Supporting Information

Measuring Ultra-Weak Protein Self-Association by Nonideal Sedimentation Velocity

Sumit K. Chaturvedi,[†] Vatsala Sagar,[‡] Huaying Zhao,[†] Graeme Wistow,[‡] Peter Schuck[*,†]

†Dynamics of Macromolecular Assembly Section, Laboratory of Cellular Imaging and Macromolecular Biophysics, National Institute of Biomedical Imaging and Bioengineering, and

‡Section on Molecular Structure and Functional Genomics, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892, United States

Corresponding author: Peter.Schuck@nih.gov

Contents

Mathematical Methods

Nonideal Sedimentation Coefficient Distributions

The nonideal sedimentation coefficient distribution $c_{NI}(s_0)$ was recently introduced as a tool to simultaneously measure polydispersity and colloidal interactions in highly concentrated samples in sedimentation velocity[1]. Briefly, it is defined as the concentration c of species with sedimentation coefficients between s and $s+ds$, such that the sum of their sedimentation signals (given extinction coefficient ε and optical pathlength d) equals in a least-squares sense the experimental signals $a(r,t)$ over the entire range or radius r and time t of the experiment

$$a(r,t) \cong \varepsilon d \int_{s_{\min}}^{s_{\max}} c(s)\chi_1\left(s, D(f_r, s), r, t\right) ds \tag{1}$$

and where the spatio-temporal evolution of each individual species' concentration distribution at unit loading concentration $\chi_1(r,t)$ is described by the Lamm partial differential equation[2]

$$\frac{\partial \chi_i}{\partial t} = -\frac{1}{r}\frac{\partial}{\partial r}\left(s_i \left\{1 - k_S \chi_{tot}(r,t)\right\} \omega^2 r^2 \chi_i - D_i \left\{1 + k_D \chi_{tot}(r,t)\right\} \frac{\partial \chi_i}{\partial r} \right) \tag{2}$$

with sedimentation coefficients s_i and diffusion coefficient D_i. As in conventional c(s) analysis, a hydrodynamic scaling law for compact particles was used to relate the diffusion coefficients to sedimentation coefficients, based on an average frictional ratio as scaling parameter.[3,4] The curly brackets describe the hydrodynamic and thermodynamic nonideality of sedimentation and diffusion, with the nonideality coefficients k_S and k_D, respectively. In a mean-field approximation, the magnitude of nonideal interactions depends on the local total concentrations $\chi_{tot}(r,t)$. The latter dependency makes this a nonlinear problem.

In the software SEDFIT (version 16.1) Eqs. 1 and 2 are solved iteratively, with optimization of fitting parameters k_S, k_D, as well as ancillary parameters (such as baselines and meniscus positions) folded into the computation. Size distribution analysis was regularized based on maximum entropy on a confidence level of 0.68 or 0.95. All computation was carried out on Dell office workstations or on a Macbook Pro running virtual Windows 10.

As described previously[1], modeling experimental SV data provides relatively precise information on k_S , but k_D is much less well determined. In the studies shown in the present work, k_D refined to very small values. Therefore, in the k_D was fixed at a value of 1 ml/g.

Binding Isotherm Analysis

The integral of the sedimentation coefficient distribution across the sedimentation coefficient range of all interacting species

$$s_w = \frac{\int_{s_1}^{s_2} c(s)s\,ds}{\int_{s_1}^{s_2} c(s)\,ds} \tag{3}$$

provides a weight-average sedimentation coefficient s_w that relates directly to the mass balance of the sedimenting system at loading concentrations[5,6]. Therefore, the isotherm of s_w as a function of concentration can be modeled using mass action law, and for the monomer-dimer self-association takes the form

$$s_w(c_{load}) = \frac{s_1 c_1 + 2 s_2 c_2}{c_1 + 2 c_2} \tag{4}$$

with s_1 and c_1 denoting the molar monomer concentration and its sedimentation coefficient, s_2 and c_2 denoting the corresponding parameters of the dimer, and the concentrations linked through $c_2 = K_A \times c_1^2$.

Integration of the differential sedimentation coefficient distributions obtained at different concentrations was carried out in GUSSI[7] and isotherm analysis was carried out in SEDPHAT (version 15.2). In the least-squares fit the parameters subjected to refinement were s_1, s_2 (constrained to 1.4–1.6fold s_1, as can be expected for proteins of globular shape), and the association equilibrium constant K_A.

Statistical analysis was carried out by using F-statistics and the error projection method to determine confidence intervals on the 95% confidence level [8,9].

Lamm Partial-Differential Equations for Coupled Sedimentation/Reaction/Diffusion in the Centrifugal Field

For single-component self-associating systems the evolution of sedimentation, diffusion, and chemical reactions can be described with the extended Lamm equation for each species *i*

$$\frac{\partial \chi_i}{\partial t} = -\frac{1}{r}\frac{\partial}{\partial r}\left(s_i \left\{1 - k_S \chi_{tot}(r,t)\right\} \omega^2 r^2 \chi_i - D_i \left\{1 + k_D \chi_{tot}(r,t)\right\} \frac{\partial \chi_i}{\partial r} \right) + q_i(r,t) \quad (5)$$

where $q_i(r,t)$ accounts for chemical reactions. For example, in the description of sedimentation of a monomer self-associating to form a dimer, Eq. 5 takes the form

$$\begin{aligned}
\frac{\partial \chi_1}{\partial t} &= -\frac{1}{r}\frac{\partial}{\partial r}\left(s_1 \left\{1 - k_S \chi_{tot}(r,t)\right\} \omega^2 r^2 \chi_1 - D_1 \left\{1 + k_D \chi_{tot}(r,t)\right\} \frac{\partial \chi_1}{\partial r} \right) + 2k_{off}\chi_2(r,t) - k_{on}\chi_1^{\,2}(r,t) \\
\frac{\partial \chi_2}{\partial t} &= -\frac{1}{r}\frac{\partial}{\partial r}\left(s_2 \left\{1 - k_S \chi_{tot}(r,t)\right\} \omega^2 r^2 \chi_2 - D_1 \frac{s_2}{2s_1} \left\{1 + k_D \chi_{tot}(r,t)\right\} \frac{\partial \chi_2}{\partial r} \right) - k_{off}\chi_2(r,t) + \frac{1}{2}k_{on}\chi_1^{\,2}(r,t) \\
\chi_{tot}(r,t) &= M\left[\chi_1(r,t) + 2\chi_2(r,t)\right]
\end{aligned} \quad (6)$$

where χ_1 and χ_2 are molar monomer and dimer concentrations, interconverting with a chemical on-rate constant k_{on} for monomers to form dimers and a chemical dissociation rate constant k_{off} for dimers to dissociate into monomers.[6] (In Eq. 6 we have made use of the Svedberg relationship to relate the dimer diffusion coefficient to that of the monomer.) In this description, the nonideality coefficients are in g/L units and – reflecting experimental information content – are approximated to be equal for monomer and dimer. For dimer lifetimes $1/k_{off}$ much shorter than the experimental timescale of hours in SV, the chemical reaction terms can be replaced by a model of instant local equilibration of chemical species.

Eq. 4 was solved in SEDPHAT (version 15.2) and globally fitted to experimental data sets at 4 different concentrations, refining as global parameters the buoyant molar mass, the monomer and dimer sedimentation coefficient (the latter constrained to be within 1.4 – 1.6-fold the monomer *s*-value), the equilibrium dissociation constant K_D, and the nonideality parameters k_S and k_D – the latter expressed as $B_2 = (k_S + k_D)/2$. In addition, refined local fit parameters are the loading concentration, meniscus position, and time-invariant and radially-invariant baseline noise parameters. In order to suppress bias from systematic errors in data acquisition, statistical weights were attached to the data sets such that each experiment contributes equally to the global analysis[10].

Statistical analysis was carried out by using F-statistics and the error projection method to determine confidence intervals on the 95% confidence level [8,9].

Experimental

Proteins

Lysozyme from chicken egg white was purchased from Sigma Aldrich, St. Louis, USA (catalog number L6876). Lyophilized powder was resuspended in designated buffer and dialyzed.

Recombinant chicken γS-crystallin was expressed and purified as previously described [11].

Analytical Ultracentrifugation

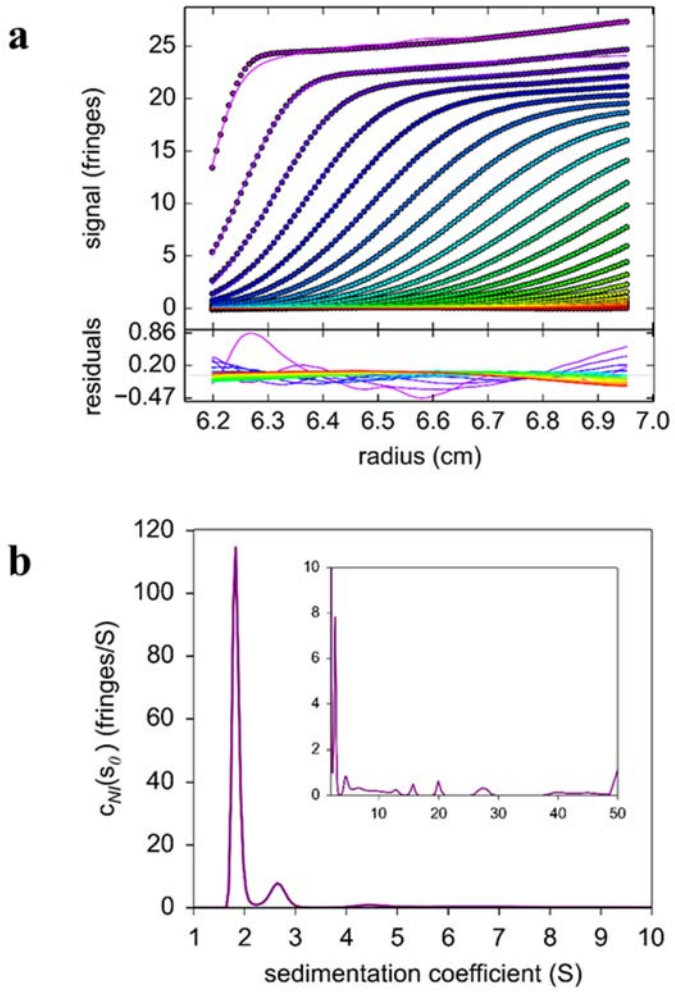
Analytical ultracentrifugation was carried out in a ProteomeLab instrument (Beckman Coulter, Indianapolis, IN) following standard methods.[12]

For sedimentation velocity experiments, samples were filled in either charcoal-filled Epon centerpieces with optical pathlengths of 12 mm or 3 mm (Beckman Coulter, Indianapolis, IN). For the two highest concentrations of chicken γS crystllin 3D printed 1.75 mm pathlength Microfine Green (Protolabs, Maple Plain, MN) centerpieces[1,13] were used (**3dprint.nih.gov model #3DPX-009261;** in the same design as described previously[1] except for a pathlength of 1.75 mm instead of 1.5 mm), at volumes to generate an approximately 12 mm high solution column. (Alternatively, short optical pathlengths may be achieved using commercial 1.5 mm titanium centerpieces (Nanolytics GmBH, Potsdam, Germany).) After assembly of AUC cells and insertion in an 8-hole analytical rotor, the rotor was placed into the ultracentrifuge and temperature was equilibrated to a set-point of 20 °C for 2-3 hours. After acceleration to 50,000 rpm data acquisition using the Rayleigh interference optical detector was started. Scan data were sorted and corrected for time-stamp errors in the software REDATE.[14]

Sedimentation equilibrium experiments were carried with 6 mm long solution columns, using time-optimized rotor speed profiles[15] to attain SE at 25,000 rpm. SE data were globally modeled in SEDPHAT using the INVEQ method by Rowe[16] to account for nonideality with the second virial coefficient.

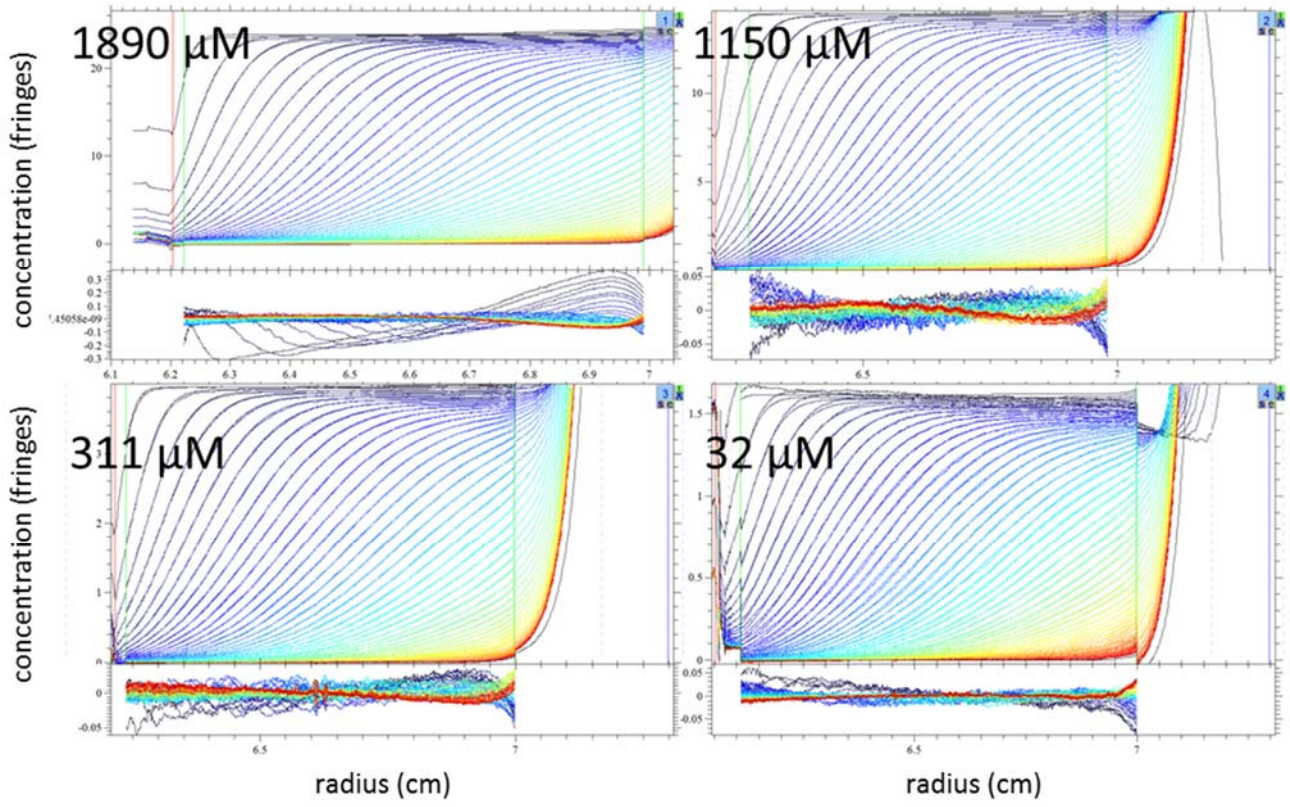
Supporting Figures

Figure S1: Sedimentation data of polydisperse hen egg lysozyme preparation



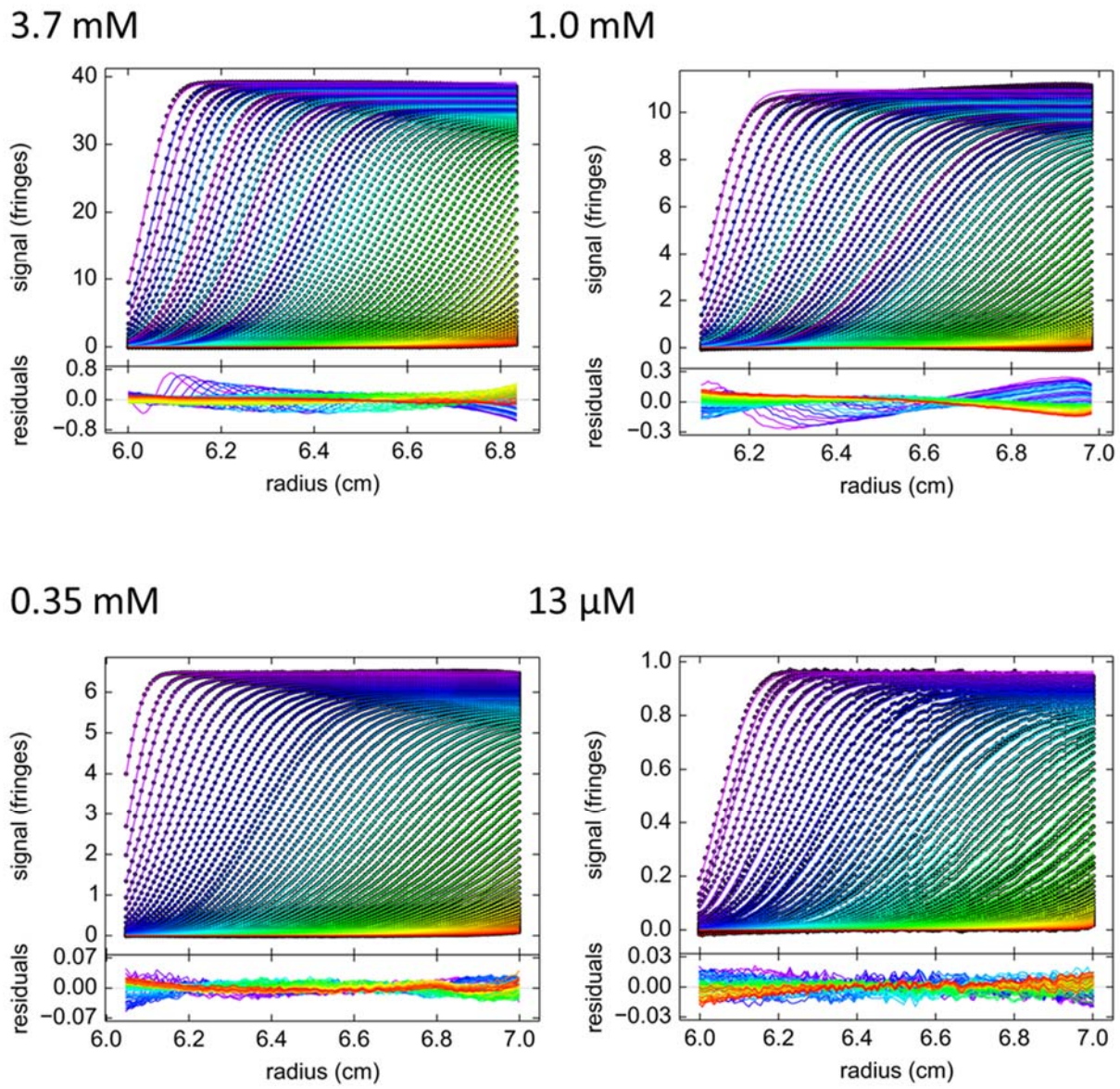
Supporting Figure 1: Example for sedimentation velocity experiment with more polydisperse sample. This was achieved by dissolving lyophilized HEL in 10 mM sodium acetate, pH 4.6, with 300 mM NaCl followed by truncated dialysis, leading to incompletely suspended HEL and/or a transient state of precipitation in excess salt. (a) Evolution of concentration with time (circles, only every 10th data point in every 3rd scan is shown for clarity), and best-fit concentration distributions from $c_{NI}(s_0)$ analysis (solid line), with residuals of the fit attached in the lower panel (rmsd = 0.082 fringes). The presence of a wide size-range of particles sedimenting faster than the majority population can be visually discerned from the sloping plateaus in early scans. (b) Corresponding best-fit $c_{NI}(s_0)$ presented for the range from 1 S to 10 S, with the inset highlighting larger particles across the range from 2 S to 50 S.

Figure S2: Global fit of hen egg lysozyme at high salt with explicit LPDE for monomer-dimer interaction



Supporting Figure 2: Global model of sedimentation data from HEL in high salt conditions with explicit Lamm equation model for instantaneous monomer-dimer self-association. Shown is a screenshot of the global fit in SEDPHAT, where each panel represents data from a different experiment at the loading concentration indicated. The best-fit buoyant molar mass is 3,843 Da, sedimentation coefficients for the monomer and dimer are 1.74 S and 2.61 S, respectively (not corrected to standard conditions), K_A is 53 M^{-1}, and non-ideality parameters are k_S = 3.3 ml/g and B_2 converged to 1.8 ml/g. The rmsd of the model is 0.055 fringes, 0.012 fringes, 0.009 fringes, and 0.009 fringes for data with highest to lowest loading concentration.

Figure S3: Global fit of chicken γS crystallin with explicit LPDE for monomer-dimer interaction



Supporting Figure 3: Global model of sedimentation data from 82 mg/ml chicken γS crystallin with explicit Lamm equation model for instantaneous monomer-dimer self-association. The best-fit buoyant molar mass is 6.2 kDa, sedimentation coefficients for the monomer and dimer are 2.14 S and 3.1 S, respectively (not corrected to standard conditions), K_A is 42 M^{-1}, and non-ideality parameters are k_S = 4.5 ml/g and B_2 converged to 1.36 ml/g. The rmsd of the model is 0.128 fringes, 0.065 fringes, 0.011 fringes, and 0.0058 fringes for data with highest to lowest loading concentration.

References

(1) Chaturvedi, S. K.; Ma, J.; Brown, P. H.; Zhao, H.; Schuck, P. Measuring Macromolecular Size Distributions and Interactions at High Concentrations by Sedimentation Velocity. *Nat. Commun.* **2018**, *9* (1), 4415.

(2) Lamm, O. Die Differentialgleichung Der Ultrazentrifugierung. *Ark. Mat. Astr. Fys.* **1929**, *21B(2)*, 1–4.

(3) Schuck, P. *Sedimentation Velocity Analytical Ultracentrifugation: Discrete Species and Size-Distributions of Macromolecules and Particles*; CRC Press: Boca Raton, FL, 2016.

(4) Schuck, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys. J.* **2000**, *78* (3), 1606–1619.

(5) Schuck, P. On the Analysis of Protein Self-Association by Sedimentation Velocity Analytical Ultracentrifugation. *Anal. Biochem.* **2003**, *320* (1), 104–124.

(6) Schuck, P.; Zhao, H. *Sedimentation Velocity Analytical Ultracentrifugation: Interacting Systems*; CRC Press: Boca Raton, FL, 2017.

(7) Brautigam, C. A. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Methods Enzymol.* **2015**, *562*, 109–133.

(8) Bevington, P. R.; Robinson, D. K. *Data Reduction and Error Analysis for the Physical Sciences*; Mc-Graw-Hill: New York, 1992.

(9) Johnson, M. L. Why, When, and How Biochemists Should Use Least Squares. *Anal. Biochem.* **1992**, *225*, 215–225.

(10) Zhao, H.; Schuck, P. Global Multi-Method Analysis of Affinities and Cooperativity in Complex Systems of Macromolecular Interactions. *Anal. Chem.* **2012**, *84* (21), 9513–9519.

(11) Chen, Y.; Sagar, V.; Len, H. S.; Peterson, K.; Fan, J.; Mishra, S.; McMurtry, J.; Wilmarth, P. A.; David, L. L.; Wistow, G. γ-Crystallins of the Chicken Lens: Remnants of an Ancient Vertebrate Gene Family in Birds. *FEBS J.* **2016**, *283* (8), 1516–1530.

(12) Schuck, P.; Zhao, H.; Brautigam, C. A.; Ghirlando, R. *Basic Principles of Analytical Ultracentrifugation*; CRC Press: Boca Raton, FL, 2015.

(13) Desai, A.; Krynitsky, J.; Pohida, T. J.; Zhao, H.; Schuck, P. 3D-Printing for Analytical Ultracentrifugation. *PLoS One* **2016**, *11* (8), e0155201.

(14) Zhao, H.; Ghirlando, R.; Piszczek, G.; Curth, U.; Brautigam, C. A.; Schuck, P.; Jakoby, W.; Zhao, H.; Ghirlando, R.; Piszczek, G.; et al. Recorded Scan Times Can Limit the Accuracy of Sedimentation Coefficients in Analytical Ultracentrifugation. *Anal. Biochem.* **2013**, *437* (1), 104–108.

(15) Ma, J.; Metrick, M.; Ghirlando, R.; Zhao, H.; Schuck, P. Variable-Field Analytical Ultracentrifugation: I. Time-Optimized Sedimentation Equilibrium. *Biophys. J.* **2015**, *109* (4), 827–837.

(16) Ang, S.; Rowe, A. J. Evaluation of the Information Content of Sedimentation Equilibrium Data in Self-Interacting Systems. *Macromol. Biosci.* **2010**, *10* (7), 798–807.