Quantifying Radiation Damage in Biomolecular Small-Angle X-ray Scattering: Supporting Information

Authors

Jesse B. Hopkins[a]* and Robert E. Thorne[b]

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

[b]Department of Physics, Cornell University, Ithaca, NY, 14853, United States

Correspondence email: jbh246@cornell.edu

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 $M^{-1}cm^{-1}$ (Kozak, 2006) and 45660 $M^{-1}cm^{-1}$ (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 Å), 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 Å (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 $Å^{-1}$ to 0.2822 $Å^{-1}$ (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 × 250 μm^2 the FWHM of the beam was 190 (H) × 196 (V) μm^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 absolute intensity of the scattering profile – The integrated absolute intensity was calculated as

$$\int |I(q) - I_0(q)| dq + \int |I_0(q)| dq \qquad \text{(S1)}$$

where $I(q)$ is the scattering profile at a given dose, and $I_0(q)$ is the first scattering profile measured for the sample. Using the absolute value of the difference ensures that increases and decreases in the scattering profile do not cancel out in the integrated intensity metric. The addition of the initial absolute integrated intensity allows the standard normalization to the initial measured value (described later).

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 reports '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}) was defined in (Jeffries et al., 2015) using a known literature value for the R_g , R_g^u , and then calculating the R_g^{ps} for each 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 ΔR_g^{ps} Gy^{-1}, defined as the change in R_g^{ps} over the first 10 kGy of dose. In (Jeffries et al., 2015) ΔR_g^{ps} 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, as we already use $I(0)$ as a damage metric, we did not use it 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 estimated, Poisson counting statistics, 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:

$$A = 1 - \exp(-\mu l) \qquad \text{(S2)}$$

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 protein 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 cm^2/g. The addition of buffer components and protein changes μ_{pe}/ρ slightly (5.008 cm^2/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

$$A = 1 - \exp((\mu_{pe}/\rho)(l\rho)). \quad \text{(S3)}$$

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 (S3) to calculate A, or the transmission itself can be

adjusted according to the difference between the mass attenuation and photoelectric absorption coefficient as

$$A_{meas} = 1 - T_{meas} \exp\left[(\mu / \rho - \mu_{pe} / \rho) \rho l_m \right] . \quad \text{(S4)}$$

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

$$A = 1 - \exp(-(\mu / \rho)(l\rho)) \approx (\mu / \rho)(l\rho) . \quad \text{(S5)}$$

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 ~3%, and the dose by ~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 ~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

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

where $F(t,x,y,z)$ is the fluence (ph/μm^2) (without any dose correction, equal to ft 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, Equation (3) from the main paper, was used. These assumptions give

$$\frac{\partial F(t,x,y)}{\partial t} = D\left(\frac{\partial^2 F(t,x,y)}{\partial x^2} + \frac{\partial^2 F(t,x,y)}{\partial y^2}\right) + B(x,y)\,. \tag{S7}$$

Equation (S7) was numerically solved using Mathematica, with hard edge boundary conditions such that $F(t, x_{edge} \vee y_{edge}) = 0$ 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,

$$W(t) = \frac{\int B(t,x,y)F(t,x,y)dxdy}{\int B(t,x,y)dxdy}\,. \tag{S8}$$

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

$$C(t) = \frac{W(t)}{W_0(t)}\,. \tag{S9}$$

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 dimension(s).

Correction factors were calculated for each of the three proteins. Diffusion coefficients were measured by dynamic light scattering (Malvern Zetasizer Nano ZS, Malvern Instruments, Worchestershire, UK) to be: lysozyme, 1.3×10^{-6} cm^2/s (literature 1.13×10^{-6} cm^2/s (Price et al., 1999)); glucose isomerase, 0.6×10^{-6} cm^2/s (literature 0.53×10^{-6} cm^2/s (Oberthuer et al., 2012)); xylanase, 1.0×10^{-6} cm^2/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 ~1 μm beam instead of a ~16 μm 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 μm 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~0.1 Å^{-1}. Each sample condition (a given protein, concentration, temperature, and degassed state) was measured with 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 S1were 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 ~75% of their staring 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^{-1}, accounting for beam shape, for (Jeffries et al., 2015). Using this, we converted the initial damage rate of the pseudo radius of gyration, $\Delta R_g{}^{ps}$, 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, $S_{rg} = -4.6 \cdot 10^{-5}$ %/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 weight then increased to the average weight 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 weight 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 changed (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 $S_{rg}^{gi} << S_{mw}^{gi}, S_{I}^{gi}$, 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, O_2^{*-}, 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 due 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

$$\delta T_{ad} = \frac{DR}{c} , \tag{S10}$$

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

$$t_d = \frac{L^2}{\alpha} , \tag{S11}$$

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

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

$$\Delta T_{ad} = \delta T_{ad} t_d = \left(\frac{DR}{c} \right) \left(\frac{L^2}{\alpha} \right) , \tag{S12}$$

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:

$$T(r) = \frac{DRr_1^2}{2k}\log\left(\frac{r_2}{r}\right) + \frac{DRr_1^2}{2r_2h} + T_{ambient} \ . \qquad \text{(S13)}$$

Here, DR is the dose rate in $\mathrm{Wm^{-3}}$, $k = 0.5$ $\mathrm{Wm^{-1}K^{-1}}$, $h = 300$ $\mathrm{Wm^{-2}K^{-1}}$ (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

$$\Delta T_{ss} = T(r_1) - T_{ambient} \ . \qquad \text{(S14)}$$

For our experimental parameters, $\Delta T_{ss} = 0.13$ 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 0.2×0.2 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 mm x 0.2 mm is the FWHM). The sample was a binary mixture of water and 2,6-Lutidne (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 $\delta T_{ad} = 1$ K. The steady state temperature rise is $\Delta T_{ss} = 1.7$ 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

$$t^* = 43\left(\frac{k\nu}{g\gamma\alpha^{1/2}q'}\right)^{2/3}, \quad \text{(S15)}$$

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

$$t^* = 43\left(\frac{0.533\ \mathrm{s^{3/2}W/m}}{q}\right)^{2/3}. \quad \text{(S16)}$$

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), $q' = 0.19$ W/m. This gives $t^* = 89$ 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

$$\frac{1}{A}\frac{dW}{d\theta} = k_L(C_g - C_L), \quad \text{(S17)}$$

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), $k_L = 0.4$ 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 mg/L = $1.28\cdot10^{-5}$ g/cm^3. 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

$$\frac{dW}{d\theta} = Ak_L\left(C_g - \frac{W}{V}\right). \quad \text{(S18)}$$

This differential equation has the solution

$$W(\theta) = \frac{C_g}{V}\left(1 - e^{-\frac{Ak_L\theta}{V}}\right), \tag{S19}$$

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

$$\frac{C_L(\theta)}{C_g} = \left(1 - e^{-\frac{Ak_L\theta}{V}}\right). \tag{S20}$$

The time constant for this equation is given by $Ak_L / V = 0.53$ h^{-1} ($1.5 \cdot 10^{-4}$ 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

Every scattering profile, calculated dose, calculated parameter (Rg, molecular weight, etc.), and calculated P(r) function used in this paper, along with all data contained in the figures (a subset of the total dataset), is available from the Cornell eCommons, via the permanent link: http://hdl.handle.net/1813/43137. 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

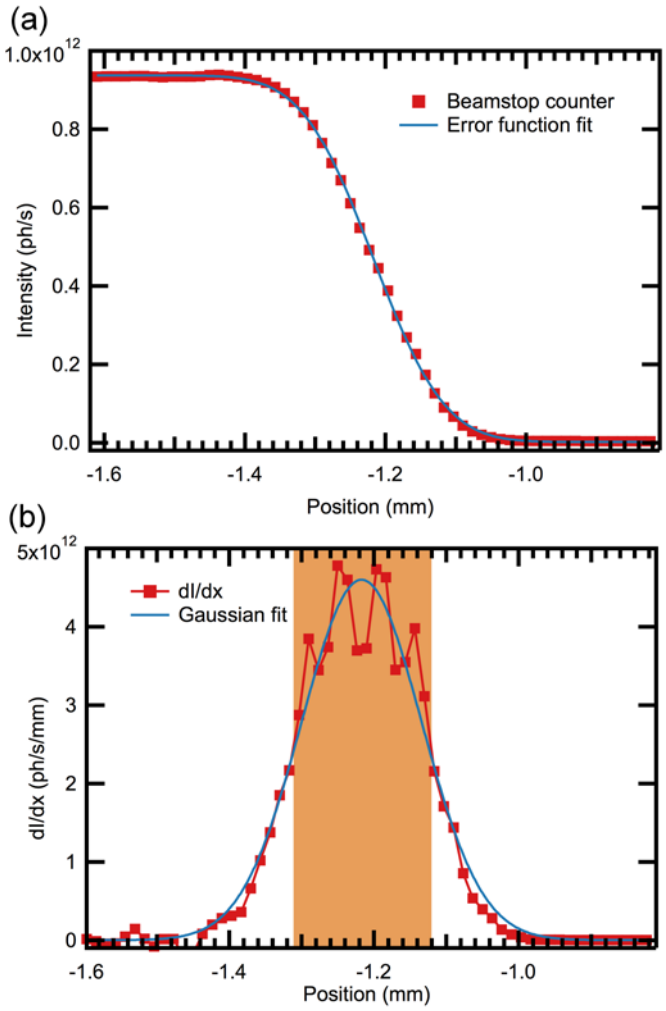

Figure S1 a) The beamstop counter output (converted to flux) as a function of the position of a knife edge being scanned through the beam in the *x* direction. A Gaussian beam should give an error function shape. The blue line shows an error function fit with FWHM of 188 μm. b) The numerical derivative 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.

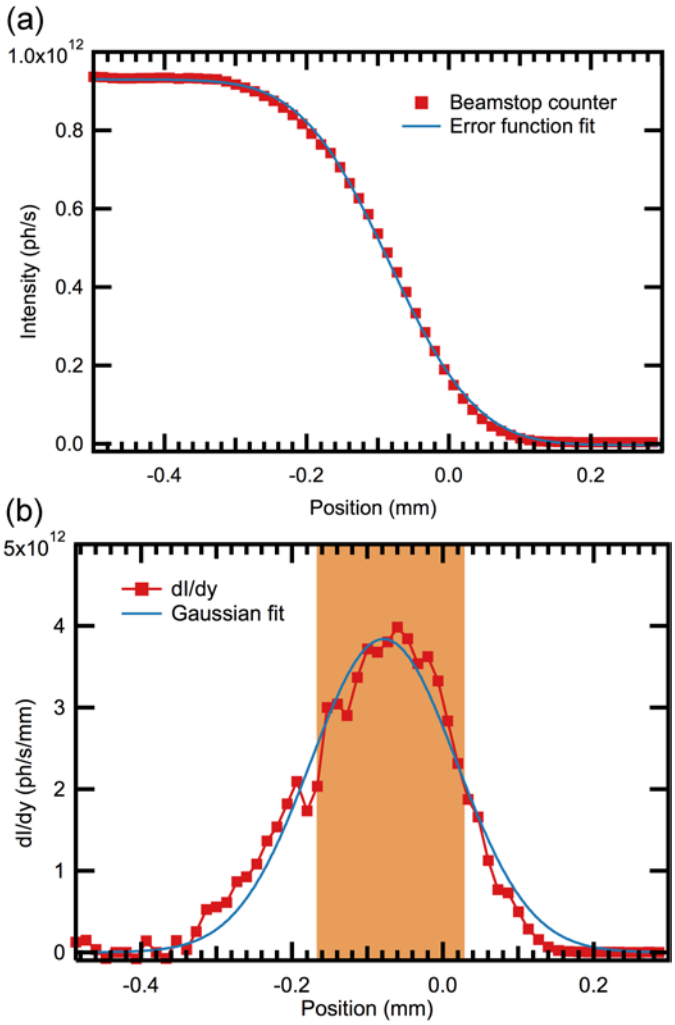

Figure S2 a) The beamstop counter output (converted to flux) as a function of the position of a knife edge being scanned through the beam in the *y* direction. A Gaussian beam should give an error function shape. The blue line shows an error function fit with FWHM of 229 μm. b) The numerical derivative 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.

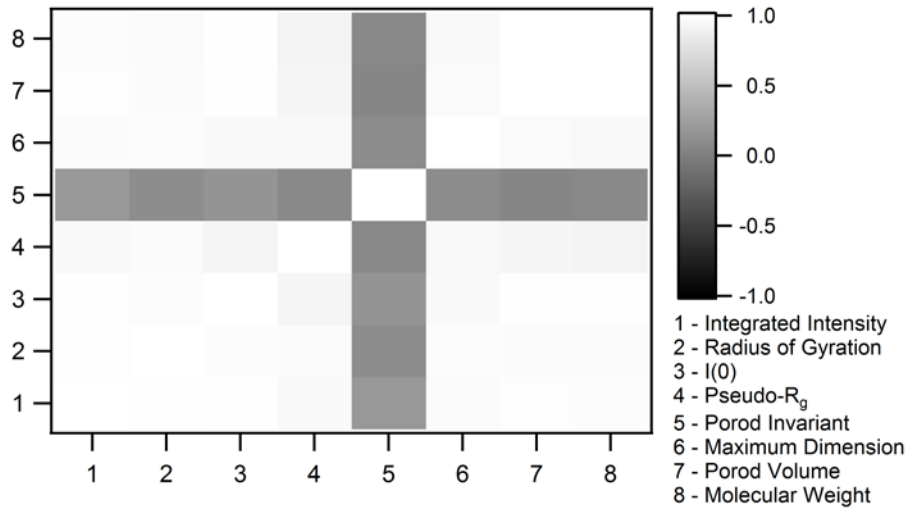

Figure S3 Plot of the average pairwise Pearson's correlation coefficient, *r*, for seven of the parameters calculated for each scattering profile for 4.1 mg/mL lysozyme. This shows that every parameter except the Porod invariant is strongly correlated.

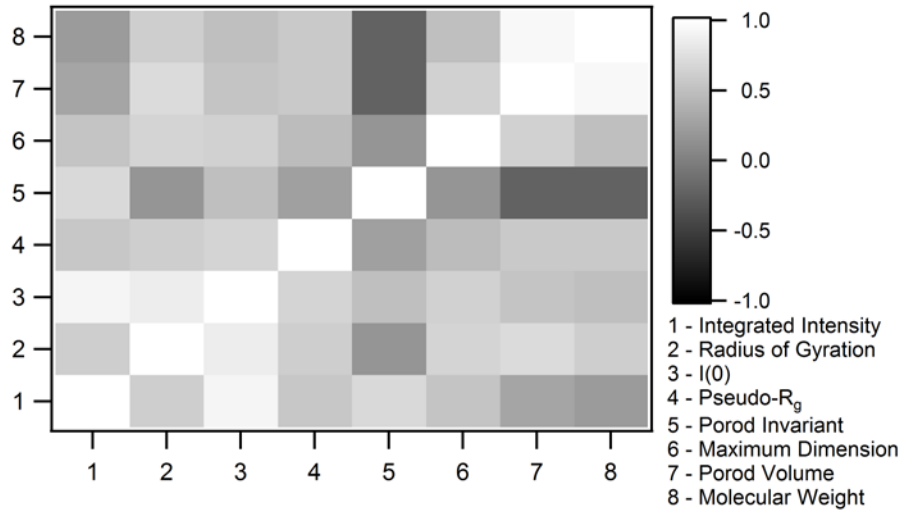

Figure S4 Plot of the average pairwise Pearson's correlation coefficient, *r*, for seven of the parameters calculated for each scattering profile for 4.9 mg/mL xylanase. This shows much less correlation than Figure S3.

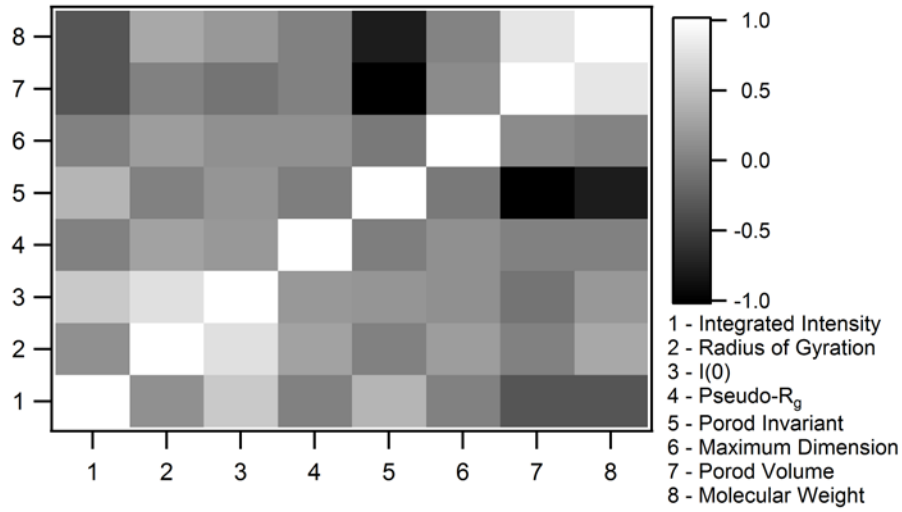

Figure S5 Plot of the average pairwise Pearson's correlation coefficient, *r*, for seven of the parameters calculated for each scattering profile for 1.2 mg/mL glucose isomerase. The only strong correlations seen are for the molecular weight and Porod volume.

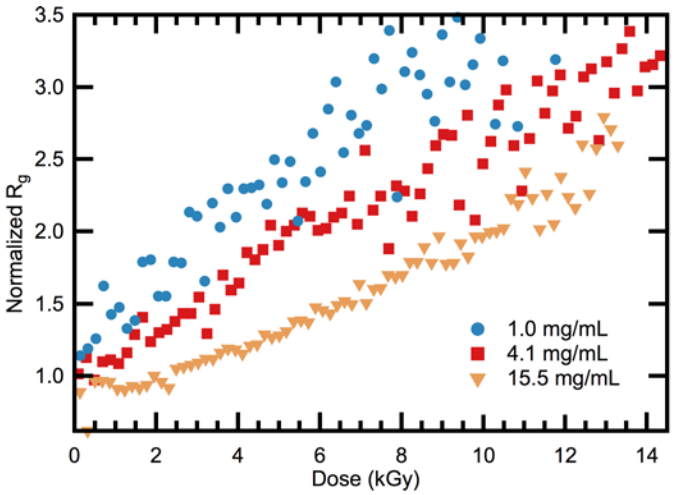

Figure S6 Plot of the normalized R_g dose curves for lysozyme at 1.0, 4.1, and 15.5 mg/mL. The lower concentration shows more scattering in the value of the R_g . The 15.5 mg/mL curve has an initial plateau, a delay to the onset of observable damage.

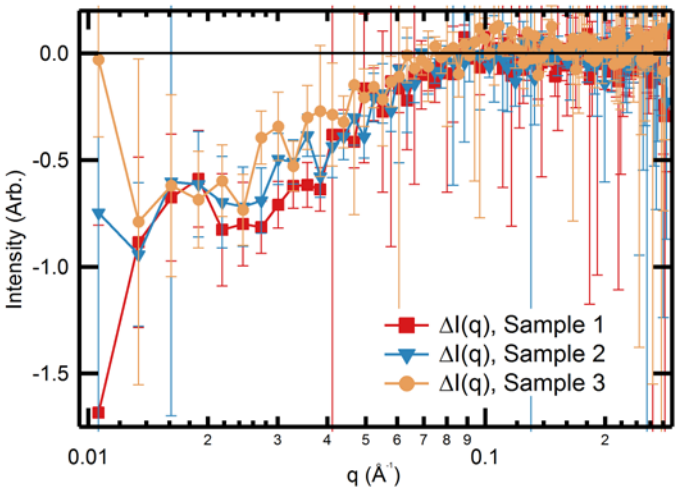

Figure S7 The difference in intensity, $\Delta I(q)$ between the average of the first 10 images (~20 kGy) and the last 10 images (~380 kGy) for the three identically prepared glucose isomerase samples. All three samples show a very similar change in the scattering profile with dose, showing that the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

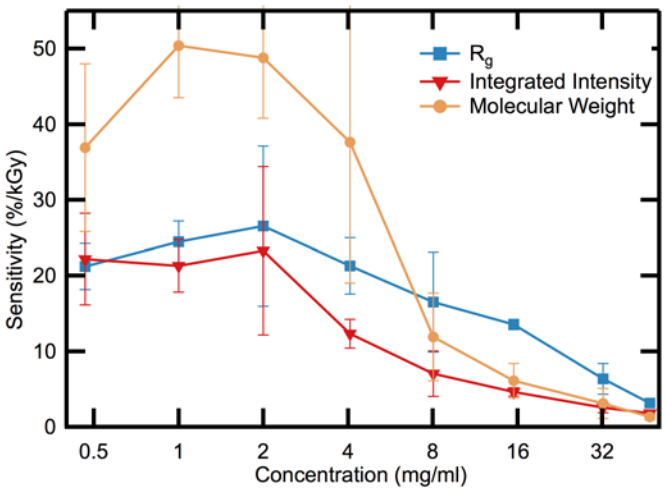

Figure S8 The R_g, molecular weight, and integrated intensity sensitivity as a function of lysozyme concentration. There is a region of either increasing or flat sensitivity at low concentration, followed by a decrease in sensitivity with concentration.

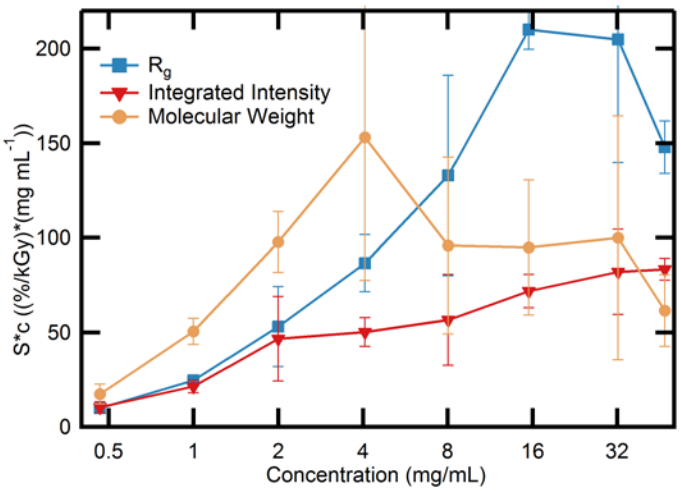

Figure S9 Plot of sensitivity times concentration vs. concentration for S_{rg}, S_{mw}, and S_I for lysozyme. No clear plateau is seen in the plot.

References

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