6

How to Use Dynamic Light Scattering to Improve the Likelihood of Growing Macromolecular Crystals

Gloria E. O. Borgstahl

Summary

Dynamic light scattering (DLS) has become one of the most useful diagnostic tools for crystallization. The main purpose of using DLS in crystal screening is to help the investigator understand the size distribution, stability, and aggregation state of macromolecules in solution. It can also be used to understand how experimental variables influence aggregation. With commercially available instruments, DLS is easy to perform, and most of the sample is recoverable. Most usefully, the homogeneity or monodispersity of a sample, as measured by DLS, can be predictive of crystallizability.

Key Words: Solubility; aggregation; hydrodynamic radius; monodispersity; polydispersity; dynamic light scattering; DLS.

1. Introduction

Along the road to atomic structure determination by X-ray crystallography, a major challenge is the growth of high-quality crystals of the macromolecule of interest. The level of interest in the macromolecule seems to correlate well with the level of difficulty in its crystallization. It is helpful to optimize the composition of the protein sample and the experimental conditions, such as buffer components and temperature, to increase the likelihood of crystallization. To optimize these parameters it is necessary to analyze the aggregation state and stability of a macromolecular sample. There are many methods that can be used to measure size or aggregation state, including sedimentation equilibrium, size exclusion chromatography, native gel electrophoresis, and light scattering. Of these methods, dynamic light scattering (DLS), otherwise known as quasielastic light scattering, is the easiest to implement, the quickest to perform, and the least destructive to the sample (**Fig. 1**). For complex cases a combination of these

From: *Methods in Molecular Biology, vol. 363: Macromolecular Crystallography Protocols: Volume 1: Preparation and Crystallization of Macromolecules* Edited by: S. Doublié © Humana Press Inc., Totowa, NJ

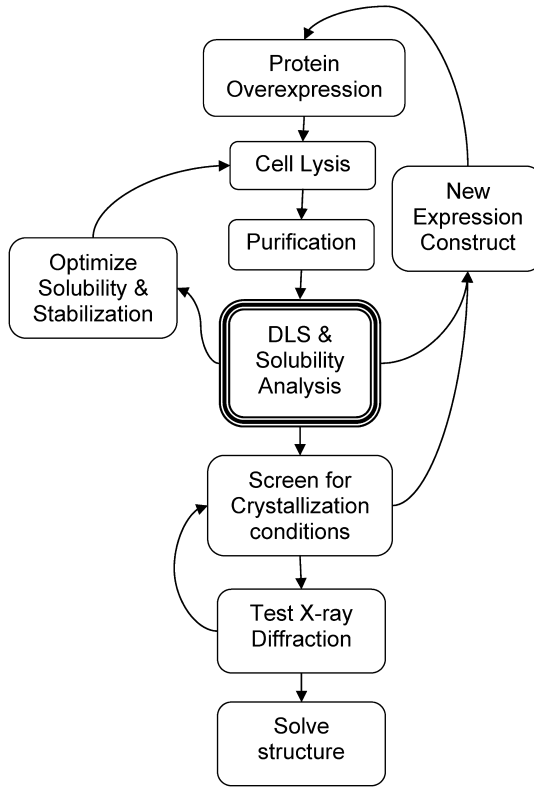


Fig. 1. Flow diagram of methods involved in growing protein crystals. By placing dynamic light scattering and solubility analysis in the center of the process the chances of growing crystals are optimized

methods may be needed to interpret the data. Samples that are monodisperse in solution, as measured by DLS, are much more likely to crystallize ***(1,2)***. Along with this chapter, there are several other useful chapters written on the use of DLS analysis in crystallization ***(3–5)***.

The DLS instrument is easy to use and detailed knowledge of the underlying physics of molecular sizing is usually not needed. Therefore, only a brief explanation is included here. A microcuvet of protein solution is illuminated by laser light (**Fig. 2**). The molecules in solution are undergoing Brownian motion and cause fluctuations in the scattered light intensity. This change in light intensity is measured by a detector placed at a 90° angle to the incident laser light. The translational diffusion coefficient D_T is derived from these data using an autocorrelation function. In general, small particles diffuse "faster" than large particles. A hydrodynamic radius (R_H) of the molecules in solution can be calculated from D_T. In general, particles must differ in R_H by 50% or more to be well

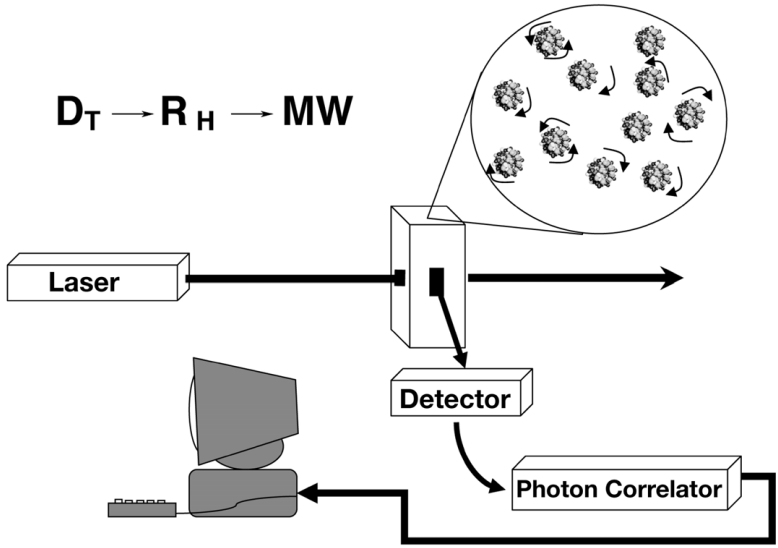


Fig. 2. Simplified schematic of the dynamic light scattering experiment.

separated by DLS (**Fig. 3**). Molecular weight (MW) can also be estimated, if the shape of the molecule is assumed, e.g., spherical or elongated. DLS as an estimator of MW is not recommended and must be used with caution. The shape definitions used to estimate MW may not accurately represent the particles in solution, and for polydisperse samples the R_H and MW will be based on a weighted average of more than one species. Multiple-angle static light scattering downstream from size-exclusion chromatography is the recommended light-scattering method to measure absolute MW of molecules in solution. In this chapter, DLS is primarily used to assess the aggregation state of a sample and to measure polydispersity, which is predictive of crystallizability.

2. Materials

1. Protein Solutions DynaPro MS/X instrument with temperature control (0–60°C) installed and correctly configured.
2. Dynamics software v6 installed on a compatible PC.
3. Protein Solutions 12-µL quartz cuvets.
4. Protein Solutions microfiltration system.
5. Syringe tip (0.2-µm filters).
6. 20–30-µL Protein sample.
7. Water, 1% Triton X-100, and a range of appropriate buffer and salt solutions, all 0.2-µm filtered.
8. Compressed air, either house air or in a can.
9. Lens paper.
10. Small ultrasonic cleaner for cleaning cuvets.

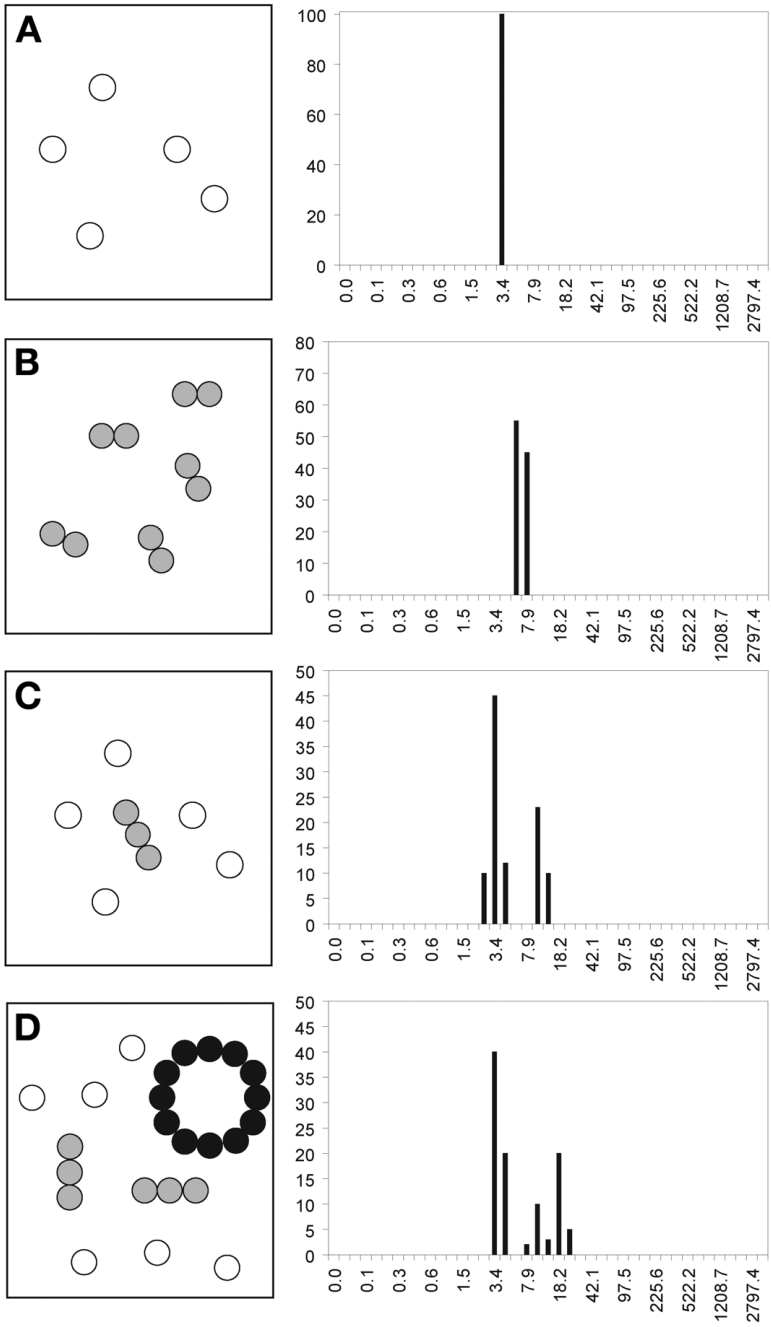

A
B
C
D
0.0
0.1
0.3
0.6
1.5
3.4
7.9
18.2
42.1
97.5
225.6
522.2
1208.7
2797.4

Fig. 3. Examples of solutions that differ in composition (left) and fake dynamic light scattering regularization histograms (right). The relative amount of light scattered by each bin, percentage of intensity, is plotted against the discrete particle sizes, R_H, in nanometers on a log scale. **(A)** A monodisperse, monomodal solution of monomers (with mean R_H of 3.4 nm, %Pd of 10%); very likely to crystallize. **(B)** A monodisperse, monomodal solution of dimers (with mean R_H of 6.8 nm, %Pd of 14%). The polydispersity is greater than the monomeric solution but it is still very likely to crystallize. **(C)** A bimodal solution of monomers contaminated by trimers (with mean R_H of 3.4 and 10.2 nm). Less likely to crystallize so put in less screening effort. **(D)** A multimodal solution of monomer, trimer, and dodecameric aggregates (with mean R_H of 3.4, 10.2, and 19.3 nm). Unlikely to crystallize but you might as well give it a little try.

Note: the mean R_H is defined by the weighted average of the number of bins comprising the peak. The polydispersity of each peak is indicated by the width. Species that differ in R_H by more than 50% are separable if their polydispersity is small. For example, the dimers and trimers differ in size by only 33%. Thus, if solution B was mixed with solution C there would be a single peak (monomodal) and it would be very broad ranging from 2.6 to 13.8 nm. This mixture would be monomodal with mean R_H of 5.9 nm, very polydisperse with %Pd of around 50%, and highly unlikely to crystallize.

11. Sterile plastic transfer pipets.
12. 100-µL Pipetteman with capillary pipet tips.
13. Table-top microcentrifuge and appropriate tubes.
14. Microconcentrators.
15. Good buffers (sodium salts) for use in Mueser's solubility screen, each at 100 m*M*: MES-NaOH, pH 5.8; PIPES-HCl, pH 6.5; HEPES-HCl, pH 7.5; and TAPS-HCl, pH 8.5.
16. Chloride salts for use in Mueser's solubility screen, each at 100 m*M*: NH_4Cl; NaCl; KCl; LiCl; $MgCl_2$; and $CaCl_2$.
17. Ammonium salts for use in Mueser's solubility screen, each at 100 m*M*: NH_4 formate; NH_4 acetate; NH_4 cacodylate; $(NH_4)_2\ SO_4$; $(NH_4)_3\ PO_4$; and NH_4 citrate.

3. Methods

The methods described next outline (1) how to prepare the sample, (2) how to take DLS measurements, and (3) how to interpret the results. The DynaPro instrument is run by the Dynamics software package (*see* **Note 1**). The methods and strategies described in this chapter can be applied to any DLS instrument, although the specific details of the instrument and software may be different.

3.1. Preparation of Instrument

3.1.1. Measuring DLS Data

While preparing for the experiment, many DLS measurements should be taken to check the state of the instrument. First, turn on the DynaPro MS/X

instrument, the temperature control unit, and start the Dynamics software. Place the quartz cuvet containing your sample or water in the sample holder on the optics block. The frosted side of the cuvet must point to the left side of the holder, as marked on the instrument. In the software, open a "New" experiment. On the left side of the Experiment window is the "Tree View." The Tree View is used to select groups or categories of information for viewing in the display side of the Experiment window. There are three main nodes in the Tree View: Hardware, Parameters, and Measurements. Connect the instrument using the Hardware node. Set the temperature using the Parameters → Instrument node. Then proceed to take DLS measurements by clicking on the green "Start" button on the Experiment window tool bar. Subcategories in the Measurements node are the individual measurements (Meas no.), each of which is further broken down into Acquisitions (Acq no.) and Readings (Read no.). The display format for the information in the Measurement node is dependent on which view button is selected in the Experiment window tool bar (e.g., Datalog grid or Regularization graph). By default, each data acquisition is accumulated over a 10-s window of time. During the course of an experiment, each acquisition collected is displayed in the Measurements node along with the corresponding calculated data. Ten or more acquisitions are recommended per measurement. Each acquisition is the average of 10 readings. The Cumulants, displayed in the Datalog grid, and Regularization analysis data is calculated by the software and can be displayed at any level of detail, including the average over all the data, the average over each acquisition, or over each reading. To stop data collection, click on the red "Stop" button in the Experiment window tool bar.

3.1.2. Cuvet Cleaning and Clean Water Count

The DLS cuvet and apparatus must be very clean to ensure good data quality. Therefore, it is highly recommended that the first step of the DLS experiment is to measure a good clean water count. The clean water count also checks the condition of the instrument components.

1. Fill the cuvet with 0.2-μm filtered deionized water. Avoid air bubbles.
2. Take DLS measurements to get a clean water count rate.
3. If the clean water count is reasonable and steady, you can start taking DLS measurements on your sample. For the DynaPro MS/X instrument counts less than 25,000 are good. Less sensitive instruments, such as the DynaPro 801, may have counts less than 10,000 for clean water. *See* **Example 1** for an example of acceptable clean water data. Reasonable clean water count rates are similar to those made when the cuvets and instrument were new. The clean water count rates should be noted in the instrument log book. If it is too high the cuvets and/or the microfilter kit need to be thoroughly cleaned or, in the worst case, should be replaced.

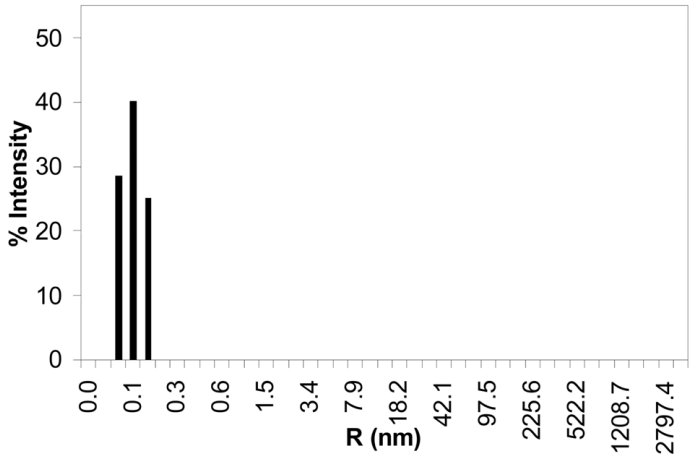
Remember, any large particles, such as dust, will scatter light intensely and will interfere with the DLS signal from the molecules of interest. Therefore, it

Example 1
Acceptable Clean Water Count Data

Data log grid

Item	Time (s)	Temp (C)	Intensity (Cnt/s)	R (nm)	%Pd	MW-R (kDa)	Amp	Baseline	SOS
Acq 1	9.5	25	16340	0.3	0.1	0	0.026	1.002	165.138
Acq 2	19.5	25	18,240	0.1	0.1	0	0.174	1.002	117.940
Acq 3	29.5	25	18,195	0.1	0.1	0	0.159	1.001	100.560
Acq 4	39.5	25	18,559	0.2	0.1	0	0.100	1.001	71.775
Acq 5	49.5	25	19,285	0.6	0.1	1	0.013	1.001	118.706
Acq 6	59.5	25	17,181	0.5	0.2	1	0.025	1.001	150.398
Acq 7	69.5	25	17,748	0.1	0.1	0	0.297	1.002	149.658
Acq 8	79.6	25	19,150	0.1	0.1	0	0.232	1.000	99.076
Acq 9	89.6	25	16,615	0.5	0.1	1	0.057	1.000	124.772
Acq 10	99.7	25	15,131	0.1	0.1	0	0.762	1.000	138.819
Mean									
Meas 1	99.7	25	17,644.4	0.26	0.1	0	0.185	1.001	123.684

Regularization graph



Regularization results summary

Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
☑ Peak 1	0.1	21.2	0	100.0	100.0

Interpretation: this is excellent water count data for the DynaPro MS/X. The regularization data at R_H < 0.5 nm is owing to noise in the detector. Some people call it the solvent peak, but it is really because of after pulse noise in the photodiode. This noise surfaces from time to time in the data if the protein concentrations are too low. In this case, it can be ignored or filtered out of the regularization fit by clicking on the check mark for that peak in the Regularization Results Summary. Note that the SOS errors are high with water count data because there are no macromolecules present.

is necessary to thoroughly clean the cuvet before and after use. Be careful not to scratch the cuvets. Clean all dust off the outer surface. The following cuvet-cleaning procedure is recommended (*see* **Note 2**):

1. Using a sterile plastic transfer pipet, flush the cuvet multiple times with a 1% Triton X-100 solution.
2. Rinse the cuvet three to five times with sterile-filtered, deionized water.
3. The interior of the cuvet can be dried using compressed air. Alternatively, if you have more time, invert the cuvet and allow it to dry.
4. Polish the exterior surface with lens paper and remove dust with compressed air. Note: tissues and other wipes should not be used as they can scratch the surface of the cuvet. The cuvet cap must also be dust free.
5. Repeat the clean water count and cleaning procedure until a reasonably low clean water count is obtained. The clean water count must be stable for 2–3 min.

3.2. Preparation of Sample

3.2.1. Estimation of Minimal Concentration

In order to make efficient use of your samples, it is helpful to know the minimum protein concentration needed for DLS measurements. Usually, it is safe to assume that DLS measurements made at low protein concentration represent fairly well the sample at the much higher concentrations typically used by crystal farmers. Also, if needed, DLS measurements can be repeated at higher concentrations. The minimal protein concentration needed is dependent on the MW of the macromolecule and the particular instrument. Smaller proteins will need to be at a higher concentration. For example, for the DynaPro MS/X the minimal concentration for a 10-kDa protein is approx 0.6 mg/mL and for a 100-kDa protein is 0.06 mg/mL. To estimate the minimum protein concentration that is needed for any DynaPro model use Tools→Calculations→Optimization in Dynamics v6. To obtain higher quality DLS data use a protein concentration two- to threefold higher than that recommended by the Calculator.

3.2.2. Sample Preparation

Before DLS measurements can be taken the sample must also be cleaned of any dust or other particles. All buffers must be also cleaned by filtration. This can be done by using the microfilter kit provided by Protein Solutions or by centrifugation. Centrifugation is the easiest. The microfilter kit is more difficult to use but has the advantage of removing particles by pore size (*see* **Note 3**). The procedure used to clean the sample for DLS must also be used to prepare the sample before crystallization.

1. Centrifugation procedure:
   a. Prepare the sample (e.g., thaw frozen stock, dilution of stock, or mixed components of a complex) in a suitable clean, dust-free microcentrifuge tube. The microcentrifuge tubes can be purged of dust with compressed air.

b. Centrifuge 5–10 min at 15,000*g* in a table-top microcentrifuge.
c. Transfer supernatant with a 100-µL pipetman with clean capillary tips to a dust-free, clean microcentrifuge tube or pipet the supernatant directly into the clean DLS cuvet. Do not disturb the pellet. Remember, only the top portion of the sample is dust free after centrifugation.

2. Microfilter kit procedure:
   a. Disassemble the microfilter system and syringe completely. Thoroughly clean all parts by rinsing/soaking in deionized water and then air-dry the parts. If needed, ultrasonication or 1% Triton X-100 can be used to clean the parts followed by thorough rinsing with deionized water.
   b. Partially reassemble microfilter system by fitting the Teflon housing into the metal housing and seat the O-rings properly in each half.
   c. Using the tweezers, place a filter disk into the "needle" half of the metal housing on top of the O-ring (*see* **Note 4**).
   d. Tightly screw the two metal housing pieces together. The filter disk will be held in place by the two O-rings.
   e. Reassemble the syringe and load with filtered water. Insert the syringe needle into the Teflon needle guide in the filter housing. Filter water through the microfilter system by pressing gently on the syringe plunger. Dispense into the cuvet and take clean water count data again to ensure that the microfilter system is clean. Keep passing water through the microfilter system until the clean water count test is passed.
   f. Load the syringe with filtered buffer. Thoroughly and gently rinse the filter with buffer before filtering your protein sample.
   g. Remove the syringe from the housing and dispense any remaining buffer to waste.
   h. If you are using the 12-µL cuvet, then load the syringe with 20 µL or more of protein sample and reinsert the syringe needle into the filter housing. Approximately 5–8 µL of sample will be lost to the filter system.
   i. Gently depress the syringe plunger to dispense one to two drops onto a paper. These two drops are sufficient to displace any remaining buffer in the needle that was used to wet the filter disk.
   j. The sample can now be directly loaded into the cuvet from the microfilter needle. Remove any air bubbles that are created, for example by sucking them back into the needle. Place the cap on the cuvet.
   k. Disassemble the microfiltration system and thoroughly clean and dry all components before placing them back into the case.

3.3. DLS Data on the Sample

3.3.1. Measuring DLS Data on the Sample

Place the quartz cuvet containing your sample in the sample holder on the optics block. Set the temperature (*see* **Note 5**) and proceed to take DLS measurements. At least 10–20 acquisitions should be taken for each solution condition. To check reproducibility, the measurements should be done in duplicate or triplicate if there is enough sample. After measurement, the sample can be recovered

using a pipetman and capillary-style pipet tips for use in crystallization screens or other experiments.

3.3.2. Interpretation of DLS Data

Before you start to analyze your DLS results make sure the solvent and sample conditions are properly entered into the Dynamics software. The viscosity and refractive index of the buffer you are studying can adversely affect the estimate of R_H and D_T, respectively. The Dynamics software has a pull-down menu of frequently used buffers under the Parameters → Solvent node. These are fairly accurate if your protein concentrations are low (*see* **Note 6**).

When interpreting the DLS data there are several things to keep in mind. First, know the limitations of the instrument. For example, the Protein Solutions MS/X has a lower limit of 0.5 nm and an upper limit of 1-μm particle size. Be aware that you will not be able to deconvolute the intensity measurements coming from your sample of interest, the buffer system the protein is in, or impurities in the solution. If you are using a complicated buffer system, DLS data on the buffer alone can be helpful in interpreting the results (*see* **Example 2**). Data from other experiments such as native polyacrylamide gel electrophoresis, electron microscopy, and size-exclusion chromatography can be helpful in interpreting results ***(6)***.

In studying the results from Dynamics, first study the measurements statistics table to evaluate the quality of your data and the modality of your sample (**Fig. 3**). Outliers can be filtered out using the "Data Filter" or individually marked manually. Both methods are accessed by a right-click on the Measurements Datalog grid. Guidelines for marking outliers are given in the Help software. Guidelines for the interpretation of the statistics in the Measurements Datalog grid are outlined in **Table 1** and described next.

1. Use the baseline parameter to judge if your sample is monomodal, bimodal, or multimodal. The quality of the fit of the data to a given autocorrelation function is indicated by the baseline value. Monomodal distributions are defined by a baseline ranging from 0.997 to 1.002. Bimodal distributions have a baseline range of 1.003 to 1.005. Baselines greater than 1.005 are from multimodal samples, dust, or noise.
2. The Dynamics software determines the uniformity of sizes through a monomodal (single particle size with a Gaussian distribution) curve fit analysis called Cumulants. The quality of the data is represented in the sum of squares (SOS) error statistic reported for each sample acquisition (a single correlation curve) in the Datalog grid view of Dynamics v6. The SOS error is the SOS difference between the measured data and the Cumulants-calculated intensity correlation curves. The SOS errors less than 20.0 are good and errors less than 5.0 are considered negligible (*see* **Note 7**) and probably represent the best samples. We have noted that, with the higher sensitivity provided by the DynaPro MS/X, the SOS errors on polydisperse samples tend to be higher than on less sensitive instruments such as the DynaPro 801, nevertheless the rules in **Table 1** still hold true.

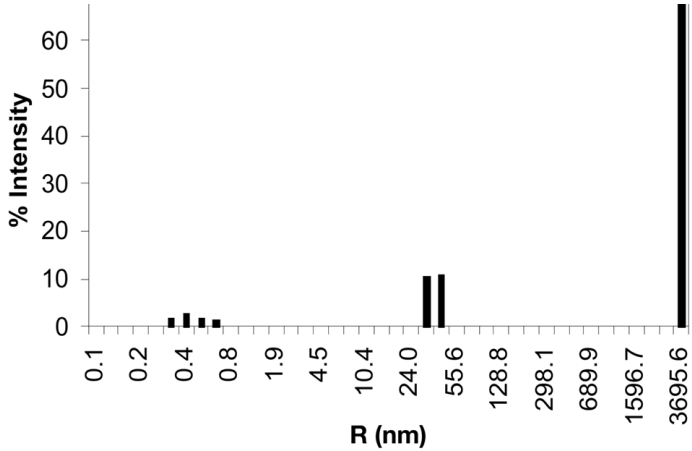
Example 2
Part A: DLS Data on Replication Protein A Buffer

Data log grid[a]

Item	Time (s)	Temp (C)	Intensity (Cnt/s)	R (nm)	%Pd	MW-R (kDa)	Amp	Baseline	SOS
Acq 1	10	25	80,513	52.2	120.5	35238	0.24	1.003	302.977
Acq 2	20	25	73,262	51.6	120.1	34,243	0.26	1.01	508.992
Acq 3	30	25	59,220	32.1	146	11,281	0.19	1.003	518.801
Acq 4	40.1	25	71,035	47.1	118.3	27,704	0.25	1.006	478.842
Acq 5	50.1	25	59,948	36.6	159.8	15,337	0.19	1.003	461.535
Acq 6	60.1	25	66,565	41.9	154.4	20,973	0.19	1.009	399.663
Acq 7	70.1	25	91,962	83.1	144.2	104,370	0.3	1.067	492.541
Acq 8	80.1	25	94,199	113.1	208.5	214,887	0.32	1.131	609.86
~~Acq 9~~	~~90.1~~	~~25~~	~~13,3398~~	~~326.3~~	~~244.7~~	~~2,562,320~~	~~0.534~~	~~1.344~~	~~1531.05~~
~~Acq 10~~	~~100.1~~	~~25~~	~~37,1576~~	~~1072.3~~	~~50.9~~	~~414,494,00~~	~~1.046~~	~~2.359~~	~~458.088~~
Mean									
Meas 1	220.3	25	299261	103.7	188.9	175181	0.29	1.15	531.697

[a]For this sample, 20 data points were taken but only the first 10 are shown. The following data filter was applied before regularization of the data: minimum amplitude 0, maximum amplitude 1, baseline limit 1 ± 1, maximum SOS 1000. Thus, data for Acq 9 & 10 did not pass the filter and were not used as indicated by the strikethrough. For this data 13% did not pass the filter.

Regularization graph



Regularization results summary

Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
☑ Peak 1	0.4	30.6	0	8.4	100.0
☑ Peak 2	37.1	13.8	15,783	21.7	0.0
☑ Peak 3	3695.6	0.0	749,619,000	69.9	0.0

Interpretation: The SOS and baseline are very high because there is no macromolecule present and the buffer is polydisperse. Compared with **Example 1,** Peak 1 can be attributed to noise in the detector. Peaks 2 and 3 are from components in the buffer and are probably due to cHEGA10 detergent micelles.

(Example 2 continues)

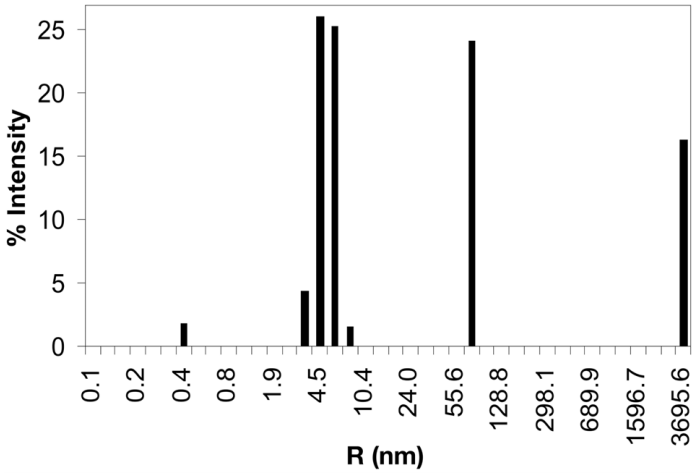
Example 2 *(continued)*
Part B: DLS Data on RPA Heterotrimer at 0.9 mg/mL

Data log grid[a]

Item	Time (s)	Temp (C)	Intensity (Cnt/s)	R (nm)	%Pd	MW-R (kDa)	Amp	Baseline	SOS
Acq 1	9	25	464,697	11.9	104.1	1101	0.588	0.999	938.809
Acq 2	19	25	366,039	7.6	61.6	385	0.587	1.003	150.868
Acq 3	29.1	25	383,781	8	74.2	439	0.574	1.006	194.406
~~Acq 4~~	~~39.1~~	~~25~~	~~1,295,610~~	~~148.0~~	~~241.4~~	~~402739~~	~~0.233~~	~~1.319~~	~~1136.620~~
~~Acq 5~~	~~49.1~~	~~25~~	~~1,255,210~~	~~147.0~~	~~191.2~~	~~396358~~	~~0.304~~	~~1.327~~	~~1558.530~~
~~Acq 6~~	~~59.1~~	~~25~~	~~610,538~~	~~19.3~~	~~115.9~~	~~3426~~	~~0.529~~	~~1.017~~	~~1780.420~~
Acq 7	69.1	25	485,968	12.4	101.9	1228	0.594	1.004	803.41
Acq 8	79.1	25	648,094	15.9	103.8	2180	0.53	1.018	838.817
Acq 9	89.2	25	393,460	8.9	91.2	553	0.623	1.002	288.664
Acq 10	99.2	25	374,938	8.3	75.5	470	0.627	1	235.677
Mean									
Meas 1	249.4	25	517,431	10.1	95.3	761	0.574	1.013	529.

[a]Greater than 20 data points were taken, only the first 10 are shown. The following data filter was applied before regularization of the data: minimum amplitude 0, maximum amplitude 1, baseline limit 1 ± 1, maximum SOS 1000. Thus, data for Acq 4–6 did not pass the filter and were not used as indicated by the strikethrough. For this data 20% did not pass the filter.

Regularization graph



Regularization results summary

Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
☑ Peak 1	0.4	0.0	0	1.7	98.6
☑ Peak 2	5.1	18.5	154	55.3	1.4
☑ Peak 3	73.6	0.0	78,597	23.4	0.0
☑ Peak 4	3695.6	0.0	749,619,000	19.6	0.0

(Example 2 continues)

Example 2 *(continued)*

Interpretation: the data are very polydisperse as indicated by high baseline and high SOS error. Owing to the polydispersity of the buffer, comparison to the buffer alone DLS data is needed for interpretation of the protein sample DLS data. Peak 1, observed in the buffer only data too, is due to the low protein concentration used. Peaks 4 and 3 in buffer only data, is because the polydispersity of the buffer and probably represent cHEGA10 micelles. Identification of peak 3 components is ambiguous and may be due to the buffer components, protein or both. Peak 2 is owing to the protein, in this case RPA heterotrimer (~110 kDa). DLS could be used in this case to find the minimal concentration of CHEGA10 to solubilize the protein without micelles. Note that the peak attributed to detector noise is dominating the percentage mass calculation.

Table 1
Interpretation and Use of the Statistical Parameters Calculated by Dynamics v6[a]

Parameter	Interpretation
Baseline	
0.997–1.002	Monomodal distribution
1.003–1.005	Bimodal distribution
>1.005	Multimodal distribution
Sum of squares (SOS)	
1.000–5.000	Low noise, negligible error
5.000–20.000	Background error because of noise, low protein concentration, or a small amount of polydispersity
>20.000	High noise/error owing to high polydispersity in size distribution (aggregation), irregular solvent
Normalized polydispersity	Note, this parameter should be used for monomodal distributions only.
%Pd < 15	Monodisperse solution, very likely to crystallize
%Pd < 30	A moderate amount of polydispersity, more likely to crystallize
%Pd > 30	A significant amount of polydispersity, less likely to crystallize

[a]Adapted from the DynaPro Operator Manual, Protein Solution, Inc. Note %Pd in Dynamics v6 was called C_p/R_H in older versions of the software.

3. The polydispersity statistic will tell you the likelihood of crystallizing your sample. The polydispersity (Pd or standard deviation) is indicative of the distribution in the peak or subpeak. By default, %Pd, or normalized polydispersity, is listed in the Datalog grid and the Regularization results summary in Dynamics v6.0. Here, %Pd is calculated by dividing Pd by R_H and reported as a percent. In older

versions of the Dynamics software, this statistic was called C_p/R_H. If your sample is monomodal, the mean %Pd of the sample can be read straight off the Datalog grid. If your sample is multimodal then mean R_H and %Pd for each peak can be obtained from the Regularization graph. If the sample is monomodal and the %Pd is less than 15% your sample is monodisperse and very likely to crystallize during screening (*see* **Note 8**). Go on to screen the sample for crystallization (**Fig. 1**), perhaps gauging your level of effort on the quality of the DLS analysis of that sample.

3.4. How to Use DLS Data to Find Conditions That Will Improve Crystallization Results

If your sample is multimodal, or monomodal but polydisperse (%Pd > 30%), or simply will not crystallize, the following experimental considerations can help improve your DLS and crystallization results.

1. Increasing the solubility of the sample will typically decrease the aggregation and polydispersity of the sample. Therefore, it is helpful to perform the Mueser solubility screen on your sample to optimize buffer conditions. The following protocol that has been adopted successfully several times ***(7,8)***. It is especially powerful when coupled to the DLS polydispersity analysis.
   a. Dialyze 2–5 mg protein against deionized water (no buffer, no salt). Most proteins will precipitate under these conditions.
   b. Resuspend the precipitated protein and aliquot the precipitate into 20 1.5-mL microcentrifuge tubes. Centrifuge at maximum speed in a table-top centrifuge for 2–5 min to repellet the protein. Remove the supernatant.
   c. Each tube will be a separate experiment. Add either buffer, chloride salt, or ammonium salt (20 mL of 100 m*M* solutions described in **Subheading 2.**, **steps 15–17**), and resuspend the pellet. Let stand at room temperature for 10 min, centrifuge to pellet the undissolved protein, and measure the protein concentration of the supernatant amount redissolved. If the pellets completely dissolve then use less volume or more protein.
   d. When you have information about which buffer and/or salt are best, it is suggested that you try a cross-coupled experiment. For example, if LiCl and $(NH_4)_2SO_4$ give good solubility, then perhaps $(Li)_2SO_4$ is worth trying.
   e. Test the final buffer condition for maximum solubility using mini concentrators. The best starting protein concentration for crystallization is one-half of the maximum solubility.
2. Perform a series of DLS experiments to test the effect of ionic strength, pH, protein concentration, organic solvents, detergents, and other additives on the polydispersity of your sample (*see* **Example 3**). Perform crystallization experiments at the solution condition where your sample is the most monodisperse and stable over time.
3. Test the effect of temperature on your sample. Find the temperature where your protein is the most monodisperse and then set the crystallization incubator to this temperature for crystal growth. The Event Scheduler can automate these experiments (right-click on Parameters node to access the Event Scheduler).

Example 3 Part A:
Monodisperse DLS Data on Crystallizable RPA14/32 Heterodimer, at a Concentration of 10 mg/mL

Data log grid[a]

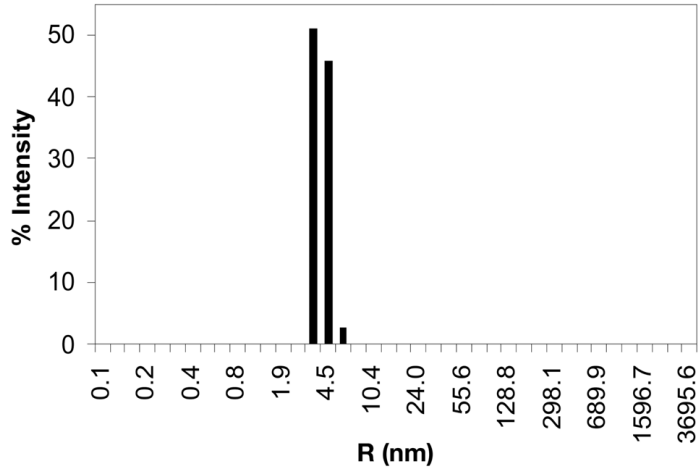
Item	Time (s)	Temp (C)	Intensity (Cnt/s)	R (nm)	%Pd	MW-R (kDa)	Amp	Baseline	SOS
Acq 1	10	25	1,872,940	3.9	22.5	80	0.496	1.000	0.812
Acq 2	20	25	1,872,590	3.9	14.7	83	0.552	1.000	1.145
Acq 3	30	25	1,888,920	3.9	17.2	81	0.544	1.000	0.761
Acq 4	40.1	25	1,894,110	3.9	24	82	0.521	1.000	1.124
Acq 5	50.1	25	1,908,650	3.9	29.1	84	0.51	1.000	1.83
Acq 6	60.1	25	1,892,830	3.9	23.5	82	0.504	1.000	1.239
Acq 7	70.1	25	1,891,400	3.9	21.5	82	0.495	1.000	0.862
Acq 8	80.1	25	1,888,290	3.9	23.6	83	0.489	1.000	1.355
Acq 9	90.1	25	1,882,720	3.9	10.9	83	0.49	1.000	1.385
Acq 10	100.1	25	1,876,270	3.9	29.7	82	0.49	1.000	1.071

Mean

Item	Time (s)	Temp (C)	Intensity (Cnt/s)	R (nm)	%Pd	MW-R (kDa)	Amp	Baseline	SOS
Meas 1	300.4	25	1,867,130	3.9	12.2	83	0.509	1.000	1.1375

[a]Greater than 20 data points were taken, only the first 10 are shown. The application of a data filter was not needed.

Regularization graph



Regularization results summary

Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
☑ Peak 1	4.0	15.8	84	100.0	100.0

Interpretation: the sample gave a monomodal fit with a baseline of 1.000 and SOS of 1. RPA14/32 is a 46-kDa heterodimer with a predicted R_H of 3.1 nm. A dimer of heterodimers has a predicted R_H of 4.1 nm. Therefore, these data indicate primarily a dimer of heterodimers in solution. With a %Pd of 16%, this sample crystallized readily out of several crystallization conditions and several space groups ***(10)***.

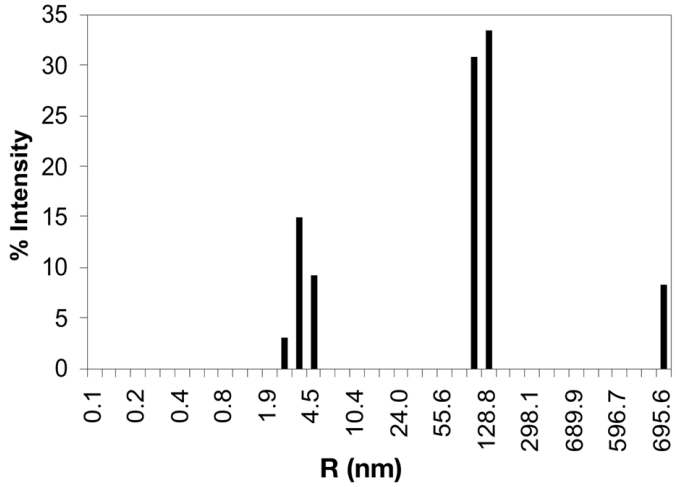
Example 3 Part B:
Multimodal DLS Data on RPA14/32 Heterodimer, at a Concentration of 5 mg/mL, diluted With 30 m*M* HEPES pH 7.8

Data log grid[a]

Item	Time (s)	Temp (C)	Intensity (Cnt/s)	R (nm)	%Pd	MW-R (kDa)	Amp	Baseline	SOS
Acq 1	10	25	3,976,320	69.5	120.1	68,813	0.336	1.010	401.397
Acq 2	20	25	2,970,080	42.4	137	21568	0.377	1.003	837.184
Acq 3	30	25	2,968,450	43.6	118.7	23036	0.384	1.003	832.709
Acq 4	40.1	25	2,327,790	21.9	115.9	4613	0.386	1.003	958.650
Acq 5	50.1	25	2,388,140	22.8	118	5087	0.371	1.003	833.651
Acq 6	60.1	25	2,560,250	29.7	116.5	9434	0.388	1.003	984.212
Acq 7	70.1	25	2,680,860	32.8	116.8	11840	0.39	1.002	902.132
Acq 8	80.1	25	2,796,570	35.8	118	14553	0.368	1.000	849.627
Acq 9	90.1	25	2,777,060	36	114.7	14790	0.365	1.001	876.540
Acq 10	100.2	25	2,600,990	27.6	116.2	7896	0.374	1.001	914.389
Mean									
Meas 1	310.5	25	2925130	38.3	118.1	17040	0.362	1.002	819.282

[a]Greater than 20 data points were taken, only the first 10 are shown. The application of a data filter was not needed.

Regularization Graph



Regularization Results Summary

Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
☑ Peak 1	3.7	17.4	70	27.4	99.9
☑ Peak 2	113.7	13.8	217490	64.3	0.0
☑ Peak 3	3695.6	0.0	749619000	8.3	0.0

Interpretation: the increased baseline and very high SOS error indicate that dilution with this buffer introduces polydispersity into the sample. The regularization fit indicates three peaks. Peaks 2 and 3 are from aggregated protein. When many data acquisitions are taken this aggregation becomes worse over time (data not shown). Therefore, this is not a suitable buffer condition for this protein.

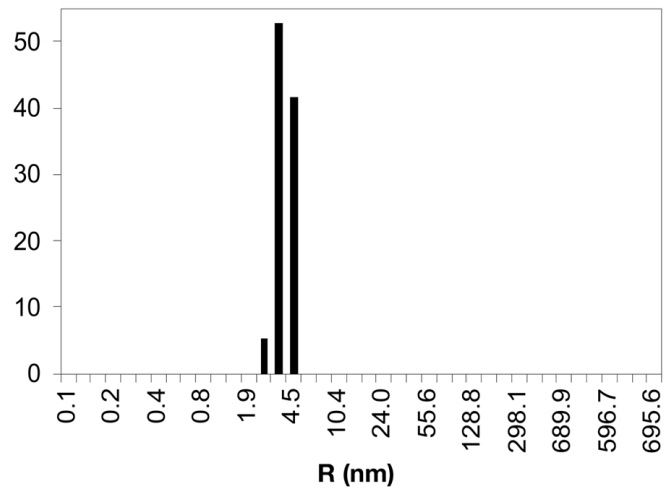
Example 3 Part C:
Monodisperse DLS Data on RPA14/32 Heterodimer at a Concentration of 5 mg/mL, Diluted with 30 m*M* HEPES pH 7.8, 200 m*M* KCl, and 10 m*M* DTT; Crystallizable in This Solution Condition

Data log grid[a]

Item	Time (s)	Temp (C)	Intensity (Cnt/s)	R (nm)	%Pd	MW-R (kDa)	Amp	Baseline	SOS
Acq 1	9	25	1,047,070	3.6	14.6	68	0.531	1.000	1.244
Acq 2	19	25	1,053,610	3.8	30.6	77	0.540	1.000	2.441
Acq 3	29	25	1,053,860	3.7	35.1	73	0.528	1.000	1.797
Acq 4	39.1	25	1,061,240	3.8	26.2	76	0.509	1.000	2.154
Acq 5	49.1	25	1,067,120	3.7	21.3	73	0.515	1.000	0.983
Acq 6	59.1	25	1,029,520	3.8	21.4	77	0.593	1.000	2.747
Acq 7	69.1	25	1,026,860	3.9	28.5	80	0.599	1.000	2.787
Acq 8	79.1	25	1,034,900	3.8	19.3	77	0.592	1.000	1.178
Acq 9	89.1	25	1,054,730	3.8	30.4	77	0.566	1.000	2.361
Acq 10	99.1	25	1,074,230	3.8	32.9	79	0.545	1.000	2.580
Mean									
Meas 1	329.5	25	1056810	3.8	24.2	75	0.532	1.000	1.581

[a]Greater than 20 data points were taken, only the first 10 are shown. The application of a data filter was not needed.

Regularization graph



Regularization results summary

Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
☑ Peak 1	3.8	15.9	77	100.0	100.0

Interpretation: when compared with **Parts A** and **B** of **Example 3**, dilution of the protein into this buffer is good. The baseline of 1.000 and SOS of 1.6 shows that the sample is monomodal. The %Pd of 16 indicates that the protein is still monodisperse after dilution. Therefore, the protein is stable in this buffer and this is a good starting point for crystallization trials.

4. Test the effect of binding partners (protein, peptides, or oligonucleotides) or substrates on R_H and monodispersity. The effect of stoichiometry of mixing can also be tested (*see* **Example 4**). Anything that makes the molecule smaller and more compact may render it more crystallizable ***(9)***.
5. Use DLS analysis to help optimize the protein purification protocol, e.g., to avoid inappropriate disulfide bond formation ***(10)***.
6. The protein itself should be considered a crystallization variable ***(11)***. Use DLS analysis to help select the best construct. For example, perhaps three deletions of different lengths are made from the N-terminus of the protein. Then put your biggest crystallization effort into the sample with the best monodispersity.
7. DLS is also helpful to test the effects of storage procedures, e.g., freezing vs refrigeration and to assess shelf life ***(12–18)***.

4. Notes

1. Protein Solutions has also written "Dynapro Data Interpretation Guide;" be sure to obtain a copy. In addition, several texts and articles have been written on the collection and analysis of DLS data ***(12–16)***.
2. If the protocol described does not clean the cuvet try placing it in a sonicating bath for 15–20 min and then clean it again. Do not use *concentrated* acids and bases to clean the cuvet as they can etch the surface. Do not use organic solvents (e.g., ethanol) to rapidly dry the cuvet as they can leave a thin residue on the surface. Always clean the cuvet thoroughly after use and before storage. If it is not stored

Example 4
DLS Data on a RPA:Rad52 Complex at a Concentration of 8 mg/mL

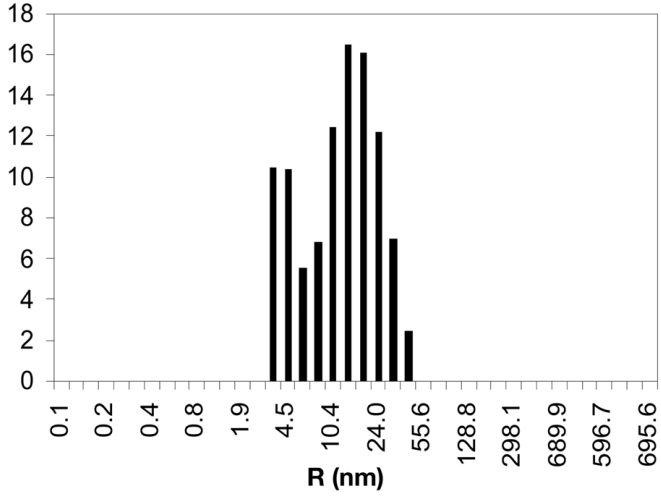
Data log grid[a]

Item	Time (s)	Temp (C)	Intensity (Cnt/s)	R (nm)	%Pd	MW-R (kDa)	Amp	Baseline	SOS
Acq 1	9	25	4,495,800	10.8	64.5	872	0.311	1.001	14.876
Acq 2	39.1	25	5,018,720	10.8	54.1	875	0.332	1.000	20.763
Acq 3	49.1	25	4,969,060	10.7	61.9	870	0.328	1.000	19.873
Acq 4	69.1	25	4,941,200	11	61.2	921	0.325	0.999	21.940
Acq 5	89.2	25	4,953,440	11.2	64.7	969	0.334	1.000	20.287
Acq 6	99.2	25	4,894,110	10.9	69.8	903	0.327	0.999	18.665
Acq 7	109.2	25	4,878,090	10.9	72.9	894	0.327	0.999	19.605
Acq 8	119.3	25	4,881,350	11	72	925	0.330	1.000	21.353
Acq 9	129.3	25	4,869,450	10.9	61.5	901	0.330	1.001	20.228
Acq 10	139.3	25	4,835,920	10.8	65.7	879	0.334	1.000	18.834
Mean									
Meas 1	289.5	25	4,939,260	10.9	61	904	0.32	1.000	18.077

[a]Greater than 20 data points were taken, only the first 10 are shown. The application of a data filter was not needed.

Example 4 ***(continued)***

Regularization graph



Regularization results summary

Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
☑ Peak 1	4.4	21.8	106	26.4	92.0
☑ Peak 2	18.0	45.5	2901	73.6	8.0

Interpretation: the sample in this case is a complex of two proteins, RPA heterotrimer (110 kDa) and Rad52 heptameric ring (~350 kDa), mixed with an approximate one to one molar ratio. The data are polydisperse as indicated by the high SOS error. The regularization analysis shows that the sample is bimodal and polydisperse. This sample did not crystallize. Possible reasons are that the molar ratio was not exact or that a portion of the molecules is inactive. The next step is to vary the stoichiometry of RPA to Rad52 to try to obtain 100% monodisperse complex in solution. If this is not successful the complex will need to be separated from free RPA by size exclusion chromatography before crystallization trials.

in a clean state, you will cause the next user many headaches. If the cuvet window becomes scratched it will need to be replaced.

3. The sample may be unexpectedly lost during the filtration process. This could be because of aggregation, unexpected quaternary structure, or unusual binding to the filter. Also, it is important to take a protein reading or run a sodium dodecyl sulfate-polyacrylamide gel electrophoresis PHAST gel (Amersham Biosciences, Piscataway, NJ) on 1 μL of sample before and after filtration to access how much, if any, is lost owing to filtration. Consider pore size and the MW of your protein. The Dynamics software includes a MW calculator that will estimate R_H for you. Use it to see which pore size to use. For example, do not use 0.020-μm size pores if your protein is larger than 150 kDa. If needed, a filter with larger pore size can be used.

4. The Whatman Anotop filters used with the microfilter kit are very brittle. Care must be taken in handling them so that they do not crack. They must also be seated properly on the O-ring so that solution cannot pass around the filter.
5. Useful information on the effect of temperature on aggregation can be gained from starting the DLS measurements at 4°C and then stepping up the temperature in 5° increments and taking DLS measurements at each temperature from the same sample. The Event Scheduler node in Dynamics software can be used to automate these measurements. To access the Even Scheduler right-click on the Parameters node. The sample should be incubated 30 min at each temperature before DLS data (10–20 measurements) are taken. The maximum temperature for the DynaPro MS/X is to 60ºC. Temperature can be used to control the aggregation of a protein ***(17)***. Ramping down from high temperature has been used in the crystallization of macromolecules, e.g., insulin ***(18)***.
6. The viscosity is influenced by protein concentration and buffer components, such as alcohols and glycerol. For precise measurements of R_H, viscosity and refractive index can be measured with a viscometer and refractive index detector, respectively, and then entered by hand into the software.
7. Regularization analysis will be able to give you some size information about your sample even if the SOS error is high and your baseline is high. Both of these statistics absorb error so the amount of trust you should place in the data in these cases should be proportionately low.
8. The %Pd is the statistic most useful for predicting crystallizability, if your sample is monomodal. A special exception is made for bi- or multimodal samples where one or more of the peaks result from something in the solvent (e.g., detergent micelles) that produces light scattering. If the protein peak can be identified from the noise through appropriate control experiments, the C_p/R_H for this peak off the regularization graph can be used to predict crystallizability. Also, the concentration of detergent can be optimized by DLS to eliminate the presence of micelles in the protein solution.

Acknowledgments

The author would like to thank Tim Mueser and Doba Jackson for intellectual contributions and for critically reading drafts of this manuscript, and Gilson Baía for providing the DLS data used in the examples. Wasantha Ranatunga and Jeff Habel helped develop the laboratory's expertise in DLS. Grants from the Ohio Cancer Research Associates, Inc., the US Army Medical Research and Materiel Command DAMD17-98-1-8251, and the American Cancer Society RSG GMC-104160 supported this work.

References

1. Zulauf, M. and D'Arcy, A. (1992) Light scattering of proteins as a criterion for crystallization. *J. Cryst. Growth* **122,** 102–106.
2. D'Arcy, A. (1994) Crystallizing proteins: a rational approach. *Acta Cryst.* **D50,** 469–471.

3. Wilson, W. W. (2003) Light scattering as a diagnostic for protein crystal growth: a practical approach. *J. Structural Biol.* **142,** 56–65.
4. Bergfors, T. M. (1999) *Protein Crystallization Techniques, Strategies and, Tips.* International University Line, La Jolla, CA.
5. Ferré-D'Amaré, A. R. and Burley, S. K. (1997) Dynamic light scattering in evaluating crystallizability of macromolecules. *Meth. Enzymol.* **276,** 157–166.
6. Ranatunga, W., Jackson, D., Lloyd, J. A., Forget, A. L., Knight, K. L., and Borgstahl, G. E. O. (2001) Human Rad52 exhibits two modes of self-association. *J. Biol. Chem.* **276,** 15,876–15,880.
7. Mueser, T. C., Rogers, P. H., and Arnone, A. (2000) Interface sliding as illustrated by the multiple quaternary structures of liganded hemoglobin. *Biochemistry* **39,** 15,353–15,364.
8. Collins, B. K., Tomanicek, S. J., Lyamicheva, N., Kaiser, M. W., and Mueser, T. C. (2004) A preliminary solubility screen used to improve crystallization trials. Crystallization and preliminary X-ray structure determination of Aeropyrum pernix Flap Endonuclease-1. *Acta Cryst.* **D60,** 1674–1678.
9. Jackson, D., Dhar, K., Wahl, J. K., Wold, M. S., and Borgstahl, G. E. (2002) Analysis of the human replication protein A:Rad52 complex: evidence for crosstalk between RPA32, RPA70, Rad52 and DNA. *J. Mol. Biol.* **321,** 133–148.
10. Habel, J. E., Ohren, J. F., and Borgstahl, G. E. O. (2001) Dynamic light scattering analysis of full-length, human RPA14/32 dimer: purification, crystallization and self-association. *Acta Cryst. D* **D57,** 254–259.
11. Dale, G. E., Oefner, C., and D'Arcy, A. (2003) The protein as a variable in protein crystallization. *J. Struct. Biol.* **142,** 88–97.
12. Pusey, P. N., Koppel, D. E., Schaefer, D. W., Camerini-Otero, R. D., and Koenig, S. H. (1974) Intensity fluctuation spectroscopy of laser light scattered by solutions of spherical viruses: R17, Obeta, BSV, PM2 and T7. I. Light-scattering technique. *Biochemistry* **13,** 952–959.
13. Camerini-Otero, R. D., Pusey, P. N., Koppel, D. E., Schaefer, D. W., and Franklin, R. M. (1974) Intensity fluctuation spectroscopy of laser light scattered by solutions of spherical viruses: R17, Obeta, BSV, PM2 and T7. II. Diffusion coefficients, molecular weights, solvation, and particle dimensions. *Biochemistry* **13,** 960–970.
14. Pecora, R. (1985) *Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy.* Plenum Press, New York, NY.
15. Brown, R. G. W. (1990) Miniature laser light scattering instrumentation for particle size analysis. *Applied Optics* **29,** 1.
16. Phillies, G. D. J. (1990) Quasieleastic light scattering. *Anal. Chem.* **62,** 1049A–1057A.
17. Ranatunga, W., Jackson, D., II, R. A. F., and Borgstahl, G. E. O. (2001) Human Rad52 protein has extreme thermal stability. *Biochemistry* **40,** 8557–8562.
18. Long, M. M., Bishop, J. B., Nagabhushan, T. L., Reichert, P., Smith, G. D., and DeLucas, L. J. (1996) Protein crystal growth in microgravity: bovine insulin, human insulin and human alpha interferon. *J. Crystal Growth* **168,** 233–243.