Introduction to single crystal X-ray analysis XI. Crystallization of proteins —Crystallizability evaluation using BioSAXS—

Akihito Yamano*

1. Introduction

The principle problem in single crystal X-ray analysis is the so-called "phase problem"; however, in the current situation in which various measurement devices, synchrotron radiation facilities and sophisticated software have become available, the largest bottleneck in practice is the acquisition of single crystals, that is, crystallization of the substance to be analyzed. Crystallization is particularly difficult for biomacromolecules. For ribosome, the largest asymmetric single molecule analyzed to date, acquisition of crystals yielding low-resolution diffraction was first reported in 1980[1], but crystals diffracting to a practical resolution could not be obtained for more than two decades. One of the reasons that it took so long was that only a limited number of researchers had addressed this challenge, as ribosome crystallization had been initially considered an unfeasible task. This was because ribosome was believed to have multiple subtypes and was thus low in homogeneity. If the homogeneity could be evaluated before crystallization, more researchers might have attempted ribosome crystallization to achieve successful ribosome crystallization a few years earlier. This article describes small-angle X-ray scattering from solution-state protein samples and its relevance to evaluation of protein crystallizability.

2. Particle size distribution and crystallizability

Although approaches and procedures used for crystallization have evolved over the years, crystallization is still the most labor-intensive step essentially performed on a trial-and-error basis. In late 1980s, crystallization screening solutions became commercially available, which practically standardized the first step of crystallization. Nevertheless, the users needed to wait and see whether crystallization occurs after mixing these screening solutions with the protein solutions. In the 1990s, researchers utilized the relationship between particle size distribution of a protein solution prior to crystallization and its crystallizability. They discovered that monodispersity of the protein solution prior to crystallization was a prerequisite for successful crystallization, which enabled rough prediction of crystallizability before crystallization screening. Bringing non-monodisperse protein solutions to monodispersity was established as a guiding principle, which could save a lot of work. Because it provides a direction to manipulate protein solution to raise crystallizability, that is to obtain monodisperse solution when it is polydisperse.

3. Evaluation of crystallizability by dynamic light scattering

The method first widely used for evaluation of monodispersity was dynamic light scattering. Devices specialized for crystallizability evaluation became commercially available, which facilitated the use of dynamic light scattering in structural biology laboratories worldwide. Dynamic light scattering determines particle size distribution based on the rate of fluctuation of scattering intensity caused by the Brownian motion of particles[2]. Fluctuation of intensity is slower in solutions containing large, slow-moving particles, while it is faster in solutions containing small, fast-moving particles. Figure 1 shows typical polydisperse and monodisperse diagrams output by a dynamic light scattering device. It is nearly impossible to obtain crystals from a solution exhibiting a polydisperse distribution (Fig. 1(a)), but by removing the aggregates by filtration (Fig. 1(b)) for example, the solution acquires the potential to produce crystals. The most impressive application of dynamic light scattering to crystallization has been reported by Burley et al; they applied the technique to protein−DNA co-crystals to demonstrate a high correlation between crystallizability and particle size distribution. They reported that crystals could only be obtained from solutions with narrow particle size distribution in which DNAs were strongly associated with proteins[3].

Some dynamic light scattering devices, such as DynaPro (Wyatt technology), are capable of measuring static light scattering intensity as well. If the second virial coefficient (A_2) can be determined by static light scattering, it can be used to estimate interaction forces acting among protein molecules[4][5]. When $A_2 > 0$, repulsive forces act among protein molecules, preventing their direct interactions required for crystallization. Meanwhile, when $A_2 < 0$, strong attractive forces act between the protein molecules to cause aggregation. Crystallization occurs when $A_2 < 0$, that is, when weak attractive forces exist. A_2 depends on

* Application Laboratories, Rigaku Corporation.

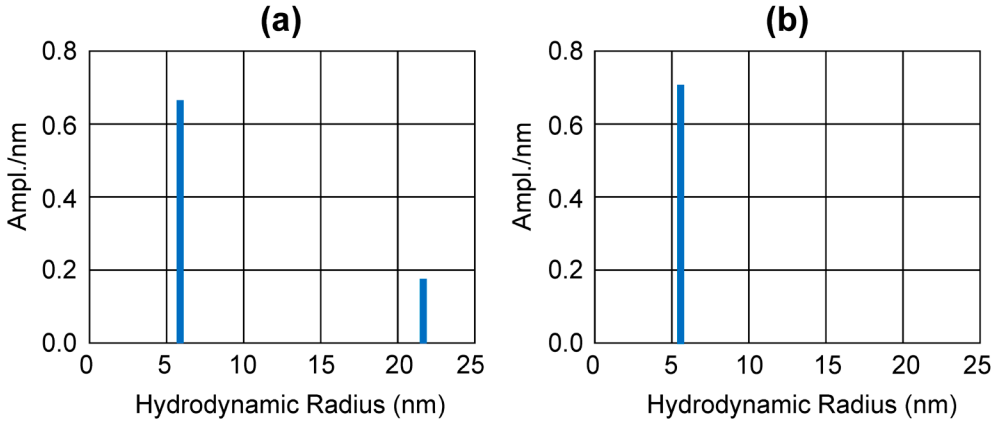

Fig. 1. Measurement of particle size distribution by dynamic light scattering.
(a) A typical polydisperse sample. Bimodal size distribution consisting of original protein particles and aggregates.
(b) A monodisperse sample obtained by filtering sample (a) to remove aggregates.

ionic strength, so it serves as an index of adequate salt concentration.

A_2 can be determined by measuring the intensity of light scattering from protein solutions of various concentrations and using the following equation:
in which K is the optical constant, C is the protein solution concentration, M is the average molecular weight, A_2 is the second virial coefficient, and R_θ is the Rayleigh ratio of scattered light to incident light.

4. Crystallizability evaluation by small-angle X-ray scattering

Since it has become relatively easy to generate constructs inducing mutation on residues that play important roles in crystallization of protein, and since crystallization robots have become popular, the frequency to use dynamic light scattering seems to be decreasing. Dynamic light scattering measurement will have less significance, unless it can offer additional advantages besides determining particle size distribution, such as providing information useful for structure determination or ease of measurement.

Guinier plots, which are based on the intensity data in the low-angle scattering region, can provide information on molecular interaction as well as on particle size and can be used in X-ray, neutron or static light scattering. Among various methods for measuring particle size distribution, small-angle X-ray scattering is relatively easy to access for structural biologists; there are some radiation facilities worldwide that can provide small-angle X-ray scattering measurements along with X-ray diffraction data. Moreover, most structural biology laboratories have diffractometers equipped with microfocus rotating anode X-ray sources for obtaining diffraction data or screening crystals, so it should be easy to additionally install small-angle X-ray scattering devices.

A Guinier plot is the logarithmic plot of diffraction intensity $I(q)$($q=4\pi \sin\theta/\lambda$, $Å^{-1}$) versus q^2 value in the low-angle region, also known as the Guinier region, shown in Fig. 2. The slope of the linear fit of the Guinier plot can be used to estimate the radius of gyration R_g,

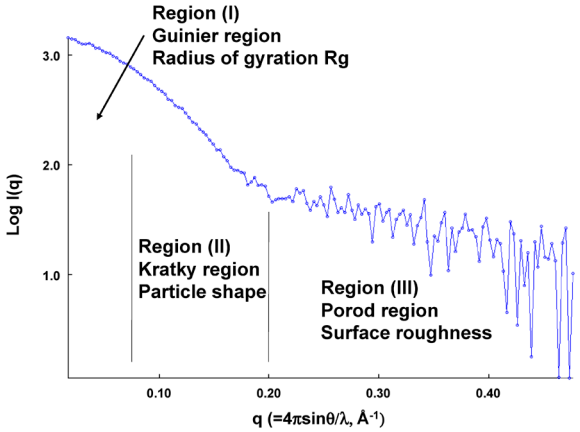

Fig. 2. Small-angle X-ray scattering curve.
Starting from the origin in increasing order of angle, the regions are named the Guinier region, Kratky region and Porod region.

which describes the substantial size of molecules in a solution.

The advantage of Guinier plotting is that it describes the association state of molecules as well as their size[6]. Among the schematic diagrams shown in Fig. 3, Fig. 3(a) represents a suitable state for crystallization. Fig. 3(b), in which the Guinier plot deviates upward from the straight line, suggests that the solution contains aggregated molecules and is thus not suitable for crystallization. Meanwhile, a downward deviation as shown in Fig. 3(c) suggests that repulsive force is acting between molecules. Such solution state is also considered unsuitable for crystallization as it lacks the intermolecular attractive force required for crystallization.

5. Other information obtained by small-angle X-ray scattering

In addition to the information on molecular interaction obtained by Guinier plotting, small-angle X-ray scattering mainly provides the following two types of information.

First, although the resolution is low, the external

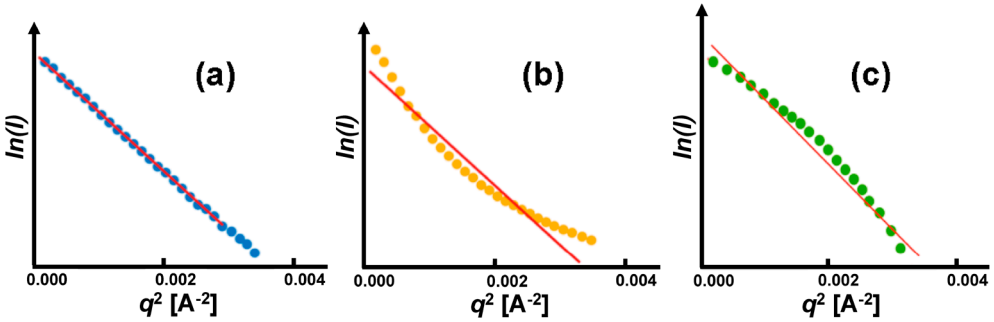

Fig. 3. Schematic diagrams of Guinier plot.
Fig. 3(a) demonstrates a state suited for crystallization, Fig. 3(b) demonstrates the occurrence of aggregation, and Fig. 3(c) demonstrates repulsive forces acting among protein molecules.

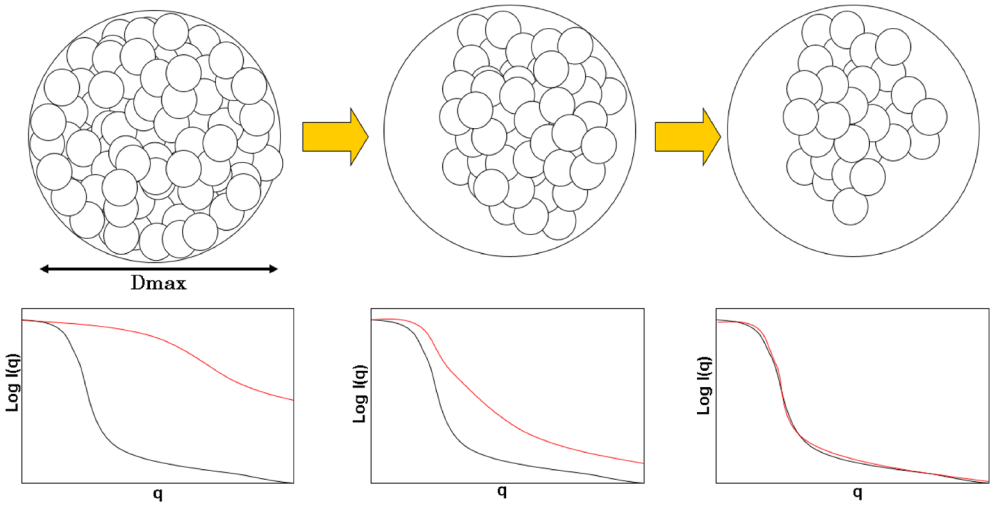

Fig. 4. Schematic diagrams describing the principle of the dummy atom model.
Various positions and numbers of small spheres packed in a sphere are shuffled to find an arrangement that best describes the scattering curve, which is assumed to reach the convergence, that is, the protein structure in solution.

form of proteins can be deduced with relatively high precision. For single crystal X-ray structure analysis, it is a prerequisite to prepare single crystals of the substance to be analyzed. However, human proteins such as target proteins for drug development are generally difficult to crystallize, and even if they could, the obtained crystals could not necessarily yield sufficient resolution for single crystal X-ray structure analysis. An advantage of small-angle X-ray diffraction is that it can be performed using protein solution and thus always gives an outcome. The use of a small angle X-ray scattering in protein structure analysis was rather unpopular due to its poor credibility; in recent years, however, an objective method for structure determination based on the dummy atom model has been put into practice, which has brought a remarkable breakthrough in this technique. In the dummy atom model, small balls relevant to amino acid residues are packed into a sphere having a diameter calculated by the distance distribution function, and positions and numbers of the small balls are changed to find an arrangement that best explains the X-ray scattering curve, which is assumed to represent the shape of the protein (Fig. 4). Although the model has the disadvantage of low resolution, it can reproduce *ab initio* characteristic structures for proteins or protein complexes of known structure, leaving no doubt in its credibility. It is assumed that there are many cases in which BioSAXS can be successfully applied to compound search in the initial stage of drug development. If a compound binds to a protein and induces a substantial change in the radius of gyration or external form of the protein, the potential binding of the compound may be evaluated by BioSAXS[7]. Moreover, it may provide some clues for protein conformation in a solution that cannot be observed in the crystalline state. Figure 5 shows the external form of a protein molecule determined by BioSAXS and the high-resolution ribbon model obtained by single crystal X-ray structure analysis.

The other information obtained by small-angle X-ray scattering is related to protein folding. The Kratky plot, obtained for the Kratky region in Fig. 6, is constructed by plotting $I(q)q^2$ against q (Fig. 6). If the plot has a peak within a relatively low q region, it demonstrates that the protein is properly folded; if the plot has no peak, the protein has a random coil conformation[8][9]. Since successful crystallization requires a protein solution dominated by properly folded protein molecules, the Kratky plot alone should be able

to provide valuable information for making the decision whether to proceed to crystallization.

6. Conclusion

BioSAXS-2000 (Rigaku Corporation) is a small-angle X-ray scattering device specifically designed for proteins, which can be mounted on the empty and opposite side of an existing diffractometer when the X-ray generator is a microfocus type (Fig. 7). As a BioSAXS-dedicated device, BioSAXS-2000 is equipped with the following characteristics:

(i) X-rays illuminate the same positions of the sample cell for both buffer and samples.
(ii) By combining an optical system having a focal point on the detector surface with a Kratky block, smearing issues can be avoided despite applying a line-shaped X-ray beam to samples for securing high radiation intensity.
(iii) A low-noise semiconductor detector is adopted to resist long exposure.
(iv) An extremely small volume (around 20 μl) of protein solution is required.
(v) A temperature control function can be installed.
(vi) It can be combined with a sample changer.

By generating Guinier plots and Kratky plots, BioSAXS provides information useful for crystallization as well as low-resolution protein structure in solutions. Although the time varies by protein size and concentration, generally, only a few hours of measurement is required to produce data sufficient for determining protein structure. Since protein structure in a solution often differs from that in a crystal, BioSAXS can provide findings that cannot be obtained by single crystal X-ray structure analysis alone. Moreover, it does not require a crystallized sample, so implementation of measurement always gives an outcome. The device is ideal to combine with single crystal X-ray structure analysis.

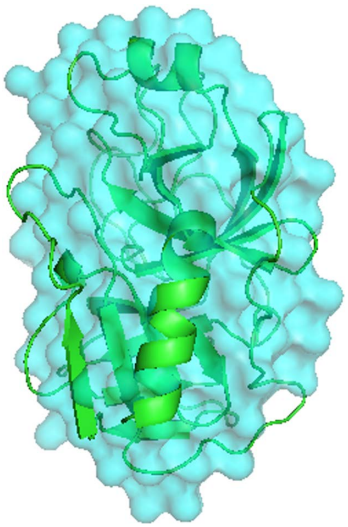

Fig. 5. An overlapping view of a low-resolution profile a protein (trypsin) molecule determined by the dummy atom model and a high-resolution ribbon model obtained by single crystal X-ray structure analysis.

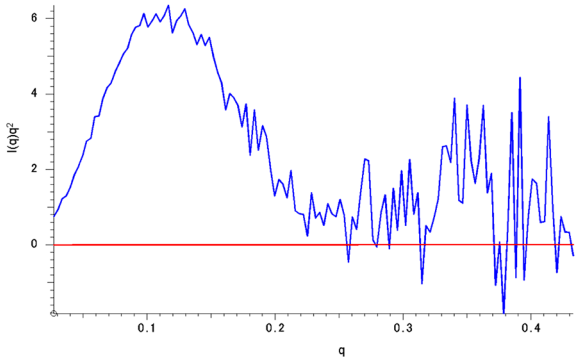

Fig. 6. A Kratky plot of a properly folded protein. There is a broad peak having a center around $q=0.12$. The peak height decreases as the protein approaches the random coil conformation.

Fig. 7. Typical installation of BioSAXS-2000.
It can be combined with a microfocus X-ray generator such as MicroMax007 and FR-X.