Please cite this article as: Parker, E. T. and Lollar, P. (2021). Measurement of the Translational Diffusion Coefficient and Hydrodynamic Radius of Proteins by Dynamic Light Scattering. *Bio-protocol* 11(20): e4195. DOI: 10.21769/BioProtoc.4195.

Measurement of the Translational Diffusion Coefficient and Hydrodynamic Radius of Proteins by Dynamic Light Scattering

Ernest T. Parker and Pete Lollar*

Aflac Cancer and Blood Disorders Center, Children's Healthcare of Atlanta; Department of Pediatrics, Emory University, Atlanta, GA, USA

*For correspondence: jlollar@emory.edu

[Abstract] Diffusion is a fundamental process in biological systems that governs the molecular collisions driving biochemical reactions and membrane and transport. Measurement of the diffusion coefficient and application of the Stokes-Einstein equation produces the hydrodynamic radius, which is a commonly used gauge of particle size. Additionally, measurement of the diffusion coefficient and the sedimentation coefficient, and application of the Svedberg equation, yields the molecular weight, which is particularly useful in the characterization of very large macromolecules. Dynamic light scattering (DLS) is the most common method to measure the diffusion coefficient of macromolecules. We describe a procedure to perform DLS measurements on monomeric bovine serum albumin (BSA) purified by size-exclusion chromatography using the Zetasizer Nano S particle size analyzer. We compare several analytical methods in existing software programs to estimate the diffusion coefficient of BSA (extrapolated to water at 20°C at infinite dilution, $D_{20,w}^{0}$) and describe a statistical method to obtain 95% confidence limits of the precision of the estimates. We compare $D_{20,w}^{0}$ estimates to literature values obtained by diffusiometry, sedimentation velocity analytical ultracentrifugation, and other DLS instruments. The method of cumulant analysis in the program SEDFIT (www.analyticalultracentrifugation.com) produced the most precise estimate, $D_{20,w}^{0}$ 6.06 ± 0.07 F (1 F = 10^{-7} cm^2 s^{-1}), which was within the range of estimates obtained by diffusiometry or sedimentation velocity. This protocol is useful for DLS method validation and quality control.

Keywords: Dynamic light scattering, Diffusion coefficient, Hydrodynamic radius, Cumulant analysis, Continuous distribution, Analysis of variance, Particle sizing

[Background] Diffusion is a process whereby solute concentration differences spontaneously decrease toward a state of homogeneity. It is a fundamental process in biological systems that governs the molecular collisions that drive biochemical reactions and intracellular transport (Berg, 1983). Strictly speaking, this process is called translational diffusion in contrast to the rotational diffusion of particles. An empirical law describing one-dimensional diffusion was described by Fick in the 19th century, which states that the rate of flow of particles is proportional to the concentration gradient at a given position. The proportionality constant is called the diffusion coefficient, D.

A molecular interpretation of diffusion was made by Einstein in a series of papers beginning in 1905 (Einstein, 1926 [English translation first]). Einstein proposed that solute particles are subjected to forces due to collisions with solvent particles that are in random, Brownian motion due to thermal energy. A

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

particle that is accelerated by a collisional force experiences a resistive force that is proportional to the particle velocity. The proportionality constant, called the frictional coefficient, f, is a function of the size and shape of the particle. Einstein's law of diffusion is

$$D = \frac{kT}{f} \tag{1}$$

where k is Boltzmann's constant and T is the absolute temperature. This equation is fundamental to the understanding of the hydrodynamic properties of macromolecules. It is important to note that it is restricted to a two-component system, *e.g.*, a solute molecule and solvent water, in dilute solution. In a three-component system (*e.g.*, a macromolecule, a low molecular weight electrolyte, and water), there are four diffusion coefficients, one each that components 2 and 3 would have in the absence of the other and two cross-diffusion coefficients (Tanford, 1961). In most dynamic light scattering (DLS) studies, measurements are made under so-called pseudobinary conditions, namely at low macromolecular concentration in a solution that contains an electrolyte at sufficiently high concentration to screen the charge of the macromolecule. Under these conditions, cross-diffusion terms are considered negligible. For a spherical particle with radius R, Stokes' law is

$$f = 6\pi\eta R \tag{2}$$

where η is the solvent viscosity. Combined with the Einstein equation, the radius of a spherical particle can be calculated if its diffusion coefficient is measured in a solution with known temperature and viscosity

$$R = \frac{kT}{6\pi\eta D}\ (Sphere) \tag{3}$$

For a particle with arbitrary shape, the hydrodynamic radius, R_h (also called the Stokes radius), can be calculated using the Stokes-Einstein equation

$$R_h = \frac{kT}{6\pi\eta D} \tag{4}$$

The hydrodynamic radius of a particle is an abstraction defined as the radius of a hypothetical spherical particle with the same diffusion coefficient of the actual particle under consideration. It is important to recognize that "particle sizing" for non-spherical particles based on measurement of the diffusion coefficient and the Stokes-Einstein equation does not refer to any actual geometric property of the particle. R_h is used widely because of the intuitive appeal to its units of length, in contrast to the units of the diffusion coefficient, which in the cgs system are $cm^2\ s^{-1}$. Additionally, because temperature and viscosity are components of Eq. 4, R_h is independent of temperature and viscosity if particle undergoes

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

no structural changes over the relevant ranges of these parameters. Subject to the same caveat, diffusion coefficients are adjusted to the standard condition of 20°C in solvent water using

$$\frac{D}{D_{20,w}} = \frac{T}{T_{20}}\frac{\eta_{20,w}}{\eta} \tag{5}$$

where T and T_{20} are the absolute experimental temperature and temperature at 20°C, respectively, and η and $\eta_{20,w}$ are the corresponding solvent viscosities (Tanford, 1961).

Diffusion measurements of particles, including macromolecules, were originally made by measuring concentration gradients (Miller, 2014). In free diffusion, a sharp boundary is formed between a solution containing particles and pure solvent in a one-dimensional diffusion cell, and the concentration gradient is measured at various positions in the cell by optical methods. The most accurate measurements have been made using Gouy interferometry and Rayleigh interferometry (Annunziata *et al.*, 2005; Miller, 2014), which are unparalleled in accuracy and precision. While conceptually simple, free diffusion measurements are extremely demanding in practice because of the care that must be taken to avoid convective mixing, the specialized optics employed, and the analytical methods used in interpreting the interference patterns. Additionally, diffusion is relatively slow and must be allowed to proceed for hours to days to achieve a suitable concentration gradient, which limits throughput.

In contrast, diffusion measurements by DLS are easy to make, but the underlying physics is conceptually more difficult. For reviews of DLS, see references (Finsy, 1994; Stetefeld *et al.*, 2016; Ogendal, 2017). Qualitatively, light interacts with a non-absorbing particle and is scattered quasi-elastically (*i.e.*, at the nearly same wavelength as the incident light) in all directions. Diffusing particles can be moving either toward or away when struck by an incident photon. As a result, the scattered light is Doppler shifted to different frequencies, producing constructive and destructive interference and a fluctuating intensity. The frequency of the fluctuations increases with increasing diffusion rate, which is the underlying principle of the DLS method.

The scattered light intensity measured a short period after some arbitrary initial time is a function of the initial intensity. This relationship deteriorates as the particles randomly move about until they finally have no memory of their initial positions. The functional relationship describing the temporal dependence of the scattered light intensity due to diffusion is called autocorrelation (Berne and Pecora, 1976) and is given by the intensity autocorrelation function

$$G_2(\tau) = \langle I(t)I(t+\tau)\rangle \tag{6}$$

where $I(t)$ and $I(t+\tau)$ are the intensity of scattered light between an initial time t and some later time τ, called the decay time. The brackets indicate averaging over many initial times. The normalized *intensity* autocorrelation function is a dimensionless quantity reported by a DLS instrument and is defined as

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

$$g_2(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)\rangle^2} \quad (7)$$

The diffusion coefficient is related to the scattered electric field, not the scattered intensity, and specifically to the normalized *field* autocorrelation function, defined as

$$g_1(\tau) = \frac{\langle E(t)E(t+\tau)\rangle}{\langle E(t)^2\rangle} \quad (8)$$

For a monodisperse system, the normalized field autocorrelation function is

$$g_1(\tau) = exp(-\Gamma\tau) \quad (9)$$

Where Γ is called the decay rate. For particles sufficiently small relative to the incident wavelength of light, the decay rate is related to the diffusion coefficient by

$$\Gamma = Dq^2 \quad (10)$$

where q is the magnitude of the scattering vector,

$$q = \frac{4\pi n}{\lambda} sin\left(\frac{\theta}{2}\right) \quad (11)$$

and where n is the refractive index of the solvent, λ is the wavelength of the incident light, and θ is the scattering angle.

Thus, $g_2(\tau)$ is measured, but $g_1(\tau)$ is needed to calculate a diffusion coefficient. The conversion is accomplished using the Siegert relationship

$$g_2(\tau) = B + \beta[g_1(\tau)]^2 \quad (12)$$

As derived B equals 1, but varies experimentally due to noise and becomes a fitted parameter. β, called the coherence factor, depends on the experimental geometry and also is a fitted parameter. Several factors complicate the interpretation of the $g_2(\tau)$ decays. Macromolecular solutions are usually polydisperse, either intrinsically (*e.g.*, the microheterogeneity of glycoproteins) or due to impurities. Additionally, they are invariably studied as multi-component systems in the presence of salts and buffers. Scattering intensity increases approximately with the 6th power of the hydrodynamic radius. Consequently, the presence of aggregates or other impurities larger than the particle of interest can dominate the DLS measurement. Finally, experimental noise is a feature of all measurements and is particularly troublesome in the interpretation of DLS results.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

The simplest approach to deal with polydispersity is assuming some finite number, k, of scattering macromolecules and fit the data to a sum of exponential decays, producing a normalized field autocorrelation function,

$$g_1(\tau) = \sum_{i=1}^{k} exp(-\Gamma_i \tau) \quad (13)$$

Although this is done in some fitting programs, it is more customary to assume that there is a continuous distribution of decays, $G(\Gamma)$, given by

$$g_1(\tau) = \int_0^{\infty} G(\Gamma) exp(-\Gamma\tau) d\Gamma \quad (14)$$

The most widely used method for fitting experimental data to this integral is cumulant analysis (Koppel, 1972). There are several methods of cumulant analysis. One approach, introduced by Frisken (2001) uses the Siegert relationship, yielding a function in which $g_2(\tau)$ vs. τ decays are fit by nonlinear least-squares regression:

$$g_2(\tau) = B + \beta exp(-2\bar{\Gamma}\tau)\left(1 + \frac{\mu_2}{2!}\tau^2 - \frac{\mu_3}{3!}\tau^3 + \cdots\right)^2 \quad (15)$$

Here $\bar{\Gamma}$, the mean (the first moment) of the distribution, is the z-average diffusion coefficient (Koppel, 1972),

$$D_z = \sum_k c_k m_k D_k \Big/ \sum_k c_k m_k \quad (16)$$

and c_k, m_k, and D_k are the total cell concentration, mass, and diffusion coefficient, respectively, of species k. μ_2, μ_3 … are the second and third moments, *etc.*, about the mean. The second moment produces the polydispersity index,

$$PDI = \frac{\mu_2}{\bar{\Gamma}^2} \quad (17)$$

Because of measurement error, even a truly monodisperse solution will produce a non-zero estimate of μ_2. Polydispersity indices of less than 0.1 commonly are given as a criterion for monodispersity (Brautigam, 2019). Due to the errors associated with higher order moments, meaningful estimates beyond μ_2 are seldom, if ever, made.

For a monodisperse system, the molar mass, M, of a macromolecule (in cgs units of g/mol, which are used for conversion to molecular weight in daltons) can be obtained using the Svedberg equation

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

$$M = \frac{sRT}{D(1 - \bar{v}\,\rho)} \tag{18}$$

Where s and $\bar{v}$ are the sedimentation coefficient and partial specific volume of the macromolecule, R is the gas constant, and ρ is the solvent density (Cantor and Schimmel, 1980). For a polydisperse system, sedimentation velocity measurements produce a weight-average sedimentation coefficient, s_w. The value of s_w, extrapolated to the condition of solvent water at 20°C and at infinite dilution, $(s_w^0)_{20,w}$, is estimated by measuring the concentration dependence of s_w and extrapolation. The corresponding value for the diffusion coefficient is $(D_z^0)_{20,w}$. Substitution of these values in the Svedberg equation yields a weight-average molar mass, M_w, of a polydisperse system (Pusey *et al.*, 1974).

Some authors state that cumulant analysis is only applicable to a distribution function that is Gaussian, which is incorrect. Other authors state that cumulant analysis is only applicable to monomodal functions. This also seems incorrect. However, the utility of the parameter estimates derived from a polydisperse system that produces a multi-modal is limited. For this reason, another approach is to try to fit the data to estimate the function, $G(\Gamma)$, in the integral in Eq. 14, which is the distribution of all scattering particles in the system. This integral is an example of a Fredholm integral that arises in many physical systems. It does not have an analytic solution and is famously ill-posed to solve numerically. Considerable effort has been made to address the problem, and various methods are available in DLS software programs to calculate it.

Sizing and characterization of the polydispersity of "nanomaterials," defined as particles having a "size" approximately between 1 and 100 nm (Bhattacharjee, 2016), has found application in many commercial areas. Consequently, low-cost DLS instruments featuring automated data collection and analysis have become widely available. This is both a blessing and a curse in academic applications because the non-specialist investigator has access to high quality, easy to obtain data but is intellectually far removed from the underlying technology and analysis. For example, the Malvern Zetasizer Nano S particle size analyzer, which is described in this protocol, does not even report a diffusion coefficient, but rather a hydrodynamic radius or diameter. Additionally, the documentation associated with the instrument provides limited information with respect to the properties of the instrument and methods of analysis of the experimental results.

As a result, as is often the case in the modern laboratory, in addition to the subject under study, DLS measurements open up a separate additional research project aimed towards understanding the instrument and its associated software. In this protocol, a method is described to validate a commercial DLS instrument with the goal of accurately estimating macromolecular diffusion coefficients. Bovine serum albumin (BSA) is chosen as a reference material because previous measurements have been made of its diffusion coefficient using free diffusion (Creeth, 1952 and 1958; Wagner and Scheraga, 1956) and other DLS instruments (Oh and Johnson, 1981; Gaigalas *et al.*, 1992).

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

Materials and Reagents

1. Bullet tubes, graduated 1.5 ml, natural (Fisher Scientific, catalog number: 05-408-129)
2. Centrifuge tubes, Corning 15 and 50 ml conical tubes (Millipore Sigma, catalog numbers: 430052 and 430290)
3. Dialysis tubing, Spectrapor Standard RC, 12-14 kD cutoff, 16 mm diameter (Fisher Scientific, catalog number: 132678)
4. Gloves, Powder-free (Cobalt, catalog number: N-193)
5. Parafilm, 4 in. (Fisher Scientific, catalog number: 13-374-12)
6. Pipet tips, Avant Premium Low Binding 1-200 µl (Midwest Scientific, catalog number: AVR2)
7. Syringes, Norm-Ject 1 ml Tuberkulin (Henke Sass Wolfe, catalog number: 4010-200V0)
8. Syringe filters, Whatman Anotop 10, 0.02 µm (Cytiva, catalog number: 6809-1022)
9. Vacuum filter system, Corning 500 ml, 0.2 µm filters (Millipore Sigma, catalog number: CLS431097)
10. Weighing dishes, Polystyrene, 67 mm (Fisher Scientific, catalog number: 02-202-101)
11. Pyrex Media Storage bottles, 1 L (Fisher Scientific, catalog number: S14255)
12. Kimwipes (Fisher Scientific, catalog number: 06-666)
13. Bovine Serum Albumin, Fraction V, RIA and ELISA grade (Millipore Sigma, catalog number: 126593)
14. Ethanol, 200 proof (Decon Labs, catalog number: 2716)
15. HCl, ACS (Carolina Biological Supply, catalog number: 867792)
16. Hellmanex III (Fisher Scientific, catalog number: 14-385-864)
17. KH_2PO_4 (EMD Millipore, catalog number: 529568)
18. Na_2HPO_4 (JT Baker 99.6%, catalog number: 3828-5)
19. NaCl (JT Baker Ultrapure, catalog number: 4058-07)
20. NaOH (solid) (Acros Organics, catalog number: 450580025)
21. 10× PBS (1 L) (see Recipes)
22. PBS (1 L) (see Recipes)

Equipment

1. Analytical balance, Mettler AE100 (Mettler Toledo, no longer manufactured, substitute similar analytical balance)
2. Centrifuge (Beckman Coulter Microfuge 18, no longer manufactured, substitute similar high speed microcentrifuge)
3. Chromatography system, AKTA Pure 150 (Cytiva Life Sciences, catalog number: 29046665)
4. Computer, Dell Precision 3431 Workstation, Windows 10 operating system (no longer manufactured, substitute similar computer)
5. Cuvette, Malvern Pananalytical ZEN 2112 (Fisher Scientific, catalog number: NC9289172)

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

6. Fraction collector, F9-R (Cytiva Life Sciences, catalog number: 29011362)
7. Freezer, -86 to -50°C, VWR model 40086D (VWR International, catalog number: 10002-786)
8. Magnetic stir bars, octagonal, PTFE, 1 in. (Fisher Scientific, catalog number: 14-513-51)
9. Magnetic stirrer, Corning PC-17 (Millipore Sigma, catalog number: CLS6795171)
10. Particle size analyzer, Zetasizer Nano S (Malvern Pananalytical)
11. pH meter, Thermo Scientific OrionStar A211 (Fisher Scientific, catalog number: 13-645-519)
12. Pipets, Gilson Pipetman, 20 µl & 200 µl (Fisher Scientific, catalog numbers: F123600 and F123601)
13. Size-exclusion chromatography column, HiLoad 16/600 Superdex 200 prep grade (Cytiva Life Sciences, catalog number: 28989335)
14. Spectrophotometer, Beckman DU650 (no longer manufactured, substitute spectrophotometer with UV absorbance capability)
15. Top loading balance, Mettler Toledo XS4002S (Mettler Toledo, no longer manufactured, substitute similar top loading balance)
16. Vacuum desiccator, Nalgene 280 mm × 251 mm (Millipore Sigma, catalog number: D2797)
17. Viscometer, Lovis 200 M (optional) (Anton Paar)
18. Water bath, FisherVersa Bath (no longer manufactured, substitute similar bath)
19. Water purification system, Synergy UV-R producing ultrapure (Type I) water (EMD Millipore)

Software

1. Zetasizer Software, version 8.01 (Malvern Pananalytical, https://www.malvernpanalytical.com/en/)
2. SEDFIT, version 15.36 (Schuck, P., http://www.analyticalultracentrifugation.com/)
3. Excel 2016 (Microsoft, https://www.microsoft.com/zh-cn/)
4. SEDNTERP, version 20130813BETA (Hurton, 2012; http://www.rasmb.org/sednterp/)

Procedure

A. Purification of monomeric BSA by size-exclusion chromatography (SEC)

Fraction V, lyophilized BSA contains BSA dimers and trimers in addition to 66 kDa BSA monomers. It also contains H_2O and other unknown low molecular weight excipients.

1. Preparation of crude BSA for SEC
    a. Add 3 ml 0.154 M NaCl, 5.6 mM Na_2HPO_4, and 1.06 mM KH_2PO_4, pH 7.4 (PBS) to a 15 ml conical tube (see Recipes).
    b. Weigh ~60 mg BSA in a weigh boat using an analytical balance.
    c. Add the BSA to the conical tube. It dissolves readily.
    d. Wash the weigh boat with the BSA solution to collect more material. Add to the conical tube.
    e. Measure the concentration of BSA.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

Make a 1/50 dilution in PBS.

Read the absorbance at 280 nm and 320 nm.

Calculate the concentration of BSA in mg/ml:

$$c = \frac{A_{280}}{\varepsilon_{280} b}$$

where b is the pathlength and using an extinction coefficient at 280 nm, ε_{280}, of 0.678 $(mg/ml)^{-1}$ cm^{-1} for BSA. Multiply by the dilution factor to get the concentration of undiluted BSA. It should be about 12 mg/ml.

f. Soak an 8 in. length of SpectraPor 16 mm diameter dialysis tubing in H_2O for ~30 min.

g. Rinse the tubing with approximately 50 ml of H_2O, then once with 10 ml PBS.

h. Knot one end of the tubing with three knots, add 3 ml of the BSA solution and seal by adding three knots to the other end of the tubing.

i. Dialyze against 1 L PBS for 6 h, then overnight against 1 L fresh PBS at room temperature. Use a 1 L beaker covered with Parafilm on a stir plate with a stir bar rotating to gently swirl the dialysis tubing.

j. After dialysis is complete, carefully cut the top of dialysis tubing with scissors, remove the BSA solution with a pipet, and add to a 5 ml test tube.

k. Measure the BSA concentration as described above to check for loss of material.

2. SEC of BSA

a. Wash and equilibrate a Superdex 200 SEC column at 1 ml/min with 60 ml H_2O, followed by 60 ml 0.5 M NaOH, 60 ml H_2O, and 375 ml PBS, all at room temperature.

b. Load ~1.5 ml of ~12 mg/ml BSA at 1 ml/min using a sample loop.

c. Collect 45 ml in a single fraction and collect 0.8 ml fractions at 1 ml/min.

d. Read the $A_{280/320}$ of peak fractions (Figure 1).

e. Pool the central eight fractions from the major peak into a 50 ml conical tube.

f. Measure the concentration spectrophotometrically, as described above.

g. Dilute to 0.53 mg/ml by adding the PBS SEC buffer.

h. Confirm the concentration spectrophotometrically.

i. Add 0.5 ml aliquots to 1.5 ml bullet tubes, place in a freezer box, and place in a -80°C freezer.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

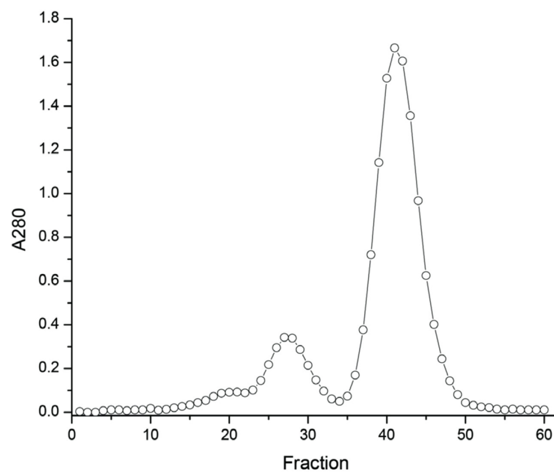


Figure 1. Superdex 200 SEC of BSA

B. Dynamic Light Scattering

The Zetasizer Nano S DLS instrument (Figure 2) uses a 632.8 nm He Ne red laser. It houses a single cuvette and measures light scattered at 173°. It is interfaced to a computer running Windows 10 and controlled with Zetasizer software. The scattering volume in the cuvette is automatically determined to allow for a large range of particle concentrations. Alternatively, the user can specify the position manually. A measurement consists of several averaged normalized intensity autocorrelation decays, called runs, each lasting 10 s. The number of measurements is specified by the user. The number of runs can either be determined automatically by the software based on the intensity of scattered light, which depends on the concentration and size of the macromolecule, or entered manually by the user. For BSA at 0.53 mg/ml, the number of runs is typically around 15. The runs are averaged in a manner that is not well described in the instrument documentation to produce a single z-average particle radius and polydispersity index. Additionally, the software produces an intensity particle diameter that is calculated using a user-specified algorithm that inverts the Fredholm integral (Eq. 14).

In this protocol, a frozen BSA sample is thawed and prepared for analysis by ultrafiltration and high-speed centrifugation to remove possible large interfering scattering particles, as described below. A trial of four measurements is made on an aliquot. This process was repeated with fresh aliquots from the same thawed bullet tube for a total of eight trials to assess the variation due to aliquoting. Additionally, another sample of BSA is thawed and also subjected to eight trials, each with four measurements to assess possible variation due to freezing, thawing, and sample preparation.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

Figure 2. Zetasizer Nano S particle size analyzer. The ZEN2112 cuvette is shown next to the open sample compartment.

1. Instrument start-up
   a. Turn on the computer.
   b. Turn on the Zetasizer Nano S by pressing the power switch on the back of the instrument.
   c. The status light on the Zetasizer Nano S will initially be red.
   d. Start the Zetasizer Nano S software.
   e. The status light on the Zetasizer Nano S will become green.
   f. Select Tools > Count Rate Meter, select a temperature of 20°C, and enter Set.
   g. Place a ZEN2112 cuvette (Figure 2) in the instrument for temperature equilibration.
   h. Wait at least 30 min for the instrument to stabilize and the temperature to equilibrate.
2. Preparation of a standard operating procedure (SOP)
   Measurements in the Zetasizer Nano S are made according to an SOP that is created by the user. The SOP specifies 1) the temperature and solvent viscosity, on which the diffusion rate depends and which are required for the calculation of hydrodynamic radius (Eq. 4), 2) the solvent refractive index, which is required for the calculation of the scattering vector (Eq. 11), 3) the number of measurements, and 4) the algorithm used in the calculation of the intensity distribution (Eq. 14). The viscosity for most common buffers can be estimated using SEDNTERP. Alternatively, it can be measured using a viscometer. The latter is preferable because errors due to the preparation of the buffer are not identified by SEDNTERP. The refractive index of buffers near physiologic ionic strength is only slightly increased from water, which is 1.333 at 632.8 nm laser wavelength and 20°C. For PBS used in this protocol, a value of 1.334 is used.
   a. Select File > New > SOP. A New SOP window appears (Figure 3):

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

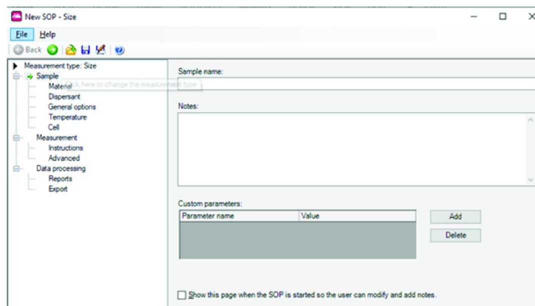


Figure 3. Zetasizer SOP window

i. Measurement type: Select Size
ii. Material: Select Protein. (This selection is irrelevant for particles with a hydrodynamic radius less than 5% of incident wavelength, as is the case for BSA. However, there is no harm in selecting the Protein option.)
iii. Dispersant: The dispersant is the solvent. A window appears with the default, solvent water, or the last used dispersant (Figure 4A).
    1) Click the "…" (three dots) icon, and the Dispersants Manager window appears (Figure 4B).
    2) Select an existing solvent. A window will appear with the temperature, viscosity, and refractive index.
    3) Alternatively, add a new solvent by selecting Add… > Simple dispersant or solvent. A Dispersant Properties window will appear (Figure 4C). Enter a name for the solvent, the temperature, solvent viscosity, and refractive index. The dielectric constant is irrelevant, but a number must be entered. Enter 50. The new solvent will appear in the Dispersants Manager window.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

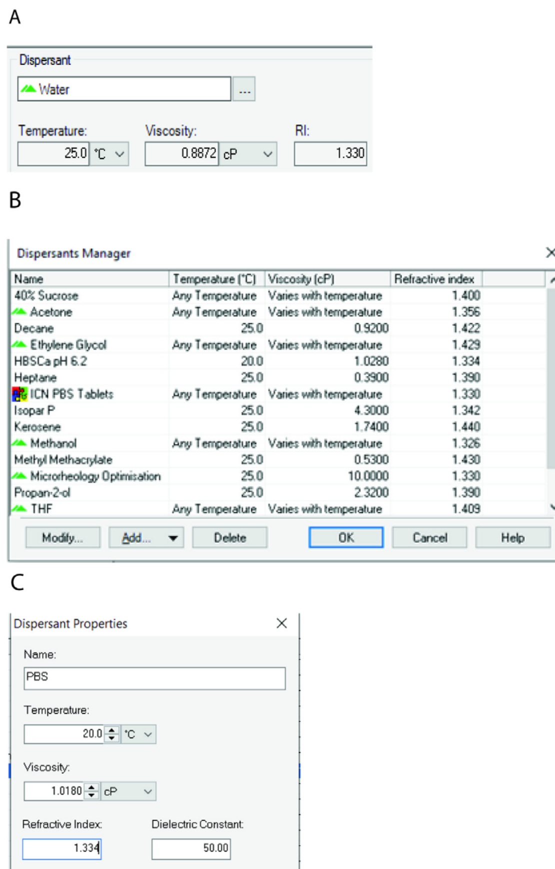


Figure 4. Dispersant properties. (A) Default Dispersant window. (B) Dispersants Manager window. (C) Dispersant Properties window.

iv. General Options: Select Use dispersant viscosity as sample viscosity. (The Mark-Houwink parameters are irrelevant for this analysis.)

v. Temperature.

1) The temperature is set by selecting the dispersant (see above).

2) Equilibration time: Enter 30 s, which is sufficient for samples sitting at room temperature close to the measurement temperature of 20°C.

vi. Cell > Cell type > Quartz cuvettes Select ZEN2112.

vii. Measurement

1) Measurement angle: Select 173° Backscatter

2) Measurement duration: Select Automatic

3) Number of measurements: Enter 4

4) Delay between measurements: Enter 0 s

5) Append measurement number to sample name: Check the box

6) Artial results: Do not check the box

7) Measurement duration > Extend duration for large particles. Enter No

8) Positioning method: Select Seek for optimum position

9) Automatic attenuation selection: Enter Yes

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

viii. Data Processing > Analysis model Select Protein Analysis
Configure: Select Display range 0.1 to 6,000, upper and lower thresholds 0.01 and 0.05 (these are the default values).

b. Save the SOP using File > Save as.
*NOTE: Be sure to add the extension *.sop, the program will not do it automatically.*

c. Place the SOPs in the folder \Malvern Instruments\Zetasizer\SOP\Protein. This folder is the default folder the program looks for to find an SOP file.

d. Before every experiment, check the SOP for the correct settings by selecting File > Open > SOP > Protein and choosing the SOP.

3. Preparation of a sample for measurement
   a. Remove a 0.5 ml aliquot of purified BSA from the -80°C freezer and thaw 15 min in a 37°C water bath.
   b. Remove the plunger of a Henke Sass Wolf 1 ml Tuberkulin syringe.
   c. Place a 0.02 µm Whatman Anotop10 filter on the syringe.
   d. Add the BSA to the barrel and filter into a 1.5 ml bullet tube.
   e. Centrifuge at 18,000 × *g* for 20 min at room temperature.
   f. Remove the top 0.45 ml and pipet into a fresh 1.5 ml bullet tube.
   g. Remove the cuvette from the Zetasizer Nano S instrument.
   h. Add 20 µl of sample by placing the pipet tip at the bottom of the cell. Do not fully expel the contents to avoid air bubbles. This volume leaves the meniscus in the viewing window of the cuvette, which is satisfactory. It is acceptable if the meniscus is uneven.
   i. Cap the cuvette and place in the Zetasizer Nano S instrument.
   j. After the measurements (see below), remove and dispel the old sample, repeating the extraction once. If repeating a trial with the same material, simply add 20 µl of fresh sample. If adding a solution containing a different macromolecule or concentration, wash the cell three times with 20 µl of the new solution, extracting the washes as described above, and then add 20 µl for measurement.
4. Making DLS measurements
   a. When the Zetasizer software is started, a window appears entitled "Example results," which lists several records. Close the window by selecting File > Close. Initially, the arrow next to Browse for SOP… on the Toolbar is greyed out (Figure 5A).
   b. Select File > New > Measurement file and give the file a name. A .dts file will be created.
   c. Save the file in the folder \Malvern Instruments\Zetasizer\Measurement Data. The small arrow next to Browse for SOP… becomes green (Figure 5B).
   d. Click on the small green arrow. A window appears for the selection of the SOP.
   e. Navigate to \Malvern Instruments\Zetasizer\SOP\Protein and select the appropriate SOP file. A Documentation window appears (Figure 5C). Enter a name for the sample. If a window appears that does not have OK at the bottom, selecting the SOP has been done incorrectly. Repeat the steps above.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

f. After selecting OK, a window appears with a big green arrow (Figure 5D). Click it to start the set of measurements. Initially, when the arrow is clicked, the measurements start after the user-specified equilibration period. On repeat trials, a message appears directing the user to insert the cell. After inserting the cell, click the big green arrow again, and the measurements will start.

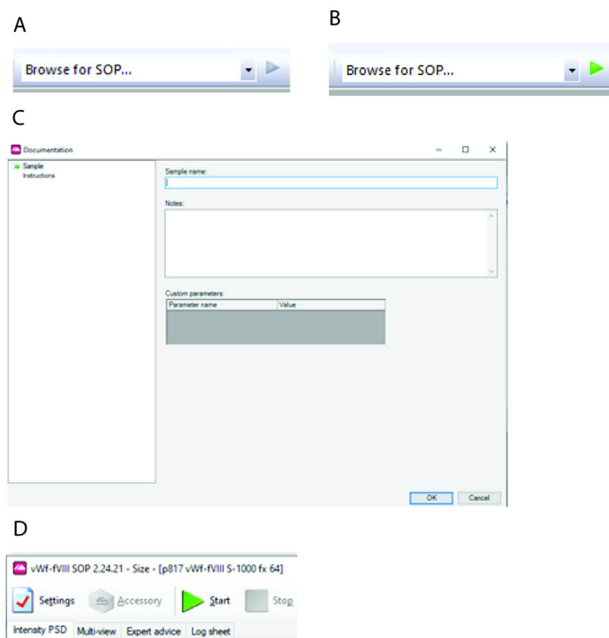


Figure 5. Making DLS measurements. (A) initial Browse for SOP window. (B) Browse for the SOP window after a new measurement file has been created. (C) Documentation window. (D) Big green start arrow.

g. As a series of runs accumulates into a measurement, select Intensity PSD, and a graph appears showing the distribution analysis as Intensity versus Size and cumulant analysis, showing the Z-average diameter and polydispersity index Figure 6A. At the completion of the run, the distribution function is calculated, and Intensity vs. Size is plotted. Then another run is executed, with continual updating of the intensity vs. size graph until the runs have been completed, which constitutes a single measurement.

h. Alternatively, during measurements, select Multi-View, and a window appears showing: 1) $g_2(\tau)-1$ versus decay time τ, which is updated as a run evolves; 2) the intensity fluctuations described in Background; and 3) the intensity distribution (Figure 6B). The intensity distribution for a monodisperse protein preparation should be monomodal and symmetric, as shown in Figures 6A and 6B. An example of polydisperse sample is shown in Figure 7, which was prepared by mixing BSA with thyroglobulin. If the intensity distribution is multimodal, the sample preparation should be repeated.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

Please cite this article as: Parker, E. T. and Lollar, P. (2021). Measurement of the Translational Diffusion Coefficient and Hydrodynamic Radius of Proteins by Dynamic Light Scattering. *Bio-protocol* 11(20): e4195. DOI: 10.21769/BioProtoc.4195.

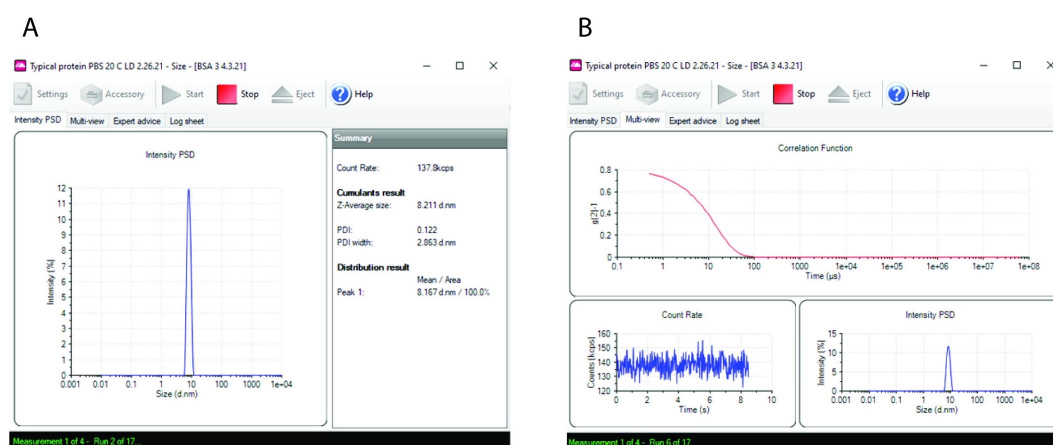


Figure 6. Measurement windows. (A) Intensity PSD. (B) Multi-view.

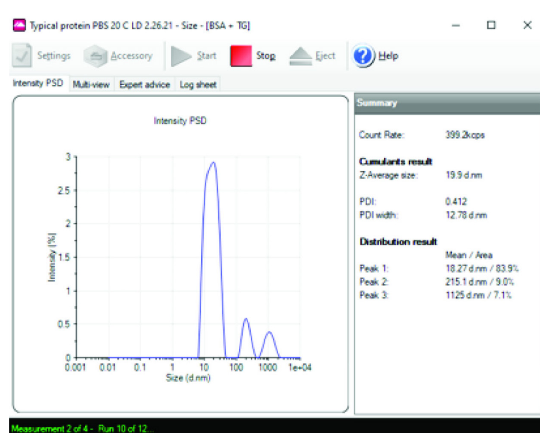


Figure 7. Sample polydispersity. BSA and crude, unfiltered thyroglobulin were mixed at concentrations of 0.25 and 0.13 mg/ml, respectively.

i. After the measurements are complete, the measurement window displays the message Insert cell and press start when ready. Close it by clicking the X in the upper right corner.
j. Close the window showing the records view by entering File > Close or clicking the red X on the upper right part of the window.
k. Start another trial using the same SOP. Select File > New > Measurement file and enter a new file name. The existing SOP should still be present. Click the small green arrow. (NOTE: If the small green arrow is not present, the measurement window probably has not been closed.) When the SOP window appears, change the name of the sample and select OK as described above. The big green arrow should appear to start the next trial.

5. Shutdown
   a. Exit Zetasizer software (File > Exit).
   b. Turn off the Zetasizer.
   c. Shut down computer.
6. Cleaning the Zen2115 cuvette
   Wear powder-free gloves. Use a 200 µl pipette and 200 µl volumes.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

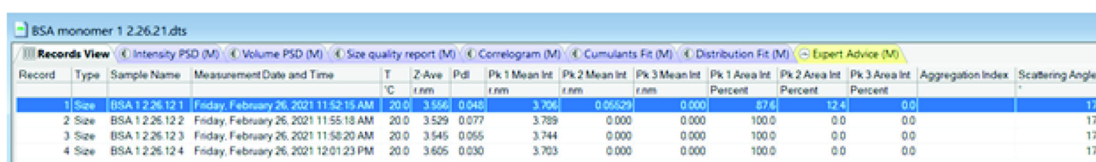
Please cite this article as: Parker, E. T. and Lollar, P. (2021). Measurement of the Translational Diffusion Coefficient and Hydrodynamic Radius of Proteins by Dynamic Light Scattering. *Bio-protocol* 11(20): e4195. DOI: 10.21769/BioProtoc.4195.

a. Withdraw the cuvette contents.
b. Wash 5× with H_2O.
c. Wash 5× with 2% (v/v) Hellmanex III.
d. Wash 5× with H_2O.
e. Wash 5× with H_2O:HCl:EtOH 2:1:1 (v:v:v).
f. Wash 3× with H_2O.
g. Wash 1× with 200 proof EtOH.
h. Dry the cuvette in a vacuum desiccator.

Data analysis

A. Zetasizer analysis

1. At the completion of a set of measurements, select the Records View tab. Each Record corresponds to a single measurement. It lists the Z-average hydrodynamic radius, R_h, (or diameter if the user prefers) and polydispersity index (PdI) obtained by cumulant analysis and the hydrodynamic radii (or diameters) associated with peaks identified using distribution analysis (Figure 8). The results can be viewed either from individual measurements or collectively by selecting all measurements with the mouse.


BSA monomer 1 2.26.21.dts

Records View | Intensity PSD (M) | Volume PSD (M) | Size quality report (M) | Correlogram (M) | Cumulants Fit (M) | Distribution Fit (M) | Expert Advice (M)

Record	Type	Sample Name	Measurement Date and Time	T	Z-Ave	PdI	Pk 1 Mean Int	Pk 2 Mean Int	Pk 3 Mean Int	Pk 1 Area Int	Pk 2 Area Int	Pk 3 Area Int	Aggregation Index	Scattering Angle
				°C	r.nm		r.nm	r.nm	r.nm	Percent	Percent	Percent		°
1	Size	BSA 1 2.26.12 1	Friday, February 26, 2021 11:52:15 AM	20.0	3.556	0.048	3.706	0.05629	0.000	87.6	12.4	0.0		173
2	Size	BSA 1 2.26.12 2	Friday, February 26, 2021 11:55:18 AM	20.0	3.529	0.077	3.789	0.000	0.000	100.0	0.0	0.0		173
3	Size	BSA 1 2.26.12 3	Friday, February 26, 2021 11:58:20 AM	20.0	3.545	0.055	3.744	0.000	0.000	100.0	0.0	0.0		173
4	Size	BSA 1 2.26.12 4	Friday, February 26, 2021 12:01:23 PM	20.0	3.605	0.030	3.703	0.000	0.000	100.0	0.0	0.0		173

Figure 8. Records View

2. Ensure that the correct solvent was used by right-clicking on a measurement file and selecting Edit Result. A window will appear with solvent (dispersant) properties. If the wrong solvent has been used, make the correction by modifying the SOP.
3. To view results as radius instead of diameter, click d.nm and change to r.nm.
4. To get statistics, select all the records and means, and sample standard deviations for the set of measurements will appear.
5. Select the Intensity PSD tab, which opens up a table that includes the count rate, measurement position in the cuvette, attenuation and size information, data relevant to the measurement (Figure 9A), and a graph of the particle size distribution (Figure 9B). The Volume PSD is not relevant to measuring the diffusion coefficient of proteins. The Size Quality Report evidently is primarily for commercial applications.
   a. Copy the z-average R_h and Pk1 Mean intensity R_h values (Figure 8) and paste into an Excel file for later use. There will be 32 values corresponding to four measurements on eight aliquots.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

A

Sample Name:	BSA 1 2.26.12 2		
SOP Name:	Typical protein PBS 20 C LD 2.26.21		
File Name:	BSA monomer 1 2.26.21	**Dispersant Name:**	Schuck PBS
Record Number:	2	**Dispersant RI:**	1.334
Material RI:	1.45	**Viscosity (cP):**	1.0180
Material Absorbtion:	0.001	**Measurement Date and Time:**	Friday, February 26, 2021 11:55:18 AM

Temperature (°C):	20.0	**Duration Used (s):**	90
Count Rate (kcps):	112.1	**Measurement Position (mm):**	4.20
Cell Description:	Low volume glass cuvette (12µL)	**Attenuator:**	11

			Size (r.nm):	**% Intensity:**	**St Dev (r.nm):**
Z-Average (r.nm):	3.529	**Peak 1:**	3.789	100.0	0.9131
PdI:	0.077	**Peak 2:**	0.000	0.0	0.000
Intercept:	0.857	**Peak 3:**	0.000	0.0	0.000
Result quality :	Good				

B

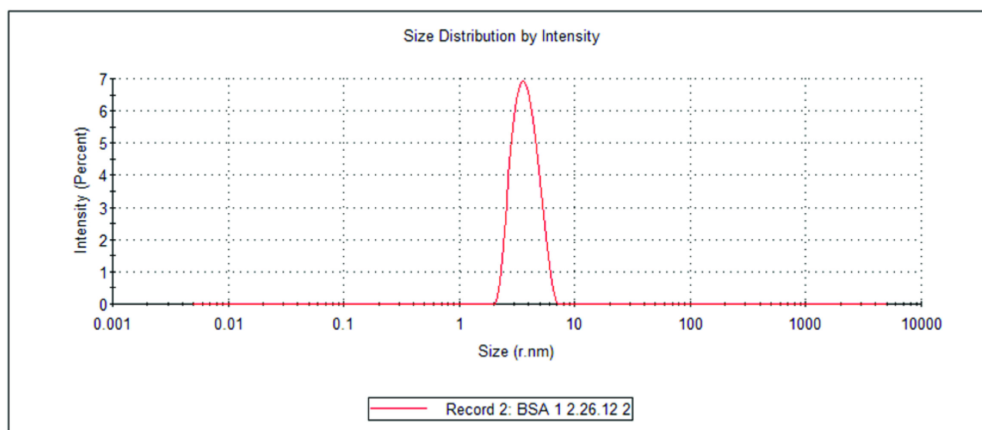


Figure 9. BSA DLS measurement results. (A) Tabulated results. (B) Particle size distribution.

6. Select Correlogram (M), which displays a graph of Correlation Coefficient versus Time (Figure 10A). “Correlation Coefficient” is the quantity $g_2(\tau)-1$. Time is the decay time, τ.
7. Select Cumulants Fit (M), and a graph appears showing G1 Correlation Function versus Time data fitted using cumulant analysis (Figure 10B). The G1 Correlation Function is the normalized field autocorrelation function, $g_2(\tau)$. Note that data are only fitted for $g1(\tau)$ values greater than 0.1.
8. Select Distribution Fit (M), and a graph appears showing G1 Correlation Function versus Time data fitted using distribution analysis (Figure 10C).

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

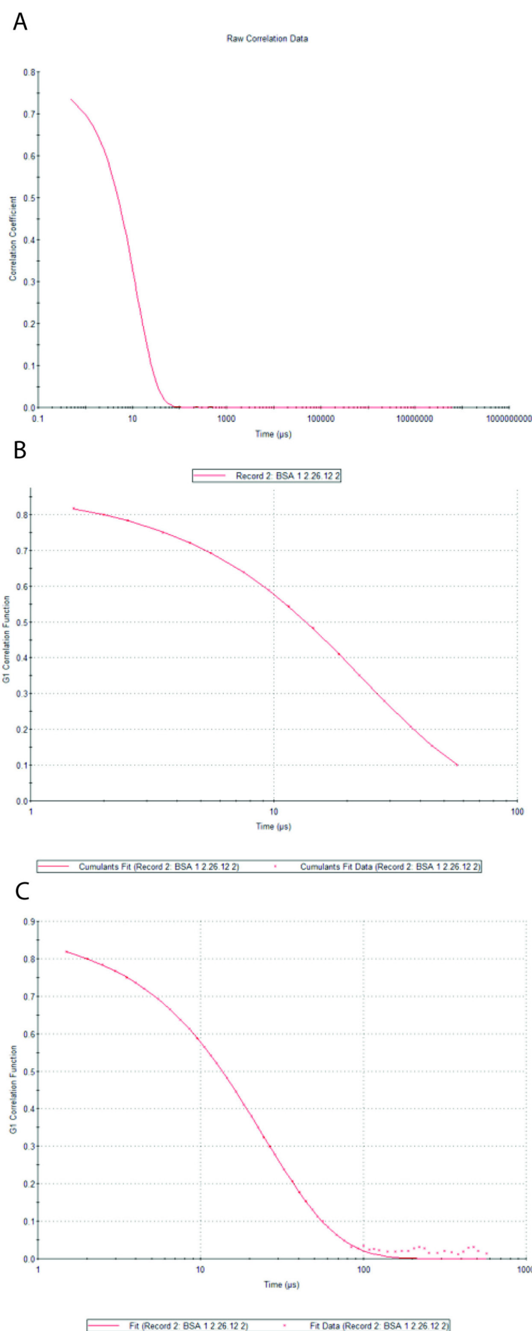


Figure 10. DLS decays. (A) $g_2(\tau)-1$ vs. τ. (B) Zetasizer cumulants fit. (C) Zetasizer distribution fit.

B. SEDFIT analysis

1. Converting Zetasizer files to SEDFIT compatible files

   Zetasizer record files, which contain the results of a single measurement, have the suffix .dts. They can be converted to files that SEDFIT can read using a macro in Zetasizer software under Tools > Macros > Make DLSDAT-v1-0.

   a. Create the .dlsdat macro

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

i. Go to www.materials-talks.com/blog/2015/10/28/new-macro-to-export-dls-data-to-sedfit/. Select Method 1.
ii. Download and unzip MakeDLSDAT-v1-0.zip by clicking on the link at the above site. MakeDLSDAT-v1-0.zmac appears in the /Downloads folder. Move it to the Desktop.
iii. Start the Zetasizer software.
iv. Go to Tools > Options > Macros > Install new Macro.
v. Select MakeDLSDAT-v1-0.zmac from the Desktop.
vi. Restart the Zetasizer software. Now under Tools > Macros, Make DLSDAT-v1-0 appears.

b. Using the dls.dat macro
i. In the Zetasizer software, select File > Open > Measurement File and navigate to the folder containing the desired .dts file. A .dts file contains a set of measurements, which appear as separate records. Each measurement can be converted to a .dlsdat file.
ii. Open a .dts file by selecting Open > Measurement file and highlight a record under Records View. Right-click Macros > MakeDLSDAT-v1-0 and enter a file name. The macro will produce a file with the extension .dlsdat for use with SEDFIT.

c. Properties of .dlsdat files
Each .dlsdat file contains a header and about 200 data pairs. Figure 11 shows an example of the header and the first few lines of a .dlsdat file.

```
A 19.998910 1 1.00 1.334000 0.010180 173.000000 6.328000e-07
 5.00000e-07 7.97300e-01
 1.00000e-06 7.67832e-01
 1.50000e-06 7.34034e-01
 2.00000e-06 7.05835e-01
 2.50000e-06 6.75456e-01
```

Figure 11. Structure of a .dlsdat file

In the header, 19.998910 is the temperature, 1.334 is the solvent refractive index, 0.010018 is the solvent viscosity in Poise, 173 is the scattering angle, and 6.328e-07 is the wavelength of the helium-neon laser in meters. SEDFIT calculates the scattering vector according to Eq 11. The first column of data is the delay time, τ, in seconds. The second column is $g_2(\tau)-1$.

2. SEDFIT cumulant analysis
a. Run SEDFIT, select Data > Load DLS data and load a .dlsdat file. SEDFIT takes the square root of $g_2(\tau)-1$ in the .dlsdata file to produce $g1(\tau)$ (which it calls g(1) field autocorrelation). The SEDFIT default is to graph and fit $g1(\tau)$ data. Instead, graph and fit $g_2(\tau)-1$ data by selecting Options > DLS Tools > Toggle Intensity <-> Field Autocorrelation Fct, which toggles the SEDFIT graph between "intensity autocorrelation function" and "field autocorrelation function".
b. Select Options > DLS Tools > Always Load Intensity Autocorrelation g(2)–1.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

c. Use the mouse to move the green vertical lines to set the fitting limits. Select the left fitting limit to include the first data point. Select a right fitting limit of about 1 log to the right of where the data has approached zero (Figure 12). Selecting more data does not improve the fitted constants.

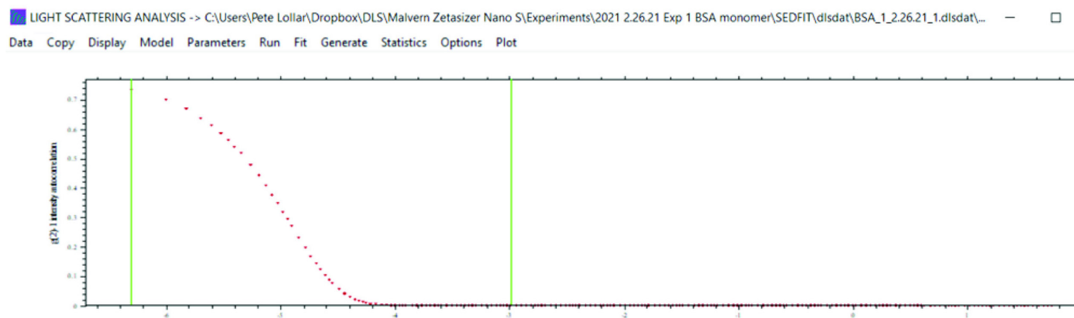


Figure 12. SEDFIT fitting limits

d. Select Model > Dynamic Light Scattering > Cumulant Analysis.

e. Select Parameters, and the default Cumulant Analysis window appears (Figure 13A). "I tot" refers to the value of $g_2(\tau)-1$ at zero τ. mu1/S^2 = Dz is $\bar{\Gamma}$ (Eq. 15). mu2 and mu3 are the moments μ_2 and μ_3 (Eq. 15). S is the magnitude of the scattering vector (q in Eq. 11). The units of the parameters are F, F^2, and F^3, respectively, where 1 F is 1 Fick and is equal to 1 $cm^2 s^{-1}$.

f. Fit I tot, Dz and mu2/S^4 and Baseline by checking the boxes and entering initial guesses. mu3/S^4 is not fitted. Enter a starting value for Dz above the anticipated value. The same value can be entered for mu2/S^4. Enter a value 1 of I tot and a value of zero for Baseline. In practice, the fitting is very robust with respect to initial estimates and reproducibly converges to the same values. The initial estimates are shown in Figure 13B.

g. Select Run. SEDFIT calculates $g_2(\tau)-1$ using the initial estimates and Eq. 15(Figure 13D). In this example, the initial fit follows the data to some extent but requires optimization, which is done by nonlinear least-squares analysis using the Marquardt Levenberg algorithm.

h. Select Options > Fitting Options > Marquardt Levenberg.

i. Select Fit. Now optimization of the parameters produces a close fit to the data (Figure 13E). (Select Display > Update to display only the graph).

j. Select Parameters to see the fitted constants (Figure 13C). In this example, D_z is 6.03 F.

k. Enter Display > Show Last Fit Info Again or type "Ctrl o," and the fitting information will be superimposed on the graph, including the root mean squared deviation (rmsd) of the fit and the polydispersity index (Figure 13F). In this example, the polydispersity is 0.073.

l. Some of the information displayed (not shown in the figure) is in reference to fitting sedimentation data and is irrelevant.

m. Record the D_z values in an Excel spreadsheet for later use. There will be 32 values corresponding to four measurements on eight aliquots.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

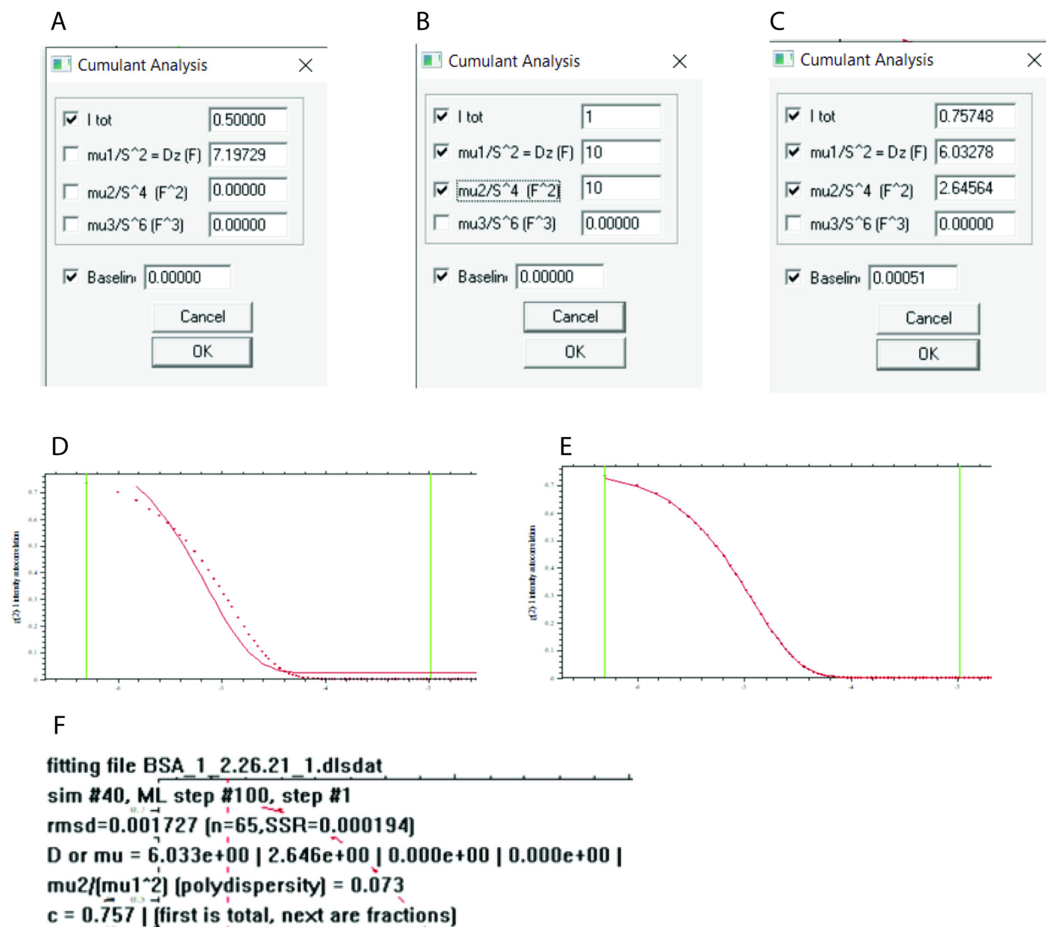


Figure 13. SEDFIT DLS cumulants fit. (A) Opening parameter window. (B) Initial parameter estimates. (C) Fitted parameters. (D) Calculated $g_2(\tau)-1$ values from initial estimates. (E) Fitted curve.

3. SEDFIT continuous I(D) distribution analysis
    a. In SEDFIT, load DLS data as described above under Data analysis B.2 and select Model > Dynamic Light Scattering > DLS:Continuous I(D)-Distribution.
    b. In contrast to cumulant analysis, in which fitting is done to $g_2(\tau)-1$ values, fitting must be done to $g1(\tau)$ in this model. Select Options > DLS Tools > Toggle Intensity <-> Field Autocorrelation Fct and toggle to "field autocorrelation function". The graph displays a plot of $g1(\tau)$ versus τ (not shown).
    c. Select Parameters. The Parameters window appears (Figure 14A).
    d. There are several possible combinations of starting parameter estimates. Select the default values equidistant log(1/tau) grid, integrate linear, and resolution equal to 100. Leave the confidence level (F-ratio) at the default value of 0.55 as recommended at www.analyticalultracentrifugation.com. The partial specific volume, buffer density, and buffer viscosity are irrelevant. Make initial starting estimates of D min and D max that are

Please cite this article as: Parker, E. T. and Lollar, P. (2021). Measurement of the Translational Diffusion Coefficient and Hydrodynamic Radius of Proteins by Dynamic Light Scattering. *Bio-protocol* 11(20): e4195. DOI: 10.21769/BioProtoc.4195.

10-fold below and above the estimated value of the diffusion coefficient. For BSA, values of 0.6 and 60 are selected (Figure 14B).

e. Select Run. An apparent idiosyncrasy of the program is that, in contrast to cumulant analysis, the Run command does a nonlinear least-squares fit. An upper window appears, showing the fit (Figure 14C). The lower window shows the fitted distribution (Figure 14D). As in the Zetasizer software, the presence of aggregates or impurities can produce several peaks indicative of polydispersity (not shown).

f. Calculate the signal-average diffusion coefficient by clicking the integrate button (Figure 14D, upper right corner) and selecting the region containing the distribution using the right mouse button as directed in the program. In this example, a signal average diffusion coefficient of 6.203 F was obtained (Figure 14D, *inset*).

g. Record the signal average diffusion coefficient values in an Excel spreadsheet for later use.

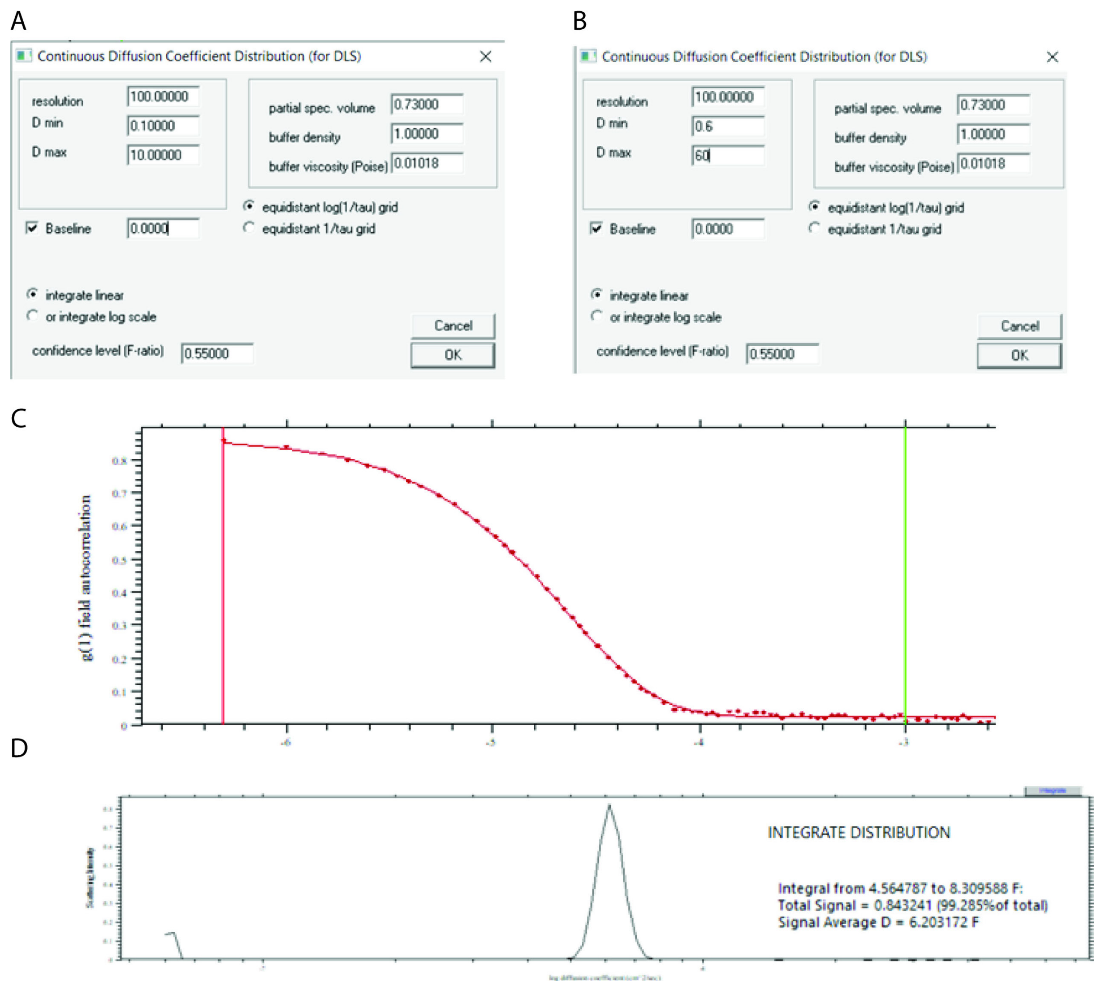


Figure 14. SEDFIT DLS continuous I(D) distribution fit. (A) Default continuous I(D) window. (B) User-selected continuous I(D) values. (C) Fitted $g1(\tau)$ decay curve. (D) Fitted continuous I(D) distribution, which is integrated to produce the signal-average diffusion coefficient.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

C. Excel Solver analysis

In this example, the Excel GRG unweighted nonlinear least-squares regression algorithm is used to fit the results from a single Zetasizer measurement of the BSA DLS decay to the first two moments of the Frisken cumulant equation (Eq. 15)

$$g_2(\tau) = B + \beta\, exp(-2\bar{\Gamma}\tau)\left(1 + \frac{\mu_2}{2!}\tau^2\right)^2$$

This protocol assumes familiarity with Excel and that the Solver add-in has been installed.

1. Run Excel and name a blank sheet "Master Sheet."
2. Create a table to calculate the magnitude of the scattering vector in cells I26:J32 (Figure 15A).

A

Scattering vector	
WL (nm)	632.8
WL (cm)	6.328E-05
Refractive index, n	1.334
theta (deg)	173
theta (rad)	3.0194
q (cm^-1)	2.64E+05

B

26	Parameter starting estimates	
27	D est (F)	20
28	D est (cm^2 s^-1)	2.00E-06
29	Gamma est (usec^-1)	1.40E-01
30	PDI est	0.4
31	Mu2 est (usec^-2)	7.82E-03

C

Parameter fits	Excel
B	1.00000
beta	1.00000
Gamma (usec^-1)	1.40E-01
Mu2 (usecs^-2)	7.82E-03
rmsd	9.840E-02
Diffusion coefficient and PDI	
D (cm^2 s^-1)	2.00E-06
D (F)	20.0000
PDI	0.4000

D

Parameter fits	Excel	SEDFIT
B	1.00051	1.00052
beta	0.75746	0.75747
Gamma (usec^-1)	4.22E-02	
Mu2 (usecs^-2)	1.29E-04	
rmsd	1.739E-03	1.740E-03
Diffusion coefficient and PDI		
D (cm^2 s^-1)	6.03E-07	
D (F)	6.0324	6.0329
PDI	0.073	0.073

E

Fit	
SSR	6.19633E-01
Data pairs	64

F

Fit	
SSR	1.93622E-04
Data pairs	64

Figure 15. Excel Solver DLS cumulant fit. (A) Scattering vector. (B) Parameter starting estimates. (C) Parameter fits before running Solver. (D) Parameter fits after running Solver with the SEDFIT fitted parameters shown for comparison. (E-F) SSR before and after running Solver.

3. In column I, enter row labels ("WL (nm)," "WL (cm)," *etc.*).
4. In column J, enter the values for the DLS laser wavelength, solvent refractive index, and scattering angle in degrees. Convert the scattering angle to radians. Use the Excel function

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

Define Name and define the names WL, RI, and theta_rad. Calculate the magnitude of the scattering vector, q (Eq. 11):

$$q = \frac{4\pi n}{\lambda} sin\left(\frac{\theta}{2}\right)$$

The Excel code for this equation using the defined names is "4*PI()*RI*SIN(theta_rad/2)/W."

5. Create a set of column names tau (s), tau (usec), g2-1, g2, and g2 calc in columns A, B, C, D, and E in row 37 (Figure 16).

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

Please cite this article as: Parker, E. T. and Lollar, P. (2021). Measurement of the Translational Diffusion Coefficient and Hydrodynamic Radius of Proteins by Dynamic Light Scattering. *Bio-protocol* 11(20): e4195. DOI: 10.21769/BioProtoc.4195.

	A	B	C	D	E
37	tau (s)	tau (usec)	g2-1	g2	g2 calc
38	5.00E-07	0.50	7.34E-01	1.7339	1.7267
39	1.00E-06	1.00	7.01E-01	1.7009	1.6968
40	1.50E-06	1.50	6.69E-01	1.6693	1.6681
41	2.00E-06	2.00	6.36E-01	1.6360	1.6407
42	2.50E-06	2.50	6.11E-01	1.6114	1.6145
43	3.00E-06	3.00	5.86E-01	1.5855	1.5893
44	3.50E-06	3.50	5.62E-01	1.5616	1.5652
45	4.00E-06	4.00	5.38E-01	1.5376	1.5422
46	4.50E-06	4.50	5.18E-01	1.5178	1.5201
47	5.50E-06	5.50	4.79E-01	1.4787	1.4787
48	6.50E-06	6.50	4.43E-01	1.4426	1.4407
49	7.50E-06	7.50	4.06E-01	1.4065	1.4058
50	8.50E-06	8.50	3.76E-01	1.3755	1.3738
51	9.50E-06	9.50	3.47E-01	1.3468	1.3444
52	1.05E-05	10.50	3.18E-01	1.3181	1.3174
53	1.15E-05	11.50	2.94E-01	1.2940	1.2926
54	1.25E-05	12.50	2.71E-01	1.2714	1.2698
55	1.45E-05	14.50	2.30E-01	1.2301	1.2295
56	1.65E-05	16.50	1.97E-01	1.1974	1.1955
57	1.85E-05	18.50	1.66E-01	1.1664	1.1667
58	2.05E-05	20.50	1.43E-01	1.1428	1.1423
59	2.25E-05	22.50	1.22E-01	1.1220	1.1216
60	2.45E-05	24.50	1.04E-01	1.1043	1.1040
61	2.65E-05	26.50	8.81E-02	1.0881	1.0890
62	2.85E-05	28.50	7.56E-02	1.0756	1.0763
63	3.25E-05	32.50	5.57E-02	1.0557	1.0562
64	3.65E-05	36.50	4.07E-02	1.0407	1.0416
65	4.05E-05	40.50	2.97E-02	1.0297	1.0309
66	4.45E-05	44.50	2.15E-02	1.0215	1.0231
67	4.85E-05	48.50	1.68E-02	1.0168	1.0173
68	5.25E-05	52.50	1.18E-02	1.0118	1.0131
69	5.65E-05	56.50	9.30E-03	1.0093	1.0099
70	6.05E-05	60.50	7.06E-03	1.0071	1.0075
71	6.85E-05	68.50	4.10E-03	1.0041	1.0045
72	7.65E-05	76.50	1.83E-03	1.0018	1.0028
73	8.45E-05	84.50	1.88E-03	1.0019	1.0018
74	9.25E-05	92.50	1.34E-03	1.0013	1.0013
75	1.01E-04	100.50	1.51E-03	1.0015	1.0009
76	1.09E-04	108.50	1.00E-03	1.0010	1.0008
77	1.17E-04	116.50	1.24E-03	1.0012	1.0007
78	1.25E-04	124.50	6.21E-04	1.0006	1.0006
79	1.41E-04	140.50	1.34E-03	1.0013	1.0005
80	1.57E-04	156.50	1.62E-03	1.0016	1.0005
81	1.73E-04	172.50	8.90E-04	1.0009	1.0005
82	1.89E-04	188.50	1.25E-03	1.0013	1.0005
83	2.05E-04	204.50	1.18E-03	1.0012	1.0005
84	2.21E-04	220.50	8.97E-04	1.0009	1.0005
85	2.37E-04	236.50	6.30E-04	1.0006	1.0005
86	2.53E-04	252.50	2.59E-04	1.0003	1.0005
87	2.85E-04	284.50	7.18E-04	1.0007	1.0005
88	3.17E-04	316.50	9.59E-04	1.0010	1.0005
89	3.49E-04	348.50	5.63E-04	1.0006	1.0005
90	3.81E-04	380.50	4.30E-04	1.0004	1.0005
91	4.13E-04	412.50	3.59E-04	1.0004	1.0005
92	4.45E-04	444.50	6.21E-04	1.0006	1.0005
93	4.77E-04	476.50	5.95E-04	1.0006	1.0005
94	5.09E-04	508.50	6.32E-04	1.0006	1.0005
95	5.73E-04	572.50	4.02E-04	1.0004	1.0005
96	6.37E-04	636.50	3.94E-04	1.0004	1.0005
97	7.01E-04	700.50	1.80E-04	1.0002	1.0005
98	7.65E-04	764.50	6.90E-04	1.0007	1.0005
99	8.29E-04	828.50	2.80E-04	1.0003	1.0005
100	8.93E-04	892.50	4.67E-04	1.0005	1.0005
101	9.57E-04	956.50	8.75E-04	1.0009	1.0005

Figure 16. DLS data and Excel Solver cumulant fits. Column A, imported values of τ; column B, conversion of τ to μs; column C, imported values of $g_2(\tau)-1$; column D, $g_2(\tau)$; column E, the fitted values of $g_2(\tau)$.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

6. Import a .dlsdat file containing $g_2(\tau)-1$ vs. τ data by selecting Data > From Text/CSV, then Load. Data will load as a single column in a new sheet.
7. Select the column of loaded data, then select Data > Text to Columns, Delimited, Space. Columns of τ and $g_2(\tau)-1$ will be created. Delete the three header rows.
8. Cut and paste τ and $g_2(\tau)-1$ data into columns A and C of the Master Sheet, respectively, starting at row 38 (Figure 16). Use the same number of data pairs as in the SEDFIT cumulant fit.
9. Convert τ from seconds to microseconds (to avoid a numerical precision problem) by multiplying the values in column A by 1E6 and place the values in col B (Figure 16).
10. Calculate $g_2(\tau)$ by adding 1 to the values in column C and place the values in column D (Figure 16).
11. Make a table of Parameter starting estimates starting in row 26 (Figure 15C). Enter a starting estimate of the diffusion coefficient in Ficks in row 26. In this example, a starting estimate of 20 F has been entered. Convert to units of cm^2 s^{-1} by multiplying by 1E-07. Calculate an estimated value of $\bar{\Gamma}$, and label it Gamma est using the first moment form of Eq. 10:

$$\bar{\Gamma} = D_z q^2$$

12. Enter a starting estimate for the polydispersity index (PDI). In this example, a starting value of 0.4 has been entered. Calculate an estimated value of the second moment, μ_2, and label it Mu2 est using Eq. 17:

$$\mu_2 = (PDI)\bar{\Gamma}^2$$

13. Make a table Parameter fits starting in row 26 and a table Diffusion coefficient and PDI starting in row 32 (Figure 15D).
14. Copy the values of Gamma_est and Mu2_est into the Parameter fits table. Enter 1 for starting estimates for both B and β. Define names for the values in the Parameter Fits table: B_, beta, Gamma, Mu_2. Figure 15C shows the Parameter Fits table before Solver has run.
15. Calculate $g_2(\tau)$ from the starting estimates using the first two moments of the Frisken equation (Eq. 15). The Excel code for this calculation is:
```
B_+beta*EXP(-2*Gamma*B38)*(1+(Mu_2*B38^2/2))^2
```
where B38 represents the value of τ in the first row of data.
16. Make a table entitled Fit starting in row 26 with rows SSR and Data pairs. Figure 15E shows the values of SSR before Solver has run.
17. Calculate the summed squared residuals (SSR) from the g2 and g2 calc columns using the SUMXMY2 function in Excel. In this example, the g2 and g2 calc values are in rows D38:D101 and E38:E101, respectively (Figure 16), and the Excel code is:
```
SUMXMY2(D38:D101,E38:E101)
```
Determine the number of data pairs using the Excel COUNT function or by entering it manually.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

18. Minimize SSR by fitting the parameters B, beta, Gamma, and Mu2 that are in cells E27:E30 using the GRG Nonlinear algorithm in Solver by selecting Data > Solver, Set Objective SSR, To Min, By Changing Variable Cells $E$27:$E$30 (Figure 17). In this example, the SSR has decreased from 1.94E-04 after Solver has run (Figure 15F). The fitted parameters and estimates of the diffusion coefficient and PDI appear in the Parameter fits and Diffusion coefficient and PDI tables (Figure 15D). The rmsd has decreased from 0.098 to 1.74E-03 (Figure 15D).
19. Cut and paste SEDFIT cumulant estimates of B, beta (called "I tot" in SEDFIT), rmsd, D, and PDI for comparison (for B, add 1 to the SEDFIT fitted baseline). The values are in close agreement (Figure 15D).

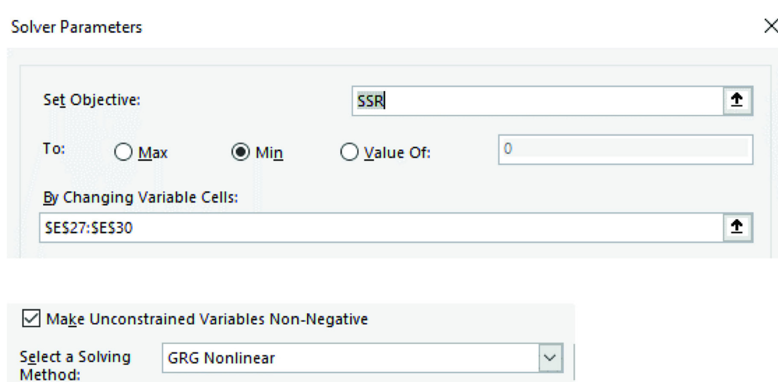


Figure 17. Excel solver window

D. Statistical analysis

1. Point estimates and confidence intervals by analysis of variance

The goal of this protocol is to assign a point estimate and a 95% confidence interval for the diffusion coefficient of BSA, corrected to the standard condition of water at 20°C, $D_{20,w}$. A trial of four measurements (replicates) is made on a sample. The process was repeated with fresh aliquots from the same tube for a total of eight trials. Additionally, another sample of BSA is thawed and also subjected to eight trials, each with four measurements, to assess possible variation due to freezing and thawing.

Possible sources of variation in the experimental results arise from: 1) the variation within a trial of four measurements; 2) variation between trials due to removing the cuvette and repeated aliquoting; and 3) variation due to freezing, thawing, and sample preparation. The variation due to replication and repeated aliquoting is assessed by one-way random effects analysis of variance (ANOVA), which produces a point estimate and 95% confidence interval for the diffusion coefficient. Although there are many sources that describe one-way random effects ANOVA, few provide derivations of the confidence interval. One example is Brownlee (1965), which provides the basis for the description that follows.

In one-way ANOVA, there are a independent groups of samples (here, aliquots), each of which has b measurements (replicates). The ANOVA model is

$$X_{j,k} = \mu + \alpha_j + \varepsilon_{j,k}$$

where $X_{j,k}$ is the measurement associated with the replicate k of group j, μ is the grand mean of the measurements of all the groups, α_j is the mean of group j, and $\varepsilon_{j,k}$ is the error associated with measurement within the group j, measurement k. It is assumed that α_j are normally distributed with zero mean and variance σ_B^2 and that $\varepsilon_{j,k}$ is normally distributed with zero mean and variance, σ_W^2. The variance of X is given by

$$\sigma_X^2 = \sigma_B^2 + \sigma_W^2$$

The terms on the right are called the components of variance. In this example, we are interested in the mean, μ(the point estimate) and confidence limits of the diffusion coefficient of BSA. The hypothesis tested in the random effects model is

$$H_0: \sigma_B^2 = 0$$

i.e., all variation is due to replication, and there is no variation because of group effects. It is equivalent to the hypothesis

$$H_0: \alpha_j = 0$$

Results are summarized using an ANOVA table (Figure 18):

One-Way ANOVA table

Source of variation	Degrees of freedom	Sum of squares	Mean square	F-statistic
Between groups	$a-1$	$SS_B = b\sum_{i=1}^{a}(\bar{X}_{j.} - \bar{X})^2$	$MS_B = \frac{SS_B}{a-1}$	$\frac{MS_B}{MS_W}$
Within groups	$b-1$	$SS_W = \sum_{j=1}^{a}\sum_{k=1}^{b}(X_{j,k} - \bar{X}_{j.})^2$	$MS_W = \frac{SS_W}{a(b-1)}$	
Total	$ab-1$	$SS_T = SS_B + SS_W$		

Figure 18. ANOVA table

In the table, $\bar{X}$, is the sample grand mean and $\bar{X}_{j.}$ is the mean of the replicates of group j. The hypotheses, H_0, is tested using F statistics where

$$F = \frac{MS_B}{MS_W} \quad (19)$$

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

with $a-1$, $b-1$ degrees of freedom. The estimator of the variance within groups is

$$\hat{\sigma}_W^2 = MS_W \tag{20}$$

The estimator of σ_B^2 is

$$\hat{\sigma}_B^2 = \frac{MS_B - MS_W}{b} \tag{21}$$

The estimator of the variance of the mean, $\bar{X}$, is

$$\hat{\sigma}_{\bar{X}}^2 = \frac{1}{n}(b\hat{\sigma}_B^2 + \hat{\sigma}_W^2) \tag{22}$$

Equivalently,

$$\hat{\sigma}_{\bar{X}}^2 = \frac{MS_B}{n} \tag{23}$$

The 95% confidence interval for μ is obtained from Student's *t*-distribution with $a-1$ degrees of freedom.

$$\bar{X} - t_{0.975}\hat{\sigma}_{\bar{X}} < \mu < \bar{X} + t_{0.975}\hat{\sigma}_{\bar{X}} \tag{24}$$

where $\bar{X}$ is mean of all measurements, $-t_{0.975}$ and $t_{0.975}$ are the values of *t*-distribution for which 2.5% of the area lies within each tail of the distribution, and $\hat{\sigma}_{\bar{X}}$ is the estimator of the standard deviation of the mean (also called the standard error of the mean). One-way ANOVA can be performed using Excel. This example assumes the Data Analysis Add-In has been installed.

2. SEDFIT cumulant analysis statistics
   a. Create a new worksheet in Excel and make eight columns for the aliquots and four rows for the measurements (replicates) (Figure 19).
   b. Select all 32 measurement values from Data analysis B.2 and calculate the sample grand mean using the Excel AVERAGE function. This is the estimate of D_z.
   c. Use the Define Name function to assign names to the number of groups and replicates in Figure 19.

	D (F)							
	Thaw 1							
Replicate	1	2	3	4	5	6	7	8
1	6.0324	6.0298	6.0461	5.9563	5.778	5.9920	5.9322	6.0597
2	6.0468	6.0723	6.0156	5.9165	5.857	6.0946	5.9673	6.0601
3	6.0417	5.9965	5.9857	5.8700	5.888	6.0877	5.9248	6.0670
4	5.9834	6.0225	5.8589	5.8660	5.917	6.0544	5.8229	6.0797
Mean	6.0261	6.0303	5.9766	5.9022	5.8599	6.0572	5.9118	6.0666
	GrandMean	5.9788						
	Groups (a)	8						
	Replicates (b)	4						
	n	32						

Figure 19. SEDFIT cumulant fits

d. Select Data > Data Analysis > Anova: Single factor.

e. In the window that appears, select the cells that contain the D values, $B$17:I20$ (Figure 20). The trials (aliquots) are grouped by columns.

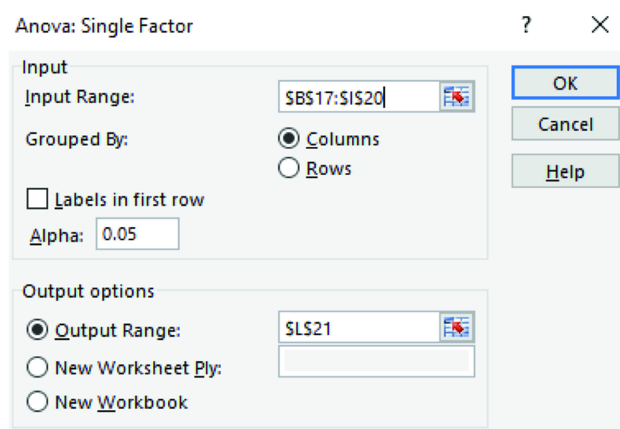


Figure 20. Excel ANOVA window

f. Select Alpha of 0.05. This is the significance level for the F test.

g. Select $L$21 for the cell where the output range begins and then select OK. Excel produces an ANOVA table (Figure 21A). The F test reveals what is evident from the inspection of Figure 19, namely, that there is more variation between trials than between the measurements (replicates). Because the aliquots are taken from the same bullet tube, the variation is not due to pipetting error. Possibilities include small bubbles in the sample, reflection off the meniscus, or variation in the position of the cuvette during the process of removing it, adding a new aliquot, and replacing it.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

A

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.1730	7	0.0247	9.80	9.7E-06	2.42
Within Groups	0.0605	24	0.00252			
Total	0.2335	31				

B

	D (F)	D20w (F)
Mean	5.98	6.06
95% confidence limits:	5.91	5.99
	6.04	6.13

Figure 21. SEDFIT cumulants ANOVA. (A) ANOVA table. (B) Point estimate and confidence limits for z-average D and z-average D_{20w}.

h. Calculate the confidence interval

i. Calculate the variance of the mean by dividing the mean square between groups MS_B (Figure 21) by n, the total number of measurements.

ii. Calculate the standard deviation of the mean by taking the square root of this value.

iii. Calculate the value of Student's t density function for a 95% confidence interval using the Excel function T.INV(probability,deg_freedom). Here, the probability is 0.975, and the degree of freedom is a−1. Since a equals 8, there are 7 degrees of freedom and T.INV(0.975, 7) = 2.365.

iv. Calculate the upper and lower limits of the confidence interval using Eq. 24.

v. Calculate $D_{20,w}$ using Eq. 5:

$$D_{20,w} = D\frac{T_{20}}{T}\frac{\eta}{\eta_{20,w}}$$

The point z-average estimates of D and $D_{20,w}$ are shown in Figure 21B.

3. SEDFIT continuous I(D) distribution analysis statistics

a. Create a new worksheet in Excel and make eight columns for the aliquots and four rows for the measurements as Data analysis D.2 (Figure 22).

	D (F)							
	Thaw 1							
	Aliquot							
Replicate	1	2	3	4	5	6	7	8
1	6.1880	6.0789	6.0602	6.0676	5.7600	6.2326	5.9428	6.2313
2	6.0044	6.1899	6.0313	5.9452	5.9197	6.1769	5.9848	6.2602
3	6.0923	6.1064	6.1780	6.0519	5.8056	6.1767	5.9266	6.2023
4	6.0868	6.1529	5.8644	5.8889	5.9799	6.1927	5.7622	6.2422
Mean	6.0929	6.1320	6.0335	5.9884	5.8663	6.1947	5.9041	6.2340
	GrandMean	6.06						
	Groups (a)	8						
	Replicates (b)	4						
	n	32						

Figure 22. SEDFIT continuous I(D) fits

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

b. Insert the measurement values from the SEDFIT continuous I(D) distribution analysis from Data analysis B.3 and calculate the grand mean. This is the signal-average diffusion coefficient.

c. Construct an ANOVA table and calculate the calculate confidence limits as described in Data analysis D.2 (Figure 23).

A

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.4888	7	0.0698	10.5	5.5E-06	2.42
Within Groups	0.1597	24	0.00665			
Total	0.6485	31				

B

	D (F)	D20w (F)
Mean	6.06	6.14
95% confidence limits:	5.95	6.02
	6.17	6.25

Figure 23. SEDFIT continuous I(D) ANOVA. (A) ANOVA table; (B) Point estimate and confidence limits for z-average D and z-average D_{20w}.

4. Zetasizer cumulant analysis statistics

a. Create a new worksheet in Excel and make eight columns for the aliquots and four rows for the measurements (replicates) as in Data analysis D.2 (Figure 24).

	Rh (nm)							
	Thaw 1							
	Aliquot							
Replicate	1	2	3	4	5	6	7	8
1	3.556	3.541	3.546	3.594	3.702	3.661	3.596	3.622
2	3.529	3.522	3.552	3.609	3.659	3.671	3.605	3.635
3	3.545	3.562	3.577	3.638	3.618	3.649	3.570	3.650
4	3.605	3.556	3.659	3.663	3.642	3.707	3.674	3.638
	GrandMean	3.61						
	Groups	8						
	Replicates	4						
	n	32						

Figure 24. Zetasizer cumulants Rh fits

b. Insert the measurement values of R_h from the Zetasizer cumulant analysis Data analysis A and calculate the grand mean.

c. Construct an ANOVA table and calculate the confidence limits for R_h as described Data analysis D.2. Convert R_h values to z-average diffusion coefficients, D and $D_{20,w}$, using the Stokes-Einstein equation (Eq. 4). The results are shown in Figure 25.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

Please cite this article as: Parker, E. T. and Lollar, P. (2021). Measurement of the Translational Diffusion Coefficient and Hydrodynamic Radius of Proteins by Dynamic Light Scattering. *Bio-protocol* 11(20): e4195. DOI: 10.21769/BioProtoc.4195.

A

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.05740	7	0.00820	7.27	1.03E-04	2.42
Within Groups	0.02709	24	0.00113			
Total	0.08449	31				

B

	Rh (nm)	D (F)	D20,w (F)
Mean	3.61	5.86	5.94
95% Confidence limits:	3.65	5.80	5.87
	3.57	5.92	6.00

Figure 25. Zetasizer cumulants ANOVA. (A) ANOVA table. (B) Point estimate and confidence limits for z-average D and z-average D_{20w}.

5. Zetasizer "Protein Analysis" distribution statistics
   a. Create a new worksheet in Excel and make eight columns for the aliquots and four rows for the measurements as in Data analysis D.2 (Figure 26).

	Rh (nm)							
	Thaw 1							
Replicate	1	2	3	4	5	6	7	8
1	3.706	3.763	3.796	3.778	3.905	3.661	3.817	3.662
2	3.789	3.725	3.772	3.839	3.925	3.671	3.894	3.635
3	3.744	3.707	3.744	3.848	3.897	3.649	3.874	3.650
4	3.703	3.638	3.843	3.915	3.83	3.707	3.914	3.638
	GrandMean	3.77						
	Groups	8						
	Replicates	4						
	n	32						

Figure 26. Zetasizer Protein analysis Rh fits

   b. Insert the measurement values of R_h from Zetasizer Protein Analysis Data analysis A and calculate the grand mean.
   c. Construct an ANOVA table and calculate the confidence limits for R_h as described in Data analysis D.2.Convert R_h values to signal-average diffusion coefficients, D and $D_{20,w}$, using the Stokes-Einstein equation (Eq. 4). The results are shown in Figure 27.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

A

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.2444	7	0.0349	20.7	1.00E-08	2.42
Within Groups	0.0405	24	0.00169			
Total	0.2848	31				

B

	Rh (nm)	D (F)	D20,w (F)
Mean	3.77	5.61	5.69
95% Confidence limits:	3.85	5.50	5.57
	3.69	5.73	5.81

Figure 27. Zetasizer Protein Analysis ANOVA. (A) ANOVA table. (B) Point estimate and confidence limits for z-average D and z-average D_{20w}.

6. Testing possible variation due to freeze-thawing followed by sample preparation
   a. Thaw a new sample and repeat the sample preparation steps described in Procedure B.3
   b. Perform DLS measurements and analysis on eight aliquots of the sample, four measurements per aliquot as described in Procedure B.4 and Data analysis.
   c. Compare the results to the first thawed sample (Figure 28A). The results show overlapping confidence intervals for all four methods of analysis, indicating there is no systematic error due to the freeze-thaw sample preparation process.
7. Summary of fitting methods and literature review
   The confidence intervals in Figure 28A are a measure of the precision but not the accuracy of the measurements and analysis. The goal of the protocol is to have a method of measurement and analysis that is free from significant systematic errors and produces a confidence interval that closely brackets a true value of the diffusion coefficient at infinite dilution at 20°C in water ($D^0_{20,w}$). It is important to address what is meant by $D^0_{20,w}$. Although diffusion coefficients are not measured at infinite dilution, the values for globular proteins typically show negligible variation at submilligram per mL concentrations, as is the case in this protocol. This specifically has been shown for BSA (Creeth, 1952). As noted in Background, the notion of a single diffusion coefficient is restricted to a two-component system, namely the macromolecule and solvent water. Because proteins are polyelectrolytes and carry net charge that varies with pH, it would seem fanciful to attempt to extrapolate from a buffered salt solution to water. However, the screening effects in solutions close to physiologic anionic strength tend to eliminate solvent-specific effects on diffusion, which again is the case with BSA (Creeth, 1952).

   Assuming it is realistic to think of a true $D^0_{20,w}$, the fundamental problem arises that it is never possible to know that all sources of systematic error have been eliminated. Consequently, measurement accuracy often is defined in terms of a reference value (Mandel, 1964). For example, the speed of light *in vacuo* is not known exactly from experiment but rather is defined as 299,792,458 m/s. Other measurements are made relative to this standard. In the absence of a reference standard, as with $D^0_{20,w}$ values, the evaluation of accuracy becomes an educated

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

guess (Mandel, 1964). There are two ways to attempt to make a reasonable estimate. The first is to attempt to identify and estimate the size of all possible systemic errors and produce a range within which the true value is thought to lie. The second is to compare the results of independent measuring processes. For the $D^0_{20,w}$ value of BSA, three independent methods are diffusiometry, DLS, and application of the Svedberg equation. In diffusiometry, the measurements by Creeth reported in 1958 (Creeth, 1958) represented an evolution of improvements in the method and can be considered the most accurate measurements that are available (Figure 28B).

The Svedberg equation (Eq. 18) can be rearranged to yield the diffusion coefficient as a function of the sedimentation coefficient, gas constant, temperature, molar mass, partial specific volume, and solvent density:

$$D = \frac{sRT}{M(1 - \bar{v}\rho)}$$

Of these, the experimental determination of the sedimentation coefficient extrapolated to water at 20°C, $s^0_{20,w}$, and the uncertainty the partial specific volume, $\bar{v}$, are the most significant sources of error. Without doing a formal propagation of errors analysis, using a measured value of $s^0_{20,w}$ of 4.44 S (Parker and Lollar, 2021), $\bar{v}$ = 0.733 ml/g using the BSA amino acid composition and the partial specific volume calculator in SEDFIT, M = 66,365 g/mol, and $\rho_{20,w}$ = 0.9982 g/ml, yields a diffusion coefficient of 6.05 F (Figure 28B), which is within the Creeth's diffusiometry confidence interval.

Of the data analysis methods in this protocol, the continuous I(D) method and Protein Analysis method yield confidence intervals that are less precise than the cumulant methods. The SEDFIT cumulant confidence intervals cover the estimates from diffusiometry and the Svedberg equation better than the Zetasizer cumulant analysis, which produces lower values. The SEDFIT cumulant analysis also is preferable because the equation it is fitting is Frisken equation, as verified independently using Excel Solver. Fitting data to the Frisken equation represents an improvement to earlier methods because it uses nonlinear least-squares regression instead of transforming the $g_2(\tau)$ logarithmically and performing a linear regression. In contrast, the Zetasizer cumulant method is not described on the software documentation. It truncates the data set at low values of $g_2(\tau)-1$, possible due to a logarithmic transform that must avoid fitting values that vary around zero.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

A

	D20,w (F)		
	Mean	95% Confidence limits	
SEDFIT cumulants			
Thaw 1	6.06	5.99	6.13
Thaw 2	6.04	6.00	6.09
SEDFIT continuous I(D)			
Thaw 1	6.14	6.02	6.25
Thaw 2	6.10	6.02	6.18
Zetasizer cumulants			
Thaw 1	5.88	5.82	5.94
Thaw 2	5.95	5.90	6.01
Zetasizer Protein Analysis			
Thaw 1	5.69	5.57	5.81
Thaw 2	5.65	5.57	5.73

B

	D20,w (F)			
	Pt. est.	Reported range		Reference
Akeley & Gosting		5.89	6.00	a
Wagner & Scheraga		5.86	5.90	b
Creeth		5.95	6.10	c
Harvey et al.		6.06	6.20	d
Oh & Johnson	5.92	5.88	5.96	e
Gaigalas et al.	6.09	6.05	6.13	f
Svedberg equation	6.05			g

Figure 28. $D^0_{20,w}$ of BSA. (A) Point estimates (Pt. est) and 95% confidence limits for the z-average $D^0_{20,w}$ this protocol. (B) Literature the $D^0_{20,w}$ values.

a: Ref. (Akeley and Gosting, 1953), diffusion cell, Gouy interferometry, 0.01 M K acetate, 0.15 M KCl, pH 4.59, corrected for impurity

b: Ref. (Wagner and Scheraga, 1956), diffusion cell, Gouy interferometry, 0.5 M KCl, pH 5.14; corrected for impurity

c: Ref. (Creeth, 1958), diffusion cell, Rayleigh interferometry, 0.01 M K acetate, 0.15 M KCl, pH 4.59, corrected for impurity

d: Ref. (Harvey *et al.*, 1979), DLS, cumulant analysis, SEC fractionated BSA; 0.1 M NaCl, 0.07 M phosphate, pH 7.0

e: Ref. (Oh and Johnson, 1981), DLS, cumulant analysis, SEC fractionated BSA, 0.1 M NaCl, 0.02 M sodium acetate, pH 4.7

f: Ref. (Gaigalas *et al.*, 1992), DLS, cumulant analysis, NIST standard BSA, various buffers

g: Svedberg equation (Eq. 18, see text)

The ranges reported refer to undefined confidence intervals or error estimates.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

Recipes

1. 10× PBS (1 L)
   1.54 M NaCl
   56 mM Na_2HPO_4
   10.6 mM KH_2PO_4
   a. Weigh out:
      For 1.54 M NaCl (MW 58.44) 90. 0g
      For 56 mM Na_2HPO_4 (MW=141.96) 7.95 g
      For 10.6 mM KH_2PO_4 (MW 136.09) 1.44 g
   b. Add the solids to ~980 ml H_2O in a 1 L beaker
   c. After dissolution, adjust volume to 1 L in a graduated cylinder
   d. Filter using a Corning 500 ml, 0.22 µm filter apparatus
   e. Store at 4°C
2. PBS (1 L)
   0.154 M NaCl
   5.6 mM Na_2HPO_4
   1.06 mM KH_2PO_4, pH 7.4
   a. The densities of 10× PBS and water at 20°C calculated using SEDNTERP are 1.0684 g/ml and 0.9982 g/ml, respectively.
   b. Add 100 ml 10× PBS gravimetrically to a 1 L Pyrex bottle (mass = density × volume = 106.8 g).
   c. Dilute with 900 ml H_2O gravimetrically (mass = density × volume = 898.4 g).
   d. The pH should be ~7.15. pH to 7.40 with 5 N NaOH using a pH meter.
   e. Filter as described above.
   f. Store at 4°C.

Acknowledgments

This work was supported by the National Institutes of Health/National Heart, Lung and Blood Institute (award number U54HL141981 to PL). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This protocol describes the validation of the use of the Zetasizer Nano S particle size analyzer to measure the protein translational diffusion coefficients described in Parker and Lollar (2021).

Competing interests

We have no competing interests to declare.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

Ethics

There are no human subjects or animal studies described in this protocol.

References

1. Akeley, D. F. and Gosting, L. J. (1953). Studies of the diffusion of mixed solutes with the Gouy diffusiometer. *J Am Chem Soc* 75(22): 5685-5696.
2. Annunziata, O., Buzatu, D. and Albright, J. G. (2005). Protein diffusion coefficients determined by macroscopic-gradient Rayleigh interferometry and dynamic light scattering. *Langmuir* 21(26): 12085-12089.
3. Berg, H. C. (1983). Random Walks in Biology. Princeton, Princeton University Press.
4. Berne, B. J. and Pecora, R. (1976). Dynamic Light Scattering with Applications to Chemistry, Biology, and Physics. Mineola, Dover Publications.
5. Bhattacharjee, S. (2016). DLS and zeta potential - What they are and what they are not? *J Control Release* 235: 337-351.
6. Brautigam, C. A. (2019). Applications and complementarity of analytical ultracentrifugation and light scattering. Chapter 11. In: *Biomolecular and Bioanalytical Techniques: Theory, Methodology and Applications* 255-278.
7. Brownlee, K. A. (1965). Statistical Theory and Methodology in Science and Engineering. In: *A Wiley Publication in Applied Statistics*. New York, John Wiley & Sons.
8. Cantor, C. R. and Schimmel, P. R. (1980). Biophysical chemistry: Part II Techniques for the study of biological structure and function. Oxford, W. H. Freeman and Company 591-642.
9. Creeth, J. M. (1952). The use of the Guoy diffusiometer with dilute protein solutions - an assessment of the accuracy of the method. *Biochem J* 51(1): 10-17.
10. Creeth, J. M. (1958). Studies of free diffusion in liquids with the Rayleigh method. III. The analysis of known mixtures and some preliminary investigations with proteins. *J Phys Chem* 62(1): 66-74.
11. Einstein, A. (1926). Investigations on the Theory of Brownian Movement. (English translation first). New York, Methuen Co. & Dover Publications. pp124.
12. Finsy, R. (1994). Particle sizing by quasi-elastic light scattering. *Adv Colloid Interface Sci* 52: 79-143.
13. Frisken, B. J. (2001). Revisiting the method of cumulants for the analysis of dynamic light-scattering data. *Applied Optics* 40(24): 4087-4091.
14. Gaigalas, A. K., Hubbard, J. B., McCurley, M. and Woo, S. (1992). Diffusion of bovine serum albumin in aqueous solutions. *J Phy Chem* 96(5): 2355-2359.
15. Harvey, J. D., Geddes, R. and Wills, P. R. (1979). Conformational studies of BSA using laser light scattering. *Biopolymers* 18(9): 2249-2260.
16. Hurton, T., Wright, A., Deubler, G. and Bashir, B. (2012). SEDNTERP. http://bitc.unh.edu.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.

17. Koppel, D. E. (1972). Analysis of macromolecular polydispersity in intensity correlation spectroscopy - method of cumulants. *J Chem Phys* 57(11): 4814-4820.
18. Mandel, J. (1964). The Statistical Analysis of Experimental Data. New York, John Wiley and Sons.
19. Miller, D. G. (2014). The history of interferometry for measuring diffusion coefficients. *J Solution Chem* 43(1): 6-25.
20. Ogendal, L. H. (2017). Light Scattering Demystified: Theory and Practice. Danmark, University of Copenhagen.
21. Oh, Y. S. and Johnson, C. S. (1981). The wave vector dependence of diffusion-coefficients in photon-correlation spectroscopy of protein solutions. *J Chem Phys* 74(5): 2717-2720.
22. Parker, E. T. and Lollar, P. (2021). Conformation of the von Willebrand factor/factor VIII complex in quasi-static flow. *J Biol Chem* 296: 100420.
23. Pusey, P. N., Koppel, D. E., Schaefer, D. W., Camerini.Rd and Koenig, S. H. (1974). Intensity fluctuation spectroscopy of laser light scattered by solutions of spherical viruses - R17, Qbeta, BSV, PM2, and T7. 1. Light scattering technique. *Biochemistry* 13(5): 952-960.
24. Stetefeld, J., McKenna, S. A. and Patel, T. R. (2016). Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical Reviews* 8: 409-427.
25. Tanford, C. (1961). Physical Chemistry of Macromolecules. New York, John Wiley & Sons.
26. Wagner, M. L. and Scheraga, H. A. (1956). Gouy diffusion studies of bovine serum albumin. *J Phys Chem* 60(7): 1066-1076.

Copyright © 2021 The Authors; exclusive licensee Bio-protocol LLC.