Analytical Ultracentrifugation in the Study of Protein Self-association and Heterogeneous Protein-Protein Interactions:

Protocols for Velocity and Equilibrium Sedimentation

Andrea Balbo and Peter Schuck

Protein Biophysics Resource, Division of Bioengineering & Physical Science, ORS, OD, National Institutes of Health, Bethesda, Maryland 20892

This protocol is part of an introduction to analytical ultracentrifugation for the study of protein-protein interactions. The introduction is to appear in Cold Spring Harbor Laboratory Press, Protein-Protein Interactions: A Molecular Cloning Manual. A preprint of this work, as well as this protocol, and additional web resources can be found at

http://www.nih.gov/od/ors/dbeps/PBR/AUC.htm

Address for Correspondence:
Peter Schuck
NIH/DBEPS
Bldg 13, Rm 3N17
13 South Drive
Bethesda, MD 20892
phone: (301) 435-1950
fax: (301) 480-1242
email: pschuck@helix.nih.gov

Sedimentation Velocity Experiment

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. 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 on the websites of the analysis software packages. Importantly, they also contain more detailed information and tutorials on their use. 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. A flowchart for the described course of the experiment is described in Figure 9. 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 website http://www.beckman.com/resourcecenter/labresources/sia/cellassy_video.asp provides video instructions of the cell assembly and a description of the terminology. 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.

MATERIALS

Special Instrumentation

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

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

Densitometer and viscometer (Anton Paar) may be required for measuring the solvent density and viscosity 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: SEDNTERP (freeware from www.jphilo.mailway.com) for calculating the protein partial-specific volumes, extinction properties, hydrodynamic shapes, and buffer properties; SEDFIT and SEDPHAT (freeware from www.analyticalultracentrifugation.com) for analyzing the sedimentation data; a spreadsheet program for plotting data and storing distribution data

METHOD

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 do not govern the choice of the detection system, use the IF system. Measure the density and viscosity of the buffer with the densitometer and viscosimeter, respectively, or calculate them with SEDNTERP.

2. Familiarize yourself with the purification 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 (see Appendix 2).

3. 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. For the ABS system, this step is not required. 10 ml of the dialysate or running buffer of the column, respectively, for use in dilutions and as optical reference buffer.

4. Establish the amounts and concentration of the proteins to be studied. The determination of the 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 1.2, 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 are for an optical pathlength of 1 cm such as measured in the spectrophotometer. Taken into account is that the pathlength is 20% longer in the ultracentrifuge, leading 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.

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.)

6. 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

1. Assemble the components of the cell assembly. The video instructions at http://www.beckman.com/resourcecenter/labresources/sia/cellassy_video.asp 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 II.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 (wipe 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. 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.

3. 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.)

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.

6. Choose the 2 channel centerpiece, and enter a new directory name for data storage. Save the scan settings file.

7. If the vacuum in the rotor chamber is < 10 micron, accelerate the rotor to full speed.. For the IF system, pre-adjust the camera exposure and delay (see step III.8 for details). Start the scanning. Take at least one scan of each cell to verify a visible meniscus for both liquid columns in the sample and reference sector. If one or both menisci are missing, this will indicate the presence of a leak (see Step II. 1.) This stage of the test run can also be used to pre-adjust the timing and exposure parameters of the IF optics in the 'Laser Setup' menu (see Step III.8. below). Stop the scan.

8. Stop the 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 f°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).

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 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 (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, the reference volume should exceed the sample volume by 5 – 10 microliters.

3. Seal the cell assemblies, and insert them into the rotor (see Step II.2 above). 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.

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. On the computer controlling the data acquisition of the analytical ultracentrifuge, retrieve the scan settings file (see above). Enter the desired rotor speed and rotor temperature.

8. For the IF system, for each cell select the 'Details' section to fine-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.

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. 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. Continue the run until no sedimentation is visible. 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. For example, in Figure 2 the main boundary has just migrated completely through the solution column, while for an optimal characterization of the slowest one the experiment should continue somewhat longer.

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 sonicated 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

Calculating a c(s) sedimentation coefficient distribution

1. 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 10th scan and counting the profiles that show a sedimentation boundary; alternatively, the choice of scans can be refined after a preliminary analysis.

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. A detailed software tutorial for this procedure can be found at www.analyticalultracentrifugation.com/getting_started.htm. In particular, the websites for loading data and the step-by-step practical application will explain this step.

3. As a first overview over 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.

4. 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 (in the 'Display' menu) and select the residuals bitmap to be displayed. This will show 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. For example, the data in Figure 2 are not entirely describing the sedimentation of the smallest species and it would be advantageous to include later scans.

6. 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, after lowering the s-min value, the distribution grows at the minimum *s*-value, this indicates a mathematical correlation of very slow sedimentation with the baseline parameters which is not detrimental and can be ignored (over-parameterization of this range of *s*-values is only problematic if the very small species themselves are subject of the analysis). In this case, reset the minimum *s*-value to the previous value.

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 (see V.2 above).

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. 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). 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). 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.

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. 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. Integrate the peaks in $c(s)$ and the complete distribution (excluding the value at s-min) to measure the loading concentration corresponding to each peak, as well as the weight-average s-value.

13. Repeat steps V.1 to V.13 for all cells. The obtained c(s) curves will be the basis for the further analysis.

Refining the analysis of independently sedimenting species

14. Following the hypothesis that all species are stable on the time-scale of sedimentation, switch to

$c(M)$ to read the molar mass that would correspond to the main peaks assuming the validity of the weigth-average f/f_0. 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.

15. 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.

16. 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 $c(M)$ in V.14), surrounded (with a spacer of ~ 1 S) by segments of continuous distributions. Fit for the molar mass and s-value of the discrete species.

17. 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, enter the s-value in SEDNTERP, 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 should be switched off in SEDNTERP. In contrast, SEDFIT does not produce solvent corrected s-values, and need to be corrected to water, $20°$C conditions in SEDNTERP. The 'Results' menu in SEDNTERP will give the frictional ratio of this particular species (with a much better precision than the weight-average f/f_0 from $c(s)$ provided the molar assignment is correct), the Stokes radius, and the dimensions of hydrodynamically equivalent ellipsoids. 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.

   Detecting Interactions

18. 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.

19. Option a: Constant peak positions with proportional areas 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. The SEDNTERP analysis of hydrodynamic shape (without further buffer correction) can follow.

20. Option b and c: 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

21. 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. 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.

22. If the model 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 equivalent hydrodynamic shape analysis in SEDNTERP), a more detailed boundary model may follow. (If not, alternative models for $s_w(c)$ should be considered.) For this analysis, load the xp-files of all SV experiments of the interaction in SEDPHAT. 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.

23. 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.

Sedimentation Equilibrium Experiment

This protocol illustrates a SE study of a heterogeneous interaction between two proteins 'A' and 'B' forming a reversible complex. The goal is the determination of the binding constant and the binding stoichiometry. It includes the determination of the molar mass of a non-interacting protein (or stable protein complex), and can be easily adapted to the study of protein self-association. Again, the details of the practical steps are given in conjunction with the workflow of the data analysis, which is closely tied to the experimental configuration. It will be assumed that a SV study of the same proteins and their mixture has taken place before, using the SV protocol above.

MATERIALS

Special Instrumentation

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 (Anton Paar) may be required for measuring the solvent density and viscosity 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: SEDNTERP (freeware from www.jphilo.mailway.com) for calculating the protein partial-specific volumes and extinction properties, as well as buffer density; MATCH (freeware from the website http://vm.uconn.edu/~wwwbiotc/uaf.html) for verifying that equilibrium has been attained; SEDFIT and SEDPHAT (freeware from www.analyticalultracentrifugation.com) for transforming and analyzing the sedimentation equilibrium data; general purpose software to store screenshots for documentation (MS Word, powerpoint or equivalent)

METHOD

I. Planning the Experiment

1. The molar mass as well as the extinction coefficient at 280 nm and the $\bar{v}$ values can usually be calculated from a known amino acid sequence (via SEDNTERP). For glycoproteins, the determination of the carbohydrate mass fraction by mass spectrometry is desirable, because a good estimate of the partial-specific volume can be obtained. For proteins that are known not to self-associate, the buoyant molar mass (essentially the $\bar{v}$) values of the individual proteins will be measured in this protocol.

2. Conduct a SV experiment to assess the purity of the sample. Generally, purity of better than 95% is required for SE. Small molar mass impurities (e.g. with masses in the range of a few percent of the proteins of interest), which frequently go undetected by SDS-PAGE, will complicate the analysis and should be removed, if possible. Irreversible small oligomeric aggregates will generally also limit the precision of the binding parameters to be measured. Therefore, size-exclusion chromatography is highly recommended as the last preparative step. Contamination with very large aggregates, distinctly larger than the molar mass of the complex, can usually be tolerated. The SV run will also allow the determination of the oligomeric state of each protein in solution and the detection of reversible self-association of the two components. It will also permit estimation of the

binding constant (K_D* – e.g., an empirical estimate of the concentration over which the $c(s)$ changes the most) and the size of the complexes formed. This will greatly help in planning the sample concentrations and the rotor speeds for the SE experiment.

3. The choices for the buffers and their relationship to the optical system are very similar to those described in I.1. of the SV protocol, and in Tables 1 and 2. They are slightly less restrictive with regard to components that can create density gradients at the higher rotor speeds of SV: For example, a small percentage of glycerol can be tolerable in SE using the ABS optics, provided the potential effect on $\bar{v}$ from preferential protein hydration and the extended equilibration time is taken into account. As in the SV experiments (Step I.3.), a chemically identical reference buffer is required when using IF optics. Measure the buffer density or use SEDNTERP to calculate the appropriate value.

4. Choose the rotor and centerpieces: If an eight-hole rotor is available (this is preferable), reserve one cell for each of the individual proteins to run as a control, and use the 5 remaining cells for mixtures. With a four-hole rotor, consider using three 6-channel centerpieces to increase the number of samples. If these options are not available and the individual components have to be studied in a separate run, note that small differences in the purity or presence of irreversible aggregates from different preparations or from different time-points can be confused with and correlate with the detection of putative heterogeneous complexes. In practice, this can be a problem, particularly for weak interactions.

5. Select the sample concentrations to span a range that will generate different relative populations of free and complex species: For hetero-associations with unknown stoichiometry, it can be advantageous to choose a constant concentration of 'A' several-fold greater than the estimated K_D*, and to titrate constant 'A' in different samples with variable amounts of 'B' to generate a range of molar ratios comprising the suspected possible stoichiometries. (The constancy of loaded amounts of 'A' can later serve as a powerful constraint on the data analysis.) This may be combined with a second set of A and B at the same molar ratios, but at a concentration of 'A' several-fold below K_D*. In contrast, if the stoichiometry is known, it can be advantageous to make a stock mixture at a concentration of > 10fold K_D* with a molar ratio equal to the complex stoichiometry, and to generate the different samples by serial dilution to span a large range of concentrations. (Again, the constancy of the molar ratio generates a useful constraint for the data analysis.) If the available maximal sample concentrations are too low (e.g. only one tenth of K_D*) adjust the following protocol to use a longer solution column, tolerating the longer equilibration time but taking advantage of the concentrating effect of the centrifugal field. The dynamic range of the optical system (see Table 1) must be considered for selecting sample concentrations.

6. The key for selecting the optical system is to generate SE profiles in such a way that the contributions from free species, which attain relatively shallower exponential distributions, and complex species, which generate relatively steeper exponentials, can all be detected and distinguished (see the introduction to SE and Appendix 1): Proteins vary greatly in the number of tyrosine and tryptophan residues, and therefore in their extinction spectrum in the far UV. Because the ABS system can take advantage of this to help distinguish between the components 'A' and 'B', ABS with multi-wavelength analysis is usually preferred. Obviously, this technique will be more powerful for protein/nucleic acid interactions, or if one of the interacting components has an intrinsic chromophore in the near UV or VIS or if an extrinsic label can be attached. This is frequently very useful if the molar masses of 'A' and 'B' differ by less than 20% or more than 80%. (An alternative in this case can be SE at multiple rotor speeds with mass conservation analysis.) For the choice of the protein concentrations for ABS detection, if consistent with the order of magnitude of K_D*, maximize the dynamic range of the ABS system at the wavelengths of 230 nm, 250 nm, and 280 nm: make an equimolar stock mixture of 'A' and 'B' and prepare a dilution series

to obtain samples with total absorbencies of 0.1 OD_{230}, 0.2 OD_{280}, 0.5 OD_{280}, and 0.5 OD_{250}. Add a second sample at 0.5 OD_{280} with a 2:1 stoichiometry of 'A' and 'B'. Note that the centrifugal field will create a gradient locally decreasing or increasing the concentrations several-fold. For the two individual protein samples to be studied in the same run as controls, choose a concentration of 0.5 OD_{280}. This will permit the determination of their extinction coefficients at 230 and 250 nm. However, the ABS system may be prohibited by the solvent conditions, or it may not provide a sufficient dynamic range of protein concentrations for the detection of weak interactions. Also, one or both of the proteins under study may not have a sufficiently high extinction coefficient to generate signals > 0.05 OD at any of the absorbance wavelengths permitted by the buffer system. In this case, the IF system can be used, with the considerations for sample concentrations as outlined above (Step I.5.). Maximum information can be gained if the IF system is used in conjunction with the ABS system, which improves the distinction between the two proteins in a multi-signal analysis. Note that the use of the IF system for SE will require the mechanical stabilization of the cell assembly (Step II.2a.). In the combination of IF with ABS data, in particular if the solution volumes and rotor speeds are used as suggested in this protocol, the radial-dependent baseline can be computationally determined and water blanks are not essential.

7. The choice of the sample volume and rotor temperature is dictated mainly by the protein size and stability. Except when studying interacting systems with very slow chemical rates (frequently encountered with very high affinity systems), the sedimentation time is governed by the diffusion of the proteins throughout the solution column. As a standard choice, use 180 microliter samples and expect equilibration at the first speed within 48 hours (usually less time is required at the following, sequentially higher rotor speeds). To avoid longer equilibration times required for larger protein complexes > 200 kDa, usually the sample volume is decreased to 150 microliter. For samples that do not possess sufficient stability for attaining equilibrium at three rotor speeds, 100 – 120 microliter samples can be used. (Alternatively, shorter column techniques are available, but require a different analytical strategy than described in this protocol (see section Experimental Approaches above.) On the other hand, when studying small proteins (e.g., < 5 kDa) or peptides, larger solution columns of up to 400 microliter may be advantageous to generate equilibrium gradients with sufficient curvature at the accessible rotor speeds. In this case, the higher diffusion coefficient of smaller species leads to an only moderate increase in equilibration time. To shorten the time to attain equilibrium, an initial overspeeding period for a few hours at threefold the first equilibrium rotor speed can be used, but the concentration profiles should be monitored so as not to exceed a value two or three times the loading concentration at the bottom of the cell (this is described in Step III.3. below). Usually, the rotor temperature is an important factor for the stability, and unless thermodynamic considerations prescribe the use of a particular temperature (the binding constant is temperature dependent following $RT\ ln(K_A) = \Delta G = \Delta H - T\Delta S)$, low temperatures such as 4 – 8 °C are recommended.

8. The rotor speeds must be selected to generate concentration gradients with shapes that include both shallow and steeper regions. It is highly useful to acquire data sequentially at multiple rotor speeds, generating first relatively shallow gradients with ratios of concentrationw at the bottom to meniscus concentrations of 3:1 – 5:1, and finally, at the highest rotor speed, a condition where the meniscus region is free of protein, termed 'meniscus depletion'. The shapes of typical profiles are shown in Figure 5. For 180 microliter samples, select the three rotor speeds according to Eq. 2, inserting for *M* the average molar masses of the components and the expected complex. (For samples of unknown mass, these values can be estimated from the previous SV experiment, or from the initial slopes in an Archibald analysis of the early experiment SE (Schuck and Millar 1998), available in SEDFIT) For different sample volumes, measure the new column height *l* from initial scans and

use the formula $rpm^* = rpm \times \sqrt{5/l \, (mm)}$ to correct the value from those of Eq. 2. A slightly higher rotor value of the highest rotor speed generating steeper gradients can be applied when using the IF optics.

9. To assist in selecting the appropriate configuration for the experiment, you can use the 'Generate' function of SEDFIT to assess the shape of the predicted SE profiles by simulating the approach to equilibrium at a given rotor speed. If this is based on realistic molar mass values and sedimentation coefficients of the free and complex species (for example, assuming equal concentrations of each), this will also indicate the expected minimal time to attain equilibrium. A more advanced simulation with SEDPHAT is also possible, in which realistic levels of noise are added to theoretical equilibrium profiles calculated on the basis of hypothetical values for the binding constants. Re-analysis of this theoretical data set can be used to judge if the SE run will provide information to determine the correct binding constants with acceptable error estimates. This requires more familiarity with the software.

II. Preparing the Analytical Ultracentrifuge and the Cell Assembly

1. Assemble the cell components according to manufacturer's instructions. For the IF optics, sapphire windows are required, for the ABS system, use quartz windows. When assembling the components, determine if the window gaskets obscure the corners at the bottom of the sample and reference sectors. Frequently, some trimming with scissors is required to eliminate any shadow in the light path from the gaskets. This is very important in SE, because otherwise the region of the steepest gradients with the highest information content in SE cannot be observed. Select double-sector charcoal filled Epon centerpieces. (There are several alternatives: six-channel centerpieces permit more samples to be studied, but only at lower sample volume; for the IF optics, centerpieces with dual filling holes are available, or six-channel external loading centerpieces, which make thorough rinsing of the cell assembly easier after the 'aging' process (see next step)).

2a. For the IF optical detection only, mechanical stabilization of the cell assembly is required. Due to the exquisite sensitivity of the IF system to optical pathlength differences, rotor-speed and time-dependent signal offsets may result from small deformation induced by the gravitational stresses. This can be largely prevented by the following pre-conditioning procedure: *a*) Assemble the cells and torque to 120 – 140 inch/lb. Fill the cells with 200 µl of water (the volume must be greater than the sample volume for the following SE experiment) and seal them. *b*) Insert the assemblies into the rotor, carefully align them, and centrifuge for 1 hour at 50,000 rpm. *c*) Stop the run and take the cell assemblies out of the rotor. Re-torque the cells, and observe if the screw rings have slightly loosened as a result of the centrifugation. *d*) Repeat steps *b* and *c* until no movement of the screw ring is required to reach the proper torque. *e*) Insert the cell assemblies in the rotor, insert the rotor in the centrifuge chamber, and install the optical arm. Accelerate the rotor to 50,000 rpm and configure the IF optical system. Set up ten alternating rounds of 2 hours centrifugation at 50,000 rpm and 1 hour centrifugation at 3,000 rpm. During each 50,000 rpm run, take a series of 10 IF scans in 5 minute intervals. This step can be conveniently set up as an 'equilibrium method' in the AUC operating software to run overnight. Use the software MATCH to establish when the scans in successive series show no change. This completes the mechanical stabilization or 'aging' of the cells. If the cells are to be stored before use, leave them filled with water.

2b. For the IF system only: Take 'aged' cells, insert the rotor and mount the optical arm, and centrifuge at the first rotor speed selected for the experiments. Adjust the parameters for the IF optics in the 'Laser setup' for each cell (for criteria see III.8. of the SV protocol). After 1 hour of centrifugation, take a set of 10 IF scans (in the velocity mode with time interval 10 sec). Repeat at all other rotor speeds. These scans will serve as 'water blanks'. Stop the run, carefully remove the

water from the cell assemblies through the filling holes, without causing any mechanical change in the assembly. For example, use suction applied through a small tube connected to a vacuum flask. Rinse both sectors of the cell with reference buffer and remove the buffer. This will prevent small changes in the buffer composition from residual water when filling the cell assembly. This step II.2b is not required in experiments with a sufficient range of rotor speeds and column length (the default recommendations) to permit computational determination of the radial-dependent baseline. However, the experimental determination can be both a safe-guard against unexpected profiles that do not meet the computational requirements, and a test for consistency of the data analysis.

2c. For the IF system only: The settings of the IF optics, and in particular the timing of the laser must remain unchanged for the remainder of the SE experiment for the water blanks to be valid. In order to reproduce these settings, make a note of each of the laser setup parameters for each cell.

III. Sample Preparation and Starting the Run

1. Mix the samples in Eppendorf tubes (see Step I.2 for sample requirements). Fill the cell assemblies with the sample at the volumes planned in Step I.7. For the ABS system, fill the reference buffer to a volume 10 μl larger than that in the sample sector. When using the IF system or a combination of ABS and IF, precisely match the reference buffer in volume (and chemical composition) to the sample. Seal the cell assemblies as described in Step II.1 of the SV protocol.

2. Insert the cell assemblies into the rotor, weigh and install the counterbalance, and align the cells at the scribe marks. Insert the rotor into the rotor chamber and mount the optical arm. Evacuate the chamber, set the desired temperature on the centrifuge panel. Determine whether or not the optical system is radially calibrated. If not, select a rotor speed of 3,000 or 4,000 rpm, and perform the radial calibration of the required optical system. (Note that this can be done with the sample in place for a SE experiment, in contrast to the SV experiment.) When the vacuum permits, accelerate the rotor to the first (this is the lowest) rotor speed selected in Step I.8. Note that the equilibrium experiment does not require the temperature equilibration period that is essential for SV.

3. In exceptional cases when studying proteins of limited stability, an initial overspeeding period may be required to shorten the experimental time. For overspeeding, increase the rotor speed to threefold the first selected equilibrium speed, and set up a sequence of scans (e.g., absorbance at 280 nm) as a velocity method to monitor the concentration gradients. After a few hours, when the absorbance signal near the cell bottom has approximately doubled, stop the scanning and drop the rotor speed to the first equilibrium speed. (Any sedimentation creating excess concentration at the bottom of the solution column will significantly prolong the experimental time and may lead to irreversible or slowly reversible aggregate formation.)

4. Prepare the scan settings file on the AUC control computer. Specify the rotor type, set the centrifugation time to 'hold' mode, and set the run temperature. Set the parameters of each cell to the absorbance (and/or interference, respectively) equilibrium mode. Set the radius interval for scanning from 6.6 to 7.25 cm (for 180 microliter columns), and the wavelength to 280 nm.

5. In the 'Options' settings, select 'Acquire intensity data instead of OD data'. This will separately store the radial-dependent light transmission through the reference and the sample sector, respectively, and provide a rational criterion for the data analysis range. The intensity scans (which will be saved in '*.ri*' files) can be transformed to absorbance scans easily at any time with SEDFIT. In the same menu, set the number of overlays to 3.

6. For each cell, specify the scan 'Details': Enter as additional wavelengths 250 nm and 230 nm. Set the radial step size to 0.001 cm and request 20 replicates in the stepping mode. Specify a new directory name for data storage. When using the IF system, set the inside and outside radius to encompass the complete solution column. Do not change the laser settings after water blanks have been acquired. Verify the 'Laser Setup' parameters are as noted in Step II.2c. Only if no water

blanks were taken (for computational determination of the radial-dependent baseline), set up the laser timing and imaging parameters in the 'Laser Setup' (for criteria see III.8. of the SV protocol) and make a note of the settings for each cell.

7. Save the scan settings file. Start a single scan to verify the quality of scans. When the ABS data acquisition is in intensity acquisition mode, two profiles will be shown for each cell: the radial-dependent reference transmission and the sample transmission. Both menisci should be visible; they can be identified as local drop in the transmitted light. The scans should extend from the air-to-air region in front of the meniscus to the shadow region at the end of the solution column. The radial position r^* near the bottom of the cell where the reference intensity starts to decline (from the shadow of components of the cell assembly) will mark the maximum radius that can be potentially considered for analysis. The transmitted intensity through the sample, at least for one of the selected wavelengths, should be between 10% and 90% of the reference intensities. With time, the intensity transmitted through the sample close to the meniscus will increase, approaching that of the reference, and will strongly decrease or disappear close to the bottom, indicating the accumulation of the protein.

8a. In the 'Methods' menu for SE, start with the lowest rotor speed and set up a sequence of scans in six hour intervals for several days ('delay condition' is 6:00, 'number of scans' setting is 1, enter the appropriate run temperature). Start a 'Method' scan. When the concentration profiles are expected to be in sedimentation equilibrium, usually after 24 – 48 hours, use the software MATCH and compare the sequence of the last several scans for each cell. For this test, use either IF data (if available) or ABS data intensity at 280 nm or 250 nm transformed to absorbance with SEDFIT. The rms difference between the scans relative to the last scan should approach a constant value. Usually the asymptotic value for the rms difference reflects the noise in the data acquisition and will be < 0.01. If the rms difference does not appear to have attained a constant value, sedimentation equilibrium is not yet attained and centrifugation at the current rotor speed must be continued. If equilibrium has been reached, stop the scanning. In the 'Methods', delete all previous rotor speeds, then, beginning at step 1, enter the next higher selected rotor speed. When using the IF optics, verify that the laser settings remain unchanged. Start a new set of scans and verify that this will accelerate the rotor. Repeat this step for the next two rotor speeds.

8b. As an alternative to 8a: For proteins that are known to be sufficiently stable, and in the absence of factors that could significantly extend the centrifugation time (e.g., high viscosity buffer, very large protein complexes > 200 kDa, very elongated protein shape leading to an unusually small diffusion coefficient, or very small chemical off-rate constants), the 'Methods' for the SE can be configured in one step to go through a sequence of rotor speeds. For 180 microliter columns, start with the lowest rotor speed and request 8 scans in 6 hour intervals. Continue with the next higher rotor speeds, for each requesting 8 scans in 6 hour intervals. At the end of the highest rotor speed, append a line with the delay condition 'hold'. Start the 'methods scan'. This will automatically collect the necessary set of scans and switch the rotor speeds during the next several days. Attainment of equilibrium can be verified with MATCH, as described above, although without the possibility of extending the sedimentation time except for the highest rotor speed.

9. After equilibrium profiles at all rotor speeds have been acquired, stop the run.

10. If in doubt about the stability of the proteins, for example, with regard to proteolytic degradation, recover the sample for SDS-PAGE. Alternatively, the solution can be re-suspended by carefully shaking, and a short-column SV experiment can be conducted with the $c(s)$ analysis indicating the solution state of the protein mixtures (this is usually not done routinely).

11. When using the IF optics: Without disassembly, carefully remove the sample and reference solutions, rinse the cell assembly with water, insert water, seal, and re-centrifuge and scan the water-filled cells at the rotor speeds of the equilibrium experiment. This will generate a new set of

water blank scans to be compared with the initial water blanks to verify the optical stability of the cells and the acquisition system. This step is not required if the computational determination of the radial-dependent baseline is chosen.

12. Clean the cell assemblies as outlined in the SV Protocol (IV.2-3), except when using 'aged' assemblies. The cleaning of these can take place without disassembly by exhaustive rinsing. This is facilitated with six-channel external loading cells and the dual filling hole double-sector cells. This allows the pre-conditioned 'aged' cells to be re-used without new 'aging' cycles, as long as they remain mechanically unaltered. Store the 'aged' cells filled with water.

IV. Data Analysis

1. To prepare the data: For the ABS data, convert the intensity scans into absorbance scans: At each wavelength and rotor speed choose the scans where the equilibrium was attained. When using 6-channel centerpieces, save the data in separate files for each sector. For the IF data, compare the water blanks before and after the SE experiment, take the average, and subtract this average from the equilibrium profiles. These steps are aided by software: The equilibrium condition can be tested with MATCH; SEDFIT utility functions (such as saving raw data with modified filenames) facilitate the other operations. When executing this data analysis for the first time, visit the website http://www.analyticalultracentrifugation.com/sedphat/sedphat.htm and familiarize yourself with the help system of SEDPHAT, the introduction to the organizational structures and the screenshots for the specific models. It can be useful to have access to the online help system while going through the following analysis steps.

Analyzing the sedimentation profiles of the individual components to determine their extinction coefficients and apparent molar mass M*

2. Open SEDPHAT and assemble the equilibrium scans for each cell and detection wavelength (and IF) into a multi-speed equilibrium experiment structure. For example, start with the equilibrium scans at 280 nm of the cell with individual component 'A' only. Of those, select and load the equilibrium profiles at all rotor speeds. Specify the known extinction coefficient at 280 nm and the buffer density. Enter the $\bar{v}$ of the component (or an estimate), or an operationally defined $\bar{v}*$. Graphically estimate the meniscus and bottom. For the left fitting limit, exclude the artifacts close to the meniscus. For the right fitting limit, chose the highest radius where the maximum absorbance is < 1.5 OD, with the radius not exceeding $r*$ determined in Step III.7. With ABS data, also avoid regions of very steep gradients. In the experiment ('xp') parameters, mark the bottom to be a floating parameter. Save the xp file. Similarly, load the scans at the other wavelengths (and IF) in multi-speed equilibrium structures. For these, extinction coefficients are usually not known: Enter an initial estimate of 0.5 and 5fold the 280nm extinction coefficient for 250 and 230 nm data (or 2.75 times the Mw for the IF 'extinction coefficient' of non-glycosylated proteins), respectively, and mark this parameter to be floated in the non-linear regression. (Irrespective of our designation of the components, the single component models use the 'extinction coefficient A' field.) Save the experiment file.

3. Make sure all xp-files of all signals are loaded. Select a self-association model from the model menu, such as the monomer-dimer model 'A + A $\longleftrightarrow$ A2'. In the global parameters: Enter an estimated molar mass, and mark it to be floated in the fit. Switch on 'mass conservation', and 'vary local concentration'. Switch off the self-association by entering a value of $\log(Ka) = 1$ (unmarked to be kept fix), and fix $\log(k-)$ at -10, switch off all non-ideality parameters. Set the concentrations in the local parameters: For the experiment at 280 nm (or that with known extinction coefficient, respectively), link the concentration to itself, and enter the loading concentration in micromolar units. For all other experiments, link the concentrations to the first experiment with known

extinction coefficients. This redirection will ensure that only one unambiguous value for the concentration for each cell is used, which will permit calculation of the extinction coefficients. Next, open the xp parameters and redirect all meniscus and bottom parameters to the data set with the known extinction coefficient. This will constrain the analysis of each cell to the same geometry, making use of the fact that the different signals are from the same cell. Save the configuration.

4. Test the starting estimates of all parameters with the 'Run' command. If necessary, change some estimates to ensure that the theoretical curves at this stage are already in the same order of magnitude as the experimental data. Fit the model. Repeat the fit with alternating optimization methods (Simplex and Marquardt-Levenberg). After the non-linear regression has converged, update the configuration, and document the fit by copying the SEDPHAT window display. Assess the quality of the fit, using the local rms deviations as a measure (they should be < 0.01, within the noise of the data acquisition), and the criteria for systematic residuals described in Figure 5.

5a. If the fit is satisfactory, note the best-fit apparent molar mass value, M^*_A. If the molar mass scale was based on an operational $\bar{v}^*$, transform the molar mass to the 'true' molar mass M_A with $M_A = M^*_A \times (1 - \bar{v}^* r) / (1 - \bar{v}_A r)$. The determined molar mass M_A should be within a few percent of the molar mass from sequence or mass spectrometry, with deviations usually originating from errors in the estimated $\bar{v}$. Note the best-fit values of the extinction coefficients. Along with M^*_A, these will be used in the subsequent interaction analysis.

5b. If the fit is not satisfactory, some of the most common problems are: 1) The amplitudes of the fitted distributions are very far from the measured profiles. If so, the loading concentrations were initialized poorly. 2) Sometimes the ABS baseline offsets are not constant and shift with rotor speed. In this case consider the 'fit RI noise' option to permit individual baselines. 3) Mass conservation is not strictly fulfilled. If so, make sure the bottom position was an adjustable parameter during the fit, and if necessary remove the mass conservation constraint. 4) The component is self-associating. (So far, this was switched off in this model by fixing $\log(Ka) = 1$.) Reconcile this possibility with the SV results. Initialize the $\log(Ka)$ at a value of $4 - 5$, and float in a new fit with the monomer-dimer model, or switch to another self-association model. 5) Impurities of small molar mass or aggregates are present; inspect SDS-PAGE, HPLC elution profiles, or $c(s)$ SV distributions of the recovered sample. 6) Thermodynamic non-ideality is affecting the sedimentation at high concentration; this will generate a distinctly lower slope of the profiles at higher concentrations.

6. Repeat Steps IV.2 – IV.5 with the second component. Use the same $\bar{v}^*$ as an operational quantity, and derive the apparent molar mass M^*_B of the second component based on the same $\bar{v}^*$ scale as with the first component. In this way, although the apparent molar mass values appear different than the known sequence molar mass, they are valid for the interpretation of the centrifugation experiment on the particular $\bar{v}^*$ scale, and they are directly additive in complex formation, for example, $M^*_{AB} = M^*_A + M^*_B$.

Determining the equilibrium binding constant of the hetero-association

7. For each cell with mixtures, assemble the scans at each wavelength (and IF) into multi-speed sedimentation equilibrium structures. Set the meniscus, bottom and fitting limits as before in Step IV.2. Mark the bottom position to be a floating parameter. Use the same operational $\bar{v}^*$ and enter the predetermined extinction coefficients for the respective wavelength (unmarked as fix parameters). Save the xp files.

8. Select the 'A + B $\longleftrightarrow$ AB Hetero-Association' model. Toggle to the 'both A,B in micromolar concentration' field in the global parameters. Load xp files of all mixtures in SEDPHAT. Not all wavelengths are required, but the loaded set should contain the most informative signals (usually

230 nm for the lower concentrations, 250 nm for the highest concentrations). IF data should be combined with at least one ABS signal from the same cell. Redirect the meniscus and bottom so that only one set of geometric parameters exist for each cell. Similarly, for each cell enter the loading concentrations of both components (in micromolar units) and establish links redirecting the loading concentrations for data sets of different signals from the same cell to point to only one 'experiment' (or 'xp' data channel) for each cell. In the global experiment parameters, enter the predetermined values of the apparent molar masses for each component, M^*_A and M^*_B. Mark the 'atot' and 'btot' field to optimize the local concentrations in each cell. Switch the mass conservation model on. Set the log(Ka) field to an expected value (for example, – 5 for binding constant of $K_D = 10 \, \mu M$). Save the configuration.

9. Use the 'Run' command to verify that the initial estimates of the parameter values produce theoretical distributions in the correct order of magnitude. Fit the model (execute the fit command with alternating optimization methods), and after parameter values have converged, save the configuration and document the result. Assess the quality of the fit. In addition to the criteria in Step IV.5b., a possible consideration is often the existence of a subpopulation of material incompetent to associate. Alternatively, different association models may be explored.

10. If an acceptable fit is achieved, note the values for the loading concentrations in all cells. For interactions between proteins which are too similar or too dissimilar in size, a high correlation can exist between the estimate of the binding constant and the ratio of the loading concentrations. If the calculated effective loading concentrations are far from the known values, enter additional constraints (this requires the mass conservation mode). Different variations are possible, which are dependent on the design of the experiment. a) In the global parameters, switch to 'A in micromolar conc and B/A molar ratio', and unmark the field 'btot' to keep the ratios [B]/[A] constant; in the local parameters, enter the known loading concentrations and the molar ratio [B]/[A] for each cell. This is a strong constraint based on the knowledge of the true concentration ratio, which is unfortunately difficult to achieve and to verify independently, and therefore this option is rarely used. b) A more relaxed version of the previous constraint is possible if the cells are set up in a dilution series prepared from the starting stock of the mixture of A and B. This requirement is frequently fulfilled. In this case, mark the field 'btot' for the molar ratio [B]/[A] to be floated (optimized in the fit), but redirect the local parameters to generate only one common molar ratio for all cells. c) If the cells are loaded as a titration series (for example, constant [A] with variable [B]), in the global parameters switch to 'both A, B in micromolar concentrations'; mark both 'atot' and 'btot' to be floated. In the local parameters, establish redirections to link all concentrations of [A] to be the same. After entering these constraints and fitting the model, save the analysis in a new configuration file with new copies of xp files.

11. Perform an error analysis on the determined binding constant. The most reliable approach is the projections of the error surface method (Bevington and Robinson 1992). Although the absolute value of χ^2 ('global red. chi-square' in SEDPHAT) should not be interpreted without absolute knowledge on the noise in the data, the relative increase of χ^2 comparing the original and the new fit is a statistical quantity that can be evaluated using the F-statistics. Calculate the critical value for χ^2 (with the SEDPHAT statistics function). Set the value of the binding constant at a non-optimal value close to the best-fit parameter (for example, increase log(K_A) by 0.1) , fix the binding constant but float all other unknown parameters in a new fit. If the original fit was the best fit, then the current fit will have a higher χ^2. If the new value exceeds the critical value, then the modified binding constant is outside of the error interval. If it is less than the critical value, fix log(K_A) at a value further from the best-fit value and re-fit. Repeat this until the critical χ^2 is exceeded, which

indicates that one of the limits of the $\log(K_A)$ error interval has been found. Reload the best-fit configuration and repeat the procedure by changing $\log(K_A)$ in the other direction. An illustration of this approach can be found in (Schuck *et al.* 1999).

Table 1

	absorbance optics (ABS)	interference optics (IF)
selectivity:	• selective detection (e.g., in the presence of non-absorbing components)	• not selective: sensitive to all solution components (including buffer salts)
linearity and concentration range:	• linear to ~ 1.5 OD, a large concentration range may be achieved by the use of multiple wavelengths	• unlimited linearity, 10^4-fold concentration range
buffer considerations	• buffer cannot contain large amounts of DTT, TRIS, HEPES, other absorbing components for use in far UV (e.g. 230 nm)	• advantageous in the presence of strongly absorbing components (e.g., nucleotides, nucleic acids), **but** requires an exact chemical match of reference buffer volume and composition (through dialysis or gel filtration)
baselines	• small time-invariant (TI) radial baseline profile[9]	• generates significant time-invariant radial-dependent (TI) and radial-invariant time-dependent (RI) baselines, unproblematic in SV, but not trivial in SE
maximum signal/noise ratio	• ~ 300	• > 3000
data acquisition	• ~ minutes/scan, may be limiting rotor speed in SV, depends on scanning mode	• few seconds/scan
windows	quartz windows	• sapphire windows
conditions for velocity sedimentation (SV) • high speed, single speed typical sample requirements: • stability for 3 hours [3] • several cells with a range of loading concentrations; for example, stock solution with serial dilutions [4]	• volume 400 microliters (as low as 150 microliters) • rotor speed high: 40 – 60,000 rpm[1] • optimal loading absorbance: 0.5 – 1.3 OD • typical minimal desirable loading absorbance ~ 0.05 OD[5] • requires thorough temperature equilibration • controlled start from 0 rpm[6] • constant baseline usually with small radial-dependent features[7] • scan settings for fast scans (continuous mode, 0.003 cm radial increment)	• volume 400 microliters (as low as 150 microliters), sample/reference precisely matched • rotor speed usually 50 – 60,000 rpm[1] • optimal loading concentration: > 0.1 mg/ml (> 0.3 fringes) • typical minimal desirable loading concentration: ~ 0.05 mg/ml[5] • requires thorough temperature equilibration • controlled start from 0 rpm, may need pre-adjustment of optics [6] • generates radial baseline profile and radial-invariant offsets in each scan, which can be computationally eliminated after modeling[8]
conditions for equilibrium sedimentation (SE) • low speed, multiple speeds typical sample requirements: • stability for 2 – 5 days [3] • use gel-filtration to remove small Mw contaminants • several cells with a range of loading concentrations; for example, stock solution with serial dilutions [4]	• volume 180 microliters sample and 190 microliters reference (150 microliters sample for Mw > 100 kDa) • two or three rotor speeds, lowest at c(b)/c(m) ~ 3, highest generating meniscus depletion c(m) ~ 0 [2] • optimal loading absorbance: 0.2 to 0.5 OD[4] • typically scan at multiple wavelengths: 280 nm, 230 nm, 250 nm • usually no prior temperature equilibration required • constant baseline usually with small radial-dependent features[9] • scan settings for slow, precise scans (step mode, 0.001 cm radial increment)	• volume 180 microliters, sample/reference precisely matched (150 microliters sample for Mw > 100 kDa) • two or three rotor speeds, lowest at c(b)/c(m) ~ 3, highest meniscus depletion c(m) ~ 0; can tolerate steeper gradients leading to higher sample concentration[2] • optimal loading concentration: > 0.1 mg/ml (> 0.3 fringes) • require 'aging' of cell assemblies, water blanks • usually no prior temperature equilibration required • radial baseline profile and radial-invariant offsets, requires water blanks or TI noise elimination from global analysis of different rotor speeds[10]

(1) Choice of rotor speed: generally as fast as possible but dependent on protein size and optical system; the acquisition of at least 5 – 10 scans during the complete sedimentation process is desirable in SV; for molar mass determination slightly lower rotor speeds may be desirable (2) The ratio of concentration at the bottom relative to the meniscus, c(b)/c(m), can be theoretically predicted by simulating the approach to equilibrium with SEDFIT. This also provides a lower limit for the time to attain equilibrium and allows assessing the concentration profiles and gradients in equilibrium; (3) Stability may depend on temperature – SV and SE can be run at 4 °C; sedimentation equilibrium can be shortened by reducing column volume. (4) Concentration choice will depend on the purpose of the experiment (5) Lower values are possible, but with deteriorating level of detail due to limiting signal/noise ratio. (6) Controlled start from 0 rpm excludes the use of a low-speed (typically 3,000 rpm) phase for adjustment of optical and scan settings or temperature equilibration prior to high-speed acceleration. (7) Ideally exhibits a constant flat baseline, but ordinarily shows some time-invariant features from imperfections in the windows, which can be computationally eliminated after data analysis (8) Computational elimination is usually unproblematic in conjunction with modeling the time-course of sedimentation. (9) Baseline may shift at different wavelengths or when using buffer components with unstable absorbance, such as DTT (may by substituted by TCEP). Radial-dependent features may be eliminated computationally in the global analysis of equilibrium at a sufficient range of rotor speeds. (10) Computational treatment of TI noise in sedimentation equilibrium depends on the use of a sufficiently large range of rotor speeds, but may be improved by global multi-signal analysis in conjunction with absorbance data.

Table 2

	detection	details and recommendation
ionic strength you need sufficient ions to screen protein charges and prevent long-range electrostatic interactions from affecting protein sedimentation	both IF and ABS	▪ *always use > 10 mM NaCl or other salt*
absorption for ABS detection, the absorption of the buffer at the detection wavelength should not exceed 0.2 OD ▪ *select appropriate detection system* ▪ *if in doubt, measure the absorption spectrum of the buffer*	ABS	▪ *always use > 10 mM NaCl or other salt* ▪ HEPES and TRIS buffers, as well as EDTA and EGTA, absorb in the far UV. At low concentrations they can be tolerable for 280 nm ABS detection (e.g., 10 mM TRIS), but this may not permit the 230 nm detection. ▪ β-mercaptoethanol or DTT at low mM concentrations are compatible with IF and ABS detection at 280 nm, but they will generate an absorbance signal which may change with time. (For the ABS detection system, this requires a radial-invariant baseline ('RI-noise') to be considered in the data analysis) ▪ the presence of nucleotides at $> 50\,\mu M$ usually prohibits the use of the ABS system
refractive index the IF system detects the sedimentation of buffer salts (and any other buffer component) and will be sensitive to even very small mismatches in the concentration between the sample and reference buffer	IF	▪ *use size-exclusion chromatography or equilibrium dialysis to change buffer, if necessary* ▪ high concentrations > 1M of buffer components with large refractive index signal, e.g. guanidine hydrochloride, CsCl, glycerol, and others, are very difficult to match optically between sample and reference. In this case, either use the ABS system, or use pure H_2O as a reference buffer and explicitly treat the buffer as a sedimenting component in the data analysis ▪ if detergents are required, if possible, non-absorbing detergents in conjunction with the ABS system is usually advantageous over the IF system
density buffer components that raise the density of the solution may create density gradients at high rotor speed	both	▪ *glycerol or sucrose should be absent, if possible* ▪ solutions of higher density decrease the sedimentation velocity and increase the time required to attain SE ▪ self-forming density gradients may be a concern for mixed solvents and high concentrations of buffer components (such as CsCl or sucrose)
viscosity buffer components that lead to increased viscosity extend the experimental time	both	▪ if glycerol cannot be avoided, multiply the time-intervals for establishing SE of 6 h with the relative viscosity of the solution. ▪ for SV experiments, use larger time intervals between scans and let the experiment continue until the protein is depleted from the solution column.
preferential hydration when using buffer components that significantly increase the solvent density, preferential binding or exclusion of water from the protein solvation shell can lead to changes in the protein buoyancy and in its sedimentation behavior	both	▪ this is usually not a concern for buffers with density close to water (1.02 g/ml), for which the density between the hydration shell and the buffer is nearly matched ▪ the effect of buffer components strongly interacting with the protein, such as chaotropic agents or detergents, needs to be considered with regard to the altered partial-specific volume of the protein (see Appendix 2). ▪ the characterization of heterogeneous interactions between proteins is usually insensitive of hydration effects, as the effective buoyant molar mass the solvated protein can be determined by sedimentation for each protein separately (see Appendix 2).

Standard phosphate buffered saline works well with regard to all necessary considerations. For differences between IF and ABS detection, see also Table 1. Information on chemical compatibility with solvents can be found at http://camis.sr.unh.edu/AUC/cell.html.

Table 3:

problem under study:	sedimentation equilibrium	sedimentation velocity
determine the molar mass of a tight protein complex (assume known partial-specific volume)	single gradient possible, but desirable are 2 – 3 rotor speeds, 2 – 3 concentrations; Model with a single exponential to extract Mw **advantage:** direct measurement, usually reliable estimate, error < 5% **disadvantage:** resolves contaminants only poorly, experiment takes up to a few days	single SV experiment is possible, but desirable is a dilution series of concentrations; Modeling with the Lamm equation is in theory possible but usually very sensitive to heterogeneity and only gives a lower limit of M; c(M) can be advantageous; optimal is a hybrid discrete continuous model describing impurities with continuous sections, floating Mw of main discrete component **advantage:** relatively tolerant of impurities outside the size-range of interest (they will be resolved), takes several hours **disadvantage:** frequently lower precision (~ 10 %)
determine the oligomeric state of a membrane protein in detergent solution	density compensation, or Edelstein-Schachman technique {Edelstein, 1967 #186}	
determine the purity of the sample, detect protein aggregates	results can be highly variable dependent on the nature of the sample	long-column, high-speed SV with c(s) analysis is the method of choice, very sensitive for detection of higher oligomers, stable complexes
determine the number of species	(variable resolution)	long-column SV with c(s) gives better resolution
presence of self-association	dilution series, powerful negative control: can all data be fit with a single species model?	dilution series, test: are c(s) peak position concentration dependent?
kinetics of association	no information possible	diagnostics: are c(s) peak positions concentration dependent, or shift only the peak heights with concentration? global modeling of sedimentation boundaries
determine the stoichiometry of a weak protein complex	dilution series, model SE globally with different stoichiometries, compare	dilution series With c(s) analysis, analyze global model isotherm of $s_w(c)$ with models of different stoichiometries **advantage:** may work better than SE when complexes cannot be populated well, shape of c(s) may give hint of complexes formed. Tolerant of some impurities. Alternative: global modeling of sedimentation boundaries; **advantage:** boundary shape can exhibit a characteristic shape for higher-order oligomerization, but is very dependent on association kinetics
determine the association constant of self-association	dilution series, model SE globally **advantage:** very direct, use prior molar mass information complication: need good estimate of partial-specific volume, in particular for weak self-association	dilution series with c(s) analysis, global model of isotherms sw(c) **disadvantage:** need to span a very large range of concentrations, since s(1) and s(n) are not known a priori (in contrast to M(1) and M(n) for a given association scheme)
determine association constant of heterogeneous and mixed protein interaction	sediment and completely characterize the sedimentation behavior of both components separately, then use dilution series of mixture, model SE globally **advantage:** very direct, use prior molar mass information; no need to know partial-specific volume	sediment and completely characterize the sedimentation behavior of both components separately, then use dilution series of mixture With c(s) analysis, global model of isotherm of sw(c) Alternative: global modeling of sedimentation boundaries; **disadvantage:** dependent on association kinetics
hydrodynamic shapes of complexes, ligand-induced conformational change	no information possible	populate complex near saturation, best in concentration series to verify limiting s-value of peak in c(s); followed by hydrodynamic modeling.