Application of dynamic light scattering in protein crystallization

Proteau, Ariane; Shi, Rong; Cygler, Miroslaw

Publisher's version / Version de l'éditeur:
https://doi.org/10.1002/0471140864.ps1710s61
Current protocols in protein science, Supplement 61, pp. 17.10.1-17.10.9, 2010-08-01

Application of Dynamic Light Scattering in Protein Crystallization

Ariane Proteau,[1] Rong Shi,[1] and Miroslaw Cygler[1,2]

[1]Department of Biochemistry, McGill University, Montreal, Quebec, Canada
[2]Biotechnology Research Institute, Montreal, Quebec, Canada

ABSTRACT

Success in determining the three-dimensional structure of a macromolecule by X-ray diffraction methods depends critically on the ability to obtain well ordered crystals of the macromolecule in question. Predisposition to crystallization correlates with the homogeneity of the molecules in solution. Dynamic light scattering (DLS) is particularly well suited for evaluating protein homogeneity under multiple conditions and at concentrations commensurate with crystallization conditions. This unit presents a typical protocol for DLS measurements of a protein sample, and describes approaches to improve protein homogeneity in solution. *Curr. Protoc. Protein Sci.* 61:17.10.1-17.10.9. © 2010 by John Wiley & Sons, Inc.

Keywords: dynamic light scattering • crystallization of proteins • crystallography

INTRODUCTION

The essential step in structure determination by the X-ray diffraction method is obtaining diffractable crystals containing the molecule of interest. Crystallization of proteins is largely an empirical procedure; in a typical screening experiment, a concentrated protein sample is subjected to a variety of conditions covering a wide range of precipitating agents, buffers, pH ranges, and additives. The success of crystallization depends critically on both the chemical purity of the protein sample and its homogeneity. The purity of the protein sample is most routinely assessed by SDS-PAGE. Several methods are available to assess protein homogeneity, such as size-exclusion chromatography (SEC), native PAGE, isoelectric focusing, and dynamic light scattering (DLS). Of these, SEC is usually performed as the final step in protein sample preparation and provides information about the initial solution state of the protein at a relatively low concentration as compared to that used for crystallization trials. The length of time required to perform this chromatographic step, and the amount of protein required, make SEC less useful than DLS for testing the association/aggregation state of the protein under a variety of conditions. DLS is particularly well suited for evaluating protein homogeneity under multiple conditions and at concentrations commensurate with crystallization conditions. The availability of high-sensitivity DLS instruments equipped with plate readers makes such a task both easy and fast. In addition, the protein is easily recoverable for other uses following the measurement.

Basic Protocol 1 in this unit details DLS measurement of protein samples using the DynaPro plate reader, while Basic Protocol 2 describes DLS to optimize protein solution behavior for crystallization.

Information Obtained From DLS Measurements

Dynamic light scattering provides certain information about protein particles present in solution. Two main properties important from the viewpoint of crystallizability are particle size distribution and average particle size. The first parameter provides information about the homogeneity of the protein solution (uniformity of particle sizes) characterized

Current Protocols in Protein Science 17.10.1-17.10.9, August 2010
Published online August 2010 in Wiley Interscience (www.interscience.wiley.com).
DOI: 10.1002/0471140864.ps1710s61

by the percentage of polydispersity. For protein solutions with low polydispersity, the radius of gyration (particle size) can be converted to a molecular mass of the particle, and provides information about the oligomeric state of the protein.

Theory

A dynamic light scattering experiment measures fluctuations in the intensity of light scattered from particles present in the solute. These fluctuations result from the coherent scattering by the particles undergoing Brownian motion. Equation 17.10.1 is a mathematical description of these fluctuations of light intensity:

$$C(\tau) = \frac{\int I(t)I(t+\tau)dt}{\left(\int I(t)dt\right)^2} = B + \beta\exp(-2\Gamma\tau)$$

Equation 17.10.1

where $\Gamma = D(c)q^2$, $q = (4\pi \cdot n/\lambda)\sin(\theta)$, $D(c)$ is the concentration-dependent particle translational diffusion coefficient, q is the scattering vector, n is the refractive index of the solvent, θ is the scattering angle, and λ is the wavelength of light. Assuming that a globular protein is roughly spherical, the apparent hydrodynamic radius of the protein is given by Equation. 17.10.2:

$$r_h = \frac{kT}{6\pi\eta D(c)}$$

Equation 17.10.2

where η is the solution viscosity. Finally, r_h is converted to molecular mass based on the empirical formula mol. wt. $= (\alpha \cdot r_h)^\beta$, with $\alpha = 1.68$ and $\beta = 2.3398$, derived from analysis of a large number of proteins (Claes et al., 1992). It should be noted that for proteins whose shape differ significantly from spherical, the estimate of molecular weight can be quite inaccurate (Claes et al., 1992).

Relationship Between DLS and Crystallizability

Dynamic light scattering has become a routine method used in many crystallography laboratories to assess the probability of protein crystallizability, applicable to both soluble and membrane proteins. It is also a useful tool to detect/confirm the formation of a protein complex. The relationship between monodispersity (homogeneity) of the protein solution and the probability of obtaining protein crystals was initially shown using relatively small subsets of proteins (Baldwin et al., 1986; Zulauf and D'Arcy 1992; Ferre-D'Amare and Burley 1994; D'Arcy 1994). Recently, structural genomics efforts have produced much more exhaustive data that clearly support the relationship between monodispersity and crystallizability (Niesen et al., 2008; Price et al., 2009). The relationship is as follows: protein solutions that were used successfully to obtain protein crystals were predominantly monodisperse; protein solutions that showed significant polydispersity led predominantly either to no crystals at all or to poorly diffracting crystals. However, monodispersity of a protein sample is a necessary but insufficient condition for successful crystallizability, i.e., monodisperse protein solutions often do not lead to crystals in high-throughput screening. Therefore, while good behavior in DLS cannot be viewed as guaranteeing crystallization success, it should definitely be considered as a tool to evaluate crystallizability and, importantly, as a tool to improve protein behavior in solution by modifying the conditions and examining the effect of various additives.

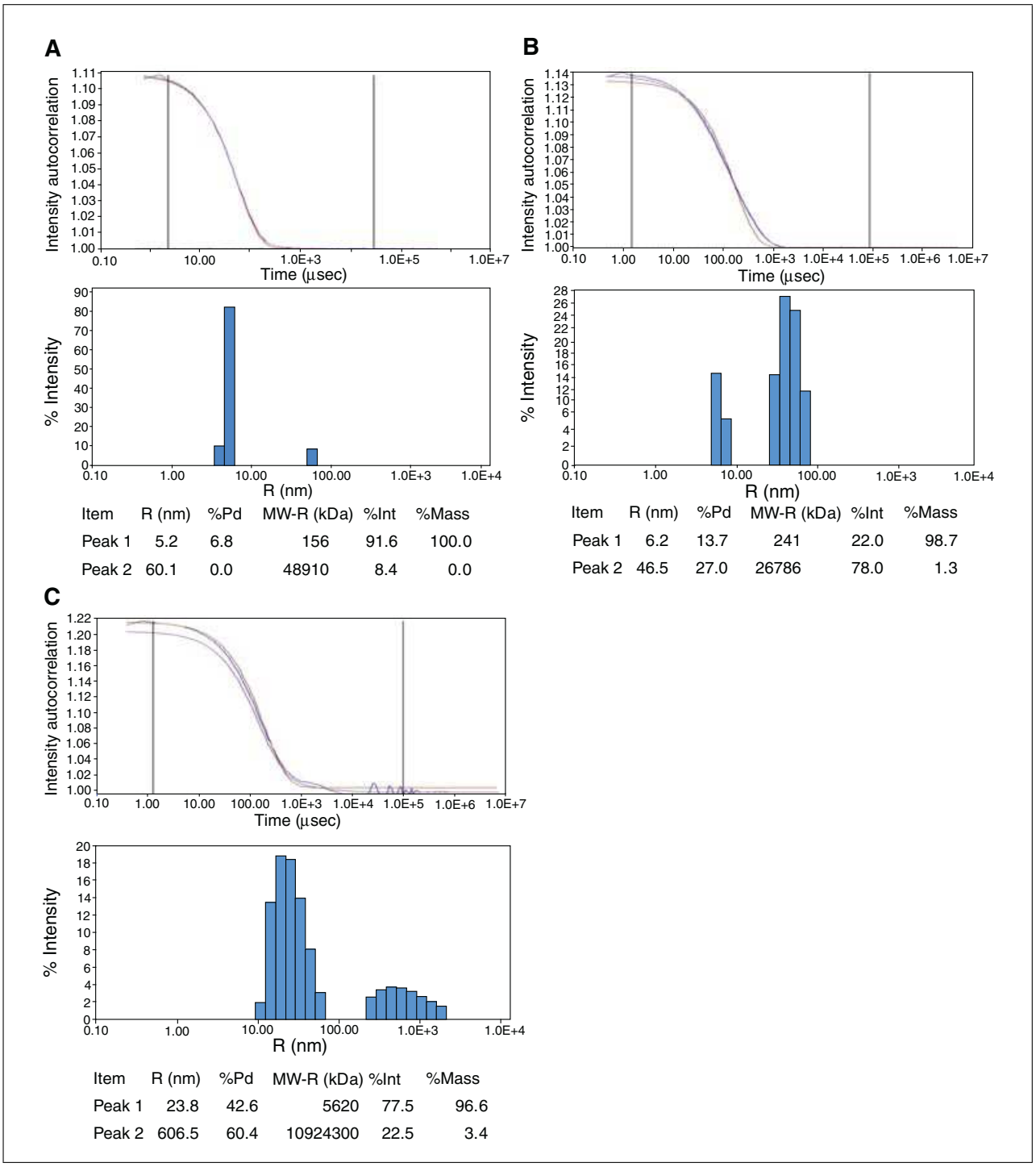

Figure 17.10.1 Correlation functions and the corresponding calculated histograms. Top row: experimental correlation function and the best calculated fit; bottom row: corresponding histograms of particle sizes. (**A**) Monomodal distribution; (**B**) Small percentage of aggregates; (**C**) Large aggregates.

Several publications are available that review DLS procedures as they apply to protein crystallization (Bergfors, 1999; Borgstahl 2007).

Acquiring DLS Data

For DLS measurements, the authors use the DynaPro Plate Reader manufactured by Wyatt Technology Corporation (*http://www.wyatt.com*). This instrument measures light scattered at a single angle, sufficient for assessment of protein behavior in solution. A more detailed, quantitative analysis requires a more versatile instrument that is capable of measuring scattered light at a variety of angles. Data acquisition and analysis for the DynaPro instrument is performed by the Dynamics software (also a product of Wyatt Technology Corp.), which is used to derive pertinent data about the protein solution, including the hydrodynamic radius (r_h), % of polydispersity, the estimated molecular weight, the relative amount of light scattered by each population (% of intensity), and the estimated relative amount of mass of each peak or species (% of mass). The distribution of molecular weights is derived from the experimental correlation function. If the profile generated by the software shows only one peak, the distribution is called monomodal, as opposed to multimodal for a distribution including two or more peaks. Examples of monomodal and multimodal distributions are shown in Figure 17.10.1.

Interpretation of DLS Results

A histogram showing particle size distribution derived from DLS data allows for a quick, qualitative assessment of the protein solution. The appearance of a single peak is an indication of a monomodal distribution of particle sizes. The width of the peak represents the broadness of the size distribution and is expressed as the % polydispersity defined as the ratio of the width of the distribution (measured at half maximal intensity) divided by the weighted average dynamic radius and expressed as percent. Low polydispersity values (<10% to 15%) indicate a homogeneous solution, while higher values (>25%) reflect a polydisperse solution, whose properties should be improved prior to crystallization screening. If the protein solution contains two species, they can be resolved into two separate peaks only when their sizes differ by a factor of 5 or more.

BASIC PROTOCOL 1

DLS MEASUREMENT OF PROTEIN SAMPLES

The first experiment is to determine the behavior of the protein sample obtained from the purification process. The DynaPro plate reader works with either 96- or 384-well plates and allows for consecutive measurement of many different proteins or the same protein under many different conditions. In each run, a range of wells to be measured is specified.

Materials

Protein solution at appropriate concentration (see annotation to step 1)
0.2- to 0.45 μm nitrocellulose filter disc (Whatman)
Refrigerated centrifuge
96- or 384-well flat-bottom plastic plate (Thermo Scientific, part no. 95040000)
DynaPro plate reader with Dynamics software (Wyatt Technology Corporation; *http://www.wyatt.com*)

1. *Recommended:* Filter the protein sample through a 0.2- to 0.45-μm nitrocellulose filter disc.

   The minimal concentration of the protein solution depends on the molecular mass of the protein. For small proteins (e.g., lysozyme, 14 kDa), the minimal concentration to obtain a measurable signal on the DynaPro instrument is ∼1 mg/ml; for a protein with a mol. wt. of ∼100 kDa, the minimal concentration is ∼0.2 mg/ml

2. Centrifuge protein solution 10 min at ∼18,000 × *g*, 4°C, to remove large aggregates.

3. Carefully deliver ~40 µl of protein solution to each well of a 96- or 384-well flat-bottom plastic plate. Take special care not to introduce air bubbles, which can be spotted by visual inspection.
4. Start and initialize the DynaPro instrument using a well that contains only deionized water or buffer. Make sure that the baseline is stable.

   The experiment is usually performed at room temperature.
5. Collect data until the counts per well are 800,000 to 1,000,000, which typically provides an acceptable ratio of signal to noise. Collect at least 10 measurements for each well.
6. When data collection is finished, perform data analysis using the Dynamics software.

   If the data analysis shows a single narrow peak with a molecular weight expected for a monomer or an oligomer and low % polydispersity, either proceed directly to crystallization or re-measure DLS at a protein concentration that will be used for crystallization (e.g., ~10 mg/ml).

   If the protein solution is polydisperse, proceed to the optimization step (below). If DLS indicates the existence of large aggregates in addition to smaller particles, consider performing size-exclusion chromatography (UNIT 8.3) using a column such as High-Load Superdex-75, Superose-12, or Superdex-200 (GE Healthcare) before continuing with the optimization process.

DLS TO OPTIMIZE PROTEIN SOLUTION BEHAVIOR FOR CRYSTALLIZATION

BASIC PROTOCOL 2

For a sample with a DLS profile showing a monodisperse distribution with low polydispersity, no further action is required. When the size distribution contains two peaks with low polydispersity, the sizes of the two types of particles are sufficiently different that they should be easily separated by SEC. Individual peak fractions from SEC could then be used for crystallization screening.

When the protein solution shows high polydispersity or contains peaks with a large r_h that correspond to aggregates, further optimization is necessary. Three aspects should be considered when attempting the optimization. First, it is important to consider what is known about the protein's properties, e.g., its function and what ligands, substrates, or cofactors it binds (Table 17.10.1). The presence of such compounds often has a stabilizing effect on the protein. The DLS experiment would then be performed in the presence of individual chemical additives or using a cocktail of several of them.

Table 17.10.1 List of Common Cofactors[a]

Common cofactors	Concentration
ATP, GTP	1-5 mM
ADP, GDP	1-5 mM
CoA	1-5 mM
NAD(P)	1-5 mM
FAD	1-5 mM
PLP	1-5 mM
SAM/SAH	1-5 mM
PAP	1-5 mM

[a]Substrates, products, or their analogs are specific for each particular protein and are not listed here.

Table 17.10.2 Buffers to Test Behavior of a Protein Sample at Different pH Values

Buffers (all 100 mM)	pH
Sodium citrate	4.5, 5.0, 6.0
Sodium acetate	4,5, 5.0
Sodium cacodylate	5.5, 6.5
Sodium phosphate	6.0, 6.5, 7.0, 7.5, 8.0
ADA	6.0, 7.0
Imidazole	6.5
Tricine	6.5, 7.5
PIPES	6.5, 7.5
HEPES	7.5, 8.0
Tris·Cl	7.5, 8.5
Bis-Tris propane	6.5, 7.5, 8.5, 9.5
TAPS	9.0

Secondly, it is known that pH is one of the most important parameters affecting protein solubility. Therefore, the next round of optimization should include testing different buffers and pH values ranging from ~5 to ~9 (Table 17.10.2). It should be kept in mind that different buffers with the same pH may affect protein behavior differently.

These tests can be performed in two ways. A protein sample can be first dialyzed against different buffers using a microdialysis system with only 50 µl of the protein sample (1 to 2 mg/ml) in different buffers at a concentration of 100 mM at 4°C overnight. After centrifugation of the dialyzed sample, a DLS experiment (Basic Protocol 1) is performed. The other option is dialysis (*APPENDIX 3B*) of a concentrated protein sample into ~10 mM buffer, pH ~7 to 7.5, followed by 1:4 dilution into prepared test buffers at 100 mM concentration for DLS measurements. The buffer showing the best protein behavior is chosen for further studies. If the protein solution still does not present a satisfying DLS profile, further optimization through the use of additives is pursued.

The additives that can influence protein behavior include different ions, glycerol, zwitterions, and detergents. A list of possible additives to test is given in Table 17.10.3. The first group includes the most common additives that should be considered first for testing. The second group includes compounds that would provide more detailed and systematic information about protein behavior.

Materials

Protein sample, 1 to 2 mg/ml
Concentrated solutions of selected buffers and additives (Table 17.10.3), if possible, 10-fold more concentrated than required for the experiments

Protein concentrator (e.g., Amicon centrifugal filter unit Ultra-4)
96- or 384-well microtiter plate

Additional reagents and equipment for measuring DLS (Basic Protocol 1)

1. Using a protein concentrator, concentrate protein to 4 to 6 mg/ml in low buffer concentration (~10 mM) and in low salt (if necessary for protein solubility).

Table 17.10.3 List of Additives and Suggested Concentration Ranges to Explore

Additive	Concentration
Group 1 additives[a]	
Sodium chloride	0.3-1 M
Glycerol	5-20% (w/v)
β-octylglucoside	5 mM
D-glucose	0.2-0.5 M
L-arginine	0.1-0.5 M
TCEP	5 mM
Group 2 additives[b]	
NDSB-201	0.1 M
Magnesium sulfate	0.1 M
Ammonium sulfate	0.1 M
Ammonium acetate	0.1 M
Ammonium phosphate	0.1 M
Ammonium cacodylate	0.1 M
Sodium sulfate	0.1 M
Lithium chloride	0.1 M
Calcium chloride	0.1 M
Magnesium chloride	0.1 M
Urea	0.5-1.5 M
Sucrose	0.25 M
EDTA	10 mM
Tween-20	50 μM
NaBr	0.1 M
Detergents	
Triton X-100	0.01% (v/v)
Dodecyl maltoside	0.1% (v/v)
CHAPS	2 mM

[a]Group 1 contains additives to try first.
[b]Group 2 contains additional possible choices to try.

2. Dilute the concentrated stock solutions of the selected buffers and additives (Table 17.10.3) to 1.25× the final concentrations. Place these 1.25× solutions (40 μl) in the wells of a microtiter plate.

3. Add 10 μl of protein solution (from step 1) to each well.

   The wells will now contain the desired final concentration of the buffer/additive and 1/5 the concentration of the protein.

4. Measure DLS as described in Basic Protocol 1.

5. Analyze DLS data and select conditions with improved protein behavior.

COMMENTARY

Background Information

As a rule of thumb, successful crystallization of a protein requires several conditions to be met: (1) high purity of protein sample preparation (>98%), (2) sufficient solubility allowing protein concentration of 6 to 10 mg/ml, (3) single protein species in solution, e.g., single oligomeric state, charge homogeneity, etc. Several methods can be used to assess the protein solution homogeneity, with DLS being very useful as the final test. The broadly observed correlation between the crystallizability and monodisperse behavior of the protein solution is a convincing reason for DLS gaining wider use in protein crystallography laboratories. This trend is even more pronounced with recent development of high-throughput instrumentation equipped with plate readers that make the measurements relatively simple as compared to the previous generation of instruments utilizing 12- to 15-μl cuvettes that suffered frequently from problems with air bubbles and required substantial amounts of time to complete one measurement.

DLS allows monitoring of protein aggregation that is difficult to follow by other methods and does so in near-real time. The possibility of performing measurements at temperatures between 4° and 40°C allows detection of temperature sensitivity of protein samples and selection of appropriate temperatures for crystallization. Finally, current instrumentation allows for rapid screening for the best conditions that support homogeneity of the protein preparation, that is, the presence of a single species: monomers, dimers, and higher-order multimers.

DLS has been successfully applied as a routine method for testing and optimizing protein samples for crystallization. Given its high sensitivity, the small amount of protein required, and its speed and convenience, a DLS experiment can easily be performed for each protein before setting up crystallization screening trials. The protocol described in this unit involves affinity purification of the expressed protein followed, if necessary, by an additional ion exchange chromatography step. The protein sample is then further purified by size-exclusion chromatography using either a Superdex-75 or Superdex-200 (or equivalent) column. The aggregates, if any, are separated from the main peak, and the time of elution indicates the protein oligomeric state. The sample from the main peak is then centrifuged or filtered to remove dust particles or any other type of nonspecific contaminants and subjected to DLS measurements. An important benefit of this approach is the full recovery of the protein sample, which can be reused for crystallization screening.

Critical Parameters and Troubleshooting

The protein concentration can be as low as 1 mg/ml, but it is recommended to check the protein behavior near the final concentration that will be used in crystallization screening.

If the number of counts in the DLS measurement fluctuates, check for air bubbles and filter solutions with a smaller-pore-size filter.

If there are too many counts, indicative of protein aggregation, lower the concentration of protein and try again.

If there are insufficient counts, concentrate the protein sample further.

Anticipated Results

Successful applications of DLS for protein solution crystallization can be found in the literature, for example the acyl coenzyme A:isopenicillin-N acyltransferase (Yoshida et al., 2005). This protein has been difficult to crystallize due mostly to aggregation after purification. A DLS analysis was performed with a combination of additives, and when 5 mM DTT, 250 mM NaCl, and 5 mM EDTA were added into the protein's buffer, the DLS behavior was improved, and the protein was crystallized and showed an excellent diffraction pattern. Other successful examples include *E. coli* MnmG (GidA), a protein involved in tRNA modification (Shi et al., 2009). Figure 17.10.1B shows the initial DLS when the protein was in solution containing 20 mM Tris·Cl, pH 8, 100 mM NaCl, 5 mM DTT, and 5% glycerol. Improved DLS was obtained by increasing the NaCl concentration to 0.8 M (Fig. 17.10.1A) and led to crystals diffracting to 2.4 Å resolution and structure determination (Shi et al., 2009). Another example is *E. coli* HypE, a protein involved in the maturation of the [NiFe] hydrogenase (Rangarajan et al., 2008). Although this protein could be successfully purified, its DLS profile showed high polydispersity. DLS optimization led to the finding that addition of 10 mM magnesium acetate to the buffer resulted in a monodisperse profile. HypE under these conditions was submitted to crystallization screening, which led to crystals that diffracted to

high resolution and subsequently to structure determination (Rangarajan et al., 2008). Another example is the Cj1293 protein from *Campylobacter jejuni*, which encodes a putative UDP-GlcNAc C6-dehydratase/C4-reductase (Goon et al., 2003). This protein could be expressed in high quantity, but the DLS after purification showed the presence of high-molecular-weight aggregates. The protein behavior was significantly improved by the addition of magnesium chloride; the presence of 5 mM $MgCl_2$ was sufficient to obtain monodisperse DLS profiles. Subsequently, crystals of this protein diffracting to 2.8 Å resolution were obtained, which allowed structure solution. Another example is an acyl-CoA transferase YdiF. This protein showed polydisperse behavior in every buffer condition tested and produced poorly diffracting crystals. The addition of 1 mM acetyl-CoA to the protein's buffer improved the DLS profile and led to crystals diffracting to 1.9 Å resolution and structure determination (Rangarajan et al., 2005).

Time Considerations

Step 1, protein concentration: 30 to 40 min per sample.

Step 2 and 3, place solutions in the plate: several minutes.

Step 4, DLS measurement: ~2 min per well.

Step 5, analysis: 1-2 min per well.

Literature Cited

Baldwin, E.T., Crumley, K.V., and Carter, C.W. 1986. Practical, rapid screening of protein crystallization conditions by dynamic light scattering. *Biophys. J.* 49:47-48.

Bergfors, T. 1999. Dynamic light scattering. *In* Protein Crystallization: Techniques, Strategies, and Tips. A Laboratory Manual. (T. Bergfors, ed.) pp. 29-38. International University Line, La Jolla, Calif.

Borgstahl, G.E. 2007. How to use dynamic light scattering to improve the likelihood of growing macromolecular crystals. *Methods Mol. Biol.* 363:109-29.

Claes, P., Dunford, M., Kennedy, A., and Vardy, P. 1992. An on-line dynamic light-scattering instrument for macromolecular characterization. *In* Laser Light Scattering in Biochemistry (S.E. Harding, D.B. Sattelle, and V.A. Bloomfield, eds.) pp. 66-76. Royal Society for Chemistry, Cambridge, U.K.

D'Arcy, A. 1994. Crystallizing proteins—a rational approach? *Acta Crystallogr. D Biol. Crystallogr.* 50:469-471.

Ferre-D'Amare, A.R. and Burley, S.K. 1994. Use of dynamic light scattering to assess crystallizability of macromolecules and macromolecular assemblies. *Structure* 2:357-359.

Goon, S., Kelly, J.F., Logan, S.M., Ewings, C.P., and Guerry, P. 2003. Pseudaminic acid, the major modification on Campylobacter flagellin, is synthesized via the Cj1293 gene. *Mol. Microbiol.* 50:659-671.

Niesen, F.H., Koch, A., Lenski, U., Harttig, U., Roske, Y., Heinemann, U., and Hofmann, K.P. 2008. An approach to quality management in structural biology: Biophysical selection of proteins for successful crystallization. *J. Struct. Biol.* 162:451-459.

Price, W.N., Chen, Y., Handelman, S.K., Neely, H., Manor, P., Karlin, R., Nair, R., Liu, J., Baran, M., Everett, J., Tong, S.N., Forouhar, F., Swaminathan, S.S., Acton, T., Xiao, R., Luft, J.R., Lauricella, A., DeTitta, G.T., Rost, B., Montelione, G.T., and Hunt, J.F. 2009. Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. *Nat. Biotechnol.* 27:51-57.

Rangarajan, E.S., Li, Y., Ajamian, E., Iannuzzi, P., Kernaghan, S.D., Fraser, M.E., Cygler, M., and Matte, A. 2005. Crystallographic trapping of the glutamyl-CoA thioester intermediate of family I CoA transferases. *J. Biol. Chem.* 280:42919-42928.

Rangarajan, E.S., Asinas, A., Proteau, A., Munger, C., Baardsnes, J., Iannuzzi, P., Matte, A., and Cygler, M. 2008. Structure of [NiFe] hydrogenase maturation protein HypE from *Escherichia coli* and its interaction with HypF. *J. Bacteriol.* 190:1447-1458.

Shi, R., Villarroya, V., Ruiz-Partida, R., Li, Y., Proteau, A., Prado, S., Moukadiri, I., Benitez-Paez, A., Lomas, R., Wagner, J., Matte, A., Velazquez-Campoy, A., Armengod, M.E., and Cygler, M. 2009. Structural flexibility and mutational analysis of *E. coli* MnmG (GidA), a highly-conserved tRNA-modifying enzyme. *J. Bacteriol.* 191:7614-7619.

Yoshida, H., Hensgens, C.M., van der Laan, J.M., Sutherland, J.D., Hart, D.J., and Dijkstra, B.W. 2005. An approach to prevent aggregation during the purification and crystallization of wild type acyl coenzyme A: Isopenicillin N acyltransferase from *Penicillium chrysogenum*. *Protein Expr. Purif.* 41:61-67.

Zulauf, M. and D'Arcy, A. 1992. Light scattering of proteins as a criterion for crystallization. *J. Crys. Growth* 122:102-106.