

Quantifying Radiation Damage in Biomolecular Small-Angle X-ray Scattering

Authors

Jesse B. Hopkins^{a*} and **Robert E. Thorne^b**

^a Cornell High Energy Synchrotron Source, Ithaca, NY, 14853, United States

^bDepartment of Physics, Cornell University, Ithaca, NY, 14853, United States

Correspondence email: jbh246@cornell.edu

Synopsis A method for analysing radiation damage experiments in biological small-angle X-ray scattering is presented. This method allows direct comparison between results obtained at different beamlines and by different experimenters.

Abstract Small-angle X-ray scattering (SAXS) is an increasingly popular technique that provides low resolution structural information about biological macromolecules in solution. Many of the practical limitations of the technique, such as minimum required sample volume, and of experimental design, such as sample flow cells, are necessary because the biological samples are sensitive to damage from the X-rays. Radiation damage typically manifests as aggregation of the sample, which makes the collected data unreliable. However, there has been little systematic investigation of the most effective methods to reduce damage rates, and results from previous damage studies are not easily compared to results from other beamlines. Here we provide a methodology for quantifying radiation damage in SAXS to provide consistent results between different experiments, experimenters, and beamlines. We demonstrate these methods on radiation damage data collected from lysozyme, glucose isomerase, and xylanase, and find that no single metric is sufficient to describe radiation damage in SAXS for all samples. We found the radius of gyration, molecular weight, and integrated SAXS profile intensity to be a minimal set of parameters that capture all types of observed behaviour. Radiation sensitivities derived from these parameters show large protein dependence, varying by up to six orders of magnitude between the different proteins tested. This work should enable consistent reporting of radiation damage effects, allowing more systematic studies of the most effective minimization strategies.

Keywords: Small-angle X-ray scattering; SAXS; radiation damage.

1. Introduction

Small-angle X-ray scattering (SAXS) provides a low resolution structural probe for biological macromolecules in solution. SAXS measurements generally require a homogeneous, monodisperse, aggregate-free solution, and that these conditions be maintained throughout data collection (Meisburger et al., 2013). X-ray induced radiation damage can cause macromolecule aggregation,

fragmentation, conformation changes, and unfolding, all of which can be detected by SAXS.

Radiation damage is therefore a major obstacle for SAXS, and descriptions of dedicated biological SAXS beamlines acknowledge the need to check for and avoid radiation damage (David & Pérez, 2009; Hura et al., 2009; Pernot et al., 2010; Blanchet et al., 2012, 2015; Martel et al., 2012; Nielsen et al., 2012; Classen et al., 2013; Kirby et al., 2013a; b; Acerbo et al., 2015). Minimizing radiation induced changes in SAXS places limits on minimum sample volumes (~10 μ L) and maximum X-ray exposure times (Dyer et al., 2014; Skou et al., 2014). Radiation damage is also an obstacle to time resolved SAXS studies, as large amounts of sample must be available to obtain damage-free low-noise scattering profiles at many time points (Pollack, 2011; Graceffa et al., 2013).

To minimize radiation damage, three strategies are commonly employed. First, exposure times for a test sample can be reduced until subsequent exposures of the test sample show no change in the scattering profile (Dyer et al., 2014; Skou et al., 2014). Second, the total sample volume irradiated can be increased, typically by flowing/oscillating the sample or defocusing the beam at the sample, to minimize dose (Fischetti et al., 2003; Lipfert et al., 2006; Classen et al., 2013). Third, small molecule compounds such as glycerol can be added to reduce changes in SAXS profile (e.g., by competitively binding with free radicals or by inhibiting aggregation) (Kuwamoto et al., 2004; Kmetko et al., 2011; Bobrowski, 2012; Jeffries et al., 2015). These approaches can be employed in parallel, and result in the limitations to sample volume and exposure time given above. Cryocooling samples to 100 K has been shown to reduce radiation damage rates in SAXS (Meisburger et al., 2013; Hopkins et al., 2015), but substantial methodological development is required before cryocooling is accepted for routine use. Despite the importance of radiation damage as a limiting factor in SAXS, early efforts using laboratory X-ray sources (Zipper & Durchschlag, 1980a; b; c, 1981; Zipper et al., 1980, 1985; Zipper & Kriechbaum, 1986), have been followed by only two systematic, quantitative studies at synchrotron sources (Kuwamoto et al., 2004; Jeffries et al., 2015). With recent and planned upgrades to already bright third generation sources and construction of high brightness fourth generation sources, understanding, quantifying, and ultimately minimizing radiation damage in biological SAXS will be essential to efficient use and full exploitation of these sources.

To aid in the development of effective strategies for minimizing radiation damage, nominally identical experiments carried out by different experimenters or at different beamlines should yield identical results. Previous studies do not give a consistent framework for quantifying damage, and neglect variables that may affect reported damage rates. For example, (Meisburger et al., 2013; Jeffries et al., 2015) use the same formula to calculate the absorbed X-ray energy, but differences in how they evaluate beam area give, if everything else were equal, a factor of six difference in absorbed energy. Such discrepancies make evaluating claimed damage mitigation effects based on comparisons between independent studies very difficult.

Here we build upon previous work (Kuwamoto et al., 2004; Meisburger et al., 2013; Jeffries et al., 2015) to develop a procedure for quantifying radiation damage in SAXS that is broadly applicable and allows comparison between different samples and beamlines. We discuss how to accurately quantify absorbed X-ray energy and dose (absorbed energy per unit mass of the sample) for static SAXS samples, accounting for variables including beam shape, diffusive exchange of protein into and out of the beam, and beam heating. Flowing or oscillating samples significantly complicate the calculation of dose, so we recommend and discuss in detail only the static experiment. Different proteins show different modes of damage such that a single metric is insufficient to quantify radiation damage. We motivate and illustrate our procedure using radiation damage data for lysozyme, glucose isomerase, and xylanase, which was collected at the MacCHESS BioSAXS user facility at the CHESS beamline G1 (Nielsen et al., 2012; Acerbo et al., 2015). Details of the data collection and processing methods are given in Section S1; standard protocols for SAXS data collection and analysis have been previously reviewed (Dyer et al., 2014; Skou et al., 2014). Accurate quantification of dose and application of suitable damage metrics will facilitate development of methods to minimize radiation damage in SAXS.

2. Overview of a radiation damage experiment

The general steps recommended to quantify radiation damage in a static SAXS experiment are as follows:

1. *Calibration*: measure X-ray flux and beam shape at the sample position. Measure the path length through the sample, or the sample transmission. If necessary, measure or calculate sample density.
2. *Measurement*: For each sample of interest, record consecutive exposures until clearly detectable damage is observed. For a given sample condition (protein concentration, buffer, temperature, etc.), measure at least 3 identically prepared samples.
3. *Dose calculation*: Calculate the X-ray dose for each exposure.
4. *Quantification*: Calculate, at a minimum, the following parameters for each scattering profile: radius of gyration, molecular weight, and absolute integrated intensity. Normalize the parameters for a given sample to their initial values. Plot the normalized parameters vs. dose.
5. *Damage sensitivity metrics*: The resulting plots will generally have an initial linear region. Fit this region to obtain a radiation sensitivity per dose for each parameter.

If these steps are carried out as described below, data from different experiments and beamlines should be directly comparable. For radiation damage experiments that use non-static samples (particularly flowing samples), consecutive exposures may not be sufficient to increase dose delivered to the sample. Other variables may need to be changed in step 2, such as flux density (by changing flux and/or beam size) or flow rate, to vary sample doses.

3. Definition of dose

Dose is the X-ray energy absorbed per unit mass of the sample, in units of Gray (Gy), where 1 Gy = 1J/kg. Radiation damage in most contexts is a strict function of dose (Holton, 2009). Dose is related to the number of absorbed or inelastically scattered photons, and for static samples is calculated as

$$\text{Dose} = \frac{ftAE_{\gamma}}{\rho l}, \quad (1)$$

where f is the flux density (ph/s/μm²), t is the exposure time, A is the fraction of incident energy absorbed, E_{γ} is the X-ray energy per photon, ρ is the sample density, and l is the X-ray path length through the sample (Kmetko et al., 2006). For comparison between different experiments on the same beamline or between different beamlines, number of incident photons and exposure time are not good proxies for dose. Dose can (and should) be calculated for non-static (such as flowing) samples, but different inputs will be needed to determine the absorbed energy and illuminated sample mass.

Accurately calculating the sample dose is critical for quantification of radiation damage. In macromolecular crystallography, tools for calculating dose (Murray et al., 2005; Paithankar et al., 2009; Paithankar & Garman, 2010; Zeldin et al., 2013) are widely used in optimizing data collection. Such tools are not available for SAXS, and previous studies (Kuwamoto et al., 2004; Meisburger et al., 2013; Jeffries et al., 2015) used different dose calculations.

4. Calibration of experiment and sample parameters

In order to accurately determine dose, the following experimental parameters must be known: buffer composition; macromolecule concentration; X-ray energy; exposure time; *flux and beam shape at the sample position; sample path length or sample transmission*; sample density; and the macromolecular diffusion coefficient in the buffer. The first four parameters are generally known for any SAXS experiment (Dyer et al., 2014; Skou et al., 2014), while the last two can be measured or calculated away from the beamline. The italicization indicates additional parameters not typically measured/reported for SAXS experiments that must be measured at the beamline for each experiment.

Beam shape at the sample position can be measured in a number of ways. These include detection of optical beam images (such as from fluorescence or scintillation); X-ray exposure of a glass slide followed by optical measurement of the resulting fogging (Meisburger et al., 2013); and scanning a knife edge through the beam in perpendicular directions, measuring the flux downstream of the knife edge (for example on an active beamstop), and then calculating the derivative of the flux vs. knife edge position. This latter method was used here, and details are given in Figures S1 and S2.

The flux at the sample position can be measured by inserting a calibrated ion chamber or PIN diode at the sample position, and then accounting for the attenuation of the upstream sample holder window.

Alternatively, flux can be measured upstream or downstream of the sample, and the attenuation of any intervening material (such as windows on the downstream flight tube) can be calculated and accounted for. Here we measured the flux downstream of the vacuum flight tube and corrected for flight tube window attenuation using XCOM (Gerward et al., 2001; Berger et al., 2010).

The sample path length can usually be obtained from a beamline scientist or by direct measurement of the sample holder. If it is unknown or not fixed, the transmission of a sample of known attenuation length (e.g., water) can be measured and used to estimate the path length, as in (Meisburger et al., 2013). As in (Jeffries et al., 2015), sample densities can be calculated, for example by using MULCh (Whitten et al., 2008), or measured. For solutions with modest protein, salt and buffering agent concentrations, densities calculated using MULCh (Jeffries et al., 2015) are within 3% of that of water, so using the density of water instead of the specific solution density will generally have a minimal effect on the dose calculation. Here we use the density of water, 1 g/cm³.

5. Measurement of radiation damage data

In a static experiment (Kuwamoto et al., 2004; Jeffries et al., 2015), the sample is held stationary with respect to the X-ray beam and a series of SAXS images are recorded. Current detectors have minimal dead times between images, allowing continuous exposure with the X-ray shutter open for the experiment's duration and the most straightforward dose calculation. Detectors with significant dead times require beam shuttering between exposures. As we will discuss later, identically prepared samples do not yield identical damage rates, for reasons currently unknown. To account for this variability, at least three samples should be measured for a given set of sample conditions (sample concentration, buffer composition, etc.).

Figure 1 shows sample SAXS scattering profiles at increasing accumulated dose (calculated as described in Section 6) for lysozyme, xylanase, and glucose isomerase, collected in static mode with continuous exposure. Both lysozyme and xylanase show an increase in intensity with dose at low q . This increase – the only previously reported effect of radiation damage in SAXS – is characteristic of aggregation: aggregates produce much more low- q scattering than an equivalent number of isolated protein molecules, and so even small amounts can be easily detected (Kuwamoto et al., 2004; Skou et al., 2014; Jeffries et al., 2015). However, lysozyme also shows a decrease in intensity with dose at high q . This is also consistent with aggregation, as scattering from large particles decays to zero at lower q than from small particles, although whether the amount of protein involved in aggregates is sufficient to explain the large- q intensity reduction (also possibly visible in Figure 1b of (Jeffries et al., 2015)) is unclear. And in contrast, glucose isomerase shows a decrease in intensity with dose at low q . The response of biomolecules to radiation, as manifested in their SAXS profiles, is thus quite diverse.

While commonly used to reduce radiation damage, flow/oscillation of the sample complicates the dose calculation. For typical flow cell diameters/widths (1.5–2 mm), flow rates (1–30 $\mu\text{l/s}$) and mean flow velocities (0.5–15 mm/s), the corresponding Reynolds' numbers of between 1 and 25 are fully in the laminar flow regime. The expected Poiseuille velocity profile is parabolic with radial position r . For typical mean flow velocities and beam sizes ($\sim 500\ \mu\text{m}$), mean sample residence times in the beam of 0.03–1 s are too short for appreciable radial diffusive mixing (see section 6.3 below.) Consequently, the Poiseuille profile results in radius-dependent biomolecule residence times in the X-ray beam. The resulting dose distribution within a flowing sample is nonuniform and depends on the geometry of the flow cell and the volume of the cell that is illuminated by X-rays. Similar behavior is expected in oscillating flow configurations. Static experiments are thus recommended for accurate quantification and comparison between beamlines. Once static radiation sensitivities are established, they can be used to estimate — and/or be directly compared to — damage in a flow/oscillation experiment. Even in static experiments, the sample plug may slowly creep or drift, bringing fresh sample into the beam; samples should be carefully observed or imaged before and after each measurement.

6. Calculation of dose

Accurate calculation of dose using Eq. (1) requires accurate determination of all parameters. The flux density and absorbed energy fraction can be subject to a variety of corrections, as outlined below.

6.1. Determining the flux density

Flux density depends on beam shape (Zeldin et al., 2013). For a robust metric of radiation damage induced change to the scattering profile, we need to consider not only the amount of damage caused to the sample at each position in the beam profile, but also the magnitude of each position's contribution to the total SAXS signal. Sample regions where the flux density and thus damage are small also contribute little to the total scattering, and so should be appropriately weighted — by the flux density — when calculating dose. The appropriate weighted flux density $\bar{f}$ for use in Eq. (1) is given by

$$\bar{f} = \frac{\int f(x,y) f(x,y) dA}{\int f(x,y) dA}, \quad (2)$$

where $f(x,y)$ is the beam profile perpendicular to the beam direction (z).

For a Gaussian beam,

$$f(x,y) = \frac{f_0}{\sqrt{2\pi\sigma_x\sigma_y}} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right], \quad (3)$$

where f_0 is the total number of incident photons per second and σ_x and σ_y are the standard deviation of the Gaussian in x and y . Plugging this into Eq. (2) gives

$$\bar{f} = \frac{f_0}{4\pi\sigma_x\sigma_y}. \quad (4)$$

The FWHM of a Gaussian is given by $\sigma \times 2\sqrt{2\ln 2}$, so the ratio of the flux density calculated by the FWHM method ($f_{FWHM} = f_0 / (FWHM_x FWHM_y)$) and using Eq. (2) is

$$\frac{f_{FWHM}}{\bar{f}} = \frac{\pi}{2\ln 2} \approx 2.27. \quad (5)$$

This shows that accounting for the beam shape will apply a significant correction to the reported dose.

In the present work, $\bar{f}$ was numerically calculated using Eq. (2) and the measured beam profile.

(Meisburger et al., 2013) assumed the incident photons were uniformly distributed within the FWHM; if their beam was Gaussian, this overestimated the dose by a factor of 2.3. (Jeffries et al., 2015) assumed the photons were uniformly distributed over the full extent of the beam ($\sim 2.5 \times \text{FWHM}$); if their beam was Gaussian, this underestimated the dose by a factor of 2.5.

6.2. Determining the absorbed energy fraction

In Eq. (1), the fraction of the incident energy that is absorbed, A , can be calculated from Beer's law,

$$A = 1 - \exp\left(-\left(\frac{\mu}{\rho}\right)\rho l\right), \quad (6)$$

where μ is an absorption coefficient and l is the sample path length. Absorption coefficients are tabulated as μ / ρ (Hubbell, 2006), where ρ is the sample density. Not all of the energy initially absorbed by the sample will stay within the volume of interest (the illuminated volume). The use of different absorption coefficients accounts for different energy loss mechanisms. A detailed discussion of absorption coefficients, given in Section S2, shows that the mass photoelectric absorption coefficient, μ_{pe} / ρ , and the mass-energy absorption coefficient, μ_{en} / ρ , are both reasonable choices of absorption coefficients for SAXS. Further, at our energy of ~ 10 keV, there is no difference between μ_{pe} / ρ and μ_{en} / ρ for water so either is acceptable. We chose to use μ_{pe} / ρ to calculate the absorption, as it can be directly calculated for each sample condition using XCOM (Gerward et al., 2001; Berger et al., 2010). We note that the addition of buffer components and protein changes μ_{pe} / ρ by only $\sim 1\%$ (from $4.944 \text{ cm}^2/\text{g}$ for water to $5.008 \text{ cm}^2/\text{g}$ for 47.5 mg/mL lysozyme in buffer),

so for our samples it would create minimal error to use the μ_{pe} / ρ of water for every sample. If the path length of the sample is unknown, for example for the windowless sample holder used in (Meisburger et al., 2013), the transmission can be used to approximate the path length as

$l \approx \mu^{-1} \ln(1/T)$. This can be used with Eq. (6) to estimate A .

6.3. Correcting for diffusive turnover

Eq. (1) for dose assumes that all sample molecules within the X-ray beam are static. However, molecules will diffuse into and out of the illuminated volume, and also diffuse between regions having different flux densities within that volume. As with oscillating the sample, this will spread the energy absorbed from the beam over a larger volume, reducing the dose. The magnitude of this effect depends on the data collection time (and, in shuttered data collection, on the detector dead time between images), which determines maximum diffusion distances; the beam size, shape, and intensity profile, which determine the distance molecules must diffuse to experience a substantially different flux density; and the molecular diffusion coefficient of the macromolecule in the buffer.

A timescale for this effect can be estimated by setting the average diffusion length $l = \sqrt{Dt}$, where D is the diffusion coefficient, equal to the half of the beam FWHM (or the smallest dimension for non-circular beams) as

$$t_d = \frac{l_{FWHM}^2}{4D}. \quad (7)$$

Since D is inversely proportional to a molecule's Stokes radius R_s , $t_d \propto R_s$. For experiments on a given sample longer than t_d we expect diffusive turnover to be large. Section S3 describes a more complete calculation to evaluate the dose correction due to diffusion. This calculation shows that at t_d , diffusion reduces the dose evaluated using Eq. (1) by a factor of 2. This validates t_d as a reasonable indicator of total exposure lengths for which diffusion effects will be important. For the SAXS data collected here for lysozyme, xylanase, and glucose isomerase — using the measured beam size of 190×196 micron FWHM and maximum total exposure times of ~ 3 , 30, and 120 s respectively (see Section S1 for details) — the maximum diffusion corrections determined by the calculations in Section S3 were 2.5%, 35%, and 43% respectively.

The timescale for diffusion is proportional to l_{FWHM}^2 , and so rapidly becomes shorter as beam size shrinks. For beams with FWHM = 100 μm and 10 μm , $t_d = 20$ s and 0.2 s respectively, for lysozyme. Figure 2 shows how much diffusion reduces the dose, calculated as in Section S3 relative to the expected dose without diffusion, for lysozyme irradiated by different beam sizes.

7. Quantifying radiation damage data

7.1. Calculate parameters from scattering profiles, normalize

High throughput SAXS beamlines are increasingly providing automated methods for assessing when radiation damage occurs (Pernot et al., 2010; Blanchet et al., 2015; De Maria Antolinos et al., 2015). These methods use statistical techniques to determine when the measured scattering profile has significantly changed relative to the initial exposure (Franke et al., 2012, 2015; Grant et al., 2015). While practical for assessing the onset of radiation damage, these do not give data that can be easily interpreted as to type and rate of damage.

Almost any parameter that can be calculated from a SAXS profile can be used as a metric for radiation damage. Using a series of consecutive images, the parameter can be normalized by its value in the first image, and then plotted versus dose, yielding a "dose curve." Previously, metrics based on the change in the radius of gyration R_g , the pseudo radius of gyration, and the scattering intensity at zero angle $I(0)$ were used to study radiation damage (Kuwamoto et al., 2004; Jeffries et al., 2015). We have investigated these and metrics based on additional parameters, including the Porod Invariant, the Porod volume, the molecular weight, the maximum dimension, and the integrated intensity. Details of these calculations are given in Section S1.

Figure 3 shows representative dose curves calculated from our sample data sets for lysozyme, xylanase, and glucose isomerase, for these parameters. These dose curves show how radiation damage changes the information available from the SAXS scattering profile.

7.2. Calculate sensitivities from fits to dose curves

Dose curves generally show a linear region of change in parameter versus dose. The slope of the linear fit to the dose dependence of parameter P then gives the radiation sensitivity S_p . Most SAXS experiments deliver doses in the 1–10 kGy range, so % change per kGy provides a convenient unit allowing a quick estimate of damage in a typical experiment. Table 1 gives experimentally determined sensitivities for lysozyme (4.1 mg/mL), xylanase (4.9 mg/ml), and glucose isomerase (1.2 mg/mL), determined as described in Section S4. Figure 4 shows that identically prepared samples may exhibit different dose curves and different sensitivities. Consequently, it is important to measure at least three identically prepared samples, and report the average and standard deviation of the sensitivity. Section S4 gives a detailed discussion of the behavior of these samples vs. dose, including possible sources of sample-to-sample variability.

7.3. Which sensitivities are important?

Given the large number of possible parameters P , we would like to identify a minimal set that accurately captures the diverse radiation responses of biomolecules in SAXS. To do this, we calculated the Pearson product-moment correlation coefficient (Pearson's r) between every parameter P and for each SAXS data set, and then averaged the r values over all identically prepared samples. Example plots of these correlation coefficients for lysozyme, xylanase, and glucose isomerase are shown in Figures S3–S5. For lysozyme, Figure S3, all of the parameters were strongly correlated ($r \geq 0.87$ for 4.1 mg/mL lysozyme) except for the Porod invariant. As seen in Figure 3a, the Porod invariant doesn't change for these samples, so in this case, any other parameter would characterize radiation damage in the system.

For xylanase, Figure S4, there was less overall correlation. $I(0)$ correlated strongly with the integrated intensity and R_g ($r \geq 0.84$), and both $I(0)$ and R_g correlated with the pseudo- R_g ($r \geq 0.63$). The molecular weight correlated strongly with the Porod volume ($r = 0.93$). D_{max} correlated with R_g , $I(0)$, and Porod volume ($r \geq 0.7$), and the Porod invariant was correlated with the integrated intensity ($r = 0.68$). This suggests that the R_g , molecular weight, and integrated intensity would serve as a minimal parameter set for examining damage.

For glucose isomerase, significant change with dose was only seen in the integrated intensity, molecular weight, Porod volume, and Porod invariant. Figure S5 shows, as expected, that the molecular weight was strongly correlated with Porod volume ($r = 0.94$), and both are strongly anticorrelated with the Porod invariant, ($r \leq -0.95$). The integrated intensity was correlated with $I(0)$ ($r = 0.57$) but was otherwise not strongly correlated ($|r| \leq 0.42$) with anything. In this case molecular weight and integrated intensity would represent the damage.

Thus, for the proteins measured here, the radius of gyration, molecular weight, and integrated intensity together capture the q -dependent changes in SAXS profiles with dose. R_g and molecular weight both provide insight into the structural changes caused by damage. Integrated intensity yields no obvious structural insight, but it is sensitive to any change in the scattering profile. Integrated intensity also has the advantage that it can be used to characterize radiation-induced changes in scatter from protein-free buffers. The integrated intensity will depend upon the q range and the instrument background, and so will not typically be useful for comparison with results from separate experiments. For this work, the entire q range (see Section S1) was used to calculate the integrated intensity. Note that the parameters used in previous work (Kuwamoto et al., 2004; Jeffries et al., 2015) - R_g and $I(0)$ - fail to capture the most important changes for glucose isomerase in Fig. 1c.

In addition to the correlation analysis described above, principle component analysis was performed on the data, and is described in Section S5. The results of that analysis generally support the results of

the correlation analysis, though the interpretation is not as straightforward. Further details of the correlation analysis are given in the same section.

7.4. Quantification in a nutshell

To summarize, first calculate the interesting parameters for each measured scattering profile. Based on our sample data, we recommend calculating, at a minimum, the radius of gyration, molecular weight, and integrated intensity. Second, normalize these parameters to the initial value and plot vs. dose to generate dose curves. Third, fit the linear region of each dose curve to find the sensitivity of that parameter.

8. Additional considerations

The preceding sections provide the framework needed to reliably quantify radiation damage in SAXS, in a way that will allow comparison between different experiments and beamlines. This section will discuss experimental parameters that are either known to or may affect quantification, and so should be considered and reported for experiments. We will also discuss the idea of a critical dose, previously used as a metric for radiation damage.

8.1. Buffer composition, temperature, macromolecule concentration, and degassing

Several additional factors could affect the radiation sensitivity measured by an experiment. The first of these is the buffer composition. There is significant evidence that certain additives, such as glycerol or ascorbic acid, reduce the rate of damage in SAXS (Kuwamoto et al., 2004; Jeffries et al., 2015). Buffering agents and salts may also have an effect. For example, Tris scavenges OH radicals (Davies et al., 1987; Audette-Stuart et al., 2005) and may be a more effective radical scavenger than glycerol (Krisch et al., 1991), while sodium nitrate scavenges aqueous electrons (Audette-Stuart et al., 2005). Having identical buffers is thus important when attempting to study the effect of other experimental variables.

While temperature may nominally affect both macromolecule and radical diffusion rates, previous and present measurements (described in Section S4) indicate that damage rates don't vary significantly between 4 and 30 °C (Kuwamoto et al., 2004; Jeffries et al., 2015), although there is a huge reduction in damage rate on cooling to 100 K (Meisburger et al., 2013). Protein concentration affects damage rate (Kuwamoto et al., 2004; Jeffries et al., 2015), so must also be properly controlled. Our data, Section S4, shows that a two order of magnitude increase in lysozyme concentration reduces the measured sensitivities by a factor of ~20.

It is sometimes mentioned that a deoxygenated environment or removal of dissolved oxygen from the solution can reduce radiation damage in SAXS (Hura et al., 2009). Deoxygenating, or, more generally, degassing the solution, is also done to reduce the chance of bubble formation upon

oscillation/flow, or to prevent the dissolved gasses from being forced out of solution when exposed to the X-ray beam (Kirby et al., 2013b). In Section S4 we report damage rates of degassed samples of lysozyme and xylanase. These degassed samples damaged slightly faster than the normal samples. However, due to the large standard deviations involved, we cannot conclusively say there was any effect from degassing. The effect may be small, and it may also be protein and/or buffer dependent. Dissolved molecular oxygen in solution has been observed to both increase and decrease radiation sensitivity of macromolecules (Saha et al., 1995). The assumed mechanism for sensitization is generation of superoxide radicals, $O_2^{\bullet-}$, and singlet oxygen (Garrison, 1987), and reaction of these species with the macromolecules (Davies, 1987; Davies & Delsignore, 1987; Davies et al., 1987). Sensitization by a factor of ~2-3 has previously been observed (Saha et al., 1995). When dissolved oxygen acts as a desensitizer, it is assumed to be due to scavenging of H^+ and e_{aq}^- by the oxygen in systems that are more sensitive to those species than the generated superoxide radicals (Garrison, 1987; Saha et al., 1995).

8.2. X-ray energy and dose rate

X-ray energy and/or dose rate could affect the damage rate in SAXS samples. In X-ray crystallography, changing X-ray energy does not change global damage rates (Murray et al., 2005; Shimizu et al., 2007), but there is some evidence from cryocooled samples that site specific damage rates are affected (Homer et al., 2011). Aggregation can be driven by site specific damage, such as the reduction of surface cysteine residues and subsequent cross-linking with cysteines on other molecules (Durchschlag & Zipper, 2007), and so might change with X-ray energy in SAXS. However, any energy-dependence of site-specific damage to proteins in solution is likely much smaller than in protein crystals: in crystals, more protein damage is due to direct interaction with photo-electron generated secondary electrons whereas in dilute solution, most protein damage is due to interaction with radicals generated in the solvent.

For low dose rates (28-162 Gy/s), increasing the dose rate was found to increase the damage rate (Kuwamoto et al., 2004), and this was attributed to diffusive turnover of the sample. It is unclear if there is an additional dose-rate effect once diffusion is accounted for.

8.3. Beam Heating

Significant heating of macromolecules could lead to deleterious effects, such as denaturation, that could be mistaken for radiation damage. Beam heating could also create temperature gradients, driving convection and mixing that might reduce the apparent rate of damage. As discussed in Section S6, we have estimated the heat diffusion timescale, the adiabatic heating rate, the adiabatic temperature rise ΔT_{ad} , the steady state temperature rise ΔT_{ss} (Warkentin et al., 2012), and a

minimum time to the onset of natural convection, t^* . Table 2 gives values calculated using our experimental parameters, and for several SAXS beamlines using parameters available online, which may not reflect the current state of the beamlines. Some beamlines may not use the maximum available flux for biological experiments, which would further reduce the calculated parameters. For our experiments, $\Delta T_{ad} = 0.08$ K, $\Delta T_{ss} = 0.13$ K, and $t^* = 86$ s, so neither beam heating nor convection is expected to influence our results. The same appears true for current experiments on other SAXS beamlines.

8.4. Critical Dose

Previous work has defined critical doses for radiation damage in SAXS (Kuwamoto et al., 2004; Jeffries et al., 2015). In (Kuwamoto et al., 2004), the critical dose is somewhat nebulously defined as the dose where damage is first observed. In (Jeffries et al., 2015), the critical dose is defined as the dose at which the pseudo R_g has changed by 0.1 nm. Applying the definition of (Jeffries et al., 2015) to our glucose isomerase data, we find a critical dose of $\sim 66,000$ kGy. However, the molecular weight shows a significant ($\sim 13\%$) change after just ~ 75 kGy. This suggests that one particular definition of a critical dose will not work for all proteins.

The critical dose for a system will depend upon the metrics/macromolecular properties of interest. For example, R_g is typically determined to a precision of ~ 0.1 – 0.2 Å (in the best cases), which for lysozyme corresponds to a 1% change in R_g . One might then define a critical dose as that which causes a 1% change in R_g , which for our lysozyme data at concentrations near 4 mg/mL is ~ 0.04 – 0.05 kGy. However, molecular weight is, as a rule of thumb, not determined in SAXS to better than $\sim 10\%$, and our lysozyme data then gives a dose limit of ~ 0.2 – 0.3 kGy. The intrinsic "noise" and precision limits on SAXS-derived parameters may vary between samples and beamlines, and these may also influence choice of critical dose. For these reasons, we recommend using radiation sensitivities for each SAXS-derived parameter rather than a single critical dose, as they provide a more universal metric for damage rates.

9. Summary

Radiation damage in SAXS is a serious concern: it limits minimum sample volumes to ~ 10 microliters; it increases the need for frequent and aggressive cleaning of sample cells, slowing down data collection; and it requires extra data analysis in high throughput experiments to ensure damage does not affect the results. The best strategies for reducing radiation damage remain oscillating or flowing the sample through the beam, to spread out the absorbed energy over a large volume (lowering the dose), and the addition of small molecule compounds to the buffer to reduce damage (Fischetti et al., 2003; Kuwamoto et al., 2004; Lipfert et al., 2006; Jeffries et al., 2015). Additionally,

cryocooling samples to 100 K, cryoSAXS, can greatly reduce damage to the samples (Meisburger et al., 2013; Hopkins et al., 2015). As X-ray sources get brighter, with the upgrade of third generation sources and the construction of fourth generation sources, the need to prevent radiation damage will get more urgent. This paper builds upon previous literature to provide a protocol for accurate quantification of radiation damage in SAXS. This should enable accurate comparison of results between different beamlines and experiments, and so enables new studies of the most effective methods for reducing radiation damage.

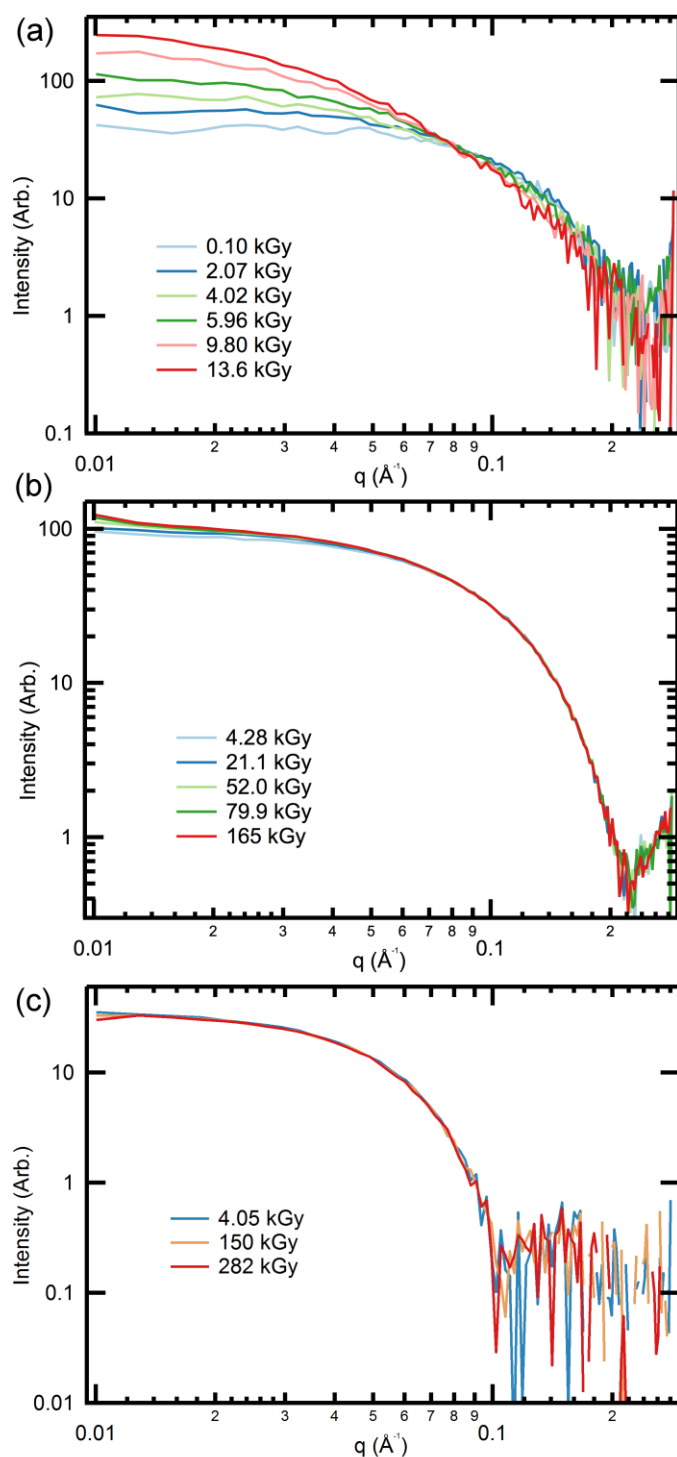


Figure 1 Selected scattering profiles of (a) 4.1 mg/mL lysozyme, (b) 4.9 mg/mL xylanase, and (c) 1.2 mg/mL glucose isomerase as a function of dose. For lysozyme and xylanase, the damage manifests primarily as an increase in intensity at low q , indicating aggregation. For glucose isomerase, only a slight downturn is observed at the lowest q , which could be attributable to damage or charging.

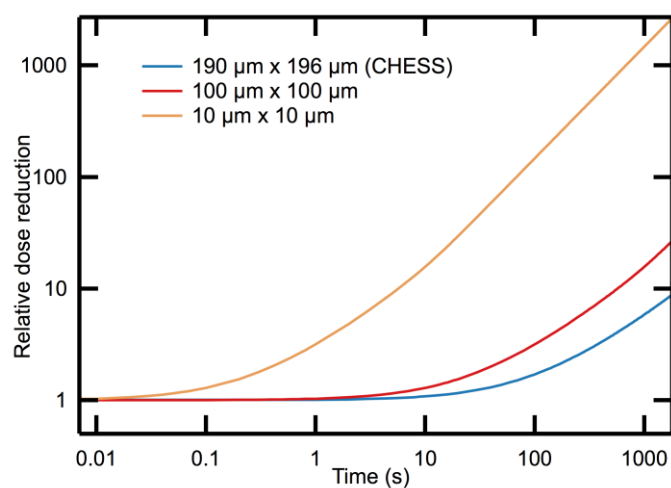


Figure 2 Dose reduction by diffusive exchange of lysozyme versus time for three beam sizes: $190 \times 196 \mu\text{m}^2$ (used for the experiments in this paper), $100 \times 100 \mu\text{m}^2$, and $10 \times 10 \mu\text{m}^2$. The vertical axis is the (unitless) factor by which diffusion reduces the dose, relative to the expected dose without diffusion. These reduction factors were calculated as described in Section S3.

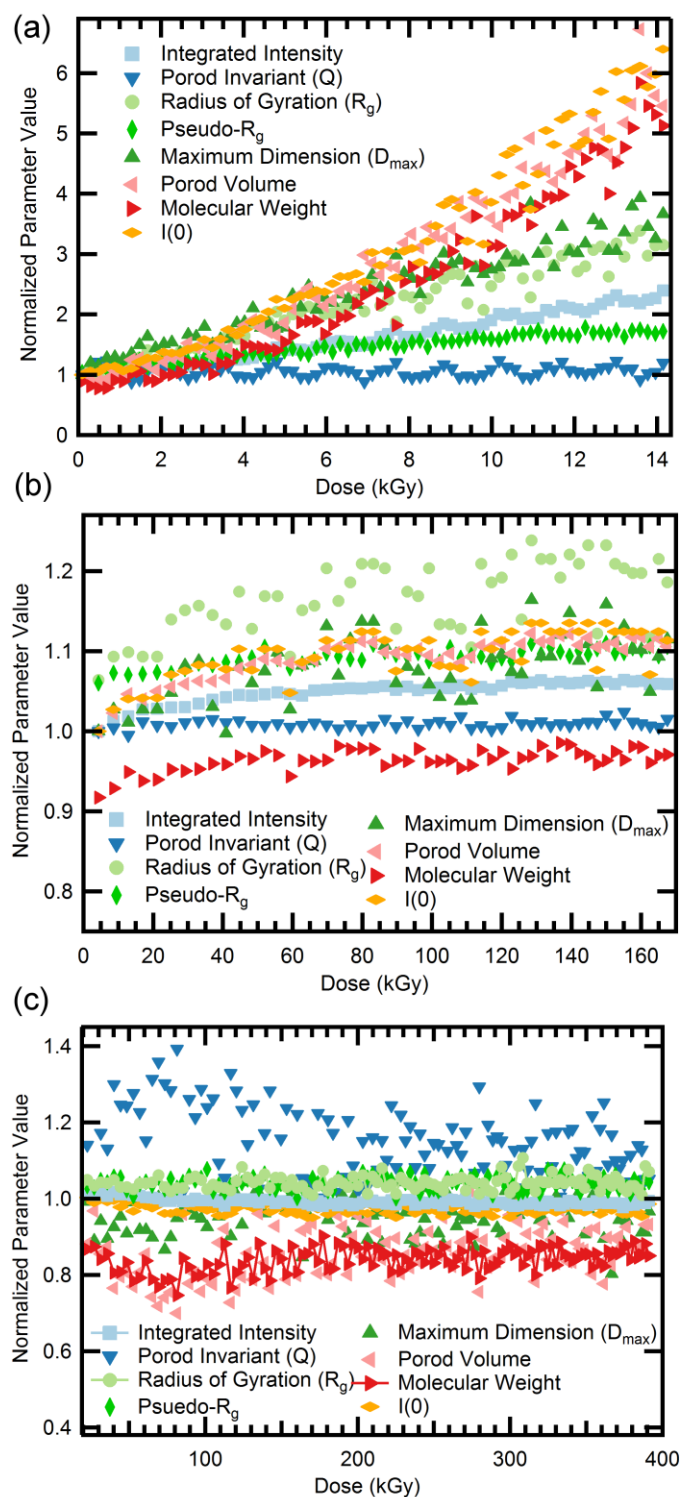


Figure 3 Plot of normalized calculated parameters vs. dose for every scattering profile measured for single samples of (a) 4.1 mg/mL lysozyme, (b) 4.9 mg/mL xylanase, and (c) 1.2 mg/mL glucose isomerase. For lysozyme and xylanase, every parameter except for the Porod invariant shows an increase with dose. For lysozyme, the integrated intensity, R_g , and maximum dimension are linear for the entire dose range, while the Porod volume, molecular weight, and $I(0)$ exhibit linear behavior only at doses below ~3 kGy. For xylanase, all of the increasing parameters show a relatively linear

region for the first ~60 kGy. Unlike lysozyme and xylanase, for glucose isomerase most parameters remain roughly constant. An increase is clearly visible for the Porod invariant, while a similar decrease is visible for the Porod volume and molecular weight. It is not clear if the changes in the glucose isomerase sample are actually reflecting structural changes in the protein or are due to another effect such as charging.

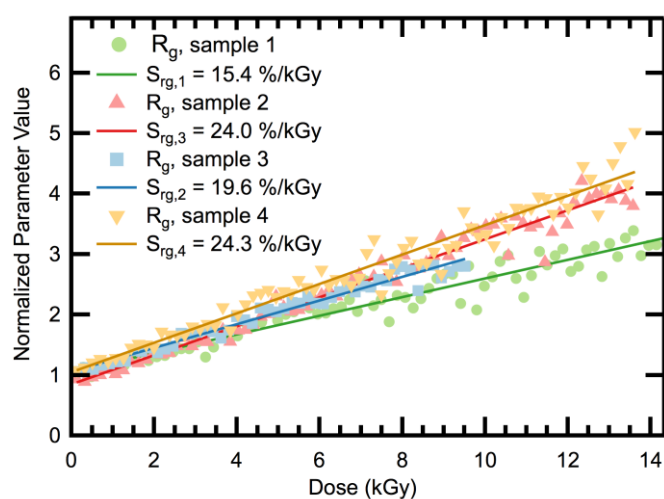


Figure 4 The change in normalized R_g as a function of dose for four identically prepared lysozyme samples. The lines represent the best linear fit, and the values of the slopes are reported in the legend. For nominally identical samples, sensitivities ranged from 15.4 %/kGy to 24.3 %/kGy, emphasizing the necessity of measuring and averaging sensitivities from multiple samples.

Table 1 R_g , molecular weight, and integrated intensity sensitivities (S_{rg} , S_{mw} and S_I) and standard deviations (σ_{rg} , σ_{mw} , and σ_I) for lysozyme (4.1 mg/mL), xylanase (4.9 mg/mL), and glucose isomerase (1.2 mg/mL). Glucose isomerase shows radically different sensitivities, $S_{rg} \ll S_{mw}, S_I$, while for lysozyme and xylanase these sensitivities agree to within a factor of 2-3, or less.

Protein	Concentration (mg/mL)	S_{rg} (%/kGy)	σ_{rg} (%/kGy)	S_{mw} (%/kGy)	σ_{mw} (%/kGy)	S_I (%/kGy)	σ_I (%/kGy)
Lysozyme	4.1	21.3	3.7	37.6	18.6	12.3	1.9
Xylanase	4.9	0.44	0.13	0.31	0.15	0.24	0.08
Glucose Isomerase	1.2	-0.000046	0.0007	-0.10	0.08	-0.023	0.009

Table 2 Estimates of beam heating at several BioSAXS beamlines. All dose and heating rate calculations assume a sample comprised of pure water. Dose calculations used Equation (1), assuming the FWHM for a Gaussian beam. When the actual sample path length could not be found, a standard sample path length of 1.5 mm, indicated by a *, was used. Calculations of t_d , δT_{ad} , ΔT_{ad} , ΔT_{ss} , and t^* used the smallest beam dimension. Flux, beam size, and energy numbers are from: BioCAT, measured April 2015; SIBYLS, (Hura et al., 2009); G1, measured (this paper); BL 4-2, (Martel et al., 2012); ID02, assuming 200 mA ring current and sample path length 2 mm (ESRF ID02); BM29, assuming 1.8 mm path length (ESRF BM29); BL45XU (BL45XU); SAXS/WAXS, (SAXS/WAXS); P12, (Jeffries et al., 2015); I22, (Diamond); SWING, (SWING), path length from (David & Pérez, 2009).

Beamline	Flux (ph/s)	Beam size, FWHM (μm^2)	Energy (keV)	Sample path length (mm)	Dose rate (kGy/ s)	Adiabatic heating rate δT_{ad} (K/s)	Adiabatic heat diffusion time t_d (s)	Adiabatic heating ΔT_{ad} (K)	Steady state heating ΔT_{ss} (K)	Time to onset of convection t^* (s)
BioCAT (APS)	1.1×10^{13}	115×122	12	1.5*	180	41	0.024	1	2	15
SIBYLS (ALS)	10^{12}	4000×1000	12	1.5*	0.05	0.01	1.8	0.02	8×10^{-5}	83
G1 – BioSAXS (MacCHES S)	9.6×10^{11}	190×196	9.96	2.0	5.1	1.2	0.06	0.08	0.1	86
BL 4-2 (SSRL)	2×10^{12}	300×500	11	1.5	3.2	0.8	0.16	0.1	0.2	46
ID02 (ESRF)	2×10^{14}	200×400	12.4	2	519	120	0.07	9	15	2.4
BM29 (ESRF)	10^{12}	700×700	12.5	1.8	0.43	0.10	0.88	0.09	0.03	80
BL45XU (Spring8)	1×10^{12}	200×400	12.4	1.5*	2.8	0.7	0.07	0.05	0.08	77
SAXS/WAXS (Australian Synchrotron)	2×10^{13}	150×250	10	1.5	127	30	0.04	1	2	10
P12 (DESY)	4×10^{12}	110×200	10	1.7	39	9	0.02	0.2	0.5	31
I22 (Diamond)	6×10^{12}	320×80	12.4	1.5*	40	10	0.01	0.1	0.3	28
SWING (SOLEIL)	8×10^{12}	450×20	7	1.6	256	60	0.0007	0.04	0.2	16

Acknowledgements This paper benefitted from discussions with Andrea M. Katz, Steve P. Meisburger, Matthew E. Warkentin and Lois Pollack about radiation damage and SAXS. Matthew E. Warkentin and Jonah Haber helped collect preliminary data (not used here) that guided later data collection. Richard Gillilan is the staff scientist in charge of the MacCHESS BioSAXS user facility and aided with set up and data collection. He also provided the silver behenate calibrant. Arthur Woll is the beamline scientist for G-Line at CHESS and helped solve various instrumentation issues at the beamline. This work was funded by the NSF (DBI-1152348). This work is based upon research conducted at the Cornell High Energy Synchrotron Source (CHESS), which is supported by the National Science Foundation and the National Institutes of Health/National Institute of General Medical Sciences under NSF award DMR-0936384, using the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is supported by award GM-103485 from the National Institutes of Health, through its National Institute of General Medical Sciences. This work made use of the Nanobiotechnology Center shared research facilities at Cornell.

References (includes supporting information references)

- Acerbo, A. S., Cook, M. J., & Gillilan, R. E. (2015). *J. Synchrotron Radiat.* **22**, 180–186.
- Ambrosini, D., Paoletti, D., & Schirripa Spagnolo, G. (2003). *Int. J. Heat Mass Transf.* **46**, 4145–4155.
- Audette-Stuart, M., Houée-Levin, C., & Potier, M. (2005). *Radiat. Phys. Chem.* **72**, 301–306.
- Berger, M. J., Hubbell, J. H., Seltzer, S. M., Chang, J., Coursey, J. S., Sukumar, R., Zucker, D. S., & Olsen, K. (2010). XCOM: Photon Cross Sections Database (version 1.5).
- BL45XU BL45XU Outline.
- Blanchet, C. E., Spilotros, A., Schwemmer, F., Graewert, M. A., Kikhney, A., Jeffries, C. M., Franke, D., Mark, D., Zengerle, R., Cipriani, F., et al. (2015). *J. Appl. Crystallogr.* **48**, 431–443.
- Blanchet, C. E., Zozulya, A. V., Kikhney, A. G., Franke, D., Konarev, P. V., Shang, W., Klaering, R., Robrahn, B., Hermes, C., Cipriani, F., et al. (2012). *J. Appl. Crystallogr.* **45**, 489–495.
- Bobrowski, K. (2012). *Encyclopedia of Radicals in Chemistry, Biology, and Materials*, C. Chatgililoglu, & A. Studer, edited by, pp. 395–432. Hoboken, N.J.: John Wiley & Sons.
- Boyd, R. D. & Vest, C. M. (1975). *Appl. Phys. Lett.* **26**, 287–288.
- Butler, J. A. V., Robins, A. B., & Rotblat, J. (1960). *Proc. R. Soc. A Math. Phys. Sci.* **256**, 1–14.
- Classen, S., Hura, G. L., Holton, J. M., Rambo, R. P., Rodic, I., McGuire, P. J., Dyer, K., Hammel, M., Meigs, G., Frankel, K. A., et al. (2013). *J. Appl. Crystallogr.* **46**, 1–13.
- David, G. & Pérez, J. (2009). *J. Appl. Crystallogr.* **42**, 892–900.
- Davies, K. J. A. (1987). *J. Biol. Chem.* **262**, 9895–9901.
- Davies, K. J. A. & Delsignore, M. E. (1987). *J. Biol. Chem.* **262**, 9908–9913.

- Davies, K. J. A., Delsignore, M. E., & Lin, S. W. (1987). *J. Biol. Chem.* **262**, 9902–9907.
- Davies, M. J. (2012). *Encyclopedia of Radicals in Chemistry, Biology and Materials*, C. Chatgililoglu, & A. Studer, edited by, pp. 1425–1457. Hoboken, N.J.: John Wiley & Sons.
- Diamond Technical Specifications - I22 - Small Angle Scattering Beamline.
- Durchschlag, H., Hefferle, T., & Zipper, P. (2003). *Radiat. Phys. Chem.* **67**, 479–486.
- Durchschlag, H. & Zipper, P. (2007). *Radiat. Phys. Chem.* **76**, 1295–1301.
- Dyer, K. N., Hammel, M., Rambo, R. P., Tsutakawa, S. E., Rodic, I., Classen, S., Tainer, J. A., & Hura, G. L. (2014). *Structural Genomics: General Applications*, Y.W. Chen, edited by, pp. 245–258. Totowa, NJ: Humana Press.
- ESRF BM29 BM29 Beamline Specifications.
- ESRF ID02 ID02 Beamline Layout.
- Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I., & Craievich, A. F. (2009). *J. Appl. Crystallogr.* **43**, 101–109.
- Fischetti, R. F., Rodi, D. J., Mirza, A., Irving, T. C., Kondrashkina, E., & Makowski, L. (2003). *J. Synchrotron Radiat.* **10**, 398–404.
- Franke, D., Jeffries, C. M., & Svergun, D. I. (2015). *Nat. Methods.* **12**, 419–422.
- Franke, D., Kikhney, A. G., & Svergun, D. I. (2012). *Nucl. Instruments Methods Phys. Res. A Accel. Spectrometers, Detect. Assoc. Equip.* **689**, 52–59.
- Garrison, W. M. (1987). *Chem. Rev.* **87**, 381–398.
- Gerward, L., Guilbert, N., Bjorn Jensen, K., & Levring, H. (2001). *Radiat. Phys. Chem.* **60**, 23–24.
- Gillilan, R. E., Cook, M., Temnykh, G., Møller, M., & Nielsen, S. (2013). *ACA Transactions 2013*, pp. 40–50.
- Graceffa, R., Nobrega, R. P., Barrea, R. A., Kathuria, S. V., Chakravarthy, S., Bilsel, O., & Irving, T. C. (2013). *J. Synchrotron Radiat.* **20**, 820–825.
- Graewert, M. A. & Svergun, D. I. (2013). *Curr. Opin. Struct. Biol.* **23**, 748–754.
- Grant, T. D., Luft, J. R., Carter, L. G., Matsui, T., Weiss, T. M., Martel, A., & Snell, E. H. (2015). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**, 45–56.
- Holton, J. M. (2009). *J. Synchrotron Radiat.* **16**, 133–142.
- Homer, C., Cooper, L., & Gonzalez, A. (2011). *J. Synchrotron Radiat.* **18**, 338–345.
- Hopkins, J. B., Katz, A. M., Meisburger, S. P., Warkentin, M. A., Thorne, R. E., & Pollack, L. (2015). *J. Appl. Crystallogr.* **48**, 227–237.
- Huang, T. C., Toraya, H., Blanton, T. N., & Wu, Y. (1993). *J. Appl. Crystallogr.* **26**, 180–184.
- Hubbell, J. H. (1999). *Phys. Med. Biol.* **44**, R1–R22.
- Hubbell, J. H. (2006). *Phys. Med. Biol.* **51**, R245–R262.
- Hubbell, J. H. & Selter, S. M. (2004). Tables of X-ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients from 1 keV to 20 MeV for Elements Z=1 to 92 and 48 Additional Substances of Dosimetric Interest.
- Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole II, F. L., Tsutakawa, S. E., Jenney Jr, F. E., Classen, S., Frankel, K. A., Hopkins, R. C., et al. (2009). *Nat. Methods.* **6**, 606–612.
- Jeffries, C. M., Graewert, M. A., Svergun, D. I., & Blanchet, C. E. (2015). *J. Synchrotron Radiat.* **22**,

273–279.

Kirby, N. M., Mudie, S. T., Hawley, A. M., Cookson, D. J., Mertens, H. D. T., Cowieson, N., & Samardzic-Boban, V. (2013a). *J. Appl. Crystallogr.* **46**, 1670–1680.

Kirby, N., Mudie, S., Hawley, A., Mertens, H. D. T., Cowieson, N., Samardzic-Boban, V., Felzmann, U., Mudie, N., & Dwyer, J. (2013b). *ACA Trans.* **2013**, 27–39.

Kmetko, J., Hussein, N. S., Naides, M., Kalinin, Y., & Thorne, R. E. (2006). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 1030–1038.

Kmetko, J., Warkentin, M., Englich, U., & Thorne, R. E. (2011). *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **67**, 881–893.

Kozak, M. (2005). *J. Appl. Crystallogr.* **38**, 555–558.

Kozak, M. (2006). *Biopolymers*. **83**, 95–102.

Kriminski, S., Kazmierczak, M., & Thorne, R. E. (2003). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **59**, 697–708.

Krisch, R. E., Flick, M. B., & Trumbore, C. N. (1991). *Radiat. Res.* **126**, 251–259.

Kuwamoto, S., Akiyama, S., & Fujisawa, T. (2004). *J. Synchrotron Radiat.* **11**, 462–468.

Kuzay, T. M., Kazmierczak, M., & Hsieh, B. J. (2001). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57**, 69–81.

Lewis, W. K. & Whitman, W. G. (1924). *Ind. Eng. Chemist.* **16**, 1215–1220.

Lipfert, J., Millett, I. S., Seifert, S., & Doniach, S. (2006). *Rev. Sci. Instrum.* **77**, 046108.

De Maria Antolinos, A., Pernot, P., Brennich, M. E., Kieffer, J., Bowler, M. W., Delageniere, S., Ohlsson, S., Malbet Monaco, S., Ashton, A., Franke, D., et al. (2015). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**, 76–85.

Martel, A., Liu, P., Weiss, T. M., Niebuhr, M., & Tsuruta, H. (2012). *J. Synchrotron Radiat.* **19**, 431–434.

Meisburger, S. P., Warkentin, M., Chen, H., Hopkins, J. B., Gillilan, R. E., Pollack, L., & Thorne, R. E. (2013). *Biophys. J.* **104**, 227–236.

Murray, J. W., Rudiño-Piñera, E., Owen, R. L., Grininger, M., Ravelli, R. B. G., & Garman, E. F. (2005). *J. Synchrotron Radiat.* **12**, 268–275.

Mylonas, E. & Svergun, D. I. (2007). *J. Appl. Crystallogr.* **40**, s245–s249.

Nielsen, S. S., Møller, M., & Gillilan, R. E. (2012). *J. Appl. Crystallogr.* **45**, 213–223.

Nielsen, S. S., Noergaard Toft, K., Snakenborg, D., Jeppesen, M. G., Jacobsen, J. K., Vestergaard, B., Kutter, J. P., & Arleth, L. (2009). *J. Appl. Crystallogr.* **42**, 959–964.

Oberthuer, D., Melero-García, E., Dierks, K., Meyer, A., Betzel, C., Garcia-Caballero, A., & Gavira, J. A. (2012). *PLoS One*. **7**, e33545.

Paithankar, K. S. & Garman, E. F. (2010). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 381–388.

Paithankar, K. S., Owen, R. L., & Garman, E. F. (2009). *J. Synchrotron Radiat.* **16**, 152–162.

Parsons Jr., J. R. & Mulligan, J. C. (1978). *J. Heat Transfer*. **100**, 423–428.

Pernot, P., Theveneau, P., Giraud, T., Nogueira Fernandes, R., Nurizzo, D., Spruce, D., Surr, J., McSweeney, S., Round, A., Felisaz, F., et al. (2010). *J. Phys. Conf. Ser.* **247**, 012009.

Petoukhov, M. V., Konarev, P. V., Kikhney, A. G., & Svergun, D. I. (2007). *J. Appl. Crystallogr.* **40**, s223–s228.

- Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V., & Svergun, D. I. (2012). *J. Appl. Crystallogr.* **45**, 342–350.
- Pollack, L. (2011). *Biopolymers*. **95**, 543–549.
- Price, W. S., Tsuchiya, F., & Arata, Y. (1999). *J. Am. Chem. Soc.* **121**, 11503–11512.
- Rambo, R. P. & Tainer, J. A. (2011). *Biopolymers*. **95**, 559–571.
- Rambo, R. P. & Tainer, J. A. (2013). *Nature*. **496**, 477–481.
- Saha, A., Mandal, P. C., & Bhattacharyya, S. N. (1995). *Radiat. Phys. Chem.* **46**, 123–145.
- Sanishvili, R., Yoder, D. W., Pothineni, S. B., Rosenbaum, G., Xu, S., Vogt, S., Stepanov, S., Makarov, O. A., Corcoran, S., Benn, R., et al. (2011). *Proc. Natl. Acad. Sci. U. S. A.* **108**, 6127–6132.
- SAXS/WAXS SAXS Beamline Technical Specifications.
- Shimizu, N., Hirata, K., Hasegawa, K., Ueno, G., & Yamamoto, M. (2007). *J. Synchrotron Radiat.* **14**, 4–10.
- Skou, S., Gillilan, R. E., & Ando, N. (2014). *Nat. Protoc.* **9**, 1727–1739.
- SWING Technical Specifications.
- ToolBox, E. The Engineering ToolBox.
- Vest, C. M. & Lawson, M. L. (1972). *Int. J. Heat Mass Transf.* **15**, 1281–1283.
- Voronov, V. P. & Buleiko, V. M. (1998). *J. Exp. Theor. Phys.* **86**, 586–590.
- Warkentin, M., Badeau, R., Hopkins, J. B., Mulichak, A. M., Keefe, L. J., & Thorne, R. E. (2012). *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 124–133.
- Whitten, A. E., Cai, S., & Trehwella, J. (2008). *J. Appl. Crystallogr.* **41**, 222–226.
- Winterbourn, C. C. (2008). *Nat. Chem. Biol.* **4**, 278–286.
- Witala, M., Han, J., Menzel, A., & Nygård, K. (2014). *J. Appl. Crystallogr.* **47**, 2780–2080.
- Zeldin, O. B., Gerstel, M., & Garman, E. F. (2013). *J. Appl. Crystallogr.* **46**, 1225–1230.
- Zhang, F., Skoda, M. W. A., Jacobs, R. M. J., Martin, R. A., Martin, C. M., & Schreiber, F. (2007). *J. Phys. Chem. B.* **111**, 251–259.
- Zipper, P. & Durchschlag, H. (1980a). *Zeitschrift Für Naturforschung. Sect. C.* **35c**, 890–901.
- Zipper, P. & Durchschlag, H. (1980b). *Radiat. Environ. Biophys.* **18**, 99–121.
- Zipper, P. & Durchschlag, H. (1980c). *Monatshefte Für Chemie.* **111**, 1367–1390.
- Zipper, P. & Durchschlag, H. (1981). *Monatshefte Für Chemie.* **112**, 1–23.
- Zipper, P., Gatterer, H. G., Schurz, J., & Durchschlag, H. (1980). *Monatshefte Für Chemie.* **111**, 981–1004.
- Zipper, P. & Kriechbaum, M. (1986). *Monatshefte Für Chemie.* **117**, 557–572.
- Zipper, P., Wilfing, R., Kriechbaum, M., & Durchschlag, H. (1985). *Zeitschrift Für Naturforschung. Sect. C.* **40c**, 364–372.

Supporting information

S1. Methods

S1.1. Sample preparation

For this study, three different proteins were used: glucose isomerase from *Streptomyces rubiginosus* (Hampton Research, Aliso Viejo, CA, HR7-100), hen egg white lysozyme (Affymetrix, Santa Clara, CA, 18645), and xylanase from *Trichoderma longibrachiatum* (Hampton Research, HR7-104). Glucose isomerase was buffer exchanged in a spin column (Amicon 30 kDa MW cutoff, EMD Millipore, Billerica, MA) into a 100 mM pH 7.0 HEPES and 1 mM magnesium chloride buffer. Lysozyme was reconstituted from powder into a 40 mM sodium acetate pH 4.0, 50 mM sodium chloride buffer and filtered through a 0.22 μm syringe filter (Millex ethylene oxide sterilized, EMD Millipore) to remove undissolved powder. Xylanase was buffer exchanged in a spin column (Amicon 10 kDa MW cutoff, EMD Millipore) into a 50 mM pH 7.4 Tris buffer. Protein concentration in stock solutions and dilutions was measured using a NanoVue Plus Spectrophotometer (GE Healthcare Bio-Sciences, Pittsburgh, PA). The NanoVue has a built in method for determining lysozyme concentration, which was used. The extinction coefficients used for xylanase and glucose isomerase were $55900\text{ M}^{-1}\text{cm}^{-1}$ (Kozak, 2006) and $45660\text{ M}^{-1}\text{cm}^{-1}$ (Kozak, 2005) respectively. Lysozyme was prepared at 0.5, 1.0, 2.0, 4.1, 8.1, 15.5, 32.2, and 47.3 mg/mL, xylanase was prepared at 4.9 mg/mL, and glucose isomerase was prepared at 1.2 mg/mL.

Degassed samples were prepared in the same way, but prior to buffer exchanging or reconstitution the buffers were degassed under vacuum for 30 minutes. After preparation, individual aliquots of the sample and buffer were degassed under vacuum for 10 minutes, and sealed while under nitrogen gas (Airgas, 99.998% pure). These aliquots were subsequently kept sealed until immediately before being loaded into the sample cell.

S1.2. Beamline Setup

SAXS measurements were carried out at the G1 beamline at CHESS using the BioSAXS user facility run by MacCHESS (Nielsen et al., 2012; Acerbo et al., 2015). For these measurements the energy (wavelength) was 9.96 keV (1.25 $\text{\AA}$), and the X-ray path length through the sample cell was 2.0 mm. The sample-to-detector distance was 1506 mm, found using silver behenate powder ($d = 58.38\text{ \AA}$ (Huang et al., 1993)) (The Gem Dugout, State College, PA). SAXS data was collected on a Pilatus 100 K detector (Dectris, Baden, Switzerland). The useful q range ($q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle and λ is the incident X-ray wavelength) was $0.0098\text{ \AA}^{-1}$ to $0.2822\text{ \AA}^{-1}$ (range in figures is slightly reduced by binning). While the setup was capable of oscillation to reduce damage, measurements were carried out in a static mode, similar to those in (Jeffries et al., 2015). Exposures were collected in a shutterless mode with individual image exposure times from 0.03 to 1 s and total

exposure times per sample of 2.25 to 120 s, depending on the protein. The following protocol was used to ensure good buffer matching and that there was no damage to (or protein adsorption on) the sample cell: measure buffer, clean sample cell, measure empty sample cell, measure protein, clean sample cell, measure empty sample cell. If two subsequent empty measurements disagreed, the sample cell was changed and the data retaken. In order to characterize any variability in the damage rate from nominally identical samples, at least three experiments were carried out for every sample condition. Unless otherwise indicated, measurements were carried out at 4 °C.

In order to accurately calculate the absorbed X-ray dose, the X-ray flux and beam size/shape at the sample position were measured. Using a pure vacuum flight path (no sample cell) the detector and beamstop were removed and a nitrogen ion chamber placed at the end of the downstream flight tube. The flux was calculated, accounting for the transmission of the flight tube window (Mylar, 0.97 transmission) and of one sample cell window (25 µm polystyrene, 0.995 transmission). The reference incident flux on the sample was 9.60×10^{11} ph/s (larger than that reported in (Acerbo et al., 2015) because of the recent upgrade to the undulator insertion device for the beamline). This provided a calibration value for measurements taken by the upstream ion chamber and the beamstop, allowing adjustments for changing ring current and any other effects that altered the incident intensity. To measure the beam size a knife edge at the sample position was scanned through the beam horizontally and vertically. These profiles are shown in Figures S1 and S2. With beam defining slits set at $250 \times 250 \mu\text{m}^2$ the FWHM of the beam was $190 \text{ (H)} \times 196 \text{ (V)} \mu\text{m}^2$.

Initial data processing at the beamline was performed using BioXTAS RAW (Nielsen et al., 2009).

S1.3. Data processing and metrics for radiation damage

Damage to proteins may manifest in a number of ways, including aggregation, fragmentation, conformational changes, and unfolding. Increasingly, high throughput SAXS beamlines include automated methods for assessing when radiation damage occurs (Pernot et al., 2010; Blanchet et al., 2015; De Maria Antolinos et al., 2015). These methods use a variety of statistical techniques to determine when the measured scattering profile has significantly changed relative to the initial exposure (Franke et al., 2012, 2015; Grant et al., 2015). While practical for assessing the onset of radiation damage, these do not give data that can be easily interpreted as to type and rate of damage. Previously, metrics based on the change in the radius of gyration, the pseudo radius of gyration (described below), and the scattering intensity at zero angle were used to study radiation damage (Kuwamoto et al., 2004; Jeffries et al., 2015). We evaluated these and additional parameters to assess changes in the scattering profile and the protein structure. All of the parameters can be calculated from a scattering profile without any external calibration.

To efficiently process the large data sets generated, a custom Python program automated calibration, masking, integration, normalization, background subtraction, and calculation of dose for each image. The program also automated calculation of the following parameters for every scattering profile:

Integrated intensity of the scattering profile – The integrated intensity was calculated by direct integration of the scattering profile.

Porod Invariant – The Porod invariant (Q) was calculated by direct integration of the q^2 weighted scattering profile.

Radius of gyration – Radius of gyration (R_g) was calculated two ways: by AUTORG and DATGNOM from the ATSAS suite (Petoukhov et al., 2007, 2012). While AUTORG also ‘quality’ of the data (which is not defined in the manual) and the q range used, this information was not used.

Pseudo R_g and initial damage rate – The pseudo R_g (R_g^{ps}) defined in (Jeffries et al., 2015) finding a known literature value from the R_g , R_g^u , and then calculating the R_g^{ps} for each subsequent image using a Guinier fit in the q range $0.8 < qR_g^u < 1.3$. The literature values are given below. In general, if a literature value for the R_g is not available, the value from the first image of the data set could be substituted to achieve the same effect from the calculation. The calculated R_g^{ps} was then used to find the initial damage rate $\Delta R_g^{ps} \text{ Gy}^{-1}$, defined as the change in R_g^{ps} over the first 10 kGy of dose. In (Jeffries et al., 2015) $\Delta R_g^{ps} \text{ s}^{-1}$ was defined as the change in R_g^{ps} for the first 5 images of the data set. Since dose is a more appropriate metric for comparison between different beamlines, we estimated their dose rate as 40 kGy/s (Gaussian beam assumed), so 5 images at their standard exposure time of 50 ms is ~ 10 kGy.

Maximum dimension – The maximum dimension (D_{max}) of the protein was obtained from DATGNOM.

$I(0)$ – The scattering intensity at zero angle ($I(0)$) was obtained from AUTORG and from DATGNOM.

Porod volume – The Porod volume is nominally the particle volume, though it is often an underestimate (Rambo & Tainer, 2011), and was calculated as $V = 2\pi I(0)/Q$ and from DATPOROD.

Molecular weight – The molecular weight was calculated from the Porod volume using the method of (Fischer et al., 2009) and was separately calculated using the method of (Rambo & Tainer, 2013).

Molecular weight can also be determined from $I(0)$ by scaling by a measured constant from a standard. However, since $I(0)$ is already included as a parameter there is no need to include that as an alternative determination of molecular weight.

As in (Kuwamoto et al., 2004) we normalized each parameter for each scattering profile by an initial value. This normalization presumes that, for example, a 1 Å change in R_g is more significant for lysozyme ($R_g = 14.3$ Å) than glucose isomerase ($R_g = 32.7$ Å). The automated calculations carried out above can yield relatively large uncertainties. To minimize the effect of random variation in the initial value, known literature values were used for normalization when available (R_g : 14.3 Å for lysozyme (Mylonas & Svergun, 2007), 32.7 Å for glucose isomerase (Kozak, 2005), 17.2 Å for xylanase (Kozak, 2006), molecular weight: 14.3 kDa for lysozyme, 172 kDa for glucose isomerase (Kozak, 2005), 21 kDa for xylanase (Kozak, 2006)); otherwise the initial measured value was used. The R_g values given here are also those used as R_g^u in the calculation of R_g^{ps} .

The parameters were observed to vary seemingly linearly with dose at low doses. For a parameter P we fit this region using a linear fit, and used the slope (in % change per kGy (%/kGy)) as the corresponding metric S_p of radiation sensitivity. Reported sensitivities and standard deviations are the average and standard deviation of the sensitivities of all identically prepared samples measured.

It is important to note that instrumental stability is key to successful quantification of radiation damage parameters. Changes in scattering from other, non-sample, sources can be misinterpreted as sample changes. Shifts in overall incident beam intensity and intensity profile within the beam can occur from upstream drifts and shifts, and these must be detected and properly removed by normalization. Intensity changes with q can result from, e.g., radiation-induced fouling and other changes in window materials. In our experiments, the acquisition of buffer and empty sample cell scattering profiles before and after each radiation damage experiment allowed us to measure the stability of our data collection. Generally, scattering profiles of empty cells and buffers taken before and after data collection subtracted to zero (within experimental uncertainty). When differences were seen, they were characteristic of fouling of the windows of the sample cell with damaged protein; the sample cell was changed and the data retaken.

S2. The attenuation coefficient

The dose calculation also requires as input the fraction of the incident energy that is absorbed, A . In principle this can be calculated from Beer's law:

$$A = 1 - \exp(-\mu l) \quad (\text{S1})$$

where μ is the absorption coefficient and l is the sample path length. This is the approach used in macromolecular crystallography (Kmetko et al., 2006). The main complication is determining the proper absorption coefficient to use. As μ depends on both the sample composition and density, which can vary for a number of reasons, values are typically tabulated for μ / ρ , the mass attenuation coefficient (Hubbell, 2006). Different absorption coefficients are calculated by assuming that some of the total absorbed energy escapes due to various physical mechanisms, such as fluorescent photon emission.

The mass attenuation coefficient μ / ρ represents every process that contributes to attenuation of the incident intensity. This includes processes such as coherent scattering that do not contribute to the absorbed energy. In order to calculate the absorption, a variety of approximate coefficients are calculated, which include some but not all of the attenuation processes (Hubbell, 1999). The important X-ray interaction processes at the energies typically used for SAXS (~5–15 keV) are elastic (coherent, Rayleigh/Thomson), inelastic (incoherent, Compton), and the atomic photoeffect absorption (Hubbell, 1999; Paithankar et al., 2009).

Elastic scattering does not contribute to the absorbed dose. Inelastic scattering can deposit some energy in the sample. However, in macromolecular crystallography this energy deposition has been approximated, and shown to be negligible for energies less than ~20 keV (Paithankar & Garman, 2010). For example, at 15 keV, the calculated contribution of inelastic scattering to total dose for a lysozyme crystal was 0.25%. Based on this, we will ignore the energy contribution by inelastic scattering.

The atomic photoeffect is the direct absorption of all of the energy of the incident photon by an atom (Hubbell, 1999, 2006), leading to the ejection of a photoelectron from the atom. Energy deposited by photoelectric absorption can be lost via secondary photon emission. One emission source that has been well characterized in the crystallography community is emission of fluorescent photons from the primary absorption event. The atom is left with an inner shell electron vacancy, and this is filled by the decay of an outer shell electron, which emits an Auger electron or a fluorescent X-ray (Paithankar et al., 2009). Energy can also be lost by radiation of secondary electrons slowing in the medium (such as bremsstrahlung processes), further cascade of fluorescent emissions after ionizing events in atomic subshells, and other processes (Hubbell, 1999). The mass energy absorption coefficient μ_{en} / ρ represents the energy absorbed if all of the energy emitted in secondary photons is lost from the sample.

The important question for calculating dose is how many of these secondary photons leave the ‘volume of interest.’ This is a nebulous concept for SAXS, as damage products can diffuse. Thus, the volume of interest is not simply the illuminated volume. Using the photoelectric absorption assumes

that none of these secondary photons escape, while μ_{en}/ρ assumes that all of them escape. The photoelectric absorption can be calculated for a mixture using XCOM (Gerward et al., 2001; Berger et al., 2010). μ_{en}/ρ for most elements and certain compounds are tabulated (Hubbell & Selter, 2004), and can be calculated approximately (but not exactly, see (Hubbell, 1999)) by summing atomic μ_{en}/ρ weighted by the mass fraction of the atomic constituents (Kmetko et al., 2006).

The proper absorption coefficient for calculating dose in SAXS is likely somewhere between the mass-photoelectric absorption coefficient μ_{pe}/ρ and the mass-energy absorption coefficient μ_{en}/ρ . The details of determining the balance of energy loss in the sample are beyond this paper. Fortunately, for water at 10 keV, the loss mechanisms that would lead to a difference between μ_{pe}/ρ and μ_{en}/ρ are negligible, and both equal 4.944 cm²/g. The addition of buffer components and protein changes μ_{pe}/ρ slightly (5.008 cm²/g for 47.5 mg/mL lysozyme in our buffer), and may change μ_{en}/ρ incommensurately. However, even if μ_{en}/ρ for a protein solution remained unchanged from that of water, the difference between it and the photoelectric absorption coefficient would be $\leq 1\%$. So, at least at energies near 10 keV, the choice of either photoelectric or mass energy absorption coefficient does not matter.

The photoelectric absorption coefficient is easily calculated using XCOM (Gerward et al., 2001; Berger et al., 2010). Thus in this work, we will determine A as

$$A = 1 - \exp((\mu_{pe}/\rho)(l\rho)). \quad (\text{S2})$$

In this work, μ_{pe}/ρ was calculated for each sample condition.

If the path length of the sample is unknown, for example for the windowless sample holder used in (Meisburger et al., 2013), the transmission can be used to approximate the path length as

$l \approx \mu^{-1} \ln(1/T)$. This can be used with Equation (S2) to calculate A , or the transmission itself can be adjusted according to the difference between the mass attenuation and photoelectric absorption coefficient as

$$A_{meas} = 1 - T_{meas} \exp[(\mu/\rho - \mu_{pe}/\rho)\rho l_m]. \quad (\text{S3})$$

Previously, Beer's law was used by (Kuwamoto et al., 2004; Jeffries et al., 2015), while (Meisburger et al., 2013) measured the transmission T and estimated $A = 1 - T$. In (Kuwamoto et al., 2004), the thin sample approximation to Beer's law, typically used in crystallography (Kmetko et al., 2006), was used, which gives

$$A = 1 - \exp(-(\mu/\rho)(l\rho)) \approx (\mu/\rho)(l\rho). \quad (\text{S4})$$

This assumes $(\mu/\rho)(l\rho)$ is small, but in SAXS sample path lengths tend to be chosen to be one attenuation length (Nielsen et al., 2012). This lead to an overestimation of the absorbed dose by 27% for the given beam parameters (Kuwamoto et al., 2004). In (Jeffries et al., 2015), the exponential form of Beer's law was used. However, they use an absorption coefficient equal to the photoelectric absorption plus the incoherent scattering. For water at 10 keV this overestimates the absorption coefficient by a factor of $\sim 3\%$, and the dose by $\sim 1\%$, so it is a minimal correction. At 10 keV, for water, the use of measured transmission in (Meisburger et al., 2013)) lead to an overestimate of the dose by $\sim 4\%$, again a minor correction.

S3. The diffusion correction

Over the course of an experiment, damaged protein will diffuse out of the illuminated volume and undamaged protein will diffuse in. In order to compensate for this, we made the following assumption: the effective dose is the dose absorbed by solution while a macromolecule is in the beam, averaged over all macromolecules in the beam. For example, if 50% of the sample has been in the beam for the entire experiment and 50% is fresh, the effective dose would be half of the nominally expected dose. This was calculated as follows. First, a source term was introduced to the diffusion equation as

$$\frac{\partial F(t,x,y,z)}{\partial t} = D \left(\frac{\partial^2 F(t,x,y,z)}{\partial x^2} + \frac{\partial^2 F(t,x,y,z)}{\partial y^2} + \frac{\partial^2 F(t,x,y,z)}{\partial z^2} \right) + B(t,x,y,z) \quad (\text{S5})$$

where $F(t,x,y,z)$ is the fluence ($\text{ph}/\mu\text{m}^2$) (without any dose correction, equal to $\dot{f}t$ from Equation (1)), and is directly proportional to the effective absorbed dose; D is the diffusion coefficient of the macromolecule; and $B(t,x,y,z)$ is the incident beam flux density. While this allows for a time dependency to the beam, the incident intensity was stable so this was ignored. One additional major simplification was made, which was to ignore the z dimension (the beam direction). While the intensity should decay in the sample as $\exp(-(\mu/\rho)(\rho z))$, including z was too computationally intensive. Additionally, to speed up the calculation a Gaussian beam profile in x and y was used Equation (3) from the main paper, was used. These assumptions give

$$\frac{\partial F(t,x,y)}{\partial t} = D \left(\frac{\partial^2 F(t,x,y)}{\partial x^2} + \frac{\partial^2 F(t,x,y)}{\partial y^2} \right) + B(x,y). \quad (\text{S6})$$

Equation (S6) was numerically solved using Mathematica, with hard edge boundary conditions such that $F(t, x_{\text{edge}} \int y_{\text{edge}}) = 0$ and with the initial condition $F(0,x,y) = 0$. The calculation was also carried out with periodic boundary conditions and yielded similar results, so the choice of boundary condition was assumed to have no significant effect.

A unitary amplitude was used, $f_0 = 1$, allowing the computation of a general correction function, $C(t)$, for a particular diffusion coefficient and beam size. To calculate $C(t)$, first the weighted average of $F(t, x, y)$ with the beam profile as the weight was calculated,

$$W(t) = \frac{\iint B(t, x, y) F(t, x, y) dx dy}{\iint B(t, x, y) dx dy} . \quad (\text{S7})$$

Here $W(t)$ is the weighted fluence. The nominal, diffusion free, weighted dose at time t is simply $W_0(t) = ft$. The correction factor is the ratio of $W(t)$ and $W_0(t)$,

$$C(t) = \frac{W(t)}{W_0(t)} . \quad (\text{S8})$$

Figure 2 in the main paper shows a plot of the correction factor calculated for lysozyme and a set of beam sizes. Smaller beams have faster turnover of sample, so for a fixed exposure time the correction is more important the smaller the beam dimensions.

Correction factors were calculated for each of the three proteins. Diffusion coefficients were measured by dynamic light scattering (Malvern Zetasizer Nano ZS, Malvern Instruments, Worcestershire, UK) to be: lysozyme, $1.25 \times 10^{-6} \text{ cm}^2/\text{s}$ (literature $1.13 \times 10^{-6} \text{ cm}^2/\text{s}$ (Price et al., 1999)); glucose isomerase, $0.60 \times 10^{-6} \text{ cm}^2/\text{s}$ (literature $0.53 \times 10^{-6} \text{ cm}^2/\text{s}$ (Oberthuer et al., 2012)); xylanase, $1.01 \times 10^{-6} \text{ cm}^2/\text{s}$. Note that this calculation does not account for the fact that damaged products may have different diffusion coefficients than the undamaged protein.

The photoelectrons and free radicals generated by the X-ray beam will also diffuse. Experiments on protein crystals at 100 K have found that using a $\sim 1 \text{ }\mu\text{m}$ beam instead of a $\sim 16 \text{ }\mu\text{m}$ beam reduced the damage rate by a factor of three, which was attributed to photoelectron escape from the illuminated volume (Sanishvili et al., 2011). At room temperature free radical diffusion lengths have been estimated to vary from $< 0.2 \text{ }\mu\text{m}$ to 1.5 mm , depending on the radical and solution condition (Winterbourn, 2008; Davies, 2012), so diffusion of radicals out of the beam may also reduce damage rates.

S4. Effect of radiation on lysozyme, xylanase, and glucose isomerase

S4.1. Measured Sensitivities

Scattering profiles at several doses are shown in Figure 1 of the main paper for lysozyme (4.1 mg/mL), xylanase (4.9 mg/mL), and glucose isomerase (1.2 mg/mL) respectively. Lysozyme shows the previously reported form of radiation damage in SAXS, a strong increase in scattered intensity at

low q with increasing dose (Kuwamoto et al., 2004; Jeffries et al., 2015). Qualitatively the same behavior is also seen for xylanase. This type of increase at low q is commonly considered to be due to aggregation of the sample. The scattering profiles of glucose isomerase did not show this same increase at low q with increasing dose. Only a small change in the scattering profiles was observed, a slight downturn at low q at large dose. Closer analysis showed that the profile decreased from the lowest q until $q \sim 0.1 \text{ \AA}^{-1}$. Each sample condition (a given protein, concentration, temperature, and degassed state) was measured for at least 3 nominally identical samples, leading to over 50 measured sets of scattering profiles. Each set contained 50–150 scattering profiles, depending on the protein, so not all collected scattering profiles could be shown.

The set of parameters given in Section S1 were calculated for each individual scattering profile, normalized, and plotted as a function of dose for each sample. We call these plots vs. dose ‘dose curves.’ Figure 3 from the main paper shows representative dose curves of parameters calculated from the scattering profiles of lysozyme, xylanase, and glucose isomerase. For lysozyme, all of the parameters except the Porod invariant increased with dose, though the magnitude of the increase depended on the parameter. For the integrated intensity, radius of gyration, and maximum dimension the increase was linear at all doses, while for the molecular weight, Porod volume, and $I(0)$ there was a linear region at low dose. Looking at the plot, we can tell that the normalized rate of change in molecular weight was strongly correlated to that of the Porod volume and $I(0)$, which makes sense as both the volume and $I(0)$ should be proportional to the molecular weight. Likewise, visual inspection shows the normalized rate of change in the radius of gyration and the maximum dimension were strongly correlated, and both measure the size of particles in solution.

The dose curves for xylanase show different behavior from that of lysozyme. Every parameter except the Porod invariant had an initial roughly linear increase until a dose of 60–80 kGy, followed by either a slowing or levelling off of the change in parameter value. In this case, from just the plot it is not clear if the parameters that showed strongly correlated changes for lysozyme are also correlated for xylanase.

Finally, most of the dose curves for glucose isomerase showed little if any change. The exceptions to this were the Porod invariant, the Porod volume, and the molecular weight. The Porod invariant increased for the first ~ 75 kGy, then decreased until ~ 200 kGy, where it levelled off, while the reverse is seen for the molecular weight and Porod volume. The molecular weight and Porod volume decreased to a minimum of $\sim 75\%$ of their starting values, and levelled out near 90% of their initial values. A decrease until ~ 200 kGy and levelling out was seen in the integrated intensity, but for a much smaller fractional change.

Data was also collected for a range of lysozyme concentrations from 0.5 to 47.3 mg/mL. Dose curves for R_g at three different concentrations are shown in Figure S6. This plot demonstrates two things. First, for lower concentrations the dose curves tended to be noisier than at higher concentrations, particularly for any parameter not calculated by direct integration of the scattering profile. Second, at higher lysozyme concentrations there was an initial plateau in the dose curves for R_g . This plateau was observed at concentrations of 15.5 mg/mL and above, while at 8.1 mg/mL a deviation from linearity was seen in the R_g dose curve at low doses. A similar plateau or deviation from linearity was seen for all of the parameters at these concentrations, but it was most distinct in the R_g dose curves.

The rate of radiation damage for all samples was quantified by a fit to the linear region of the dose curve, as described in Section 7.2 (main paper). The slope of this fit for a parameter P is the sensitivity S_P . When discussing sensitivities, we will discuss the average and standard deviation of the sensitivity for a parameter from each set of nominally identical samples. In Table 1 (main paper) we reported the R_g sensitivity S_{rg} , the molecular weight sensitivity S_{mw} , and the integrated intensity sensitivity S_I and standard deviations σ_{rg} , σ_{mw} , and σ_I for three samples. Sensitivities and standard deviations for all conditions measured are reported in Table S1 and clearly show that there can be large sample-to-sample variations in damage rate for identically prepared samples.

S4.2. Differences between proteins

The type and rate of damage progression depends on the macromolecule. The classic form of radiation damage in SAXS is shown by both lysozyme and xylanase. For these proteins, radius of gyration, molecular weight, maximum dimension, and Porod volume all increase as a function of dose. This is consistent with aggregation being the dominant damage mode detected by SAXS for these proteins, as has previously been reported for lysozyme (Durchschlag et al., 2003; Kuwamoto et al., 2004; Jeffries et al., 2015). However, for similar concentrations of protein, the ratio of S_{rg} for lysozyme (4.1 mg/mL) to xylanase is 48. This indicates that xylanase damages much more slowly than lysozyme, something that is visible in Figures 1 and 3 (main paper). The ratios of S_{mw} and S_I for lysozyme to xylanase are 121 and 52, showing the same behavior as the ratio of S_{rg} .

Damage rates for lysozyme have been previously reported in the literature (Kuwamoto et al., 2004; Jeffries et al., 2015). We estimated a dose rate of 40 kGy/s, accounting for beam shape, for (Jeffries et al., 2015). Using this, we converted the initial damage rate of the pseudo radius of gyration, ΔR_g^{ps} , from nm/s to Å/kGy, giving 24, 22, and 19 Å/kGy at 2.2, 4.4 and 8.8 mg/mL concentrations

respectively (for values read off of their Figure S2). We measured values of 11.6, 9.4 and 7.3 for 2.0, 4.1, and 8.1 mg/mL concentrations, respectively. Differences in these values could easily be caused by our estimation of the dose rate. The differences in the ratios of the values may be due to differences in buffer composition. For (Kuwamoto et al., 2004), we assumed the dose rate was actually 32% of the reported value (a reduction by a factor of 1.37 from the thin sample approximation, and 2.27 from beam shape). From their Figure 1, we estimated $S_{rg} = 34$ %/kGy for a lysozyme concentration of 4.9 mg/mL. Given the approximations made in calculating this value, and difference in buffer and dose rate, this is in reasonable agreement with our reported $S_{rg} = 21.6$ %/kGy at 4.1 mg/mL.

Glucose isomerase shows a different type of damage, evidenced in the scattering profiles by a decrease in intensity at low q . The measured R_g and D_{max} decrease very slightly, $S_{rg} = -4.6 \times 10^{-5}$ %/kGy. A larger decrease is seen in molecular weight, $S_{mw} = -0.10$ %/kGy, and the Porod volume, as seen in Figure 3 (main paper). One possibility is that the protein is getting smaller. An alternative possibility is that we are seeing the effects of charging on the macromolecules, as seen when changing ionic strength in solution (Zhang et al., 2007).

Considering the possibility of a reduction in size, since glucose isomerase is a tetramer in solution, the most likely explanation is that the tetramer is being broken up into subunits. A decrease in the rigidity of the relative positions of the remaining subunits could explain the relatively stable average size. A closer examination of the molecular weight provides additional information. The initial calculated molecular weight was ~160 kDa (172 kDa expected), which decreased to a minimum of ~140 kDa after ~75 kGy of dose. The mass then increased to the average mass at maximum dose of ~150 kDa. The other glucose isomerase samples showed similar decreases, though the change in the molecular weight was not quite as large (the change in the scattering profile is consistent, see Figure S7). The relatively flat higher dose region may represent a steady state of tetramer dissolution and recombination. If true, changing the dose rate would change the average mass in this steady state. We have insufficient information to speculate on the possible components in solution.

Charging could come from redox reactions of side chains or main chain of the protein with X-ray produced radical products. However, at least in the initial (and most well understood) steps of many of the more common radical reactions, the protein tends to remain neutrally charged (Garrison, 1987; Davies, 2012). It may be that charge is generated in the further cascade of secondary reactions. The exact biochemical mechanism is beyond the scope of this paper. Further experiments would be needed to distinguish between charging and an actual size reduction, such as whether the observed sensitivities depend strongly on ionic strength of the buffer.

In addition to the dominant observed behavior being a decrease at low q rather than an increase, the damage rate for glucose isomerase is also much smaller than that of either lysozyme or xylanase. For 1.0 mg/mL lysozyme, $S_{rg}^{lys} / S_{rg}^{gi} = -532000$, $S_{mw}^{lys} / S_{mw}^{gi} = -504$, and $S_I^{lys} / S_I^{gi} = -907$. The dramatic difference between these ratios is due to the fact that $S_{rg}^{gi} \ll S_{mw}^{gi}, S_I^{gi}$, in contrast to the results for lysozyme and xylanase. This is a good reason to use multiple metrics, as different types of radiation induced changes will manifest more quickly in different metrics. The robustness of glucose isomerase has been previously noted and attributed at least in part to the absence of disulfide bridges (Kozak, 2005), and is well known in the community.

Radiation damage results were reported for glucose isomerase by (Jeffries et al., 2015 and supporting information). Scattering profiles were not shown, and the only dose dependent information available are R_g plots for the first ~ 10 kGy of dose. From our data, we can see that very little change is expected at a dose of 10 kGy, and that is reflected in their data. They state that R_g is increasing for two of their three experimental concentrations, in contrast to our observations of an overall decrease. However, any change in R_g is not convincingly correlated to dose in their Figure S2a. Without data at higher doses, without error bars for the reported R_g values, and without characterization of the variability in damage rates by reporting results from multiple samples, it is impossible to say if their experiments show the same behavior as ours. All of their observations are also at higher concentrations and dose rates, and in different buffers than were used here, so that may explain differences in the results.

Overall, the damage observed fits well with the expected damage types, in particular aggregation and change in oligomeric state and/or fragmentation. The exact nature of damage to a macromolecule in solution will depend upon the details of the macromolecule, including presence of multiple domains or subunits, and number and type of exposed surface residues. These variables, and others, change the consequences of radical damage in ways that are not fully understood (Davies, 2012). Thus, prediction of the expected mode(s) of damage and rate of damage of a particular macromolecule is not currently possible.

S4.3. Effect of concentration

It has previously been observed, using SAXS and other methods, that the rate of damage for proteins in aqueous solutions decreases with increasing protein concentration (Butler et al., 1960; Kuwamoto et al., 2004; Jeffries et al., 2015). We have tested this over two decades of lysozyme concentration, from 0.5 to 47.3 mg/mL. Figure S8 shows a plot of S_{rg} , S_{mw} , and S_I vs. concentration, and there are two clear regions. At low concentrations, below 4.1 or 2.0 mg/mL (for S_{rg} , and S_{mw} and S_I

respectively), the sensitivity is either constant or slowly increasing. Due to the large standard deviations at low concentration, and the differences in the three metrics, we cannot reliably distinguish between these two options. At higher concentrations the sensitivity decreases. At all concentrations studied, most of the dose is delivered to the solvent, so most radicals will be generated in solvent. A given dose will therefore generate a (relatively) fixed number of free radicals, which means there is an upper limit to the concentration of protein that can be damaged by a given dose. At concentrations below this value all protein molecules are damaged. As concentrations increase beyond this maximum, the fraction of molecules damaged for a given dose decreases, even though the total number remains constant, so the measured sensitivity decreases. This drop in sensitivity has previously been reported for lysozyme, over a smaller concentration range, and other proteins (Kuwamoto et al., 2004; Jeffries et al., 2015). (Kuwamoto et al., 2004) suggest that the damage rate times the concentration, c , should be constant in this region. This is only true if the same type and number of damaged species are being generated in solution at every concentration. $S_{rg}c$, $S_{mw}c$, and S_Ic are not constant for our data, Figure S9, in contrast to their results for lysozyme between 10 and 20 mg/mL.

A plateau in sensitivity at lower concentrations follows from the model given above. If the concentration of protein is such that all of the available molecules are being damaged, then the sensitivity will not change with changing concentration. A decrease in sensitivity at lower concentrations could be due to either increased diffusion lengths for free radicals to react with proteins leading to more radical recombination in the bulk solvent, or to longer diffusion times for proteins to react with each other to create the damaged species. Our data is not clear on whether the low concentration region is decreasing or plateauing, so we cannot distinguish between these possibilities. The model also explains the delayed onset of damage seen in the higher concentration lysozyme in Figure S6. There is some minimum fraction of the population that must be damaged before damage is observable via SAXS. As the concentration goes up, this minimum fraction takes more dose to reach.

S4.4. Other parameters affecting damage rates in SAXS

We also investigated the effect of temperature and degassing on damage rates. As previously observed, there is no significant effect of temperature variation near room temperature (Kuwamoto et al., 2004; Jeffries et al., 2015). It is sometimes mentioned that a deoxygenated environment or removal of dissolved oxygen from the solution can reduce radiation damage in SAXS (Hura et al., 2009). Deoxygenating, or, more generally, degassing the solution, is also done to reduce the chance of bubble formation upon oscillation/flow, or to prevent the dissolved gasses from being forced out of solution when exposed to the X-ray beam (Kirby et al., 2013b). We measured degassed samples of lysozyme and xylanase under standard atmospheric conditions. Our degassed samples damaged

slightly faster than the normal samples. However, due to the large standard deviations involved, we cannot conclusively say there was any effect from degassing.

There are two possible explanations. Either, degassing has only a small effect on the damage rates for these proteins or the samples were not fully degassed. Section S7 shows that the samples should stay deoxygenated on the timescale of the experiments, a couple of minutes. Some oxygen may have been introduced when the samples were sealed after being degassed. Later handling, such as pipetting into the sample cell, may also have added oxygen to the solution.

Dissolved molecular oxygen in solution has been observed to both increase and decrease radiation sensitivity of macromolecules (Saha et al., 1995). The assumed mechanism for sensitization is generation of superoxide radicals, $\text{O}_2^{\bullet-}$, and singlet oxygen (Garrison, 1987), and reaction of these species with the macromolecules (Davies, 1987; Davies & Delsignore, 1987; Davies et al., 1987). Sensitization by a factor of ~2-3 has previously been observed (Saha et al., 1995). When dissolved oxygen acts as a desensitizer, it is assumed to be due to scavenging of H^+ and e_{aq}^- by the oxygen in systems that are more sensitive to those species than the generated superoxide radicals (Garrison, 1987; Saha et al., 1995).

S5. Correlation between metrics

Given the large number of possible parameters P , we would like to identify a minimal set that accurately captures the diverse radiation responses of biomolecules in SAXS. To do this, we calculated the Pearson product-moment correlation coefficient (Pearson's r) between every parameter P and for each SAXS data set, and then averaged the r values over all identically prepared samples. Example plots of these correlation coefficients for lysozyme, xylanase, and glucose isomerase are shown in Figures S3-S5. Some analysis is given in the main body of the paper, Section 7.3.

The correlations in the parameters for lysozyme were affected by lysozyme concentration. At lower concentrations, the same correlations were seen, but less strongly, likely do to the larger variation in most parameters noted previously. At 32.2 and 47.3 mg/mL the values for parameters derived from DATGNOM became increasingly uncorrelated ($|r| \lesssim 0.3$) with the other parameters, even when nominally the same parameter is being calculated in two ways, such as the R_g . This may be because the large structure factor at the higher concentrations had a different effect on the DATGNOM calculations than on the other methods.

For $I(0)$, R_g , Porod volume, and molecular weight, the parameter was determined multiple ways.

The question of which approach gives the most accurate values is beyond the scope of this work. However, it is important to understand if there is a difference in the dose response for different methods of calculation. For lysozyme, different methods of determining a given parameter yielded

highly correlated results ($r \geq 0.95$). For xylanase the dual determination of R_g ($r = 0.77$), $I(0)$ ($r = 0.93$) and molecular weight ($r = 0.96$) were strongly correlated. For glucose isomerase, only the different methods of calculating $I(0)$ ($r = 0.68$) and molecular weight ($r = 0.94$) were strongly correlated. For neither xylanase nor glucose isomerase was the Porod volume based on direct integration of Q well correlated with the value from DATPOROD.

Here we report the R_g and $I(0)$ from AUTORG, as the method of calculation is straightforward, and seemed to give more reasonable estimates of the uncertainty. It appeared that the uncertainty reported for these values from DATGNOM only accounted for the uncertainty in the calculation of the parameters from $p(r)$, not the uncertainty in the determination of $p(r)$ itself. We report the Porod volume based on direct integration of Q , but that choice was arbitrary.

In addition to calculating correlation coefficients, principle component analysis was performed on the parameters as a function of dose. While this is a ‘cleaner’ way to obtain information about the number of non-correlated components in the system, because it returns the eigenvectors of the covariance matrix it is not as simple to relate the principle components to the measured variables. For lysozyme, there was generally one axis whose eigenvalue was one or more orders of magnitude greater than any other, and this corresponded to a principle component axis which contained roughly equal contributions from all of the data axes except the Porod invariant. The second largest eigenvalue typically matched with a principle component axis with a strong contribution from the Porod invariant axis of the data, but with minimal contribution from any other axis, i.e. it was almost parallel to the Porod invariant axis. For some higher concentrations, other large eigenvalue principle axes show up corresponding to DATGNOM vs. other parameters. All of this matches with the correlation analysis, which, at lower concentrations, generally found everything to be correlated except the Porod invariant. At higher concentrations the DATGNOM derived parameters were increasingly uncorrelated from the other parameters. The PCA analysis supports the correlation analysis, indicating for lysozyme that there are only one or two axes of the data needed to describe the observed behaviour.

For xylanase, there are three strong principle component eigenvalues, all within an order of magnitude of the first, and three additional values within two orders of magnitude of the largest value. This supports our analysis from the correlation coefficients that three components could serve to describe the data set. While the principle component axes are not as clearly delineated in terms of which data axes contribute as they were for lysozyme, we do see principle component axes with strong contributions from the correlated components.

For glucose isomerase, there is one strong eigenvalue, which has a principle component axis with large contributions from the integrated intensity, Porod invariant, Porod volume, and molecular

weight data axes. The second strongest principle component axis contains significant contributions from the radius of gyration and $I(0)$ data axes, but the eigenvalue is roughly an order of magnitude less than the first component's eigenvalue. This indicates that perhaps just one parameter would be sufficient to describe the behaviour observed for this sample. The contributing data axes mostly correspond with what we expected due to the correlation analysis.

S6. Beam heating calculations.

Significant heating of macromolecules could lead to deleterious effects, such as denaturation, that could be mistaken for radiation damage. We have estimated the beam heating in two ways. First, a timescale for heat diffusion was estimated and an adiabatic assumption was made for the heating rate, which gives the adiabatic temperature rise. Second, the calculation from Appendix A of (Warkentin et al., 2012) was used to estimate a steady state temperature rise.

An upper bound on the heating rate due to X-ray illumination is obtained by assuming the illuminated volume is thermally isolated, so that heating is adiabatic. This adiabatic heating rate δT_{ad} is given by

$$\delta T_{ad} = \frac{DR}{c}, \quad (S9)$$

where DR is the dose rate and c is the specific heat capacity of the solution. Using a dose rate of 5 kGy/s (used in this work) and the specific heat of water, 4186 J/kg/K, gives an adiabatic heating rate of 1.2 K/s. An estimate for adiabatic beam heating will be most accurate for data collection times less than the timescale for heat diffusion out of the illuminated volume, which can be estimated as

$$t_d = \frac{L^2}{\alpha}, \quad (S10)$$

where L is a characteristic sample length and $\alpha \approx 0.14 \times 10^{-6} \text{ m}^2/\text{s}$ is the heat diffusion coefficient (Kuzay et al., 2001). For $L = 95 \text{ }\mu\text{m}$, half the beam FWHM, $t_d = 0.06 \text{ s}$.

The adiabatic heating in the sample, ΔT_{ad} is thus

$$\Delta T_{ad} = \delta T_{ad} t_d = \left(\frac{DR}{c} \right) \left(\frac{L^2}{\alpha} \right), \quad (S11)$$

a temperature rise of 0.08 K for our samples. Other factors may reduce this, such as heat transfer to the environment (particularly in actively cooled sample cells).

Another estimate of the beam heating can be taken from Appendix A of (Warkentin et al., 2012). We will not recreate the entire calculation here, but will simply state the main assumptions and the result. The model assumes the sample is an infinitely long cylinder whose axis is the beam axis. The sample has a radius r_2 , and the beam, also cylindrical, has a radius r_1 that is smaller than r_2 . Assuming infinite length neglects heat transfer in the direction parallel to the beam, so this model should

overestimate the temperature rise. The model also ignores beam absorption by the sample, so the intensity of the beam is always equal to the incident value. By applying appropriate initial and boundary conditions, an expression for the steady state temperature distribution is found:

$$T(r) = \frac{DRr_1^2}{2k} \log\left(\frac{r_2}{r}\right) + \frac{DRr_1^2}{2r_2h} + T_{ambient} . \quad (S12)$$

Here, DR is the dose rate in Wm^{-3} , $k = 0.5 \text{ Wm}^{-1}\text{K}^{-1}$, $h = 300 \text{ Wm}^{-2}\text{K}^{-1}$ (Kriminski et al., 2003), and $T_{ambient}$ is the ambient temperature around the sample. This expression diverges at $r = 0$, but provides an estimate for the heating at the edge of the beam, $r = r_1$. The steady state temperature rise ΔT_{ss} can be found as

$$\Delta T_{ss} = T(r_1) - T_{ambient} . \quad (S13)$$

For our experimental parameters, $\Delta T_{ss} = 0.13 \text{ K}$, in reasonable agreement with the adiabatic heating estimate above.

There is only one study in the literature on X-ray induced heating in liquid samples at modern synchrotron sources (Witala et al., 2014). The work was carried out at the Swiss Light Source, at 12.4 keV with a flux of 10^{13} ph/s into a $0.2 \times 0.2 \text{ mm}^2$ area at the sample, which gives a dose rate of 60 kGy/s using Equation (1) (assuming a Gaussian beam, and that $0.2 \text{ mm} \times 0.2 \text{ mm}$ is the FWHM). The sample was a binary mixture of water and 2,6-Lutidine (0.286 mass fraction) which has a heat capacity of 4.23 J/(g K) (Voronov & Buleiko, 1998), very similar to water. In the adiabatic approximation the heating rate is 14 K/s . The heat diffusion time is 0.07 s , giving an adiabatic warming of $\delta T_{ad} = 1 \text{ K}$.

The steady state temperature rise is $\Delta T_{ss} = 1.7 \text{ K}$. The measured local temperature rise after 60 s was 0.45 K in a temperature controlled copper sample cell near 34° C . This is in reasonable agreement with both estimates, and shows that each is likely to be an overestimate of the actual temperature rise.

It is possible that the onset of convection in the sample, due to thermal gradients, could create mixing and change the effective dose of the sample. There is significant literature about the onset of natural convection from heat sources. The most relevant geometry studied is that of a line heat source, typically treated theoretically as an infinite line. An expression for the ‘delay time’ at which significant convection is initiated from an infinite line heat source can be calculated, assuming: all temperature changes in the sample are due to the heat source, the sample is infinite, and that initially the sample is isothermal (Vest & Lawson, 1972). The delay time is

$$t^* = 43 \left(\frac{k\nu}{g\gamma\alpha^{1/2}q'} \right)^{2/3} , \quad (S14)$$

where k is the thermal conductivity, ν is the kinematic viscosity, g is the acceleration due to gravity, γ is the thermal expansion coefficient, α is the thermal diffusivity, and q' is the power

input per unit length. This expression has been shown to agree well with delay times measured for line heat sources in both water and air (Vest & Lawson, 1972; Boyd & Vest, 1975; Parsons Jr. & Mulligan, 1978; Ambrosini et al., 2003). The confinement of the sample cell geometry will make it harder to establish convection, so this represents a worst-case estimate for convection in our samples. The delay time was evaluated for water at room temperature, yielding

$$t^* = 43 \left(\frac{0.533 \text{ s}^{3/2} \text{ W/m}}{q} \right)^{2/3}. \quad (\text{S15})$$

Using our dose rate of 5.1 kGy/s, and assuming that the entire dose is distributed within the FWHM (which will create an underestimate of the delay time), $q' = 0.19 \text{ W/m}$. This gives $t^* = 89 \text{ s}$. Given that this is a lower bound, it is long enough to be irrelevant for our experiments.

Table 2 (main paper) gives beam heating estimates for ten of the twelve BioSAXS beamlines identified in (Graewert & Svergun, 2013) (one of the beamlines is no longer operational, and no data could be found for another). Data for beamlines in Table 2 is generally from publically available sources, and so may not be completely accurate. Calculations of t_d , δT_{ad} , ΔT_{ad} , ΔT_{ss} , and t^* were carried out, using the smallest beam dimension when applicable.

S7. Rate of oxygen absorption in degassed samples

The rate of oxygen absorption into a sample can be expressed as

$$\frac{1}{A} \frac{dW}{d\theta} = k_L (C_g - C_L), \quad (\text{S16})$$

where A is the area of the liquid gas interface in cm^2 , $dW/d\theta$ is the rate of absorption in g/h (grams per hour), $k_L = 0.4 \text{ cm/h}$ is the diffusion coefficient through a liquid film, C_L is the concentration of the gas in the liquid, and C_g is the saturation concentration of the gas in the liquid (Lewis &

Whitman, 1924). The solubility of oxygen in fresh water at 5°C is $12.8 \text{ mg/L} = 1.28 \times 10^{-5} \text{ g/cm}^3$ (ToolBox). The surface area of our sample plug in the sample holder is 0.04 cm^2 (Gillilan et al., 2013).

We assume that the sample started out with no dissolved oxygen, so $C_{L,0} = 0$. Additionally,

$C_L = W/V$ where V is the sample volume. Thus

$$\frac{dW}{d\theta} = Ak_L \left(C_g - \frac{W}{V} \right). \quad (\text{S17})$$

This differential equation has the solution

$$W(\theta) = \frac{C_g}{V} \left(1 - e^{-\frac{Ak_L \theta}{V}} \right), \quad (\text{S18})$$

or, in terms of percentage oxygenation relative to saturation, C_L / C_g ,

$$\frac{C_L(\theta)}{C_g} = \left(1 - e^{-\frac{Ak_L\theta}{V}} \right). \quad (\text{S19})$$

The time constant for this equation is given by $Ak_L / V = 0.53 \text{ h}^{-1}$ ($1.5 \times 10^{-4} \text{ s}^{-1}$), using a sample volume of 30 microliters. After 67 s, 1% oxygenation relative to maximum is achieved, after 702 s, 10% oxygenation relative to maximum is achieved. Thus, the samples should stay deoxygenated on the timescale of the experiments.

S8. Data

All of the scattering profiles used to generate the data in this paper, as well as dose and time of exposure for each profile is available from the Cornell eCommons, permanent link: **<to be inserted when ready>**. This is in accordance with the data access plan associated with NSF grant DBI-1152348.

Table S1 R_g , molecular weight, and integrated intensity sensitivities (S_{rg} , S_{mw} and S_I) and standard deviations (σ_{rg} , σ_{mw} , and σ_I) for all measured sample conditions. Of note is that degassing seems to slightly increase the damage rate, as does increasing the temperature. For S_{rg} and S_I these effects are small, while S_{mw} shows a larger effect.

Protein	Concentration (mg/mL)	S_{rg} (%/kGy)	σ_{rg} (%/kGy)	S_{mw} (%/kGy)	σ_{mw} (%/kGy)	S_I (%/kGy)	σ_I (%/kGy)
Lysozyme	0.5	21.2	3.0	36.9	11.1	22.2	6.0
Lysozyme	1.0	24.5	2.8	50.4	6.9	21.3	3.5
Lysozyme	2.0	26.5	10.6	48.8	8.0	23.3	11.1
Lysozyme	4.1	21.3	3.7	37.6	18.6	12.3	1.9
Lysozyme	8.1	16.5	6.6	11.9	5.8	7.0	3.0
Lysozyme	15.5	13.5	0.7	6.1	2.3	4.6	0.6
Lysozyme	32.2	6.4	2.0	3.1	2.0	2.5	0.7
Lysozyme	47.3	3.1	0.3	1.3	0.4	1.8	0.1
Lysozyme (Degassed)	4.1	24.3	1.2	49.0	2.4	15.4	1.1
Lysozyme (30 °C)	4.1	22.3	0.9	60.6	11.2	15.6	0.4
Xylanase	4.9	0.44	0.13	0.31	0.15	0.24	0.08
Xylanase (Degassed)	4.9	0.46	0.06	0.34	0.04	0.27	0.03
Glucose Isomerase	1.2	-0.000046	0.0007	-0.10	0.08	-0.023	0.009

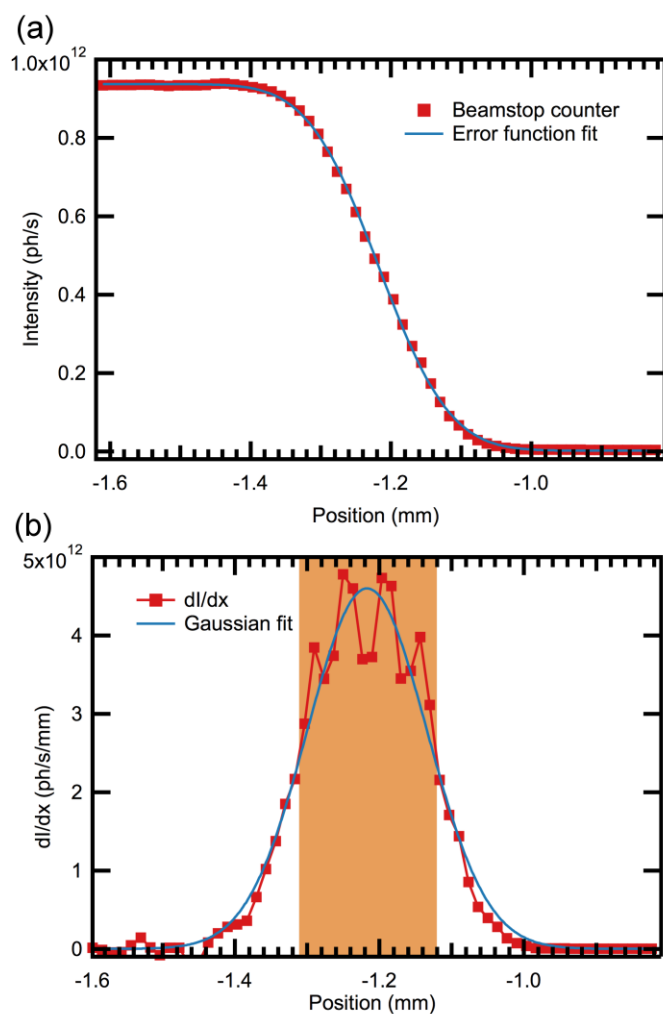


Figure S1 a) The beamstop counter output (converted to flux) as a function of the position of a knife edge being scanned through the beam in the x direction. A Gaussian beam should give an error function shape. The blue line shows an error function fit with FWHM of 188 μm . b) The numerical derivative of part a, giving the beam shape in the x direction. A Gaussian fit, FWHM 195 μm , is shown. The shaded orange region is the actual FWHM of the data, 190 μm .

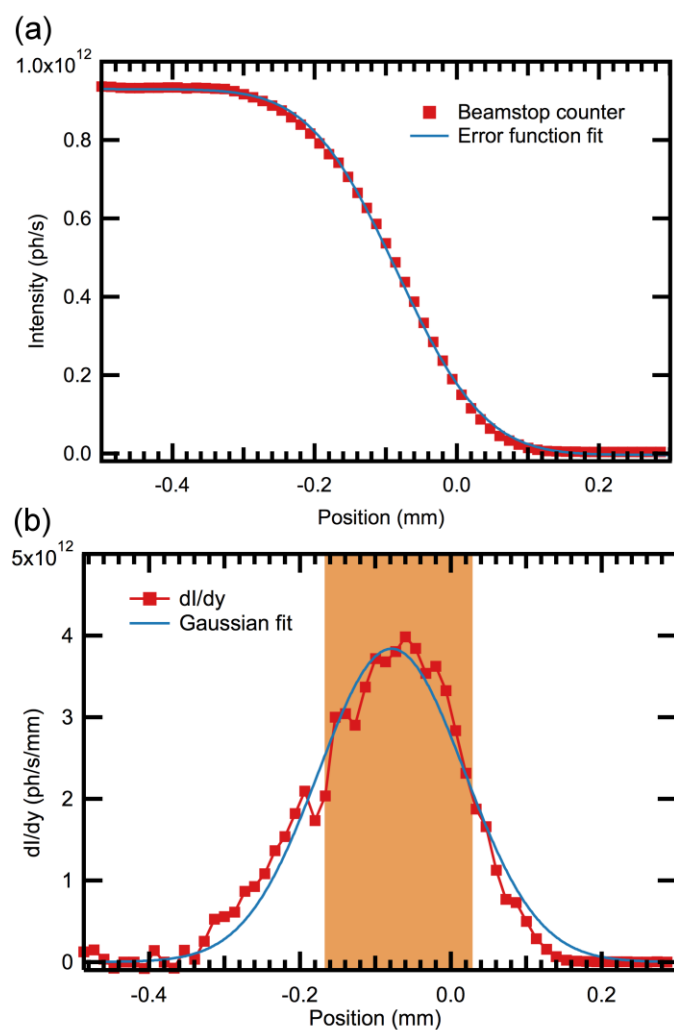


Figure S2 a) The beamstop counter output (converted to flux) as a function of the position of a knife edge being scanned through the beam in the y direction. A Gaussian beam should give an error function shape. The blue line shows an error function fit with FWHM of 229 μm . b) The numerical derivative of part a, giving the beam shape in the y direction. A Gaussian fit, FWHM 228 μm , is shown. The shaded orange region is the actual FWHM of the data, 196 μm .

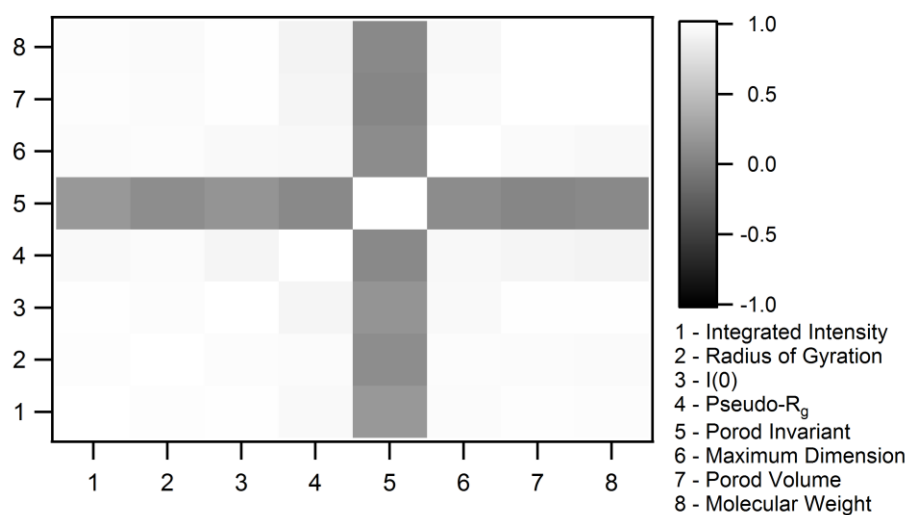


Figure S3 Plot of the average pairwise Pearson's correlation coefficient, r , for seven of the parameters calculated for each scattering profile for 4.1 mg/mL lysozyme. This shows that every parameter except the Porod invariant is strongly correlated.

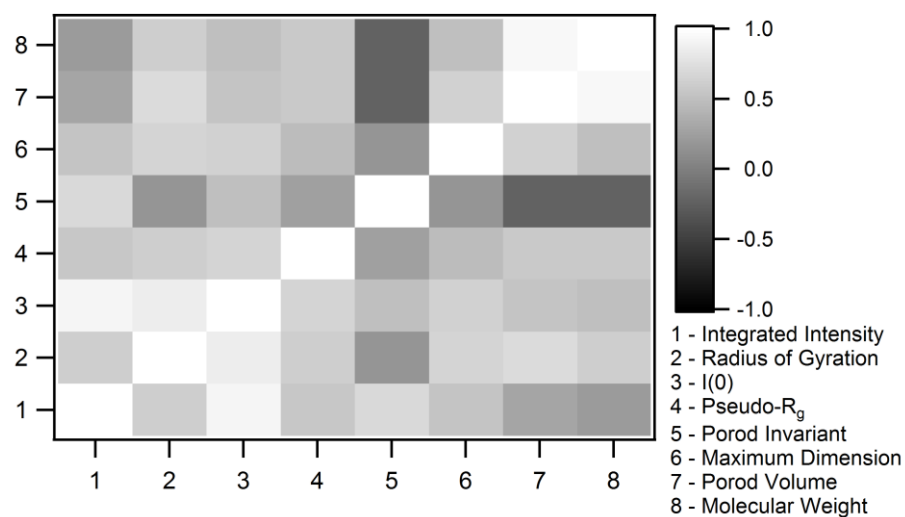


Figure S4 Plot of the average pairwise Pearson's correlation coefficient, r , for seven of the parameters calculated for each scattering profile for 4.9 mg/mL xylanase. This shows much less correlation than Figure 3.

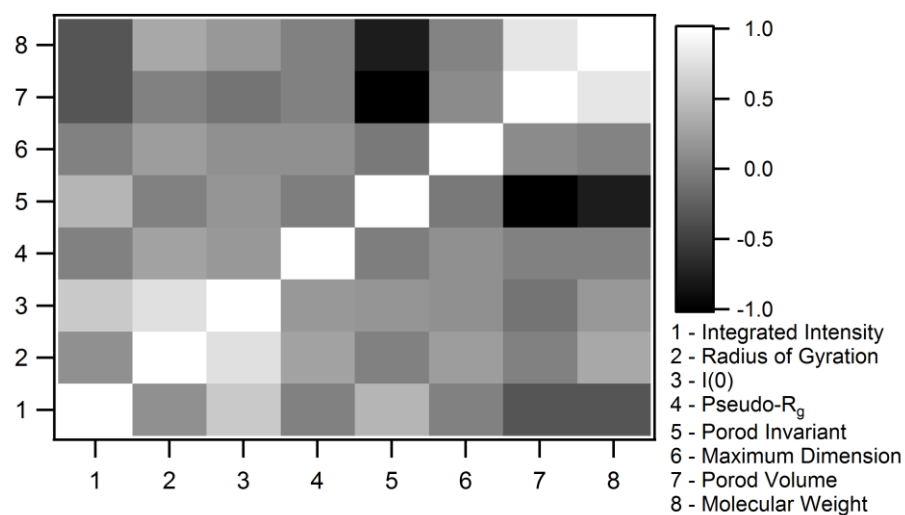


Figure S5 Plot of the average pairwise Pearson's correlation coefficient, r , for seven of the parameters calculated for each scattering profile for 1.2 mg/mL glucose isomerase. The only strong correlations seen are for the molecular weight and Porod volume.

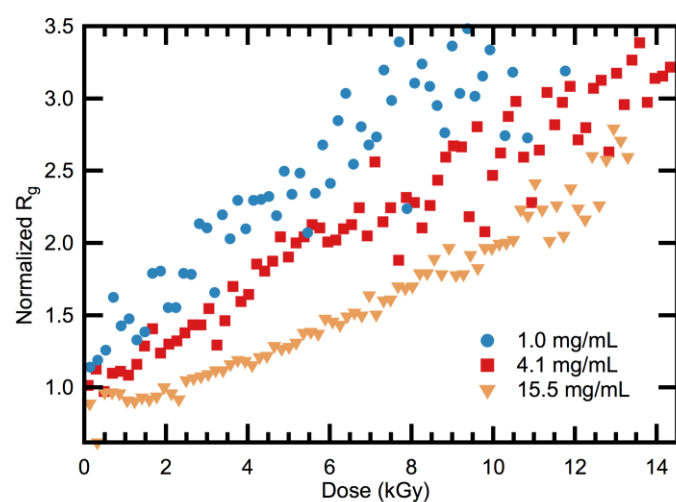


Figure S6 Plot of the normalized R_g dose curves for lysozyme at 1.0, 4.1, and 15.5 mg/mL. The lower concentration shows more scattering in the value of the R_g . The 15.5 mg/mL curve has an initial plateau, a delay to the onset of observable damage.

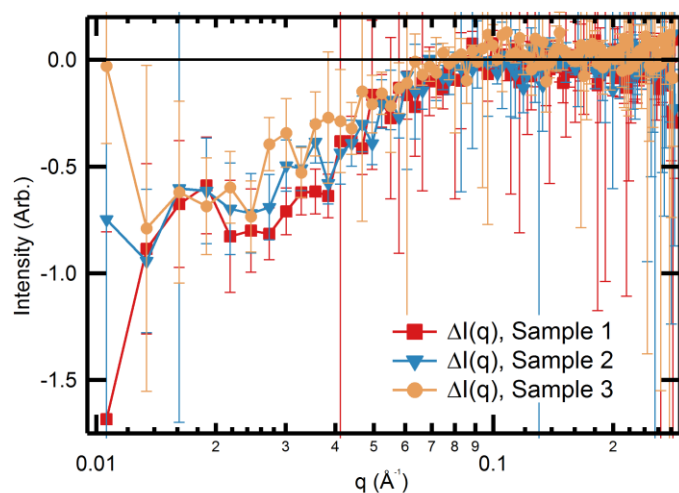


Figure S7 The difference in intensity, $\Delta I(q)$ between the average of the first 10 images (~ 20 kGy) and the last 10 images (~ 380 kGy) for the three identically prepared glucose isomerase samples. All three samples show a very similar change in the scattering profile with dose, showing that the observed decrease in the scattering profile, though unusual, is quite repeatable.

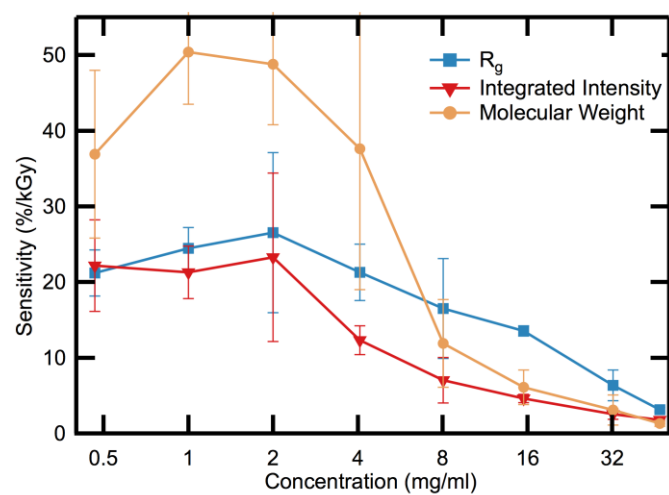


Figure S8 The R_g , molecular weight, and integrated intensity sensitivity as a function of lysozyme concentration. There is a region of either increasing or flat sensitivity at low concentration, followed by a decrease in sensitivity with concentration.

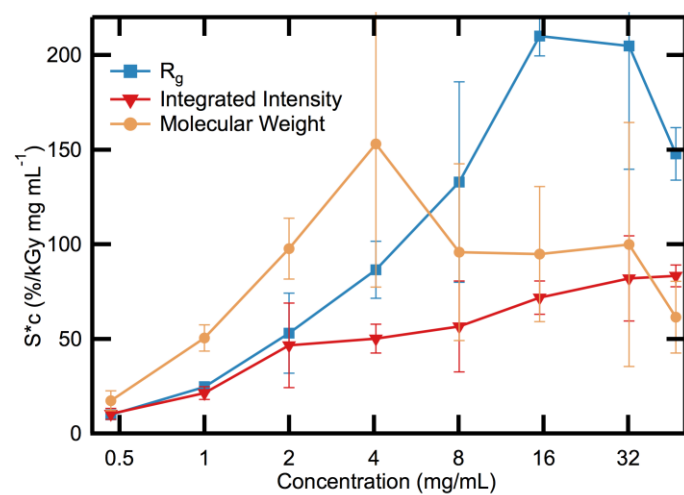


Figure S9 Plot of sensitivity times concentration vs. concentration for S_{rg} , S_{mw} , and S_I for lysozyme. No clear plateau is seen in the plot.

References

References for the supporting information are included with the references for the main paper.