SAXSANA: an interactive program for the analysis and monitoring of static and time-resolved small-angle X-ray solution scattering measurements

Yuzuru Hiragi,[a]*† Yoh Sano[b] and Tomoharu Matsumoto[a]‡

[a]Institute for Chemical Research, Kyoto University, Gokasyo, Uji, Kyoto-fu 611-0011, Japan, and [b]Department of Pharmacology, Setsunan University, Togecho, Nagao, Hirakata, Osaka-fu 573-0011, Japan.
E-mail: hiragi@scl.kyoto-u.ac.jp, yhiragi@yahoo.co.jp

An interactive analytical program, *SAXSANA*, for small-angle X-ray scattering measurements of solutions is described. The program processes scattered data without disciplined knowledge of small-angle scattering. *SAXSANA* also assists in finding the best experimental conditions, thus avoiding blind runs of experiments. *SAXSANA* consists of the following procedures: (i) determination of the centre of scattered X-rays and moment transfer Q ($Q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle and λ is the wavelength) for each measured channel; (ii) conversion of the data format to the format of Q *versus* scattered intensities $J(Q)$; (iii) truncation of unnecessary data and smoothing of scattering curves by cubic-spline function; (iv) correction of the absorption effect and subtraction of the scattered intensity of the buffer (solvent) solution from that of the sample solution; (v) creation of a data file for a three-dimensional representation of time-resolved scattering curves; (vi) determination of radii of gyration by Guinier plots; (vii) determination of persistent lengths by Kratky plots; (viii) extrapolation of the small-angle part by Guinier plots; (ix) extrapolation of the wide-angle part by Porod's & Luzzati's laws for the Hankel transformation in order to obtain the distance distribution function $p(r)$; (x) calculation of $p(r)$ and computation of the invariant, the chord length, the volume, the spherical radius, the maximum dimension D_{max} and the radius of gyration (Rg). *SAXSANA* also serves as an on-site monitor for the validity of an experimental result during the measurements.

Keywords: small-angle X-ray scattering; data analysis; synchrotron X-rays; time-resolved measurements.

1. Introduction

During the past decade, X-ray beams produced by a synchrotron or a generator combined with an advanced detector system have changed the situation for measurements of small-angle X-ray scattering (SAXS) of materials in solution. A few hundred time-resolved data as well as static data in different conditions are now collectable within a day. This progress in experimental techniques has brought about great changes in the study of biology. The structural investigation of biological macromolecules, which requires the accumulation of a large amount of data, is nowadays possible to carry out without difficulty. Studies on dissociation/unfolding and refolding/reassociation processes of oligomeric proteins GroEL (Hiragi *et al.*, 2002) and GroES (Higurashi *et al.*, in preparation), by a denaturant or temperature, and kinetic studies of structural change of proteins (Segel *et al.*, 1999; Sakash *et al.*, 2000; Hiragi & Sano, unpublished) are some examples. However, the rapid processing of the numerous data obtained is still a problem. Although experimental and analytical procedures have been established (see, for example, Glatter & Kratky, 1982; Hashizume *et al.*, 1982; Ueki *et al.*, 1983, 1985; Hiragi *et al.*, 1988; Kajiwara & Hiragi, 1996), a convenient program for on-site data analysis has not been easily available. The interactive analytical program *SAXSANA* presented here offers a solution to this problem.

In addition to the data-analysis itself, a practical difficulty when using numerous specimens is how to check the quality of the measured data at the time of the experiment. The anxiety for an experimenter is, for example, contamination of the specimen with impurities during the process of purification, or damage of a protein by X-ray irradiation at the time of measurement. It is also necessary for experimenters to set a suitable measuring time to obtain data with a good signal-to-noise ratio while avoiding radiation damage owing to too long an irradiation by X-rays. The speed of the data analyses with *SAXSANA* assists in finding the best experimental conditions in the allotted beam time without prior blind runs of experiments.

Another characteristic of *SAXSANA* is its ease of manipulation. Users with little knowledge of small-angle scattering should be able to analyse their own data after 10 min of instruction. In order to carry out analyses it is only necessary to read the instructions on the command buttons and press a button for the required procedure or select a file in a text box. Each result processed is output as a file.

To read the experimental data, *SAXSANA* uses the format employed at beamlines BL-10C and BL-15A at the Photon Factory in Tsukuba, but data formats used at other synchrotron facilities are also available with one proviso: that they are prepared in the format of Q ($Q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle and λ is the wavelength) *versus* $J(Q)$ (scattered intensities), as shown in Fig. 1.

SAXSANA is the latest version of a program originally written in Microsoft Quick Basic (Y. Hiragi, unpublished) and rewritten in Microsoft Visual Basic to carry out interactive analyses.

```
249
  1     3.21463E-04     11717352.67
  2     1.92880E-03     13210188.46
  3     3.53614E-03     16393027.05
  4     5.14348E-03     23660039.04
  5     6.75082E-03     53235087.84
  6     8.35816E-03     144720572.1
  7     9.96550E-03     345352069.8
  8     1.15728E-02     574995282.1
  9     1.31802E-02     647242901.3
 10     1.47875E-02     561531593.2
 ....     ....     ....
 ....     ....     ....
 ....     ....     ....
```

Figure 1
Data structure (.DA1) converted in process (ii) from the experimental data at the Photon Factory (.DAT). The first line indicates the measured points. The three columns give the data point number, the Q values and the scattered intensities, respectively. For time-resolved measurements, the format is identical, but an initial datum having, for example, data containing 30 frames, is separated into 30 files numbered 001.DA1 to 030.DA1.

† Present address: Kansai Medical University, 18-89 Uyama-Higashi, Hirakata 573-1136, Japan.
‡ Present address: National Institute for Physiological Sciences, Meidaiji, Okazaki 444-8585, Japan.

2. Flow of the data analysis

As seen in Fig. 2, the main menu of *SAXSANA* consists of ten sub-menus:

(i) Determination of the Q values of scattered intensities using the diffraction peaks of collagen or another calibrator material.

(ii) Data conversion of the data formats of beamlines BL-10C and BL-15A at the Photon Factory to rows of Q and $J(Q)$, where the original .DAT file is converted into the .DA1 file shown in Fig. 1. As this process includes the addition of independent data, an experimenter can add supplementary data if the initial data are not well qualified owing to a low signal-to-noise ratio. Changes in the strength of the incident beam, by the attenuation of a ring current, are automatically corrected, and the scattered intensities are converted to counts per second.

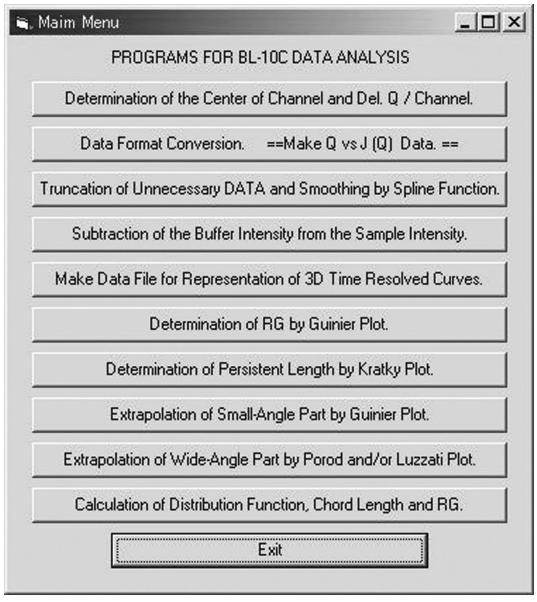

Figure 2
Main menu of *SAXSANA*.

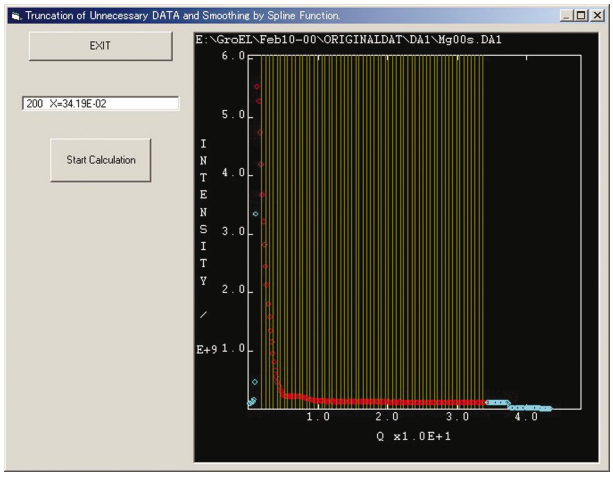

Figure 3
Truncation and smoothing of scattered data. Scattering data of the native GroEL tetradecamer is shown as an example. Vertical lines show the node position of the spline function.

(iii) Truncation of unnecessary data in the terminal low-Q area, which is hidden by the beam stop, and the high-Q area out of sensitivity of the detector (see Fig. 3). This procedure simultaneously produces both smoothed (in the .SM1 file) and native non-smoothed (.TR1) data. Unnecessary low- and wide-angle data are shown in blue; the degree of smoothing is selected by giving the number of nodes of the cubic-spline functions. Initially nodes are set at regular intervals and in multiples of 60. Users can arbitrarily set the number and position of nodes for more exact smoothing.

(iv) Subtraction of the scattered intensity of the buffer (solvent) solution from that of the sample solution (see Fig. 4). Corrections for absorption effects by high-atomic-number atoms contained in the solvent, *e.g.* chlorine, are performed at this stage using measured X-ray absorbance data.

(v) Creation of the data file for the presentation of three-dimensional time-resolved curves.

(vi) Determination of the (normal, cross sectional and thickness) radius of gyration (Rg) (see Fig. 5) by Guinier plots. The radius of gyration is determined by a linear least-squares fit. For beginners, an

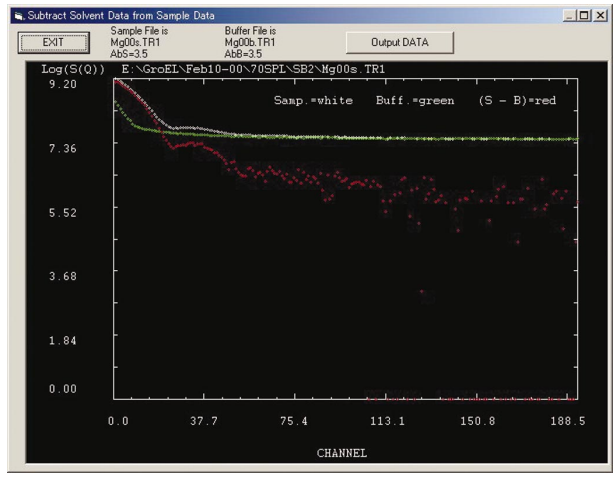

Figure 4
Subtraction of the buffer intensity from that of the sample (native GroEL) intensity.

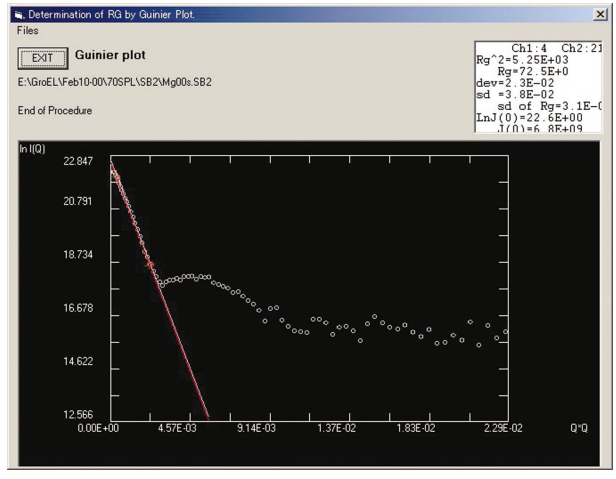

Figure 5
Guinier plot and determination of the radius of gyration (native GroEL). The white line indicates the Guinier line determined by a least-squares fit.

error value and an *Rg* value within the Guinier approximation are prepared in the text box as a clue for a reasonable *Rg* value. From our experiences by experiments and model calculations (Kajiwara & Hiragi, 1996), however, the mathematical requirement $Q \times Rg = 1$ for the Guinier approximation is not absolutely necessary. A Guinier line with a minimum error would be the correct solution. Some practical knowledge is naturally required to find the final solution.

(vii) Determination of the persistence length by a Kratky plot.

The following procedures [in sub-menus (viii) and (ix)] are prepared to obtain a better distance distribution function $p(r)$ profile.

(viii) Extrapolation of the small-angle part by the Guinier plot, in order to supply deficient data owing to the experimental restrictions.

(ix) If necessary for the calculation of $p(r)$, additional wide-angle-range data will be supplied by the method of Luzzati *et al.* (1961) based on Porod's law (Porod, 1951).

(x) Calculation of the distance distribution function $p(r)$ and the values estimable from $p(r)$ (see Fig. 6), and output of a file including the values of the invariant(s), chord length(s), volume(s), spherical radius, maximum dimension(s) (D_{max}) of the particle(s) and the radius of gyration.

In most cases, (viii) and (ix) are unnecessary if the scattered data are slightly smoothed and subtracted by (iii) and (iv) (.SB1 file). As seen in Fig. 6, where the Hankel transform in process (x) is simply performed by a Simpson integral, SAXS data obtained using synchrotron X-rays normally give satisfactory results without any further artificial manipulations. From our experience of analysing more than 500 scattered data patterns, complicated mathematical procedures are not necessary for the analysis of practical experimental data. Sophisticated computation may impede finding the origin of a problem, whether it comes from the experiment or the data treatment.

3. Example

Fig. 7 shows the results of a time-resolved stopped-flow SAXS measurement with bovine serum albumin (BSA) on cleavage of its disulphide bonds by dithiothreitol (Hiragi & Sano, unpublished). The figure, drawn using Microsoft Excel, shows the time dependence of the values calculated from the $p(r)$ functions: invariant, chord length, volume, spherical radius, maximum distance and square of *z*-average radius of gyration. In this case, the BSA solution during unfolding is a mixture of native and denatured proteins and consequently the radius of gyration is that of *z*-average, *Rgz*, and the other values are also of the average. The calculated results of the final stage of the analyses are output as a .VL0 file. As seen in Fig. 7, all the values increased, which suggests that the BSA molecule expanded. In particular, the change in invariant is the best indication of the detection of protein unfolding among the values shown in Fig. 7. A study of the denaturation process of GroEL by guanidine hydrochloride (Hiragi *et al.*, 2002) is another example of an application of *SAXSANA*.

In summary, *SAXSANA* has been practically tested and is used at BL-10C of the Photon Factory, Tottori University, the Laboratory of Biophysics, Nagaoka University of Technology, and others. At BL-10C of the Photon Factory, biochemists with little knowledge of

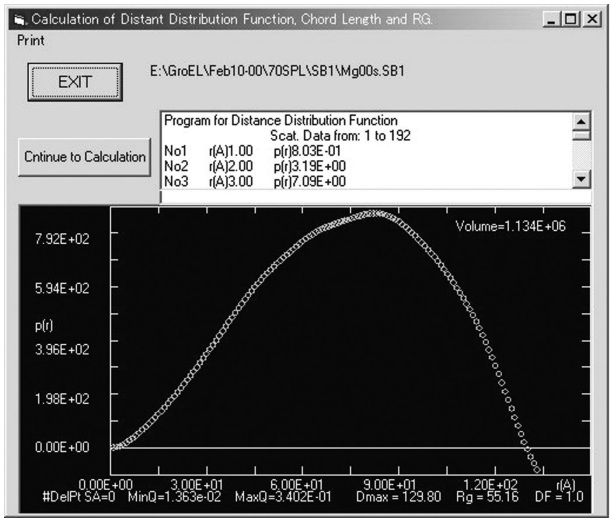

Figure 6
Display of a $p(r)$ function of GroEL.

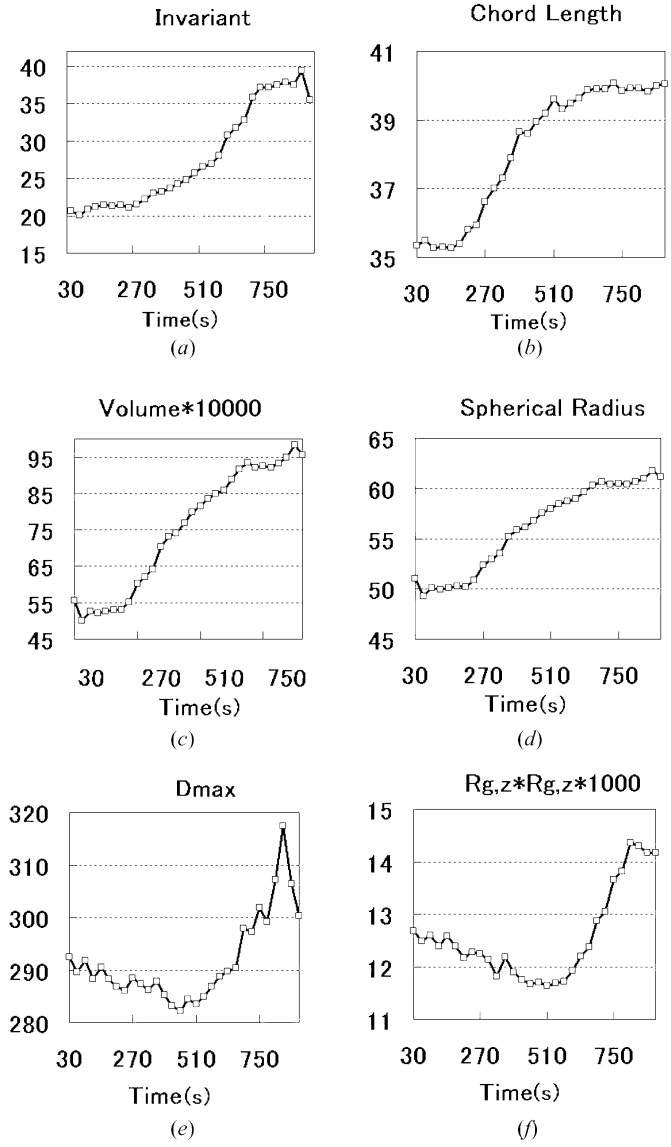

Figure 7
Time-resolved stopped-flow SAXS measurements; change in bovine serum albumin (BSA) structure during the cleavage of disulphide bonds by dithiothreitol. Plots of calculated values from $p(r)$ functions using Microsoft Excel: (*a*) invariant; (*b*) chord length (in Å); (*c*) volume (in Å^3); (*d*) spherical radius (in Å); (*e*) maximum distance; (*f*) square of *z*-average radius of gyration (in Å^2).

small-angle scattering can analyse their own data after only 10 min of instruction.

4. System requirements

SAXSANA is written in Microsoft Visual Basic, version 6.0, and runs under the Windows operating system. The program is available on request.

5. Notes for program users

If the original data, derived from either a one-dimensional or area detector, have a different format from that of the Photon Factory, they should be converted into the format shown in Fig. 1 before starting process (iii) (truncation of unnecessary data and smoothing by spline function). Automatic correction for the absorption of X-rays by a solution containing the absorber, *e.g.* guanidine hydrochloride, is not prepared in *SAXSANA* for each SAXS measuring system has a different monitor. A user having data in need of correction must have the absorption data obtained either by transmission measurements or by a monitor equipped at the measuring system (see, for example, Hiragi *et al.*, 2002).

This work was performed under the approval of the Photon Factory Advisory Committee (proposal No. 2000 G322 & 2001 G181).

References

Glatter, O. & Kratky, O. (1982). Editors. *Small-Angle X-ray Scattering.* London: Academic Press.

Hashizume, H., Wakabayashi, K., Amemiya, Y., Hamanaka, T., Wakabayashi, T., Matsushita, T., Ueki, T., Hiragi, Y., Izumi, Y. & Tagawa, H. (1982). KEK Internal Report 81–11. National Laboratory for High Energy Physics, Tsukuba, Japan.

Hiragi, Y., Inoue, H., Sano, Y., Kajiwara, K., Ueki, T., Kataoka, M., Tagawa, H., Izumi, Y., Muroga, Y. & Amemiya, Y. (1988). *J. Mol. Biol.* **204**, 129–140.

Hiragi, Y., Seki, Y., Ichimura, K. & Soda, K. (2002). *J. Appl. Cryst.* **35**, 1–7.

Kajiwara, K. & Hiragi, Y. (1996). *Applications of Synchrotron Radiation to Material Analysis*, edited by H. Saisyo & Y. Gohshi, pp. 353–404. Amsterdam: Elsevier.

Luzzati, V., Witz, J. & Nicolaieff, A. (1961). *J. Mol. Biol.* **3**, 367–378.

Porod, G. (1951). *Kolloid Z.* **124**, 83–114.

Sakash, J. B., Chan, R. S., Tsuruta, H. & Kantrowitz, E. R. (2000). *J. Biol. Chem.* **275**, 752–758.

Segel, D. J., Bachmann, A., Hofrichter, J., Hodgson, K. O., Doniach, S. & Kiefhaber, T. (1999). *J. Mol. Biol.* **288**, 489–499.

Ueki, T., Hiragi, Y., Izumi, Y., Tagawa, H., Kataoka, M., Muroga, Y., Matsushita, T. & Amemiya, Y. (1983). Photon Factory Activity Reports V7–V9, V29 and VI70–VI71. Photon Factory, Tsukuba, Japan.

Ueki, T., Hiragi, Y., Kataoka, M., Inoko, Y., Amemiya, Y., Izumi, Y., Tagawa, H. & Muroga, Y. (1985). *Biophys. Chem.* **23**, 115–124.