FIG. 1. Schematic of a coverslip containing three drops in an L-pattern. The L-pattern allows one to unambiguously identify the drops: (a) 1µl RNA (dark) plus 2 µl well (light), (b) 2 Al RNA plus 2 µl well, and (c) 2 µl RNA plus 1 µl well. The drops should be laid on the coverslip in an asymmetric pattern (e.g., an "L") to facilitate the identification of the drops (Fig. 1).

**Optimization**

In general, the factors important for protein crystallization are important for RNA crystallization, but there are additional factors that must be considered in RNA crystallization. Because most methods used to purify RNA denature the RNA, the annealing step is critical. The principal variables in the annealing are the temperature (40—70°C), incubation time at the highest temperature (1—10 minutes), cooling time (seconds to hours), buffer concentration and pH, and ionic composition (5—20 mM MgCl$_2$). Each annealing protocol could be subjected to the Sparse Matrix conditions. Alternatively, the annealing conditions can be optimized by monitoring the activity of the RNA or its chemical modification pattern as conditions are varied.

There are limitations to sampling the annealing condition. The buffers and ionic composition are partially determined by the Sparse Matrix. The annealing solutions (A.1—A.9) are usually made from desalted RNA stock solutions, but they can be made from solutions containing additional salts or polyamines. When considering the ionic composition of the solution, one should be mindful of the concentration of the phosphate groups on the RNA. Electroneutrality requires the phosphates to have a counterion (e.g., H$^+$, Na$^+$, K$^+$, Mg$^{2+}$), so even after desalting the RNA, there will be cations "bound" to neutralize the phosphates. These cations may be released when cations are added in subsequent steps. The crystallization conditions sometimes depend on the type of cations present, i.e., there can be a difference between Na and K$^+$. The temperature at which the hanging drops equilibrate is very important, so different crystallization trials should be conducted at 4°C, room temperature, and 37°C.

**Acknowledgments**

This work was funded by the Colorado RNA Center, W. M. Keck Foundation, National Science Foundation (MCB-9221307). We thank NExagen, Inc., and Ribozyme Pharmaceuticals, Inc., for chemically synthesized RNA and purification, Cheryl Grosshans for helping design the first successful RNA Sparse Matrix in this laboratory, Cindy Barnes for refining purification methods, Steve Schultz and Jodi Ryter for CIAP and HPLC advice. Mark Goldstein, Anne Gooding for use and feedback on the Sparse Matrix, and Jennifer Doudna for contributions throughout. We also thank Art Pardi, Olke Uhlenbeck, and Tom Cech for encouragement and helpful discussions.

---

**[10] Dynamic Light Scattering in Evaluating Crystallizability of Macromolecules**

**By ADRIAN R. FERRE-D’AMARE and STEPHEN K. BURLEY**

**Introduction**

Crystallization frequently represents the rate-limiting step in X-ray crystallographic studies of the structure and function of biological macromolecules. Successful crystallization trials rely on searching a large multiparameter space for supersaturation conditions, where homogeneous preparations will first nucleate and then support crystal growth. Advances in recombinant DNA technology have made substantial quantities of highly purified starting material available, and sparse matrix factorial strategies permit a directed approach to evaluating different parameters.
search of the enormous parameter space relevant for macromolecular crystallization. Microscale techniques and seeding strategies allow extensive crystallization trials even when sample supplies are limiting. Unfortunately, covalent purity (assayed commonly by various electrophoretic techniques or by mass spectrometry) coupled with exhaustive screening of crystallization conditions offers no guarantees. Empirical observations suggest that macromolecules that are monodisperse in undersaturated, "normal" solution conditions crystallize readily, whereas randomly aggregating or polydisperse systems rarely, if ever, yield crystals. Dynamic light scattering can be employed to screen crystallization candidates for monodispersity. Thereafter, crystallization efforts can be restricted to monodisperse preparations, eliminating valuable time lost on trials with unsuitable starting materials.

**Dynamic Light Scattering**

Dynamic light scattering (DLS), also known as photon correlation spectroscopy or quasi-elastic light scattering spectroscopy, is a technique for measuring the translational diffusion coefficient of a macromolecule undergoing Brownian motion in solution (reviewed in Ref. 8). Monochromatic light scattered by moving particles will display intensity fluctuations corresponding to particulate motion, and a decay analysis of the autocorrelation function of the light scattering signal can yield quantitative information about the solution behavior of the dissolved particles. If there is a single macromolecular species, the experiment provides a direct measurement of $D_r$ (translational diffusion coefficient). Although the shape and density of the solute are typically unknown, its hydrodynamic radius of gyration ($R_g$) can be calculated using the Stokes–Einstein equation, and its molecular weight ($M_2$) can be estimated from the measured value of $D_r$ using a calibration curve obtained from assorted globular proteins of known mass. More sophisticated analyses of the autocorrelation function can also provide useful data regarding sample polydispersity because the technique is exquisitely sensitive to aggregation (i.e., the intensity of scattered light is proportional to the square of the mass of the solute particle).

**Monodispersity and Crystallizability**

DLS has been used to develop models of crystal nucleation and growth and to monitor the formation of nuclei versus amorphous aggregates as a function of solvent conditions and protein concentration. Early light scattering studies of lysozyme by Feher and co-workers provided experimental support for a model in which crystallization is a process involving cooperative, stepwise addition of identical molecules to a growing ordered assembly, whereas precipitation results from nonspecific aggregation of molecules into a heterogeneous polymer. With modern DLS instrumentation and an approach using the second virial coefficient, Wilson distinguished solution conditions yielding crystals or amorphous aggregates and studied the thermodynamic properties of these two distinct processes. DLS studies of model proteins approaching supersaturation have also demonstrated that precipitants causing aggregation prior to saturation failed to produce crystals, while precipitants with which the protein remained monodisperse up to the point of nucleation gave crystals.

These observations suggest that macromolecules or macromolecular assemblies, which exist as monodisperse solutions in a single aggregation state (even under solution conditions far from supersaturation) are likely to crystallize, whereas macromolecules which aggregate randomly under these conditions are very unlikely to do so. Zulauf and D’Arcy employed DLS to analyze the aggregation state of 15 proteins, which had been subjected to extensive crystallization trials, and found that in each case that the protein existed as a monodisperse solution, conditions could be found for it to crystallize. Conversely, all proteins that existed as mixtures of oligomerization states failed to crystallize. We found a similar situation during our efforts to crystallize helix–loop–helix (HLH) transcription factor–DNA complexes.

We used DLS to screen a number of different recombinant protein expression constructs and found that protein–DNA complexes that were monodisperse under dilute conditions crystallized, whereas protein–DNA complexes that were polydisperse failed to crystallize.

Subsequently, D’Arcy has summarized the results of an even larger survey carried out at F. Hoffmann-La Roche Ltd. in Basel, Switzerland. In all, 66 proteins were examined by DLS under dilute conditions without crystallization precipitants. Of the 44 proteins that showed narrow, unimodal distributions of apparent molecular weight, 34, or 77%, yielded crystals. Ten proteins had broad unimodal distributions, and six of these, or 60%, crystallized. Twelve proteins had multimodal (polydisperse) distribu-

---


---

tions, and only one of these, or 8%, crystallized. Additional published reports of the correlation between monodispersity and crystallizability include the TNF-/3/TNF receptor complexes and α-amylase.

Other commonly available techniques that can detect polydispersity in dilute macromolecular samples far from supersaturation include analytical ultracentrifugation (reviewed in Ref. 17) and gel-filtration chromatography. Neither of these methods is as convenient nor as sensitive as DLS. Gel-filtration chromatography is readily available and easily identifies pathologically aggregated samples. However, it cannot distinguish samples with narrow unimodal distributions of apparent molecular weight (34/44 of which crystallized in the D’Arcy trial) from those with broad unimodal distributions (6/10 of which crystallized in the D’Arcy trial). In our experience, gel filtration chromatography, perhaps because of sieving involved in the separation, invariably results in an overestimate of monodispersity. The exquisite sensitivity of light scattering techniques to larger particles makes them particularly suitable for detecting even modest amounts of aggregation, which appear to interfere with crystallization. George and Wilson have proposed measurements of the second virial coefficient of candidate macromolecules by static light scattering as another predictor of crystallizability.

Crystallization Strategy

In our laboratory, we employ DLS routinely to screen macromolecules and macromolecular assemblies after purification and functional characterization, but before any crystallization attempt. Typically 1 mg of the molecule of interest is expressed and purified to near-homogeneity (>95%), and then tested for biological or biochemical activity. Having passed this stringent test, the remainder of the purified material is subject to DLS measurements in dilute solution without precipitants. Although we work at sample concentrations of only one-tenth of those typically employed for crystallization, absence of random aggregates detectable by light scattering under these conditions appears to be an excellent predictor of monodispersity en route to crystallization at higher concentration with precipitants.

If the light scattering results indicate that the preparation is monodisperse, then preliminary crystallization trials, as well as further expression and purification, are carried out. If the preparation is polydisperse, DLS

is performed on the same recycled sample under a variety of solvent conditions. Commonly varied parameters include pH, ionic strength, redox potential, temperature, presence or absence of ligands, inhibitors, metal ions, cofactors, posttranslational modifications, and limited proteolysis. With monoclonal antibodies’ or receptors in hand, the relevant complexes can also be evaluated for monodispersity.

When experimental conditions under which the sample is monodisperse cannot be identified, a different construct of the molecule of interest is prepared or the molecule is purified from a different source, and the DLS screening procedure is repeated until a preparation is found that behaves as a monodisperse entity in dilute solution. Only then are crystallization trials carried out. The DLS screening vastly improves the chances of obtaining crystals of the macromolecule or macromolecular fragment or assembly of interest. Optimization of these initial crystals can then be carried out by use of standard screening and seeding methods in conjunction with quantitative analytical tools.”

Experimental Procedures

Instrumentation

A DLS instrument can be constructed from common electronic and optical components. However, operation of such an instrument can be quite technically demanding and time consuming. The commercial availability of compact, microprocessor controlled DLS instruments has permitted routine use of this technology in laboratories undertaking macromolecular crystallization. For example, the model dp-801 DLS instrument from Protein Solutions, Inc. (Charlottesville, VA) employs a 25 mW, 780 nm solid-state laser, a -7 µl quartz flow cell, and an avalanche photodiode, detecting photons scattered at a fixed scattering angle of 90°. Apparent translational diffusion coefficients, molecular masses, hydrodynamic radii of gyration, and degree of sample polydispersity are calculated by the instrument from the autocorrelation function using the manufacturer’s software. This software calculates the diffusion coefficient from the decay of the autocorrelation function by performing a nonlinear least-squares fit of the autocorrelation coefficients to an exponential decay. The equivalent hydrodynamic


radius of gyration of a hard sphere is computed using the Stokes—Einstein equation. Sample polydispersity is expressed as the standard deviation of the distribution of apparent hydrodynamic radii computed for one set of measurements on a given sample. The apparent molecular mass is derived from a standard curve of molecular weights versus measured values of \(D_T\) obtained from a set of known globular proteins, supplied by the manufacturer.

**Sample Preparation**

The minimum sample concentration required to perform a DLS measurement depends on the instrument sensitivity and on the particle size of the sample. On the dp-801, a globular protein sample with a molecular mass of 15 kDa requires a minimum concentration of -1.5 mg/ml, and one of 100 kDa, ---0.2 mg/ml. The volume of sample required is a function of the fluidics or "plumbing" of the instrument rather than the sample cell volume. On the dp-801 instrument, a volume of about 50—100 Al is required. The first DLS measurements on a sample are typically performed in the storage buffer. Solution conditions are only limited by the chemical stability of the fluidics. The sample is injected into the instrument with a Hamilton syringe, through a disposable 200 A pore filter (Anotop-10, Whatman, Clifton, NJ) to remove dust particles. The commercial availability of these filters is another technological development that has made DLS practical as a screening technique. Massively aggregated samples will not even traverse this filter, thereby demonstrating their unsuitability for crystallization without recourse to DLS. (Incidentally, these filters are also very useful for removing dust particles from macromolecular preparations just prior to crystallization, if excessive nucleation is found to be a problem.) After the light scattering measurements have been made, the sample can be recovered from the instrument, concentrated if necessary, and used for further DLS experiments or even crystallization trials.

**Interpretation of Measurements**

DLS is the technique of choice for determination of the translational diffusion coefficient for monodisperse samples. However, for assessing crystallizability, the actual value of \(D_T\) is of little consequence. It is sample polydispersity that matters! The dp-801 instrument outputs a table of \(D_T\), \(R_H\), \(M_c\), and polydispersity index (expressed as the standard deviation of \(R_H\) in units of nanometers) for successive measurements taken on the same sample every few seconds. Because the irradiated volume of the sample cell is very small, aggregates are not necessarily detected in every measurement.

Therefore, we typically collect a few dozen measurements on a sample (5 minutes unattended operation). Well-behaved samples with narrow unimodal distributions of apparent molecular weight have a polydispersity index that is typically less than 30% of the mean value of \(R_H^2\).

A more useful means of estimating polydispersity involves examining histograms showing the variation of \(D_T\) during repeated measurements from the same sample. Highly monodisperse preparations show little or no variation in \(D_T\) (and \(R_H\) and apparent \(M_c\)) over the course of a few dozen measurements. These data are the hallmark of a protein or macromolecular complex that is most likely to crystallize (77% success rate for 44 cases in the D'Arcy trial). Samples demonstrating a broad unimodal distribution of \(D_T\) or apparent \(M_c\) are less likely to crystallize (60% success rate for 10 cases in the D'Arcy trial), and highly polydisperse samples are extremely unlikely to crystallize (8% success rate for 12 cases in the D'Arcy trial).

**Example: DLS Screening of HLH Protein—DNA Complexes**

Upstream stimulatory factor (USF), a 34 kDa HLH protein originally purified from HeLa nuclear extract, was overexpressed in *Escherichia coli*, purified to apparent homogeneity, and shown to be fully active in DNA binding. Despite very considerable efforts, no crystals of either the USF apoprotein or USF—DNA complexes were obtained. Inspection of the aggregate state of the 34 kDa protein by DLS immediately revealed severe aggregation. The distribution of apparent molecular weight is very broad and irregular, and it extends to well beyond two million daltons. A C-terminal truncation of the protein, USF(1—260), did not improve the biochemical behavior of the protein. Elimination of the N-terminal activation domain did result in a dramatic reduction in aggregation (Fig. 1a). USF(196—310) contains an intact, fully active, DNA-binding domain plus some additional C-terminal residues. USF(196—310)—DNA complexes demonstrate a broad unimodal distribution of apparent molecular weight by DLS, with no evidence of high-order aggregation. A further C-terminal truncation to yield the minimal DNA-binding unit, USF(196-260), resulted in a construct that exhibited a narrow unimodal distribution of apparent molecular weight and readily crystallized on addition of precipitants (Fig. 1a). Optimization of the crystallization conditions by conventional approaches yielded diffraction-quality cocrystals, and a 2.9 Å resolution X-ray structure of the protein—DNA complex followed.
A similar construct-screening strategy was applied to Max, another member of the HLH family of transcription factors. The Max(1–113)–DNA complex exhibited a broad unimodal distribution of apparent molecular weight when inspected by DLS (data not shown). Although some crystalline precipitates were obtained during standard crystallization trials, single crystals were not forthcoming from this construct. Elimination of 10 N-terminal residues to yield Max(11–113) aggravated the problem, producing a highly aggregated protein (Fig. 1b). In contrast, deletion of 11 more residues produced a well-behaved protein [Max(22–113)], which was still fully active in DNA-binding assays. DLS studies of the Max(22–113)–DNA complex demonstrated that the preparation was monodisperse with a narrow, unimodal distribution of apparent molecular weight (Fig. 1b). When the material used for DLS was recovered from the sample cell, concentrated, and precipitants added, single crystals were obtained immediately, and a 2.9 Å resolution X-ray structure of the complex has been reported.

Examples of the experimental data output by the dp-801 software for four of the truncated proteins mentioned above are given in Table I, and the corresponding histograms of apparent molecular weight are given in Fig. 1. Samples with a narrow unimodal distribution of apparent molecular weight show little, if any, variation in $D_T$ and $M_r$, with small polydispersity estimates (<1.0 nm). Samples with a broad unimodal distribution of apparent molecular weight show sizeable variations in $D_T$ and $M_r$, with larger polydispersity estimates (>1.0 nm).

Table I: dp-801 DLS Output for Four HLH–DNA Complexes

<table>
<thead>
<tr>
<th>Sample</th>
<th>$I$ (10^-13 m^2 s^-1)</th>
<th>$R_H$ (nm)</th>
<th>Polydispersity (nm)</th>
<th>$M_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>USF(196–310) + DNA</td>
<td>631</td>
<td>3.7</td>
<td>1.094</td>
<td>72,000</td>
</tr>
<tr>
<td>USF(196–260) + DNA</td>
<td>962</td>
<td>2.5</td>
<td>&lt;0.5</td>
<td>26,000</td>
</tr>
<tr>
<td>Max(11–113) + DNA</td>
<td>652</td>
<td>3.4</td>
<td>1.355</td>
<td>59,000</td>
</tr>
<tr>
<td>Max(22–113) + DNA</td>
<td>687</td>
<td>3.3</td>
<td>0.981</td>
<td>54,000</td>
</tr>
</tbody>
</table>

The results of three successive measurements from each sample are shown for comparison.

---

Summary and Perspectives

The DLS experiment is quick (a few minutes), is nondestructive, and requires a minimum of purified material (typically <1 mg). Macromolecular samples can be assayed routinely for monodispersity as a function of solvent conditions, the presence of ligands, inhibitors, or cofactors, or following posttranslational modifications or partial proteolysis, etc., greatly enhancing, we believe, the likelihood of successful crystallization trials. The DLS experiment can be used as the screening step in a crystallization strategy that involves the use of protein engineering to generate a number of related recombinant constructs of a macromolecule or macromolecular complex. This approach may also find some application in the challenging arena of membrane protein crystallization, where it could be used to screen different recombinant constructs and monitor various detergent solubilization strategies. Sample monodispersity is also critical for other biophysical methods, including nuclear magnetic resonance spectroscopy and small-angle X-ray and neutron scattering. DLS should be a useful screening step for samples prepared for these solution-based techniques. Finally, DLS may represent an important method for quality control during the manufacture of recombinant proteins for therapeutic use.

[1 1] Two-Dimensional Protein Crystals in Aid of Three-Dimensional Protein Crystal Growth

By ALED M. EDWARDS, SETH A. DARST, SALLY A. HEMMING, FRANCISCO J. ASTURIAS, PETER R. DAVID, and ROGER D. KORNBERG

A general method of forming single-layer-thick, or two-dimensional (2-D), protein crystals has been devised, based on adsorption of proteins to lipid layers. Adsorbed proteins are constrained in 2-D but retain lateral mobility, due to the rapid lateral diffusion of lipids, which enables crystallization. Adsorption may be accomplished in two ways: proteins are bound specifically to ligands attached to the polar head groups of lipids (e.g., nucleotide—lipid, drug—lipid, Ni<sup>2+</sup>—lipid for binding hexahistidine-tagged proteins), which allows the proteins to adopt a unique orientation at the lipid/water interface; or proteins are bound nonspecifically through electrostatic interactions with a charged lipid layer, allowing rotational as well as translational diffusion of the bound molecules. In either case, a high affinity of the proteins for the lipid monolayer ensures a high concentration of proteins at the lipid/water interface (500—1000 mg/ml), and this high concentration drives the crystallization process.

The lipid layer crystallization approach has been successfully applied to dozens of proteins, including antibodies, enzymes, and polypeptide toxins, and the following benefits have emerged. First, the method is useful for proteins that have resisted attempts at the formation of 3-D crystals. One reason may be that 2-D crystals form rapidly, often within minutes, before denaturation of labile proteins can occur. Second, the formation of 2-D crystals requires only microgram quantities of material, much less than is usually needed for the growth of 3-D crystals. Third, in all cases tested, the formation of 2-D crystals occurred under a wide range of solution conditions, including physiologic conditions. This permits structural analysis and other applications of 2-D crystals to proteins in their native states, in particular conformational states, or in fragile complexes with other molecules. Fourth, the exposed face of a 2-D crystal may be exploited for interaction with additional molecules in solution. These additional molecules may be of various types, forming multiprotein complexes, or they may be identical to the proteins adsorbed to the lipid layer, resulting in epitaxial crystal growth.

The main purpose of 2-D protein crystallization to date has been for structure determination by electron microscopy. Electron diffraction is employed for most accurate measurement of structure factor amplitudes, while Fourier transformation of electron microscope images yields phase information. While 2-D crystals formed on lipid layers may diffract to atomic resolution, as in the case of streptavidin crystallized on layers of bioinylated lipid, structure determination beyond 5—10 Å resolution is limited by loss of image contrast, and X-ray diffraction remains the method of choice for revealing atomic detail.

The formation and analysis of 2-D crystals can facilitate structure determination by X-ray diffraction in a number of ways. First, because of the

---