Experimental Protocol for Sedimentation Velocity Analytical Ultracentrifugation

Andrea Balbo, Patrick Brown, and Peter Schuck

Last updated in January 2008

This protocol describes a typical SV experiment that can be used for proteins to study the number of species in solution, the oligomeric state, and the self-association and hetero-association properties. Optimally, a few hundred micrograms at greater than 90 % purity would be used: however, this much material is not necessary for some possible experimental configurations. Typically, it takes about 3 to 4 hours to prepare and start the SV experiment, and 2 to 12 hours run time of sedimentation, depending on the protein size.

Because of the tight integration of the experiment and the analysis, the considerations that enter the data analysis are also outlined as part of the protocol, as well as the typical workflow of the analysis. More detailed information on the background of the mathematical methods embedded in the software can be found at in the tutorial and references sections of our website. Other resources that can be found here are a flow-chart for AUC experiments, as well as tables containing a grid for buffer selection with different optical systems, configurations for SE and SV experiments with different optical systems, as well as AUC strategies commonly employed for selected typical problems. Finally, hands-on workshops regarding the practical planning, conduct and analysis of SV experiments are held semi-annually at the National Institutes of Health.

The protocols may have to be modified and adapted to the special system under study based on the principles above, and should be understood only as a guide which in our experience will be successful in the large majority of cases. The basic familiarity with the ultracentrifuge equipment and the cell components as provided through the manufacturer's instruction manual is required to follow the protocol. The manufacturer's website provides video instructions of the cell assembly and a description of the terminology.

0. Instrumentation needed

* Analytical Ultracentrifuge (Beckman Coulter) equipped with absorption optical scanner (ABS) and optionally Rayleigh laser interferometer (IF) imaging system

* Spectrophotometer (preferably dual-beam, any supplier) may be required for the determination of the protein concentration via UV absorbance

* Densitometer and viscosimeter (Anton Paar) may be required for measuring the solvent density if non-tabulated buffer components are used (tabulated values are available for most common buffers and salts)

* Computer (PC) equipped with software to perform the data analysis (this should not be a PC simultaneously running the ultracentrifugal data acquisition to avoid interference):

  (1) SEDFIT (version 11.1 or later) and SEDPHAT (version 5.02 or later) for transforming and analyzing the sedimentation equilibrium data;

  (2) SEDNTERP for calculating the protein partial-specific volumes and extinction properties, as well as buffer density;

  (3) general purpose software to store screenshots for documentation (MS Word, powerpoint or equivalent)

  (4) internet connection to consult the SEDPHAT getting started tutorials, the step-by-step tutorials for sedimentation equilibrium analysis, the command reference manual, and the SEDFIT/SEDPHAT search bar.

I. Planning the Experiment

1. **Consider the buffer conditions**. Phosphate buffers are transparent in the far UV and are suitable for both ABS and IF detection. For example, PBS (phosphate buffered saline) will work well. When using other buffers, see Tables 1 and 2. If the buffer conditions and the protein extinction properties do not govern the choice of the detection system, use the IF system. Make sure you know all the components of the buffer being used. Then, measure the density and viscosity of the buffer, at the actual AUC run temperatures, with the densitometer and viscosimeter, respectively, or in most cases these values can be calculated with software such as the program SEDNTERP. As a preliminary step, it might be necessary to examine sedimentation/optical properties of the buffer itself before analyzing the actual proteins, if for example, detergents are to be added, or potentially absorbing buffer components are present when using the absorbance optics.

2. **Familiarize yourself with the purification steps** of the protein sample leading up to the AUC study. SV detects aggregates generated by filter concentration or freeze/thaw cycles with higher sensitivity than most other techniques, which may interfere with the detection and analysis of *bona fide* protein complexes. The purification of a sample to be analyzed by size-exclusion chromatography (SEC, Chapter X) as the last step prior to AUC is strongly recommended. For refining the data analysis, it is useful to have the SEC elution profile at hand, as well as a standard SDS-PAGE gel. The amino acid composition of the proteins being studied is required to assess the partial-specific volume, using

SEDNTERP. If one or more of the proteins is glycosylated, a measurement of the carbohydrate content by mass spectrometry is advantageous.

3. **Buffer exchange:** In preparation for the IF system, bring the protein sample into chemical equilibrium with the reference buffer either by over-night equilibrium dialysis or by gel-filtration after exhaustive equilibration of the column. (If this is not possible, or buffer matching will be very difficult to achieve, for example, when working with very high concentrations of co-solutes, be prepared to model buffer signals as extra components.) For the ABS system, this step is not required. Have available at least 10 ml of the dialysate or running buffer of the column, for use in dilutions and as an optical reference buffer.

4. **Establish the amounts (volumes) and concentrations** of the stock proteins to be studied. The determination of concentration with a dual-beam absorption spectrophotometer is advantageous (scan from 210 to 350 nm) to assess possible aggregation and nucleic acid contamination: nucleic acid will be detected at 260 nm, aggregates as a broad signal in the range of 320 – 350 nm). Theoretical extinction coefficients can be calculated based on the amino acid composition using SEDNTERP. If the stock protein amount is limited for the ideal setup described in the following step, reduce the volume from 400 microliters to 200 or 300 microliters (sacrificing some hydrodynamic resolution).

5a. **To study self-association of a single pure protein:** Plan for 3 samples of 400 microliters each in a 3 – 4 fold dilution series. The highest concentration should usually not exceed 1 – 2 mg/ml to minimize non-ideal sedimentation, and the lowest concentration should be 2 – 3 fold above the detection limit. For example, for the IF system, use 2, 1, 0.3 and 0.07 mg/ml. For the ABS system, dependent on the protein extinction and not exceeding 2 mg/ml, use three samples at concentrations producing 1.2 OD at 250 nm, 0.5 OD at 280 nm, and 0.2 OD at 230 nm, respectively. The individual Beckman AUC instruments differ in their precision and reproducibility of the monochromator covering multiple wavelengths in the same experiment. A problem with reproducibility can be detected by large jumps in the wavelengths for subsequent scans, causing instable amplitudes of the signals from the sedimentation boundaries. If a problem is observed, instead, chose 280 nm as the detection wavelength, and use samples at 1.2 OD, 0.5 OD, and 0.1 OD. (The OD numbers given here are for an optical pathlength of 1 cm such as would be commonly measured in the spectrophotometer. Taken into account is that the pathlength is 20% longer in the ultracentrifuge when using a 12 mm centerpiece, 1.2 OD in 1 cm pathlength would lead to a signal of ~ 1.4 OD at the highest concentration. Make sure the loading signal in the AUC is below 1.5 OD, and above 0.05 OD.) To permit the study of the self-association, the concentration range chosen should cover the expected K_D in order to determine the binding constant.

Use an expected value for the binding constants (if available from previous or other types of experiments) and use the SEDPHAT calculator functions for the mass action law to make sure that the different loading concentrations planned will lead to different *relative* concentrations of all species.

If the available sample amounts are limiting, plan for using 3 mm centerpieces for the cells at the highest concentrations. This will give 4fold lower signal, but also require a factor 4 less sample. These centerpieces can be used for either absorbance or interference data acquisition.

5b. **To study hetero-association** of two proteins ‘A’ and ‘B’ of already known (or absent) self-association properties: Plan one sample of each of the interacting proteins alone, and at least 3 samples of mixtures. All samples should have a final volume of 400 microliters. Different schemes for preparing the concentration series of A and B are common. However, especially for interactions with unknown stoichiometry, use a constant concentration of A at 3 – 5 fold of the expected K_D, and vary the concentration of B from 0.1 to 10-fold K_D.

Note that the IF system does not require a red calibration counterbalance during the experiment, allowing 4 and 8 samples to be run in the 4-hole and 8-hole rotors, respectively.* The rotor still has to be balanced. A red calibration counterbalance is required for the ABS system, reducing the number of sample to 3 and 7, respectively.

> * We have routinely found this method to work well on our instruments for ABS detection also, providing the delay calibration option is deselected after having first performed the calibration without protein sample in the same rotor. However, the utility of this option may be instrument-dependent and user's should therefore explore this with individual instruments. Feasibility of this method for ABS detection relies on capabilities of the AUC better than the 'official specs'. Critical also are the condition of the rotor magnet and the overspeed disk, as well as cell alignment.

As for self-associating systems, use an expected value for the binding constants (if available from previous or other types of experiments) and use the SEDPHAT calculator functions for the mass action law to make sure that the different loading concentrations planned will lead to different *relative* concentrations of all species.

Again, if the available sample amounts are limiting, plan for using 3 mm centerpieces for the cells at the highest concentrations. This will give 4fold lower signal, but also require a factor 4 less sample. These centerpieces can be used for either absorbance or interference data acquisition.

6. **Select the rotor speed:** For proteins or anticipated protein complexes with molar mass between 30 and 300 kDa, a rotor speed of 50,000 rpm should be chosen. For larger complexes, use a lower rotor speed (30,000 to 40,000 rpm), and for smaller proteins a higher rotor speed (60,000 rpm). Use a rotor temperature of 20 °, unless the known temperature-dependence of the interaction or the limited stability of the protein dictates another choice.

7. **Optionally: simulate the experiment** with the 'generate' function of SEDFIT, using the planned protein concentrations, known extinction properties and molar mass values of the protein species, sedimentation coefficients based on estimated frictional ratios (e.g., 1.3), different values for the K_D of the interaction, and a typical value of random noise (0.005 OD or fringes, respectively). This can help to predict if information on certain aspects of the system under study can be extracted under optimal conditions from the experiment, and allows the optimization of the loading concentrations and the rotor speed.

II. Preparing the Analytical Ultracentrifuge and the Cell Assembly

1. **Assemble the components of the cell assembly.** Video instructions at the manufacturer's website describe the terminology and the assembly. (Note that the mark of the windows should be aligned with the keyway of the window holder). Also consider the information on storage, cleaning and assembly in section IV. 3. below. Use charcoal filled Epon centerpieces. Make sure that the centerpiece and the windows are free of dust. In our laboratory the charcoal filled Epon centerpieces are routinely taken to rotor speeds up to 60,000 rpm without breakage despite the manufacturer's lower rating. (Caution: A leak will damage the Epon centerpiece at the high speeds – how to check for leaks is described in Step III.4.) Sufficiently high rotor speeds are important to ensure optimal resolution. For the IF system, use sapphire windows, and for the ABS system, use quartz windows. Grease the screw ring and its washer occasionally (we use Spinkote, from Beckman Coulter, wiping off any excess to leave only a virtually invisible surface film) to ensure the torque can be correctly adjusted to 120 – 140 inch-pounds. Exceeding the torque will cause breakage of the windows, while insufficient torque causes leaks and mechanical instability of the assembly. Check the reflection of the divider between reference and

sample sector for a tight seal: Hand-tightened it should show colorful Newton-rings and after torquing it should be solid black.

2. **Optional Test Run:** When studying very large particles that sediment so fast that there will not be sufficient time to adjust the optics before significant depletion has taken place, do a test run. Do a test run also when you are not sure about possible leakage and when you may need extra time to adjust the optics.

For the test run, fill both the reference and sample sector with 400 microliters of the selected buffer, using long tapered pipette tips. (If the cell is oriented such that the screw ring is facing you, the reference sector is on the left, the sample sector is on the right). Seal each cell by placing two red polyethylene plug gaskets over each filling hole, and hand-tighten the housing plug screws. Insert the cell assemblies into the rotor according to the manufacturer's instructions with a properly balanced counterbalance. Note the orientation of the cells (the screw ring facing up and the filling holes facing the center of rotation), and carefully align the cells and the counterbalance using the scribe marks on the cells and rotor. The alignment is critical to avoid convection during the experiment. If the cells fit too loosely in the rotor, make sure the filling hole plug is tight, and if necessary, insert a spacer (of the thickness and stability of, e.g., a human hair) between the cell and the rotor hole.

Then start a test AUC run. Insert the rotor in the rotor chamber, and install the optical arm according to manufacturer's instructions. Make sure the UV filter is not in the light path for UV ABS detection. The optics must be attached firmly so that the light source will be reproducibly aligned with the detection system. Push the vacuum button on the ultracentrifuge and wait until vacuum is established (< 100 'micron'). Check if the AUC has in recent previous runs undergone a radial calibration for the optical system to be used. This calibration has to be performed at low rotor speed (e.g., 4,000 rpm) to avoid significant rotor stretching. If necessary, perform the radial calibration: For the ABS system, select the absorption optics for cell 1 in a velocity mode, choose the radial calibration in the 'Options', and start a single scan. For the IF system, set up the laser interferometer configuration for the counterbalance, and proceed as specified by the manufacturer's instructions. (It is not advisable to do the calibration and the test run with the proteins loaded, as they will already sediment during the procedure, generating systematic errors in the data analysis. In this case, the run should be stopped, the proteins re-suspended by careful mixing of the cells, and restarted as described below.)

At the end of the test run, take out the cell assemblies, and re-insert the empty rotor and the optical arm into the AUC chamber. Set the temperature to 2 – 3 °C below the intended run temperature and evacuate (turn on the vacuum pump): if the vacuum pump is off, the rotor chamber will heat up. This will pre-equilibrate the temperature of the rotor and the AUC. Carefully remove the buffer completely from both sectors of the cell assembly (for example, by applying suction through a small tube connected to a vacuum flask).

3. **Temperature equilibration.** Because the sample cells must be temperature EQ to the run temp for at least one hour before starting the run, it is a good idea to begin with and maintain throughout setting up the run a cool AUC chamber along with the rotor and optics arm. A few hours before run time, even the night before, turn on the AUC and load the chamber with the run rotor, counterbalance (if being used), and optics arm. Set the AUC temperature to 2-3 degrees below the run temperature and turn on the vacuum. For example, if the run temperature is 20c, pre-cool to 17c. This will shorten the pre-run temperature equilibration time, since the rotor temperature (indicated when the vacuum drops below 100 microns) will already be closer to the run temperature. It will also cut down on large temperature gradients within the rotor.

4. **Prepare the scan settings file:** Enter the planned rotor, rotor speed and temperature. Set the centrifugation time to 'hold' mode. In the section for each cell, set the scan mode to 'velocity' and

check the optical system (IF or ABS). In the 'Options' menu, uncheck all options and set the overlay to 2. The 'Methods' menu controls the timing of the scans: choose no delay, request nominal time increments of 1 min, and 900 scans (300 for the ABS system). The real time for the sedimentation experiment will need to be adjusted depending on the system under study.

5a. **For the ABS system**, for each cell set 'Rmin' to 6.0 (this assumes 400 microliter samples) and 'Rmax' to 7.2, avoiding excess scanning but ensuring that both menisci and the bottom of the solution column are covered. Set the desired wavelength (e.g. 280). In the 'Details', choose no other wavelength, radial step size of 0.003 cm, 1 replicate, and the continuous scanning mode.

5b. **For the IF system**, for each cell go into the 'Details' parameter box and choose no blank scan subtraction, no alignment of data. Do not change the default pixel per fringe setting, and do not do the blank setup. (Blank scans are not needed in SV as the radial-dependent baseline can be computed with higher precision, see below.) The radial range as well as the fine-adjustment of the laser setup will be entered later, at the start of the run. Choose the 2 channel centerpiece, and enter a new directory name for data storage. Save the scan settings file.

6. At this point, the rotor and optics arm are in the chamber under vacuum and the temperature is reading 2-3 degrees below the actual run temperature. The scan file is set up and saved. Now, it is time to prepare the samples, load the cells and start the run.

III. Sample Preparation and Starting the Run

1. For each cell, **prepare the protein mixtures** and/or dilutions in separate Eppendorf tubes. If dilutions are necessary to reach the planned experimental concentrations, use the reference buffer. For slow interactions, consider an incubation period of several hours at the planned experimental temperature prior to starting the SV experiment (provided the proteins are sufficiently stable).

2. **Fill the cells:** If the cell is oriented such that the screw ring is facing you and the filling holes are on top, the reference sector is on the left, the sample sector is on the right. Fill the reference sector of the empty cell assembly with buffer, using a long tapered pipette tip, very carefully and fully aspirating the precise volume. Then fill the sample sector with the sample, using the same tip (this helps to prevent small volume differences that could occur when changing tips). Avoid air locks which would cause the solution to bubble over the top. For the IF optics, the volumes in the reference and the sample sectors should exactly match. Note that this cannot be visually adjusted, but can be achieved with reliable pipetting. When using the ABS optics only, the reference volume should exceed the sample volume by 5 – 10 microliters.

3. **Seal the cell assemblies**, and insert them into the rotor. Seal each cell by placing two red polyethylene plug gaskets over each filling hole, and hand-tighten the housing plug screws. Insert the cell assemblies into the rotor according to the manufacturer's instructions with a properly balanced counterbalance. Note the orientation of the cells (the screw ring facing up and the filling holes facing the center of rotation). Very carefully adjust the angular alignment of the cell assemblies using the scribe marks on the rotor and cell barrels, such that the walls of the centerpiece sectors are concentric with the rotor axis. The alignment is critical to avoid convection during the experiment. If the cells fit too loosely in the rotor, make sure the filling hole plug is tight, and if necessary, insert a spacer (of the thickness and stability of, e.g., a human hair) between the cell and the rotor hole.

4. **Insert the rotor** into the ultracentrifuge and mount the optical arm. Evacuate the rotor chamber and observe the vacuum. Once the pressure has reached a level of several hundred microns, leaks may

become evident: If a sudden rise of pressure by one or two hundred microns is encountered, it is usually a sign of sample evaporation, and the run should be stopped, the leaking cells cleaned, and re-filled. During this time with pressures above 100 micron, the temperature reading of the ultracentrifuge only imprecisely reflects the true rotor temperature, and it is usually advisable to adjust the temperature set point on the centrifuge control panel to a value below the current temperature reading, so as to maintain cooling of the chamber.

5. **When the pressure drops below 100 microns**, adjust the temperature set point to the desired temperature of the run. Observe the actual temperature reading and notice when the actual temperature coincides with the desired run temperature. Counting from this time, let the rotor temperature equilibrate an additional hour. This is important to avoid temperature driven convection during the experiment, which would introduce systematic errors in the results.

6. **To this point, the rotor is still at rest:** Indeed, any rotation prior to this point will lead to partial sedimentation and introduce errors in the data interpretation. The pressure gauge should read 0 microns. Make sure the rotor acceleration control is set to the highest value. From the control of the ultracentrifuge accelerate the rotor to reach the required rotor speed for the experiment. A slight drop in temperature is usually noticeable, resulting from adiabatic stretching of the rotor.

7. **Prepare for data acquisition:** On the computer controlling the data acquisition of the analytical ultracentrifuge, retrieve the scan settings file (see above). Confirm the desired rotor speed and rotor temperature.

8. For the IF system, for each cell select the '**Details**' section to adjust the '**Laser Setup**' according to the manufacturer's instructions. Some details are dependent on the particular AUC models. The fine-adjustment of the laser delay should adjust the timing of the laser pulse to the center of the visible compartment, which can be achieved by slowly exploring the limits of the range of delay angles that produce a fringe pattern, followed by centering the delay. Some small deviations from this optimal position are possible in order to avoid clearly visible surface features. For the camera exposure, the overall goal is to have a fringe pattern that has a sinusoidal (not rectangular) vertical intensity profile, with bright and dark stripes of similar width, and not oversaturating or underexposing the camera.. After closing the 'Laser Setup' menu, enter the radial range for which the IF data will be collected. Make sure the minimal (inside) radius is in the air-to-air region and the maximum (outside) radius is well below the bottom of the cell. In this way, the information from the complete solution column will be stored.

These adjustments should be performed as quickly as possible, remember, the rotor is spinning at run speed and the molecules are sedimenting. It is possible to pre-adjust these setting and in addition, test for leaking cells, in a test run using water filled cells (see above), which requires only make final minimal adjustments when the samples are loaded. In the absence of a test-run, it is also possible to do this in two steps: a pre-adjustment while the rotor is still accelerating, followed by final fine-adjustments at full speeds.

9. **Start the data acquisition** with the 'Start Methods Scan' button. Verify that the speed and temperature settings of the computer and of the ultracentrifuge controls match. Observe the first several scans, which will take approximately 10 minutes, and apply corrections to the settings of the ABS or IF settings if necessary (usually, stopping and restarting the scans can be done while the centrifugation continues). The goal of the data acquisition is to observe the sedimentation from the beginning of the first deviation from uniform loading concentration. Even the initial partial depletion at the meniscus contains significant information (on the average molar mass, as well as the presence or absence of larger species).

10. **Preliminary Analysis During The Run:** When the sedimentation boundary has cleared the meniscus, which might be the case after an hour (although this is strongly dependent on protein size)

sufficient information has accumulated for a first, very preliminary analysis (see below). At this point gross discrepancies from the expected sedimentation behavior can be detected and, if necessary, corrections in the experimental conditions can be made. If the run is stopped at this stage, the samples may be re-used after resuspending them uniformly (see IV.2.), for example, by carefully turning the rotor perpendicular to the centrifugal axis of rotation, to ensure uniform distribution before restarting the run at Step III.4. (a new temperature equilibration period will be required).

IV. Stopping the Run and Cleaning Cell Assemblies

1. **The end of the run:** Continue the run until no further sedimentation is visible. (With the IF optics, the assessment whether or not sedimentation is still visible may require a preliminary data analysis while the run is still going, in order to eliminate the systematic noise contributions and visualize macromolecular sedimentation only.) For example, for proteins large enough to form a migrating sedimentation boundary, the run should be stopped when the trailing edge of the diffusionally broadened boundary disappears. Dependent on the protein size and rotor speed, this may take typically between 2 to 12 hours. When signals from small degradation products or buffer salts are present, it may be necessary to conduct the run somewhat longer such that the small species' sedimentation can be better characterized (not biasing the analysis of the species of interest)

2. **After stopping the run**, the samples may be recovered from the cell assembly, but it should be noted that in many cases the protein may be altered or aggregated from the extremely high local concentration near the bottom of the solution column. Although some proteins can be successfully resuspended for further study, this is not recommended in general.

3. **Cleaning** of the windows and centerpieces is necessary. Different procedures are possible, one of them is this: First rinse the windows and centerpieces with distilled water, and then sonicate for 30 minutes in a warm solution of diluted, low residue detergent (such as Conrad 70). After several rinses with water, they should finally be rinsed in either methanol or ethanol. (Caution: Familiarize yourself with the appropriate pre-cautions necessary for working with these solvents.) The centerpieces are dried in air, avoiding contact to their surfaces. The cell components can be reassembled for storage. Before assembly, the windows are rinsed with methanol and wiped with optical grade tissue (or Kimwipes) to remove dust. The outer surfaces of the centerpieces that will be in contact with the windows can also be wiped with optical grade tissue slightly moistened with methanol before reassembly. The assembly follows the manufacturer's instructions. The assembled AUC cell can be stored wrapped in optical grade tissue, but should be retorqued before use. It is advantageous to maintain the components together as a unit, and not to interchange parts during cleaning and assembly. This helps to identify, trace, and eliminate malfunctioning parts that may obviously or subtly degrade the data quality.

V. Data Analysis

Before conducting a data analysis for the first time, it may be useful to also go through the step-by-step guide of the website, which has screenshots of a simple c(s) analysis.

Calculating a c(s) sedimentation coefficient distribution

1. **Loading Scans:** Transfer the data to the computer dedicated to data analysis. Initially, load a data subset of ~50 scans in SEDFIT, covering the complete sedimentation process in equal time-intervals. This can be achieved, for example, by loading every 10^{th} scan and counting the profiles that show a

sedimentation boundary; alternatively, the choice of scans can be refined after a preliminary analysis. When the scans are loaded, they are shown with the color temperature (blue to red) scaled proportionally from first to the last scan. The transition green to red is the color of scans roughly in the middle of the loaded set. It is not advisable to load too many late scans where the sedimentation is already over: this situation can be diagnosed by seeing only blue or blue/green boundaries, and the lines from the later scans being compressed near the baseline. (To pick the right number of scans may require loading the scans first to survey which of the scans is the last with meaningful sedimentation information. Note that moving the mouse across the scan will reveal its filename and number.)

2. **Specify graphically a tentative meniscus and bottom position**, as well as fitting limits that exclude the region of artifacts close to the meniscus and, if possible, the region of back-diffusion close to the bottom. (Back-diffusion has to be included into the fit if it extends far into the solution column, for example, for small proteins < 5 kDa and for peptides.) Optical artifacts from reflections and refraction at the ends of the solution column dominate the signal typically within 0.05 cm of the meniscus and a slightly larger range at the bottom. Setting the meniscus (red) and bottom (blue) can be achieved with left double-clicking the mouse while the control key is pressed, and the fitting limits (green) can be set by left double-clicking without control-key. Alternatively, existing limits can be adjusted by grabbing and dragging with the mouse. A detailed software tutorial for this procedure can be found in the SEDFIT 'Getting Started' section of its homepage. In particular, the websites for loading data and the step-by-step practical application will explain this step.

3. **First c(s) analysis:** As a first overview of the sedimentation properties of the sample, select the $c(s)$ distribution model, and specify in the parameter box the $\bar{v}$ of the protein, as well as the buffer density and viscosity at the experimental temperature. When working with IF optics, the systematic time-invariant noise (TI, the radial-dependent baseline offset) will be switched on by default, as well as the radial-invariant (RI) offsets. This requires no action for IF data. With the ABS optics, switch on the TI noise. A video tutorial is available at the geting started site of the SEDFIT homepage.

4. **c(s) parameters:** Set the resolution to 100, the minimum s-value of the distribution to a value of 0.5, and the maximum s-value to a value well above the maximally expected value. For proteins of < 100 kDa, a reasonable initial value can be 10. Enter a starting value for the frictional ratio f/f_0 of 1.3 for globular proteins, or 1.5 for glycoproteins or known highly asymmetrically shaped proteins. Specify f/f_0 and the meniscus to be fitted. If back-diffusion is part of the analysis, also select the bottom position to be fitted. Set the confidence level to a value of 0.7.

5. Use **the 'Run' command** to get an initial assessment of the fit. Subtract all systematic noise components (control-N or use the function in the 'Display' menu) and select the residuals bitmap to be displayed (if it is not already showing). You should see the data and fit in the upper graph, the residuals bitmap and graph in the middle, and the distribution in the bottom graph. If the data show a large number of scans with complete depletion throughout the solution column, or if the sedimentation boundary does not appear to sediment through the entire solution column, reselect a new data subset.

6. **Some initial adjustment may be needed:** If the distribution is non-zero at the minimum or maximum s-value, chose a higher value for s-max and a lower value for s-min, respectively, and execute the 'Run' command again. Repeat this until the $c(s)$ distribution at the maximum and minimum s-value vanishes and all peaks are displayed within the distribution range. If there is a large half-peak at s-min, this usually indicates signals from buffer. In this case, it is possible to get a better fit with the model of c (s) with an extra discrete species, using the default values for s and M of the discrete species to describe sedimentation signal of buffer salts, and checking these values to be optimized later.

7. **Refine the model by using the 'Fit' command**, which will optimize the weight-average f/f_0 value and the meniscus position. A visualization of this process can be found in the step-by-step tutorial at the 'getting started' website of SEDFIT.

The initial interpretation and assessing the quality of the fit

8. **Assess the quality of the fit** by considering the rms deviation (this should be well below 0.01, except for very high loading concentrations, or noisy ABS data), the randomness of the superposition of the residuals, and the structures in the residuals bitmap. Another factor is the best-fit meniscus position which should be consistent with the graphical display of the range of optical artifacts (if it is not, restrict the range of minimum and maximum positions to reflect the experimental uncertainty and re-fit).

9. **If a good fit is achieved**, the *c*(*s*) distribution displays the most parsimonious distribution that is consistent with the data under the assumption that all species are stable on the time-scale of the sedimentation. If this is true, the peaks reflect sedimenting species (on the level of resolution provided by the signal/noise ratio of the data). The assessment whether some detailed features are statistically significant can be made, for example, by re-executing the analysis ('Run' command) with a higher confidence level.

10. **If no good fit is achieved, there are many possible reasons**. A wizard function may indicate possible problems. Among the most common are: 1) residual convection (for example, this would typically produce very large systematic residuals in the first few scans close to the meniscus). Take a look at the temperature plot in Sedift to see if the temperature of the run fluctuated significantly after starting acquisition. Explore the possibility of the cell being misaligned in the rotor or of scratches along the inside walls of the centerpiece. 2) with IF optics, a chemical mismatch of the buffer and the sample solvent. (To correct this, excess small Mw solutes like buffer salts can be taken into account as a separate species in the '*c*(*s*) with 1 discrete component' model with an apparent *s* of ~ 0.2 – 0.3 S, and an apparent *M* of ~ 200 – 300 Da). 3) with IF optics, a mismatch of the volume of sample or reference may have occurred (this can be diagnosed by the observation of diagonal sections of the experimental scans between sample and reference meniscus, pointing down to the sample meniscus). To address this problem at the stage of the data analysis, use the buffer mismatch model of SEDFIT. 4) A fast chemical reaction may occur during the sedimentation process. In this case, the analysis can continue: this factor will be more apparent in the comparison of the samples at different loading concentrations, see below. If the fit is not good, it is useful to examine the bitmap and the residual plots, and to determine if the deviation originates from certain regions of the cell or if it persists systematically migrating with the sedimentation boundary, which may be helpful to differentiate between technical problems in the detection system (e.g., local loss of fringe contrast due to scattering), and systematic errors in modeling the migration of the proteins. 5) The standard assumption of a single weight-average frictional ratio representing the distribution well does not apply within the errors of the experimental data. This is uncommon, but can be addressed either by (1) using the model for bimodal c(s), such that two different f/f_0 values are used for different segments of c(s) – this is promising only if multiple boundaries corresponding to the different s-regions are clearly visible from the data; or (2) by supplying SEDFIT with more detailed information about the relationship between s and M, if available, using one of the "c (s) with prior knowledge" models, such as that of a worm-like chain, proteins with conformational change, or that using a freely user-defined table.

11. **The f/f_0 value** displayed after the fit should also be noted: It 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/sec) chemical reactions.

12. **Press control-M for obtaining the molecular weight** estimates associated with all peaks. This will produce buttons in the vicinity of each peak. Press the button of a peak of interest, and the coloring in the sedimentation profiles window will change to highlight the signal contributions of that species. Also, a box will appear with some available information about the estimated average properties of the particle within that peak, including molecular weight, weight-average s-value, and signal contribution in

units of loading concentration.

13. **Document the fit and the parameters** displayed in the SEDFIT window (for example, by copying and storing an image of the SEDFIT window), and copy the distribution table into a spreadsheet for further analysis. From "Display" select "show peak MW (in cS)". Note the MW value appears above each peak. Click on each MW box to display an Integration window containing MW, weight-average *s*-value, S, and concentration in signal units. As described in the introduction to SV above and in more detail in Appendix 1, this can give only an orientation for the corresponding molar mass values. It is frequently within 5 – 10 % of the correct molar mass if there is a clearly defined main peak, and the data quality sufficient to establish a well-defined determination of f/f_0.

14. Repeat steps V.1 to V.12 for all cells. The obtained c(s) curves will be the basis for the further analysis.

Refining the analysis of independently sedimenting species

15. **Bayesian pure species analysis**: Assuming each peak represents a monodisperse, independently sedimenting species, use the Bayesian analysis model to implement this hypothesis into the regularization. This is achieved by pressing control-X, which will use the automatic integration to locate the best estimate for the single species, and then run another c(s). Finally, a notepad window will appear that documents how much signal contribution was found in each of the sharp peaks, as well as outside these peaks. The latter may represent impurities, or indicate that the hypothesis of hydrodynamic monodispersity is not fully consistent with the data

16. **Export the data to SEDPHAT**, saving an xp-file containing the experimental data files and related information, and switch to the hybrid discrete/continuous model.

17. **Molar mass values in SEDPHAT:** Under the assumption that the $c(s)$ peaks represent stable sedimenting species, substitute the peaks with discrete species with the same *s*-value and an estimated molar mass (for example, as derived from the automatic integration in Step V.12) surrounded (with a spacer of ~ 1 S) by segments of continuous distributions. Fit for the molar mass and *s*-value of the discrete species.

18. **Hydrodynamic shapes:** If the molar mass of a species can be identified as that of a known molecular species, such as the monomer, or a particular oligomer, use the calculator function of SEDFIT to calculate axial ratios of hydrodynamically equivalent ellipsoids, using the *s*-value, along with the known molar mass (from sequence or mass spectroscopy) and $\bar{v}$. If the *s*-value was determined in SEDPHAT, buffer corrections are already applied, and the buffer density in the calculator function should be that of water at 20°C (0.99823 g/ml) and the same for the viscosity (0.01002 Poise). In contrast, SEDFIT normally does not do provide buffer corrections (unless requested with the specific calculator function) and the experimental values should be entered.

19. **Global analysis:** For species that cannot be readily identified, more detailed information may be obtained from global analyses of the data from different cells, and/or from different rotor speeds. To accomplish this, load several xp files (one from each cell) into a single SEDPHAT window and search for a model that globally fits all data sets.

Detecting Interactions

20. **In SEDFIT, integrate (control-I)** over the s-range comprising all species possibly participating in an interaction. This will give you the weight-average s-value. For the data from all cells at different

concentrations, if the weight-average s-value does not change with concentration, an interaction is absent or too weak to be detectable. If the weight-average s-value increases with concentration, there is a binding reaction taking place. If the weight-average s-value decreases with concentration, there is repulsive non-ideality taking place. Non-ideality is usually observed only at high concentrations > 1 – 2 mg/ml or with very elongated molecules. (There is the possibility of combined non-ideality and binding reactions, in which case the weight-average s-value could be constant and decrease.)

21. **Plot the superposition of the *c*(*s*) distributions for all cells**. Alternatively, load the xp-files of all experiments into SEDPHAT and execute the hybrid discrete/continuous model with a single continuous segment covering the entire range of s-values. This will display all independent *c*(*s*) distributions superimposed. Decide on the likelihood that one of the following different types of characteristic behavior applies, taking into account that the distributions will naturally broaden with lower concentration: a) peaks that are proportional in area to the loading concentration and do not change in position; b) peaks that are constant in position but different in area ratio; c) peaks at changing positions.

22. Option a: **Constant peak positions with proportional areas** (allowing for some broadening within the signal-noise limits of the detection*) would be expected for a protein mixture in the absence of interactions. In this case, a global analysis can be performed identifying the *c*(*s*) peaks with discrete species with global *s* and *M* valid for all data, and with surrounding continuous segments for trace impurities that may be different for different cells. This is the most rigorous determination of *s* and *M* by SV, but relying on the mono-dispersity of each species.

*A refined approach is to use the c(s) distribution obtained at the highest concentration as a prior expectation in the analysis of the lower concentration data, and execute a Bayesian analysis. If there is any deviation in terms of additional peaks or peak shifts, there is some interaction present.

23. **Option b and c - Interactions**: Peaks that are constant in position but different in area ratio indicate a slow association. In contrast, changing peak positions indicate associations which are fast on the time-scale of sedimentation. A first back-of-the-envelope order of magnitude estimate of the K_D of the interaction can be obtained as the concentration range with the steepest change. In particular for small proteins, where diffusion is dominating the sedimentation profiles, it can be useful to consider – if available – independent information from other techniques on the reaction kinetics (stopped flow methods, optical biosensors, gel permeation chromatography). For interactions, the first and most important assignment is the reaction scheme. For slow interactions, but not for fast interactions, the *c*(*s*) analysis above may give valuable insights in the number and size of participating species. Also consider the results on the average molar mass values measured in SE at different loading concentrations, and assess if the covered range of average molar mass values is consistent with the population of free and complex species expected for a particular reaction scheme.

Determining binding parameters for interacting systems in SV

24. **Weight-average s-values:** After identifying the presence of protein interactions in the preceding Steps V.18 – 20, binding parameters, such as the binding energy, and the gross shape of the complex can be determined. Independent of the reaction kinetics, the weight-average *s*-value that was determined in Step V.12 from integration of the entire distribution of the associating species (excluding non-participating impurities, such as aggregates or small molecular weight fragments) can be analyzed in the form of an isotherm of $s_w(c)$ as a function of the loading concentration/composition. This can be done, for example, by generating a two or three-column ASCII file and loading the isotherm in SEDPHAT (see this website for further information: http://www.analyticalultracentrifugation.com/sedphat/isotherm_analysis.htm). After specifying the

interaction model, the $s_w(c)$ isotherm can be fitted, optimizing the *s*-values of the individual associating species, as well as the equilibrium constant(s). For self-associations, any prior knowledge on the *s*-values of the smallest species or a complex derived either from available crystal structures and hydrodynamic bead modeling, or from experiments under solvent conditions promoting or prohibiting the interaction can be highly valuable for increasing the precision of the K_D of the interaction. This is true, in particular, if a very large concentration range (and thus the plateaus of the isotherms) cannot be reached in the SV experiments. For hetero-associations, the *s*-values of the individual components can be fixed to those determined from separate experiments.

25. **More detailed analysis strategies:** If the isotherm model leads to an acceptable fit of $s_w(c)$ with realistic estimates of the *s*-values of the free species and the complexes (this can be assessed via the calculator functions of SEDFIT), a more detailed boundary model may follow. (If not, alternative models for $s_w(c)$ should be considered.) This allows determination of more precise binding constants and s-values (hydrodynamic shapes) of the complexes. The first approach exploits the peak-structure of c(s) to gain more information. This is a robust and quick approach. Finally, for the most detailed analysis of the sedimentation boundary shapes, direct Lamm equation modeling can be performed. It is useful for the latter to do the isotherm analysis, first, in order to get good starting values for better convergence of the fit.

26. **Isotherm analyses:** For the isotherm analysis of the peak structure obtained in c(s), we need to distinguish again slow from fast interactions (see above). For slow interactions, each species shows its own peak always at the same s-value but at a concentration-dependent magnitude. From integration, read the signal magnitude of these peaks and write them into a text-file, along with the loading concentrations, in the required format for the isotherm (see website). Then load all available isotherm files (e.g., from different signals or different runs) as population isotherms into a single SEDPHAT window, supply each with the required information, such as extinction coefficients and optical pathlengths, and fit the isotherms with the appropriate model. For fast two-component heterogeneous interactions, there should be two boundary components – the undisturbed boundary and the reaction boundary (the latter being at a higher s-value and usually shifting with concentration). Identify and integrate each peak from all concentration data, and assemble the signals along with the loading concentrations into a text-file (see website for formatting requirements). Load all available isotherms into a single SEDPHAT window as Gilbert-Jenkins isotherms, supply the respective ancillary information, and fit with the appropriate model.

27. For the most detailed boundary shape analysis with **direct Lamm equation modeling**, sequentially load the xp-files of all SV experiments of the interaction into a single SEDPHAT window. Supply the molar starting concentrations, and add information on the molar extinction coefficient for each experiment (save the xp-file with this new information). Select the model, enter the starting values for the species *s*-values and the equilibrium constant from the isotherm analysis, and estimate the chemical off-rate constant (for example, $\log_{10}(k_{off}) = -3$ for rapid interactions, or -4 to -5 for slow interactions on the time-scale of sedimentation, as assessed above.) Fit this model, first floating (i.e. allowing the optimization algorithm to adjust) only the starting concentrations, but subsequently also floating the equilibrium binding and reaction rate constants and the species *s*-values. This will exploit the complete information from the boundary shapes, and an estimate for k_{off}, if the reaction rate constant is between 10^{-2} and 10^{-5}/sec. Note that this last stage may not be possible or advisable if impurities contribute significantly to the sedimentation profiles. The quality of the fit, considering both the rms deviation and the residuals distribution, should be taken as a criterion for the validity of this model.

28. For any of the analysis methods, **error analyses** should be applied to the quantitative analysis of *s*-

values, molar masses, and binding constants. This can be done by Monte-Carlo techniques, or by mapping projections of the error surface using F-statistics, as specified in the software manual. An example for using the error surface projection method is given in the SE Protocol Step IV.11.