Chapter 10

Analytical Ultracentrifugation

Elena Krayukhina and Susumu Uchiyama

Abstract

Analytical ultracentrifugation (AUC) is a very useful technique to characterize macromolecular interactions. In AUC, a centrifugal force of up to about 250,000 g is applied to a solution of macromolecules, and the progression of sedimentation over time is monitored using an optical detection system. Significant advances in both hardware and software over the past few decades have greatly improved the applicability of AUC for the study of protein–protein interactions. The purpose of this chapter is to provide experimental strategies for the analysis of protein–protein interactions using AUC, including the determination of the association constant of self-associations, binding stoichiometry, and equilibrium binding constant of heterogeneous protein–protein associations. An overview of the method and software packages available for AUC data analysis and optimal protocols for the characterization of protein–protein interactions will be described.

Keywords Sedimentation velocity, Sedimentation equilibrium, Self-association, Hetero-associations, Isotherm analysis, SEDFIT, SEDPHAT

1 Introduction

AUC is an extremely useful technique for studying protein–protein interactions. It can be applied to broad molecular weight distributions (10^2–10^8 Da) to extract parameters such as equilibrium binding constant and binding stoichiometry. It is also a powerful method to assess protein stability and purity.

AUC experiments can be conducted in two basic modes of operation: sedimentation velocity (SV) and sedimentation equilibrium (SE). Regarding data collection, the major advantage of SV over SE is that the required run time is much shorter. Until recently, SE has been used to determine the buoyant molecular weight of the solutes and to estimate the stoichiometry and equilibrium constants of protein–protein interactions [1]. Recent advances in computational approaches for the analysis of SV data have made it possible to extract a wide variety of information from the SV runs [2–6]. Nonetheless, in cases where the number of species involved in the interaction is limited, SE remains the most accurate method to determine the equilibrium constant [7].

Toshiya Senda and Katsumi Maenaka (eds.), *Advanced Methods in Structural Biology*, Springer Protocols Handbooks,
DOI 10.1007/978-4-431-56030-2_10, © Springer Japan 2016

Therefore, prior to SE, the sample purity and aggregation properties should be characterized with SV. The protocols for SV and SE will be described in Sect. 3. First, some general approaches for experimental design applicable to both SV and SE will be described.

2 General Experimental Setup

2.1 Optical Detection Systems

There are three different optical detection systems available for AUC. Considerations associated with each system are briefly summarized in Table 1.

Table 1
Optical detection systems for AUC

	Absorption optics	Interference optics	Fluorescence detection system
Selectivity	High (only components absorbing at the selected wavelength are detected)	Low (all components, including buffer salts, are detected)	High (only fluorescently labeled components are detected)
Loading concentrations	Concentrations producing 0.1 to ~1.5 OD at selected wavelength	Lower limit: concentrations producing a signal above the noise of acquisition (in general, ~0.1 mg/mL) Upper limit: concentrations below those causing nonideality effects (steep concentration gradients causing Wiener skew are to be avoided)	100 pM–1 μM
Scanning speed	~1 min per 1.2 cm solution column; radial scanning across solution column	Whole solution column imaged at once, ~10 s delay between scans	~90 s per 1.2 cm solution column; radial scanning across solution column
Signal-to-noise ratio	~300	>1000	Can be adjusted by changing the photomultiplier tube (PMT) voltage
Sample/reference volume matching	Not required	Exact same volumes should be loaded in sample and reference channels	Not required
Sample/reference component matching	Not required	Exhaustive dialysis, size-exclusion chromatography, or spin columns should be used to chemically equilibrate sample and reference	Not required

2.1.1 Absorption Optics

Absorption optics is the most commonly used optical detection system for AUC as it provides highly sensitive and selective protein detection. Typically the acceptable concentration range is from a few to several hundred micromolar, depending on the absorption coefficient and molecular weight of the protein of interest. The use of different wavelengths combined with 3 mm centerpieces could be employed to extend the applicable concentration range, and successful experiments have been conducted on 24 mg/mL (160 μM) samples [8]. Several important points should be considered when using absorption optics. To maximize the signal-to-noise ratio, the highest possible intensity of the xenon flash lamp is required (Fig. 1). Oil leaking from the vacuum pump can accumulate on the lamp surface and diminish the light output. To ensure the best performance of the lamp, the emission spectrum should be acquired periodically, and the lamp should be cleaned if a decrease in the emission of the peak at 230 nm is detected. Another concern associated with absorption optics is that at the selected wavelength, the total absorbance of the sample placed in the centrifugal cell should be within the dynamic range of the detector. In general, the detected signal should be linear with respect to the concentration of the solute up to 1.5 OD, but it depends on the intensity of the lamp at a particular wavelength. Thus, care must be taken to account for the relative contribution of various components of the solution, including the buffer (see Sect. 2.2) to the total signal. As such, the absorbance of the buffer should be measured against a water blank to determine its absorbance profile. An additional issue concerning absorption optics is that the wavelength accuracy of the monochromator incorporated into the AUC absorbance system is within 1 nm. When a wavelength from the steep portion of the spectrum is chosen for detection, the unpredicted shift of the wavelength during AUC experiment affects the quality of the recorded data and can result in the signal exceeding the dynamic range. The impact of wavelength imprecision

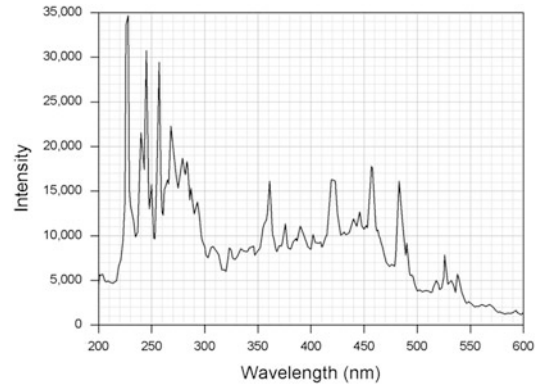

Fig. 1 The intensity profile of xenon flash lamp

can be reduced by using a relatively flat portion of the absorption spectrum such as the maximum of an absorption peak. Most proteins have an absorption peak at around 280 nm attributed to the absorption of aromatic amino acids such as tryptophan, tyrosine, and phenylalanine. Additionally, peptide bond absorption at around 220–235 nm can be used for the data acquisition of a solution with a low absorption at 280 nm. Thus, for the AUC measurements, the recommended wavelengths are 230 and 280 nm for low and medium concentrations, respectively. Highly concentrated solutions can be monitored at around 290 nm. Nevertheless, in highly concentrated solutions, the Wiener skewing effect [9] caused by the large difference in the refractive index between the solvent and solution interferes with the accurate monitoring of sedimentation profiles.

The noise of data acquisition is usually 0.005–0.01 OD, and, considering the upper limit of the dynamic range of 1.5 OD, the maximum achievable signal-to-noise ratio is approximately 300.

2.1.2 Interference Optics

In interference optics, the signal detection is based on the difference of the refractive index between the sample and reference. All components in the solution, including the buffer, contribute to the signal detected by interference optics, and different salt distributions in the sample and reference can affect the recorded signal. To obtain high-quality data, it is imperative to allow the sample and respective reference solvent to chemically equilibrate. This can be achieved through exhaustive sample dialysis against the solvent solution. Another approach is to use the sample after elution from a gel filtration column with the mobile phase being used as the reference solvent. Spin gel filtration columns have also been successfully applied for a similar purpose. Despite these technical challenges, the temporal and radial resolution of data recorded using interference optics is significantly better compared to absorbance data. An entire solution column is imaged at once with the radial step size of approximately 0.002 cm, and the time delay between consecutive scans is only 10 s. There is no specific upper concentration limit; however, significantly steep gradients should be avoided. Solvents containing strongly absorbing compounds, such as ATP, do not pose limitations on signal detection using interference optics.

2.1.3 Fluorescence Optics (Fluorescence Signal Detection System)

Recent developments of fluorescence signal detection system [10] have made it possible to use AUC for the analysis of high-affinity interactions. In addition, such system enables the detection of the sedimentation of the component of interest in complex solutions such as blood serum where other light-absorbing species are present [11]. The covalent attachment of fluorescent dyes required for fluorescence-detected AUC analysis can potentially affect the sedimentation behavior of a molecule due to modifications in its size or shape. Therefore, the impact of labeling on the structure, activity,

or associations of the macromolecule should be examined. Fluorescent emission is detected in the wavelength range of 505–565 nm using laser excitation at 488 nm. Extrinsic dyes with the same excitation wavelength, such as fluorescein, Alexa Fluor 488, Oregon Green, and green fluorescent protein, can be used to label target molecules. At low concentrations, the adsorption of the protein of interest to the windows and centerpiece can potentially interfere with the analysis; thus, for low concentrations, the addition of a "carrier" protein is recommended [10, 12]. Low concentrations (0.1 mg/mL) of ovalbumin, serum albumin, and kappa casein have been used for this purpose.

2.2 Buffers

Buffers used in AUC experiments should contain sufficient salt concentrations to shield unfavorable electrostatic repulsions between molecules. If possible, no gradient forming additives, such as glycerol or sucrose, should be added to the buffer solution. For uncommon solvents, chemical resistances (http://www.uslims.uthscsa.edu/compatibility.php) should be evaluated to select a suitable centerpiece that is compatible with the solvent. In general, it is preferable to use nonabsorbing buffers. For samples with reducing agents, it should be noted that most reducing agents demonstrate significant absorption in the near-UV range that changes in a time-dependent manner. TCEP (tris(2-carboxyethyl) phosphine) is recommended to maintain the reduced state of cysteine residues during the AUC measurement.

3 Methods

3.1 Sedimentation Velocity (SV)

3.1.1 Introduction

SV is a hydrodynamic method that provides information on the size and shape of the solute. SV is applied for the determination of the solute's sedimentation coefficient distribution and to gain limited information on the hydrodynamic shape of the solute.

In SV, the solute sediments under a strong gravitational field and the sedimentation and diffusion fluxes govern the behavior of the particle. The partial differential equation describing the evolution of concentration profiles $C(r,t)$ at each radial position r and time t during the sedimentation process is the Lamm equation [13]:

$$\frac{\partial C}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(rD\frac{\partial C}{\partial r} - s\omega^2 r^2 C\right), \tag{1}$$

where s and D are the sedimentation and diffusion coefficient of the solute, respectively, and ω is the angular speed. The Lamm equation is derived from equations governing sedimentation and diffusion transport processes combined with the balance equation of centrifugal, buoyant, and drag frictional forces acting on the solute molecule. For a mixture of non-interacting solutes, the total concentration

of all solutes can be represented by a sum of Lamm equation solutions L for each solute in the mixture multiplied by the partial concentration c_n:

$$C(r,t) = \sum [c_n L(s_n, D_n, r, t)] \tag{2}$$

Analysis of the sedimentation data by the Lamm equation can provide information about solute sedimentation and the diffusion coefficient. Unfortunately, the Lamm partial differential equation has no general analytical solution. However, the recent availability of powerful computers has favored the development of computer programs for the numerical analysis of sedimentation experiments.

Sedimentation coefficient s (Svedberg units, $1S = 10^{-13}$ s) corresponds to speed u at which the solute molecule moves in the centrifugal field $\omega^2 r$:

$$s = \frac{u}{\omega^2 r} == \frac{M(1-\overline{v}\rho)}{Nf} = \frac{MD(1-\overline{v}\rho)}{RT}, \tag{3}$$

where M is molecular mass, $\overline{v}$ is the partial specific volume, f is the translational frictional coefficient, ρ is the buffer density, T is the absolute temperature, R is the universal gas constant, and N is Avogadro's number.

The diffusion coefficient D can be conveniently expressed through the frictional ratio f/f_0 by using the Stokes–Einstein relationship:

$$D = \frac{RT}{18\pi N\left(f/f_0\eta\right)^{3/2}\sqrt{\frac{s\overline{v}}{2(1-\overline{v}\rho)}}}, \tag{4}$$

where η is the buffer viscosity. A frictional ratio is defined as the frictional coefficient of a protein f divided by the frictional coefficient f_0 of a non-hydrated sphere of equal mass and indicates the degree of globularity of the proteins. While a non-hydrated sphere has a frictional ratio equal to 1, most globular proteins have f/f_0-values in the range 1.2–1.8. For elongated molecules, frictional ratio values can be greater than 2.

The molecular mass M can be derived from the obtained parameters (s, D) using the Svedberg equation:

$$M = \frac{sRT}{D(1-\overline{v}\rho)} \tag{5}$$

3.1.2 Experimental Design and Execution

Protocol 1

1. Choose the appropriate sample concentration. To determine if the protein of interest self-associates, the initial SV runs should be performed with at least three different protein concentrations. The initial cell-loading concentrations should cover an approximately tenfold concentration range. To study the hetero-association of two proteins (A and B), the SV

experiments should be conducted with at least one concentration of A and B alone and at least three mixtures prepared with different concentrations of A and B. In general, the mixtures are prepared in the following manner: the concentration of A is kept constant within a few folds of the expected k_d, and the concentration of B is varied approximately tenfold below and above the expected k_d.

2. Choose the appropriate optical detection system. The choice depends on the concentration range and the nature of the protein (for details, see Table 1 and Sect. 2.1)).
3. Choose the appropriate solvent: see Sect. 2.2.
4. Choose the appropriate centerpieces and load samples into cells. In most cases, standard double-sector centerpieces can be utilized. The sectors are filled with 400–450 μL of the sample. It should be noted that longer solution columns produce higher hydrodynamic resolution and better quality data can be collected for a longer amount of time. In cases where absorption optics is used, the reference sector should be filled with the buffer solution, the volume of which should exceed the sample volume by 5–10 μL to avoid complications caused by signal from the solvent meniscus. When interference optics is utilized, the volumes of the sample and reference should match. Preferably, meniscus-matching centerpieces should be used. However, if the sample and reference menisci are not precisely matched, this can be accounted for computationally during data analysis using SEDFIT software [14].
5. Choose the appropriate temperature. The sample must be stable at the experimental temperature over the course of the experiment. For most applications, 20 °C is appropriate. For the special cases, temperatures between 4 and 40 °C are available using Beckman Coulter XL-A/I ultracentrifuges. Before the run, carefully equilibrate the rotor with the samples loaded at 0 rpm for at least 30 min after the rotor reaches the target temperature. It is important to avoid convection at the beginning of the run, which is caused by the mixing of the solution layers of different temperatures.
6. Choose the appropriate rotational speed and scan interval. A speed should be chosen so that at least 40 scans can be recorded before the sedimentation of the sample is complete. Simulations available in SEDFIT [15] or UltraScan [16] software packages estimate the optimum speed and consequently an approximate time to complete sample sedimentation (http://www.analyticalultracentrifugation.com/generating_simulated_data.htm; http://www.ultrascan3.uthscsa.edu/manual/astfem_sim.html). The scan interval should be as short as possible, but the sample should completely sediment before the maximum number of the scans (999) is reached.

7. Start the method scan. Collect data until the sample sedimentation is complete, which generally requires between 2 and 12 h depending on the solute size and rotational speed.
8. Stop the run. In principle, after the AUC experiment, the samples can be recovered from the cell assembly. However, due to possible changes in the structure and aggregation state of the solutes, this is not generally recommended.
9. Clean the components of the cell assembly. It is a good practice to use the same combination of cell housing, windows, and centerpiece during cleaning and assembly. In this manner, the defective components affecting the quality of the data can be easily detected and eliminated.

3.1.3 Data Analysis

3.1.3.1 Determination of Sedimentation Coefficient Distribution

The most commonly used approach to initial data analysis is the sedimentation coefficient distribution, C(s), implemented in the SEDFIT software [3, 15]. This method requires no prior knowledge of sample properties and can be conveniently used to determine the number of sedimenting species, sedimentation coefficients, and molecular masses. C(s) is a direct least-squares method for modeling experimental data using numerical solutions of the Lamm equation. To calculate diffusion coefficients, Eq. 4 is used, where it is assumed that all sedimenting species have the same frictional ratio f/f_0. This assumption is based on the lower size dependence of diffusion relative to sedimentation and weak shape dependence of the frictional ratio. The weight-average f/f_0 value can be optimized in a nonlinear regression during C(s) analysis. For heterogeneous systems, where multiple species with different shapes are present at comparable concentrations, a single frictional ratio is not suitable to describe all the components and results in skewed molecular mass determinations. However, when a single peak is seen in the C(s) distribution, the molecular mass estimation can be expected to be within 10 % of the true value.

Protocol 2

1. Load scan files into SEDFIT. The data are color-coded according to the acquisition time: scans recorded at the beginning of the experiment are shown in black and the latest scans are indicated in red. Select the appropriate number of scans so that the transition from a green to red color is seen in the middle of the loaded data set.
2. Specify the meniscus, bottom position, and fitting limits. Set meniscus (red line) to the midpoint position of the absorbance spike corresponding to the air–sample boundary. Set the bottom position (blue line) to the maximum signal corresponding to optical artifacts at the end of the solution column. Set the left and right data analysis limits (green lines) to exclude the region of optical artifacts close to the meniscus and bottom.

3. Choose continuous C(s) distribution from the "Model" menu. In the "Parameter" box, input the minimum (s_{min}) and maximum (s_{max}) expected sedimentation coefficient values. Input the resolution. This parameter corresponds to the number of species with different s-values between s_{min} and s_{max} in which relative abundance will be determined in the C(s) analysis. Input the initial value for the frictional ratio: 1.2 for globular proteins, 1.5 for antibodies or other asymmetrically shaped proteins, and 2.0 or higher for rod-shaped and unfolded proteins, fibrils, and DNA. Input the values for partial specific volume (vbar), solvent density, and viscosity. Set the confidence level to 0.68. Check the boxes for the frictional ratio, baseline, meniscus, and time-independent noise (and radial-independent (RI) noise when interference optics is used for the data acquisition) in order to optimize these parameters.
4. Use the "Run" command to estimate the initial guesses for the parameters entered in the previous step. If the distribution significantly deviates from zero at the minimum or maximum s-value, select a higher value for s_{max} and a lower value for s_{min}, respectively. Execute the "Run" command with refined parameters. Repeat until there are no peaks at the maximum and minimum s-value in the C(s) distribution.
5. Optimize the initial parameters by executing the "Fit" command. Assess the quality of the fit by verifying that the root mean-square deviation (rmsd) does not exceed 0.1 % of the total loading signal value. The randomness of the residuals can be ensured by the absence of visible diagonal lines at the residuals bitmap. If a good quality optimization is achieved, the peaks in the resulting C(s) distribution correspond to the sedimenting species. The displayed fitted frictional ratio should be consistent with the known properties of the sample (folded/unfolded chains) and should always be >1. Values <1 indicate extra boundary broadening not originating from diffusion, but likely from rapid ($k_{off} > 0.01/s$) chemical reactions.
6. Estimate the molecular weights of the detected species by choosing "Display Mw peaks in C(s)" from the "Display" menu or by clicking Ctrl-M. The obtained values should be interpreted with care (see Sect. 2.1).

3.1.3.2 Isotherm Analysis

The isotherm of weight-average sedimentation coefficients, s_w, as a function of protein concentration is constructed. The experiments performed at different protein concentrations are analyzed to elucidate if reversible self-association is present. Available methods for data analysis include $g^*(s)$ [17], van Holde–Weischet analysis [18], and two-dimensional spectrum analysis [4], with the C(s) analysis being the method of choice. Even though the C(s) analysis is based

on the assumption that all solutes are non-interacting, the integration of the size-distribution profiles over the entire sedimentation coefficient range provides a correct weight-average sedimentation coefficient, from which the equilibrium constant for a self-association or hetero-association of the solutes can be characterized.

In addition to the determination of the presence of associations, C(s) allows for the estimation of the kinetics of those interactions. If the peaks are broad and their positions are concentration dependent, then there is a fast reaction taking place. In contrast, for a slow reversible system, the peaks would be sharper and at constant positions, and only the relative peak heights would vary with concentration.

Protocol 3

1. Analyze the collected data according to Protocol 2. Integrate the area under the corresponding peaks by selecting "Integrate distributions" under "Size-distributions options" under the "Options" menu of the SEDFIT main window or simply by clicking Ctrl-I. Note the weight-average sedimentation coefficients and write them in a second column in a tab-delimited text file, with the first column representing the loading concentrations. Alternatively, the signal-average sedimentation coefficient isotherms can be conveniently constructed using GUSSI software (http://biophysics.swmed.edu/MBR/software.html).
2. Load the isotherm file into the SEDPHAT window. In the "Experimental parameter" box, input the partial specific volume, buffer density and viscosity, extinction coefficient, and optical path length.
3. Choose the appropriate model from the "Model" menu and execute the "Fit" command. To increase the precision of the determined k_d, prior knowledge of the sedimentation coefficients of either individual components or complexes can be incorporated in the analysis. In self-associating systems, the sedimentation coefficients can be derived from available crystal structures by constructing hydrodynamic bead models using SOMO [19] or HYDROPRO [20]. For hetero-associating systems, the sedimentation coefficients of A and B can be derived from the experiments performed using the individual components.

An example of isotherm analysis conducted to study the self-association of semaphorin 6A (Sema6A) receptor-binding fragment is presented in Fig. 2a [21]. Figure 2b provides an example of the monomer–dimer–tetramer equilibrium of wild-type hemoglobin.

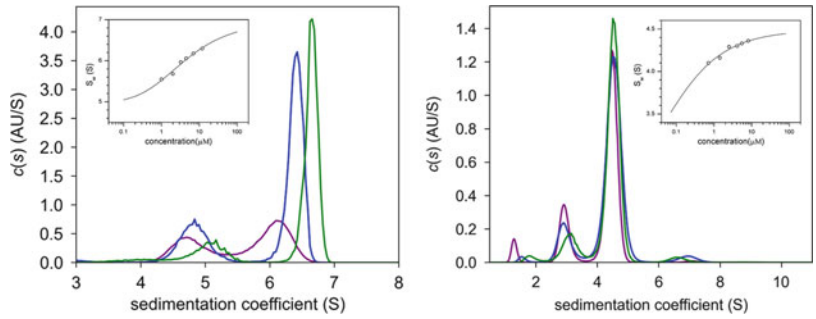

Fig. 2 Examples of isotherm analysis conducted to determine the dissociation constant of self-associating proteins. (**a**) C(s) distribution from SV experiments performed at different concentrations of the semaphorin 6A receptor-binding fragment Sema6A$_{SP}$. For the clarity of presentation, only the distributions calculated for 1 (*purple*), 3.25 (*blue*), and 12 μM (*green*) data are plotted. The concentration-dependent change observed in the sedimentation coefficient distribution indicates the presence of a monomer–dimer equilibrium. The isotherm analysis of the weight-average sedimentation coefficients yielded a k_d value of 3.5 μM (Adapted from ref. 21). (**b**) C(s) distribution from SV experiments performed at different concentrations of wild-type human hemoglobin. For the clarity of presentation, only the distributions calculated for 2.5 (*purple*), 7.5 (*blue*), and 10 μM (*green*) data are plotted. The concentration-dependent changes observed in the areas of the peaks indicate the presence of a dimer–tetramer equilibrium. The isotherm analysis of the weight-average sedimentation coefficients yielded a k_d value of 0.1 μM

3.1.3.3 Direct Boundary Modeling of SV Data: Using Prior Knowledge from Non-denaturing Mass Spectrometry

In SV, information about the molecular mass of the species is obtained using the frictional ratio parameter, which is extracted from modeling the sedimentation boundary spreading. For multi-component solutions which contain reactive species with a broad range of sizes or shapes, the determination of molecular masses is often difficult, as in addition to diffusion, the shape of the sedimentation boundary is dependent on both conformational heterogeneity and reaction kinetics [22]. Consequently, if the model applied for the data analysis does not account for either of the factors, the estimates of the obtained parameters may be incorrect. Likewise, incorporating all factors in the fitting model can significantly complicate the analysis and potentially compromise the results.

An alternative approach to SV is mass spectrometry (MS) which is capable of providing the most accurate molecular mass determination. Nonetheless, nonspecific interactions occurring during the electrospray ionization process can affect the distribution of oligomeric species. Therefore, the combination of SV and MS may be useful for the characterization of complex protein solutions.

The study of the assembly states of the nucleosome assembly protein 1 (NAP-1) reported by Noda et al. [6] highlights the utility of proposed technique. Prior to SV, the oligomeric states of NAP-1 were characterized by MS under non-denaturing conditions. The results indicated that the primary oligomeric unit of NAP-1 was a dimer, and a portion of the dimers further assembled into higher oligomers. Then, the assembly states of NAP-1 in solution were characterized using SV. The initial data analysis performed using

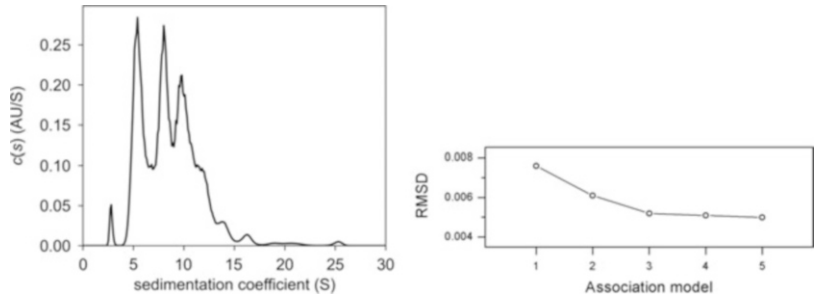

Fig. 3 Analysis of SV data from the study of the assembly states of the nucleosome assembly protein 1 (NAP-1). (**a**) C(s) distribution of human NAP-1 at 150 mM NaCl. (**b**) Plot of RMSD values from the results of "Hybrid local continuous distribution and global discrete species" analysis by the program SEDPHAT of human NAP-1. The association model number indicates *1* 1-2-4-6-8-mer model, *2* 1-2-4-6-8-10-mer model, *3* 1-2-4-6-8-10-12-mer model, *4* 1-2-4-6-8-10-12–14-mer model, and *5* 1-2-4-6-8-10-12-14-16-mer model (Adapted from Ref. 21)

the C(s) model of SEDFIT allowed accurate determination of the sedimentation coefficients and relative concentration of each oligomeric species (Fig. 3a). The assignment of molecular mass to the peaks detected in the C(s) distribution, however, was complicated by the heterogeneity of the sample and the single weight-average f/f_0 value was not suitable to describe each component individually. Thus, the findings from the non-denaturing MS measurements were incorporated as prior knowledge in the SV data analysis using SEDPHAT. "Hybrid local continuous distribution and global discrete species" analysis using a number of different models including 1-2-4-6-8-mers, 1-2-4-6-8-10-mers, 1-2-4-6-8-10-12-mers, 1-2-4-6-8-10-12-14-mers, and 1-2-4-6-8-10-12-14-16-mers was performed. With the increasing number of oligomeric species included in the model, the rmsd value decreased demonstrating a higher-quality fit (Fig. 3b). The 1-2-4-6-8-10-12-mers, 1-2-4-6-8-10-12-14-mers, and 1-2-4-6-8-10-12-14-16-mers models showed similar rmsd values, which indicated that the 1-2-4-6-8-10-12-mers model was the most appropriate for the data set analysis, according to the principle of parsimony.

3.1.3.4 Direct Boundary Modeling of SV Data: The Estimation of Kinetic Information for Systems with Reversible Associations

The sedimentation coefficients and equilibrium constants obtained from isotherm analysis can be further refined using the direct Lamm equation modeling approach.

Protocol 4

1. Load xp-files of the SV experiments into SEDPHAT. These files can be prepared while analyzing data in SEDFIT to construct the isotherm of the weight-average sedimentation coefficients. Detailed instructions on the preparation of xp-files are available elsewhere (http://analyticalultracentrifugation.com).

2. Select the model and enter the starting values for the species *s*-values and the equilibrium constant from the isotherm analysis. Estimate the chemical off-rate constant $\log_{10}(k_{off}) = -3$ for rapid interactions or -4 to -5 for slow interactions relative to sedimentation.
3. Fit this model by first optimizing only the starting concentrations. At the next step, allow the algorithm to perform the optimization of the equilibrium binding and reaction rate constants and the species *s*-values.
4. Evaluate the fit by noting the rmsd value and randomness of the residual distribution.
5. Different models can be tested. The one producing the lowest rmsd value coupled with a random distribution of residuals should be considered as the most appropriate.

3.1.3.5 Multi-signal SV (MSSV)

Multi-signal SV (MSSV) is a SV technique utilized in the study of heterogeneous protein interactions. A detailed description of this method is available in [23]. MSSV enables the investigation of binary and ternary complexes formed in mixtures of three different proteins. To resolve interacting components in MSSV, the components must show sufficiently different spectral signatures. To evaluate whether MSSV is a suitable approach for a particular mixture, the value of D_{norm} [24] is calculated based on the known extinction coefficients. Successful examples of three protein-component mixtures analyzed by MSSV are described in [25, 26].

3.2 Sedimentation Equilibrium (SE)

3.2.1 Introduction

SE experiments are conducted at lower rotational speeds than SV experiments. The sedimentation flow is opposed by counterflow diffusion that is generated according to the derivative of the concentration at a radial position. At the equilibrium state, the sedimentation force applied to the solute is balanced by the diffusion force, leading to the formation of a steady-state exponential concentration gradient. SE provides information on the total profile of detectable solute with a selected optical detection system, and therefore high purity samples containing a small number of species are preferred. Analysis of the sample by SV should be carried out prior to the SE to confirm the absence of impurities.

SE experiments provide information about solute buoyant molar mass, association constants, association stoichiometries, and second viral coefficient related to the thermodynamic nonideality of the solution. Similar to SV, the behavior of the particle in the cell is described by the Lamm equation. Unlike SV, in SE the system is studied at equilibrium, and thus the total flux, comprised of sedimentation flux and opposing diffusion flow, equals 0:

$$s\omega^2 rC - D\frac{\partial C}{\partial r} = 0 \tag{6}$$

The solution of this equation corresponds to the exponentially increasing concentration profile:

$$C(r) = C_0 e^{\frac{s\omega^2}{2D}\left(r^2 - r_0^2\right)} \tag{7}$$

where C_0 is concentration at a radial reference point r_0 in the concentration gradient. By inserting the Svedberg Eq. 5, the following expression is derived:

$$C(r) = C_0 e^{M(1-\overline{\nu}\rho)\frac{\omega^2}{2RT}\left(r^2 - r_0^2\right)} \tag{8}$$

Thus, the steepness of the concentration gradient at any particular rotor speed is determined by the buoyant molecular mass $M_b = M(1 - \nu_{\text{bar}}\rho)$. In contrast to SV, the molecular shape of the solute has no effect on the result of SE experiments within ideal solutions. The buoyant molecular mass thus can be obtained from SE experiments, and the weight-average molecular weight of the macromolecule of interest, M, can be estimated given an accurate partial specific volume. The partial specific volume can be determined experimentally by measuring the concentration dependence of the protein solution or by using density contrast in mixtures of light and heavy water [27, 28] or theoretically from the amino acid composition of the protein using SEDNTERP (http://sednterp.unh.edu/).

For a mixture of solutes, the total equilibrium concentration gradient is expressed by the following equation:

$$C_{\text{total}}(r) = \sum\nolimits_i C_{0,i} e^{M_i(1-\overline{\nu}_i\rho)\frac{\omega^2}{2RT}\left(r^2 - r_0^2\right)} + \text{baseline}, \tag{9}$$

where C_0 of the complex can be described using the C_0 values of each component and the equilibrium constant of the interaction between or among the components. In the nonlinear fitting of SE data, the C_0 values, baseline, and k_d are set as variable parameters, while M_i and ν_i are typically calculated based on the amino acid composition and are set as fixed parameters.

3.2.2 Experimental Design and Execution

3.2.2.1 Self-Association by SE (Example $A + A = A_2$)

Protocol 5

1. Choose the appropriate sample concentration. In order to determine the association constant, a broad concentration range with multiple loading concentrations should be used. At low concentrations, monomers will primarily contribute to the signal, while at high concentrations the signal will be dominated by oligomeric forms. Prior to SE experiments, it is highly preferable to characterize the sample by SV according to Protocol 1 and Protocol 2. The sample should be well purified (typically more than 95 % purity) and chemically equilibrated with its reference solvent if interference optics is utilized.

2. Choose the appropriate sample volume. Usually 3 mm solution columns (100–120 μL) are sufficient for SE experiments. Longer columns require longer times to reach equilibrium; however, concentration gradients extending longer distances provide better parameter precision. For low molecular mass proteins, higher volumes may be required to produce a concentration gradient with a sufficient length of curvature. Before performing the experiment, it is recommended to simulate data using the "Estimate equilibrium rotor speeds" option under the "Calculator" menu of SEDFIT.
3. Choose the appropriate optical detection system. The choice depends on the concentration range and the nature of the protein including the amino acid composition (for details, see Table 1 and Sect. 2.1).
4. Prepare cells for sample loading. If interference optics is chosen, before loading the samples, the assembled cells should be mechanically "aged" (for details, see http://analyticalultracentrifugation.com) to minimize the impact of time-independent noise, which can change over the time course of the SE experiment due to mechanical micro-movements of the assembly parts. Similar to Protocol 1, the same (interference optics) or 5–10 μL larger solvent volumes (absorption optics) should be loaded in the reference sector.
5. Choose the appropriate temperature. The sample should be stable at the experimental temperature during the equilibrium run (depending on the settings, the run might require 1 week or longer). For most applications 20 °C is an appropriate choice. For special cases, temperatures between 4 and 40 °C are available using the XL-A/I ultracentrifuge. In contrast to the SV run, there is no need to equilibrate the rotor with the samples loaded at the target experiment temperature.
6. Choose the appropriate rotational speed. A single speed cannot distinguish interacting and non-interacting species when a sample solution with a single concentration is measured. Therefore, three rotor speeds should be chosen for the experiment. The slowest rotational speed provides a shallow gradient resulting in information about the largest species in the sample. At the highest rotational speed, meniscus depletion should be achieved and a steep concentration gradient should be observed. This data set provides information about the smallest species. Simulations available in SEDFIT or UltraScan software packages allow for the convenient estimation of the best speed (see http://www.analyticalultracentrifugation.com/generating_simulated_data.htm, http://www.ultrascan2.uthscsa.edu/manual2/finsim.html for details).

7. Collect the data. Start collecting multispeed data from the lowest speed chosen in the previous step. Data are collected every 6 h and successive scans are compared by the SEDFIT or WinMatch program. Equilibrium is attained when the subtraction of two consecutive scans produces no systematic difference. The minimum time required to reach equilibrium can be estimated using the "Calculator" menu of SEDFIT. Once equilibrium has been attained, the data can be collected at the next speed.
8. After equilibrium is attained at the highest rotor speed and all the data have been collected, stop the run and clean the components of the cell assembly.

3.2.2.2 Hetero-associations by SE (Example A + B = AB)

Protocol 6

1. Prepare a series of sample concentrations. Each component should be measured individually and a mixture of components should be prepared as a dilution or titration series. To avoid nonideality, which complicates data interpretation and analysis, the concentration range should be chosen within 0.1–10 × k_d, producing an absorbance signal within 0.1–0.75 OD or larger than 0.1 fringes. Again, if interference optics is chosen, the sample should be free from impurities and equilibrated with its reference solvent.
2. Choose the appropriate sample volume (see Protocol 5).
3. Choose the appropriate optical detection system. The detection at multiple wavelengths (230, 250, and 280 nm) combined with interference allows for a wide range of suitable loading concentrations.
4. Load cells with the sample, choose the appropriate temperature and rotational speed, and collect the data according to Protocol 5.

3.2.3 Data Analysis

Protocol 7

1. In SEDFIT, preprocess the data for further analysis using SEDPHAT. In the "Loading Options" menu of SEDFIT, choose "Sort EQ data to Disk" and convert equilibrium data to (*.xp) files suitable for the SEDPHAT analysis (for details, refer to http://www.analyticalultracentrifugation.com/seprotocols.htm).
2. Analyze the data collected for the individual components. In SEDPHAT, load the xp-files associated with only one component of the interacting system. In the "Model" menu, select "A (single species of interacting system)."
3. Analyze the data acquired for the mixtures of components. In SEDPHAT, load the xp-files associated with the mixtures of

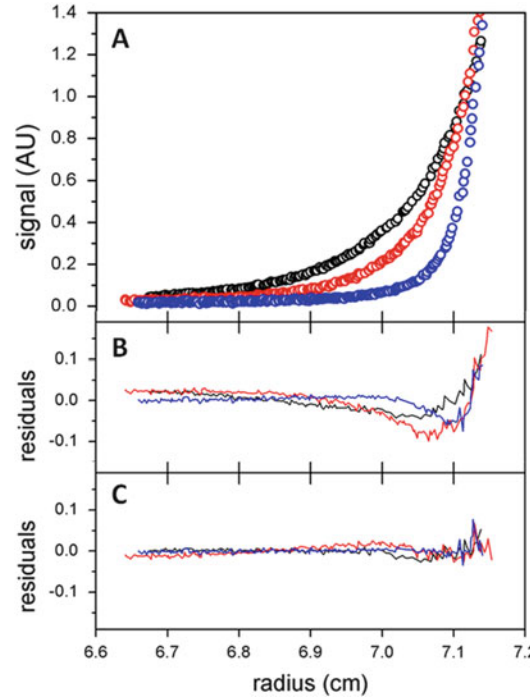

Fig. 4 SE analysis of antibody and antigen interaction. (**a**) SE concentration gradient for mAb (antibody) to NP and NP-conjugated BSA (antigen) mixed solutions each at 3.3 μM (equimolar condition). (**b**) Nonrandomly distributed residuals and high chi-squared value of 0.0231112 indicate that 1:1 interaction model is inadequate in this case. (**c**) Randomly distributed residuals and significant improvement of chi-square (0.00555707) support the 1:2 interactions

interacting components, and in the "Model" menu, choose one of the models for heterogeneous A and B interactions. To initiate the analysis, use the parameters obtained during the previous step. At this stage, time-independent noise decomposition should not be attempted as it can correlate with the model used. Different models should be tested, and the model providing the best quality fit, which is evaluated using the rmsd values for each xp-file and the randomness of the residuals, should be considered the most appropriate. Then, include "TI noise" and allow the algorithm to optimize the parameters. This should result in a decrease of the rmsd value. Ensure a relatively flat TI-noise profile with no apparent curvature.

4. Alternatively, the stoichiometry and k_d can be estimated from the nonlinear least-squares fitting of acquired data to Eq. (8) by a homemade program using software equipped with a nonlinear fitting algorithm, such as Mathematica. An example of SE analysis of antibody and antigen interaction is presented in Fig. 4.

4 Note

Recent findings suggested that time stamps recorded in the sedimentation scan files by AUC software were incorrect, leading to errors in sedimentation coefficients and molecular weight estimations [29, 30]. Even though it was discussed that the binding constants obtained from the application of isotherm analysis are unaffected by the incorrect time stamp, the absolute values of the sedimentation coefficients will be incorrect. Therefore, the use of SEDFIT (version 14.0c or later) software is recommended to compensate for possible errors.

References

1. Kato K, Sautes-Fridman C, Yamada W et al (2000) Structural basis of the interaction between IgG and Fcgamma-receptors. J Mol Biol 295:213–224
2. Philo JS (2000) A method for directly fitting the time derivative of sedimentation velocity data and an alternative algorithm for calculating sedimentation coefficient distribution functions. Anal Biochem 279:151–163
3. Schuck P (2000) Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophys J 78:1606–1619
4. Brookes E, Cao W, Demeler B (2010) A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. Eur Biophys J 39:405–414
5. Oda M, Uchiyama S, Noda M et al (2009) Effects of antibody affinity and antigen valence on molecular forms of immune complexes. Mol Immunol 47:352–364
6. Noda M, Uchiyama S, McKay AR et al (2011) Assembly states of the nucleosome assembly protein 1 (NAP-1) revealed by sedimentation velocity and non-denaturing mass spectrometry. Biochem J 436:101–112
7. Oda M, Uchiyama S, Robinson CV et al (2006) Regional and segmental flexibility of antibodies in interaction with antigens of different size. FEBS J 273:1476–1487
8. Nishi H, Miyajima M, Nakagami H et al (2010) Phase separation of an IgG1 antibody solution under a low ionic strength condition. Pharm Res 27:1348–1360
9. Svensson H (1954) The second order aberrations in the interferometric measurement of concentration gradients. Optica Acta 1:25–32
10. Kingsbury JS, Laue TM (2011) Fluorescence-detected sedimentation in dilute and highly concentrated solutions. Methods Enzymol 492:283–304
11. Demeule B, Shire SJ, Liu J (2009) A therapeutic antibody and its antigen form different complexes in serum than in phosphate-buffered saline: a study by analytical ultracentrifugation. Anal Biochem 388:279–287
12. Cole JL, Lary JW, P Moody T, Laue TM (2008) Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. Methods Cell Biol 84:143–179
13. Lamm O (1929) Die differentialgleichung der ultrazentrifugierung. Ark Mater Astr Fys 21B: 1–4
14. Zhao H, Brown PH, Balbo A et al (2010) Accounting for solvent signal offsets in the analysis of interferometric sedimentation velocity data. Macromol Biosci 10:736–745
15. Schuck P (2005) Diffusion-deconvoluted sedimentation coefficient distributions for the analysis of interacting and non-interacting protein mixtures. In: Scott DJ, Harding SE, Rowe AJ (eds) Analytical ultracentrifugation: techniques and methods. RSC Publishing, Cambridge, pp 26–49
16. Demeler B (2005) Ultrascan: a comprehensive data analysis software package for analytical ultracentrifugation experiments. In: Scott DJ, Harding SE, Rowe AJ (eds) Analytical ultracentrifugation: techniques and methods. RSC Publishing, Cambridge, pp 210–229
17. Schuck P, Rossmanith P (2000) Determination of the sedimentation coefficient distribution g*(s) by least-squares boundary modeling. Biopolymers 54:328–341
18. Demeler B, van Holde KE (2004) Sedimentation velocity analysis of highly heterogeneous systems. Anal Biochem 335:279–288
19. Brookes E, Demeler B, Rosano C, Rocco M (2010) The implementation of SOMO

(SOlution MOdeller) in the UltraScan analytical ultracentrifugation data analysis suite: enhanced capabilities allow the reliable hydrodynamic modeling of virtually any kind of biomacromolecule. Eur Biophys J 39:423–435

20. Ortega A, Amorós D, Garcia de la Torre J (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. Biophys J 101:892–898
21. Nogi T, Yasui N, Mihara E et al (2010) Structural basis for semaphorin signalling through the plexin receptor. Nature 467:1123–1127
22. Dam J, Velikovsky CA, Mariuzza RA et al (2005) Sedimentation velocity analysis of heterogeneous protein-protein interactions: Lamm equation modeling and sedimentation coefficient distributions c(s). Biophys J 89: 619–634
23. Padrick SB, Deka RK, Chuang JL et al (2010) Determination of protein complex stoichiometry through multisignal sedimentation velocity experiments. Anal Biochem 407:89–103
24. Padrick SB, Brautigam CA (2011) Evaluating the stoichiometry of macromolecular complexes using multisignal sedimentation velocity. Methods 54:39–55
25. Houtman JC, Yamaguchi H, Barda-Saad M et al (2006) Oligomerization of signaling complexes by the multipoint binding of GRB2 to both LAT and SOS1. Nat Struct Mol Biol 13:798–805
26. Barda-Saad M, Shirasu N, Pauker MH et al (2010) Cooperative interactions at the SLP-76 complex are critical for actin polymerization. EMBO J 29:2315–2328
27. Edelstein SJ, Schachman HK (1967) The simultaneous determination of partial specific volumes and molecular weights with microgram quantities. J Biol Chem 242:306–311
28. Brown PH, Balbo A, Zhao H et al (2011) Density contrast sedimentation velocity for the determination of protein partial-specific volumes. PLoS One 6:e26221
29. Zhao H, Ghirlando R, Piszczek G et al (2013) Recorded scan times can limit the accuracy of sedimentation coefficients in analytical ultracentrifugation. Anal Biochem 437:104–108
30. Ghirlando R, Balbo A, Piszczek G et al (2013) Improving the thermal, radial, and temporal accuracy of the analytical ultracentrifuge through external references. Anal Biochem 440:81–95