Simulations were performed on an Athlon PC running Slackware Linux 9.0. Sedimentation velocity experiments were performed on a Beckman Optima XL-I using UV absorbance optics at 260 nm. The samples used for the velocity experiments were restriction digests of the pPOL-1 208-12 plasmid [7].

Sample 1 was generated by digesting with *EcoRI* (Pro-mega) to completion according to the manufacturer's recommendations. Sample 1 was centrifuged for 210 min at 50,000 rpm at 20 °C in Tris buffer with 150 mM NaCl. Sedimentation was performed in two-channel charcoal/epon centerpieces using the AN-60 rotor. Sample 2 was designed to contain a wide range of DNA fragments. It was derived from a portion of the *EcoRI* digest obtained in sample 1 that was first repurified with a phenol extraction and ethanol precipitation. The DNA was redissolved in enzyme buffer and split into two equal volumes. The first part was then subjected to digestion with *BanII* (NEB), and the second part was digested with *PvuI* (Invitrogen), according to the manufacturers' recommendations. All digestions were allowed to go to completion, which was verified using a 1% agarose gel (data not shown). Equal portions of the *PvuI* digest and the *BanII* digest were then mixed. Thus, the final mixture contained the fragments from digestions by all three enzymes. Because *PvuI* cuts only in the "plasmid" portion of the plasmid and *BanII* cuts only in the inserted repeats, all of the *EcoRI* fragments are also present in the final mix. Predicted fragments, fragment lengths, their relative abundance, and their predicted relative concentrations are shown in Table 1. Sample 2 was sedimented in 20 mM Tris buffer, pH 8.0, containing 150 mM NaCl, at 20 °C, 60,000 rpm, in a two-channel aluminum centerpiece for 120 min. For both samples, absorbance scans at 260 nm were taken at 0.003 cm resolution with no delay between scans.

Enhanced algorithm for sedimentation analysis: experimental design considerations

Together with the analysis enhancements described below, a number of experimental design considerations should be observed:

1. Maximize the solution column length.
2. Acquire scans as soon as the final rotor speed has been reached and continue to scan as rapidly as possible, with minimal delay between scans so as to maximize the usable data.
3. Collect scans until most of the material is pelleted.
4. Select a speed just fast enough to collect at least 30–40 scans. However, the speed should not be too low because lower speeds sacrifice resolution in sedimentation coefficient [2].

Description of the enhanced algorithm

Plateau concentration estimation

With the availability of a data set collected as described above, the first step is to estimate a reasonable plateau concentration for all scans. A first approximation for the plateau concentration can be obtained by: (i) identifying those earlier scans that have not lost a stable plateau region due to pelleting or back-diffusion and (ii) finding a horizontal region of 50–100 data points in each early scan between the moving boundary and the bottom of the cell. The region can be averaged to provide an initial estimate for the plateau concentration for each early scan. To obtain theoretical plateau values for all other scans that do not have a stable plateau concentration, an extrapolation needs to be made. But extrapolation of what? Unfortunately, the radial dilution equation cannot be used for extrapolation at this point because, in the general case, multiple components result in a multiexponential radial dilution over time, and individual sedimentation coefficients for each species in the system initially are not known:

$$C_{p,i} = \sum_{j=1}^n C_{0,j} \exp(-2s_j\omega^2(t_i - t_0)). \quad (2)$$

Table 1
Restriction digest of plasmid pPOL-1 208-12 with *EcoRI*, *BanII*, and *PvuI*

Fragment	Number of base pairs	Copies	Molecular weight (Da)	TRC (%)	ORC (%)	Sedimentation coefficient (est., $\times 10^{-13}$ s)
Plasmid <i>EcoRI</i> fragment	2823	1	1.655×10^6	27	26	12.7
<i>EcoRI</i> fragment 2	12	11	7.051×10^3	3	—	1.2
<i>BanII</i> fragment 1	46	12	2.698×10^4	5	4	3.5
<i>PvuI</i> fragment 1	1681	1	9.853×10^5	16	16	10.6
<i>PvuI</i> fragment 2	896	1	5.252×10^5	8	8	9.0
<i>EcoRI</i> fragment 1	196	12	1.149×10^5	22	—	—
<i>BanII</i> fragment 2	150	12	8.793×10^4	17	44 ^a	5.7
<i>PvuI</i> fragment 3	246	1	1.422×10^5	2	—	—

Note. TRC, theoretical relative concentration (i.e., the expected relative concentration based on the molar fraction of base pairs for each fragment); ORC, observed relative concentration (i.e., the relative concentration of each fragment observed in the integral vHW distribution as well as the fragment's sedimentation coefficient). Note the close agreement between measured and predicted partial concentration for each fragment. The molecular weight was estimated based on nucleic acid sequence. Generation of this fragment mixture is described in Materials and methods.

^a The 246-, 196-, and 150-bp fragments were not resolved into individual discrete peaks. The entire peak represented 44% of the total concentration.

Here, $C_{p,i}$ is the plateau concentration for each scan i , $C_{0,j}$ is the partial concentration of each component j , and s_j is each component's sedimentation coefficient; other symbols are as in Eq. (1). For a first approximation of the plateau concentrations, it is sufficient to fit those plateau concentrations that can be determined from the experimental data to a sum of suitable basis functions such as a polynomial. The order of the polynomial depends on the number of scans available for fitting but should not be so large as to overdetermine the available information. A linear fit is suggested when only a few scans with a stable plateau region are available; higher accuracy can be obtained by fitting with higher order polynomials only for experiments where a sufficient number of scans with stable plateau regions are present. Using the fitted coefficients from this polynomial, extrapolations can be made to estimate both an initial total concentration and the plateau concentrations for each scan for which an experimentally determined value is not available. The resulting polynomial is extrapolated to $t = 0$ to obtain an estimate for the initial concentration, and plateau concentrations can be estimated for those scans that do not have a stable plateau concentration. In a second iteration described later, this initial estimate for each scan's plateau concentration is further refined.

Baseline determination

The baseline value is best determined from a horizontal region near the meniscus obtained from the final scan in the experiment. This baseline value is used for the determination of boundary fractions from all scans.

Determination of apparent s values for early scans that have not cleared the meniscus

Boundary fractions for scans that have partially cleared the meniscus are assigned in an analogous fashion to boundary fractions for those scans that have cleared the meniscus. However, boundary fractions that intercept the meniscus instead of those portions of the boundary that exhibit finite sedimentation are excluded from the extrapolation.

Identification of boundary segments for scans without a stable plateau

At some point during the velocity experiment, rapidly sedimenting components will have pelleted at the bottom of the cell and cause a decrease in the apparent plateau concentration. At this time, it is difficult to estimate the correct spacing for the boundary fractions because a reliable plateau concentration cannot be determined directly from the scan. Instead, in a first approximation, the plateau concentration extrapolated from the initial polynomial fit described above can be used to define the upper limit. The upper limit is revised in the algorithm presented later.

Identification of boundary segments influenced by back-diffusion

As mentioned earlier, the vHW analysis does not deal with back-diffusion at the bottom of the cell. Therefore, regions in the boundary that have been affected by back-diffusion from the bottom of the cell need to be excluded from the analysis. Because a theoretical plateau value is available for all scans (see above), the correct boundary fractions can be assigned for extrapolation so long as the boundary shape is not influenced by back-diffusion or lost altogether due to pelleting. In the original method, such scans are excluded from the analysis entirely, even though portions of the boundary do contain information suitable for the s^* extrapolation. The difficulty is to identify the position in the boundary where back-diffusion starts to affect the boundary shape. Back-diffusion will be determined largely by the component with the largest diffusion coefficient. In a first approximation, the s value determined from the first boundary fraction (presumably pointing to the smallest s value in the distribution) is used to estimate a diffusion coefficient of a sphere with the same s value. In general, this will be the largest possible diffusion coefficient observed in the sample and, hence, will be a conservative estimate of how far the fastest diffusing particle can diffuse back into the solution column from the bottom of the cell. The distance it can diffuse during the scan's time can be estimated with the following equation, which approximates the concentration gradient of a particle under conditions of diffusion into a semi-infinite medium, with one direction restricted at $x = 0$ while material is also being transported to $x = 0$:

$$\frac{\Delta C_b}{C_0} = \frac{\omega^2 \bar{x} \sqrt{t}}{\sqrt{D}} \left\{ 2 \operatorname{ierfc} \frac{x}{2\sqrt{Dt}} \right\}. \quad (3)$$

Here, $\Delta C_b = C_b - C_0$ in the back-diffusion region at distance x from the cell bottom, C_0 , is the loading concentration, t is the time, D is the diffusion coefficient, ω is the radial velocity, ierfc is the first integral of the complementary error function, and \bar{x} is the average distance between the center of rotation and the meniscus and bottom of the cell (for an explanation of the coordinate system that has been chosen to accord with earlier analyses of limited diffusion, see Fig. 1). The derivation of Eq. (3) is shown in [supplemental material](#). Assuming a small user-selected value for the concentration gradient, $\Delta C_b C_0^{-1}$ typically on the order of experimental noise, and the estimated diffusion coefficient of the most rapidly diffusing component, as well as the time of the scan, it is possible to predict a conservative estimate of the distance x that the particle may have diffused back into the solution column and could have provided a positive contribution (the selected value $\Delta C_b C_0^{-1}$) to the concentration of the plateau. Because the user can adjust the value of this gradient, the user can also control the sen-

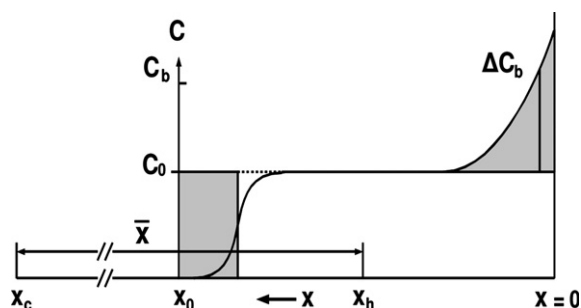


Fig. 1. Illustration of coordinate system used in the derivation of the diffusion equation. To be consistent with the nomenclature of Crank [8], the radius is chosen with $x = 0$ at the bottom of the cell and with x_c at the center of rotation. x_0 is the meniscus position, and x_h is the halfway point between the meniscus and the bottom of the cell where $\bar{x} = x_c - x_h$, and $x_h = x_0/2$. ΔC_b is the concentration in the back-diffusion region at point x where $\Delta C_b = C_b - C_0$ and C_0 is the loading concentration. The areas shaded in gray are equal and each represents the material deposited and diffused back into the solution column.

sitivity of the back-diffusion detection. Visual feedback from the vHW extrapolation plot will provide a convenient measure on how to fine-tune this value. The largest value producing good correlation coefficients for the linear extrapolations is an acceptable value for the gradient. The values of the complementary error function $\text{ierfc}(z)$ were tabulated in [8]. Subtracting the distance from the bottom of the cell position at which an acceptable value of $\Delta C_b C_0^{-1}$ is found, a limit in r is available up to which data from the boundary can be used to calculate apparent s value distributions. This value is recalculated for each individual scan by using the time from each scan and the same diffusion estimate and concentration gradient used for the first boundary fraction. Just like data from partial scans at the beginning of the experiment, partial data from the later scans are used in the extrapolation plot.

Estimation of the correct radial dilution for multicomponent systems

For single-species systems, the boundary fraction spacing is the same for all boundary fractions for a given scan. However, because of radial dilution, which is a function of s and time, the appropriate spacing for boundary fractions for dissimilar species and different scan times will be appreciably different. For multicomponent systems, leaving the spacing constant (as is done in the original vHW method) causes a significant distortion in the extrapolation plot (Fig. 2) because the radial dilution of individual components is not accounted for properly. The distortion becomes especially severe at the boundary fraction near the interface between two components with very different sedimentation coefficients. To address this problem, the following modification is made to the algorithm. The

first division, starting from the baseline, is assigned using a spacing of $(p_i - b)/n$, where n is the number of divisions, p_i is the plateau value of scan i as determined earlier in the article, and b is the baseline. A preliminary s value is calculated by extrapolation to infinite time from all apparent sedimentation coefficients obtained from this first boundary division, resulting in an estimated sedimentation coefficient corresponding to the first boundary division. Using Eq. (2) for the case of a single component, this estimated sedimentation coefficient is now used together with the initial concentration determined earlier, and the time of each scan, to recalculate a new spacing for this first boundary fraction. This results in a new extrapolation that is slightly more accurate than the previous estimate. This process is iterated three times, and the final spacings for each boundary fraction are added to the baseline offset from each scan. Therefore, each subsequent boundary division for each scan is based on the sums of the offsets calculated for the previous divisions. This algorithm ensures that all corresponding boundary fractions are made at points in the boundaries from each scan that correspond to each component's individual radial dilution. The effect of this improvement on the algorithm is best illustrated by comparing the original vHW analysis with the modified version on a sample with two components that sediment with markedly different sedimentation coefficients. Fig. 2 illustrates such a case and clearly shows the improved accuracy in the method when radial dilution corrections are applied correctly.

Creation of histogram plots from integral distributions

A histogram representation of the data contained in the integral distribution plot format has been implemented in the software. To some users, this will present a more intuitive picture of the composition of the sample. Histograms are created by dividing the range of s values covered in the distribution into a finite number of equally spaced s value bins and counting the number of s value extrapolations falling between the boundaries of each bin. The width of the bins is user-selectable. A histogram is then created by plotting the frequency of each bin against the s value. The representation of the histogram data can be refined further by plotting the envelope of the histogram. The envelope can be calculated by representing each bin as a term in a Gaussian sum, where the position of the Gaussian term is given by the center of each bin, the amplitude of the Gaussian is proportional to the frequency of the bin, and the width of the Gaussian is proportional to the width of the bin. This operation is equivalent to numerical differentiation of the integral distribution $G(s)$, followed by smoothing, to yield the familiar differential distribution $g(s)$.

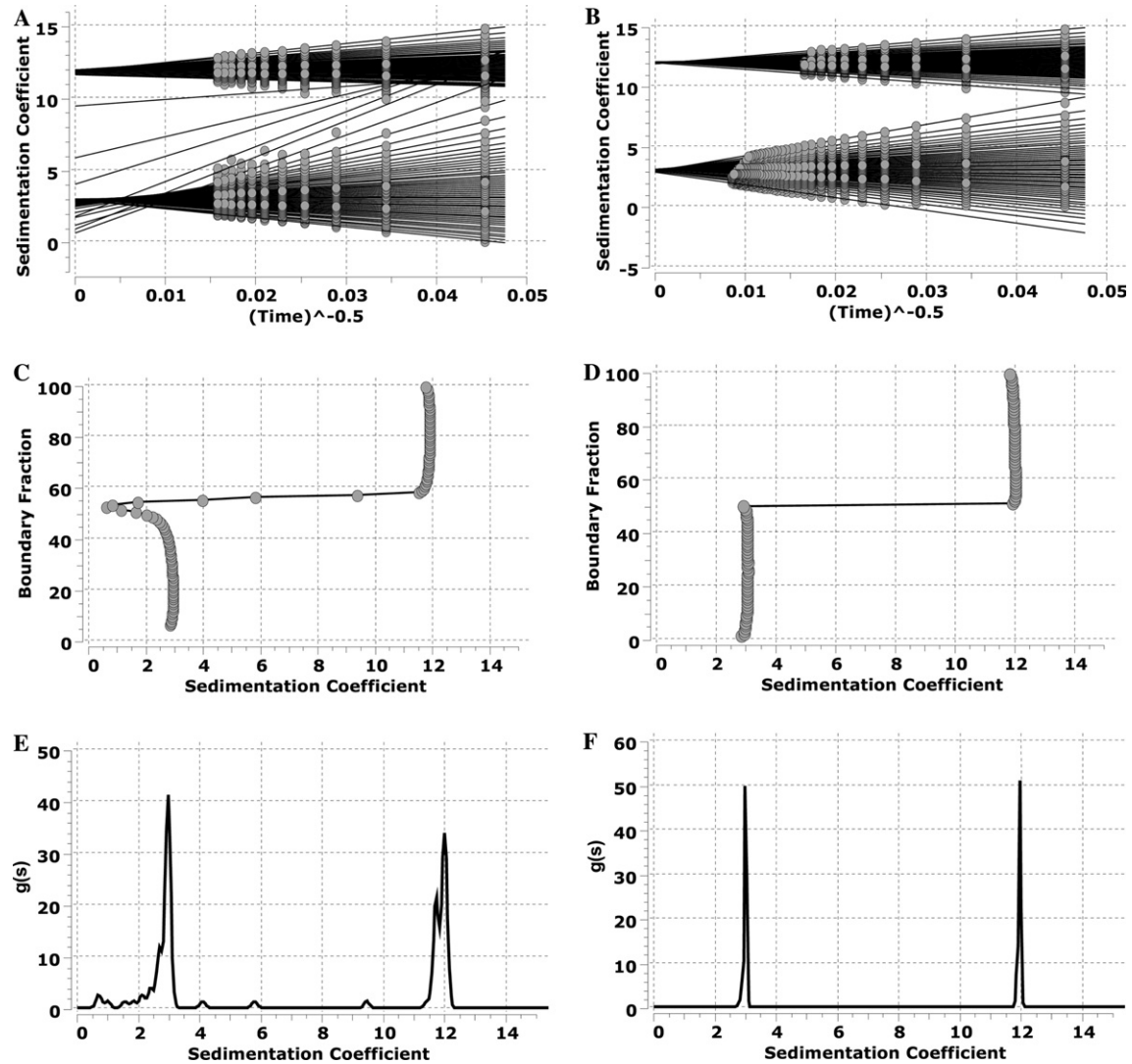


Fig. 2. Illustration of the importance of proper radial dilution corrections for heterogeneous systems. The data shown represent a finite element simulation of a sedimentation velocity experiment with equal loading concentrations of $s_1 = 3.0 \times 10^{-13}$ s, $D_1 = 7.0 \times 10^{-7}$ cm²/s, and $s_2 = 12.0 \times 10^{-13}$ s, $D_2 = 2.0 \times 10^{-7}$ cm²/s, spun for 4 h at 60,000 rpm, resulting in good separation between both components. (A) Traditional vHW extrapolation plot, (C) shows the resulting integral distribution, and (E) shows a sedimentation histogram. (B) The extrapolation plot of the identical data when radial dilution corrections are taken into consideration properly (see text), (D) shows the corresponding integral distribution plot, and (F) shows the associated histogram. The dramatic improvement in accuracy near the interface between the two sedimenting species can clearly be seen.

Results

Tests of the enhanced algorithm with highly heterogeneous systems

Results from simulation studies

The most pronounced benefit from the aforementioned modifications to the vHW algorithm is evident when strongly heterogeneous systems are analyzed. As was pointed out in [2], the best resolution for heterogeneous systems is obtained when the data are collected at maximum speed, which emphasizes separation due to sedimentation but suppresses boundary spreading due to diffusion. With the increased separa-

tion, the fast components will have nearly completely sedimented when the slow components will have barely cleared the meniscus, making it difficult to obtain sedimentation signals useful for the vHW analysis for all components simultaneously. In other words, the early scans will provide a good signal for the fast components, and the late scans will provide a signal for the slow components, but very few scans, if any, encompass information for *all* components. This is illustrated in Fig. 3, which shows the sedimentation data for a simulated velocity experiment of a polymerizing system with seven assembly states (assumed to be not reversibly interacting). For simplicity, each state has been arbitrarily assigned to have a frictional ratio of 1.0,

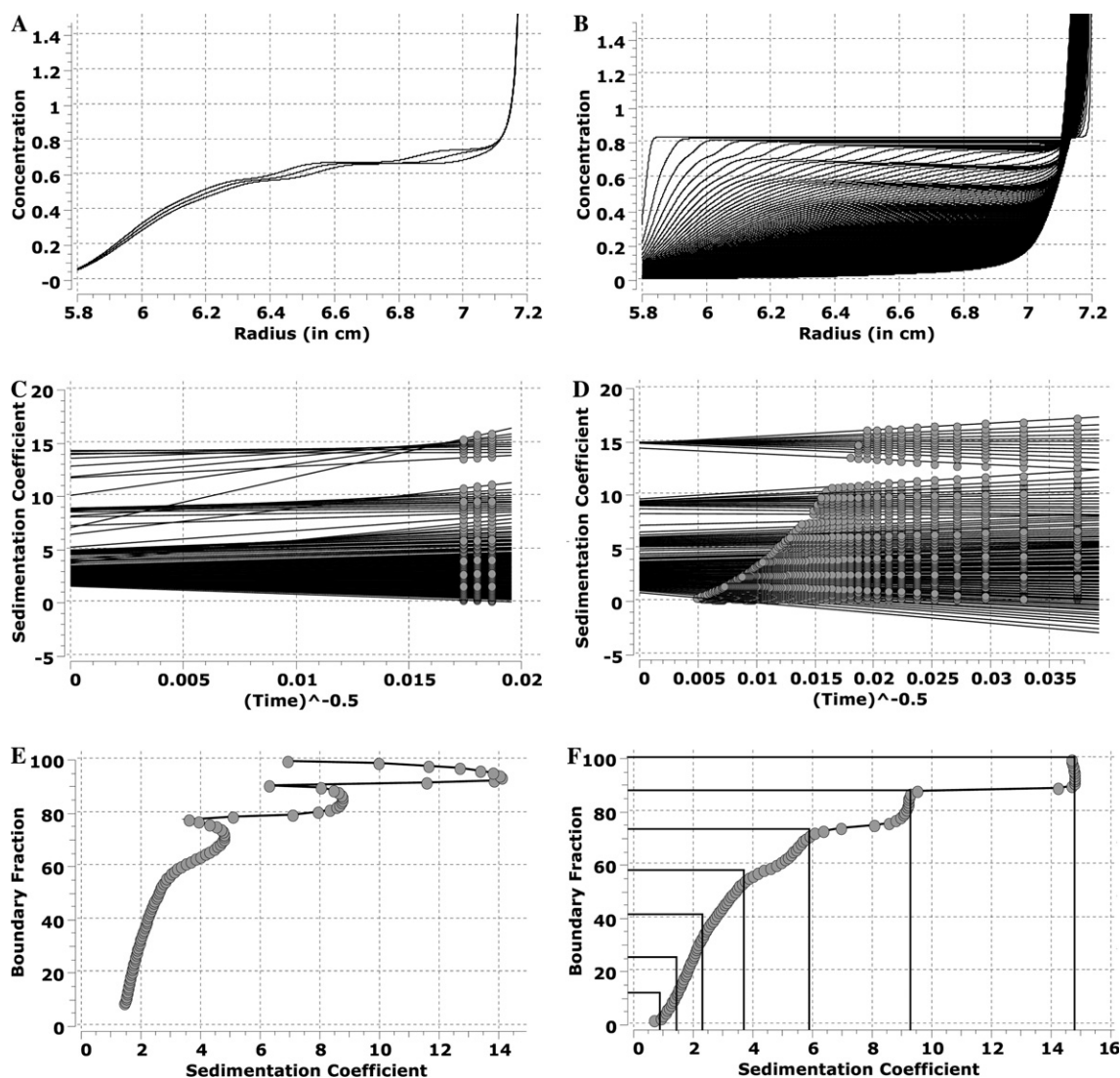


Fig. 3. Illustration of the vHW analysis for a strongly heterogeneous system. The data shown represent a finite element simulation of a velocity experiment for a system with the seven components listed in Table 2. Panels on the left refer to the standard vHW analysis, and panels on the right refer to the enhanced vHW analysis. Panels A and B show those scans that were suitable for analysis in each method. Because of severe heterogeneity, only five scans were suitable for analysis in the traditional vHW analysis (A). When additional scans are included, the portion of the analyzed boundary fraction needs to be reduced to ensure that all boundary fractions measured have intersects with two or more scans to permit extrapolation to infinite time. In the enhanced method (B), all scans could be used in the analysis. The extrapolation plot for the traditional vHW method is shown in (C). Note the poor correlation of the linear fits in the extrapolation and the scatter due to incorrect positioning of boundary fractions resulting from varying degrees of radial dilution (see text for an explanation). Because of the heterogeneity in the sample, only the top 94% of the boundary could be analyzed, cutting off signal from the 6% of the material sedimenting most slowly. (D) The extrapolation from the enhanced vHW method, which offers very good correlation on all extrapolations, produces accurate distributions entailing all components, and allows analysis of 100% of the boundary. Integral distribution plots in (E and F) clearly show the difference in resolution between the two methods. Vertical bars identify the expected s values, and horizontal lines represent the expected relative concentrations of the components.

representing a sphere. For clarity, all data have been simulated without experimental noise. The hydrodynamic parameters used to simulate this experiment are listed in Table 2. This system incorporates a strong heterogeneity in s that is impossible to capture with the traditional vHW method. Even when the analysis is restricted to scans 4–6 so as to exclude boundary effects from the bottom of the cell and after limiting the analysis to the top 80% of the boundary fractions, the anal-

ysis fails to capture the s value of most components in the system (Fig. 3A), whereas the improved version renders the sedimentation coefficients and partial concentrations much better. Because of very large diffusion of the smaller components, individual components are resolved into individual peaks only for s values above 3 s; however, partial concentrations and s values correlate very well with the input model. Fig. 4 depicts a histogram of the results from this analysis.

Table 2

Hydrodynamic parameters and partial concentrations for a simulated polymerizing system of seven assembly states

Component (mer)	Molecular weight (Da)	Partial concentration	Fractional concentration (%)	$s \times 10^{-13}$ s	$D \times 10^{-7}$ cm ² /s
1	4,000	0.100	.12	0.92	20.12
2	8,000	0.117	.14	1.47	16.00
4	16,000	0.129	.16	2.33	12.70
8	32,000	0.133	.16	3.70	10.06
16	64,000	0.129	.16	5.87	7.98
32	128,000	0.117	.14	9.31	6.34
64	256,000	0.100	.12	14.78	5.03

Note. Particles are assumed to be spherical in shape.

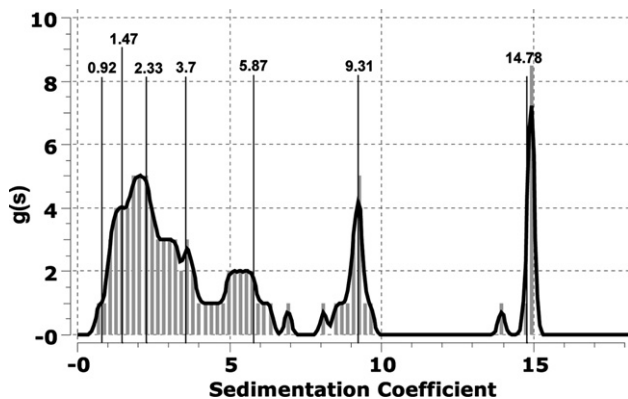


Fig. 4. Histogram and histogram envelope plot of sedimentation data shown in Figs. 3B, D, and F. Vertical bars indicate the s values of the simulated species in the system. Histogram bars refer to the relative frequencies of occurrence of a small s value range. The envelope is generated by using sums of Gaussian terms, as explained in the text.

Experimental tests of the algorithm

The tests described above have depended entirely on simulated data. We considered it imperative to also include a direct experimental comparison of the methods. Comparison of both analysis methods of the sedimentation velocity experiment shown in Fig. 5 demonstrates a significant difference in the resolution of the two meth-

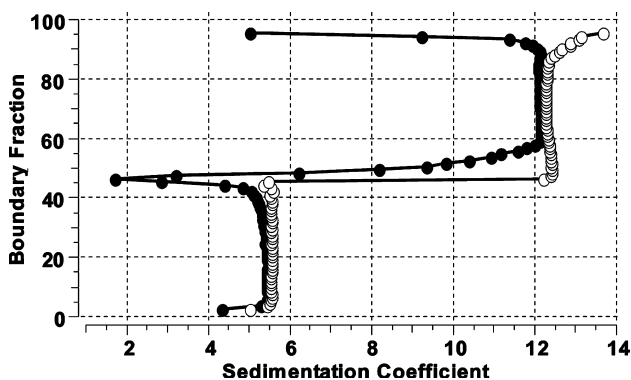


Fig. 5. Integral distribution plot showing the traditional vHW analysis (closed circles) and the enhanced vHW analysis (open circles) of a velocity experiment of sample 1 (*EcoRI* digest of pPOL-208-12). See text for details.

ods. As predicted by the simulation shown in Fig. 2, a well-separated two-component system exhibits a significant distortion of $G(s)$ in the interface region between the two components when analyzed with the traditional method. In addition, distortion at the top of the boundary and failure to distinguish a faster sedimenting component in the top 10% of the boundary by the traditional vHW method both are evident. For the analysis, identical scans and analysis settings were chosen, such that only scans with a stable plateau and cleared meniscus were included in the analysis. The system chosen for this study was plasmid pPOL 208-12 cut with *EcoRI* endonuclease (sample 1). This digestion, when performed to completion, generates three fragments (2823, 196, and 12 bp). The 2823- and 196-bp fragments are present at roughly equal optical density given that the 196-bp fragment is present as 12 copies in the plasmid but there is only one copy of the 2823-bp fragment. The 12-bp fragment is too small and too dilute to be detected in this experiment. When sedimented at moderately high speed (50,000 rpm), the two fragments get completely separated during sedimentation, with each fragment exhibiting a markedly different radial dilution. The homogeneity of each component is clearly demonstrated, and the relative concentrations of the components are represented correctly. To provide a more stringent test, we further digested the *EcoRI*-treated sample with endonucleases *Bam*II and *Pvu*II to generate additional fragments (sample 2) (see Materials and methods). The fragment sizes, sedimentation coefficients, and expected and measured partial concentrations are listed in Table 1. The integral distribution plots of both the traditional vHW analysis and the enhanced vHW analysis are shown in Fig. 6A. A histogram is depicted in Fig. 6B. As in the case of sample 1 and the simulated data, this heterogeneous sample also clearly shows the improvement in the analysis by eliminating the distortions at the interface regions between pairs of components in the sample. In addition, the faster moving components in the top 80% of the boundary are completely missed in the traditional vHW method and are incorrectly included with the slower moving components. Small amounts of slower moving material (the 12-bp fragment) were not detected by either

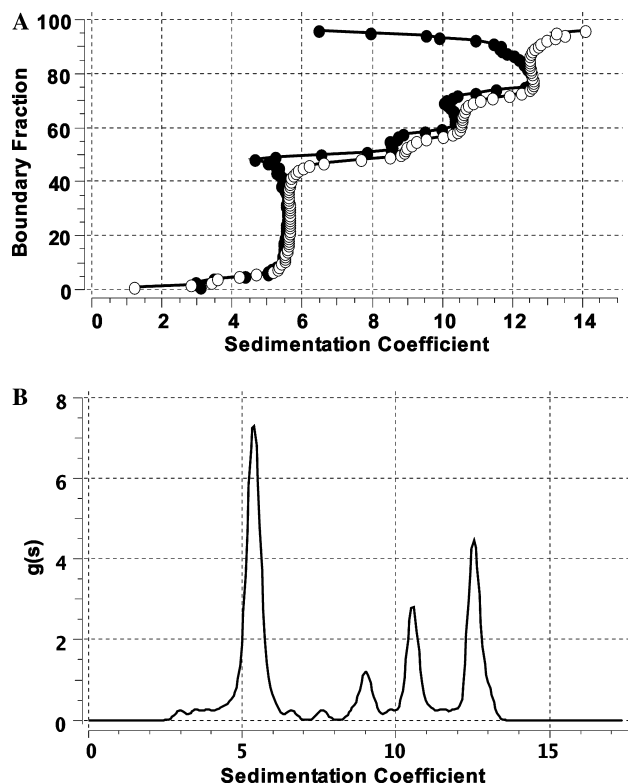


Fig. 6. (A) Integral distribution plot showing the traditional vHW analysis (closed circles) and the enhanced vHW analysis (open circles) of a velocity experiment of sample 2 (*Eco*RI digest of pPOL-208-12, followed by *Pvu*I/*Ban*II digestion as explained in Materials and methods). (B) Histogram plot showing the sedimentation coefficient distribution for the enhanced vHW analysis of the velocity experiment of sample 2. Peaks correspond to the components visible in the integral distribution plot shown in (A).

method. Closely spaced fragments (246, 196, and 150 bp) appear as one band and cannot be resolved with either method.

Discussion

The advantage of the vHW method has always been its ability to separate the sedimentation signal from the diffusion signal and to accomplish this using a model-independent approach, which does not impose any user bias on the data analysis. The new algorithms presented in this article maintain the mathematical rigor and model independence of the original method and provide additional rigor by including a correction that compensates for different degrees of radial dilution of different components during the sedimentation process and by including additional data points in the analysis to provide better statistics for the extrapolations. From the examples shown, the improved accuracy and resolution afforded by the modifications of the method are evident, in particular for the case of highly heterogeneous systems. The integral distribution plot illustrated in

Fig. 3F shows a very accurate reproduction of the sedimentation coefficients in the system, although the relative concentration is somewhat underestimated for the slower sedimenting species. Because of the large diffusion in the slower sedimenting species, the resolution of the slower components is slightly compromised. Even though the range of s values is reliably reproduced, the relative amount of each component can be resolved accurately only if the diffusion is not too large. Another factor leading to a loss of resolution may be the reduced number of data points used for the extrapolations in the boundary fractions in the lower portion of the moving boundary. The experimental results for the heterogeneous DNA sample 2 reproduced the expected component concentrations very well. With the current research focus shifting away from the simple biophysical characterization of pure and isolated systems, the availability of such an analysis tool will be important for the emerging study of mixed systems, hetero- and self-associating systems, determinations of stoichiometry, and applications in quality control in synthetic polymer science. Interested readers can simulate any conceivable system with finite element solutions of the Lamm equation [9] and analyze the resulting sedimentation distributions with the enhanced vHW method so as to assess its ability to resolve individual components under various conditions. Both a finite element simulator and the enhanced vHW method data are implemented in the UltraScan software [6].

Acknowledgments

This work was supported by the National Science Foundation through Grant DBI-9974819 to B. Demeler. We thank Mandy Rolando for preparing the DNA mixtures and Virgil Schirf for performing the analytical ultracentrifugation experiments (both in the Department of Biochemistry at the University of Texas Health Science Center at San Antonio).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.ab.2004.08.039](https://doi.org/10.1016/j.ab.2004.08.039).

References

- [1] K.E. van Holde, W.O. Weischet, Boundary analysis of sedimentation velocity experiments with monodisperse and paucidisperse solutes, *Biopolymers* 17 (1978) 1387–1403.
- [2] B. Demeler, H. Saber, C.H. Jeffrey, Identification and interpretation of complexity in sedimentation velocity boundaries, *Biophys. J.* 72 (1997) 397–407.

- [3] B. Demeler, H. Saber, Determination of molecular parameters by fitting sedimentation data to finite element solutions of the Lamm equation, *Biophys. J.* 74 (1998) 444–454.
- [4] P. Schuck, Sedimentation analysis of noninteracting and self-associating solutes using numerical solutions to the Lamm equation, *Biophys. J.* 75 (1998) 1503–1512.
- [5] B. Demeler, J. Behlke, O. Ristau, Determination of molecular parameters from sedimentation velocity experiments: whole boundary fitting using approximate and numerical solutions of the Lamm equation, *Methods Enzymol.* 321 (1999) 38–66.
- [6] B. Demeler, UltraScan 6.2: An Integrated Data Analysis Software Package for Sedimentation Experiments, Department of Biochemistry, University of Texas Health Science Center at San Antonio, 2004. Available from: <www.ultrascan.uthscsa.edu>.
- [7] P. Georgel, B. Demeler, C. Terpening, M.R. Paule, K.E. van Holde, Binding of the RNA polymerase I transcription complex to its promoter can modify positioning of downstream nucleosomes assembled in vitro, *J. Biol. Chem.* 268 (1993) 1947–1954.
- [8] J. Crank, *The Mathematics of Diffusion*, Clarendon, Oxford, UK, 1956.
- [9] O. Lamm, Die Differentialgleichung der Ultrazentrifugierung, *Ark. Mat. Astron. Fys. B* 21 (1929) 1–4.
- [10] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards, Washington, DC, 1964.
- [11] W.J. Cody, Rational Chebyshev approximations for the error function, *Math. Comput.* 23 (1969) 631–637.
- [12] Netlib Repository (a collection of mathematical software, papers, and databases). Available from: <www.netlib.org/specfun/erf>.