Supplementary Material for:

Limiting the sedimentation coefficient range for sedimentation velocity data analysis: PBM and *g(s*)* approaches revisited

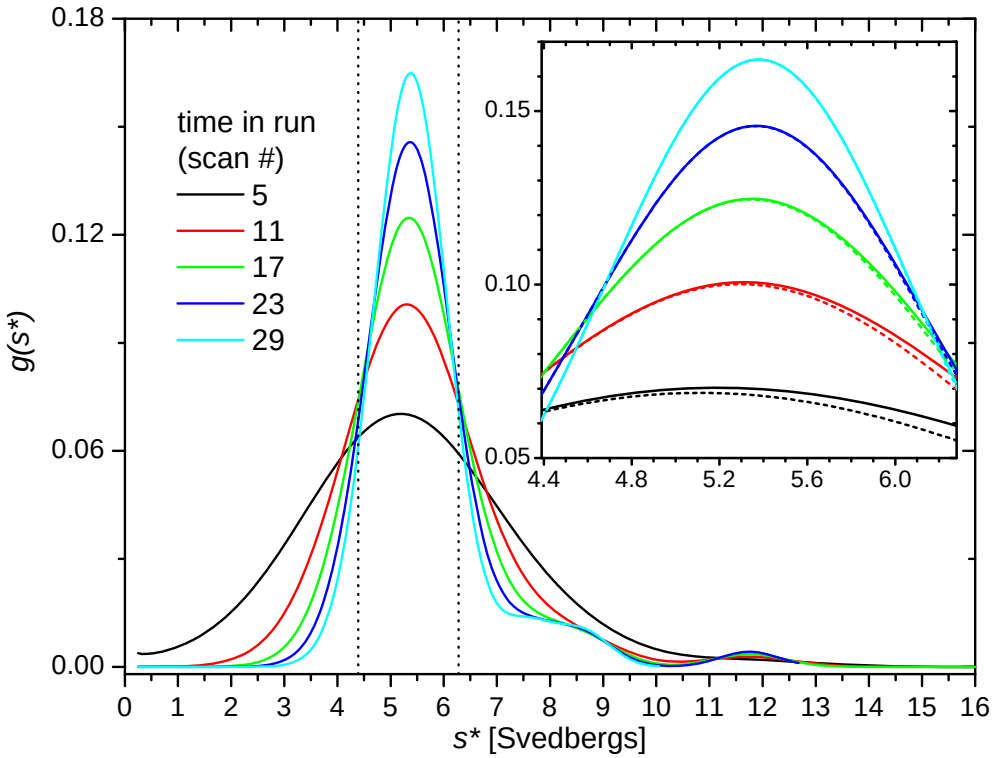
Analytical Biochemistry volume 412 (2011), issue 2, pages 189-202

John S. Philo
Alliance Protein Laboratories, 3957 Corte Cancion, Thousand Oaks, CA 91360 U.S.A.

Variation in the resolution of TRAP aggregates from the 11-mer at different times in the velocity run

Figure 4B in the main paper demonstrates that when the *g(s*)* distributions from the TRAP experiment simulation are fitted the apparent stoichiometry of the main species falls well below the true value (11-mer) when the *g(s*)* distribution is derived from scans very early in the run. It will be shown here that this effect arises because the aggregates are poorly resolved from 11-mer early in the run (when the boundaries have not moved very far from the meniscus). Therefore fitting only a sedimentation coefficient range of 4.39 to 6.28 S will not be effective at removing the influence of the aggregates (the fitted data simply cannot be properly described by a single species), whether the fitting is done using *g(s*)*, PBM, or any other approach.

Figure S1 shows *g(s*)* distributions calculated at time points corresponding to scans 5, 11, 17, 23, and 29 in the TRAP simulation. (These distributions were calculated from a noise-free simulation with a 10-fold higher scan rate.) Early in the run the main peak is very broad, and



Supplementary Fig. S1. *g(s*)* distributions for a noise-free TRAP experiment simulation calculated at different times (scan numbers) in the run. The dotted vertical lines in the main panel indicate the sedimentation coefficient limits used for the fitting. The inset shows an expanded version of this region, and the dashed curves show the *g(s*)* curves after removing the contributions from the aggregates.

the 22-mer aggregate does not even show as a distinct shoulder until about scan 17. The inset in the figure shows only the portion of the main peak between 4.39 and 6.28 S (the part that was fitted). The inset also shows as dashed lines the *g(s*)* distributions resulting when all of the aggregate contributions are removed. Thus the fact that the solid and dashed curves do not completely overlap (especially early in the run) demonstrates that the aggregate signals still contribute to the data being fitted. Note in particular that even though the 22-mer aggregate has a sedimentation coefficient of 7.57 S, at the time of scan 5 its influence extends to below 4.4 S. The heterogeneity due to aggregate contributions then produces a best-fit mass below the true 11-mer value.

Note that even when the aggregates are making significant contributions the shape of the *g(s*)* distribution is still essentially a Gaussian, especially near the top of the peak (the only region that is being fitted early in the run in this example). Thus for real experiments (when there is noise in the raw data) the residuals from a single-species fit do not show a discernible systematic pattern indicative of a poor fit.

When is the concentration high enough to allow a reversible oligomer to be treated as a non-dissociable species? And how does one show experimentally this has been achieved?

To investigate these questions simulations were carried out for rapidly-reversible monomer-octamer and monomer-dimer associations, where the rates of association and dissociation are very high so the time scale for mass-action equilibration is effectively instantaneous compared to the duration of the sedimentation velocity run. The simulations were created using instantaneous reversible association models in SEDFIT version 11.3 for a 75 kDa monomer with a sedimentation coefficient of 4 S and over a range of concentrations for a fixed association strength. Random noise was added at a root mean square level 200-fold below the loading concentration (a constant signal/noise ratio) to help distinguish when the differences between different concentrations would be experimentally observable.

These simulated scans were then fitted as single species using *g(s*)* analysis (using DCDT+ version 2.2.3). The scans for computing the *g(s*)* distribution were chosen at a time point when the oligomer was approximately in the center of the cell. Two different fits were conducted of each *g(s*)* distribution, one in which the entire distribution was fitted as a single species, and a second where the fit was limited to the central half of the main peak (50% peak height). This second fit then mimics the common situation of trying to isolate the oligomer from contributions of the monomer or other minor components.

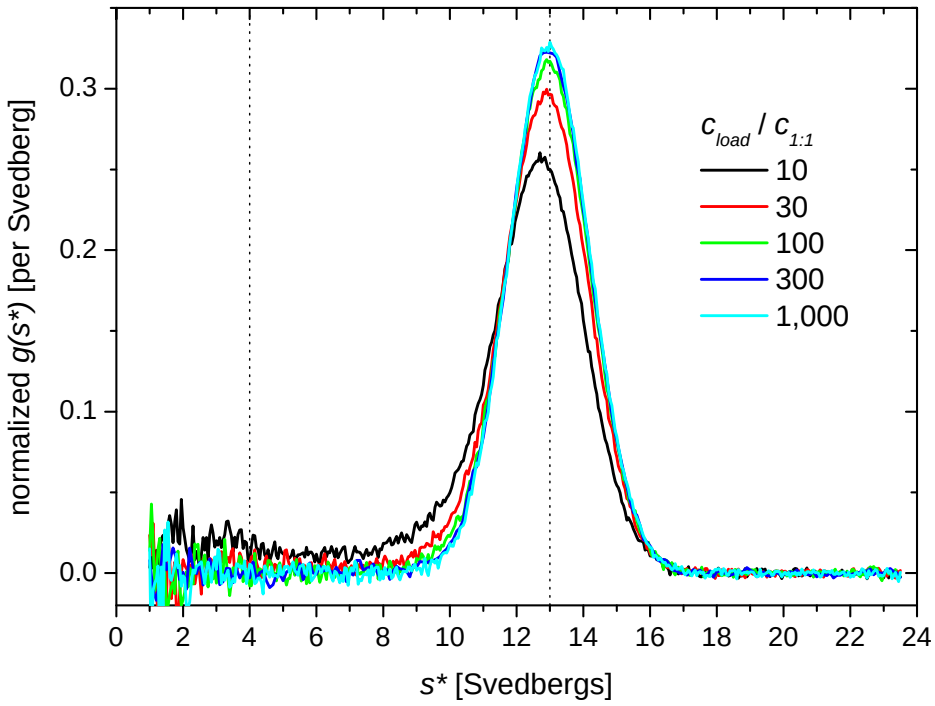
monomer-octamer association

These simulations assumed an octamer sedimentation coefficient of 13 S (slightly below the value expected for a cubic octamer of spherical monomers sedimenting at 4 S) and a rotor speed of 32,000 rpm. It is useful to characterize the extent of association by a unitless ratio of the loading concentration (weight concentration) relative to whatever weight concentration gives equal molar concentrations of monomer and octamer, $c_{1:1}$. For a molar association constant K_a the equilibrium equation is

$$[\text{octamer}] = K_a \times [\text{monomer}]^8$$

and thus when [octamer] = [monomer] and with a monomer molar mass M we get

$$c_{1:1} = 9 \times M \times K_a^{-1/7}$$

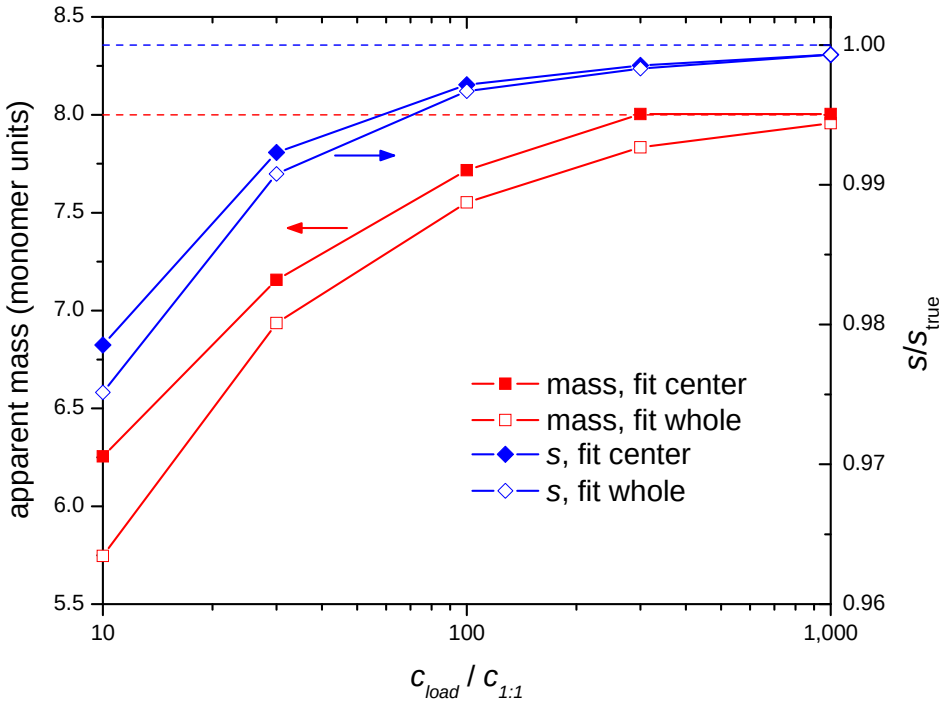


Supplementary Fig. S2. *g(s*)* distributions for simulations of an instantaneous monomer-octamer equilibrium calculated at different ratios of the loading concentration relative to the concentration that gives a 1:1 molar ratio of monomer to octamer. The assumed signal/noise ratio was held fixed at 200:1. The dotted vertical lines indicate the sedimentation coefficients for pure monomer and octamer.

Figure S2 shows the *g(s*)* distributions calculated at loading concentrations of 10, 30, 100, 300, and 1,000 times higher than $c_{1:1}$. At a loading concentration of 10 × $c_{1:1}$ there is still a weak peak near the monomer position, and a raised baseline between the monomer and the main peak. The main peak is also distinctly skewed on the low sedimentation coefficient side. Those features are nearly gone after a further 3-fold concentration increase, and the main peak shifts significantly closer to the true octamer position. At concentration ratios of 100 and higher there is little further change in the shape or position of the main peak. Obviously the same pattern of changes would also occur if the association strength increases at a constant loading concentration.

Figure S3 summarizes the fits to these distributions, giving both the best-fit stoichiometry and the ratio of the best-fit sedimentation coefficient to the true value. Not surprisingly the fits which use only the central portion of the peak (the solid points in the figure) do give results closer to the octamer value than fitting the entire distribution (the open points), particularly at the lower concentration ratios. Also not surprisingly the difference between the two types of fits is more significant for the apparent mass than for the sedimentation coefficient.

These results indicate that for the central-peak fit (the approach that would likely be used in real experiments, where aggregates or other minor components may also be present) the correct stoichiometry will be obtained (the error is below 0.5 subunit) when the concentration ratio exceeds ~50, and further than under those conditions the sedimentation coefficient will be within ~0.5% of the correct octamer value.



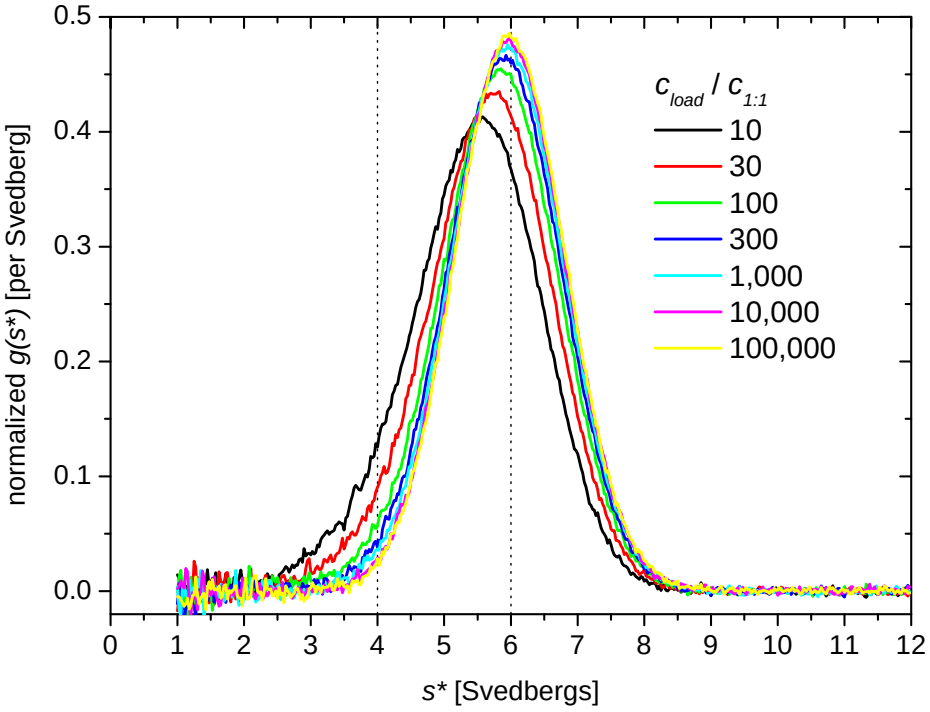
Supplementary Fig. S3. Apparent mass (red) and sedimentation coefficient (blue) values returned by fits of the monomer-octamer *g(s*)* distributions shown in Fig. S2 as a function of the loading concentration. The solid points show results from fits limited to the central portion of the main peak (50% height). The open points show results when the entire distribution is fitted as a single species. The horizontal dashed lines indicate the correct (as simulated) values.

monomer-dimer association

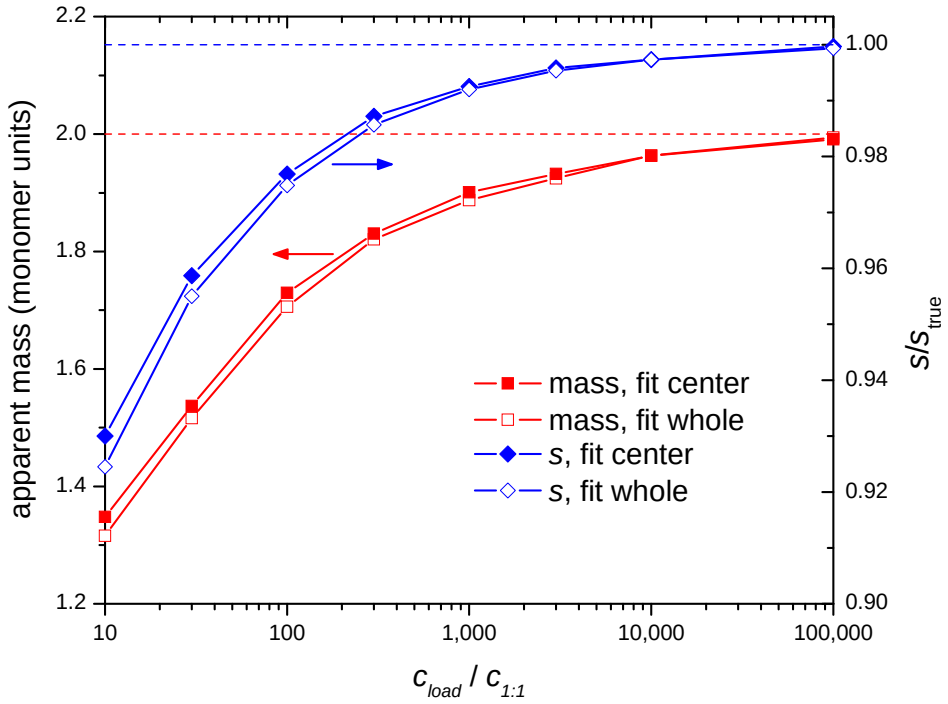
A similar set of simulations and fits was done for an instantaneous reversible monomer-dimer association, with an assumed dimer sedimentation coefficient of 6 S and at a rotor speed of 45,000 rpm. In this case the equilibrium equation is [dimer] = K_a × [monomer]^2 and when monomer and dimer are equimolar the concentration is $c_{1:1}$ = 3 × M × K_a^{-1}. Because the concentration dependence for a monomer-dimer system is much weaker than for the octamer case these simulations included concentration ratios of up to 100,000.

Figure S4 shows the simulated *g(s*)* distributions as a function of concentration ratio. At a concentration ratio of 10 there is only one peak, but that peak is moderately asymmetric and its center falls at a sedimentation coefficient more than 7% below the true value for dimer. Concentration ratios of 30 and 100 produce a narrower and more symmetric peak as well as shifts to a higher sedimentation coefficient. Above ratios of 300 the changes are small, but still discernible (at this signal/noise ratio) until the ratio reaches ~10,000 (a ratio of 100,000 is visually indistinguishable from 10,000).

The results from the fits of the monomer-dimer simulations are summarized in Fig. S5. Although the weaker concentration dependence of a monomer-dimer system means that significantly higher concentration ratios are required to obtain an equivalent precision (percent error) for the oligomer mass or sedimentation coefficient than for the monomer-octamer case, fortunately the precision needed for correct identification of stoichiometry is also much lower (distinguishing 1 from 2 requires much less precision than 7 from 8). Consequently the stoichiometry is actually correct for ratios of 30 or higher, and the accuracy of the sedimentation coefficient is 5% or better at such ratios. To get a mass accuracy better than 10% however requires a concentration ratio of at least 300.



Supplementary Fig. S4. *g(s*)* distributions for simulations of an instantaneous monomer-dimer equilibrium calculated at different ratios of the loading concentration relative to the concentration that gives a 1:1 molar ratio of monomer to dimer. The assumed signal/noise ratio was held fixed at 200:1. The dotted vertical lines indicate the sedimentation coefficients for pure monomer and dimer.



Supplementary Fig. S5. Apparent mass (red) and sedimentation coefficient (blue) values returned by fits of the monomer-dimer *g(s*)* distributions as a function of the loading concentration. This figure includes results for a concentration ratios of 3,000 in addition to those shown in Fig. S4. The solid points show results from fits limited to the central portion of the main peak (50% height). The open points show results when the entire distribution is fitted as a single species. The horizontal dashed lines indicate the correct (as simulated) values.

As might be expected, for this monomer-dimer case the differences between fitting the entire distribution versus only the central half of the peak are small.

An experimental test for whether the concentration is high enough to treat the oligomer as an independent component

In practice if the experimenter is trying to measure the association stoichiometry via a sedimentation velocity experiment then presumably the association constant is also unknown. Therefore it will be necessary to run a dilution series and judge whether the changes upon dilution are small enough (the association is strong enough) so that treating the oligomer or complex as a non-dissociating independent component will give sufficient accuracy to correctly determine the stoichiometry. These simulations can also provide some guidance for evaluating the results of the dilution series.

One possible ‘rule of thumb’ test would be to evaluate how much the apparent mass and/or sedimentation coefficient changes for a dilution of ~3-fold. Based on these simulations, the results obtained for the sample at the higher concentration will be reliable (correct stoichiometry) when a 3-fold dilution produces a decrease in the apparent mass of the oligomer of less than ~10% or a reduction in its sedimentation coefficient of less than ~0.3%.