Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins

Elizabeth Mathew, Ahmed Mirza and Nick Menhart*

Illinois Institute of Technology, 3101 South Dearborn Street, Chicago, IL 60616, USA. E-mail: menhart@iit.edu

Small-angle X-ray scattering and size-exclusion chromatography have been combined within a unified experimental set-up to obtain molecular size information. Besides providing simultaneous corroborative data bearing on the same question from two distinct experimental techniques, passing the samples over a gel filtration column immediately prior to illumination by X-rays provides both a more homogeneous sample and a continuous set of data as the concentration is extrapolated to zero. This greatly facilitates analysis of data from oligomerizing or aggregating proteins and increases the reliability of the results.

Keywords: SAXS; liquid chromatography; Guinier; aggregation.

1. Introduction

We have combined two techniques for determining the size and aggregation status of proteins, gel filtration and small-angle X-ray scattering (SAXS), in order to provide more reliable data, especially with samples that exhibit non-ideal behavior due to association or aggregation. Gel filtration is perhaps the most widely used technique for determining the size of non-denatured proteins, owing to its ease and the inexpensive nature of the equipment (Barth *et al.*, 1994; Irvine, 1994). In this technique, molecules partition between the mobile phase and the chromatographic support depending on their ability to fit into nanoscale crevices in specially constructed solid phase compounds. Small molecules easily penetrate into the support, thus experiencing large partition coefficients and being strongly retained, whereas large molecules, which are too big to penetrate and are excluded, experience small partition coefficients and elute early, with the mobile phase.

However, it is difficult to construct supports with a strictly defined structure, thus there is no completely unambiguous way to directly relate an elution position (and so a partition coefficient) to an exact molecular size. This is not to say that the technique is not useful; there has been extensive work devoted to the manufacture of suitable supports and they are in wide use. However, the calibration of the size data relies on empirical comparison to molecules of known sizes, rather than directly from some physical principles.

In contrast, SAXS provides molecular size (and shape) information from fundamental physical principles, with no calibration needed (Chu & Hsiao, 2001; Doniach, 2001; Kratky & Pilz, 1972; Svergun & Koch, 2002). In this technique the angular dependence of Raleigh scattering provides information about the distribution of scattering centers. In principle, this dependence contains information about the spatial distribution of scattering centers (*i.e.* for X-rays, primarily the electrons) in the molecule. Interpretation of this information is complicated by spherical averaging of isotropic molecules in solution at modest angles, and complete analysis relies on intricate mathematical simulation and other techniques. However, at very low angles the Guinier approximation states that logarithm intensity decays with linear dependence on the square of the angle, the slope of which is the second moment of the distribution, *i.e.* the radius of gyration, *Rg*. This relationship relies only on mathematical relationships and the physics of scattering, and so is free from empirical uncertainty.

However, this technique is not as widely used owing to the complexity and expense of the experimental equipment, as well as limited sensitivity: typically, solutions of several tens of mg ml^{-1} have been needed to achieve sufficient scattered intensity. The advent of synchrotron sources and improved detector sensitivity has ameliorated this problem, but still, sample concentrations of >1 mg ml^{-1} are routinely used. Under these conditions, aggregation of the molecules may become a significant problem. Since the total scattered intensity is proportional not only to the concentration but also to the molecular mass, aggregates are disproportionately weighted in the data, and even relatively small amounts may seriously distort the data.

Typically, the presence of aggregates is detected by an upward curvature of the Guinier plot at very low angles, and that portion of the curve is disregarded. This discrimination is easiest with very large almost macroscopic aggregates, whose SAXS signals are concentrated at very low q; these types of aggregates are also the easiest to remove by filtration or centrifugation. However, if the sample experiences association on a smaller scale, and is polydisperse with size differences of less than an order of magnitude, then either the individual regions are incompletely resolved yielding only an average *Rg*, or the plots are curved with no definite slope evident and so are unusable. In fact, without prior knowledge of the proportions and sizes of the species present, it is virtually impossible to analyze such curves. Also, since the intensity is molecular mass weighted, this becomes a problem for SAXS data with smaller amounts of aggregates than would be a problem for mass-averaged techniques; for instance, techniques that are detected by UV absorption.

This is a problem with some proteins we have studied, for which we observed a smoothly changing slope despite our best efforts to remove aggregates by, for instance, filtration, centrifugation of further purification. In some cases we observed irreproducible results, which we reasoned might be due to the presence of varying fractional composition of oligomers in various samples. Traditionally, for reversible oligomerization this problem can be dealt with by obtaining data at several concentrations and extrapolating to a concentration of zero, at which oligomerization is prevented. However, on an individual basis, this procedure is at best time-consuming; furthermore, if the samples exhibit some sort of non-equilibrium or irreversible aggregation, this procedure will not work.

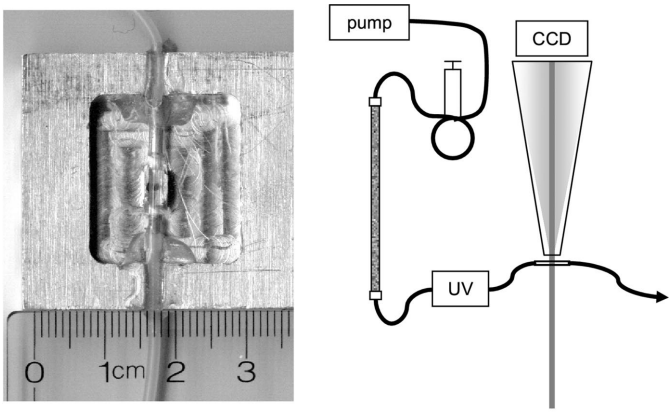
As such, we were inspired by recent developments in visible multi-angle light scattering of biological macromolecules (Wen *et al.*, 1996; Wyatt, 1991), which is now largely performed on flowing samples immediately after size exclusion chromatography. In this fashion, aggregates are removed on a continual basis immediately before measurement, and a continuous extrapolation to zero concentration can be performed as the chromatographic peak rises and then falls. We thus constructed a continuous-flow SAXS cell for use in conjunction with molecular sieve chromatography.

2. Experimental

2.1. Apparatus configuration

The experimental set-up consisted of an FPLC (Fast Protein Liquid Chromatography; Amersham Pharmacia Biotechnology, Piskataway, NJ, USA) instrument with a 0.5 cm $\times$ 10 cm Superpose-6 column, coupled to a home-made SAXS flow cell made of a machined block of aluminium and a 100 µm wall thickness 1 mm ID quartz tube, as shown in Fig. 1. Samples of $\sim$10 mg were injected into the column and eluted with a 5 m*M* Na phosphate, pH 7.4, plus 150 m*M* NaCl (PBS) buffer at a flow rate of 0.25 ml min^{-1}. The eluant was directed

  DOI: 10.1107/S0909049504014086



(*a*) (*b*)

Figure 1
The experimental apparatus consisted of an FPLC system coupled to a home-made SAXS cell made from a quartz capillary and a block of aluminium, (*a*). The configuration, shown in schematic in (*b*), was such that the sample was passed over a molecular sieve column and monitored by UV absorption immediately before it was passed through the X-ray beam. Scattered photons were then detected by a phosphor-coupled CCD camera placed behind an evacuated 2 m flight tube.

through a UV detector to monitor the chromatography, and then routed through the capillary in the SAXS flow cell, as shown in Fig. 1.

SAXS data were acquired at the BIOCAT beamline at Argonne National Laboratory with a two-dimensional CCD detector (Phillips *et al.*, 2002) after the scattered X-rays were passed through a nominally 2 m evacuated flight tube. We used the detector with 4×4 binned elements, pixels, so that the image pixels were 48 µm in size, and formed a 1798×1024 array covering the $8.6\ \text{cm} \times 4.9\ \text{cm}$ detector area. The distance from the detector to the sample was set to be about 2 m, and was measured with each set-up to 1 mm. This exact distance was confirmed by measurement of a powder pattern of a silver behenate sample, which also served to locate the exact center of the direct beam. A $\sim$6 mm beamstop was needed to shield the CCD from the brunt of the direct beam so, with the beam center located near a corner of the rectangular detector area, a range of $0.044\ \text{nm}^{-1} < q < 1.83\ \text{nm}^{-1}$ was sampled when 103 nm X-rays were used, where q is the wavelength-independent angle, $q = 4\pi \sin(\theta)/\lambda$. Images were acquired every 20 s with a nominal 2 s exposure as the chromatographic eluant passed through the SAXS cell. The exposure times varied, owing to control network latencies, but were measured to 0.001 s, and the data were normalized to account for this variation. Since the sample was continuously flowing at $0.25\ \text{ml}\ \text{min}^{-1}$ through a 1.0 mm capillary, any given portion of the sample took only <0.04 s to traverse the 200 µm-wide beam. This limited the radiation damage to the protein, and was within exposure limits previously determined to result in detectable radiation damage in the more sensitive shape-dependent high-q region (Fischetti *et al.*, 2003).

2.2. Samples

For purposes of testing this protocol, we ran and analysed samples consisting of:

(i) Cytochrome *c*: a common SAXS standard known to be resistant to aggregation and having a well characterized *Rg* of 1.3 nm. Cytochrome *c* was obtained from the Sigma Chemical Company (St Louis, MO, USA).

(ii) Plasminogen: MW 80 kDa, *Rg* reported to be variable between 3.0 and 5.0 nm, depending on conditions. We have experienced unstable and curved Guinier plots for this protein. Plasminogen was either obtained from Enzyme Research Laboratories (South Bend, IN, USA) or purified from human plasma (Castellino & Powell, 1981) from LifeSource (Chicago, IL, USA). Both preparations exhibited similar behaviors, and appeared as homogeneous single bands by SDS-PAGE.

(iii) A mixture of macromolecules: cytochrome *c*, bovine serum albumin (MW 65 kDa, $Rg = 3.0$ nm) and blue dextran (MW > 1 MDa; *Rg* unknown but presumably >10 nm). Bovine serum albumin and blue dextran were both obtained from the Sigma Chemical Company (St Louis, MO, USA).

2.3. Data analysis

The two-dimensional data images were radially integrated using *Fit2D* (Hammersley, 1997) to create a one-dimensional data set corresponding to the scattering profile at each image time. Direct beam intensity was monitored by a gas ionization detector inserted in-line before the sample chamber, and used to normalize the data and so compensate for beam fluctuations. Solvent and background scatter were assessed by averaging three to five data sets near the start and end of each run, before any proteins had eluted from the column and again after all the material had eluted (as confirmed in each case by the lack of any UV absorbance). This blank scattering profile was then subtracted from each individual data set to yield the excess scatter due to macromolecules only.

Guinier analysis was performed to determine an apparent *Rg*. Since the data still exhibited some curvature, it was important to objectively define an appropriate region. The Guinier approximation arises since the first two terms of the scattering profile, $I(q)$, of an arrangement of arbitrary scatterers, $\rho(\mathbf{r})$, match exactly with the power series expansion of $\exp[-(Rg\,q)^2/3]$. As such, in the limit of small q, the two functions approach each other, and *Rg* can be determined. Just how small q has to be depends on the rate at which the higher-order terms become important. For a spherical arrangement, the third term also matches exactly, and a limit of $Rg\,q_{\max} < 1.3$ (Guinier & Fournet, 1955) was determined, and is widely employed to this day (Svergun & Koch, 2002). Of course, without this third term, the approximation breaks down that much sooner, and for non-spherical particles a lower limit may be employed. However, since the function contains only even powers of q, this only becomes significant for particles of significantly large aspect ratios, and for globular molecules the value of 1.3 is widely used. As such, the analysis region was varied for each data set by increasing or decreasing the range in q so that $q_{\max} = Rg/1.3$, with *Rg* being determined by the local slope up to this $q_{\max}$. The lower bound of the region used to determine the slope was taken to be $2q_{\max}/3$. We also tested $q_{\max}/2$ and $q_{\max}/3$, with substantially the same results, except for a gradual loss of *Rg* resolving power in the mixed samples. We also tested a more stringent limit of $Rg\,q_{\max} < 1$ and also found no significant changes in the conclusions presented.

In addition, this Guinier line was extrapolated to $q = 0$ to yield an approximation of the aggregate-free forward scatter, $I'(0)$. To be considered reliable, an *Rg* value had to arise from a slope that was negative, on a trace with $I'(0) > 10$ events per pixel. The forward scatter is proportional to the product of the molecular mass and the concentration of excess scattering centers. As such, $I'(0)$ yields a trace that is comparable with the UV trace, except that the peaks for the larger proteins appear amplified and those for the smaller proteins appear attenuated.

3. Results

In some cases, the subtracted $I(q)$ trace, reflecting macromolecular scatter, still retained some non-ideal deflections at very low q. The positive deflections were undoubtedly at least partly the result of some small proportion of aggregate remaining in the sample. In some cases, negative deflections at very low q were observed in the subtracted traces. These may arise either as numerical artefacts owing to the inexact subtraction of the two large solvent scatter signals at very low q, subtle changes in water structure at high solute concentrations, or beam attenuation at higher concentration. In any case, such regions were not considered in our analysis owing to our utilization of a lower limit of $q_{max}/3$ for the Guinier analysis.

3.1. Cytochrome *c*

We found that the system was capable of yielding a stable and appropriate (Wu & Chen, 1988) *Rg* of 1.3 nm for cytochrome *c* throughout the peak region. This protein is a commonly employed SAXS standard, since it is highly soluble, monomeric at high concentration and has a well characterized *Rg*. Data obtained using this sample are shown in Fig. 2, in which $I'(0)$ and *Rg* are plotted as a function of elution volume. The UV signal rapidly saturated in this run, owing to its high extinction coefficient due to the heme moiety, and the relatively high concentration needed to achieve sufficient scattering owing to the modest molecular mass of this protein. However, the chromatographic elution can be monitored by $I'(0)$, which is proportional to the product of the number of scattering centers per molecule (nearly equivalent to the molecular mass) and the molecular concentration. For monodisperse samples, with uniform molecular mass, this is proportional to the mass of the sample in the beam. This monodispersity is confirmed by the *Rg* measurement, which remains fairly constant at ~1.3 nm across this peak, with an average value of $\langle Rg \rangle = 1.32 \pm 0.04$ nm obtained for the peak as a whole.

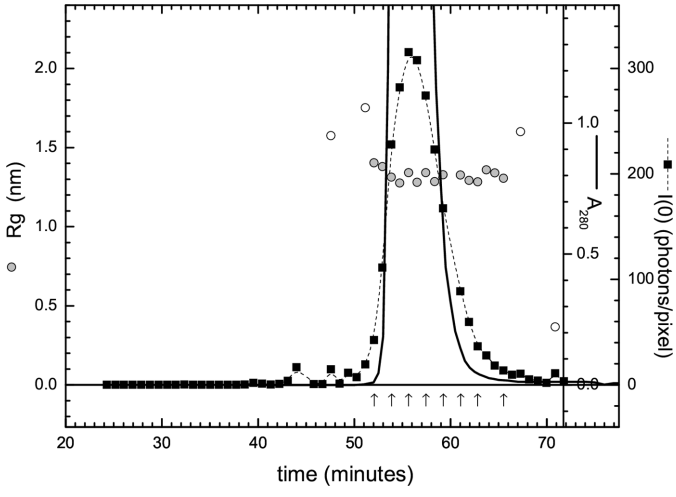


Figure 2
Chromatogram of the scattering as the protein, cytochrome *c*, eluted from the column. The forward scatter was extrapolated from the Guinier lines, and provides a measure of the total mass in the beam, weighted by its molecular size. We see a single major peak emerging at approximately 56 min. The radius of gyration of the eluted material is constant at 1.3 nm throughout this peak, when a cut-off of 10 photons per pixel is used as a reliability criterion. Small amounts of aggregate or other high-molecular-mass impurities are seen in the 43–50 min range; as expected, larger *Rg* values were yielded from these data sets, although the scattered intensity is so low that these results are not deemed especially reliable (the symbols are left open for these points). The arrows in the lower part of the figure indicate the regions from which the individual Guinier plots of Fig. 4 were obtained.

The ability of the system to clean up samples is hinted at by two features in this figure. First of all, the leading edge of the peak, ~52 min, shows an elevated *Rg*, and several spurious points in the region preceding the peak (45 to 50 min) exhibit signals of low $I'(0)$ and high *Rg*. This indicates that some high-molecular-weight impurities or aggregates were being removed from the sample. When we examine *Rg* as a function of $I'(0)$ in Fig. 3, we see the anomalous nature of the leading edge point [$I'(0) = 19.4$, $Rg = 1.75$, $t = 51.2$ min] much more clearly. However, as expected for cytochrome *c*, the protein exhibits little concentration dependence, and an extrapolated value of $Rg_0 = 1.34$ nm is obtained, consistent with the value averaged over all concentrations. To further illustrate the uniformity of this data, several selected Guinier plots (the locations of which are indicated by the arrows in Fig. 2) throughout this region are shown in Fig. 4; they are all clearly quite linear and quite parallel.

3.2. Plasminogen

Next, we tested the protein that we had previously found troublesome, plasminogen. In this case we observed quite different behavior. Again, a peak in $I'(0)$ corresponding to the bulk of the material was observed; however, substantial signal before this peak was also seen. In this experiment we were able to monitor the UV absorbance as well. We can see that although the main peak exhibits good concordance between $I'(0)$ and A_{280}, the leading material shows little UV absorption in proportion to its scattering intensity, and so is of higher molecular mass. This is confirmed by the much higher *Rg* observed for images in this region, as shown in the 20–30 min region of Fig. 5

The main peak, however, does show a fairly consistent *Rg* of ~3.3 nm, which is similar to previously reported values of 3.1 nm for this protein (Marshall *et al.*, 1994; Ponting *et al.*, 1992). When we examine the concentration dependence of *Rg*, Fig. 5, we see substantial non-ideality. This is the case whether A_{280} or $I'(0)$ is used as a measure of concentration. Both proxies yield similar extra-

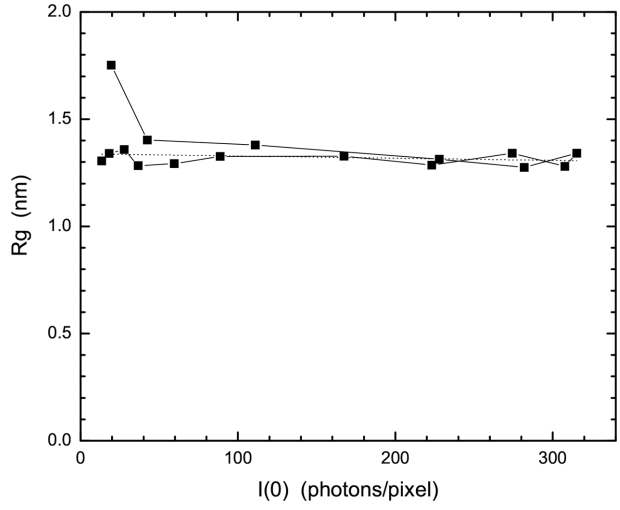


Figure 3
Rg extrapolation to zero concentration, using $I'(0)$ as a proxy of concentration, shows that cytochrome *c* does not display any non-ideal behaviour over the range tested (about 6 mg ml^{-1}), and exhibits a constant *Rg*, with a mean value of 1.32 ± 0.04 nm. One data point, in the upper left of the figure can be seen to exhibit an anomalously high *Rg*; this point is from the image at 52 min in Fig. 1, and probably contains a very small amount of impurity or aggregate.

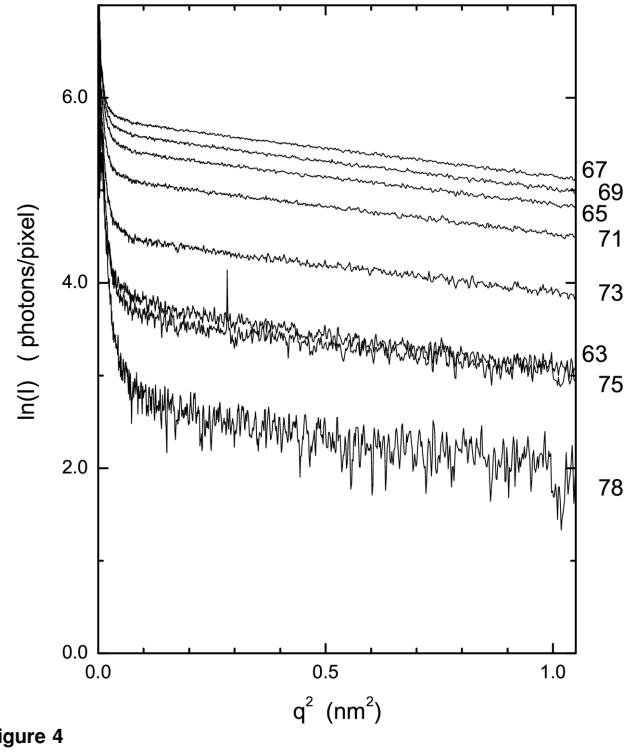


Figure 4
Guinier plots of individual images along the chromatogram shown in Fig. 2 (indicated by the arrows in that figure and ordered by the image numbers to the right of this figure) show good linearity, and are strongly parallel, indicating that one unique species is present. Note that the lowest-intensity data set spans intensities of <10 photons per pixel, and yet still yields a consistent and reliable *Rg*. This corresponds to concentrations <200 µg ml^{-1}.

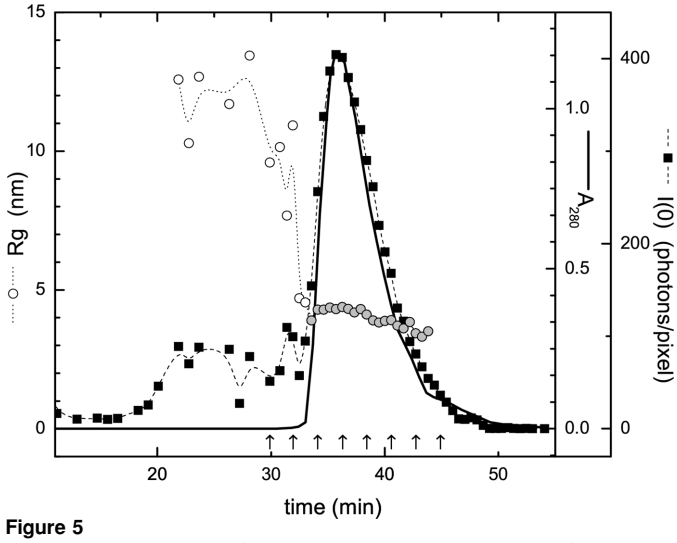


Figure 5
A chromatograph as plasminogen eluted from the column, showing not only scattering parameters *Rg* and $I'(0)$ but also the UV absorption at 280 nm as a measure of concentration. A major peak at 36 min was observed in both A_{280} and $I'(0)$, but substantial high-molecular-mass species were observed in the $I'(0)$ trace in the region immediately preceding this, 20–36 min. Since $I'(0)$ is proportional to both molecular mass and concentration, whereas A_{280} is proportional to only concentration, the differences in these two traces may be ascribed to differences in molecular mass. This is confirmed in the *Rg* data, which shows somewhat of a plateau at $\sim$3.5 nm in the major peak region, but much higher and more irregular values in this putative oligomer region. The filled symbols for the *Rg* trace correspond to the major peak, as defined by the region where the $A_{280}/I'(0)$ ratio remains constant, while the open symbol correspond to the putative aggregate region. The points are similarly delimited in Fig. 7.

polated values of 3.4 nm at zero concentration, with $Rg_{\rm apparent}$ values almost 25% higher, approximately 4.2 nm, at even the modest concentration of A_{280} $\sim$1 mg ml^{-1} ($\sim$1 mg ml^{-1}).

These changes in *Rg* are clearly seen in the individual Guinier plots shown in Fig. 6 (from positions indicated by the arrows in Fig. 5), which show non-parallel slopes. These plots are also not as strictly linear over as wide a region as was seen for the Guinier plots of cytochrome *c*. Many data sets show the slight upward curvature that is indicative of polydisperse samples. This could be due to incomplete

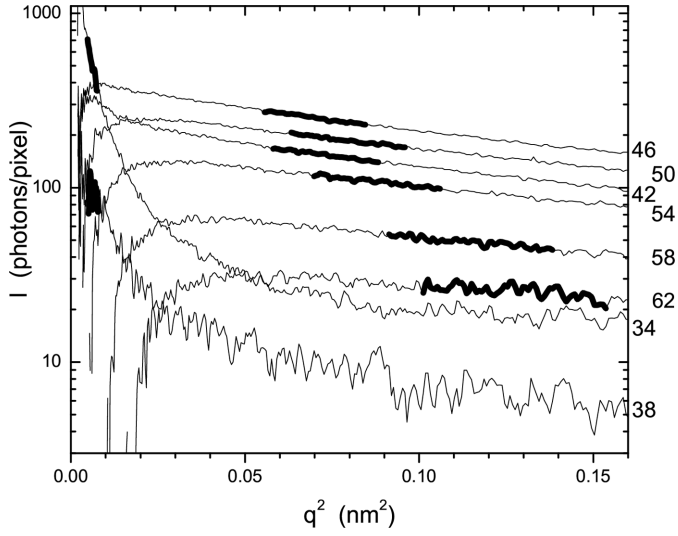


Figure 6
Guinier plots of data at selected time points indicated in Fig. 5 (indicated by the arrows in that figure). The region used for *Rg* determination, with an upper bound of q^2, such that $q_{\rm max} = 1.3/Rg$, and a lower bound of 66% of this value, is indicated by the heavy lines in each trace, while the rest of the data is indicated by thin lines. Note that, although most of the major peak Guinier plots (images 42 through 58) appear to exhibit slight upward curvature, they contain little if any aggregate as indicated by a steep peak near $q = 0$. The Guinier regions (heavy lines) are also clearly not parallel as they were in the case of the Guinier plots of cytochrome *c*, Fig. 4.

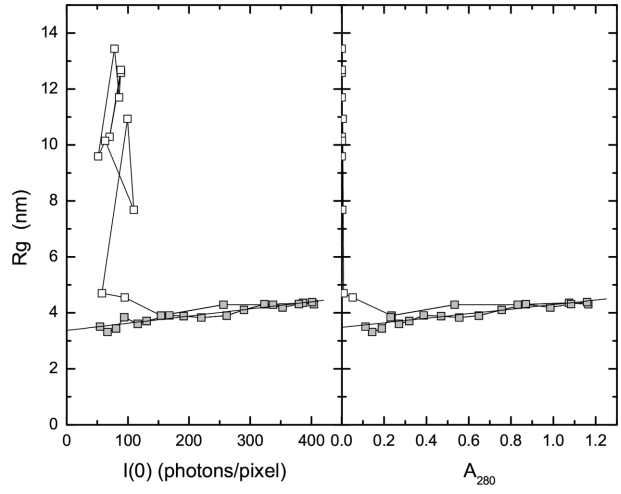


Figure 7
The major plasminogen peak shows strongly non-ideal behaviour, with *Rg* varying as the concentration changes. This can be seen when either $I'(0)$ or A_{280} are used as a proxy of concentration. The filled symbols are data points in the major peak region of Fig. 5, and were used to perform the linear extrapolation to [plasminogen] = 0. The open symbols are data points in the earlier aggregate region. The two different proxies for concentration yield a consistent Rg_0 = 3.4 nm [for A_{280}: 3.48 $\pm$ 0.07; for $I'(0)$: 3.37 $\pm$ 0.06].

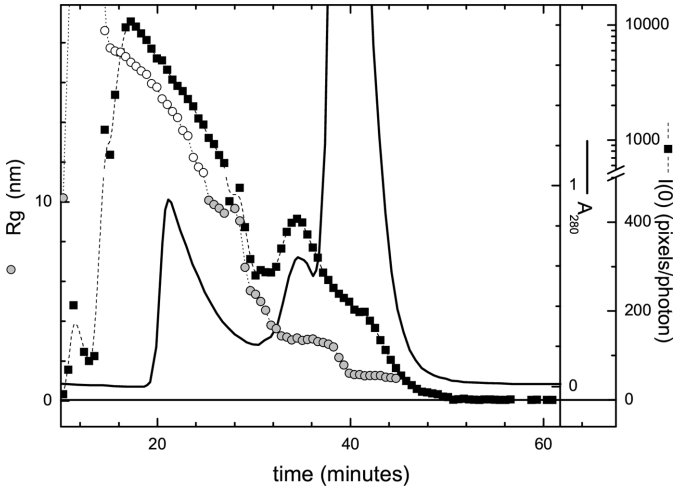


Figure 8
The ability of the technique to provide reliable data from impure samples and mixtures was illustrated by a run with a sample containing three components: cytochrome *c*, Rg = 1.3 nm, bovine serum albumin, Rg = 3.0 nm, and blue dextran, a carbohydrate molecule of very high molecular mass (>1 MDa) and uncertain but undoubtedly high Rg. This latter compound is a common chromatographic standard used to represent limitingly large molecules, and here serves as a proxy for aggregates. It can be seen that reliable plateau regions were observed for both proteins, even though the chromatographic separation was marginal. The Rg data for the blue dextran are not reliable, since our experimental set-up did not have the angular resolution to fully capture the Guinier region for molecules of this size. As such, the open circles do not satisfy the criterion $Rg\,q_{\max} = 1.3$ used everywhere else in this paper, and are presented only as guide to the increasing size of this molecule, as would be expected from its chromatographic position. Typically, in this region, $Rg\,q_{\max}$ was $\sim$2.5.

resolution of similar species by the molecular sieve column, or due to the presence of a dynamic equilibrium of various species or conformations. The latter is likely, since efforts to prepare clean samples by filtration (through 0.2 µm syringe filters) or centrifugation (13000 r.p.m., 30 min) were unsuccessful.

3.3. Mixture

The ability of this technique to facilitate SAXS analysis on impure samples was illustrated by the mixed sample, shown in Fig. 8. In this test, two proteins with different values of Rg were mixed with a high-molecular-mass macromolecule to simulate aggregates. Even though the resolution of these compounds was incomplete as judged by UV absorption, the Rg trace showed well resolved plateaus at the appropriate values of 3.0 for bovine serum albumin and 1.3 for cytochrome *c*, and efficiently separated the very high molecular mass material so that it did not hamper analysis.

4. Summary

The combination of size exclusion chromatography with SAXS promises to provide a more accurate and more sensitive method for measuring Rg, particularly for proteins prone to association or oligomerization. An automatic extrapolation to zero concentration is performed in a single experiment in a matter of minutes, increasing the fidelity of the subtraction of solvent scatter and allowing accurate detection of macromolecular scatter and more reliable Guinier analysis.

Another beneficial feature of this experimental set-up is its increased sensitivity. Accurate data were achieved from images in which $I'(0) < 10$ events per pixel. The phosphor coupling in the detector produces an amplification, so that each X-ray photon produced approximately ten events, and the radial integration was performed with an arc width of about 1 pixel, so that, for most of the data in which the arcs spanned the short dimension (1024 pixels) of the rectangular detector area, the intensity at each q point was derived from a total of only $\sim 10^3$ excess photons. Referring back to the UV absorption profiles, we can see that this translates to a concentration of <200 µg ml^{-1}, depending on the specific protein, which is an order of magnitude better than has been typical for SAXS of proteins samples. Part of this increase in sensitivity is due to the highly sensitive detector; however, the rapid almost contemporaneous acquisition of a cohort of sample images flanked by similar cohorts of buffer images is needed to produce accurate subtractions and accurately realise the excess scatter from such dilute solutions.

Accurate data were obtained at concentrations well below 1 mg ml^{-1}, depending on the specific molecular mass of the protein. Furthermore, this technique promised to provide an easily accessible accurate measure of the true aggregate-free forward scatter, $I'(0)$, which may also be used for molecular size analysis. This promises to expand the utility of this technique to proteins that are available in lesser amounts. Also, the coupling of SAXS measurement to a separatory technique allows measurements to be performed on unstable and impure proteins.

In the future, we believe that this technique will also be applicable to improve shape measurements taken at higher q. Chromatography of the sample necessarily dilutes it to some extent, which will be felt most strongly in the high-q region. However, for proteins for which adequate sample is available, so that modest concentrations can be loaded initially, this technique may improve the quality of the data. This technique may also be used to acquire a true aggregate-free estimate of the forward scatter, $I(0)$.

Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science, under contract No. W-31-109-ENG-38. BioCAT is a National Institutes of Health-supported Research Center, RR-08630.

References

Barth, H. G., Boyes, B. E. & Jackson, C. (1994). *Anal. Chem.* **66**, 595R–620R.
Castellino, F. J. & Powell, J. R. (1981). *Methods Enzymol.* **80**, 365–378.
Chu, B. & Hsiao, B. S. (2001). *Chem. Rev.* **101**, 1727–1761.
Doniach, S. (2001). *Chem. Rev.* **101**, 1763–1778.
Fischetti, R. F., Rodi, D. J., Mirza, A., Irving, T. C., Kondrashkina, E. & Makowski, L. (2003). *J. Synchrotron Rad.* **10**, 398–404.
Guinier, A. & Fournet, G. (1955). *Small-Angle Scattering of X-rays.* New York: Wiley.
Hammersley, A. P. (1997). *FIT2D: An Introduction and Overview.* ESRF Internal Report ESRF97HA02T. ESRF, Grenoble, France.
Irvine, G. B. (1994). *Methods Mol. Biol.* **32**, 267–274.
Kratky, O. & Pilz, I. (1972). *Q. Rev. Biophys.* **5**, 481–537.
Marshall, J. M., Brown, A. J. & Ponting, C. P. (1994). *Biochemistry*, **33**, 3599–3606.
Phillips, W. C., Stewart, A., Stanton, M., Naday, I. & Ingersoll, C. (2002). *J. Synchrotron Rad.* **9**, 36–43.
Ponting, C. P., Holland, S. K., Cederholm-Williams, S. A., Marshall, J. M., Brown, A. J., Spraggon, G. & Blake, C. C. (1992). *Biochim. Biophys. Acta*, **1159**, 155–161.
Svergun, D. I. & Koch, M. H. (2002). *Curr. Opin. Struct. Biol.* **12**, 654–660.
Wen, J., Arakawa, T. & Philo, J. S. (1996). *Anal. Biochem.* **240**, 155–66.
Wu, C. F. & Chen, S. H. (1988). *Biopolymers*, **27**, 1065–1083.
Wyatt, P. J. (1991). *Biochem. Soc. Trans.* **19**, 485.